Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação
Departamento de Engenharia Biomédica

DETECÇÃO DE CONTORNOS DE VENTRÍCULO ESQUERDO EM IMAGENS DE MEDICINA NUCLEAR UTILIZANDO REDES NEURAIS

Cícero Ferreira Fernandes Costa Filho

Orientador: Prof. Dr. Lincoln de Assis Moura

Tese apresentada à Faculdade de Engenharia Elétrica e de Computação, FEEC - Unicamp, como requisito parcial para obtenção do título de DOUTOR EM ENGENHARIA ELÉTRICA.

91-22582

Novembro - 1996
Campinas - SP

Este exemplar corresponde à redação final da tese defendida por Cícero Ferreira Fernandes Costa Filho e aprovada pela Comissão julgadora em ___/___/___.

Lincoln de Assis Moura Jr
Orientador

C823d
32316/BC

UNICAMP
BIBLIOTECA CENTRAL

UNIDADE BC
N.º CHAMADA:
T/Unicamp
C823d
V. Ex.
TOMBO BC/32316
PROC. 281/97
C [ ] D [✓]
PREÇO R$ 11,00
DATA 28/12/97
N.º CPD

CM-00104027-6

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

C823d

Costa Filho, Cícero Ferreira Fernandes
Detecção de contornos de ventrículo esquerdo em imagens de medicina nuclear utilizando redes neurais / Cícero Ferreira Fernandes Costa Filho.--Campinas, SP: [s.n.], 1996.

Orientador: Lincoln de Assis Moura.
Tese (Doutorado) - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Processamento de imagem. 2. Redes neurais (Computação). 3. Reconhecimento de padrões. 4. Medicina nuclear. 5. Coração - Ventrículos. 6. Inteligência artificial. I. Moura, Lincoln de Assis. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.

DETECÇÃO DE CONTORNOS DE VENTRÍCULO ESQUERDO EM IMAGENS DE MEDICINA NUCLEAR UTILIZANDO REDES NEURAIS

RESUMO

Neste trabalho apresentamos um método que faz uso de uma rede neural e informação geométrica para a detecção automática do contorno do ventrículo esquerdo (VE) em imagens de medicina nuclear. Embora o método tenha sido desenvolvido para detecção de contornos de VE, ele pode ser estendido a outras classes de estruturas e imagens. O áprendizado é feito alimentando-se o sistema com uma série de imagens e seus correspondentes contornos de VE traçados por um operador. O sistema extrai tanto informações geométricas como informações de vizinhança, que são utitilizadas para treinar a rede neural. Uma vez treinada, a rede é capaz de automaticamente detectar o contorno do VE. São apresentados detalhes da operação e performance da rede neural, cujo desempenho é avaliado através de três medidas de erro. Um importante detalhe do método é a detecção automática do centro do VE, cuja exatidão tem uma grande influência na minimização dos erros. Ao lado de apresentar erros que são compatíveis com os erros de outros métodos automáticos, o método presente tem a clara vantagem de armazenar informações geométricas e de intensidade de pixel que são aprendidas através de exemplos.

USING NEURAL NETS FOR AUTOMATIC DETECTION OF LV CONTOURS IN NUCLEAR MEDICINE IMAGES

ABSTRACT

In this paper we present a method that makes use of an artificial neural network for the automatic detection of left ventricle (LV) contours in nuclear medicine images. The images were obtained from 60 male patients using a Gamma Camera. Although the method has been developed for LV contour detection, it can be extended to other classes of structures and images. Learning is carried out by feeding the system with a series of images and their corresponding LV contours drawn by an operator. The system extracts both pixel value and geometrical information that is used for training the neural network. Once trained the network is able to automatically detect LV contours. In this thesis we present details of the neural network operation and performance, which is evaluated through the use of three error measurement criteria. One important feature of the method is the automatic detection of the LV center, which plays an important role in minimizing the error values. Apart from presenting errors that are compatible with several other automatic detection techniques, the present method has the clear advantage of being able to store geometrical and pixel intensity information that is learned from examples.

AGRADECIMENTOS

O trabalho de uma tese não é realizado apenas por uma pessoa. Muitos contribuem e devem ser lembrados nessa hora. Em primeiro lugar, gostaria de agradecer ao orientador, Prof. Lincoln Moura. Fomos depositários de sua confiança inicial e, ao longo desses três anos de convívio, recebemos o seu apoio incondicional, seja através de discussões, de palavras de incentivo e de um relacionamento rico e construtivo.

Grande parte desse trabalho foi desenvolvido em um computador em nossa residência. Isso só foi possível graças a um convívio não muitas vezes interessante para Júlia e João Pedro, que tiveram que suportar muitos fins de semana sem passeios ou diversões. A eles, minhas desculpas e meus agradecimentos pelo carinho que recebo e pela dádiva de tê-los.

À companheira que comigo reparte as mesmas preocupações, envolvida que está em um trabalho similar, reservo os meus mais profundos agradecimentos. Tenho a certeza de que aprendemos muito da vida e de que fortificamos os nossos laços ao invés de destruí-los. Desempenhamos muitos papéis nessa caminhada. Revesamo-nos no cuidado das crianças, da casa e das teses. Obrigado pela sua compreensão e pelo carinho.

Também gostaria de agradecer ao Professor Edmundo da Silva Braga, que nos acolheu no princípio como se acolhe dois filhos.

Uma palavra de agradecimento a meus pais, incentivadores primeiros desta caminhada. Colocaram-me na estrada através do jardim da infância e me forneceram exemplos que procurei imitar. Quisera pudesse realizar uma obra como a de vocês!

Finalmente, agradeço ao Instituto do Coração - InCor/HC/USP pelo suporte técnico e por me ceder as imagens com as quais realizei este trabalho.

ÍNDICE

RESUMO ii
ABSTRACT iii
AGRADECIMENTOS iv
LISTA DE TABELAS viii
LISTA DE FIGURAS ix

CAPÍTULO I: INTRODUÇÃO 1

CAPÍTULO II: REVISÃO BIBLIOGRÁFICA 7
II.1 INTRODUÇÃO 7
II.2 REDES NEURAIS 7
II.2.1 *PERCEPTRON* MULTICAMADAS E ALGORITMO DE PROPAGAÇÃO REVERSA 11
II.2.2 REDE AUTO-ORGANIZÁVEL DE KOHONEN 15
II.2.3 REDE ART1 16
II.3 APLICAÇÃO DE REDES NEURAIS NA ÁREA MÉDICA 18
II.3.1 DIAGNÓSTICO 19
II.3.2 CLASSIFICAÇÃO E RECONSTRUÇÃO DE SINAIS 23
II.3.3 CONTROLE NEUROMUSCULAR 24
II.4 IMAGENS MÉDICAS 25
II.4.1 IMAGEM DE RAIOS X 25
II.4.2 IMAGEM DE ULTRA-SOM 27
II.4.3 IMAGEM DE RESSONÂNCIA MAGNÉTICA NUCLEAR 29
II.4.4 IMAGEM DE MEDICINA NUCLEAR 31
II.5 MÉTODOS DE DETECÇÃO DO CONTORNO DO VENTRÍCULO ESQUERDO 33
II.5.1 MEDIDAS DE AVALIAÇÃO DE DESEMPENHO 43

CAPÍTULO III: MATERIAIS E MÉTODOS 51
III.1 INTRODUÇÃ0 51
III.2 AQUISIÇÃO E PRÉ-PROCESSAMENTO DAS IMAGENS 54

III.3 DETERMINAÇÃO AUTOMÁTICA DO CENTRO DO VE..... 59
III.4 SELEÇÃO DAS IMAGENS PARA O TREINAMENTO DA REDE NEURAL 63
III.5 ARQUITETURA DA REDE NEURAL TIPO PERCEPTRON MULTICAMADAS..... 65
III.6 VARIÁVEIS DE ENTRADA E TREINAMENTO DA REDE NEURAL 67
III.7 UTILIZAÇÃO DA REDE NEURAL E CRITÉRIOS DE SELEÇÃO DOS PIXELS DE CONTORNO..... 80
III.8 RESTRIÇÕES GEOMÉTRICAS E FILTRAGEM 81
III.9 AVALIAÇÃO DO DESEMPENHO DO MÉTODO 83
III.10 SOFTWARE E HARDWARE UTILIZADOS..... 83

CAPÍTULO IV: RESULTADOS 87
IV.1 INTRODUÇÃ0..... 87
IV.2 CONTORNOS DE VE OBTIDOS AUTOMATICAMENTE 87
IV.3 COMPARAÇÃO ENTRE OS CRITÉRIOS DE SELEÇÃO DOS PONTOS DE CONTORNO..... 93
IV.4 ERROS OBTIDOS COM A VARIAÇÃO DA POSIÇÃO DO CENTRO 93
IV.5 CONHECIMENTO GEOMÉTRICO ARMAZENADO PELA REDE 95

CAPÍTULO V: DISCUSSÃO.....97
V.1 INTRODUÇÃ0 97
V.2 ANÁLISE DOS RESULTADOS..... 97
V.3 DISCUSSÃO DAS PRINCIPAIS CARATERÍSTICAS DA METODOLOGIA PROPOSTA PARA A DETECÇÃO DO CONTORNO DO VE. 100
V.4 DISCUSSÃO DAS CARACTERÍSTICAS DO TREINAMENTO 102

CAPÍTULO VI: CONCLUSÕES.....104
VI.1 INTRODUÇÃO..... 104
VI.2 CONCLUSÕES..... 104
VI.3 SUGESTÕES DE TRABALHOS FUTUROS..... 106

APÊNDICE A **109**

REFERÊNCIAS BIBLIOGRÁFICAS **115**

LISTA DE TABELAS

Tabela 2.1: Parâmetros das modalidades de imagens estudadas 32
Tabela 2.2: Resumo dos trabalhos sobre aplicação de redes neurais na área médica 46
Tabela 2.3: Resumo dos métodos de detecção do contorno do ventrículo esquerdo 49
Tabela 3.1: Erros percentuais no processo de determinação do centro 62
Tabela 3.2: Percentual de imagens nos grupos resultantes da classificação através da rede auto-organizável de Kohonen 64
Tabela 3.3: Valores médios normalizados e desvios padrões das variáveis de entrada da rede para as 60 imagens de VE e para 8 orientações: (a) orientações 0,8,16 e 24. (b) orientações 32,40,48 e 56 77
Tabela 3.4: Valores do teste t para as oito orientações mostradas na tabela 3.3: (a) orientações 0,8,16 e 24 e (b) orientações 32,40,48 e 56. 78
Tabela 3.5: Número médio de ciclos por treinamento e tempo de treinamento para a rede B (em segundos) em função da região do ventrículo 79
Tabela 3.6: Módulos do programa completo 86
Tabela 4.1: Valores dos erros percentual e erro soma para as imagens da figura 4.1 (valores adimensionais em termos percentuais) 90
Tabela 4.2: Valores dos erros percentual e soma para as imagens da figura 4.2 90
Tabela 4.3: Erro percentual e erro soma médios para as imagens dos grupos A e B 95
Tabela 4.4: Comportamento do erro percentual com a variação da posição do centro 95

LISTA DE FIGURAS

Figura1.1: Esquema das imagens do coração em medicina nuclear: (a) imagem de perfusão do miocárdio na posição anterior oblíqua esquerda e (b) imagem de blood-pool na posição anterior oblíqua esquerda. S - septo, VD - ventrículo direito, LP - lateral posterior, API - apical inferior. 3
Figura1.2: Curva de variação do volume do ventrículo esquerdo durante um ciclo cardíaco. VFD - volume de final de diástole, VFS - volume de final de sístole, PER - taxa de esvaziamento máxima (Peak Emptying Rate), PFR - taxa de enchimento máxima (PeakFillingRate) 4
Figura 2.1:Neurônio de McCulloch e Pitts 8
Figura 2.2:Ilustração do processo de classificação do perceptron 9
Figura 2.3: Arquiteturas de redes neurais:(a) redes de fluxo único;(b) redes recorrentes;(c) redes celulares 10
Figura 2.4: Arquitetura de um perceptron de 3 camadas 12
Figura 2.5: Tipos de regiões de decisão (Lippmann, 1987) 14
Figura 2.6: Espaço bidimensional com 11 regiões separáveis e duas classes (Mirchandani et alia, 1989). 14
Figura 2.7: Arquiterura de uma rede auto-organizável de Kohonen 15
Figura 2.8: Arquitetura da rede ART1 17
Figura 2.9: Exemplos de curvas de ROC com valores de área A na faixa entre 0,5 e 0,95. 21
Figura 2.10:Imagem de raio-x do tórax 26
Figura 2.11: Imagem de angiografia de subtração digital 27
Figura 2.12:Obtenção da imagem de CT 28
Figura 2.13:Corte tomográfico do coração. 28
Figura 2.14:Imagem de ultra-som de um feto com 5 semanas 29
Figura 2.15:Imagem de ultra-som do coração 30
Figura 2.16: Imagem de ressonância magnética nuclear do coração 31
Figura 2.17: Subdivisão do intervalo R-R em 32 quadros para aquisição de imagens dinâmicas sincronizadas 32
Figura 2.18: Diagrama em blocos de um sistema de visão computacional para detecção de contornos de VE 34
Figura 2.19: Detecção do contorno utilizando procura em grafos 37
Figura 2.20: Exemplo de aproximação de curvas por segmentos utilizada no trabalho de Coppini et aia (1995). 41
Figura 2.21: Ilustração da distância entre os pixels i,s dos contornos manual e automático para o cálculo do desvio médio de pixel ε. 44
Figura 3.1:(a)Representação esquemática de um contorno-θ com 8 orientações traçadas e (b)Representação esquemática de um contorno não θ com 8 orientações traçadas. 51
Figura 3.2: Assinatura de um contorno genérico 53

Figura 3.3: Posições da gama câmara: oblíqua anterior esquerda e oblíqua anterior direita 55
Figura 3.4: Vista do miocárdio na posição oblíqua anterior esquerda: VD-ventrículo direito, S-septo, LP - posição lateral posterior do VE, API - posição apical inferior do VE 55
Figura 3.5: Reconstrução do ciclo cardíaco por sincronismo reverso: (a) passagem direta; (b) passagem reversa e (c) união de 2/3 de (a) com 1/3 de (b) 56
Figura 3.6: Histogramas referentes a uma imagem 64x64. (a) histograma da imagem original; (b)histograma suprimindo os pixels com intensidade igual a zero e (c)histograma ajustado. 57
Figura 3.7: Histograma genérico utilizado no processo de ajuste do histograma da imagem original de 64x64. 58
Figura 3.8: (a) Imagem original com tamanho de 64x64; (b) imagem resultante do ajuste do histograma com tamanho 64x64 e (c) imagem resultante da interpolação bilinear com tamanho 256x256 59
Figura 3.9: (a) Imagem de blood-pool ampliada de um VE; (b) mediatriz, pixels com intensidade superior ao limiar de 200 e estimativa inicial do centro; (c) segmento de reta passando pelo centro estimado do VE e (d) centro refinado do VE e pontos 1,2,3 e 4 utilizados no seu cálculo. 61
Figura 3.10: Histograma médio dos pixels localizados sobre o segmento de reta mostrado na figura 3.10 (c) e sobre 5 segmentos adjacentes situados acima do mesmo e 5 segmentos adjacentes situados abaixo do mesmo. 62
Figura 3.11: Exemplos de ventrículos com formas diferentes: (a)forma alongada e (b)forma arredondada 63
Figura 3.12: Arquitetura da rede auto-organizável de Kohonen utilizada para a classificação das formas dos VE. 63
Figura 3.13: Percentual de acertos na classificação dos contornos automáticos pela rede auto-organizável de Kohonen, para n=2,3,4 e 5. 65
Figura 3.14: arquitetura da rede neural tipo perceptron multicamadas utilizada 66
Figura 3.15:Ilustração das distâncias dos pixels utilizadas no treinamento 68
Figura 3.16: Gráfico dos valores médios normalizados das distâncias em relação ao centro do VE e dos desvios padrões dos pixels interno, do contorno, e externo, para as 60 imagens de VE e para as 64 orientações. (a) pixel interno: 0,8 vezes a distância do pixel de contorno e pixel externo: 1,2 vezes a distância do pixel de contorno. (b) pixel interno: 0,7 vezes a distância do pixel de contorno e pixel externo: 1,3 vezes a distância do pixel de contorno. 71
Figura 3.17: Gráfico dos valores médios normalizados das intensidades e dos desvios padrões dos pixels interno, do contorno e externo, para as 60 imagens de VE e para as 64 orientações. (a) pixel interno: 0,8 vezes a distância do pixel de contorno e pixel externo: 1,2 vezes a distância do pixel e contorno. (b) pixel interno: 0,7 vezes a distância do pixel de contorno e pixel externo: 1,3 vezes a distância do pixel de contorno. 72
Figura 3.18: Gráfico dos valores médios normalizados do gradiente linear dos pixels interno, do contorno e externo, para as 60 imagens de VE e para as 64 orientações. (a) pixel interno: 0,8 vezes a distância do pixel de contorno e pixel

externo: 1,2 vezes a distância do pixel de contorno. (b) pixel interno: 0,7 vezes a distância do pixel de contorno e pixel externo: 1,3 vezes a distância do pixel de contorno. 73
Figura 3.19: Gráfico dos valores médios normalizados da magnitude do gradiente dos pixels interno, do contorno e externo, para as 60 imagens de VE e para as 64 orientações. (a) pixel interno: 0,8 vezes a distância do pixel de contorno e pixel externo: 1,2 vezes a distância do pixel de contorno. (b) pixel interno: 0,7 vezes a distância do pixel de contorno e pixel externo: 1,3 vezes a distância do pixel de contorno. 74
Figura 3.20: Gráfico dos valores médios normalizados das distâncias em relação ao pixel do contorno da orientação anterior e dos desvios padrões dos pixels interno, do contorno e externo, para as 60 imagens de VE e para as 64 orientações. (a) pixel interno: 0,8 vezes a distância do pixel de contorno e pixel externo: 1,2 vezes a distância do pixel de contorno. (b) pixel interno: 0,7 vezes a distância do pixel de contorno e pixel externo: 1,3 vezes a distância do pixel de contorno. 75
Figura 3.21: Divisão do VE em 4 regiões. 76
Figura 3.22: Variação do erro médio quadrático para um treinamento da rede correspondente a orientação de 22,5° 79
Figura 3.23: Saída da rede A para a orientação de 22,5°. 81
Figura 3.24: Exemplo de detecção de contorno: (a) imagem pré-processada de VE com 16 orientações superpostas;(b) imagem mostrando os pixels candidatos;(c) imagem com os pixels de contorno detectados através do critério #3 e (d) imagem com os pixels de contorno obtidos após a filtragem e interpolados linearmente. 85
Figura 3.25: Exemplo de detecção do contorno do VE: imagem com o contorno do VE após a suavização. 86
Figura 4.1: Exemplos de detecção do contorno automático pela rede neural A. O traço escuro representa o contorno automático, enquanto que o traço claro representa o contorno manual. 88
Figura 4.2: Exemplos de detecção de contorno automático pela rede neural B. O traço escuro representa o contorno automático, enquanto que o traço claro representa o contorno manual. 89
Figura 4.3: Exemplos de VE com diferentes tamanhos: (a) $25{,}35cm^2$, (b) $44{,}47cm^2$ e (c) $53{,}55cm^2$. 91
Figura 4.4: Exemplos de VE com diferentes formas: (a) forma ovalada, (b) forma arredondada e (c) forma irregular. 92
Figura 4.5: Exemplos de proximidade do músculo papilar em relação à região lateral posterior do VE. 93
Figura 4.6: Exemplos de contornos de VE com características especiais: (a) Na região superior do VE a região da borda não é muito bem definida, havendo a possibilidade de se traçar múltiplos contornos; (b) Na região inferior do septo a separação entre os ventrículos não é muito bem definida, havendo a possibilidade de se traçar múltiplos contornos e (c) Na região superior do

septo a separação entre os ventrículos não é muito bem definida, havendo a possibilidade de se traçar múltiplas bordas..... 94
Figura 4.7: Contornos médios. (a) contorno médio obtido pela rede neural utilizando uma imagem média obtida conforme descrito no texto; (b) contorno médio obtido pela média das coordenadas normalizadas dos pixels de contorno das imagens do grupo de treinamento..... 96
Figura 6.1: Formas e inclinações de VE. (a) e (b): VE com mesmas inclinações, porém com formas diferentes; (a) e (c) :VE com mesmas formas, porém com inclinações diferentes..... 107
Figura A.1: Rede neural tipo perceptron com 2 camadas de neurônios..... 110
Figura A.2: Rede neural tipo perceptron com três camadas de neurônios..... 112

CAPÍTULO I

INTRODUÇÃO

Em imagens médicas, os processos de aquisição e processamento de imagens registraram enormes progressos nas últimas décadas. Exemplos que atestam esse progresso são o processo de aquisição tomográfica de imagens e o avanço da aquisição de imagens funcionais de órgãos através da medicina nuclear. Paralelamente a isso, os processos de reconhecimento e interpretação de imagens médicas tiveram um grande desenvolvimento. Acreditamos, no entanto, existir ainda um largo campo de pesquisa a ser explorado nessa área.

No processamento de imagens podemos distinguir entre técnicas de segmentação de imagens e técnicas de interpretação ou reconhecimento. A segmentação decompõe a imagem em regiões, enquanto que o processo de reconhecimento gera uma descrição da imagem, atribuindo marcadores a essas regiões.

Imagens médicas representam estruturas biológicas com uma grande variedade de formas, que não podem ser facilmente descritas em termos de imagens padrões rotacionadas ou escalonadas, sendo difícil a segmentação das mesmas.

Na literatura, encontramos um vasto campo de aplicação de técnicas de segmentação em imagens médicas: segmentação de estruturas do cérebro (Badran et alii, 1990); segmentação do baço (Karssemeuer et alii, 1990); segmentação de células (Wu et alii, 1995); segmentação de nódulos em imagens pulmonares (Bae et alii, 1991); segmentação do ventrículo esquerdo (Costa et alia, 1995; Coppini et alii, 1995) e segmentação do fígado (Bae et alii, 1993).

Nesses trabalhos é utilizado um grande número de técnicas, que abrangem desde aplicação de limiar (Badran et alii, 1990) até técnicas sofisticadas de inteligência artificial, como redes neurais (Costa et alia, 1995).

Gonzalez et alia (1987) classificam os métodos de segmentação de acordo com a propriedade dos pixels de contorno que eles exploram: descontinuidade ou similaridade. Na primeira categoria, a imagem é segmentada em função de mudanças abruptas que ocorrem no nível de cinza dos pixels de borda da imagem. As principais técnicas referentes a primeira categoria exploram a detecção de pontos isolados de borda e de linhas de borda. Na segunda categoria, a imagem é segmentada em função das semelhanças entre os pixels. As principais técnicas da segunda categoria são baseadas em limiar, crescimento de regiões e *splitting and merging*.

Os algoritmos tradicionais de segmentação baseados em descontinuidade podem ser vistos como uma associação entre técnicas locais e técnicas globais. As técnicas locais têm por objetivo calcular valores significativos para os pixels, que os caracterizam como pertencentes ou não a uma borda. Esses valores são calculados através de operadores diferenciais como os operadores de Sobel,

operador Laplaciano, etc. As técnicas globais, por outro lado, têm por objetivo obter uma borda contínua integrando as informações fornecidas através dos operadores locais. Como exemplo de técnicas globais citamos técnicas que utilizam funções de custo e técnicas que utilizam grafos.

Os algoritmos de segmentação podem ser classificados ainda como manuais, semi-automáticos e automáticos. Nos algoritmos manuais a segmentação é efetuada inteiramente por um operador humano. Nos algoritmos de segmentação semi-automáticos, o operador humano fornece apenas pontos característicos da estrutura a ser segmentada ou delimita uma região onde a mesma é encontrada. Na segmentação automática, a intervenção do operador humano é reservada apenas a eventuais correções da segmentação obtida através de um algoritmo automático.

Mais recentemente, na segmentação de imagens médicas, observamos a utilização de métodos de inteligência artificial como técnicas *fuzzy* (Boudraa et alii, 1993; Feng et alii, 1993) e redes neurais (Costa et alia, 1995; Coppini et alii, 1995).

Embora seja desejável que o processo de segmentação automática dependa exclusivamente de informações provenientes da própria imagem, a tendência é os processos de segmentação utilizarem informações comuns a um tipo particular de imagens, implicando numa dependência entre os processos de segmentação e reconhecimento. Como exemplo, para compensar a baixa qualidade dos dados da maioria das modalidades de imagens médicas, pode-se utilizar, no processo de segmentação, informações "a priori" provenientes de modelos descritivos das imagens (Lilly et alii, 1989; Friedland et alia, 1989).

Entre as áreas de pesquisa em imagens médicas, um campo extremamente dinâmico na atualidade é a segmentação de órgãos e estruturas em imagens cardíacas. Particularmente, a segmentação do ventrículo esquerdo (VE) é intensamente pesquisada, devido a importância do mesmo na avaliação da função de bombeamento do coração.

Entre as modalidades de imagens cardíacas destacamos: medicina nuclear, ultra-som, ressonância magnética nuclear e tomografia computadorizada. No capítulo 2 apresentamos as características de cada uma dessas modalidades de imagem.

Nesse trabalho de tese utilizamos imagens de medicina nuclear. Essas imagens são geradas a partir da utilização de fármacos previamente associados com isótopos radioativos, formando os radiofármacos. Quando administrados a um indivíduo, esses radiofármacos apresentam uma biodistribuição preferencial, ou seja, fixam-se em órgãos específicos do corpo, como por exemplo tireóide e fígado. Os radioisótopos emitem radiação localmente e esta radiação é detectada utilizando-se uma gama câmara. A imagem resultante mostra o nível de radioatividade em cada ponto do órgão.

Dois radioisótopos são mais utilizados nas imagens de medicina nuclear do coração: o tálio-^{201}Tl e o tecnécio - ^{99m}Tc. O tálio é utilizado para obtenção das imagens de perfusão do miocárdio (imagens do músculo cardíaco), enquanto que o tecnécio é utilizado para obtenção das imagens de *blood-pool* (imagens do sangue nas cavidades cardíacas).

Na figura 1.1(a) mostramos uma imagem de perfusão do miocárdio com um diagrama ilustrativo correspondente. A parte mais clara dessa imagem corresponde ao músculo cardíaco. A principal aplicação dessa classe de imagens é no estudo da perfusão miocárdica.

Na figura 1.1(b) mostramos uma imagem de *blood-pool* do coração com um diagrama ilustrativo correspondente. As partes mais claras dessa imagem correspondem as cavidades dos ventrículos esquerdo e direito.

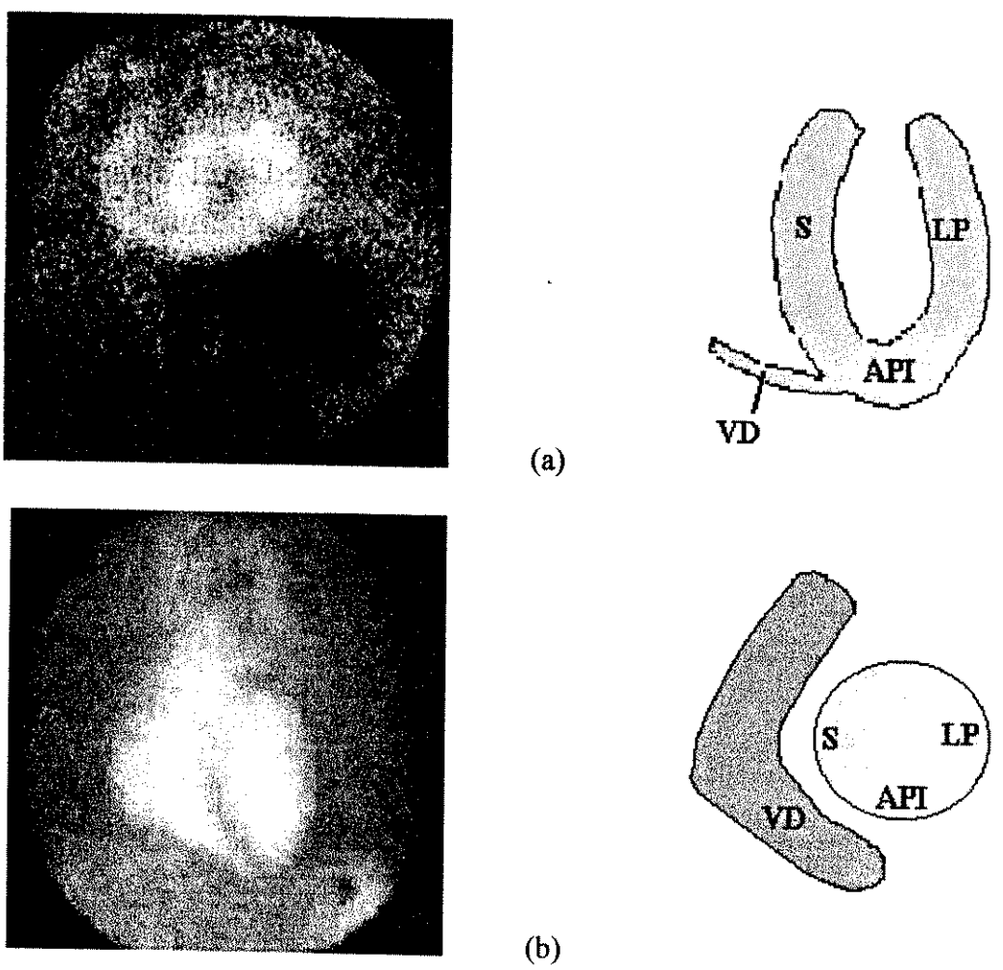


Figura 1.1: Esquema das imagens do coração em medicina nuclear: (a) imagem de perfusão do miocárdio na posição anterior oblíqua esquerda e (b) imagem de blood-pool na posição anterior oblíqua esquerda. S - septo, VD - ventrículo direito, LP - lateral posterior, API - apical inferior.

Nas imagens de *blood-pool*, a segmentação do ventrículo esquerdo (VE) é de fundamental importância para avaliação clínica de pacientes. Através da determinação da área do VE segmentado em diferentes instantes do ciclo cardíaco, são calculados parâmetros que permitem a avaliação da função cardíaca. O objetivo é quantificar a função de bombeamento do coração. Entre os parâmetros derivados do cálculo da área do VE segmentado, citamos: a fração de ejeção, a taxa máxima de

enchimento (Peak Filling Rate - PFR) e a taxa máxima de esvaziamento (Peak Emptying Rate - PER) (Gelfand et alia, 1988).

Na figura 1.2 mostramos uma curva idealizada da variação do volume ventricular durante um ciclo cardíaco, onde se pode ver as regiões correspondentes a diástole (VE relaxado) e a sístole (VE contraído) e as definições dos parâmetros acima mencionados.

A maioria dos métodos de segmentação do VE encontrados na literatura (conforme será visto no capítulo de revisão bibliográfica), é feita sob medida para uma determinada classe de imagens. Dito de outra maneira, métodos que são eficientes na detecção do contorno do VE em imagens de ecocardiografia, provavelmente não se adequam para efetuar a mesma segmentação em imagens de medicina nuclear

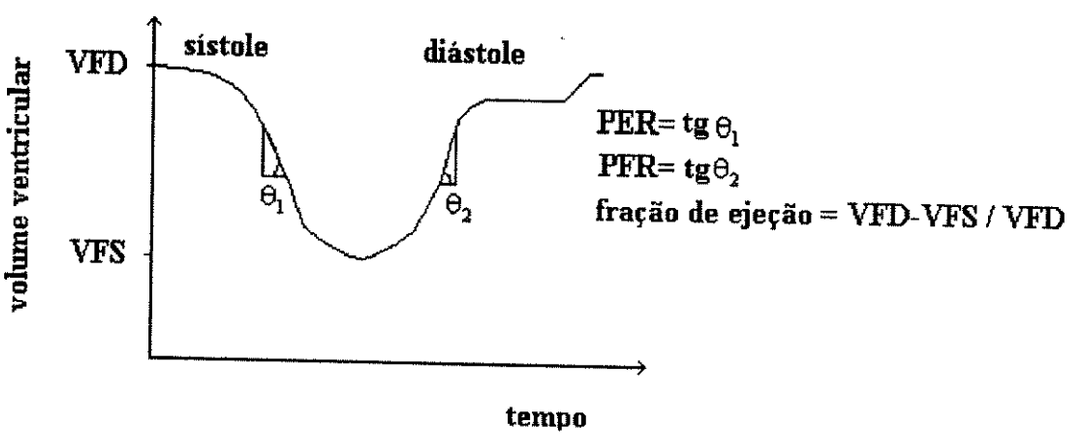


Figura 1.2: Curva de variação do volume do ventrículo esquerdo durante um ciclo cardíaco. VFD - volume de final de diástole, VFS - volume de final de sístole, PER - taxa de esvaziamento máxima (Peak Emptying Rate), PFR - taxa de enchimento máxima (Peak Filling Rate).

Nessa tese pretendemos superar essa limitação de especificidade dos métodos existentes para segmentação do VE. O principal objetivo a que nos propomos nessa tese é:

- A obtenção de um método automático para a segmentação do Ventrículo esquerdo em imagens médicas do coração.

A metodologia que propomos para segmentação do VE caracteriza-se por duas etapas principais:

- utilização de uma rede neural tipo *perceptron multicamadas* para a seleção de **pontos candidatos** ao contorno.

- utilização de critérios heurísticos para seleção dos **pontos do contorno** a partir dos **pontos candidatos** ao contorno.

Para validarmos essa metodologia de segmentação do VE utilizamos as imagens de *blood-pool* anteriormente referidas.

O conteúdo dessa tese é subdividido nos seguintes tópicos:

- Revisão Bibliográfica;
- Material e Métodos;
- Resultados;
- Discussão;
- Conclusões;
- Apêndice A

O capítulo de Revisão Bibliográfica é caracterizado por quatro momentos distintos. No primeiro momento, revisamos o conceito de redes neurais e fazemos uma breve análise das arquiteturas das redes neurais utilizadas nesse trabalho. Num segundo momento, revisamos as principais aplicações das redes neurais à área médica. Nessa altura incluímos um resumo das características dos trabalhos revisados sobre essas aplicações. O resumo é apresentado em forma de tabela. Num terceiro momento, apresentamos as diversas modalidades de imagens médicas de ventrículo esquerdo, incluindo a ecocardiografia, a tomografia computadorizada de raios X, a ressonância magnética e a medicina nuclear. Por fim, num quarto momento, revisamos os principais métodos existentes na literatura para segmentação do VE. Novamente, incluímos um resumo em forma de tabela das principais características desses métodos.

No capítulo de Material e Métodos apresentamos as características do método proposto e suas limitações. Em seguida analisamos o processo de aquisição e pré-processamento das imagens. Posteriormente, apresentamos o método automático utilizado para a determinação do centro do VE e a forma pela qual foram selecionadas as imagens para o treinamento da rede neural tipo *perceptron multicamadas*. A seleção das variáveis de entrada dessa rede neural é outro tópico de grande importância nesse capítulo. Por fim, analisamos as características do treinamento da rede neural e os critérios utilizados para a seleção dos *pontos do contorno* do VE.

No capítulo de Resultados comparamos inicialmente os contornos de VE obtidos através do uso de duas redes neurais *tipo perceptron* diferentes: uma com 4 variáveis de entrada e outra com 3 variáveis na camada de entrada. Em seguida mostramos contornos de VE de diferentes tamanhos e formas, obtidos automaticamente. São mostrados também contornos automáticos de VE com características especiais: ventrículos que apresentam regiões de difícil detecção da borda e ventrículos onde ocorre uma grande aproximação entre o músculo papilar e a parede dos mesmos. São mostradas ainda tabelas de comparação do desempenho de três critérios de seleção dos *pontos do contorno* e tabelas dos erros cometidos na detecção do contorno do VE ao se mudar a posição do centro do mesmo.

O capítulo de Discussão apresenta dois momentos distintos. Num primeiro momento analisamos os resultados apresentados no capítulo anterior. Posteriormente, analisamos a importância da estratégia adotada para seleção das variáveis de entrada da rede neural tipo *perceptron multicamadas*, e o papel fundamental que atribuímos às variáveis geométricas no desempenho dessa rede.

No último capítulo apresentamos as conclusões e as sugestões. Nas sugestões, duas vertentes são apresentadas para a continuidade desse trabalho: uma referente a aplicação do método para segmentação do VE e outra referente ao aprimoramento desse método.

No apêndice A demonstramos como o método de aprendizado de retro-propagação minimiza o quadrado das diferenças entre os valores de saída obtidos e os valores de saída desejados de uma rede neural tipo perceptron, através do gradiente descendente.

As imagens de *blood-pool* do VE utilizadas nessa tese foram fornecidas pelo Serviço de Radioisótopos do Instituto do Coração do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo.

CAPÍTULO II

REVISÃO BIBLIOGRÁFICA

II.1 INTRODUÇÃO

Neste capítulo iniciamos com uma revisão sobre redes neurais. Nesta revisão fazemos um breve histórico sobre o desenvolvimento das redes neurais, citamos as principais classificações dos tipos de redes existentes e analisamos a arquitetura e algoritmo de treinamento de três redes: *perceptron* multicamadas, rede auto-organizável de Kohonen e ART1. As duas primeiras foram utilizadas no trabalho da tese. A terceira foi mencionada por apresentar características especiais que a tornam bastante úteis em aplicações de classificação de padrões e encontra algumas aplicações interessantes na área médica. Em seguida expomos os resultados da pesquisa bibliográfica efetuada sobre a aplicação de redes neurais à área médica, ressaltando as aplicações na área de diagnóstico, classificação e reconstrução de sinais e controle neuromuscular. Posteriormente, analisamos as principais características das seguintes modalidades de imagens médicas: ultra-som, medicina nuclear, ressonância magnética e tomografia computadorizada de raios X. O capítulo é finalizado com uma análise dos métodos utilizados na literatura para detecção do contorno do ventrículo esquerdo em imagens médicas. Nesta análise descrevemos as principais características de um método de detecção do contorno do ventrículo esquerdo, avaliamos as principais técnicas encontradas na literatura sobre a detecção deste contorno e, a partir dos trabalhos revisados, deduzimos as características de um método "ideal" de detecção do contorno do ventrículo esquerdo.

II.2 REDES NEURAIS

Redes neurais são sistemas computacionais com largo espectro de aplicações na pesquisa acadêmica e na atividade industrial. A principal característica destes sistemas é serem formados por um grande número de elementos computacionais simples, denominados "neurônios". A função desempenhada por uma rede é determinada pelas seguintes características estruturais e funcionais: número total de neurônios, topologia da conexão entre os neurônios, valor dos pesos entre as conexões e tipo de processamento realizado em cada neurônio. Embora o processamento realizado por cada neurônio seja simples, o conjunto de neurônios pode realizar tarefas muito complexas, como reconhecimento de padrões, classificação, etc. As redes neurais podem ser encaradas sob duas concepções distintas: a primeira delas é encará-las como elucidativas das propriedades básicas e do processamento realizado pelo cérebro; a segunda é concebê-las como sistemas que podem realizar tarefas importantes sem nenhuma conexão com os sistemas biológicos que inspiraram a sua criação. Neste trabalho adotaremos esta segunda concepção.

O marco inicial no desenvolvimento das redes neurais foi estabelecido por McCulloch e Pitts (1943), quando propuseram o modelo do neurônio mostrado na figura 2.1. Este neurônio é uma unidade simples de processamento. Se existem **n** entradas x_i existirão **n** pesos w_i associados a essas entradas. O modelo do neurônio calcula uma soma ponderada das entradas. Ele calcula o produto da primeira entrada pelo peso correspondente a essa entrada, então faz a mesma coisa para a segunda entrada, e assim por diante, somando todos no final. Esta soma tem que ser comparada com um certo valor de limiar do neurônio. Se a soma é maior do que o valor do limiar, a saída é levada a 1. Se a soma é menor do que o valor do limiar, a saída é levada a 0. O valor do limiar é conhecido também como o *offset* ou polarização do neurônio. Podemos encarar esse processo como uma classificação, onde um padrão de entrada é classificado como pertencente a uma classe 1 ou a uma classe 0.

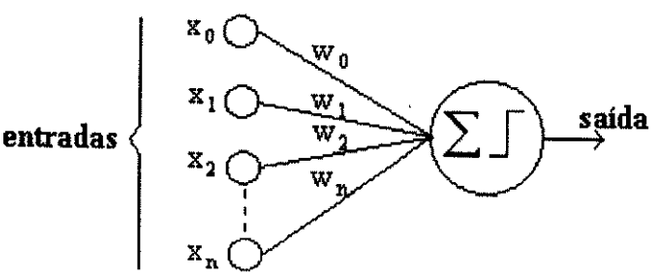


Figura 2.1:Neurônio de McCulloch e Pitts.

Uma estrutura de neurônios conectados foi denominado de *perceptron* por Rosenblatt (1962). Um *perceptron* é essencialmente uma rede linear de uma camada de neurônios sem realimentação. O processo de classificação do *perceptron* é uma extensão do processo de classificação do neurônio simples descrito anteriormente. Na figura 2.2 mostramos graficamente o resultado de um processo de classificação de padrões realizado por um *perceptron* em um espaço bidimensional com padrões linearmente separáveis. A grande limitação do *perceptron* é que ele só pode separar classes que sejam linearmente separáveis. Tal fato, demonstrado por Minsky e Papert (1969), paralisou por aproximadamente duas décadas a pesquisa na área de redes neurais..

Uma importante etapa na utilização de redes neurais é o treinamento das mesmas. O treinamento está associado à capacidade das redes neurais aprenderem a realizar uma determinada tarefa. Em termos do *perceptron,* o treinamento consistiria de um procedimento automático para determinação dos pesos w_i, a fim do mesmo determinar a melhor reta que separa as classes A e B.

Embora sistemas multicamadas de unidades de neurônios sejam muito poderosos, com capacidade de calcular qualquer função booleana, não existe nenhum algoritmo conhecido que possibilite o seu treinamento, ou seja, o mesmo é incapaz de aprender. Por outro lado, o *perceptron* pode ser treinado utilizando, por exemplo, a regra de aprendizado proposta por Hebb (1949) ou a regra delta proposta por Widrow-Hoff (1960).

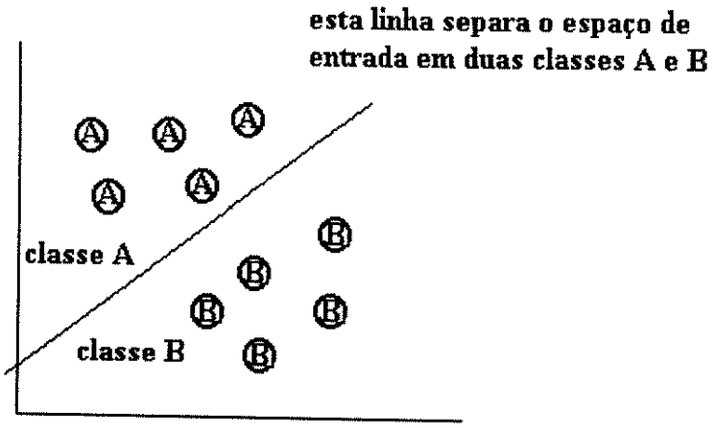


Figura 2.2:Ilustração do processo de classificação do perceptron.

Rumelhart et alia (1986), assumindo um processamento não linear do neurônio através de uma função sigmóide, propuseram um algoritmo de treinamento para uma estrutura de neurônios multicamadas. Isso só foi possível graças a introdução de uma função continuamente derivável no processamento do neurônio, a função sigmóide. No modelo do neurônio de McCulloch e Pitts era utilizada uma função com derivadas descontínuas, a função degrau. No apêndice A, mostramos este algoritmo detalhadamente. Após o trabalho de Rumelhart assistimos a uma verdadeira explosão na utilização de redes neurais nas mais diversas áreas de aplicação do conhecimento. As características das redes neurais que as tornaram tão atrativas são mencionadas a seguir:

- estrutura paralela, permitindo alta velocidade de computação com unidades simples de processamento.

- processamento distribuído por toda a estrutura, tornando o sistema tolerante a elementos com defeitos e com baixa sensibilidade a pequenas discrepâncias no desempenho das unidades individuais de processamento.

- o aprendizado é efetuado através de técnicas adaptativas, ao invés de serem impostos a priori valores para os pesos das conexões. Esta característica é particularmente importante quando tarefas complicadas têm que ser realizadas e onde as regras de decisão são difíceis de serem formuladas explicitamente, como por exemplo, o diagnóstico de doenças por um especialista da área médica.

A arquitetura das redes neurais pode ser dividida grosseiramente em três tipos: redes de fluxo único (*feedforward networks),* redes realimentadas ou recorrentes, e redes celulares. Esses três tipos são ilustrados na figura 2.3.

Nas redes de fluxo único, os neurônios da rede são arranjados em forma de camadas, N_n, não havendo realimentação. Esse tipo de rede produz um padrão na sua saída como resposta a um

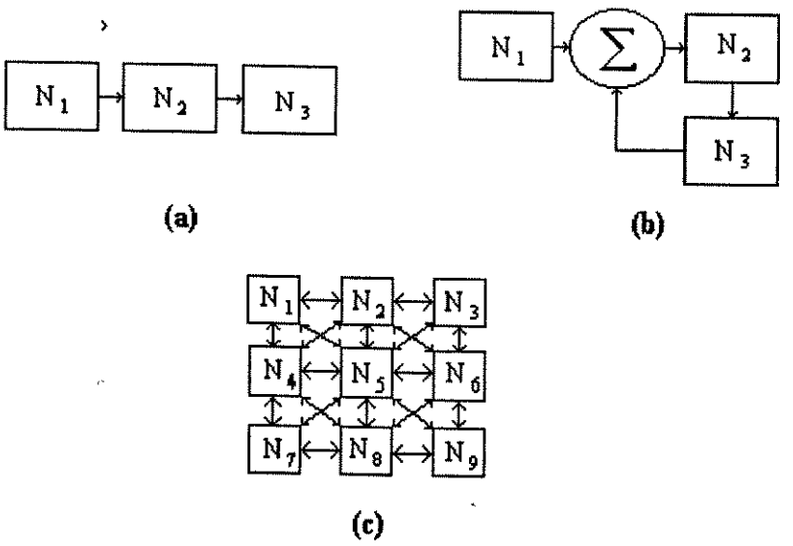


Figura 2.3: Arquiteturas de redes neurais:(a) redes de fluxo único;(b) redes recorrentes;(c) redes celulares.

padrão na entrada. Uma vez treinada, a resposta na saída a um padrão na entrada será a mesma, independente de qualquer ativação anterior da rede. Isso significa dizer que as redes de fluxo único não apresentam nenhuma dinâmica real e, portanto, não apresentam problemas de estabilidade. O mapeamento da entrada para a saída é realizado através de equações estáticas não lineares.

Nas redes recorrentes a dinâmica é mais complexa, uma vez que os seus neurônios possuem um processamento dinâmico (por exemplo, integradores e unidades de retardo). As propriedades dinâmicas destas redes são descritas através de equações diferenciais ordinárias não lineares ou equações de diferença.

Nas redes celulares, os neurônios são organizados numa matriz bidimensional, com arquitetura retangular ou hexagonal. Nestas redes não existem camadas e cada neurônio comunica-se diretamente com outros neurônios que estejam na sua vizinhança.

O exemplo mais conhecido de rede de fluxo único é o *perceptron* multicamadas. As redes celulares têm o seu melhor exemplar nas redes auto-organizáveis de Kohonen (1989). As redes de Hopfield (1984) são uma variante da estrutura celular mostrada na figura 2.2(c), onde todos os neurônios estão interligados. Um excelente exemplo de redes realimentadas são as redes ART1 (ART- Adaptive Ressonance Theory) e ART2 (Grossberg et alia, 1987 e Carpenter et alia, 1988).

Geralmente, uma rede neural é caracterizada não apenas pela sua arquitetura, mas também pelo algoritmo utilizado no seu treinamento e pela forma de operação. Os algoritmos utilizados no treinamento das redes neurais classificam-se em três tipos: supervisionados, não supervisionados e mistos. O treinamento supervisionado requer um conjunto de dados de treinamento com pares de entrada e saída. Durante a fase de aprendizado, os exemplos são apresentados à entrada da rede e o

resultado da saída é comparado com o valor desejado. Um termo de erro é calculado e utilizado para modificar os pesos da rede. Um exemplo de algoritmo supervisionado é o de propagação reversa (Rumelhart et alia, 1986).

Memórias associativas são outro exemplo de redes que utilizam algoritmo de treinamento supervisionado. A rede consiste de dois conjuntos de neurônios representando as camadas de entrada e saída. Todos os neurônios da camada de entrada estão conectados a todos os neurônios da camada de saída. Ajustando-se os pesos entre os neurônios é possível, na fase de treinamento, associar padrões de entrada a padrões de saída. Na fase de uso da rede, padrões apresentados à entrada da rede irão produzir padrões na saída correspondentes a um padrão "mais próximo" previamente aprendido.

O treinamento não supervisionado não requer uma saída desejada. Exemplos podem ser encontrados nas redes auto-organizáveis de Kohonen (1989) e Hopfield (1984). O treinamento da rede de Hopfield consiste de um processo de minimização de uma função de energia. As redes que utilizam algoritmos não supervisionados caracterizam-se por apresentarem uma estrutura constituída por uma camada de entrada e uma camada intermediária de neurônios. A rede aprende associando diferentes padrões de entrada a diferentes grupos de neurônios na camada intermediária. A teoria da ressonância adaptativa (ART) de Grossberg et alia (1987) é um outro exemplo de algoritmo de treinamento não supervisionado.

Um exemplo de algoritmo misto de treinamento é oferecido pelas redes de propagação contrária (*counter-propagation networks).* Estas redes utilizam tanto o treinamento supervisionado como o não supervisionado ([1] Hecht-Neilsen, 1987). A rede de propagação contrária consiste de três camadas de neurônios e é treinada em dois estágios. Primeiramente, o treinamento não supervisionado de Kohonen é empregado para ajustar os pesos da primeira camada entre os neurônios de entrada e os neurônios intermediários. Numa segunda etapa, o algoritmo *outstar* de treinamento supervisionado de Grossberg é utilizado para ajustar os pesos entre os neurônios da camada intermediária e da camada de saída.

A seguir analisaremos mais detalhadamente três combinações de arquitetura e algoritmo de treinamento: *perceptron* multicamadas com algoritmo de treinamento de propagação reversa; rede ART1 e rede auto-organizável de Kohonen.

II.2.1 *PERCEPTRON* MULTICAMADAS E ALGORITMO DE PROPAGAÇÃO REVERSA

Na figura 2.4 mostramos a arquitetura de um *perceptron* de três camadas: camada de neurônios de entrada, camada de neurônios intermediária e camada de neurônios de saída. Alguns autores, para caracterizar uma rede, preferem se referir ao número de camadas de pesos ao invés de número de camadas de neurônios. De acordo com esta visão, a rede da figura 2.4 seria referida como uma rede de duas camadas (camadas de pesos) ao invés de três camadas (camadas de neurônios). Neste trabalho estaremos nos referindo sempre ao número de camadas de neurônios.

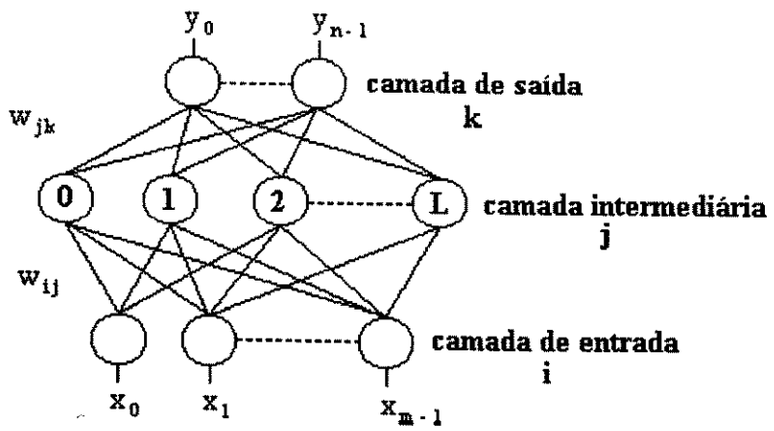


Figura 2.4: Arquitetura de um perceptron de 3 camadas.

O algoritmo de propagação reversa, conhecido também por "regra delta generalizada", é um algoritmo de gradiente descendente, cujo objetivo é minimizar o erro médio quadrático entre a resposta da rede e a saída desejada. Esse algoritmo é apresentado detalhadamente no apêndice A. A seguir, baseando-se na arquitetura da figura 2.4, descreveremos resumidamente este algoritmo. Assumiremos como função de ativação do neurônio a função sigmóide, definida como:

$$f(\alpha) = 1 / (1 + e^{-(\alpha-\theta)}) . \qquad (2.1)$$

1) Iniciar todos os pesos da rede com valores pequenos e aleatórios.

2) Apresentar um par de entrada-saída à rede.

3) Calcular a saída da rede utilizando a função sigmóide, considerando os seguintes valores de α:

$$\begin{cases} \alpha = \sum_{i=0}^{m-1} w_{ij} \cdot x_i \text{ , para os neurônios da camada intermediária } (0 \le j \le L-1) \\ \alpha = \sum_{j=0}^{L-1} w_{jk} \cdot o_j \text{ , para os neurônios da camada de saída } (0 \le k \le n-1) \end{cases} \qquad (2.2)$$

onde:

o_j = saída do neurônio j da camada intermediária .

4) Ajustar os pesos de acordo com a seguinte expressão:

$$\left.\begin{array}{l} w_{jk}(t+1) = w_{jk}(t) + \eta.\delta_k.o_j \\ \text{onde:} \\ \quad \delta_k = y_k.(1-y_k).(d_k - y_k) \\ \quad d_k = \text{saída desejada para o nó k} \end{array}\right\} \text{atualização dos pesos } w_{jk} \qquad (2.3)$$

$$\left.\begin{array}{l} w_{ij}(t+1) = w_{ij}(t) + \eta.\delta_j.x_i \\ \text{onde:} \\ \quad \delta_j = x_i.(1-x_i).\sum_k \delta_k.w_{jk} \end{array}\right\} \text{atualização dos pesos } w_{ij} \qquad (2.4)$$

onde:

η = taxa de aprendizado

5) Retornar ao passo 2.

Este algoritmo finaliza quando o erro quadrático médio entre a saída calculada pela rede e a saída desejada for menor do que um valor pré-definido. A arquitetura do *perceptron* multicamadas associada ao algoritmo de propagação reversa apresenta os seguintes inconvenientes: não existe uma forma sistemática de se estabelecer o número de neurônios ótimo da camada intermediária para resolver um problema específico; a matriz de pesos aprendida pela rede não possibilita um entendimento claro do funcionamento da mesma, ou seja, não existem regras explícitas que justifiquem o seu comportamento, e a convergência do algoritmo de propagação reversa é lenta, implicando em tempos de treinamento longos.

Tem havido um grande esforço no sentido de entender a relação entre a arquitetura do *perceptron* multicamadas e a classe de problemas que eles podem resolver. Lippmann (1987) apresenta um quadro que ilustra a relação entre *perceptrons* de 2,3 e 4 camadas de neurônios e a respectiva habilidade de classificação de cada uma dessas arquiteturas. Na figura 2.5 reproduzimos este quadro.

A princípio, como vemos na figura 2.4 o *perceptron* de 4 camadas é capaz de separar quaisquer classes. A complexidade das formas é limitada apenas pelo número de neurônios na rede. Isso é expresso formalmente na teoria através de um teorema provado por Kolmogorov em 1957 (Sprecher,1965) e popularizado por [2] Hecht-Nielsen (1987). Este teorema afirma que qualquer função, independentemente da sua complexidade, pode ser representada por um *perceptron* multicamadas de não mais do que 4 camadas de neurônios.

Mirchandani et alia (1989) trabalhando com *perceptrons* de 3 camadas de neurônios, definiram a função da primeira camada de pesos como a de estabelecer o número de regiões num espaço d-dimensional e a função da segunda camada a de selecionar estas regiões para definir classes. Um exemplo de um caso bidimensional é mostrado na figura 2.6, onde as regiões são numeradas de 1 a 11 e as classes são definidas como *A* e *B*. Mirchandani et alia demonstram que o

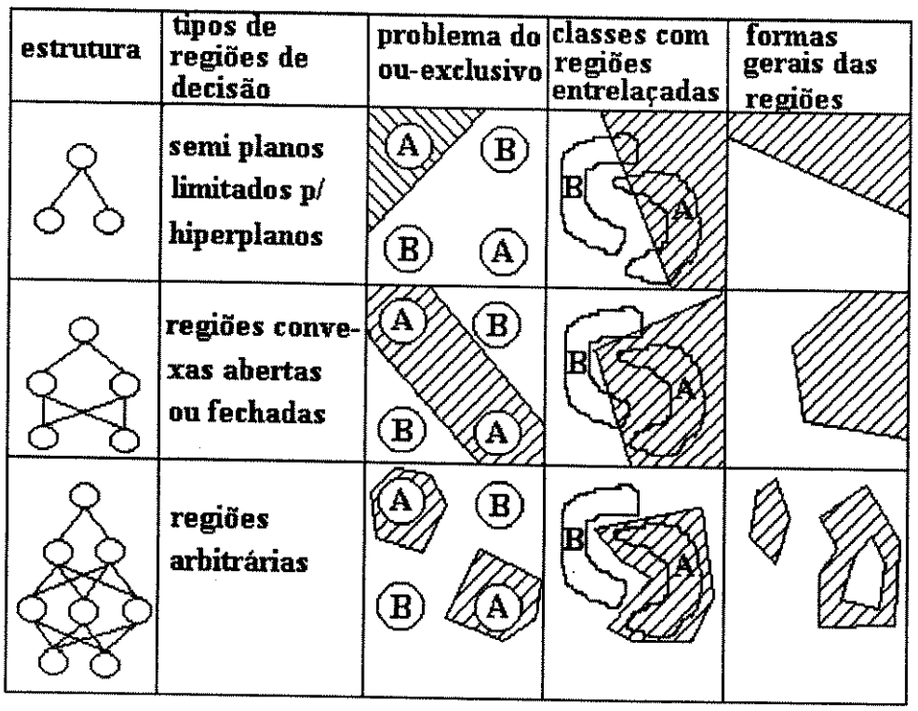


Figura 2.5: Tipos de regiões de decisão (Lippmann, 1987).

número máximo de regiões, M que são linearmente separadas utilizando H neurônios na camada intermediária, é dado por:

$$M(H,d) = \sum_{k=0}^{d} \binom{H}{k} \tag{2.5}$$

Baum (1988) ao analisar a potencialidade do *perceptron* multicamadas, mostrou que para um conjunto de N vetores num espaço de d dimensões, um *perceptron* com apenas uma camada intermediária de neurônios necessita no mínimo de N/d neurônios nesta camada para definir qualquer mapeamento entre estes N vetores de entrada e uma saída binária.

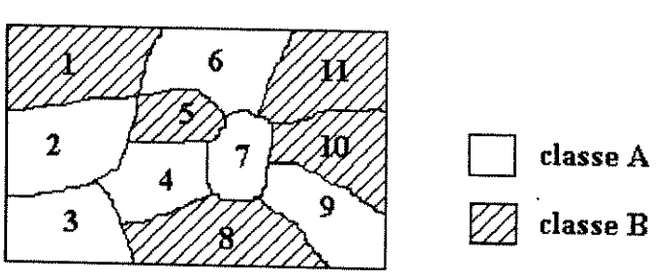


Figura 2.6: Espaço bidimensional com 11 regiões separáveis e duas classes (Mirchandani et alia, 1989).

Com o objetivo de minimizar o tempo de treinamento, há um grande número de técnicas propostas na literatura, entre as quais citamos: *quickprop* (Fahlman, 1988); adaptação global da taxa de aprendizado (Vogl et alii, 1988) e adaptação local da taxa de aprendizado (Jacobs, 1988). Cichocki et alia (1993) fornece-nos uma revisão apropriada destas técnicas.

II.2.2 REDE AUTO-ORGANIZÁVEL DE KOHONEN

A arquitetura de uma rede auto-organizável de Kohonen é mostrada na figura 2.7. A principal característica da arquitetura da rede auto-organizável de Kohonen é que os neurônios da camada de saída estão organizados numa grade plana e todos os neurônios da entrada se ligam a todos os neurônios da saída. A realimentação é restrita a interconexões laterais com vizinhos próximos na camada de saída. Cada um dos neurônios da grade já é um neurônio de saída.

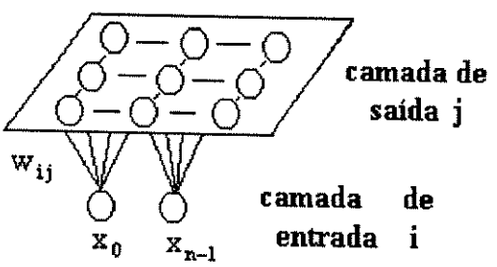


Figura 2.7: Arquitetura de uma rede auto-organizável de Kohonen.

O algoritmo de aprendizado da rede auto-organizável de Kohonen, denominado "algoritmo de quantificação vetorial" (*vector quantization)*, organiza os neurônios na grade em grupos que agem como classificadores dos dados de entrada. O mapa topográfico da grade é organizado sem supervisão, através de um processo cíclico de comparação entre os vetores apresentados à entrada da rede e os vetores de pesos associados aos neurônios de saída . Nenhuma resposta é especificada para uma entrada de treinamento. Onde ocorrer uma maior similaridade entre um vetor apresentado à entrada da rede e um vetor associado a um neurônio de saída, o mesmo é selecionado como representando um valor aproximado para o vetor apresentado à entrada da rede, e é denominado de neurônio vencedor. O aprendizado se dá obedecendo o seguinte algoritmo:

1) iniciar os pesos das N entradas para as M saídas com valores pequenos e aleatórios;

2) apresentar um novo vetor à entrada da rede;

3) computar as distâncias d_j entre o vetor apresentado à entrada da rede e os vetores de peso associados aos neurônios de saída usando a expressão:

$$d_j = \sum_{i=0}^{n-1} (x_i(t) - w_{ij}(t))^2$$

onde:

$$0 \leq j \leq m-1 \qquad (2.6)$$

5) selecionar o neurônio j^* com a menor distância d_j;

6) atualizar os pesos associados ao neurônio j^* e a todos os neurônios numa vizinhança $Vj(t)$ de j^*, de acordo com a seguinte expressão:

$$w_{ij}(t+1) = w_{ij}(t) + \eta(t).(x_i(t) - w_{ij}(t))$$

onde:

$$0 \leq i \leq n-1 \qquad (2.7)$$

$$j \in Vj^*(t)$$

7) voltar para o passo 2.

O algoritmo é concluído quando, de um ciclo para outro, não mais se observa mudanças em relação aos neurônios vencedores para os vetores apresentados à entrada da rede. O resultado deste algoritmo é que a partir de um conjunto de neurônios da grade organizados aleatoriamente, obtém-se um mapa de neurônios com representação local e auto-organizado. Uma análise mais aprofundada deste algoritmo pode ser encontrado em Kohonen (1989), onde o autor faz uma extensa abordagem de redes auto-organizáveis e de memórias associativas, procurando associar estas estruturas a comportamentos biológicos similares aos encontrados nas estruturas do cérebro humano.

II.2.3 REDE ART1

Outro exemplo de rede auto-organizável com características singulares é a rede ART1 proposta por Grossberg et alia (1987). Uma descrição em blocos simplificada da mesma é mostrada na figura 2.8.

A principal característica da rede ART1 é a forma encontrada para solucionar o dilema estabilidade-plasticidade. Para ilustrar este dilema recorreremos ao *perceptron* multicamadas treinado com o algoritmo de propagação reversa. Uma vez que o *perceptron* multicamadas já tenha sido treinado e desejássemos acrescentar um novo padrão ao conjunto de treinamento, isso só seria possível se alterássemos todos os pesos já estabelecidos no treinamento efetuado. Em outras palavras, a rede muda inteiramente a representação interna dos padrões armazenados, para se adaptar às mudanças introduzidas. Considerando que o tempo de treinamento do algoritmo de propagação reversa é longo, o acréscimo de novos padrões de treinamento é desaconselhado. Na rede ART1 é possível, após um treinamento inicial, armazenar novos padrões sem alterar a representação dos padrões previamente armazenados.

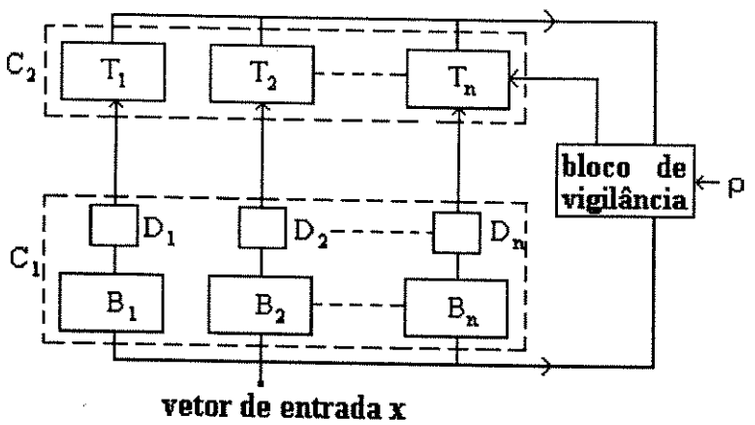


Figura 2.8: Arquitetura da rede ART1.

As principais fases do algoritmo de treinamento/classificação da rede ART1 são o reconhecimento, a comparação e a busca. A seguir descrevemos este algoritmo:

1) iniciação:

$$t_{ij}(0) = 1;\ b_{ij}(0) = 1/1+N$$

onde:

$$T_i = (t_{i0}, t_{i1} \ldots\ldots, t_{im-1});\ B_i = (b_{i0}, b_{i1}, \ldots\ldots b_{im-1});$$
$$0 \le i \le N-1; \qquad (2.8)$$
$$0 \le j \le M-1;$$
fixar o valor de ρ: $0 \le \rho \le 1$

O parâmetro de vigilância ρ fixa o grau de similaridade que deve existir entre um padrão de entrada representado pelo vetor x e um padrão armazenado T_i, para serem considerados como pertencentes a uma mesma classe.

2) aplicar uma nova entrada

3) reconhecimento: calcula-se o valor dos produtos internos, D_i:

$$D_i = \sum_{j=0}^{M-1} b_{ij}.x_i,\ X = (x_0, x_1, \ldots. x_{m-1}) \qquad (2.9)$$

seleciona-se o D_i^* com valor máximo: $D_i^* = \max\{D_i\}$.

4) comparação:

calcular $\|X\| = \sum_{j=0}^{M-1} x_j$ e (2.10)

$$\|T.X\| = \sum_{j=0}^{M-1} t_{ij} x_j \qquad (2.11)$$

se $\|T.X\| / \|X\| \rangle \rho$ ir para 6. Caso contrário ir para 5.

5) busca: A saída do produto interno D_i* é inibida e não toma mais parte do processo de maximização do passo 3. Vá para o passo 3.

6) adaptar os pesos dos vetores T_i * e B_i * correspondentes a D_i *, de acordo com as expressões:

$$t_{ij}^*(t+1) = t_{ij}^*(t).x_i \quad \text{e} \qquad (2.12)$$

$$b_{ij}^*(t+1) = \frac{t_{ij}^*(t).x_i}{0{,}5 + \sum_{j=0}^{M-1} t_{ij}^*(t).x_j} \qquad (2.13)$$

Se no final deste processo, nenhum padrão T_i é próximo o suficiente do vetor de entrada X segundo o critério dado no passo 4, é criado um novo neurônio nas camadas C_1 e C_2 para representá-lo. Uma apresentação pormenorizada deste algoritmo é feita por Grossberg et alia (1987).

II.3 APLICAÇÃO DE REDES NEURAIS NA ÁREA MÉDICA

A aplicação de redes neurais à área médica tem crescido rapidamente na última década. Elas têm progressivamente substituído os sistemas especialistas em aplicações de diagnóstico médico e muitos trabalhos têm sido efetuados nas áreas de processamento e interpretação de imagem e de sinais. Na área de diagnóstico médico, uma revisão abrangente é feita por Stubbs (1990), enquanto que Miller et alii (1992) exploram com profundidade as outras áreas citadas. Em ambos os trabalhos, a literatura revisada é anterior a 1992. Na revisão da literatura que empreendemos incluímos trabalhos posteriores. Um resumo das características dos trabalhos revisados é mostrado na tabela 2.2 no final do capítulo. Nesta tabela discriminamos o autor do trabalho, a categoria em que se insere, a arquitetura e o algoritmo de treinamento da rede neural, e o número de neurônios nas camadas da mesma. Os trabalhos apresentados na tabela 2.2 podem ser classificados genericamente nas seguintes categorias: diagnóstico médico, supervisão e uso de medicamentos, classificação de tecidos, reconstrução e classificação de sinais, controle neuromuscular e de revisão.

A maior parte dos trabalhos citados na tabela 2.2 referem-se a área de diagnóstico médico. As principais aplicações em diagnóstico são: doença de Alzheimer, câncer (de mama e melanoma), doenças pulmonares, doenças coronarianas e eletromiografia.

Os trabalhos sobre supervisão do uso de medicamentos utilizando redes neurais objetivaram prevenir iterações perigosas entre medicamentos ministrados a pacientes submetidos a tratamentos complexos (Gray et alii, 1991 e Gray et alia,1992). O sistema, implementado para utilização em

tempo real, faz uma advertência ao médico e/ou enfermeira sempre que uma iteração potencialmente perigosa for detectada entre os medicamentos prescritos a um paciente. Para se instalar esse sistema de supervisão em um hospital, é necessário a infra-estrutura de um sistema de informação hospitalar.

Nos trabalhos sobre classificação de tecidos, observa-se uma tendência de comparação entre técnicas de classificação supervisionadas e não supervisionadas. O tipo de imagem analisada nestes artigos é de ressonância magnética do cérebro (Hall et alia, 1992 e Clarke et alii, 1990).

Na área de classificação e reconstrução de sinais, as principais aplicações são: análise de sinais de EEG (Moreno et alii,1995), classificação de sinais de eletromiografia (Hassoun et alii,1994) e reconstrução de espectro de raios X (Boone, 1990).

Em relação ao controle neuromuscular, encontramos duas classe de aplicações: Na primeira delas, a ênfase é a utilização de redes neurais para interpretação de sinais de eletromiografia de pacientes paraplégicos, provenientes de regiões acima da lesão medular, com a finalidade de controlar membros paralisados (Graupe et alia, 1995). Dependendo do padrão destes sinais, a rede identifica a ação que o paciente deseja realizar, como mover a perna direita, mover a perna esquerda, sentar, etc. Na segunda classe, a ênfase é a utilização de redes neurais para o controle propriamente dito dos membros inferiores. São utilizadas redes recorrentes e simulações de modelos músculo-esqueleto (Abbas et alia, 1995).

Uma análise de todos os trabalhos foge ao escopo deste trabalho. Restringiremos a nossa análise a trabalhos selecionados nas áreas de diagnóstico, análise e interpretação de sinais e controle neuromuscular.

II.3.1 DIAGNÓSTICO

O processo de diagnóstico médico envolve normalmente as etapas de coleta, análise, reconhecimento e classificação de dados. Os dados do paciente são extraídos através de entrevistas, exames e testes. O médico, utilizando a experiência acumulada, infere um diagnóstico a partir deste conjunto de informações.

As ferramentas de inteligência artificial têm sido amplamente utilizadas no auxílio ao diagnóstico médico. As duas abordagens mais empregadas são os sistemas especialistas e redes neurais. Os sistemas especialistas requerem que um extenso conjunto de regras seja formulado por um médico trabalhando em conjunto com um profissional da área de informática. Um exemplo de sistema especialista é aquele que utiliza um conjunto de regras tipo *if-then* transformando os dados de entrada em uma saída (diagnóstico). Embora esta abordagem seja muito bem sucedida, conforme pode ser visto nos exemplos oferecidos por Shortliffe et alia (1990), o processo de enumeração das regras é demorado e tedioso. Ainda mais, em áreas que lidam com complexos critérios de decisão, como radiologia, onde um alto grau de subjetividade existe, pode ser difícil verbalizar uma lista de regras que possam ser utilizadas numa tarefa de diagnóstico. A característica de aprendizado não

algorítmica das redes neurais é extremamente atrativa para solucionar questões deste tipo. Devido a isso, observa-se um intenso interesse no uso de redes neurais na área de diagnóstico médico.

A seguir analisaremos alguns trabalhos que utilizam informações provenientes de imagens para o diagnóstico de doenças. Esses trabalhos podem ser classificados em duas categorias. Na primeira delas, os dados de entrada da rede são extraídos diretamente dos valores das intensidades dos pixels das imagens. Na segunda categoria, a informação para a entrada da rede é fornecida pelo médico, através de uma interpretação da imagem e do histórico do paciente. Essa informação é fornecida à entrada da rede neural em forma de parâmetros digitais e/ou analógicos. Como exemplo da primeira categoria, analisaremos trabalhos referentes a doença de Alzheimer e um trabalho referente a câncer de mama. Como exemplo da segunda categoria, analisaremos um trabalho referente a doenças pulmonares em recém-nascidos.

Os trabalhos sobre doença de Alzheimer utilizaram imagens de PET (Kippenham et alii, 1992) e SPECT (Dawson et alii, 1994 e Chan et alii, 1994). No trabalho usando imagens de PET o metabolismo médio de 8 regiões do cérebro é calculado a partir das intensidades dos pixels e utilizados como entradas da rede. Esses dados são normalizados na faixa de 0 a 2, a fim de adequá-los à função sigmóide. Assim, a entrada da rede corresponde a um vetor normalizado de oito dimensões. A rede foi treinada de tal forma que uma saída alta (próxima a 1) representa pacientes normais, enquanto que uma saída baixa (próxima a 0) representa pacientes com doença de Alzheimer.

Os resultados da classificação da rede são analisados através das curvas de ROC (receiver-operating characteristics). Essas curvas são construídas determinando-se pares de valores de sensibilidade-especificidade para diferentes valores de limiar na saída da rede, variando na faixa entre 0 e 1. A área sob a curva de ROC foi utilizada como figura de mérito. A área sob a curva de ROC, **A** mede o desempenho de um sistema de diagnóstico para vários valores do critério de decisão e é uma representação mais fidedigna do desempenho do sistema do que, por exemplo, um único par de valores de sensibilidade e especificidade. Pode ser mostrado (Swets, 1988) que a área sob a curva de ROC corresponde a probabilidade de uma resposta correta em um teste forçado com duas alternativas de resposta. Uma curva de ROC ideal apresenta um valor de área **A**=1. Na figura 2.9 mostramos o traçado de várias curvas de ROC com valores de área **A** variando na faixa entre 0,5 e 0,95.

Os resultados obtidos através da rede neural são comparados com a avaliação de um especialista e com um método de análise discriminante. São feitos estudos de auto-validação (teste da rede com os dados utilizados no treinamento) e de validação cruzada (teste da rede com dados não utilizados no treinamento). Com o objetivo de avaliar a importância do valor médio das entradas sobre a classificação, foi subtraído de todas as entradas da rede o valor médio das 8 regiões. As conclusões dos autores são que a subtração da média dos valores de entrada de rede tem influência mínima nos resultados da classificação e que os resultados da classificação obtidos com a rede são superiores aos obtidos com a análise discriminante, porém inferiores ao diagnóstico do especialista. Valores obtidos para a área sob a curva de ROC em estudos de validação cruzada foram: especialista, 0,89; rede neural, 0,85 e análise discriminante, 0,80.

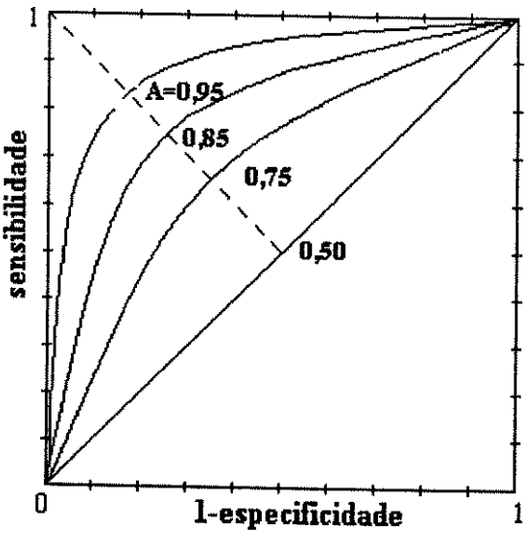


Figura 2.9: Exemplos de curvas de ROC com valores de área A na faixa entre 0,5 e 0,95.

No trabalho de Dawson et alii (1994) com SPECT, a entrada da rede é extraída a partir da contagem total de 14 regiões do cérebro. Os dados dessas contagens são normalizados através de uma divisão por 10.000, de tal forma que eles se situam na faixa de 0 a 12. A rede tem duas unidades de saída. Uma dessas unidades foi treinada para gerar uma saída com valor 1 para pacientes com a doença de Alzheimer e uma saída com valor 0 para pacientes normais. A outra unidade foi treinada para gerar uma saída com valor 0 para pacientes com a doença de Alzheimer e uma saída 1 para pacientes normais. São efetuados estudos de auto-validação e de validação cruzada. Os resultados do estudo de auto-validação são expressos através de um gráfico bidimensional, onde cada um dos eixos representa o valor de cada uma das unidades de saída, com uma variação na faixa entre 0 e 1. Observa-se neste gráfico uma separação linear entre pacientes normais e anormais. No estudo de validação cruzada a rede atingiu um índice de acerto na classificação de 72,5% ao classificar um total de 40 pacientes (16 pacientes normais e 24 pacientes com a doença de Alzheimer).

No trabalho de Chan et alii (1994), os dados de entrada da rede são extraídos a partir de 120 regiões do córtex, considerando os valores médios dessas regiões. Esses valores são então normalizados na faixa entre 0 e 1. A rede tem apenas uma saída e classifica os pacientes numa faixa entre 0 e 1. Valores de saída próximos a 1 correspondem a pacientes doentes, enquanto valores próximos a 0 correspondem a pacientes normais. Além de comparar a perfusão de pacientes com a doença de Alzheimer com pacientes normais, os autores também comparam a perfusão de pacientes drogados com cocaína com pacientes normais. O tempo de treinamento para cada rede foi de 10 a 15 minutos numa estação de trabalho SPARC2 da Sun Microsystem. O critério de convergência adotado foi que o erro total somado para todas as imagens fosse menor do que 0,001. Os resultados são analisados através de curvas de ROC, selecionando 10 valores de limiar na saída da rede para discriminar entre pacientes doentes e pacientes normais. A área sob a curva de ROC foi utilizada como figura de mérito para a análise do desempenho da rede. Na diferenciação entre pacientes com a

doença de Alzheimer e pacientes normais, conseguiu-se um valor médio para esta área de 0,93 (validação cruzada), enquanto que na diferenciação entre pacientes drogados com cocaína e pacientes normais, conseguiu-se um valor médio para essa área de 0,89 (validação cruzada).

A observação que fazemos sobre estes trabalhos de diagnóstico da doença de Alzheimer é que a área sob a curva de ROC obtida no trabalho com PET (0,89) foi inferior a área obtida no segundo trabalho com SPECT (0,93). Esse resultado é interessante, pois a modalidade SPECT é mais acessível em termos de custo que a modalidade PET.

O trabalho referente a câncer de mama que analisaremos foi realizado por Costa et alia (1995). Os autores utilizaram uma técnica de memória associativa, o filtro de novidades, para detecção de câncer de mama em imagens de cintilografia com MIBI- Tc^{99m}. O filtro de novidades foi treinado para reconhecer as características do tecido mamário normal. A conseqüência é que, em imagens com tumor, o mesmo é diferenciado das regiões normais. O conjunto de treinamento é constituído por 20 imagens cintilográficas de mamas normais, as quais são segmentadas automaticamente e padronizadas anatomicamente, de tal forma a se ajustarem às dimensões, forma e orientação de uma mama padrão. Posteriormente, são ortogonalizadas através do processo de Gram-Shmidt, a fim de constituírem uma base de vetores ortogonais. Outro conjunto com 34 imagens de mamas padronizadas é apresentado ao filtro de novidades para teste. O método descreve essas imagens como uma combinação linear dos vetores da base. Assumindo que as imagens da base sejam normais, não há diferenças significativas entre uma imagem de teste normal e sua correspondente descrição como combinação linear das imagens da base. As lesões de mama, no entanto, são identificadas como uma componente ortogonal ao sub-espaço gerado pelos vetores da base de mamas normais (novidade). Os resultados do trabalho são analisados através de curvas de ROC. Os valores de sensibilidade e especificidade obtidos com o método proposto foram de 94,12% e 88,24%, respectivamente.

Como exemplo da segunda categoria analisaremos o trabalho de Gross et alii (1990). Neste trabalho os autores utilizam redes neurais para interpretar radiografias do tórax de recém-nascidos. Para o treinamento da rede foram utilizadas 77 radiografias de pacientes com menos de 48 horas de vida, enquanto que para o teste foram utilizadas 103 radiografias de pacientes com o mesmo tempo de vida. A entrada da rede é alimentada com um vetor de 21 parâmetros fornecidos por um radiologista a partir da análise da imagem. Muitos desses representam informações binárias, indicando a presença (1) ou ausência (0) de algum sintoma. Outros parâmetros, como por exemplo infiltração pulmonar, assumem valores analógicos de 0,33, 0,66, e 1,0, que representam infiltração pulmonar pequena, moderada e severa, respectivamente. As saídas da rede são em número de 12, representando 12 possibilidades de diagnóstico distintas. Para cada imagem, o radiologista atribuiu um peso de 0 a 6 para cada um desses diagnósticos. A ausência de um diagnóstico para uma imagem era representado por 0, enquanto que uma forte indicação de um diagnóstico era representado por 6. Para uma comparação com a saída da rede esta faixa de 0 a 6 foi convertida para a faixa de 0 a 1. A rede foi treinada em um equipamento Microvax II da Digital Equipment. Na tentativa de minimizar o erro na saída, treinou-se a rede com 40.000 ciclos. Os resultados do diagnóstico fornecido pela rede são analisados de duas formas: a primeira delas considera que se uma das saídas da rede for diferente de 0 (por exemplo, 0,001), a mesma está sugerindo o diagnóstico correspondente àquela saída. Na segunda delas é estabelecido um limiar a partir do qual o diagnóstico é considerado. O limiar ótimo estabelecido foi de 0,032. Sem o estabelecimento do limiar, os resultados obtidos não foram

satisfatórios, tendo o desempenho da rede equiparado-se ao de um processo de escolha ao acaso dos diagnósticos. Com a utilização do limiar de 0,032, o desempenho da rede melhorou substancialmente. À exceção de 4 diagnósticos, que tinham baixa freqüência de ocorrência, a concordância entre os diagnósticos sugeridos pela rede e o diagnóstico de dois radiologistas experientes foi melhor que a concordância entre os diagnósticos dos dois radiologistas.

II.3.2 CLASSIFICAÇÃO E RECONSTRUÇÃO DE SINAIS

Nesta seção analisaremos um trabalho de classificação de padrões em eletromiografia e um trabalho de reconstrução do espectro de raios X .

[102] Hassoun et alii (1994) utilizam uma arquitetura tipo *perceptron* de 4 camadas para a classificação de potenciais de ação de uma unidade motora (PAUM). Devido à falta de conhecimento a priori da morfologia dos sinais de PAUM, a decomposição dos sinais de eletromiografia tem que ser realizada de forma não supervisionada. Neste trabalho, os autores propõem um algoritmo de treinamento não supervisionado baseado no algoritmo de propagação reversa. O método de decomposição consiste de três etapas. Na primeira etapa é utilizado um algoritmo de pré-processamento para identificar os picos existentes no sinal de entrada. Cada pico representa a localização de um PAUM. Em cada pico identificado é centrada uma janela de 48ms, de onde são extraídos 48 pontos para o treinamento da rede. Esse vetor é então normalizado por sua norma euclidiana. Para cada PAUM, além dos 48 pontos extraídos, entra-se na rede também com o valor da norma euclidiana utilizada para normalização e com o valor da amplitude do sinal. Na segunda etapa, utiliza-se a rede neural para aprender o número de padrões distintos no sinal de entrada e a morfologia destes padrões (treinamento). Na terceira etapa, a rede identifica a ocorrência destes padrões no sinal de entrada (classificação). O resultado é uma descrição completa do sinal de entrada como um conjunto de padrões únicos e uma tabela de ocorrência dos mesmos. As novidades do algoritmo de treinamento proposto pelos autores em relação ao algoritmo de propagação reversa são as seguintes: a) utiliza uma inclinação dinâmica na função de ativação dos neurônios da primeira camada intermediária, a qual satura durante a fase de treinamento; b) utiliza uma taxa de aprendizado amortecida e c) utiliza um conjunto de treinamento progressivamente restrito. Os objetivos dessas modificações são, respectivamente: a) saturação das funções sigmóides de ativação dos neurônios da primeira camada intermediária, limitando o número de padrões que podem ser armazenados na rede; b) diminuição da ênfase do aprendizado da rede em exemplos apresentados mais recentemente e c) remoção de padrões não corriqueiros, que podem dificultar o aprendizado da rede. Essas operações ocorrem automaticamente durante o treinamento. A rede é utilizada para a classificação de PAUM de pacientes normais e de 3 tipos de doenças: doença da junção neuromuscular, doença do músculo e doença do neurônio motor (padrão de disparo irregular). Em todos os processos de classificação foi extraída uma janela de 800ms de um sinal contendo 10s de gravação.

Na fase de classificação ocorre um processo iterativo. Quando um padrão de entrada é apresentado à rede, o mesmo propaga-se para a saída. O vetor presente na saída é próximo de um padrão de PAUM ideal previamente armazenado. A saída da rede é então realimentada para a entrada e uma nova saída é calculada. Este processo é repetido e em cada passo a saída é comparada

com a entrada. Quando a diferença entre duas saídas consecutivas da rede torna-se mínima, diz-se que a rede atingiu um valor estável e o PAUM presente em sua saída representa um padrão ideal aprendido pela rede na fase de treinamento.

O desempenho da rede é avaliado utilizando um sinal sintético contendo 4 padrões diferentes de PAUM. O principal parâmetro utilizado para a avaliação do desempenho é a razão de classificação, definida como a razão entre o número de padrões classificados corretamente e o número total de padrões. Os melhores resultados foram obtidos com uma rede com 8 e 12 neurônios na primeira e segunda camada intermediária, respectivamente, tendo sido obtido valores da ordem de 90% de acerto. O método estabelecido parece-nos bastante robusto e com grandes chances de aplicações clínicas.

O trabalho que analisaremos para a reconstrução do espectro de raios X é de Boone (1990). A rede utilizada foi tipo *perceptron* multicamadas e o algoritmo de treinamento o de propagação reversa. A rede neural foi treinada usando um espectro gerado pelo modelo de Birch-Masrshall, combinado com dados de atenuação calculados a partir do espectro através de integração numérica. Os dados de atenuação calculados constituem a entrada da rede e o espectro corresponde à sua saída. Neste estudo, o alumínio foi utilizado como material atenuante. O espectro foi subdividido em intervalos de 2kevs na faixa entre 10 e 120kev, resultando em 55 níveis de intervalos amostrados. O espectro assim dividido foi então normalizado, com o maior valor sendo 1. A maioria das sessões de treinamento envolveu de 80.000 a 250.000 iterações durante mais de 12 horas num computador Micro Vax II associado a um processador vetorial. A rede foi testada utilizando dados gerados da mesma forma que no treinamento, só que empregando espectros e características de filtragem não presentes nos dados de treinamento. Para análise do desempenho da rede, o espectro reconstruído foi superposto ao espectro original, sendo efetuada uma avaliação visual. Embora na maioria das aplicações poucos parâmetros sejam suficientes para a caracterização de um feixe de raios X, o espectro completo é algumas vezes necessário, como por exemplo na otimização da taxa sinal-ruído em função da dose aplicada a um paciente.

II.3.3 CONTROLE NEUROMUSCULAR

Graupe et alia (1995) utilizaram uma rede neural tipo ART1 para o controle da estimulação funcional elétrica (EFE) em pacientes com danos na medula e com os membros inferiores paralisados. A rede neural proposta discrimina sinais temporais de eletromiografia registrados acima da área injuriada da coluna. O objetivo é, a partir dessa discriminação, ativar funções de locomoção do paciente, através de estimulação funcional elétrica. A escolha da rede ART1 deveu-se ao fato da mesma poder aprender novos padrões sem necessariamente ter de modificar os demais padrões já aprendidos (plasticidade). Os padrões de eletromiografia podem variar gradualmente, devido a fadiga ou outras mudanças nas condições do paciente, exigindo um reaprendizado constante por parte da rede neural.

Durante a fase de treinamento da rede, quatro categorias básicas são criadas, correspondendo as seguintes funções de locomoção: dar um passo para a esquerda, dar um passo para a direita, permanecer em pé e sentar-se. Uma entrada da rede correspondendo a uma dessas situações é

composta por seis sinais temporais e dois sinais de energia, com dez coordenadas cada. Três sinais temporais são coletados através de um eletrodo no lado direito do tórax, enquanto que os outros três sinais são coletados através de um eletrodo no lado esquerdo do tórax. A entrada da rede constituiu-se então em uma matriz de 8x10.

A saída da rede identifica uma das quatro funções de locomoção desejada pelo paciente e ativa um módulo onde são geradas as funções para o controle da estimulação funcional elétrica. A arquitetura proposta pelos autores possibilita uma iteração do paciente com o sistema. Duas situações onde essa iteração se faz necessária são mencionadas a seguir: a primeira delas quando nenhuma função de estimulação elétrica é gerada a despeito do paciente ter desejado realizar uma determinada função e quando uma função de estimulação elétrica é erroneamente gerada. Em ambos os casos um controle manual do paciente realiza as devidas correções. O sistema proposto pelos autores foi aprovado para uso por paraplégicos nos Estados Unidos em abril de 1994.

Concluindo esse tópico de aplicações de redes neurais à área médica observamos que os trabalhos sobre aplicação de redes neurais que utilizam imagens não fazem uso de informação geométrica como dado de entrada da rede. Todos baseiam-se apenas em informações derivadas da intensidade do pixel. Os pixels são tratados apenas como uma lista de valores, sem nenhuma relação espacial entre eles. No trabalho de tese que ora apresentamos mudamos essa abordagem e incluímos explicitamente informações geométricas ao trabalharmos com os pixels.

II.4 IMAGENS MÉDICAS

Os processos de geração de imagens médicas permitem aos profissionais envolvidos com a área de saúde visualizarem as estruturas e funções internas de um organismo vivo. Atualmente há um rol de modalidades de imagens disponível, entre as quais abordaremos: Raios X, Ressonância Magnética Nuclear (MR), Ultra-Som e Medicina Nuclear.

II.4.1 IMAGEM DE RAIOS X

Uma imagem típica de raios X é produzida projetando-se um feixe de raios X através do corpo de um paciente, sobre um filme sensível aos raios X. Ao se observar uma radiografia (imagem produzida através desse processo) do tórax, certos detalhes anatômicos são imediatamente aparentes. As costelas, por exemplo, aparecem como estruturas claras, pois elas atenuam o feixe de raios X mais fortemente que os tecidos moles adjacentes, de tal forma que o filme fotográfico é menos exposto na região correspondente à sombra das costelas. Na figura 2.10 mostramos uma radiografia típica do tórax. A radiografia digital aplica os mesmos princípios, porém as imagens são produzidas digitalmente, sendo armazenadas na memória do computador, ao invés de serem registradas em filme.

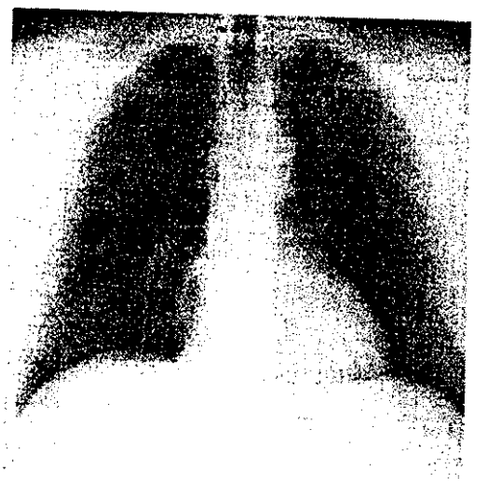


Figura 2.10:Imagem de raios X do tórax.

A imagem produzida pelos raios X apresenta dois inconvenientes:

- é inadequada para distinguir tecidos moles. Os filmes sensíveis aos raios X permitem visualizar contrastes da ordem de 2% (Webb, 1988), o que não é suficiente para distinguir o sangue nos vasos sangüíneos e outros detalhes de tecidos moles, como por exemplo a anatomia do cérebro.

- não fornece informações de profundidade, ou seja, as estruturas tridimensionais são projetadas no filme bidimensional.

A fim de solucionar o problema de superposição das estruturas, desenvolveu-se a radiografia de contraste, que consiste da utilização de contrastes radiopacos que evidenciam áreas de interesse do corpo. Uma importante aplicação desse princípio é a angiografia de subtração digital, utilizada para visualização dos vasos sangüíneos. Nessa técnica são obtidas inicialmente duas imagens dos vasos sangüíneos: uma antes da injeção do contraste e outra após a injeção do contraste no paciente. A imagem final, mostrada na figura 2.11, é obtida subtraindo-se a imagem após a injeção do contraste da imagem anterior a injeção do contraste.

Uma solução genérica para resolver o problema da superposição de estruturas é a tomografia computadorizada (CT). Com essa técnica uma fatia plana do corpo é definida e um feixe de raios X é projetado em uma direção paralela e contida no plano da fatia, conforme mostrado na figura 2.12. Nenhuma parte do corpo fora da fatia é atravessada pelo feixe de raios X. A imagem de CT é como se uma fatia de alguns milímetros de espessura fosse retirada do corpo e então radiografada passando-se um feixe de raios X perpendicularmente à mesma. As imagens resultantes mostram a anatomia do corpo em fatias com aproximadamente 1mm de espessura e com discriminação de densidade melhor do que 1%.

Ao invés de apresentar a imagem como resultado de um parâmetro medido diretamente (a absorção de raios X quando eles passam pelo corpo), a técnica de CT apresenta uma imagem funcional, reconstruída a partir dos dados de atenuação do feixe de raios X, medidos em diferentes

Figura 2.11: Imagem de angiografia de subtração digital.

ângulos de projeção, conforme mostrado na figura 2.12. As técnicas de reconstrução classificam-se em dois grandes grupos: métodos de convolução e projeção reversa e métodos iterativos. Cada um desses métodos apresenta suas vantagens. Os métodos iterativos, por exemplo, encontram grande aplicação em tomografias computadorizadas de emissão, onde a contagem de fótons é baixa. A técnica tomográfica é também aplicada para as outras modalidades de imagens estudadas nesta seção. Na figura 2.13 mostramos uma imagem de CT do coração. Os contornos dos ventrículos esquerdo e direito traçados através de um método automático são também mostrados (Philip et alii, 1994). O corte tomográfico foi obtido perpendicularmente ao eixo maior do ventrículo esquerdo. Uma característica das imagens de CT do coração assim obtidas é que o ventrículo esquerdo aparece nitidamente, porém, como o corte tomográfico abrange toda a região peitoral, ele contém muita informação não relevante, tornando difícil a segmentação do ventrículo esquerdo.

Devido ao fato do coração ser uma estrutura em movimento, o processo utilizado para a aquisição de imagens do mesmo é um processo dinâmico, conforme será mostrado no item II.4.4.

II.4.2 IMAGEM DE ULTRA-SOM

Para se obter imagens de ultra-som (US) das estruturas do corpo, utiliza-se pulsos de freqüência ultra-sônica. À medida em que as ondas sonoras encontram as interfaces entre os tecidos no interior do corpo, uma parte da onda é refletida e outra parte continua a propagar-se. O tempo requerido para que o eco retorne é proporcional à distância dentro do corpo na qual ele é refletido. A amplitude da onda refletida depende das propriedades acústicas dos tecidos encontrados e é representada na imagem como informação de brilho. O aparelho de ultra-som constrói imagens bidimensionais mostrando o eco de múltiplos caminhos adjacentes e unidimensionais. Essas imagens podem ser armazenadas em memória digital ou em videoteipe.

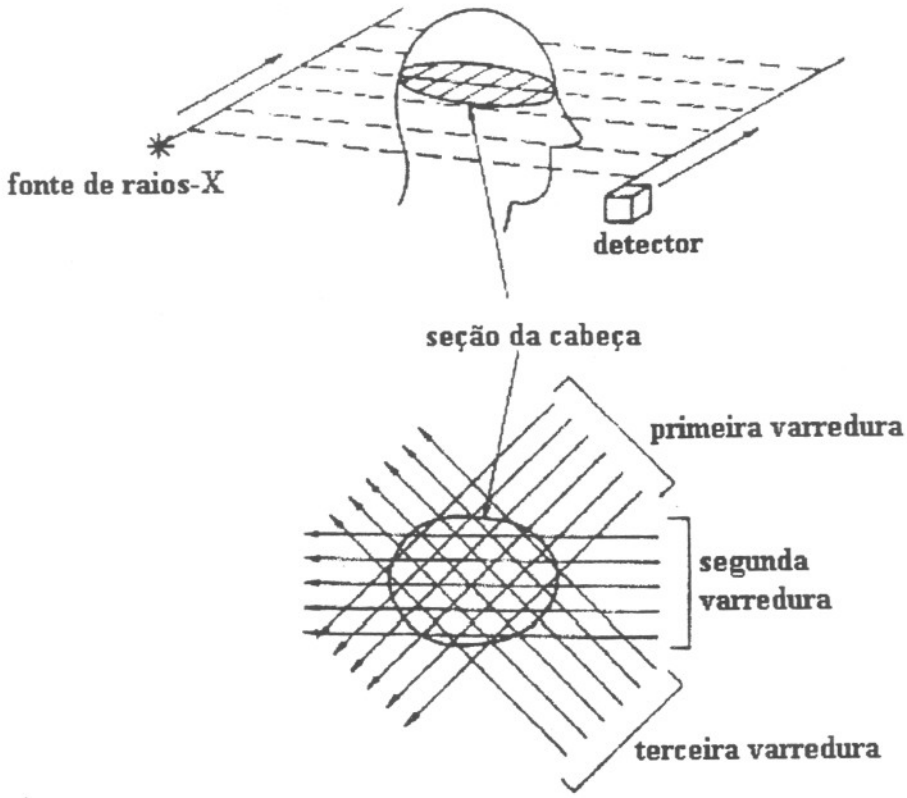


Figura 2.12:Obtenção da imagem de CT.

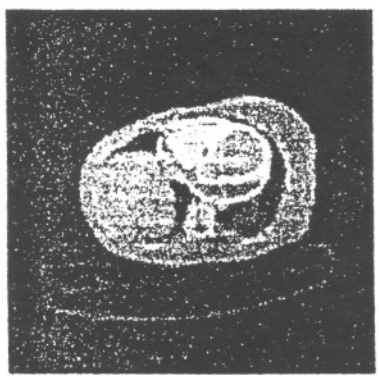


Figura 2.13:Corte tomográfico do coração.

As imagens de ultra-som são mais atrativas como ferramenta clínica do que outras modalidades de imagens que expõem o paciente a altos níveis de radiação nuclear ou de raios X. Outras vantagens das imagens de ultra-som são: baixo custo, mínimo desconforto para o paciente e processamento em tempo real. As imagens de ultra-som são utilizadas como ferramenta de diagnóstico em várias modalidades médicas. Na obstetrícia, em particular, o seu uso é de extrema importância, uma vez que o ultra-som é certamente a mais segura e menos invasiva das técnicas de obtenção de imagens. O ultra-som é considerado seguro para a aquisição de imagens fetais, onde encontra aplicações na confirmação de gravidez, na avaliação da idade gestacional, e no diagnóstico de malformação. Na figura 2.14 mostramos uma imagem de ultra-som de um feto com 5 semanas de desenvolvimento.

Outra importante área de aplicação do ultra-som é a cardiologia. Para a avaliação do sistema cardiovascular utiliza-se tanto o sistema de varredura tipo M como o tipo B. A varredura tipo M é especialmente adequada para a investigação no movimento de válvulas e das paredes cardíacas. A varredura tipo B é utilizada, entre outras aplicações, para caracterização da estrutura e da função

Figura 2.14:Imagem de ultra-som de um feto com 5 semanas.

do VE. Na figura 2.15 mostramos uma imagem de ultra-som do VE resultante de uma varredura tipo B, com os contornos do endocárdio e do epicárdio delineados. O interior do VE aparece nessa imagem como uma grande área escura. A detecção das bordas do endocárdio e do epicárdio nessa modalidade de imagem é uma tarefa difícil, devido a pobre qualidade das imagens, resultante das limitações intrínsecas das mesmas: baixo contraste e ruído especular.

II.4.3 IMAGEM DE RESSONÂNCIA MAGNÉTICA

Um campo recente de desenvolvimento em imagens médicas são as imagens de ressonância magnética (MRI). Seu princípio de operação baseia-se nas propriedades magnéticas do núcleo atômico. Todos os núcleos dos átomos têm um momento magnético, o que significa que eles agem

Figura 2.15:Imagem de ultra-som do coração.

como pequenos ímãs. Quando se coloca um paciente na presença de um campo magnético intenso, os núcleos dos átomos do corpo alinham-se na direção deste campo, rotacionando em torno de um eixo com uma freqüência que é função do tipo de núcleo e da intensidade do campo magnético. Ao aplicar-se um pulso de rádio freqüência numa direção perpendicular à direção do campo magnético estacionário, aqueles núcleos com uma freqüência igual àquela do pulso de rádio freqüência irão entrar em ressonância com o pulso e absorver energia. O estado de energia mais alto fará com que o núcleo mude a orientação de rotação em relação ao campo magnético fixo. Quando o pulso de rádio freqüência é removido, o núcleo volta ao seu estado de alinhamento original, emitindo um sinal de rádio freqüência que é detectado por bobinas especiais. A intensidade e duração deste sinal depende de vários parâmetros relacionados com a densidade e o tipo de núcleo. Tais diferenças podem ser medidas e esses valores podem ser utilizados para gerarem imagens que mostram diferenças biológicas entre tecidos.

A principal característica das imagens de MRI é o potencial de realizarem uma análise química não invasiva de regiões do corpo. Átomos diferentes respondem diferentemente à aplicação de pulsos de rádio freqüência. Um dos átomos mais fáceis de ser visualizado em MRI é o hidrogênio, que também está entre os mais comuns encontrados no corpo humano. Como a densidade de hidrogênio varia em função do tecido (Mansfield et alia, 1978), é possível caracterizar os tecidos, podendo-se em muitos casos diferenciar entre tecidos normais e cancerígenos (Stack et alia, 1990).

Na figura 2.16 mostramos duas imagem de MRI obtidas próximas à base do coração. O interior do VE nessa modalidade de imagem aparece brilhante. A segmentação de imagens cardíacas em MRI é um problema difícil, pois as intensidades em tais imagens dependem da velocidade do sangue (Schutthes, 1989). Embora o sangue nas câmaras cardíacas apareça como regiões brilhantes, ele não tem intensidade homogênea. Como resultado, a transição do sangue para o músculo do miocárdio é, em muitos casos, ambígua.

Figura 2.16: Imagem de ressonância magnética nuclear do coração.

II.4.4 IMAGEM DE MEDICINA NUCLEAR

Em medicina nuclear (MN) as imagens são geradas a partir da utilização de fármacos previamente associados com isótopos radioativos, formando um radiofármaco. Quando administrados a um indivíduo esses radiofármacos apresentam uma biodistribuição preferencial, ou seja, irão se fixar em órgãos específicos do corpo humano. Os isótopos emitem radiação localmente, sendo essa radiação detectada utilizando-se uma gama. câmara. A imagem resultante de medicina nuclear mostra o nível de radioatividade em cada ponto. Diferentemente da técnica de raios X, a radiação não atravessa o corpo, mas é gerada nele mesmo. As imagens produzidas em medicina nuclear podem ser de quatro tipos: estáticas, dinâmicas, dinâmicas cíclicas e tomográficas.

Nas imagens estáticas a quantidade de radiação incidente em cada ponto do detector é contada em um intervalo de tempo fixado e armazenada em um registro. Nas imagens dinâmicas, uma seqüência contínua de imagens estáticas é adquirida em intervalos de tempos iguais. Os estudos dinâmicos são utilizados para monitorar as mudanças no tempo de distribuição do radiofármaco em um órgão específico. Nos estudos dinâmicos cíclicos o objetivo é obter imagens funcionais do ciclo cardíaco através da sincronização da aquisição de imagens cintilográficas com o sinal eletrocardiográfico (ECG). O ECG proveniente do paciente é enviado ao computador, que divide o intervalo entre duas ondas R consecutivas do complexo QRS em quadros (16, 32 ou 64). O intervalo R-R corresponde ao ciclo cardíaco, ou seja, compreende as fases de diástole e sístole do músculo cardíaco. Na figura 2.17 ilustramos um processo de aquisição sincronizado com o eletrocardiograma, através da subdivisão do intervalo R-R em 32 quadros. Nesta figura são mostrados dois ciclos cardíacos. Os marcadores R_1, R_2 e R_3 indicam o início de um novo ciclo de aquisição.

A cada intervalo R-R temos a produção de mais radiação em cada quadro, ocorrendo uma somatória de contagens detectadas em cada intervalo para a obtenção da imagem final.

Conforme mencionado no capítulo 1, as imagens de medicina nuclear do coração podem ser de dois tipos: imagens de perfusão do miocárdio (imagem do músculo cardíaco) e imagens de blood-pool (imagem do sangue nas cavidades cardíacas). Nas figuras 1.1(a) e 1.1(b) mostramos esses dois tipos de imagens.

As imagens médicas podem ser caracterizadas por alguns parâmetros que expressam as suas qualidades. Os parâmetros mais úteis são a resolução espacial e a resolução de profundidade. A resolução espacial está associada com a capacidade de uma modalidade de imagem distinguir pontos

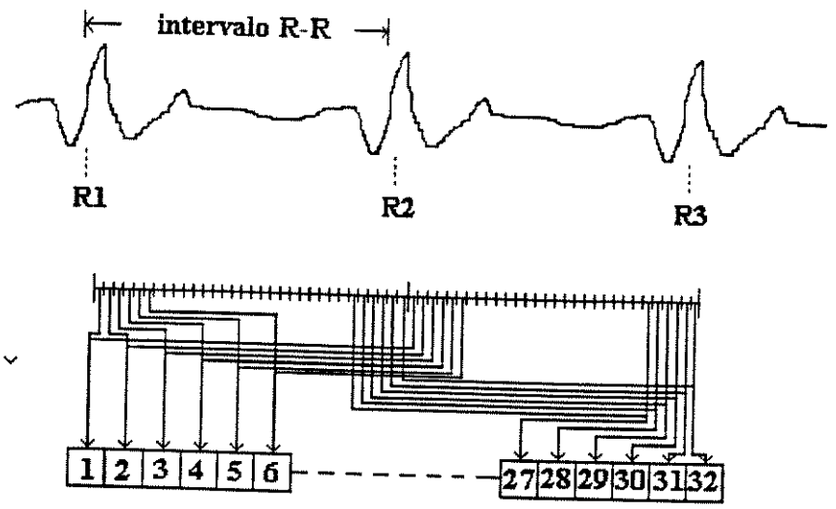


Figura 2.17: subdivisão do intervalo R-R em 32 quadros para aquisição de imagens dinâmicas sincronizadas.

do objeto que estão muito próximos. Para as imagens digitais a resolução espacial é determinada pelo número de pixels na área da imagem. A resolução de profundidade é uma medida da habilidade de uma modalidade de imagem distinguir pequenas diferenças de intensidade. Para as imagens digitais, o número de *bits* por pixel determina a resolução de profundidade. Na tabela 2.1 comparamos esses parâmetros para as diversas modalidades de imagem estudadas. Como pode ser visto, a modalidade de medicina nuclear apresenta os menores valores de resolução espacial e de profundidade. A dificuldade em se detectar o contorno do VE em imagens de medicina nuclear está associada a esses baixos valores de resolução espacial e de profundidade, bem como ao ruído presente nas mesmas (Duncan, 1987).

Tabela 2.1: Parâmetros das modalidades de imagens estudadas

	Modalidade de imagem					
Parâmetro	CT	MRI	US	MN	Raio-X digital	Raio-X analógico
Pixels por imagem (unidimensional)	(256-512)	256	512	(64-128)	1000	2000
Bits por pixel	11--13	10	8	8	12	14
Resolução espacial	média	baixa	média	baixa	alta	alta
Resolução de profundidade	alta	baixa	baixa	baixa	alta	alta

II.5 MÉTODOS DE DETECÇÃO DO CONTORNO DO VENTRÍCULO ESQUERDO

A detecção de contornos do ventrículo esquerdo (VE) em imagens médicas constitui-se em um problema relativamente complexo. Esta complexidade advém de diversos fatores, entre os quais ressaltamos:

- os ventrículos diferem significativamente de paciente para paciente quanto ao tamanho, forma e orientação.

- devido ao baixo contraste em algumas regiões, a borda do VE pode-se tornar praticamente invisível. Em imagens de tomografia computadorizada por raios X, por exemplo, com 256 níveis de cinza, a diferença de intensidade entre o sangue e a parede do ventrículo é da ordem de apenas 7-10 níveis em imagens de boa qualidade e 3-5 níveis na maioria das imagens (Taratorin, 1993).

- o ruído na imagem algumas vezes têm a mesma amplitude do contraste da borda, dificultando a localização da mesma. O problema de ruído é particularmente grave em imagens de ultra-som, devido principalmente à reverberação e ao espalhamento que ocorre na onda de ultra-som, quando a mesma colide com as pequenas microestruturas dos tecidos com dimensões próximas da resolução axial do aparelho (Fish, 1990).

- a intensidade com que o músculo papilar e órgãos sobrepostos ao VE aparecem na imagem é da mesma ordem que a intensidade da parede do mesmo, dificultando a localização da borda nestas regiões.

Devido aos motivos anteriormente expostos, uma abordagem simples, baseada apenas em técnicas convencionais de detecção de borda, não é suficiente para segmentar o VE. Os sistemas de visão computacional utilizados para a segmentação do VE caracterizam-se por apresentarem uma estrutura como mostrada na figura 2.18.

Na tabela 2.3, no final do capítulo, mostramos uma lista dos artigos revisados sobre segmentação do VE. Nesta tabela procuramos salientar as técnicas utilizadas na implementação das etapas mostradas no diagrama em blocos da figura 2.18.

A etapa de pré-processamento da imagem é realizada visando atingir os seguintes objetivos: eliminação de ruídos e artefatos; aumento do contraste na região de interesse e aumento do tamanho da imagem através de interpolação. A remoção de ruídos é efetuada através de filtros passa baixa espaciais e temporais. Os principais tipos são: filtro de média (Grattoni et alia, 1985), filtro gaussiano (Goshtasby et alia, 1995) e filtro da mediana (Adam et alii, 1987 e Lamberti et alia, 1990). O tamanho das máscaras varia de 3x3 até 41x41. A utilização de máscaras maiores em imagens de ecocardiografia é justificado por Chu et alii (1988) pelo fato do sinal da borda nesta modalidade de imagem ser fraco, sendo necessário mais informação a fim de se tomar uma decisão sobre sua localização. As desvantagens de se usar máscaras grandes é que bordas próximas podem ser unidas, resultando em bordas deslocadas.

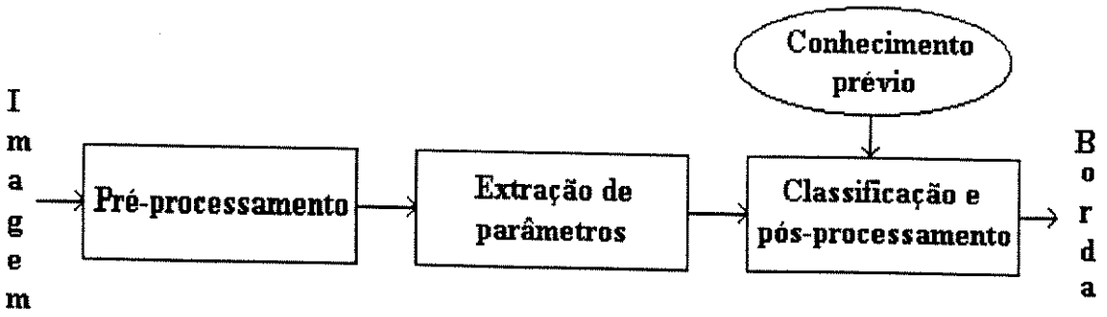


Figura 2.18: Diagrama em blocos de um sistema de visão computacional para detecção de contornos de VE.

O aumento do contraste na região de interesse é normalmente realizado através de uma operação de *window and level* (Fleagle et alii, 1990). O principal objetivo é ressaltar os tons de cinza da imagem situados nesta região.

O aumento da imagem através de interpolação atende a dois propósitos: remoção adicional de ruídos de alta freqüência e adequação da imagem a algoritmos de detecção de borda que operam melhor quando são fornecidos contornos contendo um maior número de pontos. Um exemplo disso pode ser encontrado em Duncan (1987), onde, em imagens de ecocardiografia, uma região de interesse de 32x32 pixels contendo o VE é ampliada para 64x64 pixels. A razão para isso é que o algoritmo utilizado pelo autor para quantificação regional do movimento da parede é mais eficiente com um contorno contendo de 100-150 pixels.

A etapa intermediária de extração de parâmetros, ou processamento de baixo nível, é realizada através de técnicas clássicas de processamento de imagem. O objetivo é extrair parâmetros espaciais ou temporais, que caracterizem a borda do ventrículo. Como mostra a tabela 2.1 as operações espaciais mais freqüentemente utilizadas para extração de parâmetros são: aplicação do gradiente, aplicação do limiar, obtenção dos cruzamentos por zero resultantes da aplicação do laplaciano e determinação do centro do VE.

Em relação aos eventos temporais, observamos que Jouan et alii (1990) utilizaram, para imagens de ecocardiografia, médias temporais de uma seqüência de imagens do mesmo ciclo cardíaco. O objetivo foi o de minimizar tanto os efeitos de anormalidades do movimento da parede cardíaca como os efeitos da contagem estatística durante o processo de extração da imagem. Uma alternativa à utilização de imagens de um mesmo ciclo cardíaco para obtenção de médias, é a utilização de imagens de ciclos cardíacos diferentes, obtidas através de imagens dinâmicas cíclicas (Skorton et alii, 1981).

A etapa de classificação e de pós-processamento, ou processamento de alto nível, agrupa as informações extraídas anteriormente com o objetivo de localizar a borda do VE. Em geral, procura-se identificar pixels ou segmentos com maior probabilidade de pertencerem a borda do VE.

UNICAMP
BIBLIOTECA CENTRAL

Através da revisão da literatura pareceu-nos claro que um método eficaz de detecção da borda do VE não pode repousar exclusivamente em informações derivadas da intensidade do pixel. Adicionalmente, é necessário um conhecimento prévio de parâmetros geométricos e/ou anatômicos do VE, bem como uma forma apropriada de introduzi-los no algoritmo de detecção da borda. No diagrama da figura 2.18 o conhecimento prévio é representado através de círculo acima do bloco de classificação e pós-processamento.

O rol de técnicas utilizadas na etapa de classificação e pós-processamento é bastante grande. Entre essas técnicas citamos: técnicas que utilizam procura em grafos (Thedens et alii, 1995; Pope et alii, 1985; Fleagle et alii, 1991, Moura et alia, 1992); técnicas que utilizam funções probabilísticas (Clayton et alii, 1974; Barret et alii, 1980; Dias et alia, 1996); técnicas que utilizam *simulated annealing* (Friedland et alia, 1989); técnicas que utilizam sistemas *fuzzy* (Feng et alii, 1985; Boudraa et alii, 1993); técnicas que utilizam ortogonalização de Gram-Schmidt (Jouan et alii, 1990) e técnicas que utilizam redes neurais (Coppini et alii, 1995).

Um artigo clássico que exemplifica o sistema de visão computacional delineado na figura 2.18 foi apresentado por Ezequiel et alii (1985). O objetivo deste é a obtenção da borda do endocárdio em imagens de ecocardiografia.

O algoritmo de detecção de bordas utilizado pelo autor consiste dos três estágios mostrados na figura 2.18. Na etapa de pré-processamento, com o objetivo de aumentar a precisão estatística de cada pixel, as imagens são suavizadas de forma espacial e temporal. A suavização temporal é efetuada substituindo-se cada pixel por uma média dele mesmo e dos pixels situados na mesma coordenada cartesiana, um quadro à frente e um quadro atrás no ciclo cardíaco, em um processo não recursivo. A suavização espacial é obtida substituindo-se cada pixel por uma média ponderada dos oito pixels vizinhos. A suavização espaço-tempo substitui, então cada pixel por uma média dele mesmo e de 26 outros pixels.

No processamento de baixo nível, o centro do VE é determinado utilizando-se informações de histogramas de amplitude. Após a determinação do centro subtrai-se de toda a imagem um nível de cinza correspondente ao valor de uma máscara de 3x3 pixels centrada no mesmo. A idéia explorada é que na região do centro espera-se um nível de cinza de valor zero. Assim, qualquer nível de cinza na região do centro é interpretado como ruído e assume-se está presente em toda a imagem. Devido ao fato dos pixels da borda do endocárdio serem caracterizados por uma baixa intensidade e os pixels externos terem uma intensidade maior, os autores extraem os pixels de borda associando o operador laplaciano à amplitude do pixel.

Os pixels da borda do endocárdio são identificados através de uma procura radial utilizando 32 raios igualmente espaçados. As regiões da borda do endocárdio ao longo de um raio são caracterizadas por apresentarem uma alta inclinação positiva no gráfico da amplitude versus raio, combinada com um cruzamento por zero resultante da aplicação do laplaciano. Com o objetivo de identificarem falsos pixels de borda ao longo de um raio, os autores calculam, para cada pixel de borda identificado, uma função de custo. Essa função de custo considera a semelhança entre as características de um pixel de borda e as características dos pixels de borda situados proximamente. Os termos da função de custo comparam os valores de distância radial, intensidade de pixel e valor do laplaciano de um pixel de borda com os valores dos seus vizinhos. Pixels de borda com valores

similares aos de seus vizinhos têm um baixo valor calculado para a função de custo e uma alta probabilidade de pertencerem a borda do VE. Após a identificação do pixel de borda "mais provável" em cada um dos 32 raios, os autores, com o objetivo de dobrarem o número de pixels do contorno (de 32 para 64), utilizam interpolação linear através de polinômios cúbicos. A borda do endocárdio é então determinada unindo-se os 64 pixels de borda. Os autores não comparam os contornos obtidos automaticamente com contornos traçados manualmente.

Os métodos de detecção da borda do VE podem ser divididos em três grandes grupos: manuais, semi-automáticos e automáticos. Na última década observa-se um grande esforço no sentido de obter-se métodos automáticos. Entre os principais motivos para este esforço citamos:

- embora o método manual seja o mais robusto, ele é tedioso, demorado e apresenta grande variabilidade intra e inter-observador (Thedens et alii, 1995).

- a extração manual consome tempo, limitando a quantidade de dados que pode ser processada. Conseqüentemente, a reconstrução 3-D da forma do VE, que depende da detecção do contorno do mesmo em seqüências de imagens, torna-se uma tarefa inexeqüível em tempo hábil.

Os métodos manuais datam da década de 70, quando ainda havia muita precariedade da aparelhagem utilizada no processamento de imagens médicas (quando comparado aos equipamentos disponíveis atualmente). Desta época pinçamos o trabalho de Alderman et alii (1973), onde as imagens são gravadas num vídeo disco com uma capacidade de 400 quadros, numa taxa de 15 quadros/segundo. O monitor de vídeo mostra as imagens num monitor de vídeo dinamicamente ou estaticamente. Os contornos do VE são extraídos utilizando-se uma caneta ótica na tela do monitor de vídeo. A partir deste contorno, um programa faz a estimativa da área e do volume do VE. A grande vantagem do método proposto, de acordo com os autores, é que com o vídeo disco as imagens não precisavam mais ser gravadas em filme, o que permitia uma análise da imagem logo após a sua obtenção. A ênfase destes primeiros trabalhos é mais na descrição do equipamento em si do que no método de detecção do contorno.

Os métodos semi-automáticos dependem de uma interação com o operador. Esta interação pode ser de três tipos:

- o operador traça o contorno manualmente na diástole e o método detecta o contorno nos outros instantes do ciclo cardíaco, utilizando o contorno inicial como referência (Adam et alii, 1987).

- o operador identifica pixels característicos na imagem, como o centro do VE e a posição das válvulas mitral e aórtica. (Fleagle et alii, 1991 e Lilly et alii, 1989).

- o operador delimita uma área de procura do contorno do VE (Clayton et alii, 1974) ou seleciona um limiar ótimo para segmentação do VE (Yang et alia, 1994).

Todas essas interações introduzem um certo grau de subjetividade na determinação do contorno do VE, que levam a uma variação intra e inter-observador. Acreditamos, no entanto, que mais importante do que analisar os métodos de detecção do contorno do VE sob a ótica da

automatização é analisá-los considerando os algoritmos utilizados no processamento de alto nível, os quais realmente diferenciam um método do outro.

Uma das abordagens de alto nível com excelentes resultados na tarefa de detectar o contorno do VE utiliza procura dinâmica em grafos. De uma forma elegante, esta abordagem incorpora no algoritmo informações locais e globais para determinação do contorno ótimo. A utilização de grafos para detecção de contornos é ilustrado na figura 2.19. Um grafo é constituído por nós (pixels) e ligações. Nesta abordagem, a suavização do contorno é introduzida limitando-se o número de nós na próxima coluna a que um nó de uma coluna anterior pode se ligar. Na figura 2.19 este número é limitado aos 3 nós mais próximos. Duas colunas têm um significado especial em um grafo: a coluna inicial e a coluna final. Um caminho em um grafo é definido como um conjunto de ligações que conecta um nó da coluna inicial a um nó da coluna final. Um valor de custo é associado a cada nó. Em geral, quanto menor o custo de um nó, maior a probabilidade do mesmo pertencer ao contorno. O custo de um nó em problemas de detecção de contorno é inversamente proporcional ao valor resultante da aplicação de um detector de borda de baixo nível, como o gradiente. A cada caminho em um grafo corresponde também um custo, que é o resultado da soma dos custos dos nós individuais associados a este caminho. No problema de detecção de contorno, o mesmo é associado ao caminho com menor custo, designado de caminho ótimo.

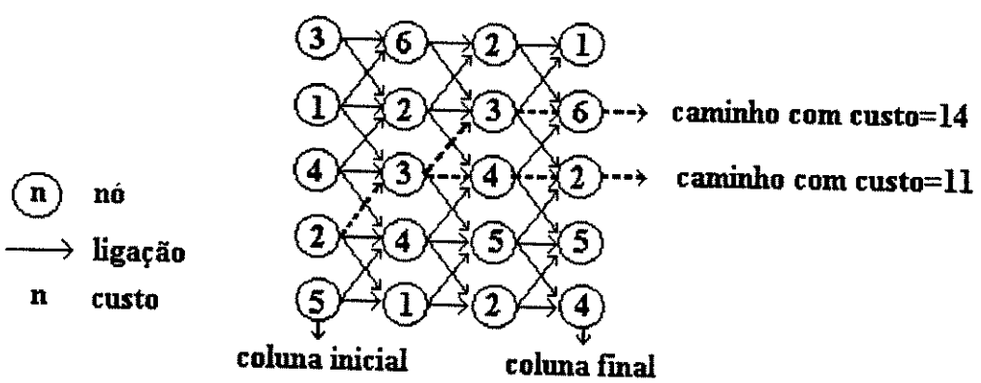


Figura 2.19: Detecção do contorno utilizando procura em grafos.

Uma excelente aplicação da metodologia anterior pode ser encontrada em Pope et alii (1985). Neste trabalho, os autores determinam inicialmente as coordenadas do centro do VE (xc,yc). A partir do centro traçam-se, então raios em direção a borda do VE. Cada um desses raios define uma das linhas de uma matriz de borda definida como:

$$\text{matriz}(\rho,\theta) = \text{imagem}(x,y) \tag{2.14}$$

onde: $\rho = ((x - xc)^2 + (y - yc)^2)^{1/2}$

$\theta = \arctan((y - yc)/(x - xc)) + \theta_0$

θ_0 = ângulo de ajuste que permite a extração da matriz começar na região da válvula aórtica.

O detector de bordas utilizado para atribuir custos aos pixels é um filtro radial casado, K(ρ), cujos pesos foram empiricamente determinados através de vários perfis de bordas do VE e consiste de nove números: 6/5,5/5/3/0/-3/-5/-5,5/-6. Através de uma operação de convolução unidimensional deste filtro com a matriz de borda encontra-se a matriz de probabilidade, definida como:

$$\text{probabilidade}(\rho,\theta) = \text{matriz}(\rho,\theta) * K(\rho) \tag{2.15}$$

Efetuando-se, então uma procura dinâmica em grafos sobre a matriz de probabilidade, determina-se o contorno do VE.

A estratégia de determinação do contorno do VE através de uma procura radial é encontrada com freqüência na literatura. Os pontos do contorno são selecionados em cada raio utilizando-se funções de custo (Lilly et alii, 1989), intersecção de conjuntos de restrições (Taratorin et alia, 1993) ou técnicas fuzzy (Feng et alii, 1991).

Uma reconstrução da borda baseada em procura radial requer que o algoritmo proposto seja capaz de estimar a borda em regiões onde a mesma não aparece claramente e em regiões em que o contorno de órgãos vizinhos se sobrepõem ao contorno do VE. Em Lilly et alii (1989) este problema é solucionado através de duas abordagens: a primeira delas é a introdução de termos na função de custo que levam em conta a proximidade de um pixel de borda num determinado ângulo com os pixels de borda de ângulos adjacentes. A segunda delas é a correção do contorno final obtido através da comparação com modelos de contornos de VE previamente calculados. Em Taratorin et alii (1993) existem vários conjuntos de restrições que procuram solucionar este problema. Um desses conjuntos assume que o contorno do VE não pode passar em algumas regiões, como por exemplo, a região de baixa intensidade de pixels do pulmão. Outro conjunto de restrições assume que a borda seja suave, ou seja que a mesma não pode variar bruscamente. Em Feng et alii (1991) limita-se, para cada ângulo, a região de escolha dos pixels de contorno em função da proximidade de pixels de contornos previamente determinados em ângulos anteriores.

A introdução de informação redundante de uma seqüência espacial ou temporal de imagens pode melhorar substancialmente a qualidade do contorno determinado através de uma técnica automática de detecção de contorno do VE. A minimização de ruídos na imagem através da média de uma seqüência temporal de imagens é, conforme já citado, uma maneira de explorar essa redundância. A seguir exploraremos alguns métodos que fazem uso de informação temporal para melhorar a qualidade dos contornos de VE detectados.

Zhang et alia (1982) detectaram contornos de VE em imagens ecocardiográficas através da aplicação de limiar em áreas de movimento na imagem. Um limiar ótimo é determinado para uma imagem a partir de uma matriz de co-ocorrência temporal, onde o elemento (i,j) representa o número de pixels que em um quadro do ciclo cardíaco tinham uma intensidade *i* e no quadro seguinte tinham uma intensidade *j*. O limiar ótimo *L* é definido como aquele que maximiza o valor de P(L) dado por:

$$P(L) = \sum_{i=1}^{L} \sum_{j>L}^{k} M_{ij} + \sum_{i>L}^{k} \sum_{j=1}^{L} M_{ij} \tag{2.16}$$

onde:

M: matriz de co-ocorrência temporal.

k: nível máximo de quantificação da imagem.

O primeiro termo da equação anterior representa aqueles pixels que num quadro *i* tinham intensidade menor do que *L* e no quadro *i+1* tinham intensidade maior do que *L* . O segundo termo representa aqueles pixels com um comportamento inverso. Os pixels que não se enquadram em nenhum desses casos são considerados estacionários e não pertencem ao contorno. A grande deficiência do algoritmo ora descrito é a não incorporação pelo mesmo de qualquer relação espacial entre os pixels e de nenhum conhecimento prévio de parâmetros geométricos e/ou anatômicos do VE.

Outro esforço no sentido de incluir informação temporal foi efetuado por Friedland et alia (1989), utilizando *simulated annealing*. Os autores utilizam uma procura radial e constróem empiricamente uma função de energia constituída por 4 elementos: U_1, U_2, U_3 e U_4. U_1 é um termo inversamente proporcional ao valor numérico resultante da aplicação de um detector de borda radial ótimo. U_2 é um termo que considera a descontinuidade do contorno. Seu valor numérico é resultante da comparação da distância de um pixel de contorno em relação ao centro do VE e as distâncias dos pixels de contorno de raios adjacentes em relação ao mesmo centro. Quanto mais suave o contorno menor o seu valor. U_3 é um termo inversamente proporcional ao volume da cavidade ventricular. Quanto maior o volume menor o seu valor. U_4 é um termo que considera a descontinuidade temporal dos pixels de contorno. Quanto menor a descontinuidade menor o seu valor. Esse último termo é dado por:

$$U_4 = \frac{\left| r_i^i - r_{i-1}^i \right|}{r_i^i} \qquad (2.17)$$

onde:

r_i^i = valor do raio *i* no quadro *i* .

r_{i-1}^i = valor do raio *i* no quadro *i-1*

Os quatro termos anteriormente citados são combinados numa expressão linear e esta expressão é então minimizada. Este algoritmo incorpora tanto informações espaciais como temporais na detecção da borda. Os autores, no entanto, não apresentam resultados comparativos com contornos traçados manualmente, o que impede uma avaliação definitiva do mesmo.

Uma abordagem utilizada para extrair informação temporal de uma seqüência de imagens é a utilização de fluxo óptico (*optical flow*). Fluxo óptico pode ser definido como a distribuição aparente do movimento de padrões luminosos em uma imagem. A expressão que relaciona as variações temporais e espaciais numa imagem é dada por (Horn et alia, 1981):

$$\frac{\delta I}{\delta x}\frac{\delta x}{\delta t} + \frac{\delta I}{\delta y}\frac{\delta y}{\delta t} = -\frac{\delta I}{\delta t} \text{ ou}$$

$$Ex \,.\, u + Ey \,.\, v = -\, Et \qquad (2.18)$$

onde:

I: intensidade do pixel
x,y: coordenadas espaciais
t: tempo
$Ex = \delta I/\delta x$; $Ey = \delta I/\delta y$; $Et = \delta x/\delta t$; $u = \delta x/\delta t$; $v = \delta y/\delta t$.

O lado esquerdo das expressões anteriores pode ser interpretado como o produto interno do vetor velocidade **V**= ***(u,v)*** pelo vetor gradiente $\nabla \mathbf{I} = (Ex, Ey)$. Computando-se $\nabla \mathbf{I}$ e **Et** a partir dos pixels de uma seqüência de imagens, é possível determinar a componente da velocidade **V** na direção do gradiente $\nabla \mathbf{I}$. É impossível calcular, no entanto, a componente da velocidade perpendicular ao gradiente. Devido a isso, restrições adicionais são colocadas em relação a variação da velocidade ao longo do objeto em movimento. Assume-se que a velocidade varie suavemente ao longo do objeto. Matematicamente, isto é implementado através de um processo de minimização do gradiente das componentes do vetor velocidade **V** ao longo dos pixels da imagem. As equações resultantes expressam as componentes *u* e *v* da velocidade em cada pixel em termos de **Ex, Ey, Et** e de valores médios de **u** e **v** computados na vizinhança de cada pixel. A solução para estas equações é buscada através de métodos iterativos. O valor de $\nabla \mathbf{I}$ é usualmente aproximado por métodos de diferenciação numérica e **Et** é obtido através de uma subtração de imagens de quadros subseqüentes. Ambos são extremamente influenciados pelo ruído, o que requer uma etapa de suavização da imagem.

A combinação do movimento da imagem com informações de contraste da mesma foi utilizada por Thompson et alii (1980) para a segmentação de objetos em movimento. Em imagens de coração, a análise do movimento das paredes do músculo cardíaco constitui-se de um meio valioso para a identificação de isquemia e infarto (Gelfand et alii, 1988). Análises do movimento do coração em imagens ecocardiográficas bidimensionais utilizando fluxo ótico podem ser encontradas em Mailloux et alii (1987), Meunier et alii (1988) e Mailloux et alii (1989).

Os métodos descritos anteriormente têm a característica de tratarem todas as áreas da imagem como tendo o mesmo contraste e nível de ruído. Na verdade, isto não ocorre. Em imagens de *blood-pool* do coração, por exemplo, a borda livre do VE oferece um excelente contraste em relação ao fundo, enquanto que a borda que faz vizinhança com o ventrículo direito é de difícil visualização. Em imagens ecocardiográficas, por outro lado, a borda do endocárdio é facilmente identificada, devido a câmara do ventrículo aparecer bastante escura. A borda do epicárdio, no entanto, é de difícil detecção. Na região inferior do VE as intensidades dos pixels ao se caminhar na direção do epicárdio variam de um brilho médio para um brilho intenso, enquanto que nos lados esquerdo e direito do VE as intensidades variam de escuro para um brilho médio. Este conhecimento é utilizado por Feng et alii (1991) como base para a detecção do contorno do endocárdio e do epicárdio em imagens ecocardiográficas. Este conhecimento é incorporado no algoritmo de detecção do contorno através de uma técnica *fuzzy*.

Apenas na década de 90 é que vamos observar a utilização de técnicas de inteligência artificial, como sistemas *fuzzy* e redes neurais, para detecção do contorno do VE. Na tabela 2.1 observamos os seguintes trabalhos utilizando essas técnicas: Feng et alii (1991), Boudraa et alii (1993) e Coppini et alii (1995). Este último trabalho é de particular interesse, à medida em que foi o

único trabalho que encontramos utilizando redes neurais para a detecção do contorno do VE. A seguir procederemos a uma análise do mesmo.

O objetivo do trabalho de Coppini et alii (1995) é a reconstrução da forma 3-D da superficie do VE a partir da detecção do contorno do mesmo em várias vistas. O nosso interesse nesta análise está centrado na obtenção do contorno do VE. O algoritmo de detecção do contorno pode ser sintetizado em três passos: a) determinação de pixels candidatos do contorno; b) agrupamento dos pixels candidatos em curvas c) aproximação destas curvas por segmentos candidatos e utilização de uma rede neural para classificação destes segmentos como segmentos do contorno ou ruído. A detecção dos pixels candidatos do contorno é realizada de acordo com a teoria de Marr e Hildreth (1980). Após a aplicação do operador laplaciano do gaussiano, os pontos candidatos do contorno são identificados através da detecção dos cruzamentos por zero. Através de um operador gradiente, associa-se a cada pixel candidato do contorno três parâmetros: o módulo do gradiente, *Me* e valores *nx e ny* da direção *ne* do gradiente. Em seguida, os pixels candidatos são conectados formando curvas, através de um algoritmo proposto pelo próprio autor. Na figura 2.20 mostramos o exemplo de uma dessas curvas. Finalmente, através de um procedimento recursivo proposto por Duda e Hart (1973), estas curvas são aproximadas por segmentos de reta, conforme pode ser visto na figura 2.20. São os parâmetros associados a estes segmentos de reta que os autores utilizam como entrada da rede neural. Cada segmento de reta é caracterizado na entrada da rede através de 10 parâmetros: coordenadas *x e y* do centro do segmento; valor médio *Ms* do módulo do gradiente *Me* dos pixels que constituem o segmento; valores médios *nsx e nsy* dos valores *nx e ny* da direção do gradiente dos pixels que constituem o segmento; comprimento do segmento *Ls* e um código binário de quatro dígitos identificando a que vista pertence o segmento.

Figura 2.20: Exemplo de aproximação de curvas por segmentos utilizada no trabalho de Coppini et aia (1995).

A rede neural utilizada é tipo perceptron de 4 camadas, com 20 e 10 neurônios nas camadas intermediárias e um neurônio na camada de saída. Se a saída da rede é maior que um determinado limiar, o segmento é caracterizado como pertencente ao contorno do VE, caso contrário o mesmo é considerado ruído. Um total de 67991 segmentos foram extraídos de 500 imagens ecocardiográficas de 7 pacientes em quatro ângulos diferentes (0°, 45°, 90°e 135°). As imagens de 4 pacientes foram utilizadas para o treinamento, enquanto que as imagens dos outros 3 foram utilizadas para testar a rede treinada. Um especialista classificou, durante 20h, 14485 segmentos extraídos do conjunto de imagens de treinamento como sendo ruído e 1056 como segmentos pertencentes a contornos. Este conjunto de 15541 segmentos foi utilizado para o treinamento da rede, que demorou em média 45h e exigiu um total de 7000 iterações numa estação RISC 6000/320. Na fase de teste, o limiar utilizado para decidir se um segmento pertence ou não ao contorno foi variado na faixa entre 0,1 e 1,0, e o

resultado da classificação de 29139 segmentos foi analisado através de curvas de ROC. O valor ótimo do limiar, estabelecido visando otimizar a especificidade da classificação, foi de 0,8. Com este limiar a rede classificou corretamente 96,57% dos segmentos testados. Entre estes, 99,73% representando ruído e 53,72% de segmentos pertencentes a contornos.

Finalizando a exposição daquele trabalho faremos os seguintes comentários: a) o autor não apresenta dados comparativos com contornos traçados manualmente, o que impede uma análise definitiva do método proposto; b) a etapa de treinamento envolve um esforço considerável por parte de um especialista na classificação dos segmentos (20h). Esta classificação é subjetiva e está sujeita a erros; c) Não existe nenhuma etapa de padronização dos tamanhos dos VE e nem de alinhamento dos seus centros. Em face disso, é surpreendente que a rede neural consiga classificar corretamente um segmento utilizando como referencial geométrico as coordenadas cartesianas do centro do mesmo. Devido às diferentes formas, tamanhos e posições dos VE, essas coordenadas em um determinado VE podem situar-se sobre o contorno, porém, em outro ventrículo diferente, podem situar-se no seu interior. A utilização de coordenadas polares associadas a um alinhamento prévio dos centros dos ventrículos talvez fosse mais apropriado para este tipo de abordagem.

A maioria dos métodos de detecção de bordas mostrados na tabela 2.3 é feita sob medida para uma determinada classe de imagens. Dito de outra forma, métodos que são bons para detectar o contorno do VE em imagens de ecocardiografia, provavelmente não se adequam para detectar o mesmo contorno em imagens de medicina nuclear. Pelo exposto nesta revisão podemos concluir que características um bom método de detecção do contorno do VE deve apresentar:

- detecção automática da borda sem a necessidade de intervenção de um operador.
- capacidade de processamento rápido, a fim de poder servir de base num sistema clínico operacional.
- incorporação de informações espaciais e temporais que explorem a redundância de uma seqüência de imagens, tornando o algoritmo mais confiável.
- incorporação de conhecimento prévio constituído por informações anatômicas e geométricas que permitam estabelecer a continuidade do contorno em regiões de difícil detecção.
- exploração da vizinhança do VE como uma região não uniforme, constituída por várias sub-regiões de contrastes distintos.
- possibilidade de aplicação em modalidades de imagens diferentes.

Ao se buscar um novo método de detecção do contorno do VE deve-se, à medida do possível, englobar um maior número das características anteriormente apresentadas. Um método otimizado seria, a nosso ver, aquele que englobasse todos esses aspectos.

II.5.1 MEDIDAS DE AVALIAÇÃO DE DESEMPENHO

Com o objetivo de comparar o desempenho dos métodos de detecção do contorno do VE , é necessário a definição de uma medida de desempenho universalmente aceita e uma base de dados padrão que possa ser utilizada por todos. Infelizmente, até o momento, tais padrões inexistem.

Em muitos trabalhos (Clayton et alii, 1974; Feng et alii, 1991; Friedland et alia, 1989) o desempenho do algoritmo de detecção do contorno do VE não foi julgado quantitativamente. O julgamento foi feito em termos de uma observação visual, eivada de subjetividade.

Os métodos de quantificação do desempenho dos algoritmos de detecção do contorno do VE podem ser classificados em dois grupos: o primeiro caracteriza-se por comparar os resultados em termos de parâmetros medidos para o VE, como volume ventricular ou fração de ejeção (Boudraa et alii, 1993; Yang et alia, 1994). O segundo envolve uma comparação da área delimitada pelo contorno, ou do comprimento do contorno, ou da posição dos pontos do contorno, com parâmetros correspondentes obtidos através de contornos traçados manualmente (Pope et alii, 1985; Lilly et alii, 1989 e Barret et alii, 1980).

Embora o objetivo de um método de detecção de contorno do VE seja, em última análise, obter informações funcionais do mesmo, como fração de ejeção, um algoritmo automático ou semi-automático de detecção de contorno deve ser avaliado em termos absolutos, a fim de estabelecer a sua precisão.

A seguir mostraremos algumas medidas utilizadas na literatura para avaliar em termos absolutos o desempenho de um método de detecção do contorno do VE. Pope et alii (1980) e Barret et alii (1985) utilizaram como medida de desempenho o desvio médio de pixel ε, que envolve uma comparação pixel a pixel entre o contorno traçado manualmente e o contorno obtido automaticamente:

$$\varepsilon = \sum_{i=1}^{L} \frac{\text{distância entre os pixels i's dos contornos manual e automático}}{2L} \qquad (2.19)$$

onde:

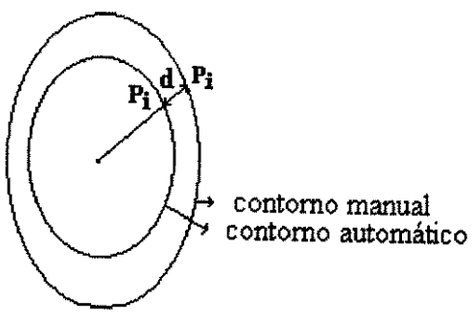
L = número de pontos do contorno

A distância entre os pixels i's dos contornos manual e automático é ilustrada através da figura 2.21.

Pope et alii (1980) obtiveram valores de ε entre 0,33 pixels e 0,52 pixels, enquanto que Barret et alii (1985) obteve valores de ε da ordem de 1,4 pixels.

Outro método reportado na literatura (Barret et alii, 1980) informa a percentagem de pixels do contorno automático que se situam dentro de intervalos de proximidade do contorno traçado manualmente. Por exemplo, os autores determinaram que, em média, para um conjunto de 10 imagens de teste, 87,2% dos pixels do contorno automático situaram-se num intervalo de ±3 pixels

em relação ao contorno manual. Uma vantagem desta abordagem é que ela indica mais precisamente a percentagem de pixels do contorno que precisam de correção.



P_i - i-ésimo pixel de um contorno
d - distância entre o i-ésimo pixel do contorno manual e do contorno automático

Figura 2.21: ilustração da distância entre os pixels i's dos contornos manual e automático para o cálculo do desvio médio de pixel ε.

Duas medidas que comparam valores de área envolvida pelo contorno foram utilizadas por Lilly et alii (1989). A primeira medida, definida como erro percentual, é dada por:

$$\text{erro percentual} = \frac{|M| - |A|}{|M|} \cdot 100 \qquad (2.20)$$

onde:

M e $|M|$ = conjunto de pixels dentro do contorno manual e área correspondente a estes pixels, respectivamente.

A e $|A|$ = conjunto de pixels dentro do contorno automático e área correspondente a estes pixels, respectivamente.

A segunda medida, definida como o erro soma, é dada por:

$$\text{erro soma} = \frac{|M \cap \overline{A}|}{|M|} \cdot 100 + \frac{|\overline{M} \cap A|}{|M|} \cdot 100 \qquad (2.21)$$

onde:

$\overline{A}$ = complemento do conjunto de pixels A.

$\overline{M}$ = complemento do conjunto de pixels M.

A primeira dessas medidas considera apenas a diferença absoluta das áreas definidas pelos contornos manual e automático, enquanto que a segunda delas considera todas as regiões onde não houve superposição das áreas definidas por ambos os contornos. Os melhores resultados obtidos pelos autores para o erro percentual médio num grupo de 25 imagens encontram-se em torno de 6,5% e do erro soma na faixa de 14%.

Tabela 2.2: Resumo dos trabalhos sobre aplicação de redes neurais na área médica

ordem	autor	categoria	1.arquitetura e 2.algoritmo de treinamento	número de nós
1	Baker et alia (1995)	diagnóstico de câncer de mama	1.perceptron multicamadas; 2.propagação reversa	18--10--1
2	Kippenham et alia (1992)	diagnóstico da doença de Alzheimer	1.perceptron multicamadas; 2.propagação reversa	8--4--1
3	Dawson et alia (1994)	diagnóstico da doença de Alzheimer	1.perceptron multicamadas; 2.propagação reversa	14--(0,5,10,15,20)--2
4	Hamilton et alia. (1995)	diagnóstico de perfusão em imagens de miocárdio	1.perceptron multicamadas; 2. propagação reversa	8--4--1
5	Hall et alia (1992)	classificação de tecidos	1.perceptron multicamadas; 2.correlação em cascata ("cascade correlation")	3--2--6
6	Gross et alii (1990)	diagnóstico de doenças pulmonares	1.perceptron multicamadas; 2.propagação reversa	21--15--12
7	Asada et alii (1990)	diagnóstico de doenças intersticiais do pulmão	1.perceptron multicamadas; 2.propagação reversa	20--6--5
8	Scott et alii(1993)	diagnóstico de ventilação-perfusão do pulmão	1.perceptron multicamadas; 2.propagação reversa	25--(10,15,20)--1
9	Fujita et alii (1994)	diagnóstico de doenças das artérias coronarianas	1.perceptron multicamadas; 2.propagação reversa	256--(5,140)--8
10	Chan et alii (1994)	diagnóstico da doença de Alzheimer, diagnóstico de abuso de cocaína	1.perceptron multicamadas; 2.propagação reversa	120--8--1 120--28--4--1
11	Boone (1990)	reconstrução do espectro de raio x	1.perceptron multicamadas; 2.propagação reversa	9--55--55 16--55--55

ordem	autor	categoria	1.arquitetura e 2.algoritmo de treinamento	número de nós
12	Wu et alii (1990)	diagnóstico de câncer de mama	1.perceptron multicamadas; 2.propagação reversa	45--(1-10)--1
13	Clarke et alii (1990)	classificação de tecidos (resumo)	—	—
14	Nelson et alii (1991)	localização da ativação do miocárdio usando ECG (resumo)	—	—
15	Garg et alia (1991)	detecção de nódulos em mama (resumo)	1.perceptron multicamadas; 2.propagação reversa	—
16	[1] Boone et alii (1990)	detecção de nódulos simulados	1.perceptron multicamadas; 2.propagação reversa	25--5--1
17	Ercal et alii (1994)	diagnóstico de melanoma	1.perceptron multicamadas; 2.propagação reversa	14--7--1 8--4--1
18	[1] Hassoun et alia (1994)	extração de padrões em eletromiografia, apresentação de um método	1.perceptron multicamadas; 2.aprendizado pseudo-supervisionado	—
19	[2] Hassoun et alii (1994)	extração de padrões em eletromiografia	1.perceptron multicamadas; 2.aprendizado pseudo-supervisionado	48--6--4--48 48--6--8--48 48--6--12--48 48--8--16--48
20	[2] Boone et alii (1990)	revisão:discute aplicações de redes neurais em radiologia	—	—
21	Scott (1993)	revisão: aborda o uso de inteligência artificial em diagnóstico médico	—	—
22	Miller et alii (1992)	revisão: aborda artigos de aplicação de redes neurais em imagens médicas e processamento de sinais	—	—
23	Moreno et alii (1995)	avaliação da maturação do cérebro usando EEG	1.perceptron multicamadas; 2.propagação reversa	5--5--2
24	Pattichis et alii (1995)	diagnóstico em EMG	1.perceptron multicamadas, rede auto-organizável de Kohonen; 2.propagação reversa	14--(10,40,100)--(5,10)--3 6x6,10x10,12x12

ordem	autor	categoria	1.arquitetura e 2.algoritmo de treinamento	número de nós
25	Graupe et alia (1995)	controle da estimulação elétrica funcional em paraplégicos	ART1	8x10--4
26	Abbas et alia (1995)	controle dos sistemas de estimulação neuromuscular	1.perceptron; 2.regra de "Hebbian" modificada	—
27	Su (1994)	diagnóstico de diabete	1.composição hiperretangular; 2.aprendizado de decisão direcionado e supervisionado	—
28	Gray et alia (1992)	supervisão do uso de medicamentos (apresentação de um método)	1.perceptron multicamadas; 2.algoritmo de treinamento baseado na álgebra de Boole	—
29	Gray et alii (1991)	supervisão do uso de medicamentos	1.perceptron multicamadas; 2.algoritmo de treinamento baseado na álgebra de Boole	—
30	Nekovei et alia (1995)	detecção de vasos sanguíneos em angiogramas	1.perceptron multicamadas; 2.propagação reversa	121--17--2
31	Adler et alia (1994)	reconstrução de imagens através de tomografia de impedância elétrica	1.Adaline;2.regra de aprendizado de "Widrow-Hoff"	—
32	Costa et alia (1995)	diagnóstico de câncer de mama	1.filtro de novidades	—

Tabela 2.3: Resumo dos métodos de detecção do contorno do ventrículo esquerdo

ordem	autores	classe de imagem	segmentação	pré-processamento	extração de características	classificação e pós-processamento
1	Goshtasby et alia (1995)	MR	automática	filtro gaussiano	cruzamentos por zero, limiar.	modelagem dos pontos do contorno através de curvas elásticas
2	Lilly et alii (1989)	ang. digital	semi-automático	—	gradiente,limiar.	funções de custo, correção do contorno através de modelos dos ventrículos
3	Grattoni et alia (1985)	ang. digital	automático	filtro passa baixa: média 9x9	gradiente:sobel	seguidor de borda heurístico, interpolação linear
4	Feng et alii (1991)	ultra-som	automático	—	cruzamentos por zero, estimativa do centro do VE	procura radial, técnica "fuzzy"
5	Taratorin et alia (1993)	CT	automático	filtro passa baixa: média 3x3	gradiente	procura radial, intersecção entre conjuntos de restrições
6	Thedens et alii (1995)	MR	semi-automático	"window and level"	gradiente	procura em grafos, uso de uma seqüência espacial de imagens
7	Friedland et alia (1989)	ultra-som	semi-automático	filtro passa baixa: média 7x7	transformada de Hough, estimativa do centro, detector de borda usando um filtro ótimo	minimização de uma função de energia utilizando "simulated annealing"
8	Yang et alia (1994)	SPET	semi-automático	—	limiar	—
9	Zhang et alia (1982)	ultra-som	automático	—	limiar	detecção de pontos móveis em seqüência temporal de imagens, matriz de co-ocorrência temporal
10	Clayton et alii (1974)	ang. digital	semi-automático	filtro gaussiano casado	—	método probabilístico, utiliza informação temporal
11	Barret et alii (1980)	ang. digital	automático	—	filtro-casado de gradiente	método probabilístico, utiliza informação temporal
12	Pope et alii (1985)	ang. digital	automático	—	filtro-casado de gradiente	método de procura em grafos
13	Fleagle et alii (1991)	MR	semi-automático	"window and level"	gradiente	método de procura em grafos
14	Alderman et alii (1973)	ang. digital	manual	—	—	caneta ótica
15	Jouan et alii (1990)	ang. digital	automático	—	expansão de Karhunen-Loeve, média temporal, cruzamento por zero	processo de ortogonalização de Gram-Schmidt, morfologia matemática
16	Chu et alii (1988)	ultra-som	semi-automático	filtro passa baixa gaussiano 41x41	cruzamento por zero	procura radial, interpolação, filtro da mediana
17	Duncan (1987)	ang. digital	automático	filtro passa baixa espacial e temporal, interpolação linear	gradiente, processo de afinamento	método utilizando funções de custo
18	Adam et alii (1987)	ultra-som	semi-automático	filtro da mediana 9x9, amplificação de área contendo o VE	valores médios de regiões pré-definidas	procura radial, algoritmo heurístico de rastreamento de borda
19	Boudraa et alii (1993)	cintilografia	semi-automático	filtro passa baixa, normalização pela mediana	extração da fase através de uma análise de Fourier	método que utiliza uma técnica "fuzzy":algoritmo "isodata"
20	Lambert et alia (1990)	ultra-som	semi-automático	filtro passa baixa:mediana e filtro de médias móveis	laplaciano do gaussiano, técnica de difusão anisotrópica	o objetivo do trabalho é apenas evidenciar as linhas do contorno, sem efetuar uma extração

ordem	autores	classe de imagem	segmentação	pré-processamento	extração de características	classificação e pós-processamento
21	Dias et alia (1996)	ultra-som	automático	—	—	método probabilístico, procura em grafos
22	Coppini et alii (1995)	ultra-som	automático	—	detecção de cruzamentos por zero, gradiente	rede neural
23	Bister et alii (1990)	MR	automático	filtro passa baixo, limiar	obtenção do centro de gravidade de micro-regiões através da transformada em distância	segmentação piramidal, técnica heurística denominada detector de cavidade
24	Ezekiel et alii (1985)	ecocardiografia	automático	suavização através de médias temporais e espaciais	centro do endocárdio, laplaciano	funções de custo, interpolação através de polinômios cúbicos
25	Skorton et alii (1981)	ecocardiografia	automático	suavização através de médias de múltiplos quadros num mesmo instante do ciclo cardíaco	limiar, gradiente:sobel	comparação da segmentação obtida através do limiar como o mapa derivado do operador de sobel
26	Skorton el alii (1985)	tutorial:abordagem técnicas de detecção de borda em ultra-som	—	—	—	—

CAPÍTULO III

MATERIAIS E MÉTODOS

III.1 INTRODUÇÃO

O método para detecção do contorno do VE que propomos apresenta duas características fundamentais que são a utilização de uma técnica de inteligência artificial na determinação do contorno, redes neurais, e a utilização de informação geométrica, centro do órgão e distâncias relativas a este centro.

Embora o método proposto neste trabalho tenha sido validado na detecção do contorno do VE em imagens de medicina nuclear, o mesmo pode ser generalizado para outras classes de imagens e outros tipos de órgãos e estruturas. Os pressupostos desse método são:

- a existência de uma região da borda onde a mesma seja bem definida;
- uma regularidade na localização do órgão na imagem a ser segmentada.
- o contorno a ser segmentado tem que ser um contorno arredondado ou contorno-θ (Moura et alia, 1992).

Um contorno-θ é representado pela coordenada de seu centro e por uma série de raios R(θ), onde θ é o ângulo entre um dado raio e o eixo x. A única restrição a esse tipo de contorno é que, para um dado θ, exista apenas um raio R(θ). Não necessariamente o contorno tem que ser convexo, porém não pode ter dobras que interceptem um raio. Na figura 3.1(a) mostramos um exemplo de contorno-θ, enquanto que na figura 3.1(b) mostramos um exemplo de contorno não θ.

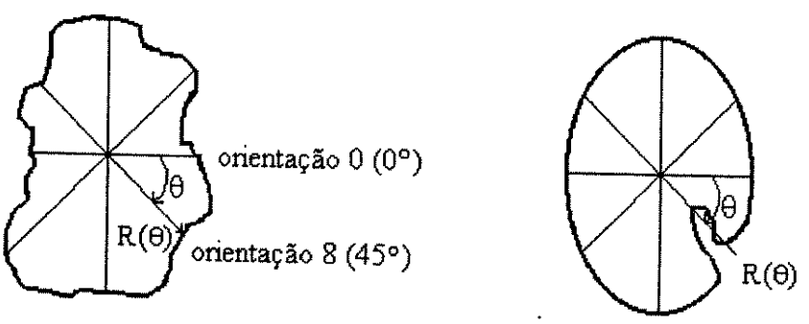


Figura 3.1:(a)representação esquemática de um contorno-θ com 8 raios traçados e (b)representação esquemática de um contorno não θ com 8 raios traçados.

A representação unidimensional de um contorno através de seu raio $R(\theta)$ é denominado por Gonzalez de uma "assinatura do contorno". Na figura 3.2 mostramos a assinatura de um contorno-θ genérico.

Em termos genéricos, sem nos referirmos a uma classe particular de imagens, as etapas desse método podem ser citadas seqüencialmente, como:

- determinação de um grupo de variáveis que diferenciem a região do contorno do órgão de sua vizinhança. Essas variáveis constituirão das entradas de uma rede neural;
- determinação da posição geométrica do centro do órgão.
- procura radial do contorno em 64 orientações (tendo como origem o centro do órgão), utilizando, em cada uma das orientações, uma rede neural tipo *perceptron* multicamadas para determinação de *pontos candidatos* ao contorno;
- utilização de critérios heurísticos para a seleção dos pontos do contorno a partir dos pontos candidatos selecionados pela rede neural.

Ao implementar-se esse método para a segmentação do VE, as etapas do mesmo podem ser classificadas de acordo com o esquema proposto no diagrama em blocos da figura 2.1, conforme detalhamos a seguir:

- pré-processamento: ajuste do histograma da imagem e ampliação da imagem original de 64x64 para 256x256.
- extração de parâmetros: intensidade do pixel; gradiente linear; distância do pixel de contorno em relação ao centro do VE e determinação do centro do VE.
- classificação e pós-processamento: procura radial; uso de redes neurais para determinação de pontos candidatos do contorno; uso de critérios heurísticos para selecionar pontos do contorno entre os pontos candidatos; filtragem e suavização.

Na etapa de pré-processamento, realizamos inicialmente uma redistribuição dos níveis de cinza da imagem através de um ajuste do histograma da mesma. Após esse ajuste, efetuamos uma ampliação da imagem de 64x64 para 256x256, através de interpolação bilinear.

Em seguida, utilizando um algoritmo original, efetuamos a determinação do centro do VE. Como poderemos perceber ao longo deste capítulo, a determinação precisa do centro é um requisito fundamental para o bom desempenho do método de detecção do contorno. Esse centro é considerado como origem de uma seqüência de 64 raios igualmente espaçados, denominados de orientações. Cada orientação é definida por seu ângulo θ. Na figura 3.1(a) mostramos um contorno genérico com 8 dessas orientações mostradas.

O processamento para a determinação do pontos do contorno é realizado através de uma procura radial em cada uma das 64 orientações anteriormente definidas. Ao conjunto de pixels

situados sobre o raio de uma orientação, atribuímos o nome de $\mathbf{Orient}_i$, onde $0 \leq i \leq 63$.

Os pixels candidatos ao contorno em cada orientação são determinados através de uma rede neural tipo *perceptron* multicamadas. Uma característica do método proposto é que a cada orientação está associada uma rede neural. Assim, para a determinação dos pixels candidatos nas 64

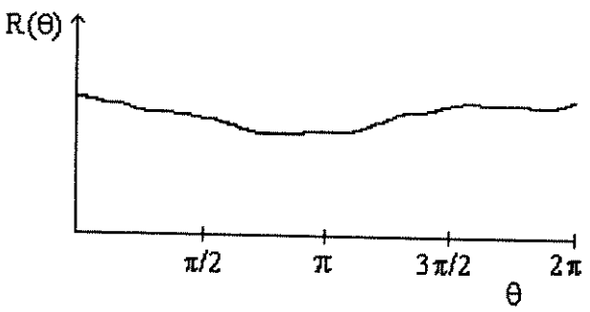


Figura 3.2: assinatura de um contorno genérico.

orientações, teremos um total de 64 redes neurais. As razões que nos levaram ao utilizar 64 redes neurais ao invés de apenas uma rede neural foram as seguintes:

- é mais fácil monitorar o treinamento de uma rede dedicada a apenas uma orientação do que a 64 orientações. No caso da rede não convergir, as alterações necessárias podem ser feitas rapidamente, e o treinamento reiniciado;

- é possível otimizar o tempo de treinamento, utilizando taxas de aprendizado maiores em regiões do ventrículo onde a borda é mais facilmente identificada.

As entradas da rede neural foram determinadas através de um procedimento que será descrito adiante. Foram testadas duas redes: uma rede **A** com 4 entradas e uma rede **B** com 3 entradas. As variáveis de entrada da rede **B** foram: distância do pixel em relação ao centro do VE, intensidade do pixel e gradiente linear. A rede **A**, além destas três variáveis, incluiu também a distância de um pixel em uma orientação em relação ao pixel de contorno da orientação anterior.

Ao calcularmos o valor de saída **s** de uma das redes neurais para todos os pixels $\mathbf{Orient}_{i'}$ de uma orientação **i'**, associamos a cada pixel $\mathbf{P}_j$ um valor $\mathbf{s}_j$, onde $0 \leq j \leq n-1$ e **n** é o número de pixels na orientação **i'**. O índice 0 corresponde ao pixel do centro do VE. A escolha dos pixels candidatos é realizada estabelecendo-se um limiar $\mathbf{s}_L$ para a saída da rede neural. Todos os pixels cujos valores $\mathbf{s}_j$ forem maiores do que $\mathbf{s}_L$ serão considerados pixels candidatos. O conjunto de pixels candidatos de uma orientação **i'** é definido como $\mathbf{Cand}_{i'}$.

A escolha dos pixels de contorno em uma orientação a partir do conjunto de pixels candidatos dessa orientação é realizada utilizando 3 critérios definidos heuristicamente. Esse critérios serão definidos posteriormente neste capítulo. Após a determinação dos pixels de contorno nas 64 orientações, teremos um conjunto de pixels de contorno **Cont**, definido como:

$$\text{Cont} = \{P_0, P_1, P_2, \ldots\ldots, P_{63}\}$$

onde:

$$P_i \in \text{Cand}_i \qquad (3.1)$$

Após a determinação dos pixels de contorno, o mesmo é filtrado e suavizado, adquirindo sua forma final. Essas operações serão descritas adiante.

A seguir, mencionamos a seqüência de tópicos que será abordada nesse capítulo: aquisição e pré-processamento das imagens; determinação automática do centro; seleção das imagens para o conjunto de treinamento da rede neural; arquitetura da rede neural tipo *perceptron* utilizada; seleção das variáveis de entrada da rede neural tipo *perceptron* e treinamento da mesma; utilização da rede neural; restrições geométricas, filtragem e suavização do contorno obtido; avaliação do desempenho do método e *software* e *hardware* utilizados.

III.2 AQUISIÇÃO E PRÉ-PROCESSAMENTO DAS IMAGENS

As imagens de *blood-pool* foram obtidas de 60 pacientes homens utilizando uma câmara de cintilação convencional (Siemens-Basicam-Maxidelta) equipada com colimador de baixa energia e média resolução (*low energy all purpose* - LEAP). Foi utilizada uma janela de 20% centrada em 140KeV. As imagens foram obtidas por especialistas do Serviço de Radioisótopos do Instituto do Coração do Hospital das Clínicas de São Paulo -InCor. Cada paciente recebeu uma dose de 1mg de pirofosfato de estanho seguida, 30 minutos depois, por uma aplicação no braço de 740Mbq de ^{99m}Tc – MIBI (Dupont Pharma - Cardiolite). As imagens foram adquiridas um minuto após a injeção, com uma resolução espacial de 64x64 pixels e com uma profundidade de 16 *bits*.

A separação espacial entre os ventrículos esquerdo e direito em imagens de *blood-pool* é melhor quando as mesmas são tomadas numa vista perpendicular ao septo interventricular. Na prática, essa condição é obtida com uma inclinação da câmara entre 30° e 50°, na posição oblíqua, anterior e esquerda (*left anterior oblique*-LAO). A separação entre o átrio direito e o ventrículo esquerdo é melhorada com uma inclinação da cauda da câmara entre 5° e 15° (*tilt*), sem um aumento do distanciamento entre o paciente e a câmara (Gelfand et alia, 1988). Os valores exatos desses ângulos dependem do paciente. Para cada paciente, em função de suas características anatômicas, determinam-se, na hora do exame, valores ótimos para esses ângulos. O objetivo, sempre, é melhorar a visualização do VE. Na figura 3.3 mostramos uma ilustração da câmara nas posições oblíqua anterior esquerda e oblíqua anterior direita. Na figura 3.4 mostramos um diagrama dos ventrículos esquerdo e direito obtidos na posição oblíqua anterior esquerda.

A imagens foram adquiridas no modo dinâmico cíclico, com 32 imagens por ciclo cardíaco. A aquisição das imagens foi efetuada utilizando-se intervalos de duração fixa. Para este tipo de aquisição, um valor médio para a duração do ciclo cardíaco é previamente estimado. Para ciclos maiores do que este valor médio, os dados dos quadros posteriores ao quadro 32 são descartados. Por outro lado, ciclos com duração menor do que este valor médio não contribuem com dados para

os últimos quadros. O resultado disso, é que os últimos quadros do ciclo cardíaco ficam com uma contagem reduzida em relação aos quadros iniciais. Para minimizar o efeito da contagem decrescente nos últimos quadros, o equipamento utilizado faz uso de uma técnica denominada de "sincronismo reverso" (*backward gating).* A contagem dos pulsos nos quadros é realizada a partir de um marcador R_i no início de um ciclo cardíaco, em direção ao marcador R_{i+1} no início do ciclo cardíaco

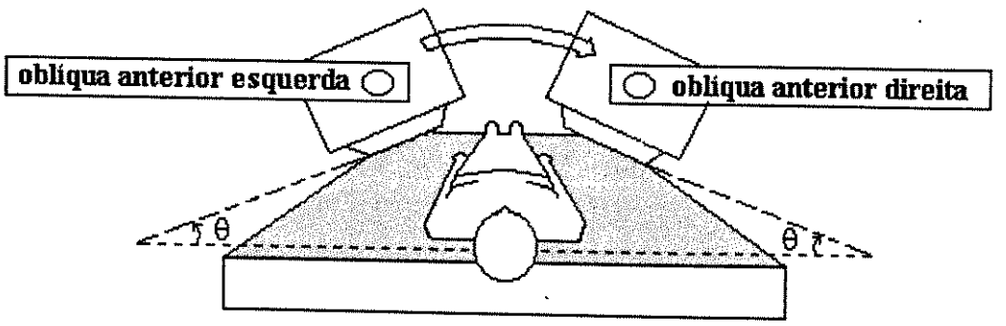


Figura 3.3: posições da gama câmara: oblíqua anterior esquerda e oblíqua anterior direita.

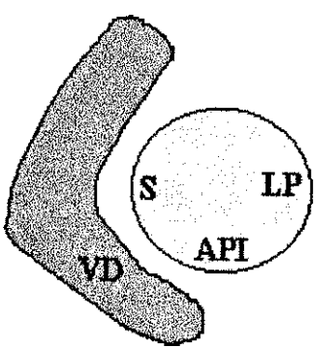


Figura 3.4: Vista do miocárdio na posição oblíqua anterior esquerda: VD-ventrículo direito, S-septo, LP - posição lateral posterior do VE, API - posição apical inferior do VE..

seguinte (passagem direta) e reversamente, a partir do marcador R_{i+1} em direção ao marcador R_i (passagem reversa). Isso é mostrado na figura 3.5. O ciclo médio é então reconstruído combinando-se a contagem da parte inicial dos quadros da passagem direta com a contagem da parte inicial dos quadros da passagem reversa. Normalmente utilizam-se 2/3 dos quadros da passagem direta e 1/3 dos quadros da passagem reversa.

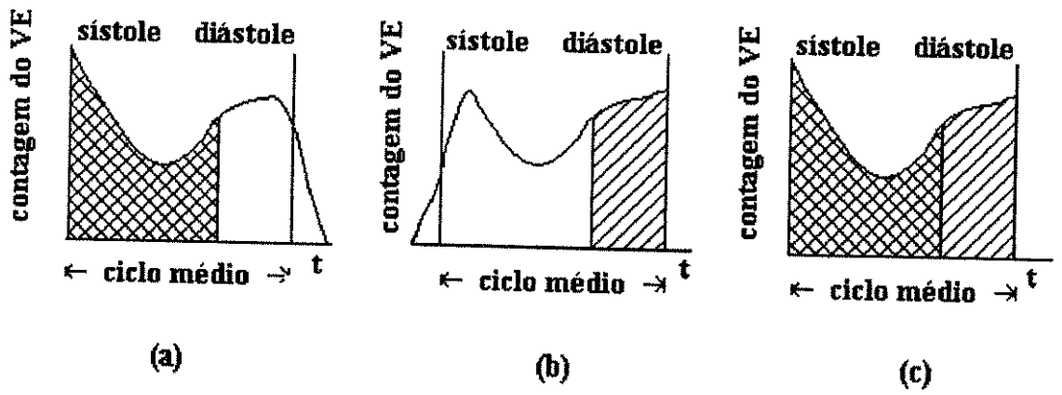


Figura 3.5: reconstrução do ciclo cardíaco por sincronismo reverso: (a) passagem direta; (b) passagem reversa e (c) união de 2/3 de (a) com 1/3 de (b).

Com o objetivo de melhorar o contraste da imagem e padronizar, foi realizado um ajuste no histograma da mesma. Com esse ajuste, objetiva-se melhorar a separação entre os níveis de cinza. A fim de realizar esse ajuste, desprezamos, no histograma da imagem original de 64x64, o nível de cinza de valor zero que, na sua grande maioria, corresponde a pixels do fundo da imagem, que não carregam nenhuma informação útil para a segmentação do ventrículo esquerdo. Na figura 3.6(a) mostramos um histograma de uma imagem original 64x64, enquanto que na figura 3.6(b) mostramos o histograma da mesma imagem, desprezando o nível de cinza zero. O ajuste do histograma consistiu de um reescalonamento linear da distribuição dos níveis de cinza da imagem. Supondo que os níveis de cinza antes do ajuste sejam representados pela variável **x** e, depois do ajuste, sejam representados pela variável **y**, a relação entre **y** e **x** é dada por:

$$y = K.x + m \tag{3.2}$$

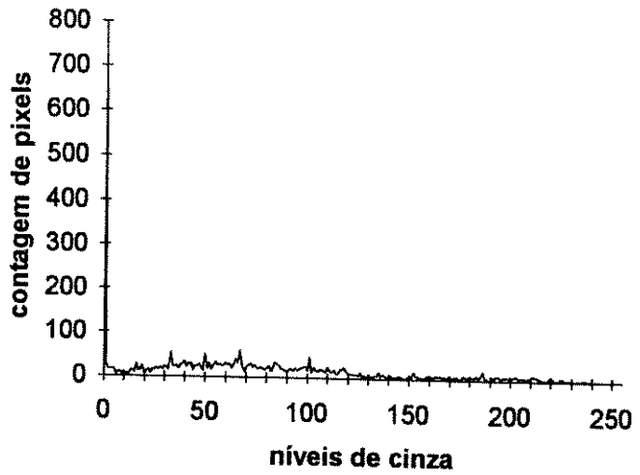
Os valores de **k** e **m** foram determinados utilizando-se 99% da área do histograma da figura 3.6 (b), conforme proposto por Moura et alia (1992). A área correspondente a 1% do histograma é removida da seguinte forma: 0,5% na extremidade inferior e 0,5% na extremidade superior. Para determinarmos **k** e **m** utilizamos as seguintes equações:

$$k = \frac{255}{C_2 - C_1} \quad \text{e} \tag{3.3}$$

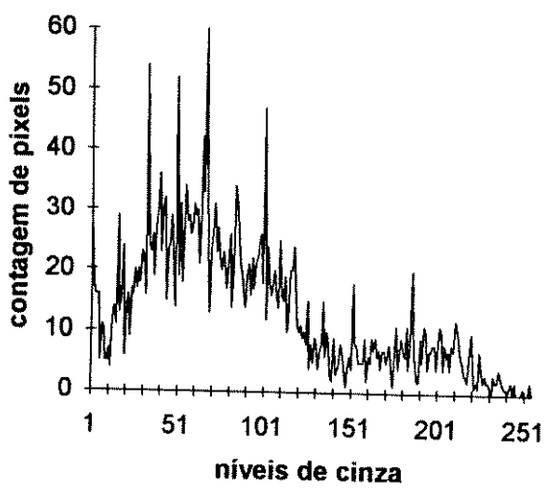
$$m = \frac{-255.C_1}{C_2 - C_1} \tag{3.4}$$

onde: C_1= valor de intensidade do pixel correspondente a 0,5% da área do histograma na extremidade inferior.

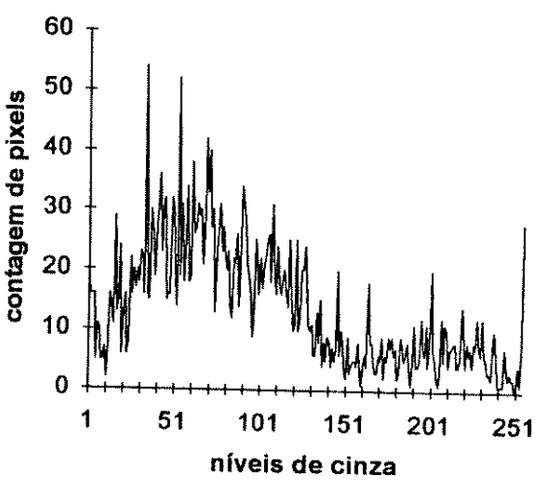
C_2= valor de intensidade do pixel correspondente a 0,5% da área do histograma na extremidade superior.



(a)



(b)



(c)

Figura 3.6: histogramas referentes a uma imagem 64x64. (a) histograma da imagem original; (b)histograma suprimindo os pixels com intensidade igual a zero e (c)histograma ajustado.

Os valores de C_1 e C_2 são ilustrados através do diagrama genérico da figura 3.7. Nesta figura, as áreas hachuradas em ambas as extremidades correspondem a 0,5% da área total do histograma.

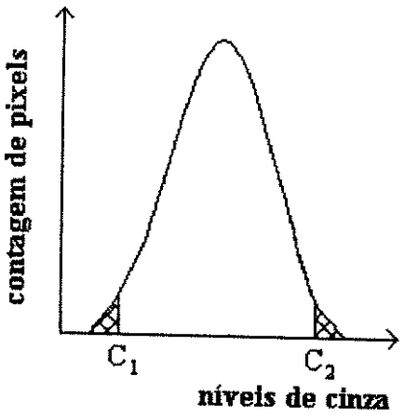


Figura 3.7: histograma genérico utilizado no processo de ajuste do histograma da imagem original de 64x64.

O reescalonamento linear aplica-se apenas aos pixels localizados na faixa entre C_1 e C_2. Os pixels situados na área hachurada inferior são levados para o nível de cinza 0, enquanto que os pixels situados na área hachurada superior são levados para o nível de cinza 255. Na figura 3.8 (a) mostramos a imagem original de 64x64, enquanto que na figura 3.8 (b) mostramos a imagem resultante do processo de ajuste do histograma anteriormente descrito. O resultado desse ajuste do histograma pode ser observado de duas formas diferentes. A primeira delas é através de uma comparação dos histogramas das figuras 3.6 (b) e 3.6(c), onde observamos que no histograma resultante do ajuste, as contagens dos pixels situados na extremidade direita são maiores que no original. A segunda forma é através da comparação das imagens 3.8 (a) e 3.8 (b), onde observamos que a imagem após a ajuste do histograma parece melhor definida do que a imagem original.

Após o ajuste do histograma, as imagens de 64x64 pixels são ampliadas para 256x256, através de interpolação bilinear (Wolberg, 1994). A interpolação bilinear utiliza uma combinação linear dos quatro valores dos pixels mais próximos, para produzir um novo valor interpolado. Dados quatro pixels na imagem original $P_0(u_0, v_0)$, $P_1(u_1, v_1)$, $P_2(u_2, v_2)$ e $P_3(u_3, v_3)$ com suas respectivas intensidades I_0, I_1, I_2 e I_3, qualquer intensidade de um pixel intermediário $P(u,v)$, pode ser calculada através da seguinte expressão:

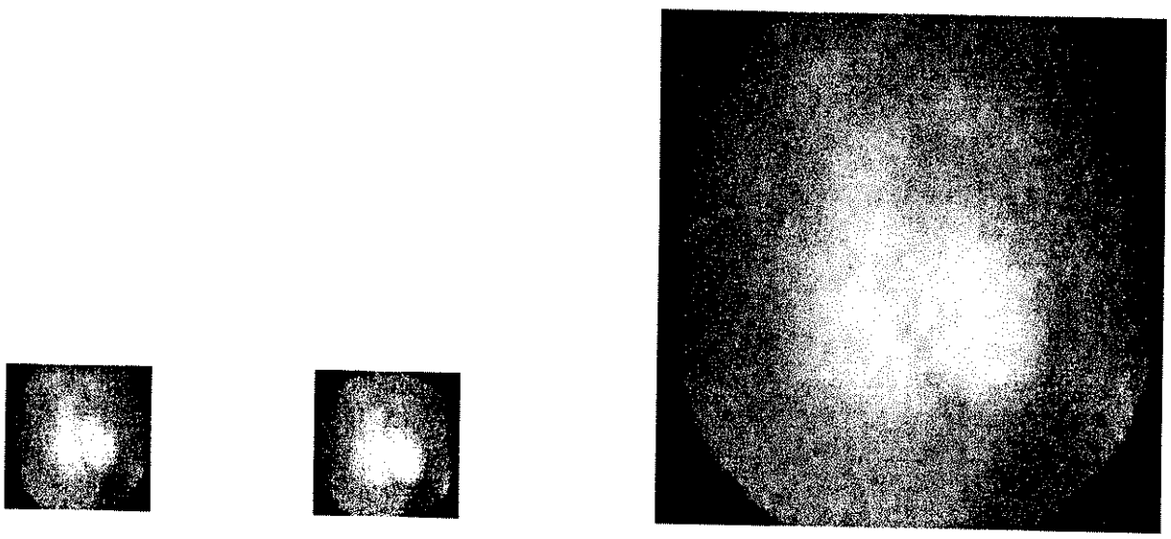
$$I(u,v) = a_0 + a_1.u + a_2.v + a_3.u.v \tag{3.5}$$

Os coeficientes $\mathbf{a_i}$ da expressão anterior são obtidos resolvendo-se o seguinte sistema de equações:

$$\begin{bmatrix} I_0 \\ I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} 1 & u_0 & v_0 & u_0v_0 \\ 1 & u_1 & v_1 & u_1v_1 \\ 1 & u_2 & v_2 & u_2v_2 \\ 1 & u_3 & v_3 & u_3v_3 \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} \tag{3.6}$$

Para resolução deste sistema de equações utilizamos o método de Gauss-Jordan (Gonçalves et alia, 1977).

Embora a interpolação bilinear não aumente a quantidade de informação, ela reduz o tamanho do pixel, aumentando a resolução quando se determina o contorno do VE. Uma conseqüência do processo de interpolação é também a minimização de ruídos de alta freqüência na imagem. Na figura 3.8 (c) mostramos o resultado da interpolação da imagem da figura 3.8 (b).



(a) (b) (c)

Figura 3.8: (a) imagem original com tamanho de 64x64; (b) imagem resultante do ajuste do histograma com tamanho 64x64 e (c) imagem resultante da interpolação bilinear com tamanho 256x256.

III.3 DETERMINAÇÃO AUTOMÁTICA DO CENTRO DO VE

O método mais simples para determinação do centro do VE consiste em procurar o pixel de maior contagem em uma área onde o VE é esperado (Gelfand et alia, 1988). A delimitação desta área é normalmente efetuada manualmente. Nesse trabalho propomos um método automático para determinação do centro do VE.

Para todas as 60 imagens de *blood-pool* utilizadas, foram extraídos manualmente os centros dos ventrículos esquerdo e direito e o contorno do VE. Os centros extraídos manualmente foram utilizados como padrão ouro de comparação para o centro obtido automaticamente, enquanto que o contorno manual do VE foi utilizado como padrão ouro de comparação para o contorno obtido automaticamente.

A partir das coordenadas dos centros dos ventrículos esquerdo e direito extraídos manualmente, calculamos valores médios para as coordenadas dos centros de ambos os ventrículos. A mediatriz do segmento de reta que une os pixels correspondentes a esses dois centros médios foi utilizada como um divisor inicial entre os pixels pertencentes aos ventrículos esquerdo e direito. O centro de massa dos pixels situados à direita dessa mediatriz e com intensidades superiores a 200, fornece-nos uma primeira estimativa para o centro do VE. Na figura 3.9 (a) mostramos uma imagem ampliada de um VE, enquanto que na figura 3.9 (b) mostramos a mediatriz anteriormente referida superposta à imagem resultante após a aplicação do limiar de 200 à imagem da figura 3.9 (a). É mostrado também o centro do VE resultante desse processo. Essa primeira estimativa do centro do VE sofre, então, um processo de refinamento.

Utilizando os contornos traçados manualmente obtivemos, para as 64 imagens de *blood-pool* do VE, as intensidades médias dos pixels situados nas extremidades lateral posterior, apical inferior e superior do VE, na posição correspondente às coordenadas $\mathbf{x}_c$ e $\mathbf{y}_c$ do centro do mesmo. Pixels com essas intensidades são marcados, respectivamente pelos números 1,2 e 4 na figura 3.9 (d). Um quarto pixel auxiliar para refinamento, situado no lado do septo, é obtido a seguir. Determinamos um segmento de reta que passa pelo centro estimado do VE e que termina em pixels cujas intensidades são iguais à intensidade do pixel 1. Esse segmento de reta é mostrado na figura 3.9 (c). Num segundo instante, obtemos o histograma médio dos pixels situados sobre esse segmento e sobre outros 10 segmentos adjacentes e paralelos, 5 deles situados acima do segmento mostrado e 5 situados abaixo do segmento mostrado. Na figura 3.10 mostramos esse histograma. Para as 60 imagens estudadas, verificamos que o ponto crítico mínimo do histograma situa-se na região do septo. Esse ponto é designado como ponto 3 na figura 3.9 (d). O valor refinado da coordenada $\mathbf{x}_c$ do centro do VE é obtido através de uma média das coordenadas **x** dos pontos 1 e 3, enquanto que o valor refinado da coordenada $\mathbf{y}_c$ do centro é obtido através de uma média das coordenadas **y** dos pontos 2 e 4. Na figura 3.10 (d) mostramos a localização do centro refinado e os pontos utilizados na sua determinação.

Na tabela 3.1 mostramos valores dos erros percentuais médios obtidos para as 60 imagens de *blood-pool* do VE. O erro $\mathbf{E}_x$ ($\mathbf{E}_y$) corresponde ao erro percentual médio obtido pela razão entre a diferença das coordenadas $\mathbf{x}_c$($\mathbf{y}_c$) do centro obtidas automaticamente e do centro obtidas manualmente, pelo raio **r** do VE, definido como a distância entre o centro do mesmo e sua extremidade lateral posterior. Esse erros são expressos através das equações:

$$E_x = \frac{1}{60} \cdot \sum_{i=0}^{59} \frac{|x_{cmi} - x_{cai}|}{r_i} \quad \text{e} \qquad (3.7)$$

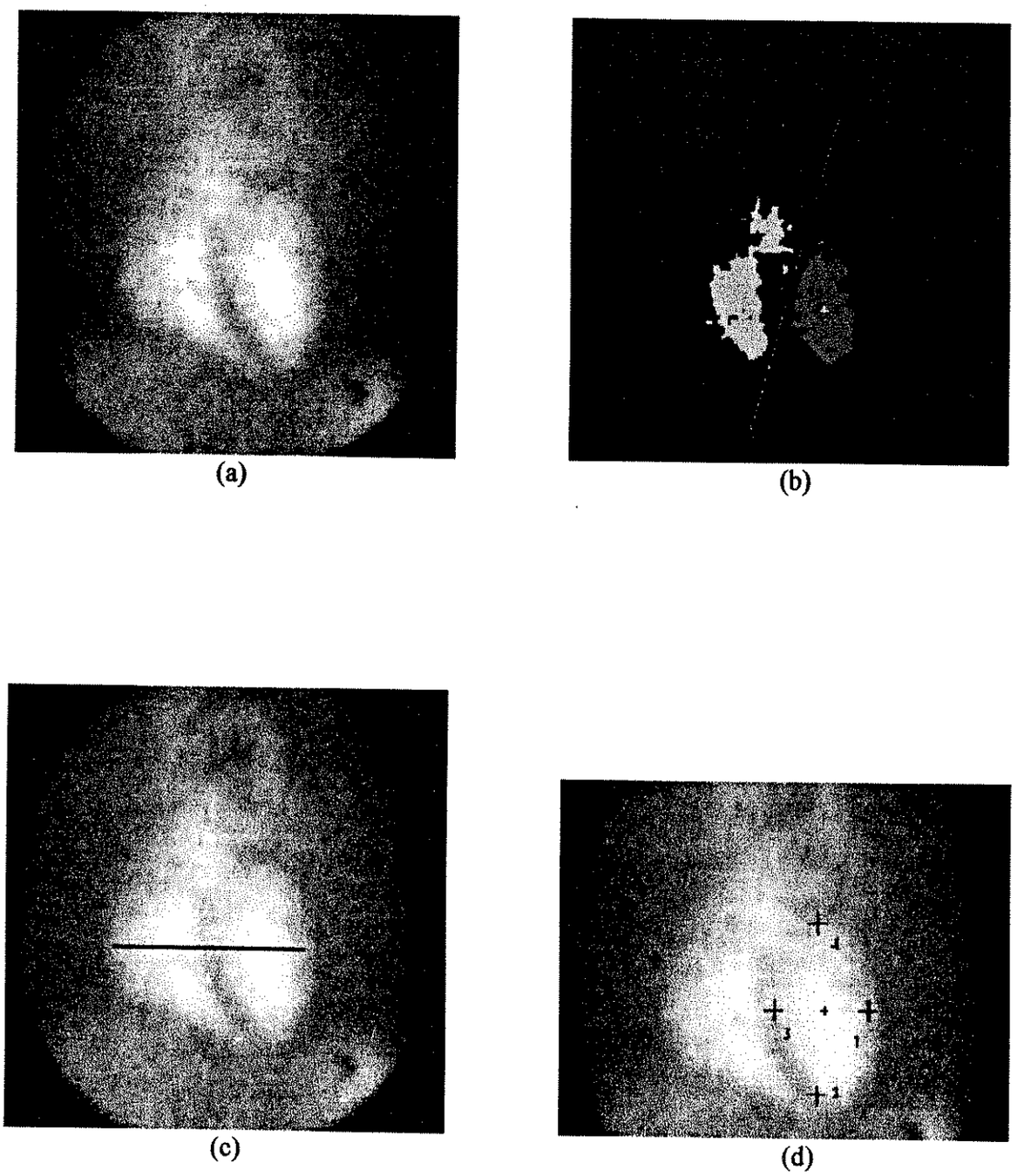


Figura 3.9: (a) imagem de blood-pool ampliada de um VE; (b) mediatriz, pixels com intensidade superior ao limiar de 200 e estimativa inicial do centro; (c) segmento de reta passando pelo centro estimado do VE e (d) centro refinado do VE e pontos 1,2,3 e 4 utilizados no seu cálculo.

$$E_y = \frac{1}{60} \cdot \sum_{i=0}^{59} \frac{|y_{cm\,i} - y_{ca\,i}|}{r} \qquad (3.8)$$

onde:

$(x_{cm\,i}, y_{cm\,i})$= coordenadas do centro manual do ventrículo **i**;

$(x_{ca\,i}, y_{ca\,i})$ = coordenadas do centro automático do ventrículo **i**;

r = raio do ventrículo.

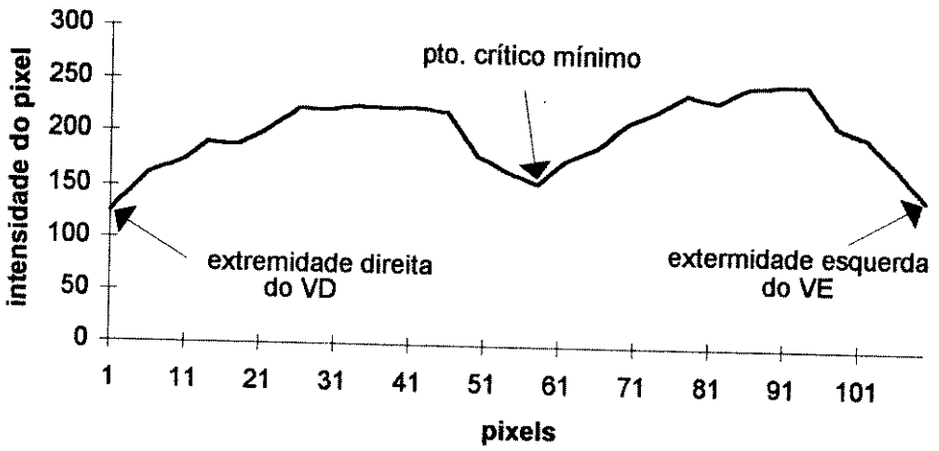


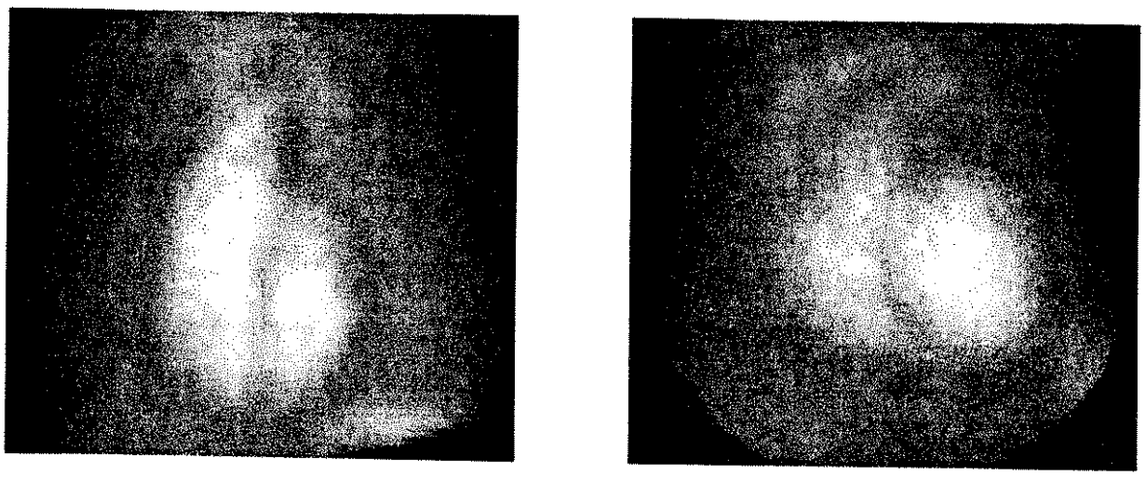
Figura 3.10: perfil médio dos pixels localizados sobre o segmento de reta mostrado na figura 3.10 (c) e sobre 5 segmentos adjacentes situados acima do mesmo e 5 segmentos adjacentes situados abaixo do mesmo.

Tabela 3.1: Erros percentuais no processo de determinação do centro.

	Ex (%)		Ey (%)	
	média	desvio padrão	média	desvio padrão
antes do refinamento	9.14	7.92	11.67	10.14
depois do refinamento	5.70	4.85	6.69	5.90

III.4 SELEÇÃO DAS IMAGENS PARA O TREINAMENTO DA REDE NEURAL

A forma do VE pode variar significativamente de uma pessoa para a outra. Para ilustrarmos esse fato, mostramos na figura 3.11 dois VE's com formas bem diferentes. O ventrículo da figura 3.11 (a) tem uma forma alongada, enquanto que o ventrículo da figura 3.11 (b) tem uma forma arredondada. Com o objetivo de garantir que no treinamento tenhamos um conjunto representativo das diversas formas do VE, adotamos o procedimento de classificação de formas descrito a seguir.



(a) (b)

Figura 3.11: Exemplos de ventrículos com formas diferentes: (a)forma alongada e (b)forma arredondada.

Para classificarmos as formas dos VE das 60 imagens de *blood-pool,* utilizamos uma rede auto-organizável de Kohonen, cuja arquitetura é mostrada na figura 3.12. O número de neurônios de entrada dessa rede é 64. Essas 64 entradas da rede correspondem as 64 orientações descritas no

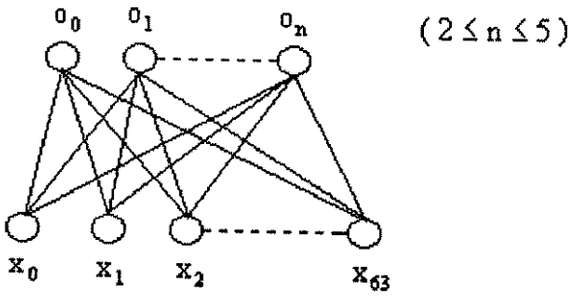


Figura 3.12: arquitetura da rede auto-organizável de Kohonen utilizada para a classificação das formas dos VE.

item III.1. Para um dado ventrículo, uma entrada x_i corresponde ao raio do contorno na orientação **i** normalizado através da seguinte expressão:

$$x_i = \frac{\text{raio da orientação } \mathbf{i}}{\text{raio da orientação } 0^\circ} \qquad (3.9)$$

Os raios utilizados na expressão (3.9) são obtidos a partir dos contornos traçados manualmente. A padronização definida na expressão 3.9 visa tornar o processo de classificação das formas dos VE independentes do tamanho do mesmo. O número de saídas da rede corresponde ao número de grupos em que se deseja classificar as formas dos ventrículos. Nesse trabalho, variamos o número de saídas **n** entre 2 e 5. O treinamento da rede foi efetuado utilizando o algoritmo apresentado no capítulo 2, com as seguintes peculiaridades: em cada ciclo de treinamento apresentamos as 60 imagens de VE; utilizaram-se dois valores para o número total de ciclos de treinamento, **nt,** 10.000 e 20.000; a atualização dos pesos da rede obedeceu o algoritmo mostrado a seguir:

A. zerar inicialmente um contador "i". Designar o número do ciclo de treinamento por nc.
B. se nc< (i+1).nt / n, ir para C. Caso contrário ir para F.
C. atualizar os pesos conectados aos nós de saída, exceto aqueles conectados aos "i" nós com os menores valores de saída.
D. incrementar nc e ir para B.
E. incrementar o valor de "i" e comparar com n. Se i<n ir para B. Caso contrário ir para F.
F. fim.

O objetivo deste algoritmo é diminuir gradativamente a vizinhança dos nós em torno do nó vencedor que tem os seus pesos atualizados. Os resultados da classificação não se alteraram quando **nt** foi aumentado de 10.000 para 20.000. Na tabela 3.2 mostramos o resultado da classificação, considerando o percentual de imagens por grupo, para n=2,3,4 e 5.

Tabela 3.2: percentual de imagens nos grupos resultantes da classificação através da rede auto-organizável de Kohonen.

	percentual por grupo				
número de grupos	1	2	3	4	5
2	42.5	57.5	0	0	0
3	20	35	45	0	0
4	12.5	15	30	42.5	0
5	10	12.5	20	27.5	30

Essa mesma rede de Kohonen foi utilizada para classificar os contornos obtidos com o método automático completo para determinação do contorno do VE desenvolvido nesse trabalho de tese. Na figura 3.13 mostramos o percentual de acerto da classificação dos contornos automáticos em termos dos grupos de classificação da tabela 3.2. Esse contornos automáticos foram obtidos

utilizando-se 20 imagens escolhidas aleatoriamente para o conjunto de treinamento da rede tipo *perceptron* multicamadas.

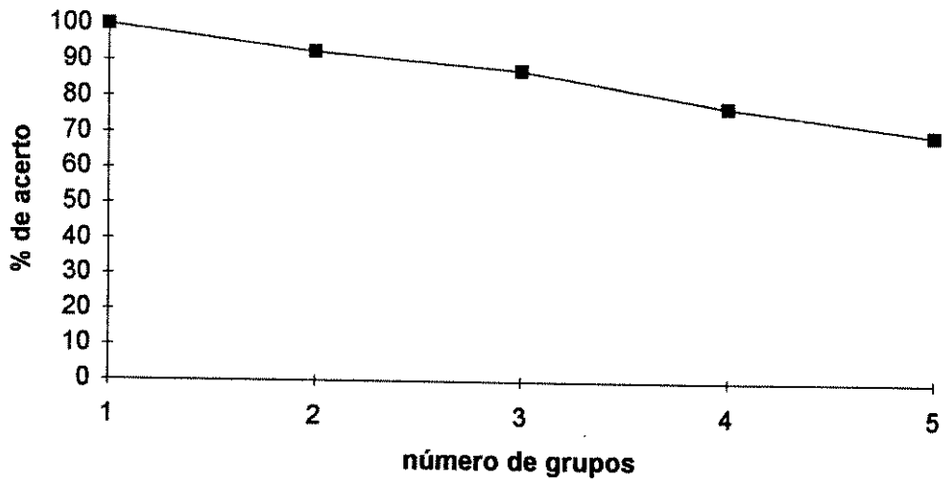


Figura 3.13: percentual de acertos na classificação dos contornos automáticos pela rede auto-organizável de Kohonen, para n=2,3,4 e 5.

Para n=3, o percentual de classificação correto foi em torno de 90%, enquanto que, para n=4 e n=5, este índice caiu abaixo de 80%. Em face disso, concluímos que as diferenças entre as formas dos contornos não justificam a escolha de mais do que 3 grupos de classificação para selecionar as imagens do grupo de treinamento da rede tipo *perceptron* multicamadas. O número de imagens desse grupo de treinamento escolhido por nós foi 20. Não existem regras para se determinar o número de amostras do grupo de treinamento de uma rede neural. O que se pode fazer apenas é variar o número de amostras e verificar se variam os resultados obtidos com a utilização da rede. As 20 imagens foram escolhidas a partir dos 3 grupos de classificação resultantes da aplicação da rede de Kohonen com n=3, sendo o número de imagens utilizadas de cada grupo definido pelos percentuais da tabela 3.2. Assim, foram extraídas 4 imagens do grupo 1, 7 imagens do grupo 2 e 9 imagens do grupo 3, totalizando as 20 imagens do grupo de treinamento. Designamos essas 20 imagens de treinamento como um conjunto A de imagens, enquanto que as 40 imagens restantes do conjunto das 60 imagens utilizadas nesse trabalho, foram designadas de grupo B. As imagens do grupo A foram utilizadas tanto para o treinamento como para o teste da rede, enquanto que as imagens do grupo B foram utilizadas apenas para o teste da rede.

III.5 ARQUITETURA DA REDE NEURAL TIPO PERCEPTRON MULTICAMADAS

A arquitetura da rede neural utilizada, mostrada na figura 3.14, é a de um *perceptron* de 3 camadas de nós. O número de nós da camada intermediária foi escolhido utilizando a estimativa de Baum (1988) apresentada no capítulo 2. De acordo com essa estimativa, o número de neurônios da camada intermediária é igual a N/d, onde N é igual ao número de vetores do conjunto de treinamento e d é a dimensão desses vetores. Para utilizarmos essa expressão, necessitamos saber o número de

vetores utilizados no treinamento de uma orientação e o número de variáveis de cada vetor. Como será visto na próxima seção, cada imagem do conjunto de treinamento contribui com 3 vetores. Multiplicando-se esse valor pelo número de imagens do conjunto de treinamento, 20 obtemos um total de 60 vetores utilizados no treinamento de uma orientação. Como será visto também na próxima seção, cada um desses vetores pode ter 3 ou 4 dimensões. Utilizando d=3, obtemos um número máximo de nós, dado por:

$$\text{número de nós na camada intermediária} = \frac{N}{d} = \frac{60}{3} = 20 \qquad (3.10)$$

Dessa forma, foram utilizados 20 nós na camada intermediária. O número de nós na camada de saída foi de apenas 1, pois a decisão envolvida é da forma "se o nó pertence ao contorno a saída vale **x**, caso contrário, a saída vale **y**".

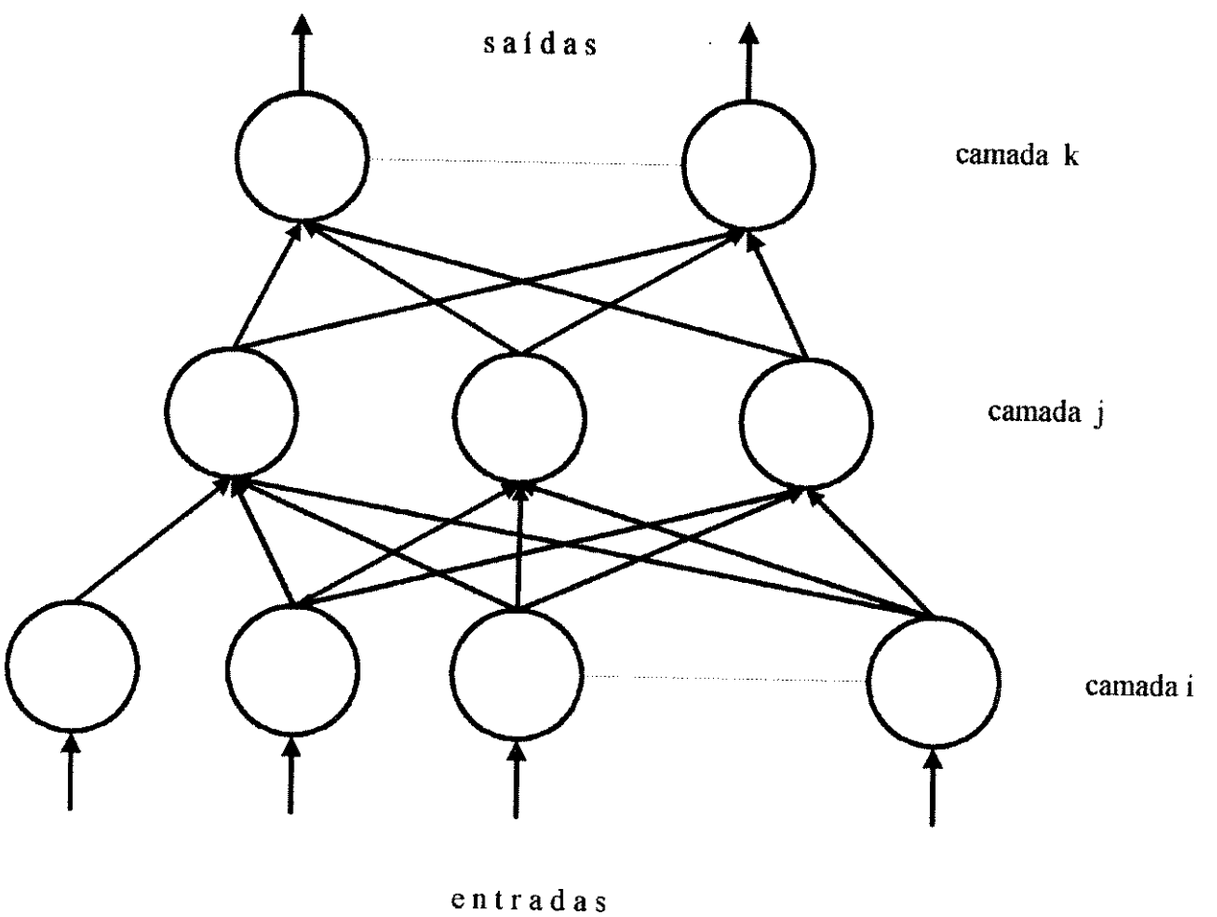


Figura 3.14: arquitetura da rede neural tipo perceptron multicamadas utilizada.

III.6 VARIÁVEIS DE ENTRADA E TREINAMENTO DA REDE NEURAL

Durante o treinamento da rede neural, são apresentados, para cada orientação e para cada imagem do conjunto de treinamento, três pixels. Dois deles não pertencem ao contorno: o primeiro reside no interior do contorno e o segundo no exterior do contorno. Um terceiro é escolhido sobre o contorno. As variáveis de entrada da rede são parâmetros extraídos desses pixels. Inicialmente, foram investigadas as seguintes variáveis: 1. distância do pixel ao centro do VE ,**d**; 2. intensidade do pixel, **I**; 3. derivada ao longo da orientação, **g**; 4. magnitude do gradiente, **m** e 5. distância em relação ao pixel do contorno da orientação anterior, **da**.

A derivada ao longo de uma orientação para um pixel P_j, foi calculada através da seguinte expressão:

$$g_j = \frac{I_{j+1} - I_{j-1}}{2} \qquad (3.11)$$

onde: I_j = intensidade do pixel P_j. O pixel P_0 de uma orientação corresponde ao pixel do centro do VE.

A magnitude do gradiente foi calculada utilizando-se os operadores de Sobel nas direções x e y e extraindo-se o módulo dos valores resultantes.

Dois importantes parâmetros a serem escolhidos na fase de treinamento são as distâncias dos pixels interno e externo em relação ao centro do VE. Nesse trabalho, expressamos essas distâncias como um múltiplo da distância do pixel de contorno em relação ao centro do VE em cada orientação. Testamos dois pares de distância: no primeiro par utilizamos uma distância do pixel interno igual a 0,8 vezes a distância do pixel de contorno e uma distância do pixel externo igual a 1,2 vezes a distância do pixel de contorno (ver figura 3.15). No segundo par utilizamos uma distância do pixel interno igual a 0,7 vezes a distância do pixel de contorno e uma distância do pixel externo igual a 1,3 vezes a distância do pixel de contorno.

Como a rede é forçada a diferenciar pixels que são do contorno de pixels que não são do contorno, é interessante visualizarmos as diferenças entre os valores das 5 variáveis anteriormente citadas para estes dois conjuntos de pixels e para os dois pares de distância escolhidos. Ao obtermos esses gráficos, mostrados nas figuras 3.16, 3.17, 3.18, 3.19 e 3.20, objetivamos dois propósitos: reconhecer quais variáveis são mais importantes na diferenciação entre os grupos de pixels que são do contorno e que não são do contorno e, portanto, quais delas são mais adequadas para o treinamento da rede, e qual par de distâncias possibilita uma melhor separação entre as variáveis correspondentes as estes dois conjuntos de pixels.

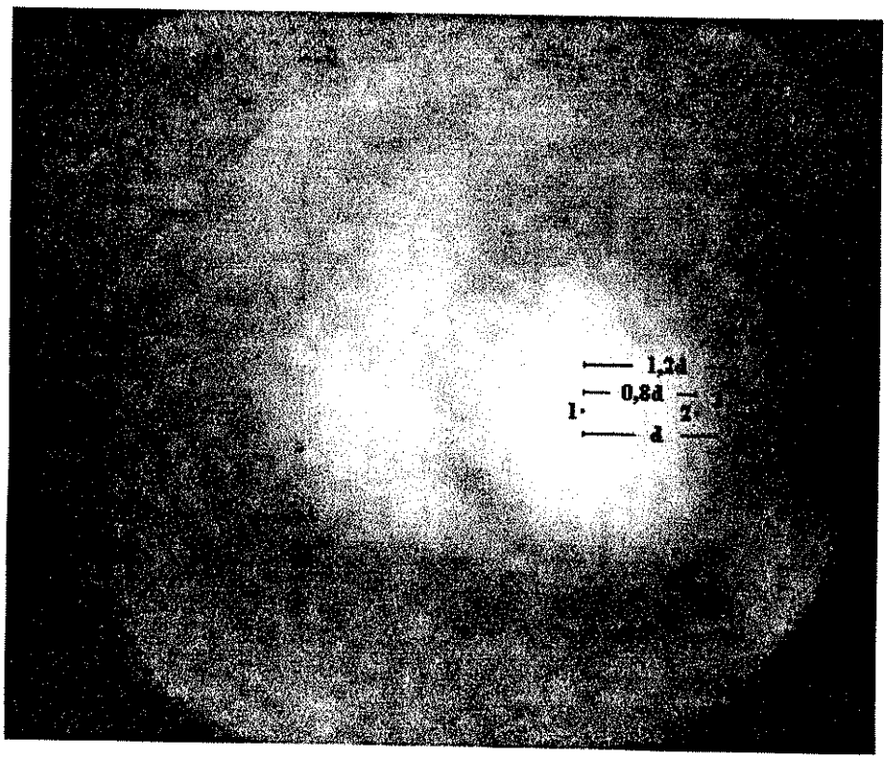


Figura 3.15:Ilustração das distâncias dos pixels utilizadas no treinamento.

Nos gráficos das figuras 3.16-3.20 mostramos valores normalizados das 5 variáveis de entrada da rede em função da orientação. Foram esses valores normalizados que utilizamos para alimentar a rede neural. A normalização teve por objetivo enquadrar todos os valores na faixa entre 0 e 1, adequando-os à entrada da função sigmóide. A normalização da distância do pixel em relação ao centro do VE foi feita obedecendo a seguinte expressão:

$$d_n = \frac{d}{\text{distância do pixel de contorno na orientaçao de } 0^\circ} \qquad (3.12)$$

onde: d = distância do pixel em relação ao centro do VE;
d_n = distância do pixel em relação ao centro do VE normalizada.

A normalização das intensidades dos pixels foi efetuada obedecendo a seguinte expressão:

$$I_n = \frac{I}{255} \qquad (3.13)$$

onde: I = intensidade do pixel;
I_n = intensidade do pixel normalizada.

Na normalização da derivada linear, da magnitude do gradiente e da distância do pixel em relação ao pixel de contorno da orientação anterior, utilizamos fatores numéricos adequados para enquadrar os valores na faixa entre 0 e 1.

Para o cálculo dos valores médios e dos desvios padrões das variáveis foram utilizadas as 60 imagens de VE. Em cada um dos gráficos das figuras 3.16, 3.17 e 3.20, referentes à distância do pixel em relação ao centro do VE, intensidade do pixel e distância de um pixel em relação ao pixel do contorno da orientação anterior, respectivamente, mostramos 9 valores em cada orientação: valor médio da variável no pixel interno, valor médio da variável no pixel interno ± 1 desvio padrão, valor médio da variável no pixel de contorno, valor médio da variável no pixel de contorno ± 1 desvio padrão, valor médio da variável no pixel externo e valor médio da variável no pixel externo ± 1 desvio padrão. Em cada orientação, então, cada pixel (interno, sobre o contorno e externo) é representado por um conjunto de 3 valores: valor médio e valor médio ± 1 desvio padrão. Nos gráficos das figuras 3.18 e 3.19, referentes a magnitude do gradiente linear e magnitude do gradiente, respectivamente, mostramos apenas 3 valores em cada orientação: valor médio da variável no pixel interno, valor médio da variável no pixel de contorno e valor médio da variável no pixel externo. Se mostrássemos nesses dois últimos gráficos os 9 valores que mostramos nos três primeiros, os mesmos tornar-se-iam muito densos, pois os valores mostrados seriam muito próximos uns dos outros. Os gráficos correspondentes a letra (a) correspondem a distâncias dos pixels interno e externo de 0,8 e 1,2 vezes a distância do pixel de contorno, respectivamente, enquanto que os gráficos correspondentes a letra (b) correspondem a distâncias dos pixels interno e externo de 0,7 e 1,3 vezes a distância do pixel de contorno, respectivamente.

A fim de melhor visualizarmos como as 5 variáveis acima citadas comportam-se nas diferentes regiões da vizinhança do ventrículo, dividimos os gráficos das figuras 3.16-3.20 em 4 seções, as quais são mostradas como setores de um círculo representando o VE, na figura 3.21. A região 1 corresponde a um ângulo θ no intervalo [-45°,45°] e engloba a posição lateral posterior do VE. A região 2 corresponde a um ângulo θ no intervalo (45°, 135°] e engloba a posição apical inferior do VE. A região 3 corresponde a um ângulo θ no intervalo (135°, 225°] e engloba a região do septo, enquanto que a região 4 corresponde a um ângulo θ no intervalo (225°, 315°) e engloba a posição superior do VE.

Na tabela 3.3 mostramos os valores dessas variáveis com seus respectivos desvios padrões, para a distância do pixel interno de 0,7 e do pixel externo de 1,3, e para oito orientações. Na tabela 3.4 aplicamos o teste t de Student para avaliarmos se os valores médios e de desvios padrões apresentados para o pixel de contorno na tabela 3.3 diferenciam-se significativamente dos mesmos valores apresentados para os pixels interno e externo. O nível de significância que adotamos foi de 0,001, com gl=118 (graus de liberdade). Para este nível de significância um valor do teste t maior do que 3,37 é significativo.

Analisando os gráficos referidos anteriormente e das tabela 3.3 e 3.4 pode-se concluir que:

- a partir dos gráficos das figuras 3.16 a 3.20 observamos que apenas as variáveis geométricas (distância em relação ao centro do VE e distância em relação ao pixel sobre o contorno da orientação anterior) levam a um conjunto de valores referentes aos pixels interno, sobre o contorno e externo bem separados nas 4 regiões em que dividimos o VE. Essa separação é maior para as distâncias do pixel interno e externo de 0,7 e de 1,3 (gráfico (b)), respectivamente.

- a partir do gráfico da figura 3.17 observamos que a variável intensidade do pixel leva a um conjunto de valores referentes aos pixels interno, sobre o contorno e externo separados nas regiões 1, 2 e início da região 3. Na região 4 ocorre uma superposição desses valores.

- as variáveis magnitude do gradiente linear e magnitude do gradiente apresentam superposição dos conjunto de valores referentes aos pixels interno, sobre contorno e externo nas 4 regiões em que dividimos o VE. Por essa razão é que, nos gráficos dessas variáveis, foram mostrados apenas 3 valores por orientação, correspondentes aos valores médios.

- a partir dos valores da tabela 4.4 observamos, para as oito orientações apresentadas, que a variável distância em relação ao centro do VE apresenta 100% dos valores com diferenças significativas; a variável intensidade do pixel apresenta 81,25% dos valores com diferenças significativas; a variável gradiente linear apresenta 87,25% dos valores com diferenças significativas; a variável gradiente linear apresenta 25% dos valores com diferenças significativas e a variável distância em relação ao pixel de contorno da orientação anterior apresenta 100% dos valores com diferenças significativas.

Em face dos valores de significância apresentados acima, resolvemos não utilizar a variável magnitude do gradiente como entrada da rede. Assim, utilizamos apenas 4 variáveis de entrada da rede: distância do pixel em relação ao centro do VE normalizada, intensidade do pixel normalizada, magnitude do gradiente linear normalizada e distância do pixel em relação ao pixel de contorno da orientação anterior normalizada. O conjunto constituído por essas quatro variáveis é designado por:

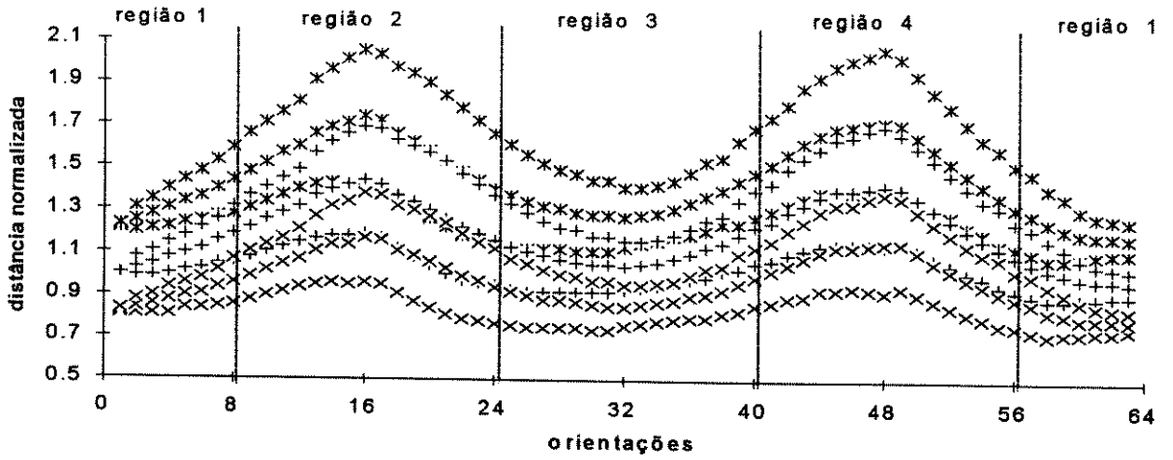
$$\mathrm{Conj}_A = \{d_n, I_n, g_n, da_n\} \tag{3.14}$$

A rede que utiliza essas variáveis de entrada é designada de rede A. Como mostraremos adiante, no capítulo de resultados (os motivos para esse comportamento serão analisados no capítulo de discussão), a rede A não apresentou bons resultados. Por isso, resolvemos utilizar uma outra rede com 3 variáveis de entrada: distância em relação ao centro do VE normalizada, intensidade do pixel normalizada e magnitude do gradiente linear normalizado. Este conjunto de variáveis de entrada é designado por:

$$\mathrm{Conj}_B = \{d_n, I_n, g_n\} \tag{3.15}$$

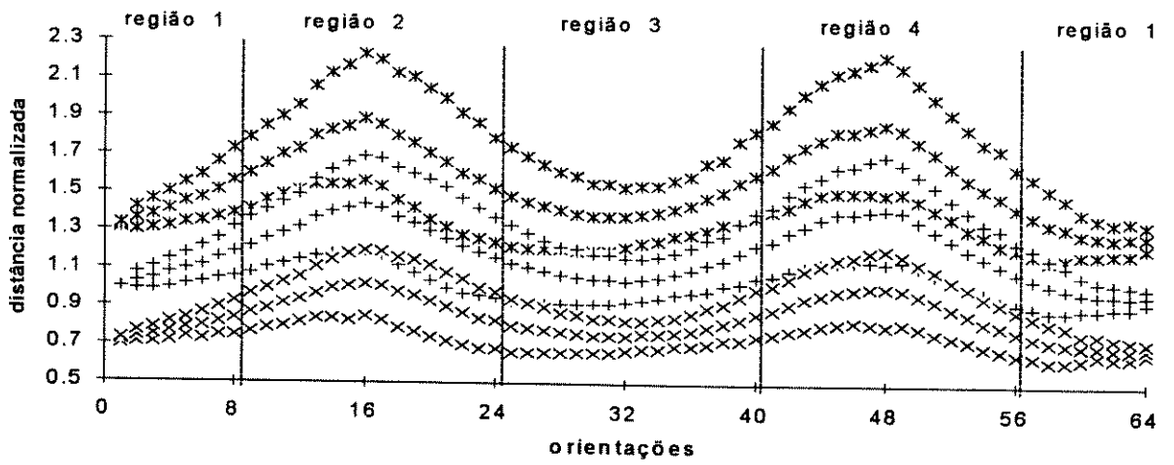
A rede que utiliza essas variáveis de entrada é designada de rede B.

No treinamento das redes, a saída foi feita bipolar (±1). Pixels situados sobre o contorno forçam a saída da rede para +1, enquanto que pixels internos ou externos forçam a saída da rede para -1. No treinamento das redes neurais utilizamos o algoritmo de propagação reversa, descrito no capítulo 2 e pormenorizado no apêndice A. Com o objetivo de evitarmos mínimos locais, as imagens foram apresentadas aleatoriamente às redes. Adotamos também um algoritmo de aprendizado tipo *batch* em que a atualização dos pesos é feita apenas após a apresentação das 20 imagens do grupo de treinamento. Com o objetivo de acelerar a convergência do treinamento, utilizamos o método de adaptação global da taxa de aprendizado (Vogl et alii, 1988). Esse método consiste em uma estratégia heurística simples para o aprendizado tipo *batch*: se o valor do erro médio quadrático na saída da rede decresce após uma iteração, aumenta-se o valor da taxa de aprendizado η (tipicamente multiplicando-se seu valor por um fator a=1,05); se o valor do erro médio quadrático na saída da rede aumenta após uma iteração, diminui-se o valor da taxa de aprendizado η (tipicamente



+ pixels do contorno, x pixels internos, * pixels externos

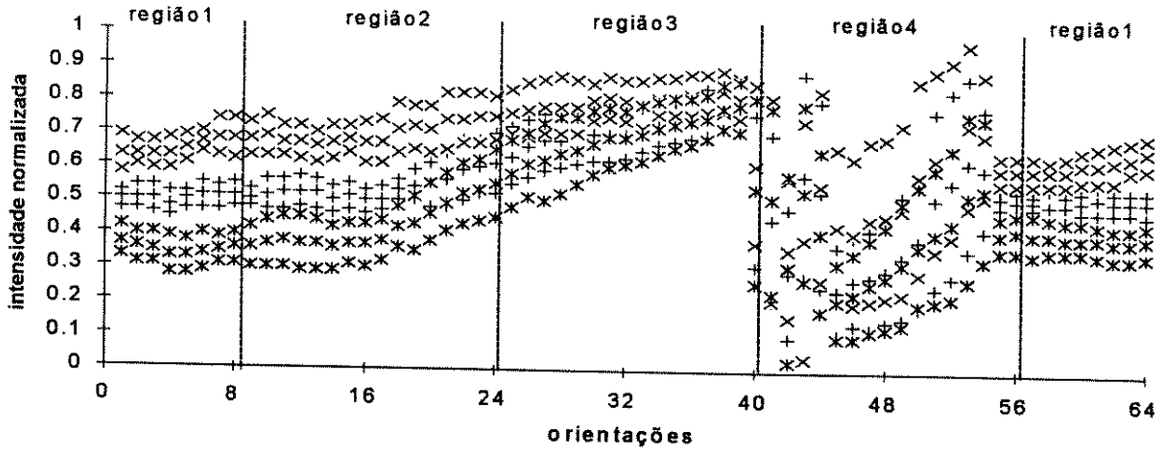
(a)



+ pixels do contorno, x pixels internos, * pixels externos

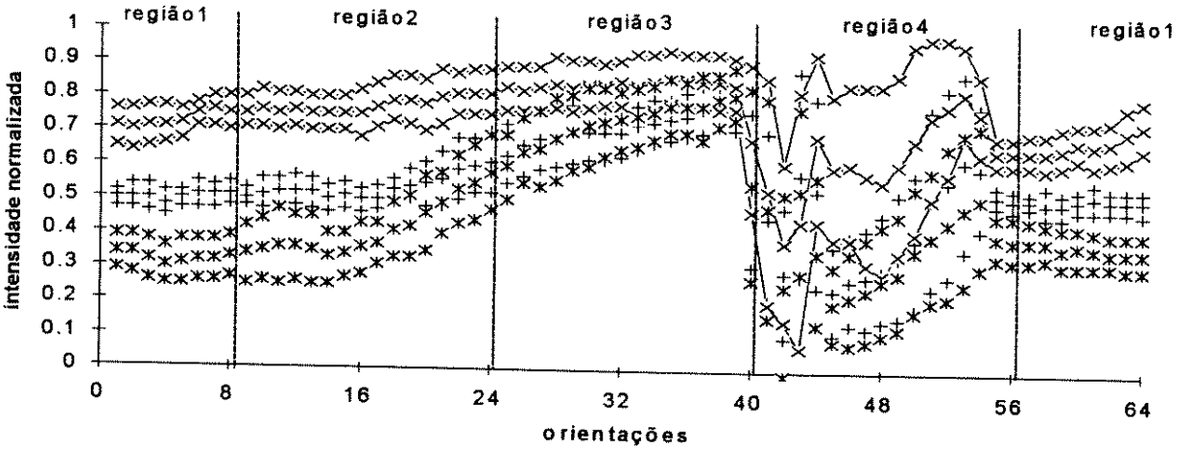
(b)

Figura 3.16: Gráfico dos valores médios normalizados das distâncias em relação ao centro do VE e dos desvios padrões dos pixels interno, do contorno, e externo, para as 60 imagens de VE e para as 64 orientações. (a) pixel interno: 0,8 vezes a distância do pixel de contorno e pixel externo: 1,2 vezes a distância do pixel de contorno. (b) pixel interno: 0,7 vezes a distância do pixel de contorno e pixel externo: 1,3 vezes a distância do pixel de contorno.



+ pixels do contorno, x pixels internos, * pixels externos

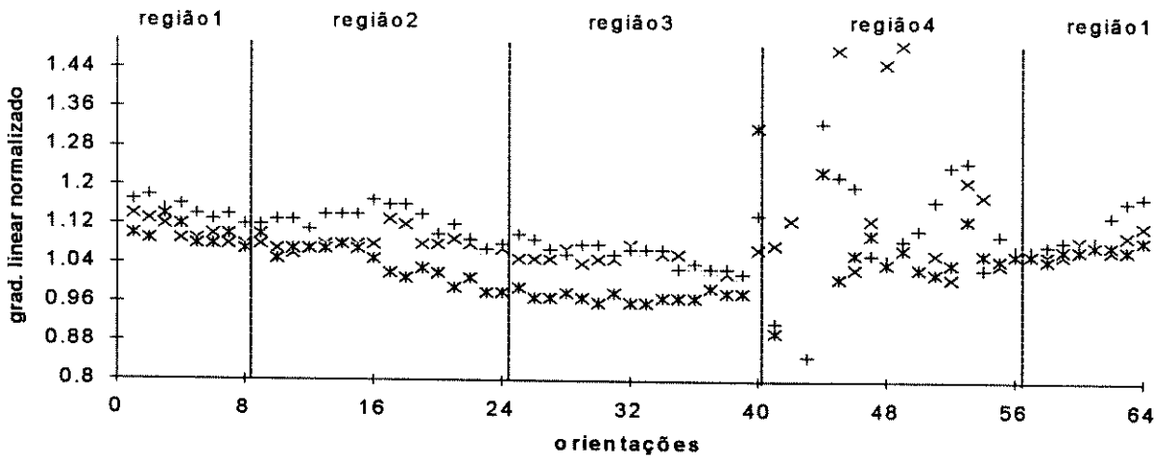
(a)



+ pixels do contorno, x pixels internos, * pixels externos

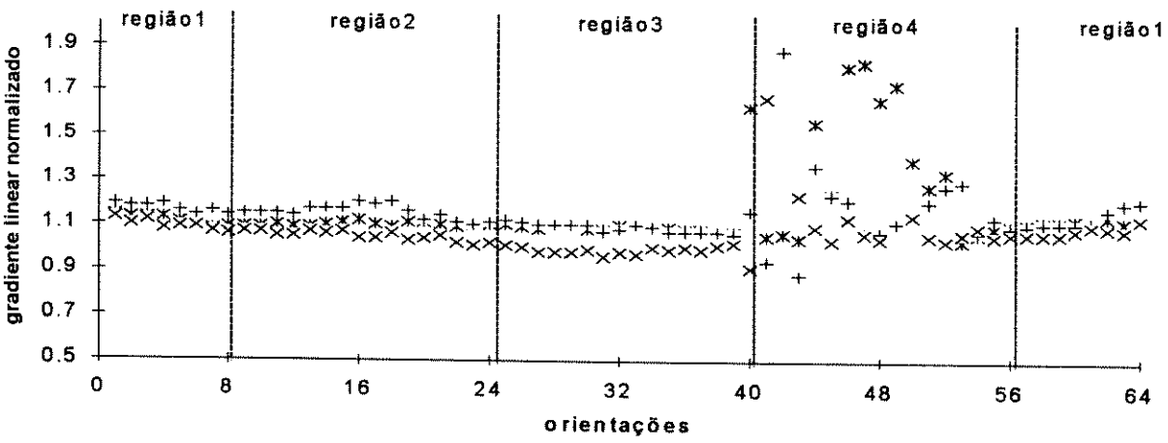
(b)

Figura 3.17: Gráfico dos valores médios normalizados das intensidades e dos desvios padrões dos pixels interno, do contorno e externo, para as 60 imagens de VE e para as 64 orientações. (a) pixel interno: 0,8 vezes a distância do pixel de contorno e pixel externo: 1,2 vezes a distância do pixel e contorno. (b) pixel interno: 0,7 vezes a distância do pixel de contorno e pixel externo: 1,3 vezes a distância do pixel de contorno.



+ pixels do contorno, x pixels internos, * pixels externos

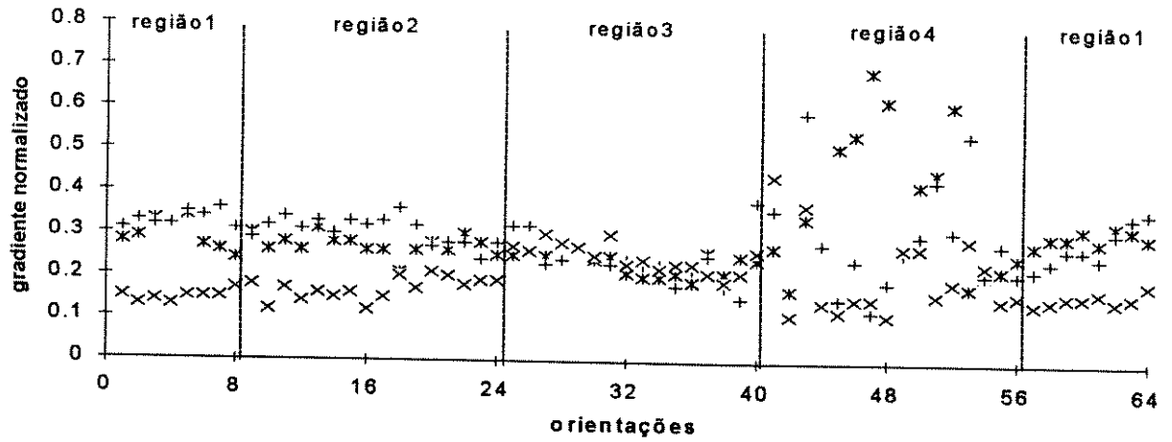
(a)



+ pixels do contorno, x pixels internos, * pixels externos

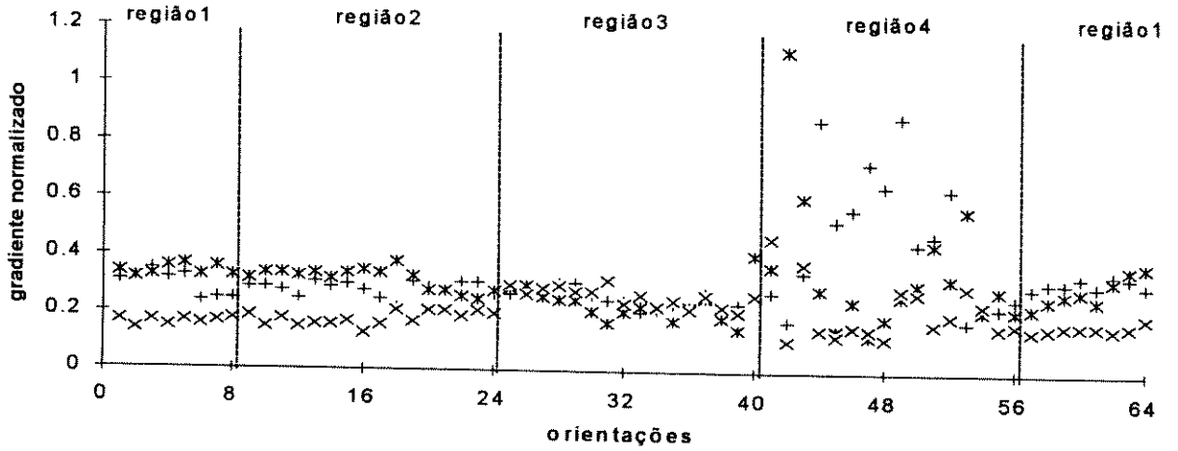
(b)

Figura 3.18: Gráfico dos valores médios normalizados do gradiente linear dos pixels interno, do contorno e externo, para as 60 imagens de VE e para as 64 orientações. (a) pixel interno: 0,8 vezes a distância do pixel de contorno e pixel externo: 1,2 vezes a distância do pixel de contorno. (b) pixel interno: 0,7 vezes a distância do pixel de contorno e pixel externo: 1,3 vezes a distância do pixel de contorno.



+ pixels do contorno, x pixels internos, * pixels externos

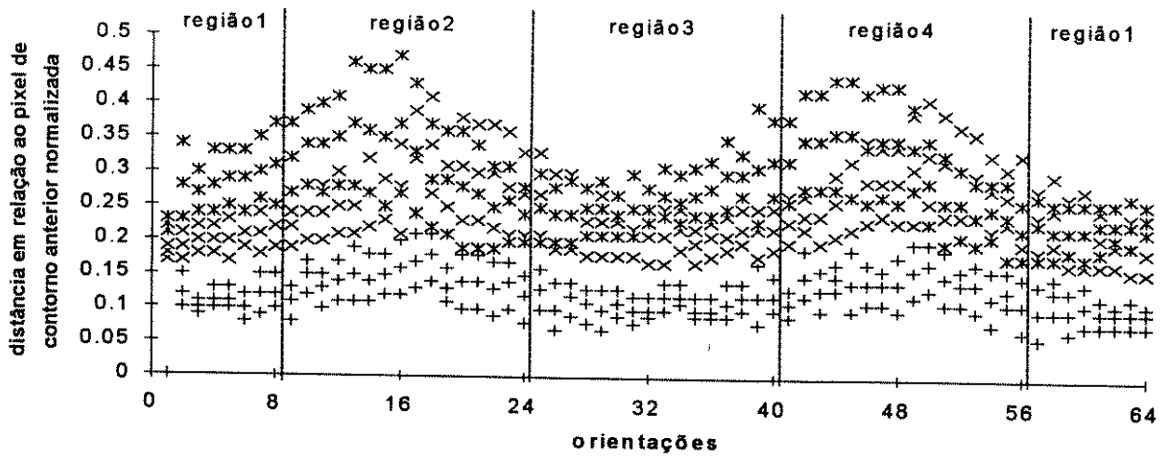
(a)



+ pixels do contorno, x pixels internos, * pixels externos

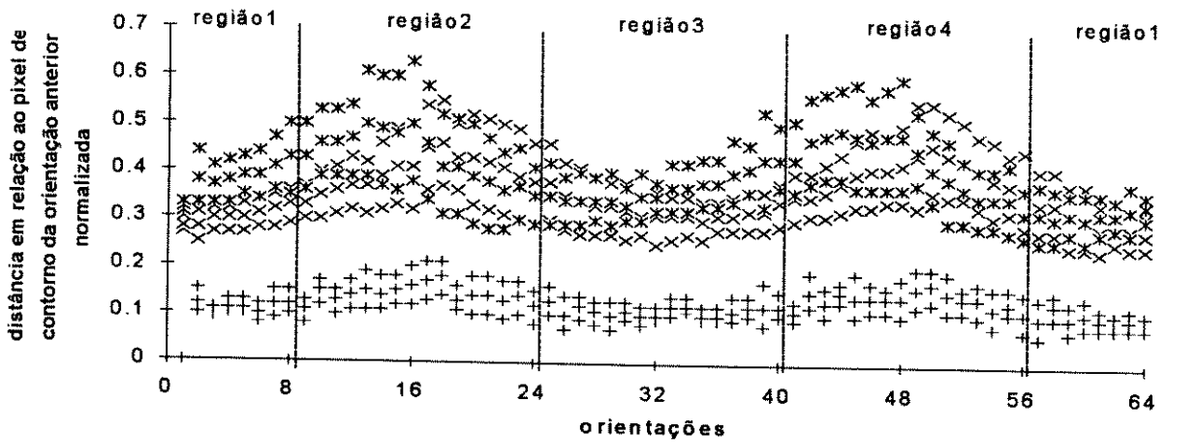
(b)

Figura 3.19: Gráfico dos valores médios normalizados da magnitude do gradiente dos pixels interno, do contorno e externo, para as 60 imagens de VE e para as 64 orientações. (a) pixel interno: 0,8 vezes a distância do pixel de contorno e pixel externo: 1,2 vezes a distância do pixel de contorno. (b) pixel interno: 0,7 vezes a distância do pixel de contorno e pixel externo: 1,3 vezes a distância do pixel de contorno.



+ pixels do contorno, x pixels internos, * pixels externos

(a)



+ pixels do contorno, x pixels internos, * pixels externos

(b)

Figura 3.20: Gráfico dos valores médios normalizados das distâncias em relação ao pixel do contorno da orientação anterior e dos desvios padrões dos pixels interno, do contorno e externo, para as 60 imagens de VE e para as 64 orientações. (a) pixel interno: 0,8 vezes a distância do pixel de contorno e pixel externo: 1,2 vezes a distância do pixel de contorno. (b) pixel interno: 0,7 vezes a distância do pixel de contorno e pixel externo: 1,3 vezes a distância do pixel de contorno.

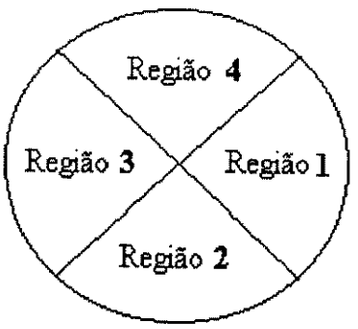


Figura 3.21: divisão do VE em 4 regiões.

multiplicando-se seu valor por 0,7) se o novo erro excede o anterior de mais do que uma razão pré-especificada (tipicamente 1,04).

Após iniciarmos a taxa de aprendizado η com um valor pequeno (tipicamente 0,0001), a modificação da mesma obedece à seguinte equação iterativa:

$$\eta^k = \begin{cases} a.\eta^{(k-1)}, \text{ se } E(w^k) < E(w^{(k-1)}) \\ b.\eta^{(k-1)}, \text{ se } E(w^k) \geq \alpha.E(w^{(k-1)}) \\ \eta^{(k-1)}, \text{ caso contrário} \end{cases} \tag{3.16}$$

onde:

k = iteração de ordem k;

$E(w^k)$ = erro médio quadrático da saída da rede para as 20 imagens

O erro médio quadrático na saída da rede, computado após a apresentação dos 60 pixels correspondentes às 20 imagens de treinamento, é definido como:

$$E(w^k) = \frac{1}{60} \cdot \sum_{j=0}^{j<60} (d_j - s_j)^2 \tag{3.17}$$

onde:

d_j = saída desejada da rede (+1 ou -1) para o pixel j;

s_j = saída real da rede para o pixel j.

Na figura 3.22 mostramos a variação do erro médio quadrático para vários valores dos parâmetros a e b, para a orientação θ=22,5°. O valor de α utilizado para obtenção desse dados foi de 1,04 e o valor inicial de η foi de 0,0001.

Tabela 3.3: Valores médios normalizados e desvios padrões das variáveis de entrada da rede para as 60 imagens de VE e para 8 orientações: (a) orientações 0,8,16 e 24. (b) orientações 32,40,48 e 56.

(a)

		orientações							
		0		8		16		24	
variáveis		média	DP	média	DP	média	DP	média	DP
distância em relação ao centro do VE normalizada	pixel interno	0,72	0,02	0,84	0,09	1,02	0,17	0,83	0,15
	pixel do contorno	1	0	1,19	0,13	1,44	0,25	1,16	0,21
	pixel externo	1,32	0,01	1,56	0,17	1,89	0,33	1,52	0,28
intensidade normalizada	pixel interno	0,77	0,07	0,81	0,09	0,81	0,13	0,85	0,12
	pixel do contorno	0,57	0,06	0,59	0,11	0,57	0,14	0,65	0,16
	pixel externo	0,42	0,09	0,41	0,13	0,45	0,17	0,65	0,17
gradiente linear normalizado	pixel interno	1,14	0,09	1,09	0,09	1,12	0,1	1,1	0,09
	pixel do contorno	1,29	0,1	1,24	0,1	1,2	0,13	1,21	0,11
	pixel externo	1,13	0,13	1,06	0,12	1,04	0,14	1,02	0,12
gradiente normalizado	pixel interno	0,31	0,14	0,25	0,16	0,28	0,15	0,27	0,17
	pixel do contorno	0,34	0,13	0,33	0,16	0,35	0,14	0,28	0,16
	pixel externo	0,17	0,12	0,18	0,09	0,13	0,11	0,2	0,14
distância em relação ao pixel de contorno da orientação anterior normalizada	pixel interno	0,28	0,01	0,33	0,04	0,41	0,09	0,38	0,09
	pixel do contorno	0	0	0,12	0,02	0,16	0,04	0,12	0,04
	pixel externo	0,32	0,01	0,43	0,08	0,5	0,12	0,35	0,06

(b)

		orientações							
		32		40		48		56	
variáveis		média	DP	média	DP	média	DP	média	DP
distância em relação ao centro do VE normalizada	pixel interno	0,75	0,08	0,87	0,12	1,01	0,2	0,78	0,12
	pixel do contorno	1,04	0,12	1,23	0,17	1,42	0,28	1,09	0,16
	pixel externo	1,38	0,16	1,6	0,22	1,87	0,37	1,43	0,21
intensidade normalizada	pixel interno	0,87	0,09	0,72	0,23	0,61	0,29	0,69	0,1
	pixel do contorno	0,74	0,11	0,6	0,25	0,38	0,2	0,56	0,11
	pixel externo	0,82	0,14	0,65	0,33	0,37	0,23	0,47	0,15
gradiente linear normalizado	pixel interno	1,1	0,1	1,63	2,48	1,66	1,03	1,1	0,08
	pixel do contorno	1,18	0,1	1,16	0,46	1,07	0,16	1,19	0,1
	pixel externo	0,98	0,11	0,91	0,33	1,04	0,21	1,06	0,12
gradiente normalizado	pixel interno	0,25	0,15	0,27	0,33	0,65	0,97	0,26	0,13
	pixel do contorno	0,21	0,12	0,41	0,66	0,19	0,14	0,22	0,11
	pixel externo	0,24	0,13	0,27	0,32	0,12	0,09	0,17	0,11
distância em relação ao pixel de contorno da orientação anterior normalizada	pixel interno	0,31	0,06	0,35	0,05	0,42	0,08	0,37	0,08
	pixel do contorno	0,1	0,01	0,11	0,02	0,14	0,04	0,11	0,04
	pixel externo	0,35	0,03	0,43	0,06	0,48	0,11	0,32	0,05

Tabela 3.4: valores do teste t para as oito orientações mostradas na tabela 3.3: (a) orientações 0,8,16 e 24 e (b) orientações 32,40,48 e 56. /S representa significativo enquanto /N representa não significativo.

(a)

		orientações			
		0	8	16	24
variáveis		teste t/sig	teste t/sig	teste t/sig	teste t/sig
distância em relação ao centro do VE normalizada	pixel contorno e interno	109,5 /S	17,1 /S	10,8 /S	9,9 /S
	pixel contorno e externo	247,9 /S	13,4 /S	8,4 /S	7,9 /S
intensidade normalizada	pixel contorno e interno	16,8 /S	12,0 /S	9,7 /S	7,7 /S
	pixel contorno e externo	10,7 /S	8,2 /S	4,2 /S	0,0 /S
gradiente linear normalizado	pixel contorno e interno	8,6 /S	8,6 /S	3,8 /S	5,4 /S
	pixel contorno e externo	7,5 /S	8,9 /S	6,5 /S	9,0 /S
gradiente normalizado	pixel contorno e interno	1,2 /N	2,7 /N	2,6 /N	0,33 /N
	pixel contorno e externo	7,4 /S	6,3 /S	9,6 /S	2,9 /N
distância em relação ao pixel da orientação anterior normalizada	pixel contorno e interno	216,0 /S	29,1 /S	20,8 /S	24,7 /S
	pixel contorno e externo	247,0 /S	29,1 /S	20,8 /S	24,7 /S

(b)

		orientações			
		32	40	48	56
variáveis		teste t/sig	teste t/sig	teste t/sig	teste t/sig
distância em relação ao centro do VE normalizada	pixel contorno e interno	15,57 /S	13,40 /S	9,34 /S	12,11 /S
	pixel contorno e externo	13,17 /S	10,30 /S	7,51 /S	9,97 /S
intensidade normalizada	pixel contorno e interno	7,08 /S	2,73 /N	5,06 /S	6,77 /S
	pixel contorno e externo	3,48 /S	0,94 /N	0,25 /N	3,74 /S
gradiente linear normalizado	pixel contorno e interno	4,38 /S	1,44 /N	4,38 /S	5,44 /S
	pixel contorno e externo	10,42 /S	3,42 /S	0,88 /N	6,44 /S
gradiente normalizado	pixel contorno e interno	1,61 /N	1,47 /N	3,63 /S	1,81 /N
	pixel contorno e externo	1,31 /N	1,47 /N	3,25 /N	2,48 /N
distância em relação ao pixel da orientação anterior normalizada	pixel contorno e interno	26,74 /S	34,52 /S	24,25 /S	22,51 /S
	pixel contorno e externo	61,24 /S	39,19 /S	22,50 /S	25,40 /S

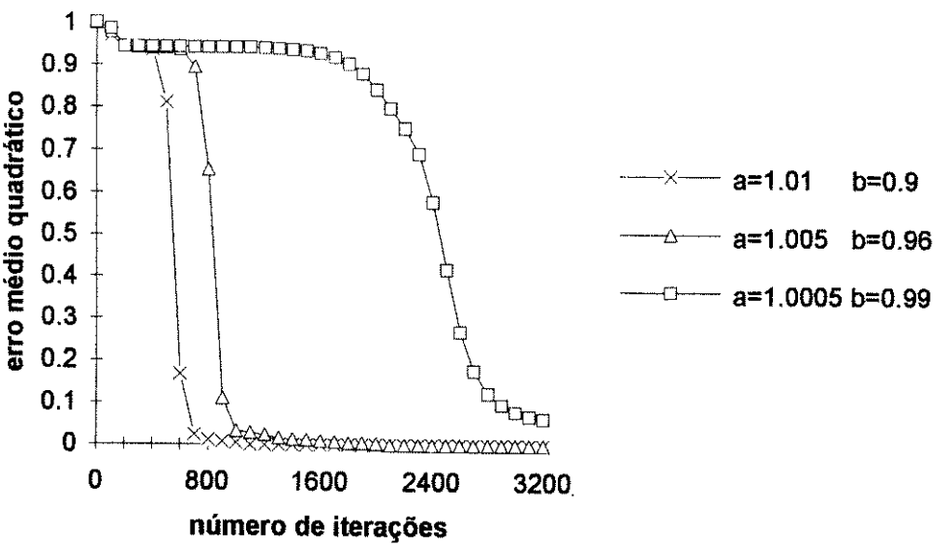


Figura 3.22: variação do erro médio quadrático para um treinamento da rede correspondente a orientação de 22,5°.

Em algumas orientações, um valor de **a** grande e um valor de **b** pequeno produziam instabilidade. Porém, quando a rede não se tornava instável, um par de valores desses produziam uma convergência mais rápida, conforme pode ser visto na figura 3.22. Objetivando minimizar o tempo de treinamento em cada orientação, adotamos a estratégia de começar o treinamento com um valor grande de **a** e um valor pequeno de **b**. Se o erro médio quadrático tornava-se menor do que 0,2, considerávamos o treinamento concluído. Por outro lado, se a rede tornava-se instável, reiniciávamos os valores dos pesos, aumentávamos o valor de a, diminuíamos o valor de b e refazíamos o treinamento para aquela orientação.

O tempo de treinamento variou muito em função da orientação. Na tabela 3.5 mostramos, para a rede B, o tempo médio e o número de ciclos por treinamento para cada uma das regiões em que dividimos o VE na figura 3.21. O tempo de treinamento da rede A foi extremamente mais rápido, sendo tipicamente menor do que 30 segundos por orientação.

Tabela 3.5: número médio de ciclos por treinamento e tempo de treinamento para a rede B (em minutos) em função da região do ventrículo.

No. médio de ciclos p/ treinamento	4000	6000	3000	6000
Tempo médio de treinamento(min)	2	3	1,5	3

III.7 UTILIZAÇÃO DA REDE NEURAL E CRITÉRIOS DE SELEÇÃO DOS PIXELS DE CONTORNO

A fim de se submeter as imagens ao processo de extração do contorno do VE, as mesmas foram primeiramente pré-processadas, conforme descrito anteriormente neste capítulo, e o centro do VE foi extraído automaticamente. As 64 orientações foram, então computadas, tendo como referência esse centro. Em seguida determinamos o raio do contorno para a orientação de 0°. Nessa orientação, as redes A e B trabalham apenas com duas variáveis de entrada: a intensidade do pixel e a magnitude do gradiente linear. A razão disso é que por ser a primeira orientação ainda não temos o valor utilizado para normalizar as distâncias dos pixels em relação ao centro do VE. Se utilizássemos, para a orientação 0°, as distâncias dos pixels em relação ao centro do VE (sem estarem normalizadas) como a terceira variável de entrada da rede neural, a mesma não convergiria, pois essas distâncias variam muito entre ventrículos esquerdos de pessoas diferentes. A variação observada nessas distâncias situa-se numa faixa entre 20 e 40 pixels.

Ao utilizarmos a rede B, obtivemos, para a orientação θ=22,5°, a saída mostrada na figura 3.23. Em uma orientação particular, conforme descrito na introdução deste capítulo, aqueles pixels que ao serem apresentados a rede levam a saída da mesma a um valor maior do que o limiar $\mathbf{s_L}$=0,8 são considerados *pixels candidatos* ao contorno. Se nenhum pixel atende a exigência anterior, subtraímos 0,1 do valor de $\mathbf{s_L}$ e uma nova busca de *pixels candidatos* é efetuada, num processo iterativo. A escolha dos pixels de contorno em uma orientação i' é efetuada a partir do conjunto dos pixels candidatos dessa orientação, utilizando um entre os três critérios:

- critério #1: o pixel de contorno é aquele que maximiza o valor de $\mathbf{c_1}$ dado por:

$$c_1 = \frac{I_{j-2} + I_{j-1}}{I_{j+2} + I_{j+1}} \tag{3.18}$$

onde: $P_j \in Cand_{i'}$ e

$I_{j-2}, I_{j-1}, I_{j+1}$ e I_{j+2} são as intensidades dos pixels $P_{j-2}, P_{j-1}, P_{j+1}$ e P_{j+2} da orientação i', respectivamente.

- critério #2: as coordenadas do pixel de contorno, $\mathbf{xcont_{i'}}$ e $\mathbf{ycont_{i'}}$ são obtidas através de uma média das coordenadas **x** e **y** de todos os pixel do conjunto de *pixels candidatos* da orientação **i'**, ou seja:

$$xcont_{i'} = \frac{1}{m} \cdot \sum_{j=0}^{j=m-1} x_j \quad e$$

$$ycont_{i'} = \frac{1}{m} \cdot \sum_{j=0}^{j=m-1} y_j \tag{3.19}$$

onde: $P_j \in Cand_i$ e

m = número de pixels em $Cand_i$.

- critério #3: as coordenadas do pixel de contorno são obtidas através de uma média das coordenadas dos pixels determinados pelos critérios 1 e 2.

O primeiro desses critérios explora as características da descontinuidade da borda em termos de níveis de cinza. A expectativa desse critério é que a borda esteja localizada em uma região onde ocorra uma transição abrupta de níveis de cinza. O segundo critério explora o fato dos pixels próximos a borda do VE terem uma tendência a apresentarem uma saída da rede neural com valor alto. Esse fato pode ser observado através do gráfico da figura 3.23, onde verificamos que na região da borda (próximo ao pixel de número 35), os valores da saída da rede têm um aumento significativo. O terceiro critério associa os dois anteriores, ao obter uma média das coordenadas dos pixels de contorno determinados por ambos. O desempenho desses critérios é apresentado no capítulo de resultados.

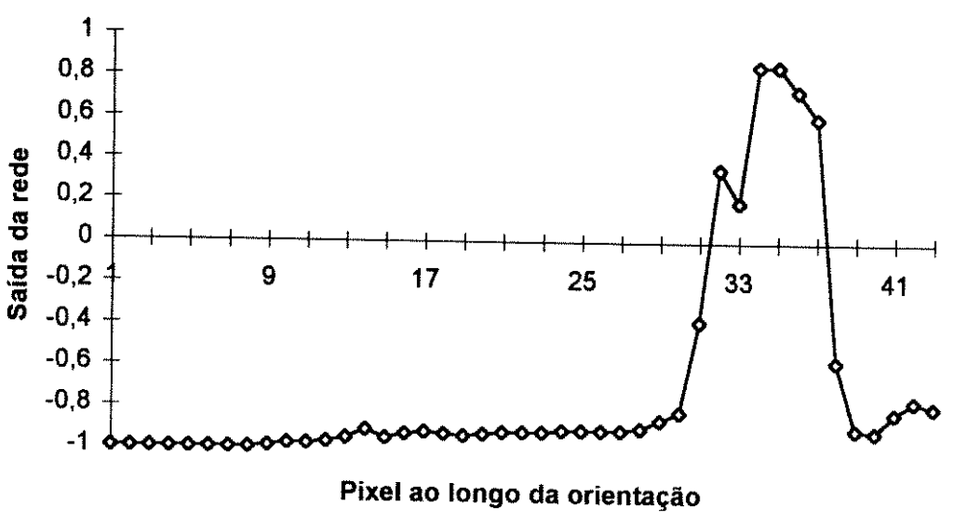


Figura 3.23: saída da rede A para a orientação de 22,5°.

III.8 RESTRIÇÕES GEOMÉTRICAS E FILTRAGEM

A procura dos pixels candidatos em algumas regiões, como a região das válvulas do VE ou a região onde o VE faz interface com o músculo papilar, é bastante difícil. Com o objetivo de direcionar essa busca, restringimos, para cada orientação, a região onde os *pixels candidatos*

podem ser procurados, através do seguinte procedimento: utilizando os contornos traçados manualmente para as 60 imagens de VE, obtivemos, para cada orientação, um valor para o raio médio normalizado do contorno e um valor para o desvio padrão. Confinamos, então a busca dos *pixels candidatos* em cada orientação a um intervalo definido pelo raio médio naquela orientação ± 3 desvios padrões. Esse intervalo engloba 99% dos pixels dos contornos traçados manualmente.

Após a determinação dos *pixels candidatos* e dos 64 *pixels de contorno* que definem o contorno-θ de um VE, efetuamos ainda um processo de filtragem e suavização.

O processo de filtragem tem por objetivo eliminar pontos do contorno que representem possível ruído. Um pixel de contorno de uma orientação **i** é considerado ruído, se a sua distância em relação ao pixel de contorno da orientação anterior **i-1** é maior do que um determinado valor pré-estabelecido. A seguir descrevemos o algoritmo que empregamos para determinar se um pixel de contorno de uma orientação **i** pode ou não ser considerado como ruído:

A. utilizando os contornos manuais das 60 imagens de VE, obtivemos, para as 64 orientações, o valor máximo da distância de um pixel de contorno numa orientação **i** em relação ao pixel de contorno da orientação anterior **i-1**. Os valores dessas distâncias foram, em cada imagem, normalizados em relação ao raio da orientação de 0°. O conjunto dessas distâncias é representado por:

$$D_{max} = \{d\max_0, d\max_1, d\max_2, \ldots\ldots, d\max_{63}\} \qquad (3.20)$$

B. para cada uma das 64 orientações de um contorno determinado automaticamente, calculamos a distância do pixel de contorno de uma orientação **i** em relação ao pixel de contorno da orientação **i-1**.

C. se essa distância for maior do que $\mathbf{dmax_i}$ ir para o passo D, caso contrário assumir que o pixel de contorno da orientação **i** não é ruído ir para G.

D. calcular a distância $\mathbf{d_2}$ do pixel de contorno da orientação **i+1** em relação ao pixel de contorno da orientação **i-1** e a distância $\mathbf{dmax_{i+1}}$ do pixel de contorno da orientação **i+1** em relação ao pixel de contorno da orientação **i**. Se $\mathbf{d_2}$ for maior que a soma das distâncias $\mathbf{dmax_i}$ e $\mathbf{dmax_{i+1}}$, ir para F, caso contrário ir para E.

E. calcular as distâncias dos pixels de contorno das orientações **i-1** e **i+1** em relação ao centro do VE. Substituir o pixel de contorno da orientação **i** por outro pixel nessa mesma orientação cuja distância em relação ao centro do VE seja uma média das duas distâncias anteriormente calculadas. Ir para G.

F. calcular a distância do pixel de contorno da orientação **i-1** em relação ao centro do VE. Substituir o pixel de contorno da orientação **i** por outro pixel de contorno nesta mesma orientação cuja distância em relação ao centro do VE seja igual a distância anteriormente calculada.

G. final do algoritmo.

Uma característica dos contornos traçados manualmente é serem suaves, enquanto que os contornos determinados automaticamente são, na maioria das vezes, marcados por descontinuidades. A fim de darmos um aspecto mais suave aos contornos que obtivemos utilizamos um filtro de média móvel, definido por:

$$R_i = (r_{i-2} + 2r_{i-1} + 4r_i + 2r_{i+1} + r_{i+2})/10, \quad \text{para } i=0,1,2,...N \tag{3.21}$$

onde:

r_i= raio a ser suavizado;
R_i= raio suavizado correspondendo ao contorno suavizado;
N= 63.

Uma vez que o contorno é fechado, observamos que $r_{-i} = r_{N+1-i}$ e $r_{N+i} = r_{i-1}$.

Após a suavização, o contorno final é obtido interpolando-se linearmente os pixels do contorno resultante dos processos de filtragem e suavização. Nas figuras 3.24 e 3.25 mostramos um exemplo da seqüência de procedimentos que resultam na detecção automática do contorno do VE. Esse contorno foi obtido utilizando-se a rede B. Na figura 3.24 (a) mostramos uma imagem de VE pré-processada, com 16 das 64 orientações superpostas a mesma. Na figura 3.24 (b) mostramos os *pixels candidatos* em todas as orientações. Na figura 3.24 (c) mostramos os *pixels de contorno* que foram detectados utilizando o critério #3. Na figura 3.24 (d) observamos o contorno do VE após a filtragem. Finalmente, na figura 3.25 mostramos a imagem final desse processo: o contorno do VE após a suavização.

III.9 AVALIAÇÃO DO DESEMPENHO DO MÉTODO

Para a avaliação do desempenho do método proposto, utilizamos as medidas de erro percentual e de erro soma, definidos através das expressões 2.20 e 2.21, respectivamente. Além de permitir uma comparação com outros trabalhos da literatura, essas medidas levam em conta a área do VE, parâmetro essencial para a avaliação da função de ejeção e para o cálculo do volume do VE.

III.10 SOFTWARE E HARDWARE UTILIZADOS

O software desenvolvido nesse trabalho foi implementado em linguagem C (Kernighan et alia, 1990), utilizando o compilador C++ da Borland. O programa completo pode ser dividido em seis módulos, mostrados na tabela 3.6 com os respectivos números de linhas. A seguir descreveremos as funções de cada um desses módulos:

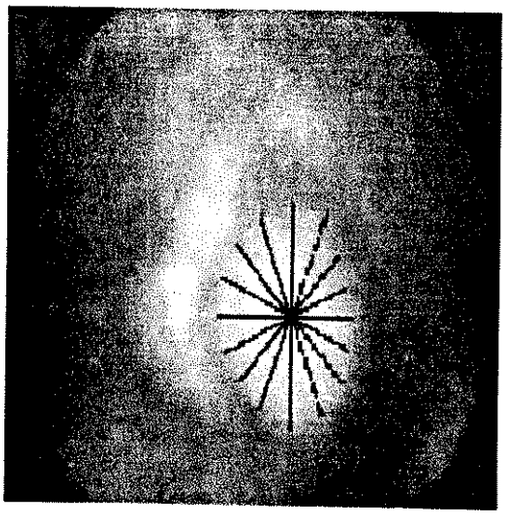
- módulo de leitura das imagens. A série de 32 imagens de *blood-pool* de VE sincronizadas com o ciclo cardíaco, são fornecidas em arquivo único, com um formato próprio da empresa que fornece o equipamento. Utilizando as informações presentes no cabeçalho desse arquivo, o módulo de leitura de imagens captura a primeira imagem da série, correspondente a diástole e a grava em um arquivo separado para ser processada pelo módulo de pré-processamento.

- módulo de pré-processamento e extração dos contornos. Esse módulo é constituído de duas partes: a primeira, correspondente ao pré-processamento, efetua as tarefas de ajustar o histograma da imagem proveniente do módulo anterior de extração e ampliar a mesma de 64x64 para 256x256; na segunda parte, correspondente a extração de contornos, efetua-se um processamento iterativo, onde o centro e o contorno do VE são extraídos pelo usuário utilizando um dispositivo apontador (*mouse*). Após a extração do contorno pelo usuário, o programa seleciona 64 pixels do mesmo, correspondendo as 64 orientações. Como resultado desse módulo são gerados dois arquivos: um arquivo *X* contendo a imagem filtrada e ampliada e o outro arquivo *Y* contendo as informações do contorno e do centro extraídos manualmente.

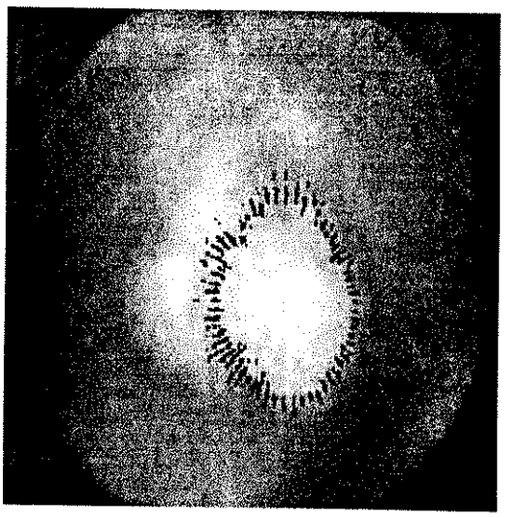
- módulo de determinação automática do centro. Nesse módulo, solicitamos que o usuário forneça inicialmente o nome da imagem. Em seguida, utilizando as informações extraídas no módulo anterior sobre o contorno e o centro do VE, calculamos o centro médio dos ventrículos esquerdo e direito e determinamos a equação da reta da mediatriz do segmento de reta que une esses dois centros médios. Através de uma operação de limiar, isolamos os pixels com intensidades maiores do que 200 e estimamos o centro do VE calculando a média das coordenadas desses pixels. Finalmente, efetuamos o cálculo do valor refinado do centro do VE. A informação do centro assim determinado é anexada ao arquivo *Y* definido anteriormente.

- módulo de treinamento e uso da rede auto-organizável de Kohonen. Nesse módulo, o usuário tem a opção de fornecer o número de grupos em que deseja classificar as 64 imagens de VE, bem como o número de interações a ser utilizado no processo de classificação. A saída desse módulo é uma relação da classificação desses contornos nos respectivos grupos.

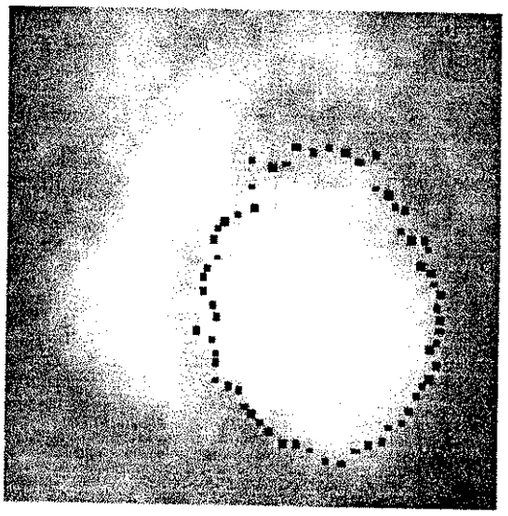
- módulo de treinamento e uso da rede tipo *perceptron* multicamadas. Fornecemos inicialmente ao usuário um *menu* com as opções de seleção das variáveis de entrada da rede e escolha entre treinamento ou uso da rede neural. No caso do usuário optar pelo treinamento, o programa permite, para o treinamento de cada orientação, a escolha dos parâmetros **a** e **b** da equação 3.16; a visualização da saída da rede após um número pré-estabelecido de iterações e recomeçar o treinamento naquela orientação caso o mesmo tenha se tornado instável. No caso do usuário optar pelo uso da rede, o programa pergunta o nome da imagem que se deseja determinar o contorno do VE; se o centro a ser utilizado é o manual ou o automático e qual o critério a ser utilizado na seleção dos *pixels de contorno* a partir dos *pixels candidatos*. O resultado da opção treinamento é um arquivo contendo os valores dos pesos da rede treinada. Como resultado da opção uso da rede, o programa fornece a imagem do contorno do VE com os pixels candidatos; a imagem do contorno do VE com os pixels de contorno; a imagem do contorno do VE filtrado; a imagem do contorno do VE suavizado e a imagem contendo a superposição dos contornos do VE extraídos manualmente e obtidos automaticamente. Além disso, obtemos também uma tabela contendo os valores do erro soma e do erro percentual resultantes da comparação entre os contornos manual e automático.



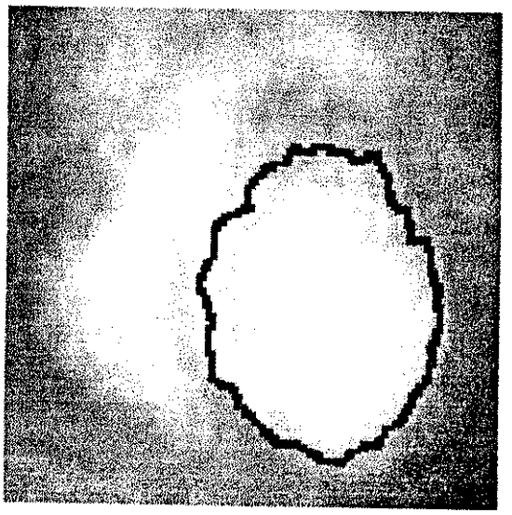
(a)



(b)



(c)



(d)

Figura 3.24: Exemplo de detecção de contorno: (a) imagem pré-processada de VE com 16 orientações superpostas;(b) imagem mostrando os pixels candidatos;(c) imagem com os pixels de contorno detectados através do critério #3 e (d) imagem com os pixels de contorno obtidos após a filtragem e interpolados linearmente.

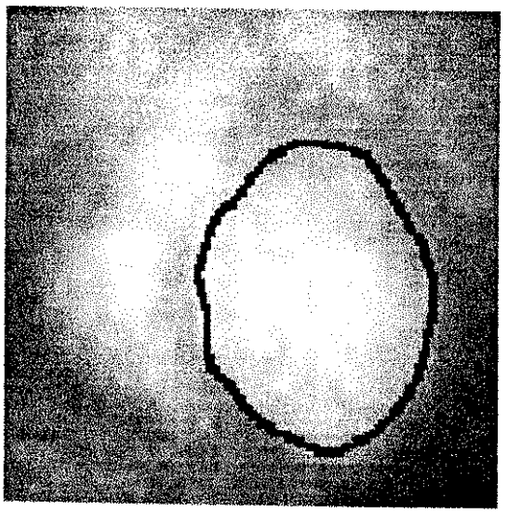


Figura 3.25: Exemplo de detecção do contorno do VE: imagem com o contorno do VE após a suavização.

Tabela 3.6: módulos do programa completo.

módulo de programa	número de linhas
1. Extração da imagem	190
2. Pré-processamento e extração dos contornos	795
3. Determinação automática do centro	959
4. Treinamento e uso da rede auto-organizável de Kohonen	257
5. Treinamento e uso da rede tipo perceptron multicamadas	2534
6. Confecção dos gráficos das variáveis de entrada da rede	530

- módulo de confecção dos gráficos das variáveis de entrada da rede: O usuário tem condições de visualizar os gráficos mostrados nas figuras 3.16-3.20. O programa requer que seja fornecido o nome da variável a ser visualizada. Na saída do mesmo é gerado um arquivo para ser utilizado pelo programa *Excel* da *Microsoft*®.

    - computador utilizado foi um PC 486 de 66Mhz, 16Mbytes de RAM e 1Mbyte de capacidade de armazenamento em meio magnético.

CAPÍTULO IV

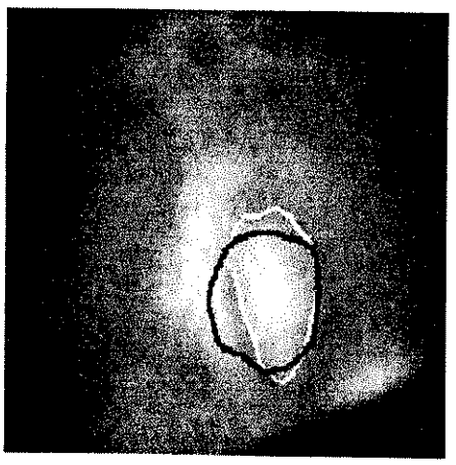
RESULTADOS

IV.1 INTRODUÇÃO

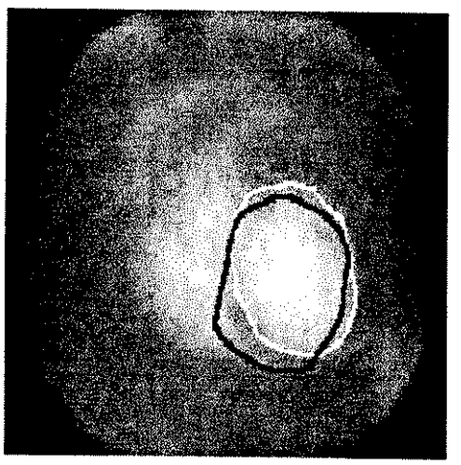
Inicialmente, mostramos contornos de VE obtidos automaticamente com a rede neural A de 4 entradas. Em seguida, mostramos contornos de VE, com tamanhos e formas diferentes, obtidos com a rede neural B de 3 entradas. Um tópico importante que focalizamos é a detecção de contornos especiais, que incluem regiões de proximidade entre o músculo papilar e o VE, regiões onde existem *bordas falsas*, com tendência de se traçar múltiplos contornos, e regiões onde existe uma grande proximidade entre ambos os ventrículos, dificultando o delineamento do contorno do VE. Outros resultados que apresentamos incluem uma tabela de comparação entre os erros obtidos através dos três critérios de seleção dos *pontos do contorno*, citados no capítulo anterior e uma tabela mostrando a influência da variação das coordenadas horizontal e vertical do centro do VE sobre a qualidade do contorno detectado. A fim de avaliarmos o conhecimento geométrico armazenado pela rede determinamos um contorno automático para uma imagem média artificialmente construída com valores constantes de intensidade de pixel e de gradiente linear, para as 64 orientações. Para efeito de comparação, mostramos, também, o contorno médio obtido através de uma média das coordenadas dos pixels de contorno das 20 imagens de treinamento.

IV.2 CONTORNOS DE VE OBTIDOS AUTOMATICAMENTE

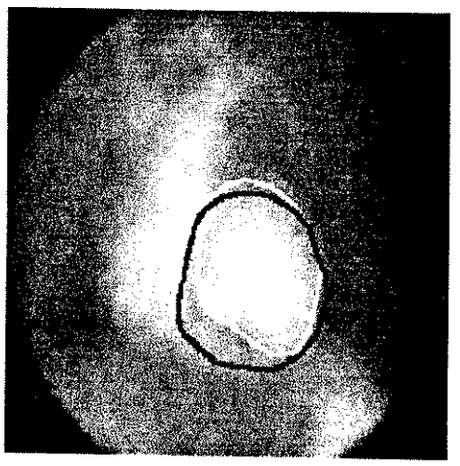
Os contornos obtidos através da rede neural B (3 entradas) delimitam melhor o VE do que os contornos obtidos como a rede neural A (4 entradas). Na figura 4.1 mostramos três exemplos de contornos automáticos (cor escura) obtidos através da rede neural A, superpostos aos contornos traçados manualmente (cor clara) e na tabela 4.1 os valores do erro percentual e do erro soma para esses três contornos. Na figura 4.2 mostramos, para as mesmas imagens de VE mostradas na figura 4.1, os contornos automáticos (cor escura) obtidos através da rede neural B, superpostos aos contornos traçados manualmente (cor clara) e na tabela 4.2 os valores do erro percentual e do erro soma para esses três contornos. A principal característica dos contornos mostrados na figura 4.1, obtidos utilizando-se a rede neural A, é que os mesmos não distinguem corretamente a borda do VE na interface entre os ventrículos, resultando em contornos que incluem regiões do ventrículo direito.



(a)

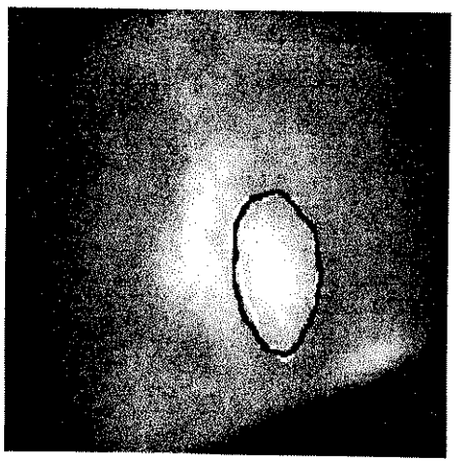


(b)

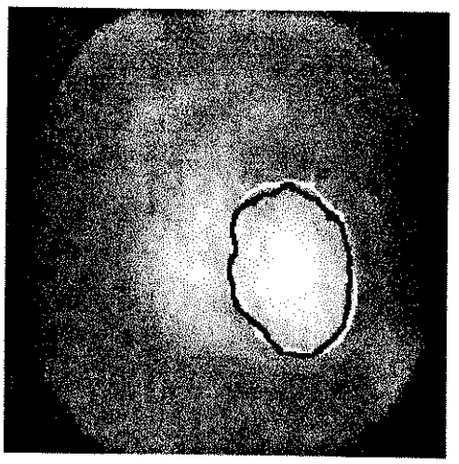


(c)

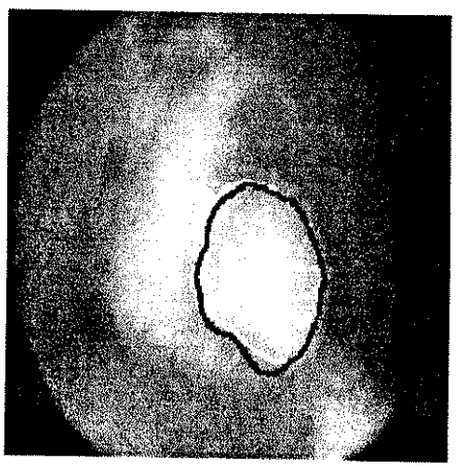
Figura 4.1: Exemplos de detecção do contorno automático pela rede neural A. O traço escuro representa o contorno automático, enquanto que o traço claro representa o contorno manual.



(a)



(b)



(c)

Figura 4. 2: Exemplos de detecção de contorno automático pela rede neural B. O traço escuro representa o contorno automático, enquanto que o traço claro representa o contorno manual.

Tabela 4.1: Valores dos erros percentual e erro soma para as imagens da figura 4.1 (valores adimensionais em termos percentuais).

	Rede neural A	
Imagem	Erro percentual	Erro soma
4.1(a)	12,1	42,11
4.1(b)	6,23	35,35
4.1(c)	27,88	38,50

Tabela 4.2: Valores dos erros percentual e soma para as imagens da figura 4.2.

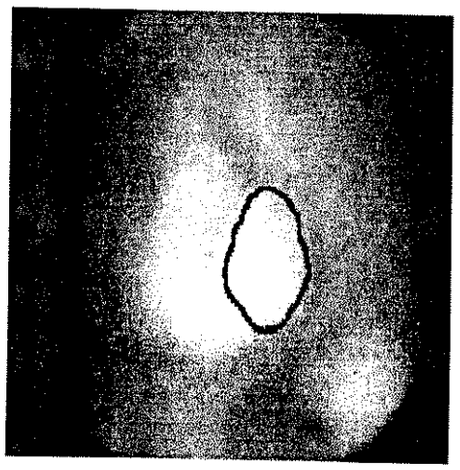
	Rede neural B	
Imagem	Erro percentual	Erro soma
4.2(a)	2,92	10,51
4.2(b)	6,66	11,04
4.2(c)	4,01	10,52

Como os resultados apresentados pela rede neural A não foram de boa qualidade, a mesma foi abandonada. Na figura 4.3 mostramos contornos obtidos para VE de diferentes tamanhos . As áreas dos ventrículos mostrados na figura 4.3 são: 4.3(a) - 25,35cm^2, 4.3(b) - 44,47cm^2 e 4.3(c) - 53,55cm^2. Na figura 4.4 mostramos contornos obtidos para VE de diferentes formas. Na figura 4.4(a) mostramos um VE de forma ovalada. Na figura 4.4(b) mostramos um VE de forma arredondada e na figura 4.4(c) mostramos um VE com forma irregular.

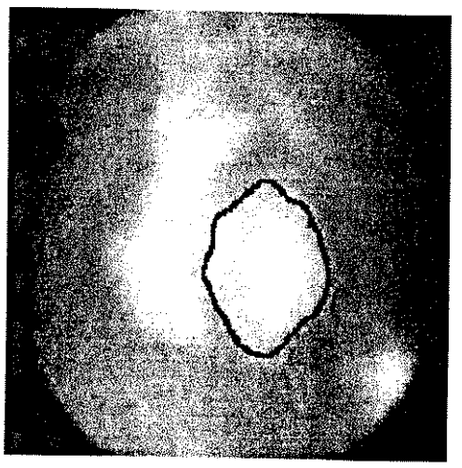
Além das imagens de contornos já mostradas, selecionamos três casos especiais que consideramos críticos para avaliação do desempenho do método proposto: imagens de VE onde o músculo papilar se aproxima muito da região lateral posterior do VE; imagens de VE apresentando bordas não definidas, com a possibilidade de se traçar múltiplos contornos e imagens de VE apresentando regiões onde não existe uma separação do mesmo em relação ao VD.

Na figura 4.5 mostramos dois exemplos de VE onde ocorre uma proximidade do músculo papilar em relação à região lateral posterior do VE. O exemplo onde esta proximidade é mais acentuada é mostrado na figura 4.5(a).

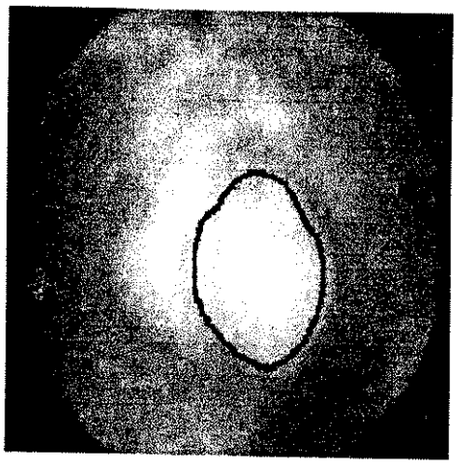
Na figura 4.6(a) ilustramos um exemplo de borda não definida na região superior do VE. Nas figuras 4.6(b) e 4.6(c) ilustramos dois exemplos de regiões onde não existe uma separação clara entre os Ventrículos esquerdo e direito. Na figura 4.6(b) esta região corresponde a parte inferior do septo, enquanto que na figura 4.6(c) esta região corresponde a parte superior do septo.



(a)

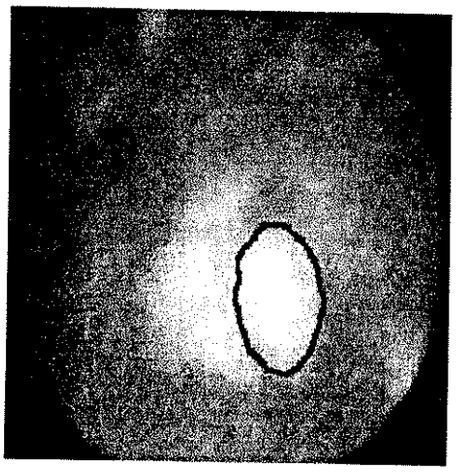


(b)

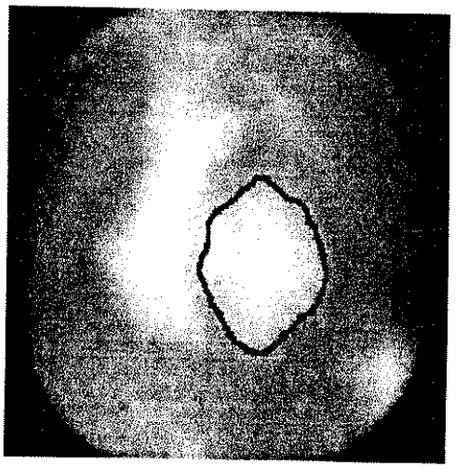


(c)

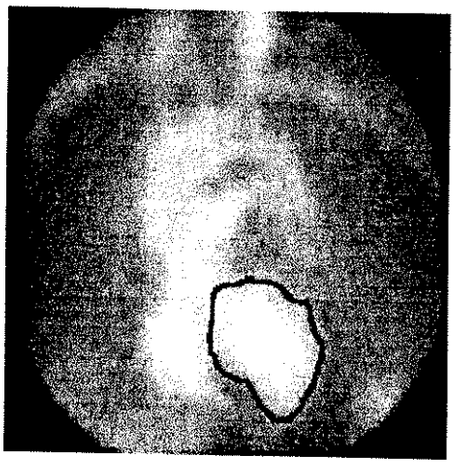
Figura 4.3: Exemplos de VE com diferentes tamanhos: (a) 25,35cm², (b) 44,47cm² e (c) 53,55cm².



(a)

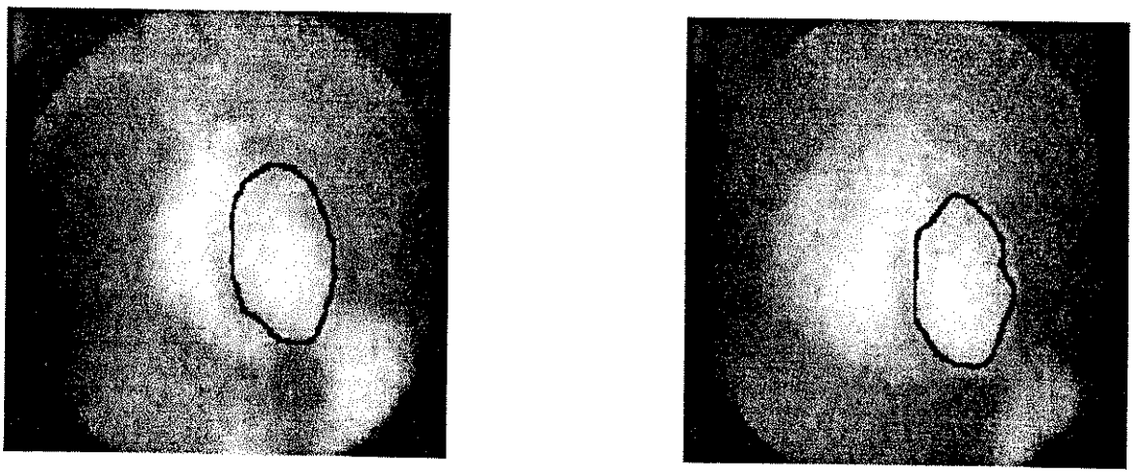


(b)



(c)

Figura 4.4: Exemplos de VE com diferentes formas: (a) forma ovalada, (b) forma arredondada e (c) forma irregular.



(a) *(b)*

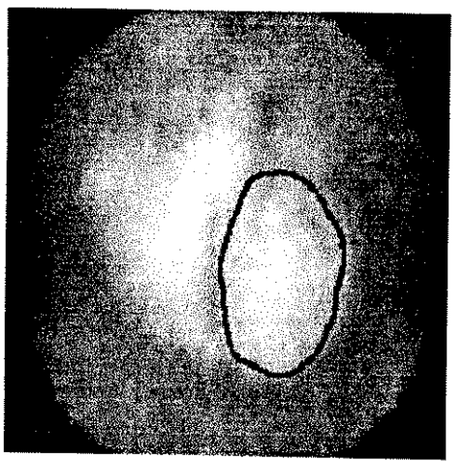
Figura 4.5: Exemplos de proximidade do músculo papilar em relação à região lateral posterior do VE.

IV.3 COMPARAÇÃO ENTRE OS CRITÉRIOS DE SELEÇÃO DOS PONTOS DE CONTORNO

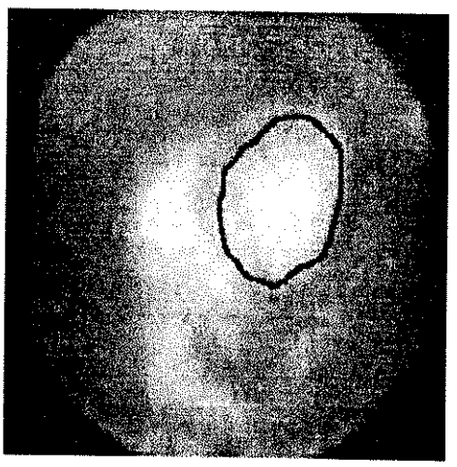
Na tabela 4.3 mostramos, para as imagens dos grupos de treinamento (grupo A) e de teste (grupo B), os valores do erro percentual e do erro soma, definidos através das equações 2.20 e 2.21, respectivamente, obtidos ao se utilizar os critérios de seleção dos *pixels de contorno* definidos no capítulo 3.

IV.4 ERROS OBTIDOS COM A VARIAÇÃO DA POSIÇÃO DO CENTRO

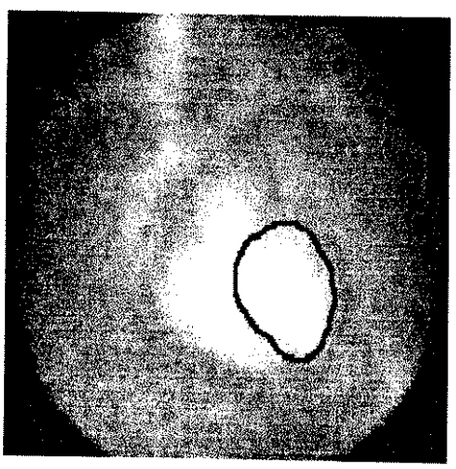
Para avaliarmos a influência da variação das coordenadas do centro sobre a qualidade do contorno obtido, adotamos o seguinte procedimento: escolhemos, entre as imagens do conjunto de teste, 4 delas com valores de erro percentual baixo, com os pontos do contorno escolhidos através do critério #3. Para essas imagens, deslocamos a posição do centro de ±4 pixels nas direções horizontal e vertical. O comportamento do erro percentual médio para essas variações da posição do centro é mostrado na tabela 4.4.



(a)



(b)



(c)

Figura 4.6: Exemplos de contornos de VE com características especiais: (a) Na região superior do VE a região da borda não é muito bem definida, havendo a possibilidade de se traçar múltiplos contornos; (b) Na região inferior do septo a separação entre os ventrículos não é muito bem definida, havendo a possibilidade de se traçar múltiplos contornos e (c) Na região superior do septo a separação entre os ventrículos não é muito bem definida, havendo a possibilidade de se traçar múltiplas bordas.

Tabela 4.3: Erro Percentual e Erro Soma médios para as imagens dos grupos A e B.

	Grupo A		Grupo B	
	Erro Percentual	Erro Soma	Erro Percentual	Erro Soma
Critério #1	5.9	12.41	4.56	12.99
Critério #2	7.47	13.98	8.89	13.81
Critério #3	4.81	11.57	4.8	11.52

Tabela 4.4: Comportamento do Erro Percentual com a variação da posição do centro.

	Erro Percentual Médio	
Deslocamento em pixels	Coordenada horizontal	Coordenada vertical
0	1.15	1.15
±1	2.58	1.53
±2	5.55	1.73
±3	9.22	2.87
±4	11.23	3.91

IV.5 CONHECIMENTO GEOMÉTRICO ARMAZENADO PELA REDE

A fim de se avaliar o conhecimento geométrico armazenado pela rede neural, fizemos uma simulação utilizando uma imagem média construída de tal forma que para cada uma das 64 orientações obtivemos, para os pixels de contorno, os valores médios da intensidade de pixel e do gradiente linear. Assumimos, então, para efeito de simulação, que todos os pixels de uma orientação teriam valores constantes de intensidade de pixel e de gradiente linear, iguais aos valores médios acima definidos. Em seguida, para essa imagem média, determinamos o contorno automático através da rede B, utilizando o critério #3 como critério de escolha *do pixel de contorno.* O raio da orientação de 0° foi fixado em um valor arbitrário R_1. A idéia que exploramos nessa simulação foi tentar isolar a influência das variáveis não geométricas na determinação automática do contorno. Para tanto, simulamos os valores das variáveis não geométricas constantes para todos os pixels de uma orientação. Assim, a única diferença de um pixel para outro em uma orientação é a distância do mesmo em relação ao centro do ventrículo esquerdo. Na figura 4.7(a) mostramos o contorno correspondente a essa imagem média. Na figura 4.7(b) mostramos o contorno médio obtido para as 20 imagens do conjunto de treinamento, através de uma média das coordenadas dos pixels de contorno normalizadas.

(a) (b)

Figura 4.7: Contornos médios. (a) contorno médio obtido pela rede neural utilizando uma imagem média obtida conforme descrito no texto; (b) contorno médio obtido pela média das coordenadas normalizadas dos pixels de contorno das imagens do grupo de treinamento.

CAPÍTULO V

DISCUSSÃO

V.1 INTRODUÇÃO

Este capítulo é marcado por três seções distintas. Inicialmente efetuamos uma análise dos resultados apresentados no capítulo anterior. Em seguida discutimos as principais características da metodologia proposta para a segmentação de órgãos e estruturas, e, em particular, para a segmentação do VE. Finalmente, discutimos as peculiaridades do treinamento da rede neural tipo *perceptron* multicamadas, onde analisamos a razão dos diferentes tempos de treinamento entre as duas redes propostas, bem como entre os diferentes tempos de treinamento das 4 regiões em que dividimos o VE.

V.2 ANÁLISE DOS RESULTADOS

O fato de se conseguir treinar com facilidade uma rede neural para detecção de contornos, nos induz a pensar que os valores das variáveis utilizadas para o treinamento para os pixels da região de contorno são bem diferentes dos valores dessas mesmas variáveis para as regiões adjacentes. Como conseqüência esperaríamos que na utilização dessa rede a mesma detectasse com êxito os contornos sobre os quais ela foi treinada. Uma conclusão aparentemente óbvia. Isso, no entanto, não foi observado nesse trabalho. Consideremos, por exemplo, os resultados da rede A. O tempo de treinamento da rede A, segundo qualquer orientação, foi da ordem de 30s. Porém, conforme mostrado através da figura 4.1, os contornos obtidos com a rede A não delimitam com precisão a área correspondente ao VE. Por outro lado, os tempos de treinamento da rede B, segundo qualquer orientação foram bem maiores que os tempos de treinamento da rede A e, conforme mostrado na figura 4.2, os contornos obtidos com a rede B delimitam com precisão a área correspondente ao VE. Procuraremos a seguir elucidar as razões que acreditamos justificar esses fatos.

Observamos que a deterioração dos contornos obtidos com a rede A ocorre normalmente no meio da região 2, corresponde a posição apical inferior, ou próximo à interface entre as regiões 2 e 3 do VE (essas regiões foram definidas na figura 3.21). Nessas localizações, verificamos a existência de muito ruído (ver, por exemplo, posição apical inferior na figura 4.1(c)). Especialmente, na interface entre as regiões 2 e 3, o contorno do VE aproxima-se do contorno do VD, criando uma

região de indefinição da borda, com a possibilidade de múltiplos contornos poderem ser traçados. Isso pode ser visto através da figura 4.6 (b).

A rede A tem como uma de suas variáveis geométricas de entrada, a distância de um pixel em uma orientação em relação ao pixel de contorno na orientação anterior. Essa variável, que não está presente na entrada da rede B, é fundamental na diferenciação entre os resultados do treinamento e do uso das redes A e B. A influência da mesma sobre o treinamento da rede A é discutida nas seções seguintes. No momento, estamos interessados apenas na discussão da sua influência sobre os resultados apresentados pelo uso da rede A. Devido a existência dessa variável na entrada da rede A, um pixel de contorno determinado em uma orientação influencia a determinação do pixel de contorno na orientação seguinte. O mesmo não ocorre na rede B, onde a determinação dos pixels de contorno em uma orientação não influencia a determinação do pixel de contorno da orientação seguinte.

Imagine-se, então, que ao se utilizar a rede A para a determinação do contorno de um VE, a mesma detecte, em uma orientação situada próxima à interface entre as regiões 2 e 3, um pixel falso de contorno, ou seja, um pixel que não esteja de fato sobre o contorno do VE, mas que esteja na região do septo, ou mesmo sobre o contorno do VD. A variável geométrica distância em relação ao pixel de contorno da orientação anterior forçará o pixel de contorno da orientação seguinte a estar próximo, pertencendo também à borda do VD ou a região do septo, pois esse pixels estarão mais próximos do pixel de borda determinado para a orientação anterior, do que um pixel da borda do VE. Dito de outra forma, a detecção de um pixel de borda falso em uma orientação, desencadeia um processo de detecção de novos pixels falsos em orientações posteriores, ou seja, temos um efeito da propagação do erro inicial. Dois fatores atenuam esse efeito em cascata. O primeiro deles é o efeito atenuador das outras variáveis de entrada da rede neural. A variável geométrica distância em relação ao centro do VE, por exemplo, forçará o contorno detectado a se manter próximo do contorno real do VE. O outro fator é o estabelecimento de uma faixa de procura *dos pixels candidatos* ao contorno. Essa faixa especifica limites absolutos para a procura dos *pixels candidatos* ao contorno. Esses limites, conforme descrito no capítulo de Material e Métodos, são definidos em função dos contornos manuais traçados para as 60 imagens de VE utilizadas nesse trabalho. Esses dois fatores, no entanto, não são suficientes para corrigir a detecção de pixels falsos que ocorre ao se utilizar a rede neural A. Como conseqüência, os contornos obtidos com a rede neural A não delimitam com precisão a área do VE.

A detecção de contornos através da utilização da rede B não apresenta a propagação do erro na detecção de pixels descrita anteriormente, e a qualidade dos contornos obtidos motivaram-nos a escrever esse trabalho. O principal espelho da qualidade desse contornos são as imagens mostradas no capítulo anterior. Para uma comparação com outros métodos existentes, é necessário, além de uma análise visual, que existam medidas quantitativas de desempenho. Nesse trabalho, conforme já salientado, escolhemos duas medidas de erro para avaliar a qualidade dos contornos: o erro percentual e o erro soma, definidos através das expressões 2.20 e 2.21. Na tabela 4.3, mostramos, para os três critérios de escolha dos *pixels de contorno,* os valores desse erros para as imagens do grupo de treinamento e para o grupo de imagens de teste. Em relação aos resultados apresentados nessa tabela, tecemos quatro observações:

- Os resultados são comparáveis aos valores obtidos por Lilly et alii (1989) para o erro percentual (entre 4% e 9%) e para o erro soma (entre 10% e 15%), ao trabalhar com imagens radiográficas do miocárdio.

- Os resultados apresentados para o Grupo A, constituído por imagens de treinamento, mostram que a rede foi capaz de aprender através de exemplos.

- Os resultados apresentados para o Grupo B, constituído por imagens de teste, são comparáveis àqueles obtidos para o Grupo A, demonstrando que a rede foi capaz de generalizar o conteúdo aprendido através de exemplos.

- Entre os três critérios propostos, aquele que apresentou melhores resultados foi o critério #3.

É mister observarmos que o erro soma é uma melhor estimativa da qualidade do contorno do que o erro percentual. Se não vejamos. Através da imagem 4.1(b) vemos que o contorno determinado pela rede A não delimita precisamente a área do VE. Ao nos reportarmos à tabela 4.1, vemos que o erro percentual para esse contorno é de 6,23, não sendo o mesmo um valor ruim de erro percentual. Isso está em contradição com a qualidade do contorno verificada através da inspeção visual da imagem 4.1(b). Na tabela 4.1 observamos que o valor do erro soma para esse mesmo contorno é de 35,35, o que representa um valor mais condizente com a qualidade do contorno obtido. Assim sendo, somente o valor do erro percentual não nos pode garantir a boa qualidade de um contorno obtido automaticamente. Essa qualidade é assegurada, no entanto, através do erro soma. Em geral, valores do erro soma em torno de 10% são comparáveis aos obtidos na literatura (Lilly et alii, 1989). A razão para valores do erro soma dessa magnitude, advém do fato de ser muito difícil a coincidência perfeita entre um contorno traçado manualmente e um contorno obtido automaticamente. Nos exemplos mostrados através da figura 4.2, as diferenças entre os contornos manual e automático levam a um erro soma da ordem de 10%.

As imagens apresentadas nas figuras 4.3 e 4.4 mostram a robustez do método na detecção de contornos de VE de diferentes tamanhos e formas. A independência da rede neural em relação ao tamanho dos ventrículos foi conseguida graças a normalização das distâncias dos pontos do contorno em relação ao centro do VE obtida através da divisão das mesmas por uma referência, a distância do ponto do contorno da orientação de 0° em relação a esse centro. As imagens apresentadas na figura 4.5 mostram a robustez do método na detecção de contorno de VE quando existe uma forte aproximação do músculo papilar do lado lateral posterior do VE, enquanto que as imagens apresentadas na figura 4.6 mostram a capacidade do método de estimar o contorno em regiões onde a borda do VE não está bem definida. Acreditamos que esta capacidade seja devida ao conhecimento geométrico armazenado pela rede neural sobre a forma do VE.

A fim de avaliar a robustez do método que propomos, em função da posição do centro, mostramos os dados da tabela 4.4. Pelos resultados apresentados na tabela 2.1 o valor máximo dos erros Ex e Ey na determinação do centro, após o refinamento, é da ordem de 7%. Isso significa que, para um ventrículo esquerdo com um raio na orientação $\theta=0°$ de 20 pixels, o erro médio máximo que cometemos na determinação do centro é de $\pm 1,4$ pixels. Analisaremos o valor desse erro em face dos resultados apresentados na tabela 4.4. Um deslocamento de $\pm 1,4$ pixels na coordenada horizontal do centro do VE faria com que o erro percentual variasse de 1.15% para um valor na

faixa entre 2,58% e 5,55%. Um deslocamento de ±1,4 pixels na coordenada vertical do centro do VE faria com que o erro percentual variasse de 1,15% para um valor na faixa entre 1,53% e 1,73%. A conclusão a que chegamos é que o método é mais sensível a erros cometidos na determinação da coordenada horizontal do centro do que a erros cometidos na determinação da coordenada vertical do centro.

Em relação aos resultados apresentados na seção 4.5, com respeito ao conhecimento geométrico armazenado pela rede, podemos comentar que a diferença da forma entre os dois contornos apresentados na figura 4.7 não é um fato significativo. Observamos que o contorno mostrado na figura 4.7(a) é apenas a saída da rede que se obtém ao se utilizar valores médios constantes das variáveis intensidade de pixel e gradiente linear nas suas entradas e não um contorno "ideal", como poderia ser imaginado. Para cada orientação e para cada par de valores das variáveis não geométricas, a rede tem armazenado um valor de distância do pixel de contorno em relação ao centro do VE.

V.3 DISCUSSÃO DAS PRINCIPAIS CARATERÍSTICAS DA METODOLOGIA PROPOSTA PARA A DETECÇÃO DO CONTORNO DO VE.

As caraterísticas mais marcantes da metodologia que propomos para a segmentação do contorno do VE são listadas a seguir:

- Utilização de uma rede neural tipo *perceptron* multicamadas como ferramenta de seleção de *pontos candidatos* ao contorno, seguida da escolha dos *pontos do contorno* através de critérios heurísticos.

- dependência fundamental do desempenho da rede em relação à variáveis geométricas.

- normalização das variáveis geométricas, de tal maneira que a determinação do contorno torna-se independente do tamanho do VE.

- modelagem da vizinhança de um órgão como constituída de sub-regiões com características distintas, sendo tratadas separadamente na hora da determinação do contorno, sendo utilizada uma rede neural distinta para cada uma das 64 orientações em que esta vizinhança é dividida.

- utilização de uma procura radial dos pontos do contorno, possível graças a uma determinação prévia do centro do VE.

Ao utilizarmos redes neurais como parte da metodologia, tivemos que sugerir também um método eficaz de determinarmos as variáveis de entrada da mesma. Nesse trabalho, esboçamos um método para a determinação de variáveis relevantes para um melhor desempenho da rede neural. Esse método baseia-se na visualização do comportamento de uma variável na região do contorno do órgão e na sua vizinhança. Essa visualização permite distingüir quais variáveis apresentam diferenças

entrada da rede neural. A visualização é feita traçando-se gráficos dos valores de uma variável na região do contorno e em regiões próximas, internamente e externamente, em cada uma das orientações onde é feita a procura radial. Para cada orientação são traçados nove valores : o valor médio e o valor médio ± 1 desvio padrão para a região do contorno e para as regiões interna e externa.

Nesse trabalho, traçamos esses gráficos para as 5 variáveis investigadas: distância do pixel em relação ao centro do VE; intensidade do pixel; magnitude do gradiente linear; magnitude do gradiente e distância do pixel em relação ao pixel de contorno da orientação anterior. Os gráficos dessas variáveis são mostrados nas figuras 3.16-3.20. Através de uma análise desses gráficos, observamos que as variáveis geométricas são as que apresentam diferenças mais significativas de valores entre a região de contorno e as outras duas regiões, ou seja, as curvas do valor médio e do valor médio ± desvio padrão, obtidas ao unir-se os valores das 64 orientações, apresentam uma superposição mínima em relação as curvas das outras duas regiões. Essa observação levou-nos a concluir que as variáveis geométricas devem assumir um papel fundamental no treinamento da rede tipo *perceptron* multicamadas e na determinação dos *pixels candidatos* ao contorno.

A normalização das variáveis geométricas permitiu uma invariância do comportamento da rede neural em relação ao tamanho do VE. Deve-se à essa normalização o mérito da rede traçar com sucesso o contorno de VE pequenos e grandes, conforme mostrado no capítulo de resultados.

O único trabalho que encontramos na literatura sobre o uso de redes neurais para segmentação do VE foi o de Coppini et alii (1995). Como salientamos no capítulo 2, os autores desse trabalho não efetuaram nenhuma padronização do tamanho dos ventrículos.

Outra idéia central da metodologia que propomos para a segmentação do VE é a utilização de uma rede neural específica para cada orientação. Subjacente à esta idéia está o fato de enxergarmos a vizinhança de um órgão como constituída de várias sub-regiões de características distintas e, como tal, deverem ser tratadas diferentemente na hora da determinação do contorno. Além desse, outros fatores motivaram-nos a utilizar 64 redes neurais ao invés de apenas uma rede neural. É mais fácil monitorar o treinamento de uma rede neural dedicada a apenas uma orientação, pois o mesmo é mais rápido, permitindo, caso necessário, recomeçar várias vezes o mesmo processo em um curto espaço de tempo. É possível otimizar o treinamento utilizando taxas de aprendizado maiores da rede neural em regiões onde o contorno é mais facilmente detectado, e taxas de aprendizado menores em regiões onde o contorno não aparece bem definido.

Acreditamos que ao propormos a utilização de uma rede neural para cada orientação dentro de uma procura radial do contorno, estejamos também realizando uma contribuição para a detecção de contornos em imagens médicas.

O sucesso da procura radial dos pontos do contorno está condicionada a uma determinação precisa do centro do VE, conforme demonstrado na seção 4.6.

V.4 DISCUSSÃO DAS CARACTERÍSTICAS DO TREINAMENTO

Nessa discussão, o parâmetro que está em relevo é o tempo de treinamento. Em primeiro lugar é importante que analisemos as diferenças entre os tempos de treinamento das redes A e B. A rede A, com quatro variáveis de entrada, apresenta tempos de treinamento menores do que 30s para todas as orientações, enquanto que os tempos de treinamento da rede B variaram em função da orientação, sendo sempre maiores do que 1,5s, conforme mostra a tabela 3.5. Duas questões devem ser levantadas. A razão dos tempos de treinamento da rede A serem bem menores que os tempos de treinamento da rede B, nas diferentes orientações, e a razão dos tempos de treinamento da rede B variarem de região para região do ventrículo esquerdo. O entendimento dessas duas questões é possível graças aos gráficos das variáveis apresentados nas figuras 3.16-3.20. Ao nos referirmos a esse gráficos, estaremos apenas nos referindo àqueles onde as distâncias dos pixels interno e externo são de 0,7 e 1,3 vezes a distância do pixel de contorno em relação ao centro do VE, já que as distâncias de 0,8 e 1,2 não levaram à convergência da rede neural em muitas orientações.

Inicialmente, procuraremos entender a primeira dessas questões. Das 4 variáveis da rede A, duas delas, a distância do pixel em relação ao centro do VE e a distância do pixel em relação ao pixel de contorno da orientação anterior, mostradas nos gráficos das figuras 3.16 e 3.20, apresentam uma distribuição onde os valores referentes aos pixels de contorno diferenciam-se significativamente dos valores referentes aos pixels internos e externos ao contorno. Esse comportamento será denominado por nós de comportamento "ideal" para o desempenho da rede. Por outro lado, das 3 variáveis de entrada da rede B, apenas uma delas, a distância do pixel em relação ao centro do VE, apresenta um comportamento similar. Para a rede A, então, 50% das variáveis de entrada apresentam um comportamento "ideal", enquanto que para a rede B, este percentual é de apenas 25%. Acreditamos que essa diferença percentual é fundamental para o fato da convergência dos treinamentos da rede A ser mais rápida que a convergência dos treinamentos da rede B.

Tratemos agora da segunda questão. Os valores encontrados na tabela 3.5 podem ser sintetizados da seguinte forma: os maiores tempos de treinamento foram observados nas regiões 2 e 4, correspondentes as regiões apical inferior e superior do VE, respectivamente. Os menores tempos de treinamento foram observados nas regiões 1 e 3, correspondentes as regiões lateral posterior e do septo, respectivamente. Vários fatores podem ter contribuído para esses diferentes tempos de treinamento. Entre esses fatores enumeramos os seguintes:

- ao analisarmos o gráfico da distribuição da variável distância do pixel em relação ao centro do VE (figura 3.16), observamos uma melhor diferenciação dos valores dos pixels de contorno em relação aos pixels interno e externo, nas regiões 1 e 3 do que nas regiões 2 e 4.

- Ao analisarmos os gráficos das distribuições das variáveis intensidade do pixel e magnitude do gradiente linear na região 4 observamos que não existe diferenciação entre os valores referentes aos pixels de contorno, interno e externo e que um valor referente a um pixel de contorno em uma orientação não tem relação com o valor referente ao pixel de contorno das orientações anterior e posterior. Acreditamos que esses fatos sejam devidos à localização das válvulas aórtica e mitral na região 4. O fluxo de sangue através dessas válvulas torna imprecisa a determinação da borda do VE.

Concluindo, parece-nos que o principal fator que contribui para um tempo de treinamento menor da rede B nas regiões 1 e 3, é o comportamento da variável geométrica nessas regiões.

A estratégia de treinamento que adotamos permitiu-nos regular a taxa de aprendizado da rede nas diversas regiões. Essa taxa de aprendizado é regulada através dos parâmetros **a** e **b** da equação 3.16. Nas regiões 1 e 3, onde é mais fácil a convergência da rede, utilizamos valores grandes de **a** e valores pequenos de **b**, enquanto que nas regiões 2 e 4, devido a dificuldades de convergência, utilizamos valores menores de **a** e valores maiores de **b**.

É útil fazermos uma comparação dos tempos de treinamento da rede que utilizamos com o tempo de treinamento citado por Coppini et alii (1995) para uma rede neural objetivando propósitos semelhantes. Esse autor revela que o tempo de treinamento da rede neural foi de 45 horas. No nosso caso, considerando um tempo de treinamento por orientação de 3 minutos (o pior tempo de treinamento por orientação mostrado na tabela 3.5), temos para as 64 orientações um tempo de treinamento de 192 minutos, ou 3 horas e 12 minutos. Esse valor é consideravelmente menor que o tempo relatado por Coppini et alii (1995). Ressaltamos também que no método que propomos não exige-se que um especialista consuma 20 horas classificando segmentos como pertencentes ou não ao contorno, conforme citado no capítulo de revisão bibliográfica.

CAPÍTULO VI

CONCLUSÕES

VI.1 INTRODUÇÃO

Este capítulo é constituído por duas seções. Na primeira seção apresentamos as conclusões que acreditamos pertinentes a essa tese e na seção de sugestões encaminhamos propostas em duas linhas diferentes: a primeira delas referente a aplicação do método proposto e a segunda, referente ao aprimoramento desse método.

VI.2 CONCLUSÕES

Na introdução, citamos que o principal objetivo desse trabalho era propor uma metodologia automática e não específica para segmentação do VE em imagens médicas. Ao caracterizarmos, no capítulo de revisão bibliográfica, o contorno do VE como pertencendo a uma classe de contornos-θ, o objetivo inicial do trabalho foi ampliado: detectar automaticamente o contorno-θ de órgãos e estruturas em imagens médicas.

O método de segmentação que propomos é caracterizado por duas etapas, a detecção de *pontos candidatos* ao contorno e a detecção de *pontos do contorno.* A detecção do contorno é realizada através de uma procura radial em 64 orientações. A etapa de detecção de *pontos candidatos* ao contorno é efetuada através do uso de uma rede neural tipo *perceptron* multicamadas.

A grande vantagem das redes neurais é a sua capacidade de aprender através de exemplos e, dessa forma, embutir conhecimento. Tanto informações geométricas quanto informações referentes a intensidade do pixel estão presentes nas entradas da rede neural. Após o treinamento, a rede tem armazenado um protótipo do contorno do VE.

Uma importante conclusão é que a facilidade com que uma rede é treinada para detecção de contornos, tem pouca relação com a qualidade dos contornos que são obtidos com o uso da mesma. Um exemplo disso são as redes A e B utilizadas nesse trabalho de tese. Outras conclusões em relação ao uso das redes neurais são apresentadas a seguir:

- mostramos a viabilidade de se utilizar redes neurais como método genérico para detecção de contornos-θ em imagens médicas.

- demonstramos que o conhecimento global da geometria de um órgão é de fundamental importância para o desempenho da rede neural.

Preocupamo-nos em propor um método eficaz para a escolha das variáveis de entrada da rede neural e para a seleção das imagens do conjunto de treinamento. Na etapa de detecção dos pontos de contorno, propomos a utilização de três critérios heurísticos de escolha, tendo um deles apresentado melhores resultados do que os outros dois.

O bom desempenho do método de segmentação que propomos repousa numa determinação precisa do centro do VE, efetuada de forma automática.

A modelagem do contorno do VE como um contorno-θ permitiu que incluíssemos informação geométrica na entrada na rede neural. O fato de cada orientação estar associada a diferentes condições de aprendizado reflete a situação normalmente encontrada em imagens médicas e raramente levada em conta pelos algoritmos de segmentação. A maioria dos sistemas que trabalham com imagens médicas considera o *background* uniforme. As imagens de *blood-pool* do VE são um excelente exemplo de que esta uniformidade pode não existir. A linha que vai do centro do VE em direção a região lateral posterior do mesmo encontra uma parede livre, caracterizada por um bom nível de contraste, enquanto que a linha que vai do centro do VE em direção a região do septo, encontra uma parede marcada por um pobre contraste.

É mister que frisemos um aspecto fundamental em relação as variáveis geométricas, a normalização das mesmas, dividindo-as pela distância do pixel de contorno da orientação $\theta=0°$ em relação ao centro do VE. Essa normalização permitiu uma invariância do comportamento da rede neural em relação ao tamanho do VE. Creditamos a essa normalização o mérito da rede traçar com sucesso o contorno de VE pequenos e grandes.

Através dos exemplos de contornos de VE detectados automaticamente mostrados no capítulo de resultados e das tabelas apresentadas para os erros percentual e soma, concluímos que o método mostrou ser preciso e robusto, com baixa sensibilidade em relação a variações da posição do centro do VE.

Os três pressupostos assumidos pelo método de detecção - a existência de uma região da borda onde a mesma seja bem definida, uma regularidade na localização do órgão na imagem a ser segmentada e o fato do contorno a ser segmentado ter de ser um contorno arredondado ou contorno-θ - não invalidam a generalidade do mesmo. Foram condições introduzidas ao se tratar com as imagens reais, que podem ser removidas em trabalhos futuros.

É nossa crença que atingimos os objetivos a que nos propusemos no início deste trabalho. Poderíamos, no entanto ter efetuado mais simulações com outros números de imagens de treinamento e também ter explorado mais a arquitetura da rede neural tipo *perceptron* multicamadas, variando o número de nós da camada intermediária. O fato de termos desenvolvido todos os programas de treinamento da rede em linguagem C, no entanto, consumiu-nos muito tempo. Se tivéssemos que recomeçar, utilizaríamos um pacote pré-existente para treinamento de redes neurais.

Acreditamos que esse trabalho contribua com uma nova visão de como se detectar o contorno de órgãos e estruturas em imagens médicas. É nossa crença também que o mesmo abre muitas frentes de pesquisa. Aquelas que divisamos são citadas na próxima seção.

VI.3 SUGESTÕES DE TRABALHOS FUTUROS

As principais linhas de pesquisa que vislumbramos para dar continuidade a esse trabalho enquadram-se em duas categorias: aplicação e aprimoramento. Na área de aplicação enxergamos duas vertentes:

- extensão do método a outras modalidades de imagem cardíacas, como imagens radiográficas, ecocardiográficas, etc. Extensão do método para a detecção do contorno de outros órgãos e estruturas.

- aplicação do método para o cálculo dos parâmetros da função cardíaca.

A extensão do método para outros órgãos e estruturas pressupõe que o contornos desses órgãos e estruturas pertençam a classe de contornos-θ.

A aplicação do método para o cálculo dos parâmetros da função cardíaca pressupõe que o mesmo seja aplicado para a detecção do contorno do VE em imagens correspondentes a outros instantes do ciclo cardíaco. Isso exige que um outros conjuntos de redes neurais sejam treinadas para esses diferentes instantes.

Em relação ao aprimoramento do trabalho atual, enxergamos três vertentes:

- exploração de novas variáveis de entrada da rede neural;

- escolha de métodos mais genéricos para a seleção dos *pontos de contorno* a partir do conjunto de *pontos candidatos*.

- exploração de outras técnicas de inteligência artifical para detecção do contorno do VE, como técnicas *fuzzy*.

Em relação a exploração de novas variáveis de entrada da rede, estamos nos referindo, principalmente, à novas variáveis geométricas. Ao analisarmos o método que propomos verificamos que o mesmo é invariante em relação ao tamanho do VE, graças ao processo de normalização da distância do pixel em relação ao centro do VE. Nenhuma alternativa foi proposta, no entanto, para tornar o método invariante em relação a forma e a inclinação do VE. Para ilustramos os conceitos de forma e inclinação reportar-nos-emos aos contornos hipotéticos de VE mostrados na figura 6.1. Nas figuras 6.1 (a) e 6.1 (b) mostramos dois ventrículos com a mesma inclinação, porém com formas diferentes, enquanto que nas figuras 6.1 (a) e 6.1(c) mostramos dois ventrículos com a mesma forma, porém com inclinações diferentes.

Essas novas variáveis geométricas poderiam se referir, por exemplo, à forma e à inclinação do VE. A inclinação do VE poderia ser representada através da inclinação do eixo maior do mesmo. Em relação a forma do VE, propomos uma variável representada pela razão entre os comprimentos dos eixos maior e menor. A determinação aproximada do comprimento desses eixos, sem uma prévia determinação do contorno do VE é uma questão ainda em aberto.

Acreditamos que com estas duas novas variáveis geométricas seria possível reduzir os valores dos erros percentual e soma obtidos nesse trabalho.

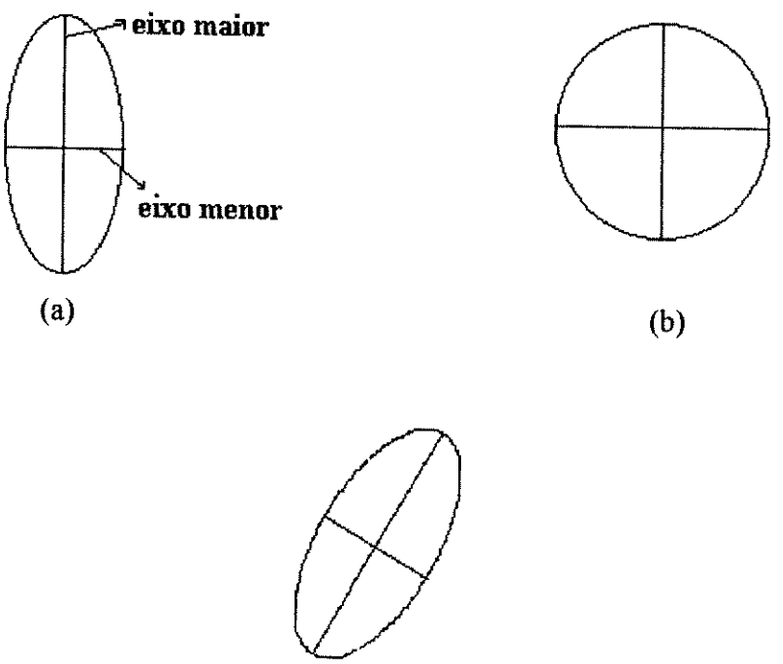


Figura 6.1: formas e inclinações de VE. (a) e (b): VE com mesma inclinação, porém com formas diferentes;(a) e (c): VE com mesma forma, porém com inclinações diferentes.

Em relação a escolha de métodos mais genéricos para determinação dos pontos do contorno, poderíamos, por exemplo, subtituir os critérios heurísticos que propomos por uma procura dinâmica em grafos. Se essa estratégia apresentasse bons resultados, estaríamos caminhando no sentido de uma maior generalização do método de segmentação do VE que propomos.

APÊNDICE A

O objetivo deste apêndice é demonstrar como o método de aprendizado de retro-propagação minimiza o quadrado das diferenças entre os valores de saída obtidos e os valores de saída desejados de uma rede neural tipo perceptron, através do gradiente descendente.

A demonstração a ser apresentada é válida apenas para redes em camadas tipo *feedforward*. Nessas redes, os neurônios de entrada correspondem à camada da base e os neurônios de saída correspondem à camada do topo. Podem existir muitas camadas de neurônios escondidos, mas a saída de cada neurônio deve conectar-se a neurônios das camadas superiores e a entrada do mesmo deve conectar-se às camadas inferiores. Dado um vetor de entrada, o vetor de saída é computado através de uma passagem direta que calcula os níveis de atividade de cada camada utilizando-se os níveis de atividade anteriormente computados das camadas inferiores. A demonstração é dividida em duas partes. Na primeira parte veremos a regra delta que se aplica a redes sem camadas escondidas, enquanto que na segunda parte veremos a regra delta generalizada, que se aplica a redes neurais com camadas escondidas. A demonstração que se segue tem a mesma estrutura daquela apresentada por Rumelhart el alia [1988].

REGRA DELTA

Na figura A.1 mostramos uma rede tipo perceptron com uma camada de neurônios de entrada, camada *i* e uma camada de neurônios de saída, camada *j*. A regra delta para atualização dos pesos após a apresentação de um par de entrada/saída é dada por:

$$\Delta_p w_{ji} = \eta.(d_{pj} - o_{pj}).i_{pi} = \eta.\delta_{pji}.i_{pi} \qquad (A.1)$$

onde:

d_{pj}= saída desejada do j-ésimo neurônio para o padrão de entrada *p*;

o_{pj}= saída obtida do j-ésimo neurônio para o padrão de entrada *p*;

i_{pi} = valor do i-ésimo elemento do padrão de entrada *p*;

$\delta_{pj} = d_{pj} - o_{pj}$;

$\Delta_p w_{ji}$= atualização a ser feita no peso que conecta o i-ésimo neurônio ao j-ésimo neurônio após a apresentação do padrão *p*.

Demonstraremos que a equação A.1 implementa o método do gradiente descendente quando os neurônios são lineares. Consideremos que:

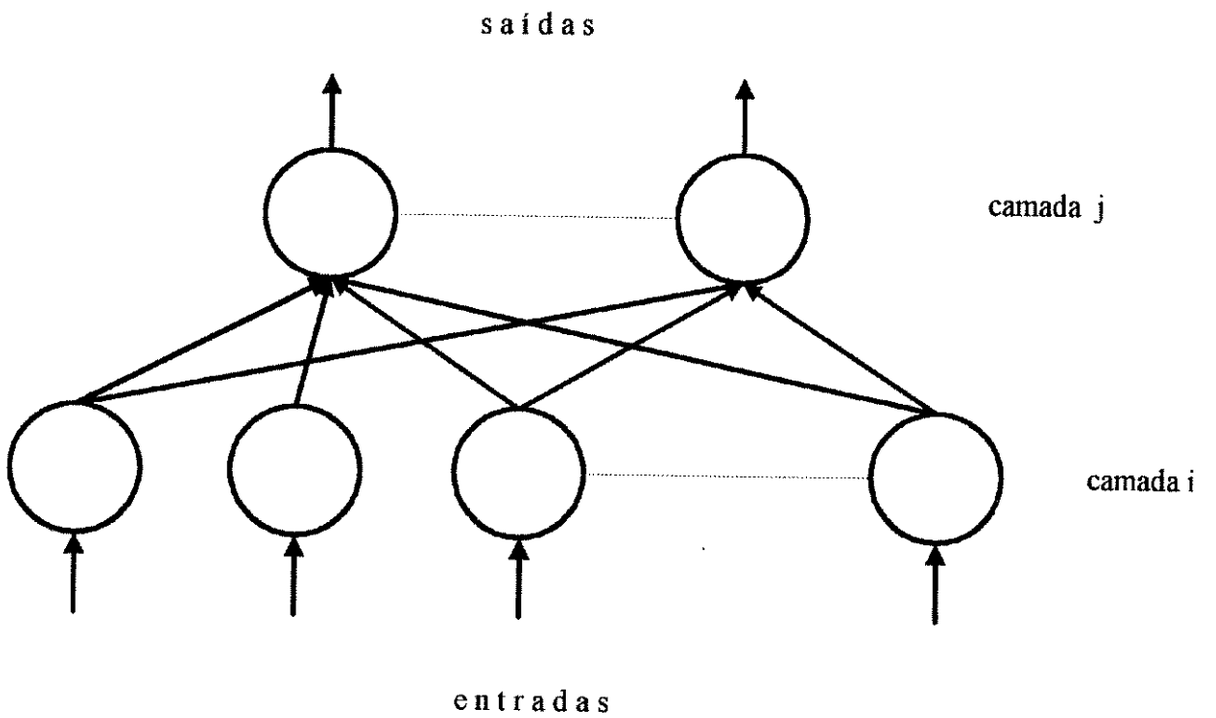


Figura A. 1: Rede neural tipo perceptron com 2 camadas de neurônios

$$E_p = \frac{1}{2} \cdot \sum_j \left(d_{pj} - o_{pj}\right)^2 \tag{A.2}$$

seja a medida de erro para o padrão de entrada *p* e que $E=\sum E_p$ seja a medida de erro global. Para mostrarmos que a regra delta implementa um gradiente descendente em E, quando os neurônios são lineares, devemos mostrar que:

$$\frac{\partial E_p}{\partial w_{ji}} = -\delta_{pj} \cdot i_{pi} \tag{A.3}$$

que é proporcional a $\Delta_p w_{ji}$ da equação A.1. O cálculo das derivadas torna-se muito fácil quando não existem camadas escondidas. Com este propósito usaremos a regra da cadeia para escrevermos a derivada como o produto de duas partes: a derivada do erro em relação a saída do neurônio vezes a derivada da saída em relação ao peso:

$$\frac{\partial E_p}{\partial w_{ji}} = \frac{\partial E_p}{\partial o_{pj}} \cdot \frac{\partial o_{pj}}{\partial w_{ji}} \tag{A.4}$$

A primeira parte mostra como o erro varia com a saída do j-ésimo neurônio e a segunda parte mostra como a saída varia em função de variações em w_{ji}. A partir da equação A.2 temos que:

$$\frac{\partial E_p}{\partial o_{pj}} = -\left(d_{pj} - o_{pj}\right) = -\delta_{pj} \tag{A.5}$$

Como assumimos neurônios lineares :

$$o_{pj} = \sum w_{ji} . i_{pi} \tag{A.6}$$

A partir da equação A.6 concluímos que:

$$\frac{\partial o_{pj}}{\partial w_{ji}} = i_{pi}$$

Substituindo esse resultado em A.4 obtemos:

$$-\frac{\partial E_p}{\partial w_{ji}} = \delta_{pj} . i_{pi} \tag{A.7}$$

como desejado. Combinando esse resultado com a observação de que:

$$\frac{\partial E}{\partial w_{ji}} = \sum_p \frac{\partial E_p}{\partial w_{ji}} \tag{A.8}$$

concluímos que a variação em w_{ji} após um ciclo completo de apresentação de padrões na entrada é proporcional a essa derivada e, assim, a regra delta implementa o gradiente descendente em E. Na verdade, isso é estritamente verdadeiro se os valores dos pesos não mudam durante um ciclo. Mudando os pesos após cada padrão ser apresentado, separamo-nos do método original do gradiente descendente em E. No entanto, uma vez que a taxa de aprendizado η seja suficientemente pequena, essa separação é desprezível e a regra delta implementará de uma maneira bem aproximada o gradiente descendente em um erro de soma de quadrados, E. Nesse trabalho de tese implementamos o método do gradiente descendente original, pois adotamos um procedimento de atualização dos pesos tipo *batch*, onde os mesmos são atualizados somente após a apresentação de todos os padrões de entrada.

Mostramos que a regra delta implementa o método do gradiente descendente em E. Nesse caso, sem camadas escondidas, a superfície de erro assume a forma de uma taça, com apenas um mínimo, de tal forma que o método do gradiente descendente encontra o melhor conjunto de pesos que minimizam E. Com camadas escondidas, no entanto, não é tão óbvio como calcular as derivadas e a superfície de erro não é côncava, de tal forma que existe a possibilidade de ficar aprisionado em um mínimo local. Na demonstração da regra delta generalizada a seguir mostraremos apenas que existe uma maneira para se calcular as derivadas.

REGRA DELTA GENERALIZADA

Na figura A.2 mostramos uma rede tipo perceptron com 3 camadas de neurônios: camada de entrada - camada *i*, camada escondida - camada *j* e camada de saída - camada *k*. Como mostrado por Rumelhart el alia [1988], neurônios da camada escondida com função linear não são muito úteis. Por isso, assumiremos que a ativação dos mesmos é semilinear. Uma função de ativação semilinear é aquela na qual a saída de um neurônio é uma função diferenciável e não decrescente da entrada total:

$$net_{pj} = \sum_i w_{ji} o_{pi} \qquad (A.9)$$

onde, $o_i = i_i$ se o neurônio *i* é da camada de entrada. Então, uma função de ativação semilinear é aquela para a qual *f* é diferenciável e não decrescente. A regra delta generalizada é válida se a rede é constituída de neurônios com funções semilineares de ativação. Observamos que neurônios com uma função de ativação de limiar como aquele mostrado na figura 2.1 não satisfaz os requisitos de uma função semilinear, pois sua derivada é infinita no limiar e zero em outros locais. Para mostrarmos que a regra delta generalizada implementa um gradiente descendente em E, devemos mostrar que:

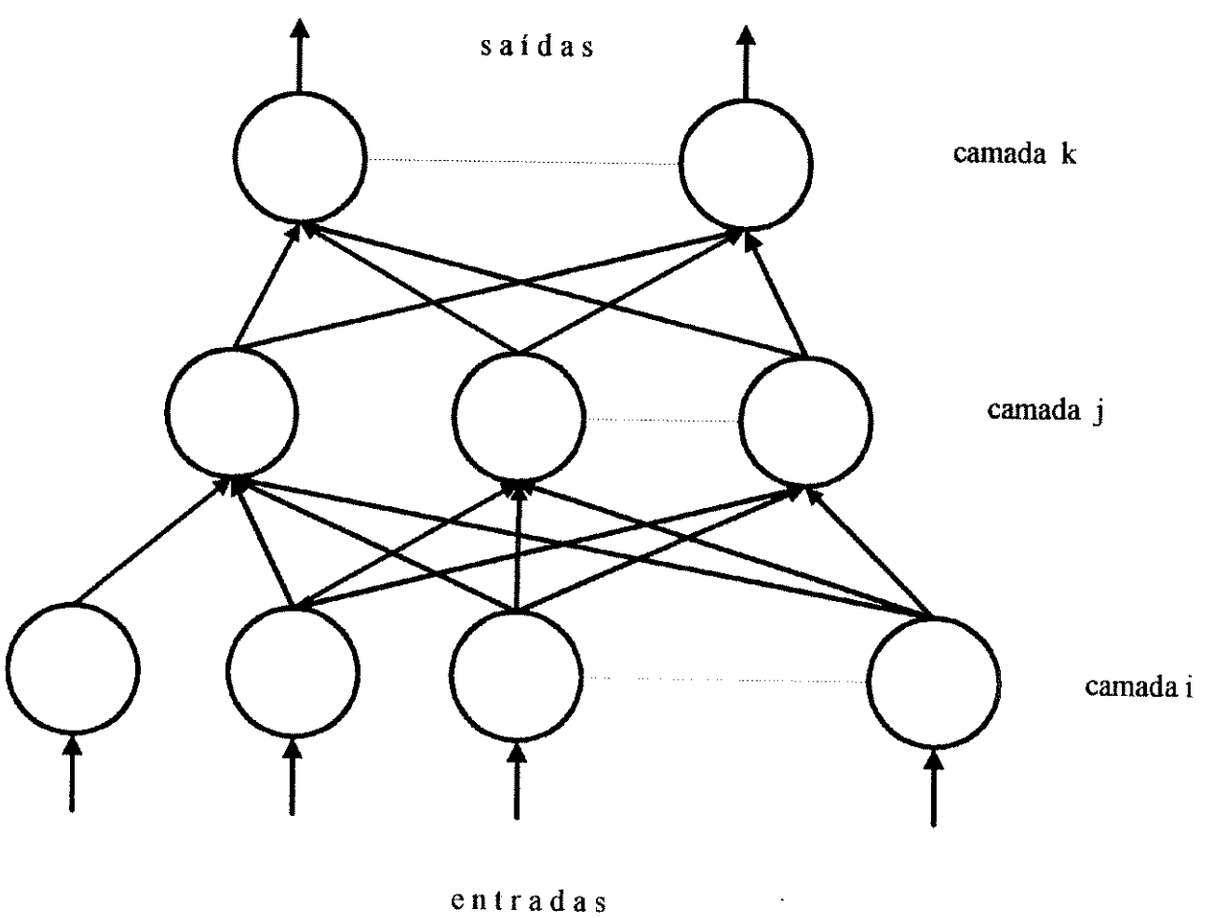


Figura A. 2: Rede neural tipo perceptron com três camadas de neurônios

$$\Delta_p w_{ji} \ \alpha \ -\frac{\partial E_p}{\partial w_{ji}}$$

Da mesma forma que na regra delta, é interessante visualizarmos esta derivada como sendo o produto de duas partes: uma parte que reflete a variação no erro como função da variação na entrada total do neurônio e outra parte que reflete o efeito da variação de um peso particular na entrada total. Assim, podemos escrever:

$$\frac{\partial E_p}{\partial w_{ji}} = \frac{\partial E_p}{\partial net_{pj}} \cdot \frac{\partial net_{pj}}{\partial w_{ji}} \qquad (A.11)$$

A partir da equação A.9 o segundo fator é dado por:

$$\frac{\partial net_j}{\partial w_{ji}} = \frac{\partial}{\partial w_{ji}} \sum_k w_{jk} o_{pk} = o_{pi}$$

Se definirmos

$$\delta_{pj} = -\frac{\partial E_p}{\partial net_{pj}}$$

obtemos a partir de A.11:

$$-\frac{\partial E_p}{\partial w_{ji}} = \delta_{pj} o_{pi}$$

Isso quer dizer que para implementarmos o gradiente descendente em E devemos fazer com que a variação do erro ocorra de acordo com a expressão:

$$\Delta_p w_{ji} = \eta \delta_{pj} o_{pi} \qquad (A.12)$$

Esse resultado é semelhante ao obtido na regra delta. A seguir mostraremos que existe uma forma computacional recursiva simples para o cálculo dos δ's, que pode ser implementada propagando-se o sinal de erro para trás através da rede. A fim de computarmos $\delta_{pj} = -\partial E_p / \partial net_{pj}$ aplicaremos a regra da cadeia a fim de escrevermos esta derivada parcial como o produto de dois fatores: um fator refletindo a variação no erro como função da saída do neurônio e o outro refletindo as mudanças na saída resultantes de mudanças na entrada. Assim temos:

$$\delta_{pj} = -\frac{\partial E_p}{\partial net_{pj}} = -\frac{\partial E_p}{\partial o_{pj}} \cdot \frac{\partial o_{pj}}{\partial net_{pj}} \qquad (A.13)$$

Para computarmos o segundo fator utilizaremos a equação A.10:

$$\frac{\partial o_{pj}}{\partial net_{pj}} = f_j'(net_{pj})$$

o qual é a derivada da função f_j para j-ésimo neurônio avaliada na entrada total net_{pj} para esse neurônio. A fim de computar o primeiro fator consideramos dois casos. O primeiro deles assume que o neurônio *j* é um neurônio da saída da rede. Nesse caso segue-se da definição de E_p que:

$$\frac{\partial E_p}{\partial o_{pj}} = -(d_{pj} - o_{pj})$$

o qual é o mesmo resultado obtido através da regra delta. Substituindo os dois fatores em A.13 obtemos:

$$\delta_{pj} = (d_{pj} - o_{pj}) f_j'(net_{pj}) \qquad (A.14)$$

se o neurônio *j* não é um neurônio de saída usamos a regra da cadeia para obtermos:

$$\frac{\partial E_p}{\partial o_{pj}} = \sum_k \frac{\partial E_p}{\partial net_{pk}} \cdot \frac{\partial net_{pk}}{\partial o_{pj}} = \sum_k \frac{\partial E_p}{\partial net_{pk}} \cdot \frac{\partial}{\partial o_{pj}} \sum_i w_{ki} o_{pi} = \sum_k \frac{\partial E_p}{\partial net_{pk}} \cdot w_{kj} = -\sum_k \delta_{pk} w_{kj}$$

Neste caso, substituindo-se os dois fatores em A.13 obtemos:

$$\delta_{pj} = f_j'(net_{pj}) \sum_k \delta_{pk} w_{kj} \qquad (A.15)$$

As equações A.14 e A.15 estabelecem um procedimento recursivo para computar os δ's para todos os neurônios da rede, que são utilizados para computar as variações nos pesos da rede, de acordo com a equação A.12. Esse procedimento constitui a regra delta generalizada para uma rede *feedforward* com neurônios cujas funções de ativação são semilineares.

REFERÊNCIAS BIBLIOGRÁFICAS

ABBAS, J.J. e CHIZECK, H.J. (1995), "Neural Network Control of Functional Neuromuscular Stimulation Systems: Computer Simulation Studies", *IEEE Transactions on Biomedical Engineering*, Volume 42, pages 1117-1127.

ADAM, D., HAREUVENI, O. e SIDEMAN, S. (1987), "Semi-Automatic Border Tracking of Cine Echocardiogram Ventricular Images", *IEEE Transactions on Medical Imaging*, Volume MI-6, pages 266-271.

ADLER, A. e GUARDO, R. (1994), "A Neural Network Image Reconstruction Technique for Electrical Impedance Tomography", *IEEE Transactions on Medical Imaging*, Volume 13, pages 594-600.

ALDERMAN, C.L., SADLER, H. , BROOKER, J.Z., SANDERS, W.J., SIMPSON, C. SIMPSON e HARRISON, D.C. (1973), "Light-Pen Processing of Video images for the Determination of Left Ventricular Volume", *Circulation*, Volume 47, pages 309-316.

ASADA, N., DOI, K., MACMAHON, H., MONTNER, S.M., GIGER, M.L., ABE, C. e WU, Y. (1990), "Potential Usefulness of an Artificial Neural Network for Differential Diagnosis of Intersticial Lung Diseases: Pilot Study", *Radiology*, Volume 177, pages 857-860.

BADRAN, A.K., DURRANI, T.S., FISHER, A.C. e PAUL, J.P. (1990), "Patient Realignment in MRI of the Head: An Algorithm Using Mathematical Morphology for Feature Extraction", *Journal of Biomedical Engineering*, Volume 12, pages 139-142.

BAE, K.T., GIGER, M.L., DOI, K. e MACMAHON, H. (1991), "Computer-Aided Detection of Pulmonary Nodules in CT Images", *Radiology*, Volume 181(P), page 144.

BAE, K.T., GIGER, M.L., CHEN, C.T. e KAHN, C.E. (1993), "Automatic Segmentation of Liver Structure in CT Images", *Medical Physics*, Volume 20, pages 71-78.

BAKER, J.A., KORNGUTH, P.J., LO, J.Y., WILLIFORD, M.E. e FLOYD, C.E. (1995), "Breast Cancer Prediction with Artificial Neural Network Based on BI-RADS Standardized Lexicon", *Radiology*, Volume 196, pages 817-822.

BAUM, E.B. (1988), "On the Capabilities of Multilayer Perceptrons", *Journal of Complexity*, Volume 4, pages 193-215.

BARRET, W.A., CLAYTON, P.D. e WARNER, H.R. (1980), "Determination of Left Ventricular Contours: a Probabilistic Algorithm Derived from Angiographic Images", *Computers and Biomedical Research*, Volume 13, pages 522-548.

BISTER, M, TARYMANS, Y. e CORNELIS, J. (1990), "Automated Segmentation of Cardiac MR Images", *Computers in Cardiology*, pages 215-218.

BOONE, J.M. (1990), "X-ray Spectral Reconstruction from Attenuation Data Using Neural Networks", *Medical Physics*, Volume 17, pages 647-654.

[1] BOONE, J.M., SIGILLITO, V.G. e SHABER G.S. (1990), "Neural Networks in Radiology: an Introduction and Evaluation in a Signal Detection Task", *Medical Physics*, Volume 17, pages 234-241.

[2] BOONE, J.M., GROSS, G.W. e HUNT, V.G. (1990), "Neural Networks in Radiologic Diagnosis: I. Introduction and Illustration", *Investigative Radiology*, Volume 25, pages 1012-1016.

BOUDRAA, A.E.Q., MALLET, J.J., BESSON, J.E., BOYOUCEF, S.E. e CHAMPIER, J. (1993), "Left Ventricle Automated Detection Method in Gated Isotopic Ventriculography Using Fuzzy Clustering", *IEEE Transactions on Medical Imaging*, Volume 12, pages 451-465.

CARPENTER, G.A. e GROSSBERG, S. (1988), "The ART of Adaptive Pattern Recognition by a Self-Organizing Neural Network", *IEEE Computer*, Volume 21, pages 77-88.

CHAN, K.H., JOHNSON, K.A., BECKER, J.A., SATLIN, A., MENDELSON, J., GARADA, B. e HALMAN, B.L. (1994), "A Neural Network Classifier for Cerebral Perfusion Imaging", *Journal of Nuclear Medicine*, Volume 35, pages 771-774.

CHU, C.H., DELP E.J. e BUDA, A.J. (1988), "Detecting Left Ventricular Endocardial and Epicardial Boundaries by Digital Two-Dimensional Echocardiography", *IEEE Transactions on Medical Imaging*, Volume 7, pages 81-90.

CICHOCKI, A e UNBEHAUEN, R. (1993), *Neural Networks for Optimization and Signal Processing*, First Edition, John Wiley & Sons, New York.

CLARKE, L.P., SILBIGER, M.S., NAYLOR W.C e BROWN, K. (1990), "Artificial Neural Net Systems for Interactive Tissue Classification with MR Imaging and Image Segmentation" *Radiology*, voll. 177(P), page 148.

CLAYTON, P.D., HARRIS, L.D., RUMEL, S.R. e WARNER, H.R. (1974), "Left Ventricular Videometry", *Computers and Biomedical Research*, Volume 7, pages 369-379.

COPPINI, G., POLI, R. e VALLI, G. (1995), "Recovery of the 3-D Shape of the Left Ventricle from Echocardiographic Images", *IEEE Transactions on Medical Imaging*, Volume 14, pages 301-317.

COSTA, C. e MOURA, L. (1995), "Automatic Detection of LV Contours in Nuclear Medicine Using Geometrical Information and a Neural Net", *Computers in Cardiology*, pages 557-560.

COSTA, M. G. F. e MOURA, L. (1995), "Automatic Assessment of Scintmammographic Images Using a Novelty Filter", *Journal of American Medical Informatics,* Symposium Supplement, pages 537-541.

DAWSON, M.R.W., DOBBS, A., HOOPER, H.R., MCEWAN, A.J.B., TRISCOTT, J. e COONEY, J. (1994), "Artificial Neural Networks that Use Single-Photon Emission Tomography to Identify Patients with Probable Alzheimer's Disease", *European Journal of Nuclear Medicine,* Volume 21, pages 1303-1311.

DIAS, J.M.B e LEITÃO J.M.N. (1996), "Wall Position and Thickness Estimation from Sequences of Echocardiographic Images", *IEEE Transactions om Medical Imaging,* Volume 15, pages 25-38.

DUNCAN, J.S. (1987), "Knowledged Directed Left Ventricular Boundary Detection in Equilibrium Radionuclide Angiocardiography", *IEEE Transactions on Medical Imaging,* Volume 6, pages 325-336.

ERCAL, F., CHAWLA, A, STOECHER, W.V., LEE, H.C. e MOSS, R.H. (1994), "Neural Network Diagnosis of Malignant Melanoma from Color Images", *IEEE Transactions on Biomedical Engineering*, Volume 41, pages 837-845.

EZEKIEL, A., GARCIA, E.V., AREEDA, J.S. e CORDAY, S.R.. (1985), "Automatic and Intelligent Left Ventricular Contour Detection from Two Dimensional Echocardiograms", *Computers in Cardiology,* pages 261-264.

FAHLMAN, S.E. (1988), "Faster-Learning Variations on Back-Propagation: an Empirical Study", *Proc. 1988 Connectionist Models Summer School,* Morgan Kaufmann Los Altos CA, pages 38-51.

FENG, J., LIN, W.C. e CHEN, C.T. (1991), "Epicardial Boundary Detection Using Fuzzy Reasoning", *IEEE Transaction on Medical Imaging,* Volume 10, pages 187-199.

FLEAGLE, S.R., THEDENS, D.R., EHRHARDT, SCHOLZ ,T.D e SKORTON, D.J. (1991), "Automated Identification of Left Ventricular Borders from Spin-Echo Magnetic Resonance Images", *Investigative Radiology*, Volume 26, pages 295-303.

FISH, P. (1990), *Physics and Instrumentation of Diagnostic Medical Ultrasound*, First Edition, John Wiley & Sons, New York.

FRIEDLAND, N. e ADAM, D. (1989), "Automatic Ventricular Cavity Boundary Detection from Sequential Ultrasound Images Using Simulated Annealing", *IEEE Transactions on Medical Imaging,* Volume8, pages 344-353.

FUJITA, H., KATAFUCHI, T., UEHARA, T. e NISHIMURA, T. (1992), "Application of Artificial Neural Network to Computer-Aided Diagnosis of Coronary Artery Disease in Myocardial SPECT Bull's-Eye Images", *Journal of Nuclear Medicine,* Volume 33, pages 272-276.

GARG, S. e FLOYD, C.E. (1991), "Artificial Neural Network Analysis for Nodule Detection in Digital Chest Radiographs", *Radiology*, Volume 181(P), page 144.

GELFAND, M.J. e THOMAS, S.R. (1988), "Effective Use of Computers in Nuclear Medicine", First Edition, McGraw-Hill, New York.

GONZALEZ, R. C. e WINTZ, P. (1987), *Digital Image Processing*, Second Edition, Addison Wesley Publishing Company, California, U.S.A.

GOSHTASBY, A. e TURNER, D.A. (1995), "Segmentation of Cardiac Cine MR Images for Extraction of Right and Left Ventricular Chambers", *IEEE Transactions on Medical Imaging*, Volume14, pages 56-64.

GRATTONI, P. e BONOMINI, R. (1985), "Contour Detection of the Left Ventricular Cavity from Angiographic Images", *IEEE Transactions on Medical Imaging*, Volume 4, pages 72-78.

GRAUPE, D. e KORDYLEWSKI, H. (1995), "Artificial Neural Network Control of FES in Paraplegics for Patient Responsive Ambulation", *IEEE Transactions on Biomedical Engineering*, Volume 42, pages 699-707.

GRAY, D.L., ASH, S.R., JACOBI, J. e MICHEL, A.N. (1991), "The Training and Use of an Artificial Neural Network to Monitor Use of Medication in Treatment of Complex Patients", *Journal of Clinical Engineering*, Volume 16, pages 331-336.

GRAY, D.L. e MICHEL, A.N. (1992), "A Training Algorithm for Binary Feedforward Neural Networks", *IEEE Transactions on Neural Networks*, Volume3, pages 176-194.

GROSS, G.W., BOONE, J.M. e HUNT, V.G. (1990), "Neural Networks in Radilogic Diagnosis: II. Interpretation of Neonatal Chest Radiographs", *Investigative Radiology*, Volume 25, pages 1017-1023.

GROSSBERG, S. e CARPENTER, G.A. (1987), "A Massively Parallel Arquitecture for a Self-Organizing Neural Pattern Recognition Machine", *Computer Vision, Graphics and Image Processing*, Volume 37, pages 54-115.

HAMILTON, D., RILEY, P.J., MIOLA, U.E. e AMRO, A.A. (1995), "A Feed Forward Neural Network for Classification of Bull's-Eye Myocardial Perfusion Images", *European Journal of Nuclear Medicine*, Volume 22, pages 108-115.

HALL, L.O., BENSAID, A.M., CLARKE, L.P., VELTHUIZEN, R.P., SILBIGER, M.S. e BEZDEK, J.C. (1992), "A Comparison of Neural Network and Fuzzy Clustering Techniques in Segmenting Magnetic Resonance Images of the Brain", *IEEE Transactions on Neural Networks*, Volume 3, pages 672-681.

[1] HASSOUN, M.H., WANG, C., e SPITZER, A.R. (1994), "NNerve: Neural Network Extraction of Repetitive Vectors for Electromyography - PartI: Algorithm", *IEEE Transactions on Biomedical Engineering*, Volume 41, pages 1039-1052.

[2] HASSOUN, M.H., WANG, C., SPITZER, A.R. (1994), "NNerve: Neural Network Extraction of Repetitive Vectors for Electromyography - PartII: Performance Analysis" , *IEEE Transactions on Biomedical Engineering*, Volume 41, pages 1053-1061.

HEBB, D. O. (1949), *The Organization of Behavior*, First Edition, John Wiley & Sons, New York.

[1] HECHT, N.R. (1987), "Counter-Propagation Networks", *Proceedings of the 1st International Conference on Neural Networks*, June, San Diego, California, USA, part2, pages 19-32.

[2] HECHT, N.R. (1987), "Kolmogorov's Mapping Neural Network Existence Theorem", *Proc. IEEE 1st International Conference on Neural Networks*, June, San Diego, Calif., pp. III-11 - III-14.

HOPFIELD, J.J. (1984), "Neural Networks and Physical Systems with Emergent Collective Computational Abilities", *Proceedings of the National Academy of Science*, Volume 81, pages 3088-3092.

HORN, B.K.P. e SCHUNCK, B.G. (1981), "Determining Optical Flow", *Artificial Intelligence*, Volume 17, pages 185-203.

JACOBS, R.A. (1988), "Increased Rates of Convergence Through Learning Rate Adaptation", *Neural Networks*, Volume 1, pages 295-307.

JOUAN, A., VERDENET, J., CARDOT, J.C., BAUD, M. e DUVERNOY, J. (1990), "Automated Detection of the Left Ventricular Region of Interest by Means of the Extraction of Typical Behaviors in Cardiac Radionuclide Angiographies", *IEEE Transactions on Medical Imaging*, Volume 9, pages 5-10.

KARSSEMEIJER, N., ERNING, L.J.T.O. e EIJKMAN, E.G.J. (1988), "Recognition of Organs in CT-Image Sequences: A Model Guided Approach", *Computers and Biomedical Research*, Volume 21, pages 434-448.

KERNIGHAN, B.W. e RITCHIE, D.M. (1990), *C: A Linguagem de Programação Padrão ANSI*, Segunda Edicão, Editora Campus, Rio de Janeiro.

KIPPENHAM, J.S., BAKER, W.W., PASCAL, S., NAGEL, S. e DURAN, R. (1992), "Evaluation of a Neural Network Classifier for PET Scans of Normal and Alzheimer's Disease Subjects", *Journal of Nuclear Medicine*, Volume 33, pages 1459-1467.

KOHONEN, T. (1989), *Self-Organization and Associative Memory*, Third Edition, Springer-Verlag. Berlin, Germany.

LAMBERTI, C. e SGALLORI, F. (1990), "A Workstation-Based System for 2-D Echocardiography Visualization and Image Processing", *IEEE Transactions on Biomedical Engineering*, Volume 37, pages 796-802.

LILLY, P., JENKINS, J. e BOURDILLON, P. (1989), "Automatic Contour Definition on Left Ventriculograms by Image Evidence and a Multiple Template-Based Model", *IEEE Transactions on Medical Imaging,* Volume 8, pages 173-185.

LIPPMANN, R.P. (1987), "An Introduction to Computing with Neural Nets", *IEEE ASSP Magazine,* Volume 4, pages 4-22.

MAILLOUX, G.E., BLEAU, A., BERTRAND, M. e PETICLERE, R. (1987), "Computer Analysis of Heart Motion from Two-Dimensional Echocardiograms", *IEEE Transactions on Biomedical Engineering,* Volume 34, pages 356-364.

MAILLOUX, G.E., LANGLOIS, F., SIMARD, P.Y. e BERTRAND, M. (1989), "Restoration of the Velocity Field of the Heart from Two-Dimensional Echocardiograms", *IEEE Transactions om Medical Imaging",* Volume 8, pages 143-153.

MCCULLOCH, W.S., e PITTS, W. (1943), "A Logical Calculus of the Ideas Immanent in Nervous Activity", *Bulletin of Mathematical Biophysics, Volume* 5, pages 115-133.

MEUNIER, J., BERTRAND, M., MAILLOUX, G.E. e PETITCLERC, R. (1988), "Local Myocardial Deformation from Speckle Motion", *Computers in Cardiology,* pages 133-136.

MILLER, A.S., BLOTT, B.H. e HAMES, T.K. (1992), "Review of Neural Network Applications in Medical Imaging and Signal Processing", *Medical & Biological Engineering & Computing,* Volume 30, pages 449-464.

MINSKY, M.L. e PAPERT, S. (1988), *Perceptrons: An introduction to Computational Geometry,* Second Edition, M.I.T. Press, Cambridge, MA.

MIRCHANDANI, G. e CAO, W. (1989), "On Hidden Nodes for Neural Nets", *IEEE Transactions on Circuits and Systems,* Volume 36, pages 661-664.

MORENO, L., PIÑEIRO, J.D., SÁNCHEZ, J.L., MAÑAS, S., MERINO, J., ACOSTA, L. e HAMILTON, A. (1995), "Brain Maturation Estimation Using Neural Classifier", *IEEE Transactions on Biomedical Engineering,* Volume 42, pages 428-432.

MOURA, L. e KITNEY, R. (1992), "Automatic Reconstruction of 3D Coronary Artery Segments", *Automedica,* Volume 15, pages 97-121.

NEKOVEI, R. e SUN, Y. (1995), "Back-Propagation Network and Its Configuration for Blood Vessel Detection in Angiograms", *IEEE Transactions on Neural Networks,* Volume 6, pages 64-72.

NELSON, T.R., JOLLA, L. e BOONE, J.M. (1991), "Neural Network Localization of Myocardial Activation Based on Electrocardiogram: Simulated Results", *Radiology,* Volume 181(P), page 144.

PATTICHIS, C.S., SCHIZAS, C.N. e MIDDLETON, L.T. (1995), "Neural Network Models in EMG Diagnosis", *IEEE Transactions on Biomedical Engineering*, Volume 42, pages 486-496.

PHILIP, K.P., DOVE, E.L., MCPHERSON, D.D., GOTTEINER, N.L. (1994), "Automatic Detection of Myocardial Contours in Cine-Computed Tomographic Images", *IEEE Transactions on Medical Imaging*, Volume 13, pages 241-253.

POPE, D.L., PARKER, D.L., CLAYTON, P.D. e GUSTAFSON, D.E. (1985), "Left Ventricular Border Recognition Using a Dynamic Search Algorithm", *Radiology*, Volume 155, pages 513-518.

ROSENBLAT, F. (1962), *Principles of Neurodynamics*, Spartan Books, New York.

RUMELHART, D.E. e MCCLELLAND, J.L.(1986), *Parallel Distributed Processing: Exploration in the Microestructure of Cognition, Volume 1:Foundations*, First Edition, M.I.T. Press, Cambridge.

SCOTT, J.A. e PALMER, E.L. (1993), "Neural network analysis of ventilation-perfusion lung scans", *Radiology*, Volume 186, pages 661-664.

SCOTT, R. (1993), "Artificial Intelligence: Its Use in Medical Diagnosis", *Journal of Nuclear Medicine*, Volume 34, pages 510-514.

SHORTLIFFE, E.H e PERREAULT, L. E.(1990), *Medical Informatics: Computer Applications in Health Care*, First Edition, Addison-Wesley, New York.

SKORTON, D.J., McNARY, C.A., CHILD, J.S., NEWTON, F.C. e SHAH, P.M. (1981), "Digital Image Processing of Two-Dimensional Echocardiograms: Identification of Endocardium", *American Journal of Cardiology*, Volume 48, pages 479-486.

SKORTON, D.J., COLLINS, S.M., GARCIA, E., GEISER, E.A., HILLARD, W., KOPPES, W., LINKER, D. e SCHWARTZ, G. (1985), "Digital Signal and Image Processing in Ecocardiography", *American Heart Journal*, Volume 1, pages 1266-1283.

SPRECHER, D.A. (1965), "On the Structure of Continuous Functions of Several Variables", *Transactions of the American Mathematical Society*, Volume 115, pages 340-355.

STACK, J. P., REDMOND, O. M., CODD, M. B., DERVAN, P. A. e ENNIS, J. T. (1990), "Breast Disease: Tissue Characterization with Gd-DPTA Enhancement Profiles", *Radiology*, Volume 174, pages 491-494.

STUBBS, D.F. (1990), "Multiple Neural Network Approaches to Clinical Expert Systems", *SPIE Applications of Artificial Neural Networks*, 1294, pages 433-441.

SU, M.C. (1994), "Use of neural networks as medical diagnosis expert systems", *Computers in Biology and Medicine*, Volume 24, pages 419-429.

SWETS, J. A. (1988), "Measuring the Accuracy of Diagnostic Systems", *Science*, Volume 240, pages 1285-1293.

TARATORIN, A.M. e SIDEMAN, S. (1993), "Constrained Detection of Left Ventricular Boundaries from Cine CT Images of Human Hearts", *IEEE Transactions on Medical Imaging*, Volume 12, pages 521-533.

THEDENS, D.R., SKORTON, D.J. e FLEAGLE, S.R. (1995), "Methods of Graph Searching for Border Detection in Image Sequences with Application to Cardiac Magnetic Resonance Imaging", *IEEE Transactions on Medical Imaging*, Volume 14, pages 42-55.

THOMPSON, W.B. (1980), "Combining Motion and Contrast for Segmentation", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 2, pages 543-549.

VOGL,T.P., MANGIS, J.K., RIGLER, A.K., ZINK, W.T. e ALLCON, D.L. (1988), "Accelerating the Convergence of the Back Propagation Method", *Biological Cybernetics*, Volume 59, pages 257-263.

WIDROW, B. e HOFF, M.E. (1960), "Adaptive Switching Circuits", *IRE Western Electric Show and Convention Record*, Part 4, pages 96-104.

WOLBERG, G. (1994), *Digital Image Warping*, Third Edition, IEEE Computer Society Press, Washington, U.S.A.

WU, K., GAUTHIER, D. e LEVINE, M.D. (1995), "Live Cell Segmentation", *IEEE Transactions on Biomedical Engineering*, Volume 42, pages 1-12.

WU, Y., GIGER, M.L., VYBORNY, C.J., DOY, K., ASADA, N., SCHMIDT, R.A. e METZ, C.E. (1990), "Application of Neural Networks to Mammographic Diagnosis of Breast Cancer", *Radiology*, Volume 177(P), page 149.

YANG, K.T.A. e CHEN, H.D. (1994), "A Semi-Automated Method for Edge Detection in the Evaluation of Left Ventricular Function Using ECG-Gated Single-Photon Emission Tomography", *European Journal of Nuclear Medicine*, Volume 21, pages 1206-1211.

ZHANG, L.F. e GEISER, E.A. (1988), "An Approach to Optimal Threshold Selection on a Sequence of Two Dimensional Echocardiographic Images", *IEEE Transactions on Biomedical Engineering*,, Volume 29, pages 577-581.