

**UNIVERSIDADE FEDERAL DE SANTA CATARINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
COMPUTAÇÃO**

ANDERSON LUIS SCHVINDT BITTENCOURT

**SEGMENTAÇÃO EM IMAGENS DIGITAIS DE
VEÍCULOS OBTIDAS POR CONTROLADORES
ELETRÔNICOS DE VELOCIDADE**

Dissertação submetida à Universidade Federal de Santa Catarina como parte dos requisitos para obtenção do grau de Mestre em Ciência da Computação

Orientador: Prof. Paulo S. S. Borges, Dr.

Florianópolis, Agosto de 2003

SEGMENTAÇÃO EM IMAGENS DIGITAIS DE VEÍCULOS OBTIDAS POR CONTROLADORES ELETRÔNICOS DE VELOCIDADE

ANDERSON LUIS SCHVINDT BITTENCOURT

Esta Dissertação foi julgada adequada para a obtenção do título de Mestre em Ciência da Computação, área de concentração sistemas de conhecimento e aprovada em sua forma final pelo Programa de Pós Graduação em Ciência da Computação.

Professor Fernando Álvaro Ostuni Gauthier, Dr.
(Coordenador do Curso)

Banca Examinadora

Professor Paulo Sérgio da Silva Borges, Dr.

Professor João Bosco da Mota Alves, Dr.

Professor Jovelino Falqueto, Dr.

Professor Luis Fernando Jacintho Maia, Dr.

Dedico este trabalho a minha mãe Cledi Bittencourt, a meu pai Ayrton Bittencourt (in memoriam), a toda a minha família e amigos e em especial a Rudimar Tellier de Freitas e família.

AGRADECIMENTOS

Agradeço ao professor Paulo S. S. Borges pelos ensinamentos, críticas, discussões, incentivo e motivação. Agradeço-lhe a imensa colaboração sem a qual este trabalho não seria possível, pela atenção dispensada e pelos conselhos que muito contribuíram para o meu crescimento e amadurecimento profissional e pessoal.

Agradeço aos meus amigos, especialmente ao colega de mestrado Tiago Terra, pelo apoio, pelas dicas, críticas, sugestões e pela presença sempre constante durante a realização deste trabalho.

Agradeço também aos meus professores, Marcos Macedo, Elisiane Macedo e Rudimar Tellier de Freitas, pelo grande incentivo que me deram no sentido de ingressar neste mestrado.

Agradeço ainda aos colegas da equipe CEF-SIAPC (Caixa Econômica Federal – Sistema de Avaliação e Provisionamento de Crédito) pela compreensão durante os últimos meses deste trabalho e a todos aqueles que de alguma forma auxiliaram na realização do mesmo.

Por fim, gostaria de agradecer a toda a minha família, minha mãe Cledi, meus irmãos Evandro e Daniela, minha cunhada Vera e meus sobrinhos Elton, Eduarda e Viviane por serem acima de tudo o meu “porto seguro”, meu ponto de apoio em todos os momentos da minha vida.

ÍNDICE

Capítulo 1: INTRODUÇÃO	1
Capítulo 2: IMAGENS OBTIDAS POR CONTROLADORES ELETRÔNICOS DE VELOCIDADE: TÉCNICAS ATUAIS.....	5
2.1 Introdução	5
2.2 Controladores Eletrônicos de Velocidade	5
2.3 Imagens Obtidas por Controladores Eletrônicos de Velocidade	6
2.4 Informações Relevantes	7
2.5 Obtenção da Informação Contida na Placa de um Veículo em uma Imagem Digital	8
2.6 Principais Abordagens Existentes.....	9
2.6.1 Sistema SIAV.....	10
2.6.2 Sistema CLPR (Character License Plate Recognition).....	10
2.6.3 Método Baseado em Campos Randômicos de Markov e Algoritmos Genéticos	11
2.6.4 Método Granulométrico para Segmentação de Placas de Veículos.....	12
2.6.5 Método Baseado em Redes Neurais Artificiais.....	12
2.7 Conclusão	13
Capítulo 3: VISÃO COMPUTACIONAL	14
3.1 Introdução	14
3.2 Fisiologia da Visão.....	15
3.3 Imagens Digitais	18
3.3.1 Captura e Digitalização de Imagens.....	19
3.3.2 Padrão de Compressão de Imagens JPEG (JPG).....	20
3.3.3 Padrão de Compressão de Imagens GIF	21
3.4 O Processo de Visão Computacional.....	22
3.5 Conclusão	23
Capítulo 4: REDES NEURAS ARTIFICIAIS.....	24
4.1 Introdução	24

4.2	O Cérebro Humano.....	25
4.3	O Neurônio Biológico	27
4.4	Redes Neurais Artificiais	29
4.5	Breve Histórico das Redes Neurais Artificiais.....	30
4.6	Neurônio Artificial.....	32
4.7	Funções de Ativação.....	34
4.1.1	Função Limiar	34
4.1.2	Função Linear.....	35
4.1.3	Função Não Linear	35
4.1.4	Algumas Variações.....	36
4.8	Topologia das Redes Neurais Artificiais.....	37
4.9	Principais modelos de Redes Neurais Artificiais	40
4.9.1	Perceptron	40
4.9.2	<i>Perceptron</i> de múltiplas camadas (<i>Backpropagation</i>)	41
4.9.3	Classificador de Carpenter-Grossberg	42
4.9.4	Rede de Kohonen.....	43
4.9.5	Rede de Hopfield.....	44
4.10	Aprendizado em Redes Neurais Artificiais	44
4.10.1	Diferentes Tipos de Aprendizado.....	45
4.11	<i>Perceptron</i> de Múltiplas Camadas	46
4.12	O Algoritmo <i>Backpropagation</i>	48
4.12.1	Problemas no Aprendizado	51
4.13	Conclusão	52
Capítulo 5: UM MÉTODO PARA SEGMENTAÇÃO DA PLACA EM IMAGENS DIGITAIS DE VEÍCULOS.....		53
5.1	Introdução	53
5.2	Equipamentos e Programas Utilizados	54
5.3	Principais Funções do MATLAB [®] Utilizadas Neste Trabalho	54
5.4	Imagens	57
5.5	Segmentação da Linha	57
5.5.1	Binarização da Linha	59
5.5.2	Preparação dos Dados de Treinamento	61
5.5.3	Configuração da Rede Neural	62

5.5.4	Treinamento da Rede Neural	64
5.6	Segmentação da Coluna.....	66
5.6.1	Binarização da Coluna.....	66
5.6.2	Preparação dos Dados de Treinamento	67
5.6.3	Configuração da Rede Neural	68
5.6.4	Treinamento da Rede Neural	69
5.7	Conclusão	70
	Capítulo 6: TESTES E AVALIAÇÕES	71
6.1	Introdução	71
6.2	Avaliação do Treinamento das Redes	71
6.2.1	Treinamento da Rede Neural para Segmentação das Linhas da Imagem	71
6.2.2	Treinamento da Rede Neural para Segmentação das Colunas da Imagem.....	75
6.3	Método de Binarização das linhas da imagem.....	76
6.4	Resultados	77
6.4.1	Segmentação das Linhas	79
6.4.2	Segmentação das Colunas.....	82
6.5	Conclusão	85
	Capítulo 7: CONCLUSÕES	87
7.1	Introdução	87
7.2	Sumário da Dissertação.....	88
7.3	Conclusões Gerais.....	89
7.4	Trabalhos Futuros.....	91
	Capítulo 8: REFERÊNCIAS BIBLIOGRÁFICAS	92

LISTA DE FIGURAS

Fig.2.1	- Funcionamento dos controladores eletrônicos de velocidade.....	6
Fig.2.2	- Imagem obtida por controlador eletrônico de velocidade.....	7
Fig.2.3	- Segmentação da área da placa.....	9
Fig.3.1	- Partes do olho humano.....	15
Fig.3.2	- Células da retina.....	16
Fig.4.1	- O Cérebro humano.....	26
Fig.4.2	- Neurônio biológico.....	28
Fig.4.3	- Neurônio artificial.....	33
Fig.4.4	- Principais funções de ativação.....	37
Fig.4.5	- RNA de múltiplas camadas não recorrente.....	38
Fig.4.6	- RNA não recorrente de uma única camada.....	39
Fig.4.7	- RNA recorrente de uma única camada.....	40
Fig.4.8	- Arquitetura MLP com duas camadas intermediárias.....	47
Fig.4.9	- Tipos de sinais em uma MLP.....	47
Fig. 5.1	- Variação tonal dos níveis de cinza em uma linha com placa.....	58
Fig. 5.2	- Variação tonal dos níveis de cinza em uma linha sem placa.....	58
Fig. 5.3	- Resultado esperado da segmentação das linhas da imagem.....	59

Fig. 5.4	- Representação do valor médio na variação dos níveis de cinza.....	60
Fig. 5.5	- Binarização de linha com placa.....	60
Fig. 5.6	- Binarização de linha sem placa.....	61
Fig. 5.7	- Topologia da rede neural utilizada para segmentar linhas da imagem.....	64
Fig 5.8	- Imagem utilizada para segmentar as colunas.....	66
Fig 5.9	- Saída da rede utilizada para segmentar as colunas.....	66
Fig 5.10	- Entrada da rede utilizada para segmentar as colunas.....	67
Fig. 5.11	- Topologia da rede neural utilizada para segmentar colunas da imagem.....	69
Fig. 6.1	- Gráfico do treinamento utilizando o conjunto LN1.....	73
Fig. 6.2	- Gráfico do treinamento utilizando o conjunto LN2.....	74
Fig. 6.3	- Gráfico do treinamento utilizando o conjunto [LN1 LN2].....	74
Fig. 6.4	- Gráfico do treinamento utilizando o conjunto CL1.....	76
Fig. 6.5	- Imagem onde o método de binarização apresenta problemas.....	77
Fig. 6.6	- Gráfico da predição – conjuntos TLC e TLS.....	81
Fig. 6.7	- Gráfico da predição – conjuntos TLC e TLS (zona nebulosa).....	82
Fig. 6.8	- Gráfico da predição – conjuntos TCC e TCS.....	84
Fig. 6.9	- Gráfico da predição – conjuntos TCC e TCS (zona nebulosa).....	85

LISTA DE TABELAS

Tab.4.1	- Quadro comparativo entre o cérebro humano e o computador.....	27
Tab.4.2	- Quadro comparativo entre computadores e neurocomputadores....	27
Tab. 5.1	- Distribuição das imagens em conjuntos de treinamento e testes.....	57
Tab. 5.2	- Composição dos conjuntos de treinamento utilizados para segmentação das linhas da imagem.....	62
Tab. 5.3	- Parâmetros de treinamento da rede neural utilizada para segmentar as linhas da imagem.....	65
Tab. 5.4	- Composição do conjunto de treinamento utilizado para segmentação das colunas da imagem.....	67
Tab. 5.5	- Parâmetros de treinamento da rede neural utilizada para segmentar as colunas da imagem.....	69
Tab. 6.1	- Valores referentes aos treinamentos da rede neural utilizada para a segmentação das linhas da imagem.....	73
Tab. 6.2	- Valores referentes ao treinamento da rede neural utilizada para a segmentação das colunas da imagem.....	75
Tab. 6.3	- Composição dos conjuntos de testes (segmentação das linhas).....	80
Tab. 6.4	- Resultados obtidos na simulação da rede treinada para segmentar as linhas (<i>threshold</i>)	80
Tab. 6.5	- Resultados obtidos na simulação da rede treinada para segmentar as linhas (zona nebulosa).....	81

Tab. 6.6	- Composição dos conjuntos de testes (segmentação das colunas)..	83
Tab. 6.7	- Resultados obtidos na simulação da rede treinada para segmentar as colunas (<i>threshold</i>)	83
Tab. 6.8	- Resultados obtidos na simulação da rede treinada para segmentar as colunas (zona nebulosa).....	84

RESUMO

Este trabalho apresenta o desenvolvimento de um método baseado em técnicas de tratamento de imagens e inteligência computacional, capaz de segmentar a área correspondente aos caracteres da placa de um veículo em imagens digitalizadas obtidas por controladores eletrônicos de velocidade. São abordados os principais aspectos teóricos das referidas áreas, entre eles os conceitos relativos à fisiologia da visão, imagens digitais, visão computacional e redes neurais artificiais, cujos conceitos dão o suporte necessário à tarefa de extrair e interpretar informações de imagens digitais. A tarefa de identificar automaticamente a informação contida nas placas de veículos em imagens digitalizadas compreende duas fases distintas de processamento. A primeira dessas fases consiste na segmentação da área da imagem em que estão localizados os caracteres da placa; e a segunda, no reconhecimento e obtenção do conteúdo semântico de tais caracteres. Nesta dissertação propõe-se uma contribuição baseada no uso de redes neurais do tipo *perceptron* de múltiplas camadas capazes de realizar a primeira das fases citadas.

ABSTRACT

This work presents the development of a method based on techniques of image processing and computational intelligence, capable of determining the region where a car license plate lies, with its proper characters, using as the source the digital images acquired by electronic speed controllers. In addition, the work presents theoretical aspects of the referred techniques, including concepts related to vision physiology, digital images, computational vision and neural networks. That conceptual review aims at providing the necessary theoretical support to extract and interpret information from digital images. The task of automatically identifying information from a car license plate is divided in two phases. The first phase consists on segmenting the image area where the characters are; the second is to recognize and obtain the semantical contents of those characters. This dissertation proposes a contribution based on the use of neural networks like multilayer perceptron contemplating a proposal for the solution of the problem associated to the first phase.

Capítulo 1: INTRODUÇÃO

O uso de computadores inicialmente foi dirigido para tarefas como cálculos numéricos e manipulação de grandes volumes de dados, ou seja, aplicações bem estruturadas e formalizadas, pois para tais atividades os mesmos apresentavam maior velocidade e precisão que os humanos. Porém, quando a área de aplicação dos computadores se estendeu às tarefas de natureza simbólica, que para os humanos são de resolução natural, como a compreensão de linguagens naturais e imagens, planejamento, dedução ou a solução de problemas que apresentam complexidade exponencial, o desempenho dos computadores foi muito baixo, devido à grande complexidade que envolve a compreensão de tais tarefas e a formulação de uma solução computacional para as mesmas. Buscou-se então criar meios de dar aos computadores capacidades semelhantes as humanas, tais como a capacidade de pensar e agir intuitivamente.

Desde a década de 80, o uso de Inteligência Artificial constitui um dos setores de maior crescimento no mercado de tecnologia de informática. Esse crescimento não é apenas devido maturação da tecnologia, mas ao significativo avanço dos recursos de *hardware*, *software* e humanos, avanço que acaba por criar uma infra-estrutura propícia ao desenvolvimento da área.

Existem vários pesquisadores que acreditam que a Inteligência Artificial (IA) é a chave para o *software* do futuro. As pesquisas nesse sentido estão relacionadas a áreas de aplicação que envolve raciocínio humano, tentando imitá-lo e realizando inferências. Seu grande objetivo, como podemos verificar, é o contínuo aumento da “inteligência do computador”, pesquisando, para isso, os fenômenos de inteligência natural. Hoje, a IA pode ser definida como sendo uma coleção de técnicas suportadas por um computador emulando algumas capacidades dos seres humanos.

As demandas para aplicação de IA são inúmeras, visto que, se levarmos em conta somente os problemas cotidianos, já temos a possibilidade de realização de muitas pesquisas e com isso muito trabalho para pesquisadores, isso sem falar nas áreas industriais e comerciais. A tarefa de extrair informação de imagens digitalizadas pode ser resolvida pelas técnicas de Inteligência Artificial juntamente com as técnicas de Visão Computacional. As atividades que fazem uso intensivo e constante de informação visual, como é o caso dos equipamentos de controle eletrônico de velocidade, são carentes deste tipo de recurso. Em geral, toda a informação extraída de imagens geradas através de tais controladores é obtida pela atividade de análise de um elemento humano, diretamente sobre a imagem digitalizada. Esse fato limita em muito a autonomia dos referidos equipamentos na medida em que os mesmos não conseguem realizar todas as tarefas pertinentes ao controle de velocidade, ficando sua atuação limitada a registrar a infração e sendo todo o processamento de informação que se segue dependente da atuação humana na interpretação dos dados registrados na imagem.

Empresas fabricantes de controladores eletrônicos de velocidade possuem projetos que visam possibilitar a autonomia de tais equipamentos. Tal intenção passa necessariamente pelo desenvolvimento de técnicas capazes de identificar o veículo através da imagem digitalizada do mesmo. A principal informação a ser considerada na identificação de um veículo é o conteúdo da sua placa, embora devem ser levadas em conta também informações referentes a demais características do veículo, tais como a sua cor e o seu modelo.

O objetivo deste trabalho é desenvolver e apresentar um método que, por meio do uso combinado de técnicas de Visão Computacional e Inteligência Artificial, desta segunda especificamente as Redes Neurais Artificiais, que devido a sua capacidade de extrair padrões de conjuntos de dados complexos, como é o caso das imagens digitais, seja capaz de localizar a placa de um veículo em uma imagem digitalizada, facilitando a tarefa de reconhecimento dos caracteres contidos na mesma.

Com o objetivo de orientar a leitura desta dissertação, apresenta-se a seguir, a organização de seus capítulos:

- O capítulo 1 contém a introdução desta dissertação, seus objetivos e considerações preliminares.
- No capítulo 2, procedeu-se uma síntese do funcionamento dos controladores eletrônicos de velocidades, as imagens por eles geradas, sendo apresentadas algumas técnicas atualmente utilizadas para segmentar a área correspondente à placa de um veículo em tais imagens.
- No capítulo 3, apresenta-se uma revisão teórica de alguns conceitos relacionados à visão computacional e imagens digitais, cujo conhecimento visa tornar mais fácil a compreensão deste trabalho.
- O capítulo 4 traz a fundamentação teórica necessária para se compreender o funcionamento das redes neurais artificiais, principal técnica utilizada neste trabalho.
- No capítulo 5 é apresentado um método desenvolvido utilizando-se o *software* MATLAB® e os respectivos módulos de redes neurais artificiais e de processamento de imagens, capaz de segmentar a área correspondente às placas de veículos em imagens digitais obtidas por controladores eletrônicos de velocidade.
- Sucessivamente, o capítulo 6 executa uma análise dos resultados obtidos nas simulações com o método proposto e descrito no capítulo 5, bem como são apresentadas algumas considerações a respeito do processo de treinamento das duas redes neurais propostas.
- Por fim, o capítulo 7 abrange a conclusão da dissertação, incluindo as observações finais, considerações relevantes sobre os assuntos

tratados e também as limitações do trabalho e recomendações para futuras pesquisas.

Capítulo 2: IMAGENS OBTIDAS POR CONTROLADORES ELETRÔNICOS DE VELOCIDADE: TÉCNICAS ATUAIS

2.1 Introdução

Este capítulo apresenta uma introdução ao controle eletrônico de velocidade com relação às imagens obtidas por estes sistemas. Neste caso, a tarefa é segmentar a região da imagem que compreende a placa do veículo e nesta parte do trabalho é feita uma revisão do assunto.

Dentro desse contexto, aborda-se tópicos relacionados ao controle de trânsito, aos controladores eletrônicos de velocidade e seu funcionamento, às imagens obtidas por tais equipamentos, às informações relevantes contidas em tais imagens e à tarefa de obtê-las. Por fim, serão tratadas algumas das principais abordagens hoje existentes a respeito do problema de se extrair informações de imagens provenientes de controladores eletrônicos de velocidade.

2.2 Controladores Eletrônicos de Velocidade

O controle de trânsito é, fundamentalmente, a supervisão do movimento de veículos, pessoas e bens, com o objetivo de garantir a sua eficiência e segurança. Uma rodovia pode ser considerada eficiente quando acomoda as necessidades para a movimentação dos seus usuários ao menor custo possível. Pode ainda, ser tida como segura, ao reduzir ou eliminar os acidentes. Desta forma, o controle eletrônico de trânsito tem como principal objetivo a organização do movimento de pessoas e bens da maneira mais segura e eficiente possível [40]. Neste contexto os sistemas de registro fotográfico de infrações de trânsito têm sido extensamente

utilizados, com grande sucesso, como um dos principais agentes para a mudança de comportamento dos condutores de veículos, principalmente em países com problemas crônicos de desrespeito a sinalização e elevados índices de acidentes.

A imensa maioria dos controladores eletrônicos de velocidade produzidos e utilizados atualmente no Brasil tem seu funcionamento baseado na utilização de sensores instalados na pista de rodagem juntamente com câmeras situadas em pontos que permitam capturar a imagem dos veículos que passam pelos sensores (fig. 2.1). A utilização desses dois mecanismos permite que o equipamento, além de fazer um controle estatístico do fluxo de veículos por hora e tamanho, calcule, através das informações obtidas pelos sensores, a velocidade do veículo ao cruzá-los. Tal informação é comparada com o limite pré-estabelecido para o local onde está instalado o equipamento e, caso este limite seja ultrapassado, o equipamento registra, através da câmera digital, a imagem do veículo infrator.

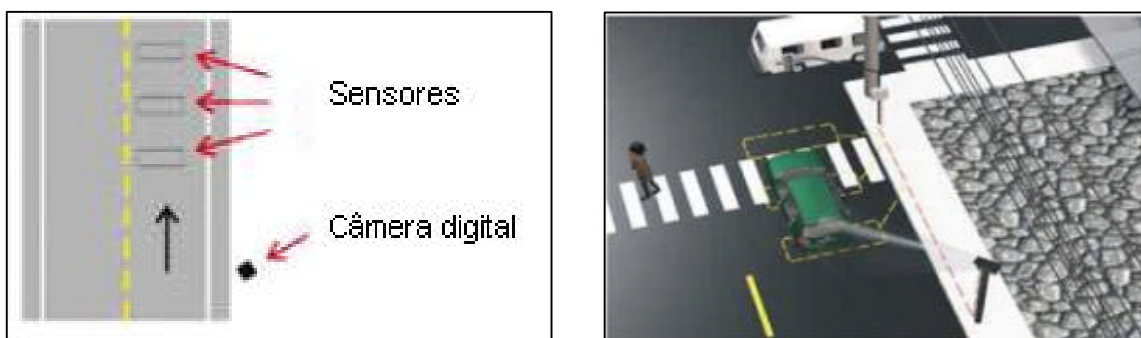


Figura 2.1: Funcionamento dos controladores eletrônicos de velocidade

2.3 Imagens Obtidas por Controladores Eletrônicos de Velocidade

As imagens capturadas pela câmera de um equipamento controlador de velocidade, tais como a imagem apresentada a seguir (fig. 2.2), possuem normalmente resolução espacial de 755x576 pixels, podendo ser coloridas ou

monocromáticas e registrar imagens frontais ou traseiras dos veículos infratores conforme a sua instalação. Tais imagens são armazenadas em um dispositivo localizado junto ao aparelho, e posteriormente, recolhidas por pessoas encarregadas de processá-las.



Figura 2.2: Imagem obtida por controlador eletrônico de velocidade

O processamento feito sobre a imagem recolhida consiste na identificação visual, feita por um elemento humano, das informações relevantes nela contidas. Tal processamento envolve a abertura, uma por uma, das imagens em um *software* gráfico como o *Imaging for Windows* da Microsoft® por exemplo, com a finalidade de obter as referidas informações. Uma vez obtidas tais informações, as mesmas são cruzadas com os dados armazenados no banco de dados dos órgãos de controle de trânsito competentes, os DETRANS, para que sejam identificadas possíveis fraudes e registradas as infrações pertinentes ao veículo em questão.

2.4 Informações Relevantes

Por informações relevantes contidas na imagem digital de um veículo entende-se todas as informações capazes de identificar o mesmo. A principal

delas é a sua placa sendo que, em seguida podem ser citadas outras tais como a sua cor, o seu tamanho, o seu fabricante e o seu modelo.

Este trabalho propõe um método para resolver parte do problema de identificação da informação contida na placa de um veículo. Outras informações, porém, são úteis e embora seu tratamento não faça parte deste trabalho, são utilizadas para identificar problemas como placas falsas, adulteradas e clonadas, entre outros, e representam um desafio aos pesquisadores que se dedicam ao assunto.

2.5 Obtenção da Informação Contida na Placa de um Veículo em uma Imagem Digital

A tarefa de obter a informação contida na placa de um veículo em uma imagem digital se divide geralmente em duas etapas. A primeira delas, na qual se insere este trabalho, compreende rotinas que permitem localizar a placa do veículo na imagem e a segunda compreende rotinas que possibilitam a obtenção do conteúdo semântico dos caracteres da mesma. Serão feitas agora algumas considerações a respeito dessas duas etapas.

A segmentação (ver item 3.4) da área da placa, como já foi dito, é o primeiro passo na tarefa de identificar o conteúdo da mesma. Em uma imagem digital, a localização da placa varia muito, o que torna consideravelmente complexa a tarefa de obtê-la. A seguir podemos observar uma imagem de um veículo apresentando o resultado esperado da fase de segmentação da placa do mesmo.



Figura 2.3: Segmentação da área da placa

A segunda fase consiste em, uma vez localizada a placa, obter o conteúdo semântico de seus caracteres. A maior parte dos trabalhos que se propõem a resolver essa tarefa, utilizam redes neurais artificiais que são capazes de reconhecer qualquer um dos 26 caracteres do alfabeto, bem como qualquer um dos 10 dígitos do sistema numérico decimal. Ambos, caracteres e dígitos, formam o conjunto possível de informação em uma placa de veículo. Mais informações a respeito destas técnicas de reconhecimento de caracteres são facilmente encontradas consultando-se as referências bibliográficas sobre o assunto, inclusive alguns trabalhos citados a seguir neste capítulo, abordam de forma mais específica este tema [42][1].

2.6 Principais Abordagens Existentes

Apesar de serem encontrados na literatura um número elevado de publicações ligadas ao problema de segmentação de imagens e de reconhecimento de padrões, encontram-se poucos artigos ligados à tarefa de localização e reconhecimento de placas de veículos em imagens digitais. A seguir serão abordados de forma resumida alguns trabalhos que se propõe a estudar o assunto, procurando dar uma visão rápida com relação às metodologias empregadas e os resultados obtidos pelos mesmo. Mais informações sobre

qualquer um dos trabalhos relacionados a seguir podem ser obtidas consultando-se diretamente a bibliografia citada.

2.6.1 Sistema SIAV

O sistema SIAV (Sistema de Identificação Automática de Veículos) [41] utiliza, para a segmentação da placa de um veículo em uma imagem digital, o método baseado na variação tonal padronizada por meio da análise do gradiente positivo na área da placa. O algoritmo de localização da placa pode ser dividido em dois procedimentos: No primeiro procedimento tem-se a localização da placa por meio da procura por variação tonal padronizada através da análise do gradiente positivo na área da placa e da binarização local adaptativa sobre a região encontrada ou, se necessário, sobre a imagem inteira. No final desta etapa, tem-se como resultado as coordenadas do provável local da placa. No segundo procedimento há a confirmação do local pela análise do número de dígitos encontrados através da procura por dígitos de tamanhos pré-definidos e por análise dos dígitos selecionados verificando a formação de algum grupo válido. Como resultado final, tem-se a confirmação do local correto da placa. Esse sistema apresenta, segundo os autores, uma taxa de acerto na segmentação da placa de 82,4% dos casos.

2.6.2 Sistema CLPR (Character License Plate Recognition)

No trabalho intitulado “License Plate Recognition, Knowledge-Based Intelligent Techniques in Character Recognition” [1] é apresentado um sistema de identificação de veículos em vias públicas holandesas. O sistema como um todo consiste em quatro unidades principais: uma unidade de segmentação, uma unidade de isolamento, uma unidade de reconhecimento e uma unidade de análise sintática. A unidade de segmentação determina o local da placa baseado

em características estruturais e algumas medidas de tamanho. A placa extraída é passada para a fase seguinte que isolará cada caracter. Os caracteres então são processados pelo reconhecedor de caracteres e, em seguida, o analisador sintático checa cada possível caracter retornado pelo reconhecedor utilizando um conjunto de regras presentes nas placas holandesas. Se estas regras não forem satisfeitas, ou um dos caracteres for irreconhecível, a imagem é rejeitada. Nesse sistema, segundo o autor, as taxas de acerto nas fases de segmentação/isolação são de 94% das imagens tratadas.

2.6.3 Método Baseado em Campos Randômicos de Markov e Algoritmos Genéticos

No trabalho intitulado “Automatic License Extration from Moving Vehicles” [2] é apresentado um estudo de extração de regiões aplicado ao problema específico de localização de placa de veículo. Nesse trabalho é apresentado uma abordagem para extrair, seguir, e binarizar os caracteres em placas dos veículos em movimento. A hipótese base é que uma placa de veículo consiste em caracteres escuros em um fundo claro. A região do texto de uma placa é localizada a partir da informação de textura. Um processo de “warping” é empregado para corrigir a distorção de perspectiva no momento da aquisição da placa. O problema de localização da placa é tratado como a localização de texto em imagens. Na solução apresentada, foi utilizado um método da variação espacial para identificar a região do texto. Para cada pixel na região marcada como texto, computam-se as derivadas parciais usando-se janelas de tamanho 5x5. A binarização dos caracteres da placa emprega de forma simultânea a informação temporal e regional sob a estrutura de campos aleatórios de Markov. O processo é otimizado utilizando-se algoritmos Genéticos. Os autores não forneceram nenhum dado numérico referente ao desempenho do sistema e à qualidade dos resultados obtidos.

2.6.4 Método Granulométrico para Segmentação de Placas de Veículos

No trabalho intitulado “Método Granulométrico para segmentação de placas de veículos” [3] é apresentado um método de localização de caracteres da placa de um veículo em imagens digitais que utiliza técnicas de morfologia matemática envolvendo reconstrução, granulometria e mapeamento a partir de resíduos morfológicos. Segundo os autores, tal método obteve sucesso, no que se refere à localização da placa, em 75% das imagens tratadas.

2.6.5 Método Baseado em Redes Neurais Artificiais

No trabalho intitulado “the neural network based method for locating car license plate” [25] é apresentado um método de localização de placas utilizando redes neurais artificiais. As redes neurais, neste trabalho, são utilizadas como filtros que produzem uma classificação baseada em janelas locais dos pixels da placa e do fundo da imagem, analisando as propriedades de cor e textura de uma placa. O pós-processamento destas imagens filtradas encontra as “bounding boxes”, que são as regiões da imagem que possuem características de cor e textura semelhantes as da placa. Duas redes neurais são utilizadas como filtros que independentemente analisam as seções horizontais e verticais da imagem. Segundo os autores, esse método é robusto por não ser afetado pelo ruído e por degradações devido à estabilidade das redes neurais. Testes com duas bases, a primeira contendo 200 imagens de veículos na entrada de um estacionamento, e a segunda composta de 200 imagens de veículos na estrada mostraram taxas de extração de 99% e de 97,5% respectivamente.

2.7 Conclusão

Neste capítulo foram abordados os temas relativos aos controladores eletrônicos de velocidade e as imagens obtidas através dos mesmos. Com a finalidade de proporcionar uma melhor familiarização a respeito da tarefa de extrair informações destas imagens, foi feita uma breve introdução a respeito do funcionamento de tais equipamentos, das imagens por ele geradas e das atuais abordagens a respeito do assunto.

Nos próximos capítulos deste trabalho serão apresentados os principais conceitos relativos à visão computacional, às imagens digitais, às redes neurais artificiais, bem como um método para localização da área correspondente à placa do veículo em imagens digitais obtidas por controladores eletrônicos de velocidade.

Capítulo 3: VISÃO COMPUTACIONAL

3.1 Introdução

A possibilidade de capacitar uma máquina a reconhecer o ambiente em que a mesma atua vem motivando cada vez mais pesquisadores a investir esforços no mais complexo dos sentidos humanos: a visão. Assim como grande parte dos processos mentais e perceptivos do homem, a visão ainda apresenta uma série de aspectos desconhecidos para a ciência. Embora nas últimas décadas tenha evoluído muito a compreensão dos mecanismos fisiológicos que permitem o seu funcionamento, o estado da arte em visão computacional, ainda carece de um modelo genérico de percepção visual passível de ser aplicado na prática[42].

Dar uma definição precisa do processo de percepção visual é uma tarefa talvez tão complexa quanto descrever seus mecanismos. Segundo BALLARD&BROWN [4], a “Visão computacional é a construção de descrições explícitas e significativas de objetos físicos através de imagens”. A visão é, antes de tudo, uma tarefa de representação e processamento de informações, sendo portanto adequada ao tratamento computacional [5] e é justamente a falta de conhecimento exato de como ela é realizada e como funcionam os mecanismos físicos e psíquicos que a compõe que a tornam um problema tão difícil de ser tratado pelas técnicas de computação.

A abordagem trabalhada hoje em dia para realizar tarefas de interpretação e obtenção de informações de imagens digitalizadas consiste na aplicação de um determinado conjunto bastante específico de técnicas e algoritmos para resolver sub-tarefas bastante limitadas dentro do universo de possibilidades que a informação contida em uma imagem digital oferece. Essas sub-tarefas e seus algoritmos compreendem, entre outras, as atividades de filtragem, segmentação, detecção de bordas de objetos, classificação, análise de texturas, representação e

identificação de formas e objetos. Todas essas técnicas são relacionadas à visão computacional, à imagem digital e suas aplicações, assuntos que, juntamente com os aspectos relativos a fisiologia da visão, serão abordados a partir de agora neste capítulo afim de dar embasamento teórico a respeito de alguns conceitos relativos à imagem digital e suas formas de tratamento computacional.

3.2 Fisiologia da Visão

O olho humano possui uma estrutura semelhante a de uma máquina fotográfica, ou vice-versa (fig. 3.1). Ele é uma câmera fechada, envolvida por um tecido branco e opaco chamado *esclera*, que reveste todo o globo ocular com exceção apenas de uma área transparente por onde a luz penetra, denominada *córnea*. A quantidade de luz que chega ao interior do olho é controlada pela *pupila* e em seguida pelo *crystalino*, onde o feixe luminoso é focalizado em direção à *coróide*. Protegida pela *coróide* está a *retina* (fig. 3.2), uma lâmina foto-sensível formada por três camadas de células nervosas sustentadas pela *membrana glial*

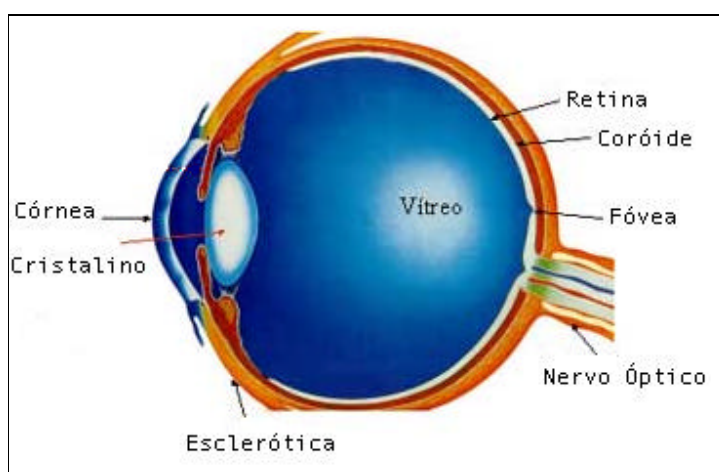


Figura 3.1: Partes do olho humano

Na primeira camada da *retina* encontramos os receptores primários, divididos em *cones* e *bastonetes* (o olho humano possui cerca de 115 a 120 milhões de *bastonetes*). Os *cones* são estimulados por diferentes comprimentos de onda e responsáveis pela percepção da cor, concentrando-se na área central da *retina* chamada *fóvea*, que é a região de máxima acuidade visual. Os *bastonetes* estão localizados, predominantemente, na periferia da *retina*, apresentando sensibilidade à intensidade luminosa. Acima das células receptoras está a camada de células bipolares, que realizam a sinapse entre os *cones* e os *bastonetes*. A terceira camada da retina é composta pelas *células ganglionares* que se ligam diretamente ao *encéfalo* a partir do ponto cego da *retina*, através do *nervo óptico*.

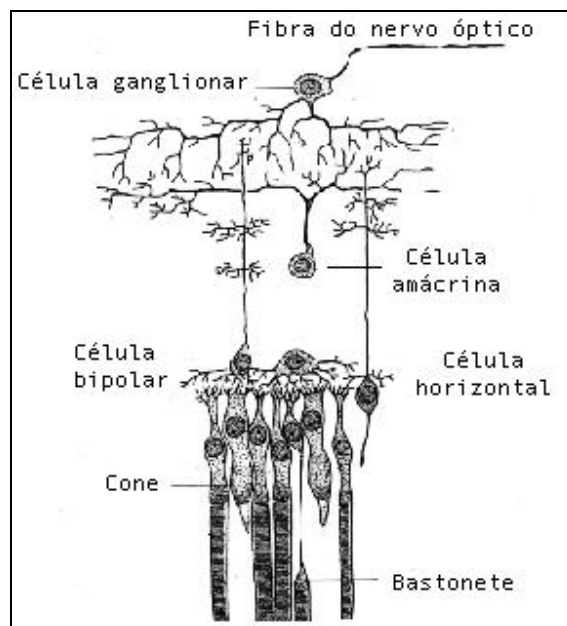


Figura 3.2: Células da retina

A *retina* apresenta uma estrutura celular de conexão entre as *células bipolares* e *ganglionares*, permitindo um elevado grau de integração entre as informações recebidas pelos receptores. As *células horizontais* e as *células amácrinas* são responsáveis pela interconexão das demais estruturas, o que confere à *retina* uma complexidade neuronal próxima a tecidos do cérebro [6].

Estudos comprovam que, apesar da maior parte da atividade organizadora da visão ocorrer no *córtex cerebral*, o processo de combinação e organização das informações se inicia na retina [7]. Dessa forma, a *retina* pode atuar como um filtro, procedendo uma pré-análise da imagem recebida e transmitindo ao cérebro apenas as informações fundamentais para seu reconhecimento [8].

Após processados na *retina*, os sinais captados pelos receptores são enviados através do *nervo óptico* até o *córtex cerebral*. Ao contrário do que usualmente se imagina, apenas parte das fibras que corresponde à porção medial da retina cruza o *quiasma óptico*, em direção ao lado do cérebro oposto ao olho de origem. As fibras da região lateral da *retina* são representadas no mesmo lado do cérebro. Feita a distribuição do feixe de fibras, estas alcançam os *núcleos geniculados laterais* do *tálamo*, onde se ligam em sinapse a um novo conjunto de neurônios, que conduzem os sinais às *áreas estriadas* do córtex. É no córtex que ocorre a abstração das propriedades dos estímulos, por meio de estruturas hierárquicas compostas de *células simples* e *células complexas* [6].

Estudos realizados com gatos apresentam aspectos interessantes da organização do sistema nervoso visual dos mamíferos no reconhecimento de linhas retas. As *células ganglionares* da *retina* se ligam em grupos às células simples do córtex, de forma que estas passam a analisar seqüências lineares presentes na imagem percebida. A atividade de uma célula simples é condicionada à presença de linhas luminosas ou escuras na imagem, sendo que a intensidade de ativação dependerá do ângulo de inclinação destas linhas. Ligadas às células simples estão as células complexas que realizam a montagem dos sinais indicadores de linhas, aplicando uma função lógica “OU”, para a composição final dos objetos visualizados. Se, durante os primeiros meses de vida, um gato é exposto apenas a padrões constituídos de linhas verticais e horizontais, ele passará a apresentar comportamento anômalo, considerando padrões de linhas oblíquas como sendo horizontais ou verticais. As células simples responsáveis pela detecção de linhas com inclinação diferente de 0 ou 90

graus não apresentarão atividade, causando a deficiência na classificação dos novos padrões [9].

Os aspectos discutidos acima mostram que as técnicas de abstração utilizadas em Visão Computacional apresentam diversas semelhanças com o funcionamento do sistema visual natural. Vimos que a *retina* exerce um papel seletivo, extraíndo da imagem informações suficientes para seu reconhecimento. Para algumas tarefas, apenas o contorno dos objetos precisa ser captado para assegurar sua classificação. A estrutura hierárquica, composta pelas *células ganglionares* da *retina* e as células simples e complexas do córtex, é responsável pela gradativa abstração das informações visuais, partindo da imagem captada à associação de conceitos, por meio da análise de características estruturais dos objetos visualizados. Metodologias que utilizam retinas de pontos devem permitir esta função durante o próprio processo de reconhecimento, como ocorre nas redes neurais, onde a caracterização dos padrões fica codificada nos pesos de conexões entre as unidades.

3.3 Imagens Digitais

Segundo WILLRICH [10], imagens digitais não são revisáveis porque seu formato não contém informações estruturais. Elas podem resultar de capturas do mundo real (via escaneamento de uma página impressa ou foto, câmeras digitais) ou elas podem ser sintetizadas pelo computador (via programas de paint, captura da tela, conversão de gráficos em imagens bitmap). Após digitalizadas, as imagens podem ser manipuladas com editores de imagens que não produzem documentos que retém a estrutura semântica.

As imagens digitais são representadas por bitmaps, ou seja, por uma matriz bidimensional de elementos de imagem chamados pixels. Um pixel é o menor elemento de resolução da imagem, ele tem um valor numérico chamado de amplitude. O número de bits disponíveis para codificar um pixel é chamado de

profundidade de amplitude ou profundidade de pixel. Exemplos típicos de profundidade de pixel são 1 (utilizado para imagens preto e branco), 2, 4, 8, 12, 16 ou 24 bits. Este valor numérico pode representar um ponto preto/branco, um nível de cinza, ou os três valores do atributo de cor do pixel de imagens coloridas.

O número de linhas da matriz de pixels é chamado de resolução vertical da imagem, já o número de colunas é chamado de resolução horizontal da imagem e o produto da resolução vertical pela resolução horizontal denomina-se resolução espacial ou resolução geométrica. A resolução geométrica ou espacial estabelece a frequência de amostragem final da imagem, desta forma, quanto maior esta resolução, mais detalhes da imagem podem ser captadas na representação matricial de pixels da imagem. Devemos considerar que a resolução geométrica dada em termos absolutos não fornece muita informação sobre a resolução real da imagem quando esta é realizada em dispositivo físico, isso porque, neste caso, fica-se na dependência do tamanho físico do pixel do dispositivo. Uma medida mais confiável de resolução é dada pela densidade de resolução da imagem que fornece o número de pixels por unidade linear de medida. Em geral utiliza-se pixel por polegada (ppi – “Pixels per inch”) que também é chamada de dpi “dots per inch”.

Imagens digitais no formato bitmap necessitam de mais capacidade de armazenamento do que gráficos e textos, isto porque, como os bitmaps ignoram a semântica da imagem, duas imagens de mesma dimensão ocupam o mesmo espaço, ou seja, um simples quadrado e uma foto digitalizada com dimensões idênticas ocupam o mesmo espaço de armazenamento.

3.3.1 Captura e Digitalização de Imagens

A captura de imagens é feita usando câmeras que focam a imagem em uma superfície foto-sensível de sensores CCD (Charge-Coupled Device), o brilho

de cada ponto (pixel) da imagem focalizada é então convertido em uma carga elétrica.

Nas imagens monocromáticas, geralmente usadas em aplicações de visão computacional, como é o caso deste trabalho, apenas um sinal de intensidade luminosa é produzido formando uma imagem em tons de cinza.

Nas imagens coloridas, três sinais de intensidade luminosa são capturados e, através destes, qualquer cor pode ser reproduzida, ou seja, para capturar uma imagem colorida uma câmera divide a luz dos pontos em seus componentes de vermelho, verde e azul (RGB), estes três componentes de cor são focalizados na câmera por sensores específicos para cada um deles e então convertidos em sinais elétricos separados para cada componente de cor.

3.3.2 Padrão de Compressão de Imagens JPEG (JPG)

O padrão JPEG (Joint Photographic Expert Group) foi desenvolvido, em 1992. Ele representa uma das melhores tecnologias de compressão de imagem. Dependendo da imagem, taxas de compressão podem alcançar 25 para 1 sem degradações notáveis. Em média, JPEG obtém uma taxa de compressão de 15:1. [10]

O algoritmo JPEG decompõe a imagem de entrada em blocos fonte de 8x8 pixels e então transforma estes blocos no domínio da frequência usando a transformada discreta de co-seno (DCT). O DCT efetua uma separação dos componentes de baixa e de alta frequência presentes numa imagem, permitindo que se faça uma seleção destas últimas, de acordo com a qualidade pretendida para a imagem comprimida.

As boas taxas de compressão alcançadas pelo formato JPEG são baseadas na exploração das limitações da visão humana, a qual apresenta sensibilidades diferentes com relação aos componentes de frequência presentes

numa imagem. Baseado no fato de que a visão humana é menos sensível as altas freqüências, estas podem ser desprezadas sem que isso resulte grandes alterações no conteúdo da imagem. O JPEG é parametrizável neste sentido, quanto maior é a compressão escolhida, menor é o número de componentes de alta freqüência da imagem desprezados.

O formato JPEG apresenta ótimas taxas de compressão para imagens fotográficas naturais multi-tonais, diminuindo consideravelmente quando aplicado a imagens gráficas com contornos e áreas bem definidas de cor ou à imagens de texto, como é o caso dos logotipos. O JPEG introduz ruído nas zonas de imagem compostas por cores sólidas, o que pode distorcer o aspecto geral da imagem. Comparado ao GIF (visto seguir neste capítulo), verifica-se que a imagem GIF comprime mais eficazmente que a JPEG e que apresenta uma melhor definição dos contornos do texto.

3.3.3 Padrão de Compressão de Imagens GIF

Atualmente esse é um dos formatos de armazenamento de imagens sem perdas que oferece as melhores taxas de compressão. Por ser sem perdas, o GIF preserva todos os dados visuais na figura, nenhuma informação é descartada ou alterada durante a compressão/descompressão o que faz com que suas taxas de compressão não sejam muito grandes (em geral 4:1). O formato GIF apenas admite o tratamento de imagens com uma profundidade de cor de até 8 bits/pixel, ou seja, imagens com um Máximo de 256 cores e por isso sua utilização é mais notada em imagens produzidas por *softwares* gráficos de desenho do que em imagens fotográficas.

3.4 O Processo de Visão Computacional

Embora constituindo um processo integrado, a visão computacional pode ser compreendida através da descrição de seis grandes etapas interligadas. Tais etapas são vistas a seguir.

Aquisição: nesta etapa, as informações visuais do ambiente são convertidas em sinais elétricos através de dispositivos ou sensores ópticos. A qualidade da imagem obtida é de fundamental importância para as demais etapas do processo. Aspectos mais detalhados a respeito da aquisição de imagens foram apresentados no item 3.3.1 deste trabalho.

Pré-processamento: As imagens obtidas na etapa de aquisição podem apresentar ruídos e pouca definição de detalhes, sendo que a utilização de técnicas de processamento digital de imagens prepara a imagem para as próximas etapas, eliminando ruídos, suavizando efeitos indesejados decorrentes da resolução espacial e quantização de intensidades luminosas, além de realçar detalhes importantes para a detecção de objetos.

Segmentação: Na etapa de segmentação, a imagem é dividida em regiões que constituem os diversos objetos nela representados. A identificação de um objeto baseia-se na detecção de descontinuidades ou similaridades na imagem, gerando uma representação abstrata de seu contorno ou da região que ocupa. Inicialmente quando uma imagem é capturada, a existência de informações pertencentes à cena que muitas vezes não interessam ao sistema, como, por exemplo, o fundo contra o qual a imagem se encontra, faz com que seja necessária a retirada destas informações através de algoritmos de segmentação de imagens.

Descrição: Cada objeto identificado no processo de segmentação da imagem é analisado para a extração de algumas de suas características. Esse conjunto de características é denominado *padrão*, e representa o objeto nas etapas seguintes.

Reconhecimento: Na etapa de reconhecimento, o padrão de cada objeto identificado é comparado com classes de padrão já conhecidas com o objetivo de se decidir a qual grupo ele pertence. Metodologias de reconhecimento devem estabelecer um compromisso entre eficiência e confiabilidade, além de possibilitarem revisão do conhecimento através da experiência adquirida.

Interpretação: Conforme a natureza do problema tratado no sistema de visão computacional, pode ser necessária uma análise de cena, onde os objetos são relacionados entre si, buscando-se uma consciência descritiva do ambiente em que os mesmos se encontram.

3.5 Conclusão

Os assuntos apresentados neste capítulo foram extraídos da bibliografia utilizada e apresentam alguns conceitos e técnicas cujo conhecimento se faz necessário quando nos propusemos a executar tarefas correspondentes tratamento de imagens digitais. Além dos aspectos fisiológicos da visão, da formação de imagens e da conceitualização de imagens digitais, um assunto muito importante aqui tratado, no que se refere ao escopo deste trabalho, é o processo de visão computacional, onde se buscou dar uma visão clara e simplificada das diversas etapas que compreendem um método de visão computacional. Com estas definições, fica mais clara e fácil a compreensão deste trabalho. No capítulo seguinte, serão abordados os conceitos relativos a Redes Neurais Artificiais, técnica em cujo emprego se baseia a solução proposta neste trabalho.

Capítulo 4: REDES NEURAIIS ARTIFICIAIS

4.1 Introdução

As redes Neurais Artificiais, mais usualmente denominadas Redes Neurais, baseiam-se no princípio de que o cérebro humano processa informações de uma forma inteiramente diferente do computador digital convencional. O cérebro humano é um sistema de processamento de informações altamente complexo, não linear e paralelo, constituído por aproximadamente 10 bilhões de elementos ou unidades estruturais chamadas de neurônios [11]. Sua capacidade para realizar certos processamentos, tais como reconhecimento de padrões, percepção, classificação e controle motor, muito mais rápido que qualquer computador digital hoje existente, tem motivado o estudo cada vez mais aprofundado das redes neurais artificiais e suas aplicações.

Consideremos inicialmente como exemplo, a visão humana, que é uma tarefa de processamento de informação [12]. A função do sistema visual é fornecer uma representação do ambiente e também a informação necessária para que possamos interagir com este ambiente. O cérebro humano realiza rotineiramente tarefas de reconhecimento visual, em aproximadamente 100-200 ms (milissegundos) ao passo que tarefas de complexidade muito menor podem levar dias para serem executadas em um computador convencional [13].

Outro exemplo que pode ser considerado é o sonar de um morcego. O sonar é um sistema ativo de localização por eco que, além de fornecer informações sobre a distância até um objeto alvo, um inseto por exemplo, também fornece informações sobre a velocidade relativa do alvo, o tamanho do alvo, o tamanho de varias características do alvo, o azimute e a elevação do alvo [14]. A complexa estrutura neural necessária para extrair essas informações do eco do alvo ocorre

no interior de um cérebro do tamanho de uma ameixa e a facilidade com que ele executa estas tarefas não tem similar em sistemas artificiais.

Baseadas nesses e em outros exemplos, as Redes Neurais Artificiais tentam reproduzir as funções das redes biológicas, tendo sua inspiração baseada nas células neurais naturais, ou seja, no neurônio biológico que é tratado mais detalhadamente no item 4.3 deste trabalho. Este capítulo abordará alguns conceitos importantes a respeito das Redes Neurais Artificiais, afim de dar o embasamento teórico necessário a compreensão deste trabalho.

4.2 O Cérebro Humano

O cérebro humano organiza-se em diferentes regiões, cada uma responsável por diferentes funções. As maiores partes são os hemisférios cerebrais que ocupam a maior parte do crânio. Tais hemisférios são organizados em camadas, sendo a camada mais externa – o córtex cerebral – a mais complexa, na qual as células nervosas estão extremamente adensadas permitindo uma grande interconectividade entre as mesmas. As funções do córtex cerebral não são ainda bem conhecidas, mas há indicadores que esta região parece ser o local onde se concentram as principais funções do cérebro e o núcleo da inteligência.

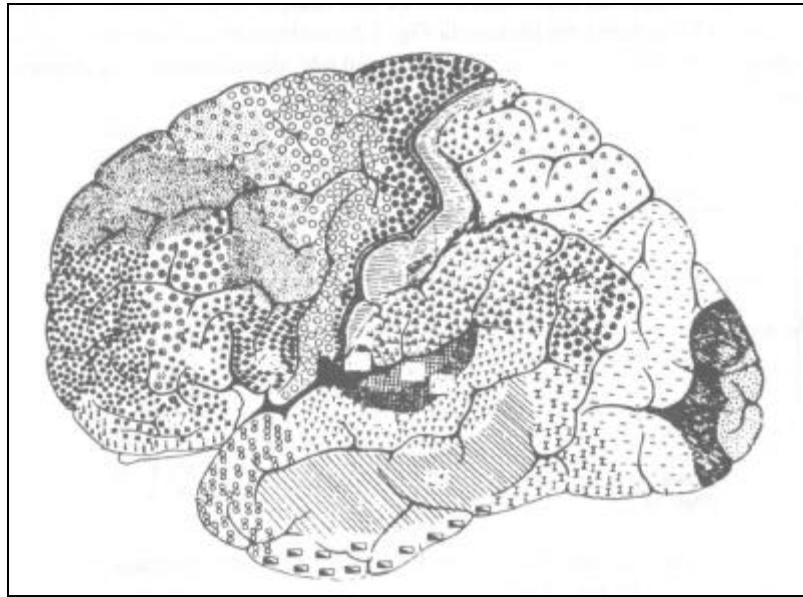


Figura 4.1: O cérebro humano

O esforço para entender o cérebro se tornou mais fácil pelo trabalho pioneiro de Ramón y Cajál [15] que introduziu a idéia de neurônios como constituintes estruturais do cérebro. Atualmente sabe-se que os neurônios são de cinco a seis ordens de grandeza mais lentos que as portas lógicas de silício. Os eventos em um circuito de silício acontecem na ordem de nanosegundos, enquanto que os eventos neurais acontecem na ordem de milissegundos. Entretanto, o cérebro compensa esta taxa de operação relativamente lenta de um neurônio pelo número realmente espantoso de neurônios com conexões maciças entre si. A eficiência energética do cérebro é de aproximadamente 10^{-16} joules por operação por segundo [16].

A tabela a seguir traça um comparativo entre o cérebro humano e o computador. Esta comparação fornece uma idéia mais clara sobre a capacidade adaptativa do cérebro humano, em contraste com a rigidez e precisão dos computadores convencionais.

Parâmetro	Cérebro Humano	Computador
Material	Orgânico	Metal e Plástico
Elementos computacionais	Neurônio simples	Processadores
Velocidade	Milisegundos	Nanosegundos
Tipo de processamento	Paralelo	Seqüencial
Confiabilidade dos elementos	Baixa	Alta
Tolerância a falhas	Grande	Quase nenhuma
Tipo de sinal	Impreciso	Preciso, simbólico
Armazenamento	Adaptativo	Estático
Controle de processos	Distribuído	Centralizado
Ligações entre elementos processados	10.000	< 10

Tabela 4.1 – Quadro comparativo entre o cérebro humano e o computador

O mesmo paralelo pode ser traçado comparando o computador com as redes neurais. Para tanto, a comparação não se dará com um computador específico encontrado no mercado, mas sim com o paradigma predominante nos computadores atuais.

Computadores	Neurocomputadores
Executa programas	Aprende
Executa operações lógicas	Executa operações não lógicas, transformações, comparações
Depende do modelo ou do programador	Descobre as relações ou regras dos dados e exemplos
Testa uma hipótese por vez	Testa todas as possibilidades em paralelo

Tabela 4.2 – Quadro comparativo entre computadores e neurocomputadores

4.3 O Neurônio Biológico

O neurônio biológico é uma célula composta basicamente por água, eletrólitos, proteínas, lipídios e carboidratos [17], sendo dividida em três seções: o corpo da célula ou soma, os dendritos e o axônio, cada uma com funções

específicas e complementares. O corpo de um neurônio é o centro dos processos metabólicos da célula, mede apenas alguns milésimos de milímetros e é a partir dele que se projetam os dendritos e o axônio. Os dendritos apresentam pouco comprimento. O axônio, contudo, pode ser mais longo e em geral, tem calibre uniforme. Os dendritos têm por função receber as informações, ou impulsos nervosos, oriundos de outros neurônios e conduzi-los até o soma. No soma, o impulso é processado, novos impulsos são gerados e então transmitidos aos dendritos de outros neurônios pelo axônio. O ponto de transferência e impulso entre um axônio e os dendritos de outro neurônio é chamado de sinapse e é por meio delas que os neurônios se unem, formando redes neurais. As sinapses funcionam como válvulas, sendo capazes de controlar a transmissão de impulsos entre os neurônios. O efeito das sinapses é variável e esta variação dá ao neurônio a capacidade de adaptação.

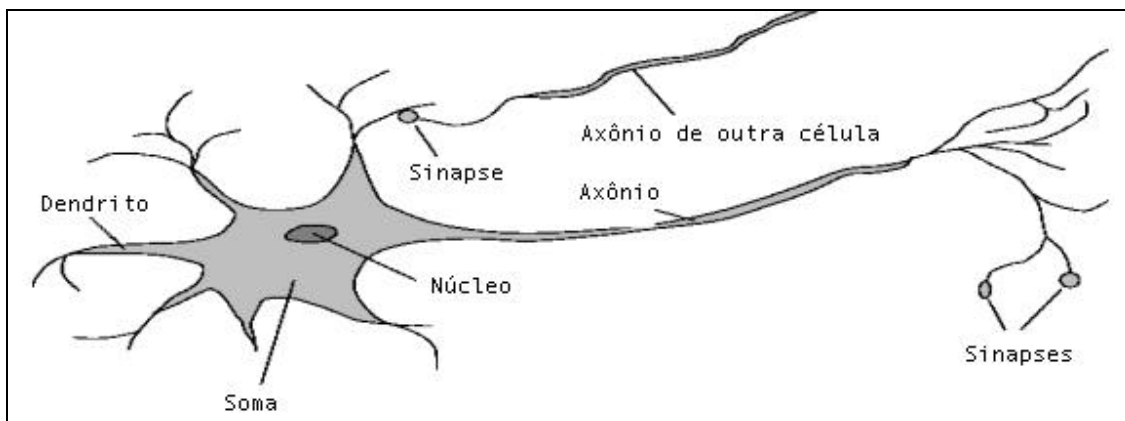


Figura 4.2: Neurônio biológico

Segundo Maria Aparecida Fernandes Almeida [17] a descrição e função de cada elemento componente do neurônio biológico é:

- **Soma (corpo da célula):** O centro dos processos metabólicos da célula nervosa, a partir do qual, projetam-se extensões filamentosas, os dendritos e o axônio.

- **Dendritos:** cobrem um volume muitas vezes maior que o próprio corpo celular e formam uma árvore dendrital, que são “dispositivos de entrada”, que conduzem sinais das extremidades para o corpo celular.
- **Axônio:** é um dispositivo de saída que transmite um sinal do corpo celular para as extremidades. As extremidades do axônio são conectadas com dendritos de outros neurônios, pelas sinapses.
- **Sinapse:** é a ligação entre as terminações axônicas e os dendritos e que permite a propagação dos impulsos nervosos de uma célula para outra.

O funcionamento da rede neural biológica pode ser resumido da seguinte forma: Os sinais oriundos dos nodos (neurônios) pré-sinápticos são passados para o corpo do neurônio (soma), onde são comparados com os outros sinais recebidos pelo corpo do neurônio. Se a soma destes sinais em um intervalo curto de tempo é suficientemente alta, a célula “dispara”, produzindo um impulso que é transmitido para as células seguintes (nodos pós-sinápticos). Esse processo simples é responsável pela maioria das funções realizadas pelo cérebro humano, sendo que a capacidade de realizar tarefas complexas vem da operação em paralelo de milhares e milhares de neurônios que compõem uma rede neural biológica.

4.4 Redes Neurais Artificiais

Uma definição possível de Redes Neurais Artificiais consiste em encará-las como “uma estrutura de processamento de informação distribuída e paralela. Ela é formada por unidades de processamento, comumente chamadas de nós, neurônios ou células, interconectadas por arcos unidirecionais, também chamados de ligações, conexões ou sinapses. Os nós possuem memória local e podem realizar operações de processamento de informação localizada. Cada célula possui uma única saída (axônio), a qual pode se ramificar em muitas ligações

colaterais (cada ramificação possuindo o mesmo sinal de saída do neurônio). Todo o processamento que se realiza em cada unidade deve ser completamente local, isto é, deve depender apenas dos valores correntes dos sinais de entrada que chegam dos neurônios através das conexões. Estes valores atuam sobre os valores armazenados na memória local da célula.”[18]

Os principais elementos utilizados na descrição de Redes Neurais Artificiais são a representação distribuída, as operações locais e o processamento não-linear[21]. Esses atributos especificam duas aplicações básicas das RNA's: situações onde poucas decisões têm que ser tomadas a partir de uma grande quantidade de dados e situações onde um complexo mapeamento não-linear deve se aprendido.

A seguir neste capítulo será apresentado um histórico bem como serão descritos os aspectos básicos das Redes Neurais Artificiais.

4.5 Breve Histórico das Redes Neurais Artificiais

Os primeiros conceitos de máquinas baseadas no comportamento das células nervosas biológicas surgiram no início da década de 40. Em 1943, Warren McCulloch e Walter Pitts apresentaram a primeira discussão sofisticada sobre “neuro-logical network” [19]. Em 1947 eles publicaram o segundo estudo, intitulado “How we Know universals” [34]. A partir destas idéias, surgiu o movimento intelectual chamado “Cybernetics”, o qual tentou combinar conceitos de biologia, psicologia, engenharia e matemática. A década terminou com a publicação do livro de Donald Hebb “The Organization of Behavior” [35].

Após a era “Cybernetics” ter aberto a perspectiva de se construir máquinas baseadas no conhecimento, surgiram, no início da década de 50, os primeiros protótipos neste campo, com arquiteturas específicas para desempenhar tarefas determinadas. Contudo, tendo em vista o fato de que os animais podiam aprender

a fazer muitas coisas para as quais eles não tinham sido “construídos” para fazer, passou-se a ter como meta a construção de máquinas que fossem capazes de aprender. Com isso, a década de 50 presenciou o surgimento de muitos sistemas que aprendiam baseados em reforço, isto é, baseados na idéia de reforçar ações que tenham sido feitas no passado com sucesso, e em formas simples de aprendizado, sendo que o primeiro deles, provavelmente foi feito por Minsky em 1951. No final dos anos 50, o campo de pesquisa em redes neurais ficou virtualmente dormente, sem nenhuma descoberta de peso na área, enquanto que importantes avanços tinham sido alcançados em sistemas especialistas.

O interesse nas redes neurais ressurgiu em 1962 com a publicação do livro de Frank Rosenblatt “Principles of Neurodynamics” [36]. Neste livro Rosenblatt definiu máquinas denominadas de “perceptrons”, e provou muitas teorias sobre as mesmas. Uma onda de entusiasmo dominou a área e muitos pesquisadores passaram a crer que os “perceptrons” seriam a base para uma inteligência artificial. Contudo, esta ilusão foi logo desaparecendo, pois constatou-se que essas redes falhavam na solução de alguns problemas simples, similares a outros em que elas tinham sucesso. O rápido entusiasmo foi sepultado em 1969 com o livro de Marvin Minsky e Seymour Papert chamado “Perceptrons” [37], onde eles provaram que as redes neurais de uma única camada de Rosenblatt eram incapazes de resolver alguns problemas elementares como o do ou-exclusivo. O prestígio e o brilhantismo de Minsky deram grande credibilidade ao seu livro, levando à quase totalidade da comunidade científica a abandonar as pesquisas na área, por um longo período.

A década de 70 e o início da de 80 foram marcadas por um silêncio quase geral na área. Alguns poucos pesquisadores como Grossberg, Kohonen, Anderson, Hopfield, Marr, Von der Malsburg e Copper continuaram pesquisando e desenvolveram trabalhos interessantes, principalmente nas áreas de memória associativa endereçável pelo conteúdo, sistema visual e reformulações do problema de aprendizado das redes. Em 1982, Hopfield deu uma importante contribuição com a introdução da função energia e pela ênfase da noção de

memórias como atratores dinamicamente estáveis. Entretanto, talvez a maior influência desta época ainda tenha vindo dos “perceptrons” [36] de Rosenblatt pois, baseados em suas teorias, vários pesquisadores foram desenvolvendo algoritmos para o ajuste dos pesos das redes neurais de múltiplas camadas, culminando com o surgimento do algoritmo de “*Backpropagation*” (retro-propagação). O “*Backpropagation*” parece ter sido primeiramente proposto por Werbos em 1974 [38] e independentemente redescoberto por volta de 1985 por Rumelhart, Hinton e Williams [39], tendo Parker e Le Cun também propostos algoritmos semelhantes.

Após a invenção do algoritmo “*Backpropagation*” o interesse na área voltou a crescer, desta vez de forma mais madura e promissora. Nos últimos anos, toda a teoria de redes neurais tem sido transformada em aplicações, sendo que novas corporações dedicadas à comercialização desta tecnologia têm aparecido. O crescimento nas atividades de pesquisa tem sido astronômico e novas áreas de atuação das redes neurais têm surgido, tais como: compressão de imagens e voz, reconhecimento de padrões, diagnósticos médicos, previsão de séries temporais, entre outros.

4.6 Neurônio Artificial

A unidade de processamento, também chamada de neurônio artificial ou simplesmente neurônio, é o componente básico das Redes Neurais Artificiais e tem por objetivo simular o funcionamento de um neurônio biológico. A figura a seguir (fig. 4.3) mostra o modelo de neurônio artificial proposto por McCulloch e Pitts [19].

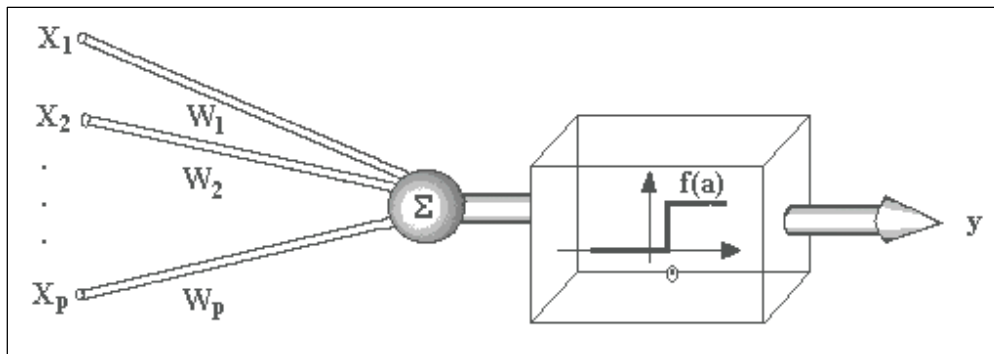


Figura 4.3: Neurônio artificial

O modelo de neurônio proposto por MCCulloch e Pitts (MCP) possui p terminais de entrada x_1, x_2, \dots, x_p , representando os dendritos do neurônio biológico, e um terminal de saída y , que representa o axônio do neurônio biológico. Para emular o comportamento das sinapses, os terminais de entrada do neurônio possuem pesos acoplados w_1, w_2, \dots, w_p , cujos valores podem ser positivos ou negativos, dependendo das sinapses correspondentes serem inibitórias ou excitatórias. O efeito de uma sinapse particular i no neurônio pós-sináptico é dado por: $x_i w_i$. Os pesos determinam o grau em que os neurônios devem considerar sinais de disparo que ocorrem naquela conexão.

O neurônio dispara quando a soma dos impulsos que ele recebe ultrapassa o seu limiar de excitação (threshold). O corpo do neurônio, por sua vez, é emulado por um mecanismo simples que faz a soma dos valores $x_i w_i$ recebidos por ele, e decide se o neurônio deve ou não disparar comparando a soma obtida ao limiar do neurônio. No modelo MCP, a ativação do neurônio é obtida através da aplicação de uma função de transferência, também chamada de regra ou função de ativação e vista mais detalhadamente a seguir, que ativa ou não a saída dependendo do valor da soma ponderada das entradas.

4.7 Funções de Ativação

A função de ativação f define a saída de um neurônio de acordo com o nível de atividade da sua entrada. Existem três tipos básicos de função de ativação [13]

4.1.1 Função Limiar

Para este tipo de função de ativação, mostrada na figura 4.4(a), tem-se que:

$$f(x) = \begin{cases} 1 & \text{se } x \geq 0 \\ 0 & \text{se } x < 0 \end{cases}$$

Equação 4.1

Correspondentemente, a saída do neurônio k que utiliza a função limiar é expressa como:

$$f_k = \begin{cases} 1 & \text{se } v_k \geq 0 \\ 0 & \text{se } v_k < 0 \end{cases}$$

Equação 4.2

Onde v_k é o nível de atividade interna do neurônio; isto é,

$$v_k = \sum_{j=1}^p w_{kj} x_j$$

Equação 4.3

4.1.2 Função Linear

No caso desta função, representada graficamente na figura 4.4(b), a saída do neurônio k é dada por:

$$f_k = c(v_k)$$

Equação 4.4

Onde v_k é o nível de atividade interna do neurônio e c é uma constante.

4.1.3 Função Não Linear

No caso de se querer imitar com maior precisão os neurônios biológicos, use-se funções de ativação não lineares. Elas caracterizam-se por não deixarem a saída exceder certos valores limites, os quais, em geral, são menores que os valores de v_k . A função não linear mais usada é a logística ou sigmóide representada graficamente na figura 4.4(c), e definida por:

$$f(v_k) = \frac{1}{1 + \exp(-av_k)}$$

Equação 4.5

Onde a é o parâmetro de inclinação da função sigmóide. Variando este parâmetro, são obtidas funções sigmóides de diferentes inclinações.

4.1.4 Algumas Variações

As funções de ativação definidas nas equações 4.2 e 4.5 estão na faixa de 0 a 1. Algumas vezes é desejável ter a faixa da função de intervalo -1 a 1 , caso em que a função de ativação assume uma forma anti-simétrica com respeito à origem. Especificamente, a função limiar da equação 4.2 pode ser redefinida como na equação 4.6.

$$f(v_k) = \begin{cases} 1 & \text{se } v_k > 0 \\ 0 & \text{se } v_k = 0 \\ -1 & \text{se } v_k < 0 \end{cases}$$

Equação 4.6

No caso da função sigmóide, a equação 4.5 pode ser substituída pela função *tangente hiperbólica*, representada graficamente na figura 4.4(d), e definida como:

$$f(v_k) = \operatorname{tg}\left(\frac{v_k}{2}\right) = \frac{1 - \exp(-v_k)}{1 + \exp(-v_k)}$$

Equação 4.7

A figura (fig. 4.4) a seguir apresenta algumas das principais funções de ativação utilizadas em redes neurais artificiais e citadas neste item.

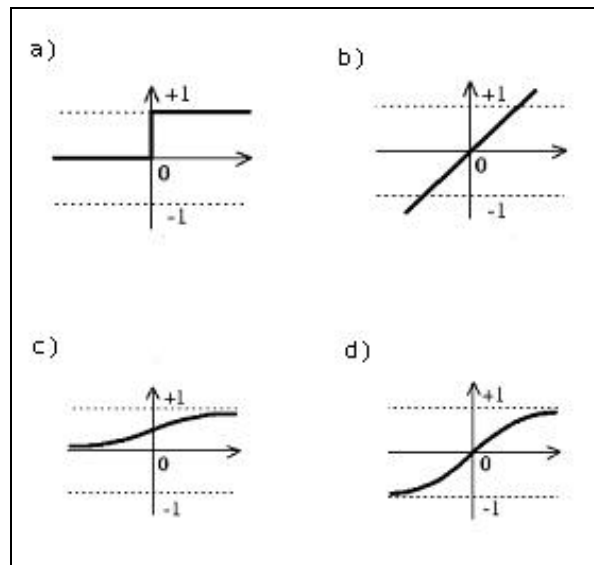


Figura 4.4: Principais funções de ativação

4.8 Topologia das Redes Neurais Artificiais

As várias topologias de redes neurais estão divididas, basicamente, em duas classes:

- *Não-Recorrentes*
- *Recorrentes*

As RNA's *não-recorrentes* são aquelas que não possuem realimentação de suas entradas. Sua estrutura pode possuir uma (RNA de camada única) ou mais camadas (RNA de múltiplas camadas). Redes neurais de múltiplas camadas contêm um conjunto de neurônios de entrada, uma camada de saída e uma ou mais camadas intermediárias. A entrada não é considerada uma camada da rede pelo fato de apenas distribuir os padrões de entrada aos neurônios da primeira camada da rede. A camada que contém os neurônios que fornecem a saída da rede é chamada de camada de saída. As camadas que não possuem ligações diretas nem com a entrada, nem com a saída são chamadas de camadas

intermediárias. Não existem conexões ligando um neurônio de uma camada a outro de uma camada prévia, nem a um neurônio da mesma camada.

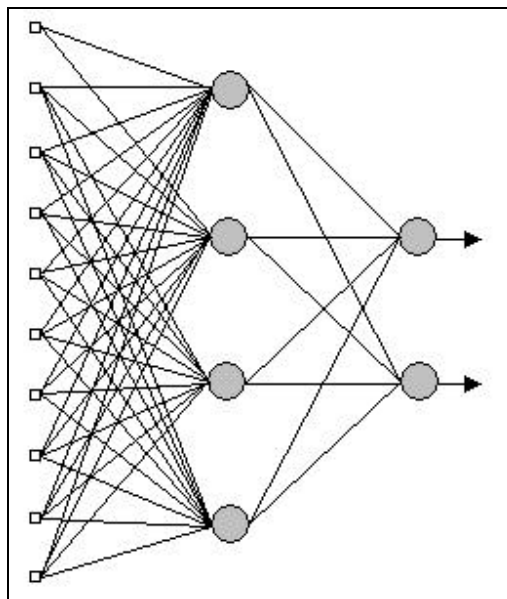


Figura 4.5: Rede neural com múltiplas camadas não-recorrente

As RNA's de uma camada, também chamadas de "*perceptrons*", são abordadas com mais detalhes no item 4.9.1 e possuem somente uma camada de neurônios entre qualquer entrada e qualquer saída.

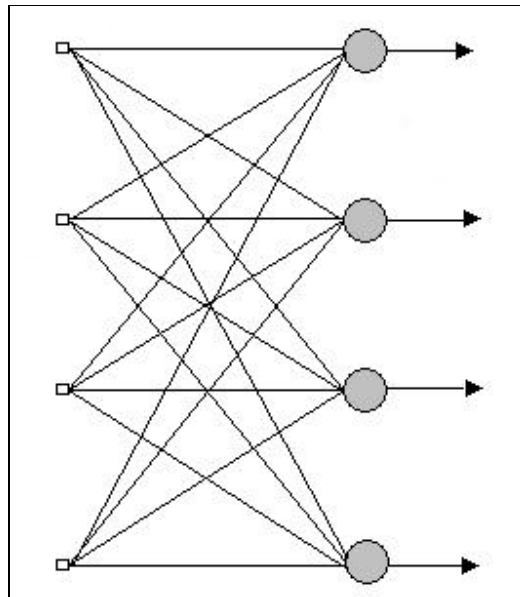


Figura 4.6: Rede neural não-recorrente de uma única camada

As RNA's *recorrentes* são redes que contém realimentação das saídas para as entradas, sendo suas saídas determinadas pelas entradas atuais e pelas saídas anteriores. Além disso, suas estruturas não são obrigatoriamente organizadas em camadas e quando são, estas redes podem possuir interligações entre neurônios da mesma camada e entre camadas não consecutivas, gerando interconexões bem mais complexas do que as RNA's não-recorrentes.

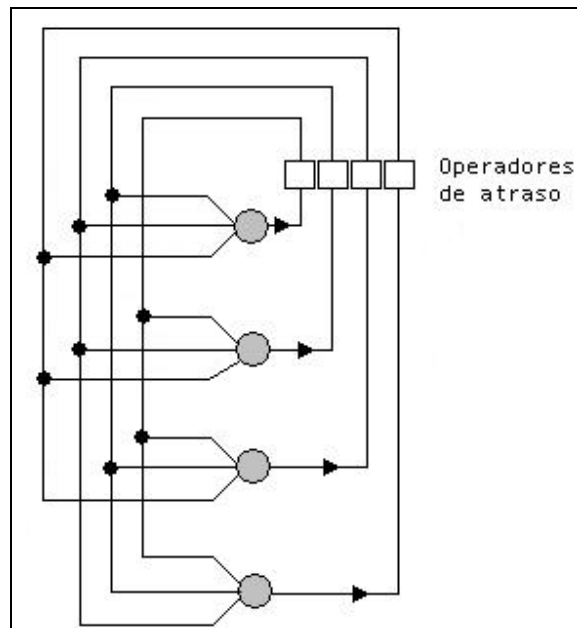


Figura 4.7: Rede neural recorrente de uma única camada

4.9 Principais modelos de Redes Neurais Artificiais

Existem muitos modelos de Redes Neurais Artificiais, e uma quantidade razoável de publicações que se dedicam a classificá-los[24][21][26]. Cinco destes modelos são apresentados a seguir juntamente com os pesquisadores que os desenvolveram.

4.9.1 Perceptron

Proposto por F. Rosenblatt em 1957, o *perceptron* de uma única camada foi o primeiro modelo de RNA desenvolvido. Suas principais características são:

- Classe de tarefas: reconhecimento de padrões (ex.: caracteres impressos);

- Propriedades dos neurônios: possuem entradas binárias e saídas que assumem os valores +1 ou -1. A função de ativação é a função degrau;
- Topologia da rede: rede não recorrente de uma única camada;
- Aprendizado: Utiliza a técnica de reforço.

Esta rede gerou muito interesse pela habilidade de aprender a reconhecer padrões linearmente separáveis. Contudo, como a grande maioria dos problemas práticos relevantes não são linearmente separáveis, o *perceptron* não tem uso generalizado[21].

4.9.2 *Perceptron* de múltiplas camadas (*Backpropagation*)

Proposto por P. Werbos, D. Parker e D.Rumelhart em 1974, são redes acíclicas, ou seja, não-recorrentes, com uma ou mais camadas de neurônios intermediários, também chamadas de camadas intermediárias, entre a camada de entrada e a camada de saída. Um algoritmo capaz de treinar os *perceptrons* de múltiplas camadas é o *backpropagation*. Suas principais características são:

- Classe de tarefas: reconhecimento de padrões (ex.: processamento de fala)
- Propriedades dos neurônios: possuem entradas com valores contínuos e função de ativação não linear, geralmente a função sigmóide;
- Topologia da rede: rede não recorrente com mais de uma camada;
- Aprendizado: Utiliza a técnica de Correção de Erros com uso da regra delta generalizada.

O algoritmo *backpropagation* foi testado em uma série de problemas clássicos, e em problemas relacionados com reconhecimento de padrões visuais.

Na maioria dos casos, ele encontrou boas soluções para os problemas propostos, apesar do algoritmo, as vezes, fornecer uma configuração de pesos correspondente a um mínimo local da função erro. Isto se dá devido ao fato do *backpropagation* utilizar o método do gradiente e, a princípio, a superfície de erro possuir uma forma qualquer.

A solução proposta neste trabalho, faz uso de redes neurais tipo *perceptron* de múltiplas camadas e uma abordagem mais específica a respeito desta topologia é apresentada no item 4.11 deste capítulo.

4.9.3 Classificador de Carpenter-Grossberg

No final da década de setenta, G.Carpenter e S.Grossberg projetaram uma rede capaz de formar aglomerados de informações (clusters), e de ser treinada sem supervisão.

- Classe de tarefas: reconhecimento de padrões (ex.: reconhecimento de sinais de radar ou sonar)
- Propriedades dos neurônios: possuem entradas binárias, podendo assumir valores contínuos. Sua função de ativação é a sigmóide;
- Topologia da rede: rede cíclica de três camadas;
- Aprendizado: Utiliza a técnica de aprendizado competitivo e Cooperativo que foi introduzida pelo próprio Grossberg.

Inicialmente o algoritmo principal seleciona a primeira entrada como um exemplo para o primeiro aglomerado. A entrada seguinte é comparada com este primeiro exemplo. Ela é agrupada com o mesmo, se a distância para o primeiro for menor que um certo limite, chamado de limiar de vigilância. Caso contrário, este exemplo formará um novo aglomerado. Este processo se repete para todas as entradas existentes. O número de aglomerados cresce em função do limiar e da

métrica da distância usada para comparar os exemplos de entrada dos aglomerados.

4.9.4 Rede de Kohonen

No começo da década de oitenta, T.Kohonen propôs uma rede onde se corroborou os estudos teóricos sobre a organização dos caminhos de sensoriamento na mente. Suas principais características são:

- Classe de tarefas: reconhecimento de padrões (ex.: reconhecimento de fala) e aprendizado da distribuição de probabilidades dos dados (ex.: auto-organização de mapas de características);
- Propriedades dos neurônios: possuem entradas com valores contínuos e função de ativação sigmóide;
- Topologia da rede: rede cíclica de duas camadas;
- Aprendizado: Utiliza a técnica de aprendizado competitivo, que foi introduzida pelo próprio Kohonen.

Segundo esta teoria, o cérebro humano foi considerado como uma coleção estruturada de neurônios. Com isto foi admitida uma ordem espacial das unidades de processamento que permitiu elaborar uma rede neural dotada de mecanismos que permitem formar representações estruturadas dos estímulos de entrada. Após o aprendizado, as unidades respondem a diferentes estímulos de maneira ordenada, formando um sistema de coordenadas de características sobre a rede.

4.9.5 Rede de Hopfield

Também na década de oitenta, o trabalho de J.Hopfield contribuiu substancialmente para o ressurgimento das pesquisas em RNAs. A rede de Hopfield são mais apropriadas quando representações binárias permitem modelar a situação desejada. Por exemplo, imagens em preto e branco, onde os elementos de entrada podem ser representados pelos valores de cada ponto da imagem (0 = branco, 1 = preto). Suas características são:

- Classe de tarefas: reconhecimento de padrões (ex.: reconhecimento de dados ou imagens completas a partir de fragmentos) e memória associativa;
- Propriedades dos neurônios: possuem entradas binárias e saídas que assumem valores +1 ou -1. A função de ativação é a sigmóide;
- Topologia da rede: rede cíclica de uma camada;
- Aprendizado: os padrões são armazenados no começo.

A Rede de Hopfield possui duas limitações quando usada como memória de acesso por conteúdo. Primeiro, apesar dos padrões armazenados, a rede pode convergir para um novo padrão diferente dos padrões-exemplo existentes. Isto pode produzir uma situação em que a rede não casa com um padrão já existente. Uma segunda limitação é que o padrão exemplo será considerado instável se ele compartilhar muitos bits com outros padrões exemplo, o que pode ocasionar uma convergência da rede para este outro exemplo.

4.10 Aprendizado em Redes Neurais Artificiais

Nesta seção é feita uma breve análise das técnicas de treinamento e aprendizado utilizadas nas RNAs e feitas algumas considerações sobre os

principais modelos encontrados na literatura, procurando destacar as suas propriedades mais relevantes.

4.10.1 Diferentes Tipos de Aprendizado

Na sua essência, o conceito de aprendizado envolve mudança associada a aperfeiçoamento[21]. Carbonell [22] define o conceito de aprendizado, dentro do campo da inteligência artificial, como a habilidade de realizar tarefas novas que não podiam ser realizadas anteriormente, ou melhorar a realização de tarefas antigas, como resultado de mudanças produzidas pelo processo de aprendizado.

Uma classificação, apresentada por Lippman [24], relacionou as redes quanto ao tipo de controle realizado durante o aprendizado (supervisionado ou não-supervisionado). No aprendizado supervisionado existem, basicamente, rótulos que especificam a classe correta para os padrões de entrada, enquanto que no aprendizado não-supervisionado não existem tais rótulos.

Diversas técnicas de aprendizado podem ser utilizadas nas RNAs, entre elas destacam-se:

- Correção de Erros: aprendizado supervisionado que ajusta os pesos das conexões entre nós, na proporção da diferença entre os valores desejados e computados de cada neurônio da camada de saída.
- Reforço: aprendizado supervisionado onde os pesos são recompensados quando o sistema executa ações apropriadas e punidos caso contrário;
- Regra de Hebb: aprendizado onde o ajuste dos pesos das conexões é realizado em função da relação de valores dos dois neurônios que ela conecta. Pode ser aplicado tanto ao aprendizado supervisionado quanto ao aprendizado não-supervisionado.

4.11 *Perceptron* de Múltiplas Camadas

As arquiteturas do tipo *perceptron* de múltiplas camadas (Multilayer Perceptron MLP) constituem os modelos neurais artificiais mais utilizados e conhecidos atualmente. Tipicamente, esta arquitetura consiste de um conjunto de unidades sensoriais que formam uma camada de entrada, uma ou mais camadas intermediárias (ou escondidas) de unidades computacionais e uma camada de saída. Os sinais de entrada são propagados camada a camada da entrada para a saída. Esta arquitetura representa uma generalização do *perceptron* apresentado anteriormente no item 4.9.1 deste capítulo.

As redes do tipo MLP têm sido utilizadas com sucesso para a solução de vários problemas envolvendo altos graus de não-linearidade. Seu treinamento é do tipo supervisionado e utiliza, como já foi dito, um algoritmo muito popular chamado retro-propagação do erro (*error backpropagation*). Este algoritmo é baseado numa regra de aprendizagem que “corrige” o erro durante o treinamento [13].

Basicamente, o processo de retro-propagação do erro é constituído de duas fases: uma fase de propagação do sinal funcional (*feedforward*) e uma de retro-propagação do erro (*backpropagation*). Na fase de propagação, os vetores de dados são aplicados às unidades de entrada, e seu efeito se propaga pela rede, camada a camada. Finalmente, um conjunto de saídas é produzido como resposta da rede. Durante esta fase, os pesos das conexões são mantidos fixos. Na retro-propagação do erro, por outro lado. Os pesos são ajustados de acordo com uma regra de correção do erro. Especificamente, a resposta da rede em um instante de tempo é subtraída da saída desejada para produzir um sinal de erro. Este sinal de erro é retro-propagado da saída para a entrada, camada a camada, originando o nome “retro-propagação do erro”. Os pesos são ajustados de forma que a “distância” entre a resposta da rede e a resposta desejada seja reduzida.

A figura 4.8 apresenta uma arquitetura do tipo MLP com duas camadas intermediárias. A rede apresentada aqui possui todas as conexões, o que significa que um neurônio em qualquer camada da rede está conectado a todos os outros neurônios na camada anterior. O fluxo de sinal através da rede é feito positivamente, da esquerda para a direita, camada a camada.

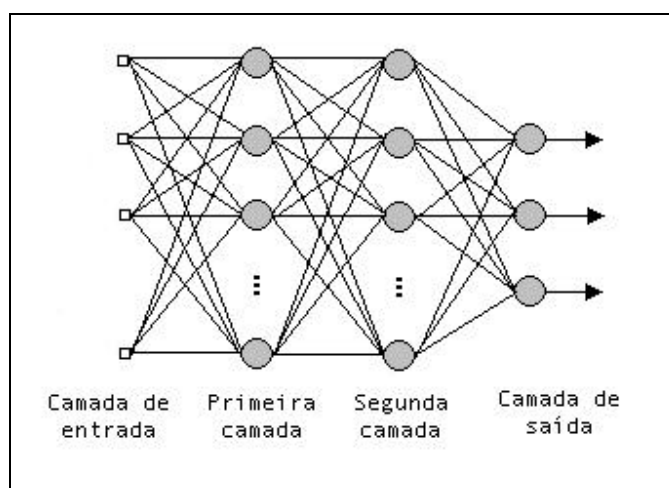


Figura 4.8: Arquitetura MLP com duas camadas intermediárias

A figura 4.9 mostra apenas uma parte da rede e nela podem ser identificados os dois tipos de sinais.

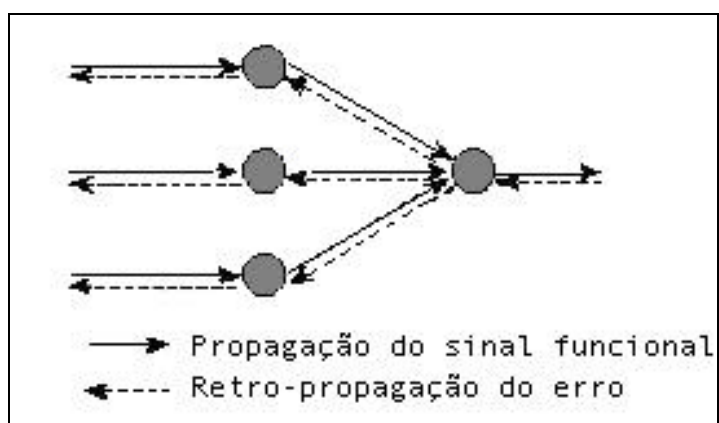


Figura 4.9: Tipos de sinais em uma MLP

- *Sinal funcional*: um sinal funcional é um sinal de entrada (estímulo) que chega na entrada e é propagado positivamente (neurônio a neurônio) através da rede, e aparece na saída como um sinal de saída.
- *Sinal de erro*: Os sinais de erro originam-se nas saídas e são retro-propagados (neurônio a neurônio) através da rede.

A camada de entrada geralmente é composta por neurônios sensoriais, ou seja, unidades que não modificam os sinais externos, apenas os distribuem para a primeira camada intermediária. As unidades de saída constituem a camada de saída da rede e as demais unidades constituem as camadas intermediárias. As camadas intermediárias são todas aquelas que não fazem parte da entrada e nem da saída.

Cada unidade (neurônio) intermediária ou de saída é responsável por duas tarefas:

- Calcular o sinal na saída da unidade, que geralmente é expresso como uma função não-linear do sinal de entrada e pesos sinápticos associados; e
- Calcular uma estimativa instantânea do vetor gradiente, que é necessário para a retro-propagação do erro através da rede.

4.12 O Algoritmo *Backpropagation*

Como já foi visto no item 4.9.2 deste trabalho, a regra de aprendizado utilizada para treinar uma rede MLP é denominada Regra Delta Generalizada ou, mais comumente, *backpropagation*. O algoritmo *backpropagation* é primordial para muitos trabalhos atuais sobre aprendizado em redes neurais [26].

O treinamento usando esta regra consiste em fornecer à rede um conjunto de pares de entradas e saídas (aprendizado supervisionado), onde para cada entrada do treinamento tem-se uma saída desejada. Este algoritmo é um método de gradiente descendente, que não garante chegar ao mínimo erro global, e que pode ser dividido em 5 passos:

- **Passo 1:** Apresente um padrão de entrada e a saída desejada.

Utilizando uma determinada estratégia de apresentação coloque um dos padrões de entrada na camada de entrada e saída desejada nas suas respectivas camadas, então ative o passo 2 (*calcule saída*)

- **Passo 2:** Calcule saída

A partir da primeira camada, permita que cada camada produza os valores de saída até atingir a camada de saída da rede e, então ative o passo 3 (*ajuste dos pesos da camada de saída*).

- **Passo 3:** Ajuste dos pesos da camada de saída

Para cada neurônio j da camada de saída, atualize todos os pesos w_{ij} conforme equação 4.8, então ative o passo 4 (*ajuste de pesos das camadas intermediárias*).

$$\Delta w_{ij} = h d_j o_i$$

Equação 4.8

Onde:

h : é uma constante positiva que determina a taxa de aprendizado, também chamada simplesmente “taxa de aprendizado”.

o_i : saída do neurônio

d_j : diferença (erro) entre a saída computada e a saída desejada do neurônio j , que pode ser calculada segundo a equação 4.9

$$d_j = o_j(d_j - o_j)(1 - o_j)$$

Equação 4.9

Onde:

d_j : saída desejada do neurônio j

- **Passo 4:** Ajuste de pesos das camadas intermediárias

Para ajustar os pesos de todas as camadas intermediárias, atualize o peso w_{ki} de um neurônio i em uma camada intermediária que está ligado a outro neurônio k na camada anterior conforme equação 4.10

$$\Delta w_{ki} = h d / i o_k$$

Equação 4.10

Onde:

h : constante positiva que determina a taxa de aprendizado, também chamada simplesmente “taxa de aprendizado.

d / i : erro relativo do neurônio i , dado pela equação 4.11

$$d / i = o_i(1 - o_i) \sum w_{ij} d_j$$

Equação 4.11

o_k : saída do neurônio k , que estimula i via w_{ki}

Depois que todas as conexões tenham sido ajustadas ative o passo 5 (*cheque magnitude do erro*).

- **Passo 5:** Cheque magnitude do erro

Para se checar a magnitude do erro, pode-se adotar várias estratégias diferenciadas quanto à forma de aceitar o erro como desprezível. Uma regra muito comum é verificar se o erro global de saída da rede em relação à saída desejada é menor que um dado erro predefinido. Se a condição adotada é satisfeita, então a rede aprendeu o conjunto de treinamento, caso contrário volte a ativar o passo 1 (*apresente um padrão de entrada e a saída desejada*).

- **Fim do algoritmo**

Este algoritmo pode ser utilizado em uma arquitetura com qualquer número de camadas. O passo 4 deve ser ativado recursivamente até atingir a camada de entrada. Para se aumentar a velocidade de convergência é possível introduzir na equação 4.9 um multiplicador chamado de *momentum*, onde este valor representa o efeito dos ajustes anteriores no ajuste atual.

4.12.1 Problemas no Aprendizado

O maior problema da utilização do algoritmo de retro-propagação ocorre quando a rede neural obtém uma solução que corresponde a um mínimo local da função do erro. Existem diversas alternativas que visam diminuir a ocorrência de mínimo local. As mais utilizadas são:

- Diminuição da taxa de aprendizado: se a taxa de alteração dos pesos é diminuída progressivamente, então o algoritmo do gradiente decrescente está possivelmente mais apto a alcançar uma solução melhor[28];

- Adição de unidades internas: é possível considerar que quando um mínimo local é apresentado como solução, duas classes diferentes são classificadas como a mesma. Para evitar isso, ao serem colocadas mais unidades na camada intermediária, a rede fará um reconhecimento melhor, diminuindo a ocorrência de mínimo local.
- Introdução de um termo de momento: termo introduzido para que se possa aumentar a taxa de aprendizado sem que ocorra oscilação entorno dos valores de saída desejados:

Outro problema enfrentado pela retro-propagação é a quebra de simetria, ou seja, se os pesos começam todos com o mesmo valor e se a solução só pode ser alcançada com pesos diferentes, o sistema não poderá aprender[29]. Isto ocorre porque o erro é retro-propagado em proporção ao valor dos pesos, o que significa que todas as unidades intermediárias conectadas diretamente às unidades de entrada e saída cometem erros iguais. Como a modificação dos pesos depende deste erro, os pesos destas unidades em relação às unidades de saída serão sempre os mesmos.

4.13 Conclusão

Esse capítulo traz um breve referencial teórico a respeito das Redes Neurais Artificiais e das suas inspirações biológicas. Foram abordados aqui os conceitos necessários para a compreensão do método proposto neste trabalho, sendo que entre eles, podemos destacar o tópico referente às principais funções de ativação utilizadas nos neurônios artificiais, e principalmente, os tópicos referentes ao *perceptron* de múltiplas camadas (4.11) e ao algoritmo de aprendizado *backpropagation* (4.12).

Capítulo 5: UM MÉTODO PARA SEGMENTAÇÃO DA PLACA EM IMAGENS DIGITAIS DE VEÍCULOS

5.1 Introdução

Com base nos conceitos abordados até aqui, este capítulo apresenta um método que, fazendo uso de redes neurais artificiais, propõe-se a fornecer uma solução para o problema de segmentação da área referente à placa de um veículo em uma imagem digital obtida por controladores eletrônicos de velocidade.

Este capítulo traz inicialmente uma descrição dos equipamentos e programas utilizados no desenvolvimento deste trabalho. Traz ainda, breve tópico referente a algumas funções do MATLAB[®], principal ferramenta utilizada no desenvolvimento e teste do método ora proposto, e que possui uma grande quantidade de funções específicas ao tratamento e manipulação de imagens digitais e ao desenvolvimento, testes e avaliações de redes neurais artificiais.

O método aqui proposto se divide basicamente em três etapas. A primeira delas compreende a tarefa de preparação dos dados e a criação de conjuntos de treinamento e testes que serão utilizados posteriormente no desenvolvimento das redes neurais aqui propostas. A segunda fase corresponde a concepção e treinamento de uma rede neural do tipo *perceptron* de múltiplas camadas capaz de identificar quais as linhas da imagem que contém a placa do veículo, bem como é apresentado um método utilizado para binarizar a linha da imagem com relação ao valor médio da intensidade de cinza da mesma. A terceira e última fase do método corresponde a concepção e treinamento de uma segunda rede neural, capaz de identificar, entre o conjunto de linhas resultantes da classificação efetuada na segunda fase, quais colunas contém a placa do veículo.

Neste capítulo se apresenta os aspectos referentes à concepção e ao treinamento de cada uma das redes propostas, sendo que os testes e avaliações serão descritos no capítulo seguinte.

5.2 Equipamentos e Programas Utilizados

Todas as tarefas descritas neste capítulo foram executadas em um microcomputador AMD ATHLON 1.4 Ghz, com 256 Mb de memória Ram, disco rígido de 20 Gb e placa de vídeo de 2 Mb.

Os programas utilizados no desenvolvimento deste trabalho foram o sistema operacional Windows® XP e o MATLAB® 6.0 com suas respectivas *toolboxes* de redes neurais e processamento de imagens.

5.3 Principais Funções do MATLAB® Utilizadas Neste Trabalho

Este tópico tem como objetivo apresentar de forma resumida as principais funções do *software* MATLAB® utilizadas neste trabalho. Maiores informações a respeito destas e de quaisquer outras funções podem ser obtidas na documentação do mesmo.

- **newff**: Cria uma rede neural tipo “*feedforward*”, inicializando os pesos das sinapses e os valores de *bias*.

Sintaxe:

$$\text{Net} = \text{newff}(\text{PR}, [\text{S1 S2...SN1}], \{\text{TF1 TF2...TFN1}\}, \text{BTF})$$

Onde:

PR: Matriz $R \times 2$ dos valores mínimos e máximos dos R elementos de entrada;

S_i : Número de neurônios da i -ésima camada, para N_1 camadas;

TF $_i$: Função de ativação da rede, padrão = 'tansig';

BTF: Função de treinamento da rede, padrão = 'traingdx';

- **train**: Treina uma rede neural, atualizando os valores dos pesos e *bias* de acordo com os conjuntos de entrada e saída apresentados como exemplo de treinamento e de acordo com as funções de treinamento, aprendizado e performance estipulados para a mesma.

Sintaxe:

$$\text{Net} = \text{train}(\text{Net}, P, T)$$

Onde:

Net: Objeto rede neural criado pela função **newff** ou outra função de criação;

P: Conjunto de entradas do exemplo de treinamento;

T: Conjunto de saídas do exemplo de treinamento;

- **sim**: Simula uma rede neural apresentando o resultado da mesma para o conjunto de entradas informado.

Sintaxe:

$$Y = \text{sim}(\text{Net}, P)$$

Onde:

Net: Objeto rede neural criado pela função **newff** ou outra função de criação;

P: Conjunto de entradas da rede;

Y: Saída da rede para o conjunto de entradas fornecido como parâmetro.

- **imread:** Carrega uma imagem de um arquivo.

Sintaxe:

$$A = \text{imread}(\text{IM}, \text{TP})$$

Onde:

A: Matriz contendo a imagem lida;

IM: Nome do arquivo de imagem a ser carregado;

TP: Tipo do arquivo de imagem a ser carregado.

- **Figure:** Abre uma nova janela de figura.

Sintaxe:

$$\text{figure}$$

- **Plot:** Traça um gráfico linear dos valores passados como parâmetro na janela de figura ativa.

Sintaxe:

$$\text{plot}(A)$$

Onde:

A: Conjunto de valores que compreendem o gráfico a ser traçado

5.4 Imagens

Foram utilizadas nesse trabalho um total de 225 imagens obtidas por controladores eletrônicos de velocidade, todas mostrando a parte traseira do veículo e obtidas durante o dia, o que lhes garante uma boa condição de iluminação. Tais imagens foram gravadas de forma monocromática utilizando-se o padrão de compressão de imagens JPEG, tratado no item 3.3.2 deste trabalho com resolução espacial de 500x755 *pixels*.

O conjunto total de imagens foi inicialmente dividido em três subconjuntos, sendo dois deles destinados ao treinamento e um destinado a testes das duas redes neurais aqui propostas conforme mostra a tabela a seguir.

Conjunto	nº . de imagens	Utilização
A	100	Treinamento
B	100	Treinamento
C	25	Testes

Tabela 5.1: distribuição das imagens em conjuntos de treinamento e testes.

5.5 Segmentação da Linha

Nesta etapa é apresentado o desenvolvimento de uma rede neural do tipo *perceptron* de múltiplas camadas capaz de identificar as linhas da imagem que contém a placa do veículo. Esta classificação baseia-se no fato de que as linhas que cruzam o local da placa possuem um tipo de “assinatura” que faz usualmente

possível distinguí-las das linhas restantes da imagem [30][43] conforme ilustram as figuras a seguir.

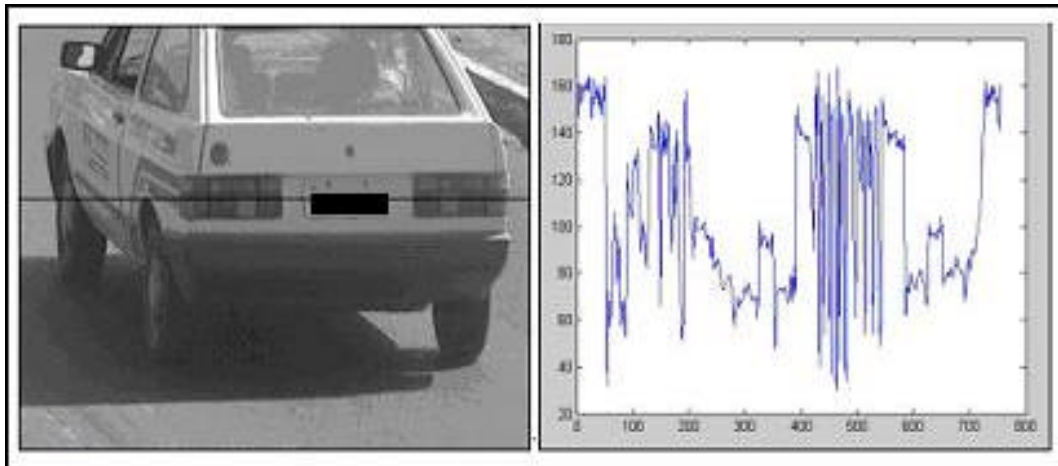


Figura 5.1: Variação tonal dos níveis de cinza em uma Linha com placa.

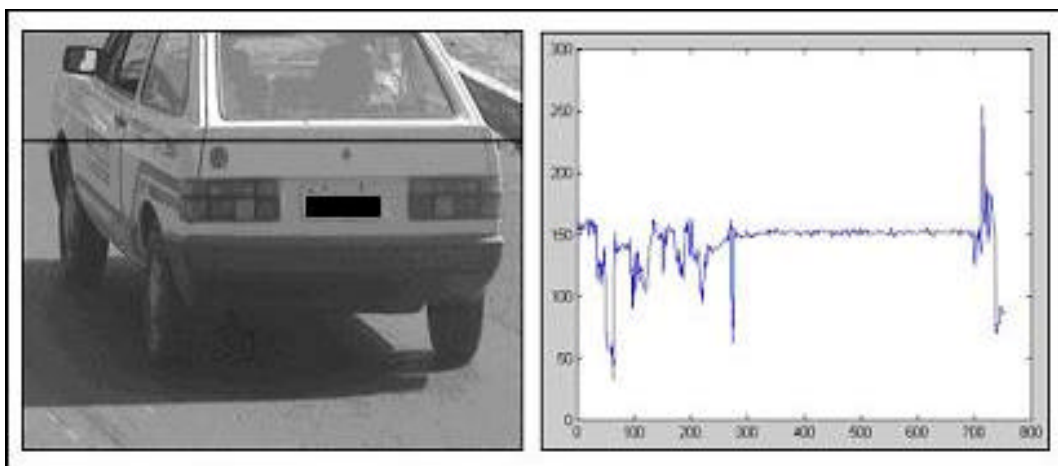


Figura 5.2: Variação tonal dos níveis de cinza de uma linha sem placa.

O principal objetivo dessa etapa é reduzir a área da imagem, mantendo apenas as linhas que apresentam variação tonal característica sendo que o resultado esperado é mostrado na figura a seguir.



Figura 5.3: Resultado esperado da segmentação das linhas da imagem.

5.5.1 Binarização da Linha

Com o objetivo de diminuir a variação dos valores de entrada da rede da escala de 0 a 255 utilizada para representar os valores de cinza da imagem, é realizado um processo de binarização da variação tonal da mesma, reduzindo a entrada da rede para valores de 0 e 1. Para tanto, é primeiramente determinado o valor médio da variação dos níveis de cinza da linha da imagem, através da seguinte equação.

$$Mt = ((Mr - Mn) / 2) + Mn$$

Onde:

Mt : Valor médio dos níveis de cinza.

Mr : Maior valor dos níveis de cinza.

Mn : Menor valor dos níveis de cinza.

Como mostra a figura a seguir.

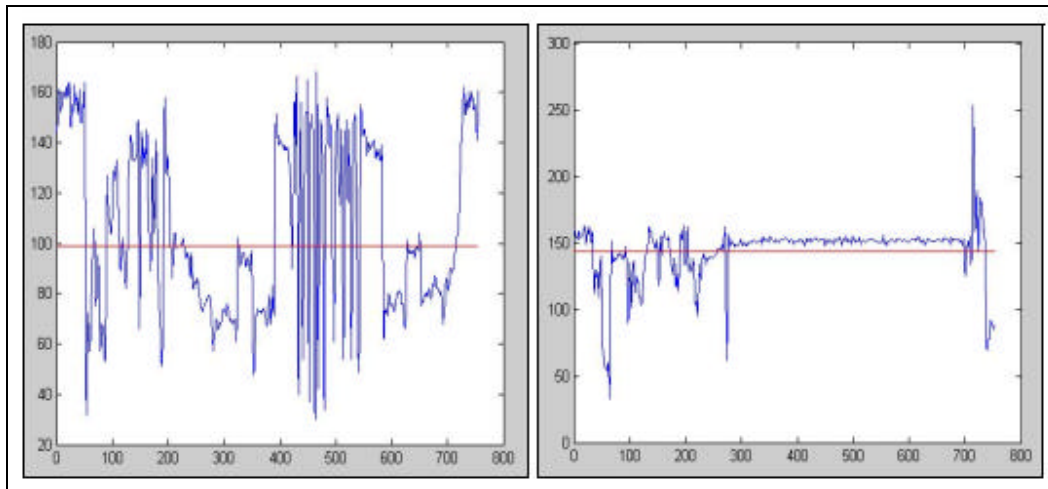


Figura 5.4: Representação do valor médio na variação dos níveis de cinza.

Uma vez encontrado o valor médio da variação tonal da linha, é criado um vetor com 755 posições, correspondente ao número de *pixels* existentes na linha da imagem. Cada vez que a variação dos níveis de cinza, transpõe este valor, é atribuído o valor 1 para a respectiva posição no vetor, e todas as demais recebem o valor 0. Tal processo resulta na informação que será posteriormente submetida à rede neural descrita no item 5.5.3.

As imagens a seguir ilustram o resultado da binarização de uma linha com placa e de uma linha sem placa respectivamente.



Figura 5.5: binarização de linha com placa.

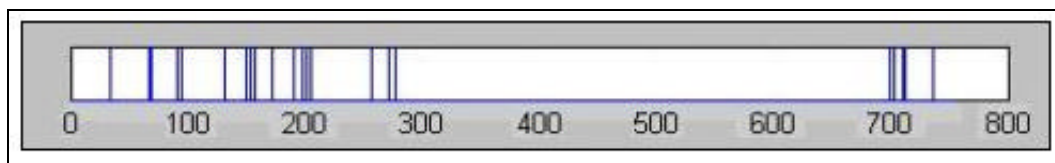


Figura 5.6: Binarização de linha sem placa.

Mais considerações a respeito deste método são apresentadas no capítulo seguinte.

5.5.2 Preparação dos Dados de Treinamento

Para treinamento de uma rede neural capaz de segmentar as linhas que correspondem à placa do veículo, foram preparados inicialmente dois conjuntos de treinamento. O primeiro é composto de linhas retiradas entre as 100 imagens do conjunto de treinamento **A**, contendo 2000 linhas que correspondem à placa e 2000 contendo outras linhas da imagem, formando assim uma matriz de 4000x755 pontos denominado LN1. O segundo é composto de linhas retiradas entre as 100 imagens do conjunto de treinamento **B**, também contendo 2000 linhas que correspondem a placa e 2000 contendo outras linhas da imagem, formando da mesma forma, uma matriz de 4000x755 pontos denominado LN2. Por fim, foi criado um terceiro conjunto de treinamento que corresponde a concatenação dos conjuntos LN1 e LN2, contendo portanto, 8000 exemplos de treinamento, sendo 4000 que correspondem a placa e 4000 que correspondem a outras linhas da imagem.

Conjunto	Exemplos com placa	Exemplos sem placa	Total de exemplos
LN1	2000	2000	4000
LN2	2000	2000	4000
[LN1LN2]	4000	4000	8000

Tabela 5.2: composição dos conjuntos de treinamento utilizados para segmentação das linhas da imagem.

Para cada um dos conjuntos de entrada citados acima foi criado um vetor de saída com 1x4000 posições contendo valores de 0,9 para linhas que contém placa, e 0,1 para as demais linhas do conjunto de treinamento (LN1S, LN2S e LN1LN2S). Optou-se pelos valores 0,1 e 0,9 a fim de evitar os extremos da função de ativação sigmóide utilizada na camada de saída das redes descritas a seguir.

Ambos os conjuntos de treinamento utilizados foram binarizados utilizando-se o método descrito no item 5.4.1 deste capítulo.

5.5.3 Configuração da Rede Neural

A definição da rede neural mais apropriada a determinado problema é essencialmente empírica [13][16]. Inicialmente foram realizados testes utilizando uma topologia simplificada da rede proposta, ou seja, uma rede contendo 125 entradas, o que corresponde ao tamanho padrão (em *pixels*) da placa do veículo nas imagens utilizadas nos conjuntos de treinamento e teste, uma camada intermediária contendo cinco neurônios e uma camada de saída contendo um neurônio que deve apresentar 0,1 para linhas que não contém placa e 0,9 para linhas que contém placa. Uma topologia semelhante foi utilizada no trabalho intitulado “The neural network based method for locating car license plate” [25]. Esse teste inicial apresentou resultados satisfatórios no que se refere a identificar blocos de 125 linhas que contém a variação tonal característica da área da placa,

sendo que seu principal problema consiste na necessidade de um algoritmo complexo para varrer toda a imagem, incrementando linha a linha, *pixel* a *pixel*, a fim de localizar conjuntos contendo as características referentes a área da placa do veículo.

No trabalho intitulado “Detecção de Volume de Tráfego de Veículos Proporcionada por Visão Computacional via Redes Neurais” [31] e posteriormente no trabalho intitulado “Wavelets e Redes Neurais Aplicadas na Medição do Fluxo de Veículos Proporcionado por Visão Computacional” [32] foram usadas topologias de redes as quais recebem um vetor com informações binárias e geram um valor de saída variando entre 0 e 1 que identifica o número de veículos em uma imagem digital. O segundo trabalho citado utiliza uma rede do tipo *perceptron* de múltiplas camadas contendo 717 valores binários de entrada, três camadas de neurônios contendo respectivamente 20, 20 e 1 neurônios e todas as camadas possuindo função de ativação sigmóide. Tal topologia serviu como base para a definição da rede ora apresentada.

Para a tarefa de classificação das linhas que contém ou não placa do veículo, foi concebida uma rede neural do tipo *perceptron* de múltiplas camadas contendo 755 entradas, que correspondem aos valores binarizados de uma linha inteira da imagem, três camadas de neurônios contendo vinte (20) neurônios na primeira, dez (10) neurônios na segunda e um (01) neurônio na terceira, também chamada de camada de saída. Nas duas camadas intermediárias a função de ativação utilizada foi a *tangente hiperbólica* e na camada de saída a função de ativação *sigmóide*. Tais funções evitam a possibilidade de ocorrerem resultados muito esparsos sendo que a função *sigmóide* utilizada na última camada, apresenta apenas valores dentro do intervalo [0,1]. Foi utilizado também um *bias* de valor 1 em cada uma das duas camadas intermediárias e na camada de saída.

A figura 5.7 representa graficamente a topologia da rede neural utilizada para segmentar as linhas da imagem, nela estão representadas: a camada de entrada; a primeira camada contendo vinte (20) neurônios, função de ativação *tangente hiperbólica*, matriz de pesos (LW1) e a matriz de *bias* (b1); a Segunda

camada contendo dez (10) neurônios, função de ativação *tangente hiperbólica*, matriz de pesos (LW2) e a matriz de *bias* (b2); e a camada de saída contendo um (1) neurônio, função de ativação *sigmóide*, matriz de pesos (LW3) e a matriz de *bias* (b3).

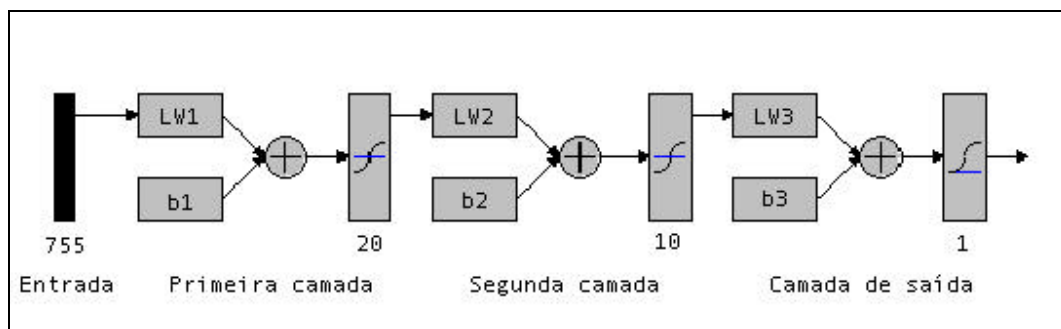


Figura 5.7: Topologia da rede neural utilizada para segmentar as linhas da imagem.

5.5.4 Treinamento da Rede Neural

Para realizar o treinamento da rede neural descrita foi utilizado o método *backpropagation*. Esse método é implementado no MATLAB[®] por meio de algumas rotinas cuja diferença entre elas está na velocidade de aprendizado dos algoritmos por elas empregados. Foi utilizado o método *traingdm* que possibilita a introdução do termo *momentum* no processo de ajuste dos pesos da rede. O termo *momentum* diminui a sensibilidade do método a pequenos detalhes na superfície de erro, ajudando a rede a evitar mínimos locais que poderiam impedi-la de encontrar um ponto de erro mínimo global.

O treinamento da rede por esse método ocorre em etapas denominadas épocas. Essas épocas consistem em se aplicar todos os dados de entrada de treinamento na rede, verificar o erro entre a saída obtida e a saída desejada para cada um destes conjuntos e ajustar os pesos das conexões para diminuir o erro

médio encontrado. Isso evita que, ao ajustar a rede para uma entrada, aumente-se o erro das outras.

O método de treinamento *traingdm* possui um vetor de parâmetros de treinamento que ajustam e definem o processo de treinamento. Tais parâmetros determinam o número máximo de épocas utilizadas no treinamento, o erro mínimo desejado, a taxa de aprendizado e o valor do termo de *momentum*. O processo de treinamento ocorre, ajustando os pesos conforme a taxa de aprendizado e o termo *momentum*, até que o número de épocas ou o erro mínimo desejado seja alcançado.

A rede aqui tratada foi treinada utilizando-se os parâmetros de treinamento apresentados na tabela a seguir.

Parâmetro	Valor
Taxa de aprendizado	0,4
<i>Momentum</i>	0,4
Épocas	10000
Erro médio	0,001

Tabela 5.3: Parâmetros de treinamento da rede neural utilizada para segmentar as linhas da imagem.

O processo de treinamento ocorreu em três etapas. Inicialmente a rede foi treinada utilizando-se o conjunto de treinamento LN1 e seu respectivo vetor de saída (descritos no item 5.4.2 deste trabalho). Em seguida, partindo da rede já treinada com o conjunto de treinamento LN1, realizou-se novamente o treinamento utilizando-se o conjunto de treinamento LN2 e seu respectivo vetor de saída. Finalmente a rede, já treinada com os conjuntos LN1 e LN2, foi treinada com ambos os conjuntos de treinamento e ambos os vetores de saída, compreendendo desta vez, 8000 exemplos de treinamento.

5.6 Segmentação da Coluna

Nesta etapa é apresentado o desenvolvimento de uma rede neural do tipo *perceptron* de múltiplas camadas capaz de identificar entre as colunas que compõem o conjunto de linhas resultantes da primeira rede, quais contém a placa do veículo. Para realizar esta tarefa a rede aqui proposta deve receber como entrada uma matriz contendo 20 linhas resultantes da primeira rede, o que corresponde ao tamanho geral das placas nas imagens utilizadas neste trabalho, somadas de 3 linhas anteriores e 3 linhas posteriores que devem compreender o fundo da placa.

O objetivo é, partindo de uma imagem já segmentada pela primeira rede, reduzir a área da mesma, mantendo apenas as colunas que apresentam os caracteres da placa, conforme mostram as figuras a seguir.



Figura 5.8: Imagem utilizada para segmentar as colunas.

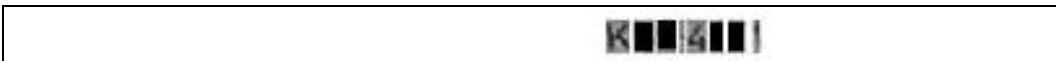


Figura 5.9: Saída da rede neural utilizada para segmentar as colunas.

5.6.1 Binarização da Coluna

Com o objetivo de diminuir a variação dos valores de entrada da rede, da escala de 0 a 255 utilizada para representar os valores de cinza da imagem, para valores de 0 e 1, é realizada a binarização da imagem através da utilização da função *im2bw*, pertencente a toolbox de processamento de imagens do MATLAB®. Essa rotina retorna, a partir de uma imagem digital em tons de cinza, uma imagem em preto e branco, representada pelos valores 0 e 1, gerada por meio da aplicação de um valor de *threshold* que varia entre 0 e 1. Neste caso foi utilizado o

valor de 0,5 o que corresponde a 50% (cinquenta por cento) do nível de branco da imagem.

A figura a seguir mostra o resultado da binarização da imagem através do uso da função *im2bw*.



Figura 5.10: Entrada da rede neural utilizada para segmentar as colunas.

5.6.2 Preparação dos Dados de Treinamento

Para o treinamento de uma rede neural capaz de segmentar as colunas que compreendem os caracteres da placa de um veículo, foi preparado a partir dos conjuntos de imagens A e B, um conjunto de treinamento contendo 800 colunas correspondentes à placa, e 800 colunas correspondentes a outras áreas da imagem, formando uma matriz de 26x1600 denominada CL1. O conjunto de colunas que contém placa compreende exemplos de todas as 26 letras do alfabeto (A a W) e os 10 dígitos numéricos (0 a 9).

Conjunto	Exemplos com placa	Exemplos sem placa	Total de exemplos
CL1	800	800	1600

Tabela 5.4: composição do conjunto de treinamento utilizado para segmentação das colunas da imagem.

Para a matriz de entrada citada acima (CL1), foi criado um vetor de saída com 1x1600 posições contendo valores de 0,9 para colunas que contém placa do veículo e 0,1 para as demais colunas do conjunto de treinamento. Optou-se pelos

valores 0,1 e 0,9 a fim de evitar os extremos da função de ativação *sigmóide* utilizada na camada de saída da rede descrita no item seguinte.

5.6.3 Configuração da Rede Neural

Para a tarefa de classificação das colunas que contém ou não a placa do veículo, foi concebida uma rede neural do tipo *perceptron* de múltiplas camadas contendo vinte e seis (26) entradas, que correspondem as vinte (20) linhas referentes ao tamanho dos dígitos, mais as três linhas anteriores e as três linhas posteriores às mesmas, três camadas de neurônios contendo vinte (20) neurônios na primeira, vinte (20) neurônios na segunda e um (01) neurônio na terceira, também chamada de camada de saída. Na primeira e segunda camadas foi utilizada a função de ativação *tangente hiperbólica* e na camada de saída a função de ativação *sigmóide*. Foi utilizado também um *bias* de valor 1 em cada uma das duas camadas intermediárias e na camada de saída.

A figura 5.11 representa graficamente a topologia da rede neural utilizada para segmentar as colunas da imagem, nela estão representadas: a camada de entrada; a primeira camada contendo vinte (20) neurônios, função de ativação *tangente hiperbólica*, matriz de pesos (LW1) e a matriz de *bias* (b1); a Segunda camada contendo vinte (20) neurônios, função de ativação *tangente hiperbólica*, matriz de pesos (LW2) e a matriz de *bias* (b2); e a camada de saída contendo um (1) neurônio, função de ativação *sigmóide*, matriz de pesos (LW3) e a matriz de *bias* (b3)..

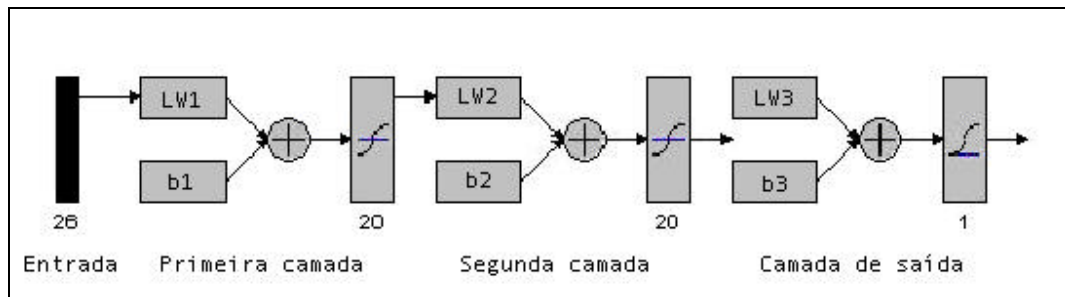


Figura 5.11: Topologia da rede neural utilizada para segmentar as colunas da imagem.

5.6.4 Treinamento da Rede Neural

Para realizar o treinamento da rede neural descrita acima, assim como na rede utilizada para segmentar as linhas da imagem, foi utilizado o algoritmo *backpropagation* através do método *traingdm* do MATLAB[®]. A rede foi treinada utilizando os parâmetros de treinamento apresentados na tabela a seguir.

Parâmetro	Valor
Taxa de aprendizado	0,4
<i>Momentum</i>	0,4
Épocas	50000
Erro médio	0,001

Tabela 5.5: Parâmetros de treinamento da rede neural utilizada para segmentar as colunas da imagem.

O treinamento ocorreu em uma só etapa onde a rede foi treinada utilizando o conjunto de treinamento CL1 descrito no item 5.6.2 deste trabalho e seu respectivo vetor de saída.

5.7 Conclusão

Nesse capítulo foi apresentado o desenvolvimento de um método para segmentar a área da placa de um veículo em uma imagem digital. Foram descritos os principais passos executados no sentido de se obter duas redes neurais artificiais do tipo *perceptron* de múltiplas camadas, uma capaz de identificar as linhas da imagem que constituem a placa e outra capaz de, baseada nas respostas da primeira, identificar quais colunas constituem a placa do veículo.

A utilização em seqüência das duas redes aqui descritas deve ser capaz de segmentar a área correspondente à placa de um veículo em uma imagem digital.

No capítulo seguinte são feitas análises e considerações a respeito do treinamento de cada uma das redes, do processo de preparação dos dados, aqui chamado de binarização das entradas da rede, bem como serão descritos os testes realizados e os resultados obtidos.

Capítulo 6: TESTES E AVALIAÇÕES

6.1 Introdução

No capítulo anterior foi descrita a concepção de um método para segmentação da área correspondente à placa de um veículo em uma imagem digital, por meio do uso de duas redes neurais artificiais do tipo *perceptron* de múltiplas camadas. Serão feitas considerações acerca do treinamento de cada uma dessas redes, bem como apresentados os resultados obtidos pelo referido método, com a aplicação do conjunto de testes.

6.2 Avaliação do Treinamento das Redes

Os itens a seguir, apresentam alguns dados importantes no que se refere ao treinamento das redes neurais aqui propostas. O objetivo principal desta apresentação é fornecer dados que possibilitem futuras comparações, tanto no que se refere a testes com outros métodos de treinamento quanto com outros parâmetros de treinamento, ou mesmo, outras topologias de redes.

6.2.1 Treinamento da Rede Neural para Segmentação das Linhas da Imagem

Utilizando o método de treinamento, os parâmetros e a topologia de rede apresentados nos itens **5.5.3** e **5.5.4** do capítulo anterior, foram executadas três (03) etapas de treinamento da rede neural utilizada para segmentar as linhas da imagem, a saber: Um treinamento com o conjunto LN1 contendo 4000 exemplos; outro treinamento com o conjunto LN2 contendo também 4000 exemplos; e por fim

um terceiro e último treinamento com ambos os conjuntos de treinamento [LN1 LN2] contendo 8000 exemplos. Todos os treinamentos realizados partiram do ponto inicial do experimento imediatamente anterior, ou seja, sem reinicializar a rede neural. Dessa forma obteve-se um total de quatro redes neurais diferentes, embora similares e possuidoras da mesma estrutura básica, sendo a última utilizada nos testes realizados neste capítulo.

O primeiro treinamento (conjunto LN1) partiu da rede inicializada, levou trinta e nove minutos (00:39') de processamento, iniciou com um erro médio de 0,360767 e ao término de 2071 épocas, atingiu o erro especificado nos parâmetros de treinamento: 0,001.

O segundo treinamento (conjunto LN2) partiu da rede já treinada com o conjunto LN1, levou uma hora e vinte e um minutos (01:21'), iniciou com um erro médio de 0.226173, e ao término de 4365 épocas, atingiu o erro especificado nos parâmetros de treinamento: 0,001.

O terceiro e último treinamento (conjuntos LN1 e LN2), levou duas horas e quarenta e oito minutos (02:48') de processamento, iniciou com um erro médio de 0.0478804, e ao término de 3044 épocas, atingiu o erro especificado nos parâmetros de treinamento: 0,001.

Pode-se observar que, no último treinamento realizado, o erro médio inicial é bem menor que os obtidos nos dois treinamentos anteriores, tendo em vista que este último partiu da rede previamente treinada com dois treinamentos anteriores. Apesar disso, o número de épocas necessárias para que se atingisse o erro desejado não diminuiu na mesma proporção, uma vez que os exemplos a serem modelados pela rede na forma de uma regra geral se tornou mais complexo. Outro fator a ser considerado foi o aumento no tempo de duração do último treinamento, que se justifica pelo aumento do número de exemplos a serem considerados em cada época de treinamento.

Conjunto	Exemplos	Horas	Erro médio inicial	Número de épocas
LN1	4000	00:39	0.360767	2071
LN2	4000	1:21	0.226173	4365
[LN1 LN2]	8000	2:48	0.0478804	3044

Tabela 6.1: Valores referentes aos treinamentos da rede neural utilizada para segmentação das linhas da imagem.

As figuras a seguir apresentam os gráficos de cada um dos três treinamentos realizados.

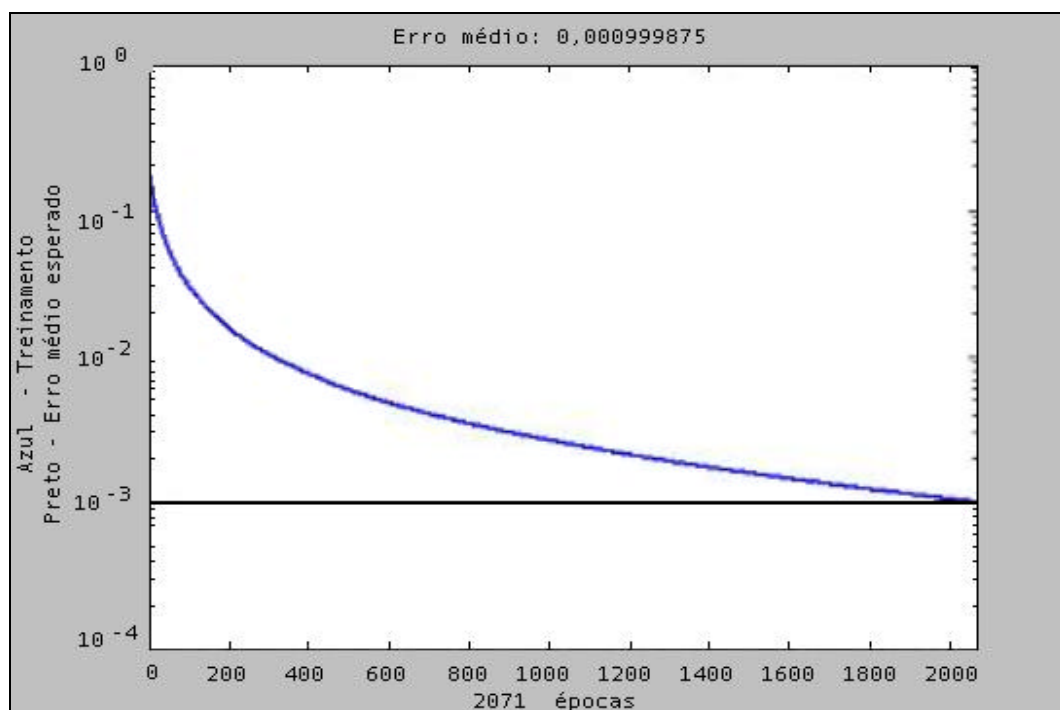


Figura 6.1: Gráfico do treinamento utilizando o conjunto LN1

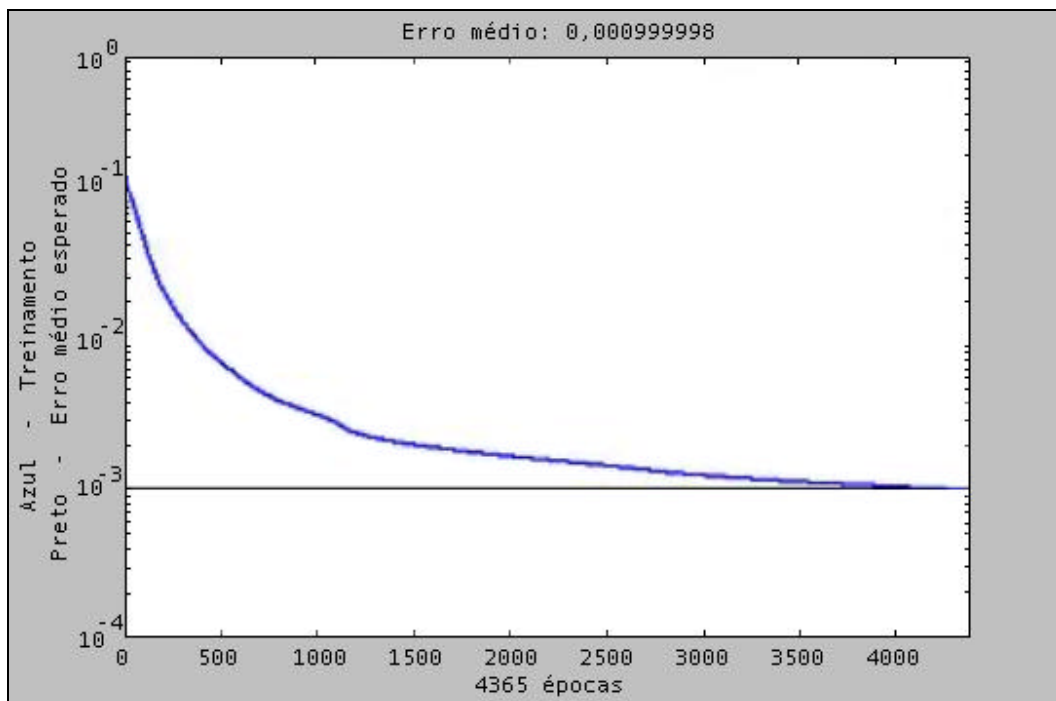


Figura 6.2 : Gráfico do treinamento utilizando o conjunto LN2

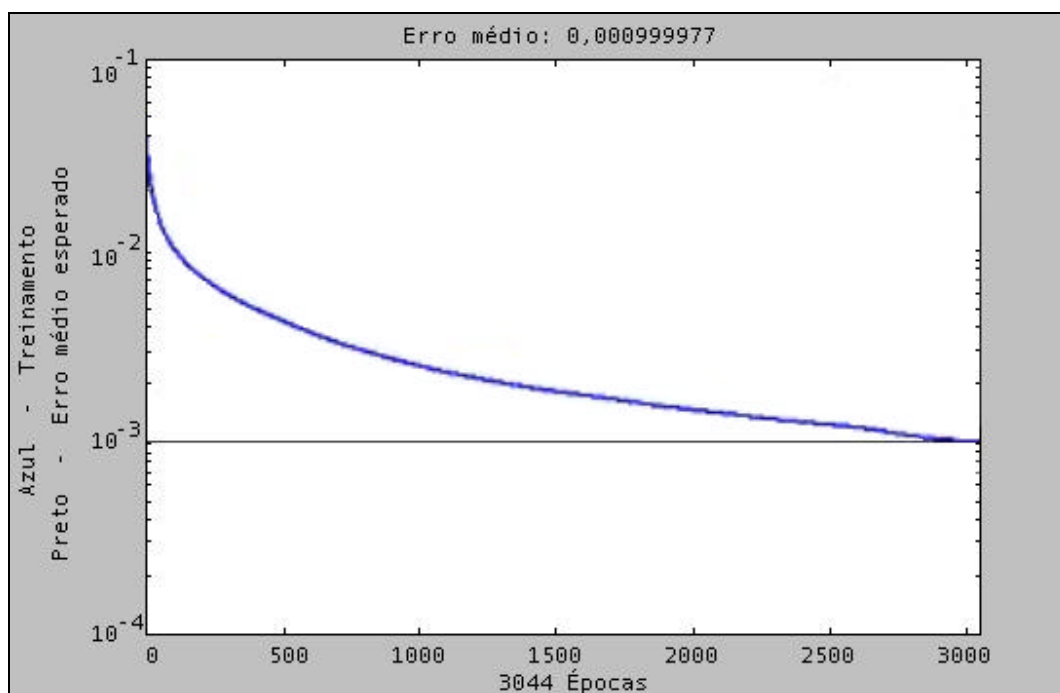


Figura 6.3: Gráfico do treinamento utilizando o conjunto [LN1 LN2]

6.2.2 Treinamento da Rede Neural para Segmentação das Colunas da Imagem

Utilizando o método de treinamento, os parâmetros e a topologia de rede apresentados nos itens **5.6.3** e **5.6.4** do capítulo anterior, foi executada uma única etapa de treinamento da rede neural utilizada para segmentar as colunas da imagem. Partindo da rede neural inicializada, foi executado o treinamento utilizando o conjunto CL1, contendo 1600 exemplos de treinamento.

O processamento durou uma hora e treze minutos (01:13'), iniciou com um erro médio de 0.283493, e ao término de 37676 épocas, atingiu o erro mínimo especificado nos parâmetros de treinamento, ou seja, 0,001. Dessa forma, obteve-se a rede neural treinada utilizada nos testes realizados neste capítulo.

Conjunto	Exemplos	Horas	Erro médio inicial	Número de épocas
CL1	1600	1:13	0.283493	33676

Tabela 6.2: Valores referentes ao treinamento da rede neural utilizada para segmentação das colunas da imagem.

A figura a seguir apresenta o gráfico do treinamento realizado com o conjunto CL1.

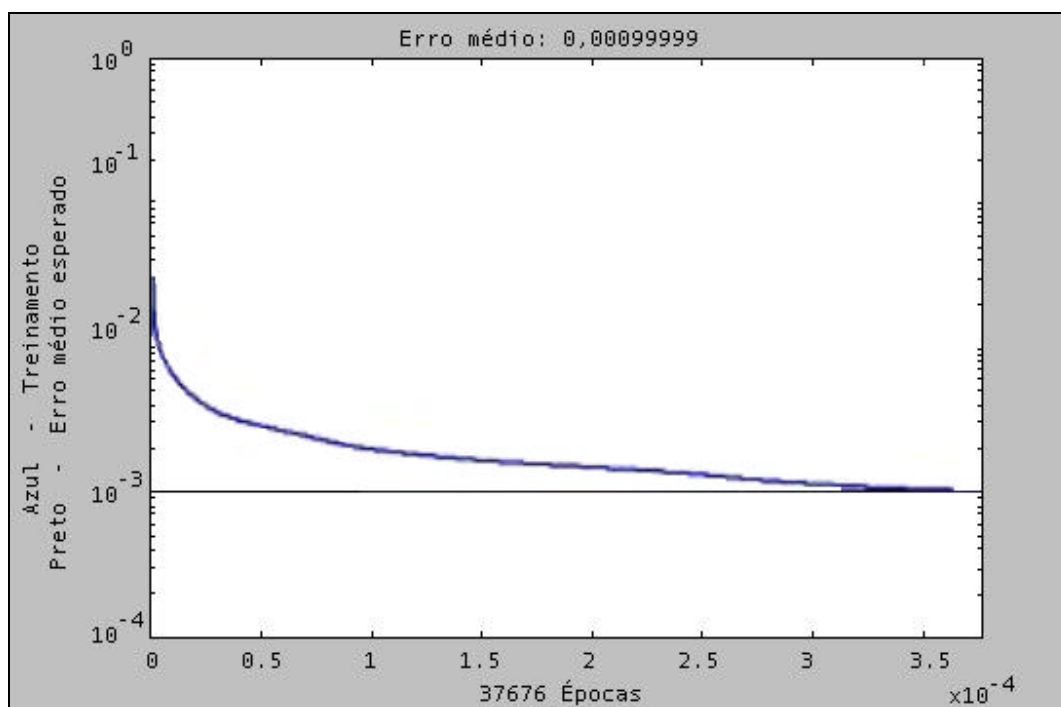


Figura 6.4: Gráfico do treinamento utilizando o conjunto CL1.

6.3 Método de Binarização das linhas da imagem

O método utilizado para binarizar as linhas da imagem que servirão de entrada da rede reduzindo a mesma para valores de 0 e 1 (item 5.5.1 do capítulo anterior), mostrou-se eficiente quando se trata de imagens com boas condições de iluminação, como é o caso das imagens utilizadas neste trabalho, mas apresentou deficiência para aquelas imagens cuja distribuição da iluminação não é homogênea, como por exemplo, as imagens obtidas durante a noite, como mostra a figura a seguir.

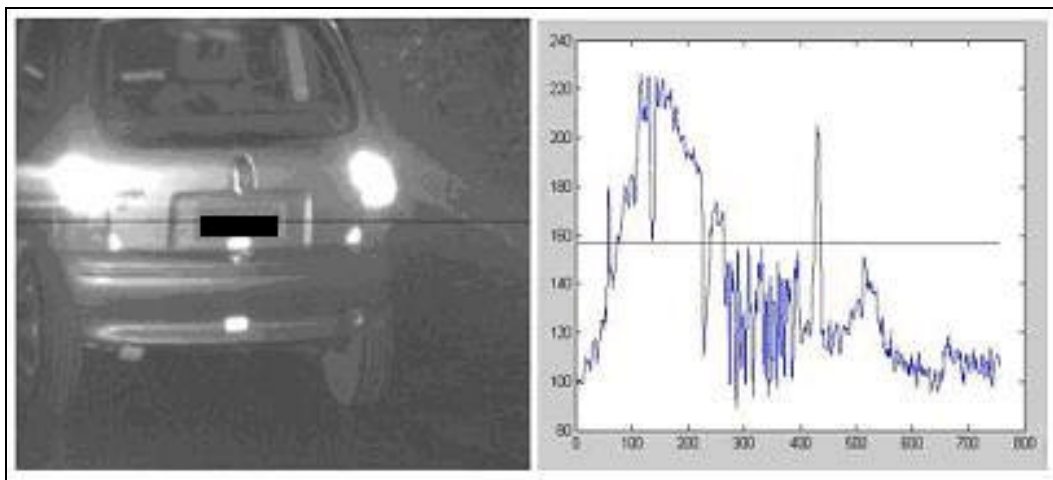


Figura 6.5: Imagem onde o método de binarização apresenta problemas.

No caso dessas imagens, o valor médio dos níveis de cinza não passa pela assinatura característica da região da placa do veículo, invalidando completamente a binarização da linha em questão.

Outros métodos de binarização, bem como outras técnicas de tratamento de imagens (aumento de contraste, detecção de bordas etc..), podem ser testadas e contribuir significativamente para uma melhora no desempenho geral do método proposto neste trabalho.

Todas as imagens que fazem parte dos conjuntos de treinamento e testes utilizados neste trabalho, foram previamente testadas com relação ao método de binarização utilizado, sendo que nem uma delas apresentou o problema descrito acima.

6.4 Resultados

Com o objetivo de registrar e avaliar o desempenho individual das duas redes neurais propostas, no que se refere a classificação correta tanto de linhas

quanto de colunas que contém ou não placa do veículo, são apresentados os resultados obtidos com a simulação de ambas as redes neurais.

Este procedimento, chamado de simulação, foi realizado por meio do uso da rotina *sim* do MATLAB[®], e executado utilizando exemplos de testes retirados do conjunto de imagens C (item 5.4). Os resultados obtidos individualmente por cada uma das redes são apresentados nos itens a seguir.

Considerando que; a resposta esperada para o problema é uma variável booleana (sim/não); as redes treinadas para realizar as tarefas de segmentação das linhas e das colunas da imagem possuem na camada de saída a função de ativação *sigmóide* com o seu intervalo de saída variando entre 0 (zero) e 1 (um); e ambas as redes foram treinadas para apresentar valores 0,9 (zero vírgula nove) para as linhas ou colunas que possuem placa e 0,1 (zero vírgula um) para as linhas ou colunas que não possuem placa, foram realizados dois conjuntos de testes.

Inicialmente, com a aplicação de um *threshold* na saída das redes, foram considerados como **1** (sim) os valores de saída maiores ou igual a **0,5** ou como **0** (não) os valores de saída menores que **0,5**, conforme demonstra a equação a seguir.

$$f(x) = \begin{cases} 0 & \text{se } x < 0,5 \\ 1 & \text{se } x \geq 0,5 \end{cases}$$

Equação 6.1

Em seguida foram realizados os mesmos testes, desta vez considerando a existência de uma zona nebulosa no intervalo de respostas das redes. Esta zona nebulosa foi definida empiricamente como correspondendo aos valores de saída situados entre 0,4 (zero vírgula quatro) e 0,6 (zero vírgula seis). Desta forma, os valores menores que **0,4** são considerados como **0** (não), os valores maiores que

0,6 são considerados como **1** (sim), e os valores situados na zona nebulosa são considerados como não identificados (*N*), conforme demonstra a equação a seguir.

$$f(x) = \begin{cases} 0 & \text{se } x < 0,4 \\ N & \text{se } 0,4 \leq x \leq 0,6 \\ 1 & \text{se } x > 0,6 \end{cases}$$

Equação 6.2

É necessário salientar que, a definição mais precisa da zona nebulosa no intervalo de respostas das redes, pode ser obtida por meio de uma análise detalhada dos erros aferidos com a aplicação dos conjuntos de testes. Essa definição pode tornar mais confiável a classificação da rede, uma vez que o intervalo de erros é identificado com precisão.

Os resultados das simulações são apresentados por meio da “matriz de confusão”, demonstrando os valores reais “versus” os valores estimados do conjunto de testes (Tabelas 6.4, 6.5, 6.7 e 6.8). As colunas da matriz mostram os valores reais e as linhas, os valores estimados. As células cujos números estão na diagonal principal, apresentam o resultado no qual o real é igual ao estimado (respostas corretas). Quanto mais próximo for a soma dessas células do total de exemplos, maior o grau de acerto do modelo.

6.4.1 Segmentação das Linhas

Para simulação da rede neural treinada para segmentar as linhas da imagem, foram realizadas simulações com dois conjuntos de testes, que contêm linhas retiradas do conjunto de imagens C. O primeiro, com 500 (quinhentos) exemplos de linhas que não correspondem a placa do veículo (conjunto TLS); e o

segundo com 500 (quinhentos) exemplos que correspondem à placa do veículo (conjunto TLC), como apresenta o quadro a seguir.

Conjunto	Exemplos c/ placa	Exemplos s/ placa	Total de exemplos
TLS	0	500	500
TLC	500	0	500

Tabela 6.3: Composição dos conjuntos de testes (segmentação das linhas)

Os resultados obtidos nos testes considerando o *threshold* na saída das redes são apresentados na “matriz de confusão” abaixo:

Predição	EXEMPLOS		Predição Total	% Erro Predição
	TLC	TLS		
Com Placa	468	18	486	3,70
Sem Placa	32	482	514	6,23
Total de Exemplos	500	500	1000	5,00
% Erro	6,40	3,60	5,00	

Tabela 6.4: Resultados obtidos na simulação da rede treinada para segmentar as linhas (*threshold*)

Analisando a “matriz de confusão”, pode-se observar que o percentual de erro na predição da rede é menor para o conjunto de linhas que não possuem placa (3,60%) do que para o conjunto de linhas que possuem a placa do veículo (6,40%). O erro de predição total é de 5%, incidente sobre os 1000 exemplos apresentados à rede, classificando corretamente 950 e erroneamente 50 do conjunto de total de testes utilizado. Do total de 486 exemplos preditos como 1 (sim), constata-se que o número de predições realmente corretas é de 468 e o número de predições erradas é 18, que demonstra um erro de predição de 3,70%.

No tocante ao total de 514 predições como 0 (não), constata-se que o número de predições corretas é 482 e o número de predições erradas é 32, que denota um erro de predição de 6,23%.

Os percentuais de erros e acertos são apresentados graficamente desta forma:

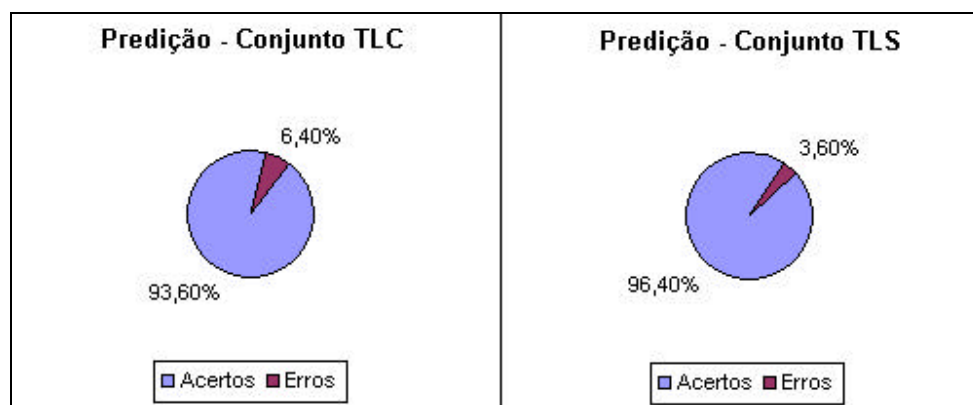


Figura 6.6: Gráfico da predição – conjuntos TLC e TLS.

Nas simulações que consideram a existência da faixa nebulosa obteve-se os resultados apresentados na “matriz de confusão” a seguir.

Predição	EXEMPLOS		Predição Total	% Erro Predição
	TLC	TLS		
Com Placa	457	15	472	3,18
Sem Placa	24	482	506	4,74
Não identificado	19	3	22	2,20
Total de Exemplos	500	500	1000	6,10
% Erro	4,80	3,00	3,90	
% Não identificado	3,80	0,60	2,20	
% Erro total	8,60	3,60	6,10	

Tabela 6.5: Resultados obtidos na simulação da rede treinada para segmentar as linhas (zona nebulosa).

Na análise incidente sobre os 500 exemplos de linhas apresentados com placa do veículo, observa-se que o percentual de erro total é de 8,6%, incluindo-se 4,8% referentes a predições erradas e 3,8% de predições não identificadas. Nos outros 500 exemplos sem placa do veículo, o percentual de erro total é de 3,6% incluindo-se 3% de predições erradas e 0,6% de predições não identificadas. O erro de predição total neste caso é de 6,10% incidente sobre os 1000 exemplos apresentados à rede, incluindo nesse percentual as predições erradas e as não identificadas. Do total de 472 exemplos preditos como 1 (sim), constata-se que o número de predições realmente corretas é 457 e o número de predições erradas é 24, que demonstra um erro de predição de 3,18%. No tocante ao total de 506 predições como 0 (não), constata-se que o número de predições corretas é 482 e o número de predições erradas é 19, que denota um erro de predição de 4,74%.

Os percentuais de erros, acertos e não identificados são apresentados graficamente desta forma:

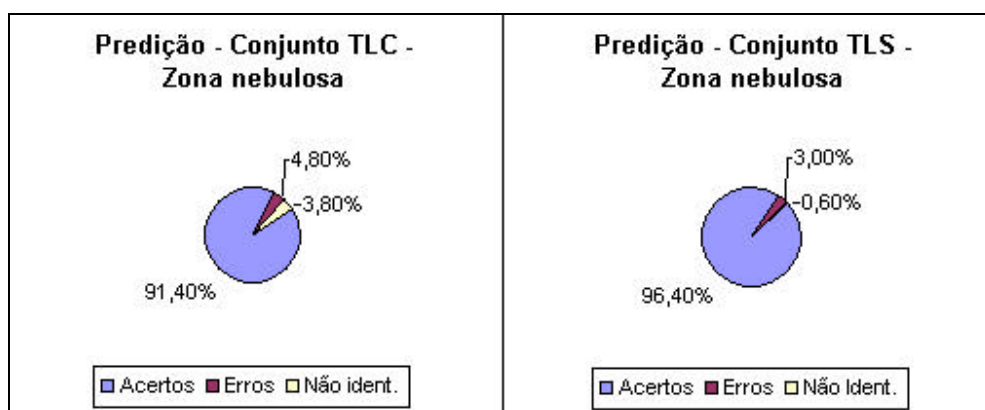


Figura 6.7: Gráfico da predição – conjuntos TLC e TLS (zona nebulosa).

6.4.2 Segmentação das Colunas

Para simulação da rede neural treinada para segmentar as colunas da imagem foram utilizados dois conjuntos de testes. O primeiro deles contendo 500

exemplos de colunas que não correspondem a placa do veículo (conjunto TCS), o segundo contendo 500 exemplos de colunas que correspondem a placa do veículo (conjunto TCC), conforme apresenta o quadro a seguir.

Conjunto	Exemplos c/ placa	Exemplos s/ placa	Total de exemplos
TCS	0	500	500
TCC	500	0	500

Tabela 6.6: Composição dos conjuntos de testes

Os resultados obtidos nos testes considerando o *threshold* na saída das redes são apresentados na “matriz de confusão” abaixo:

Predição	EXEMPLOS		Predição Total	% Erro Predição
	TCC	TCS		
Com Placa	481	37	518	7,14
Sem Placa	19	463	482	3,94
Total de Exemplos	500	500	1000	5,60
% Erro	3,80	7,40	5,60	

Tabela 6.7: Resultados obtidos na simulação da rede treinada para segmentar as linhas (*threshold*)

Analisando a “matriz de confusão”, pode-se observar que o percentual de erro na predição da rede é menor para o conjunto de colunas que não possuem placa (7,40%) do que para o conjunto de colunas que possuem a placa do veículo (3,80%). O erro de predição total é de 5,60%, incidente sobre os 1000 exemplos apresentados à rede, classificando corretamente 944 e erroneamente 56 do conjunto de total de testes utilizado. Do total de 518 exemplos preditos como 1 (sim), constata-se que o número de predições realmente corretas é de 481 e o

número de predições erradas é 37, que demonstra um erro de predição de 7,14%. No tocante ao total de 482 predições como 0 (não), constata-se que o número de predições corretas é 463 e o número de predições erradas é 19, que denota um erro de predição de 3,94%.

Os percentuais de erros e acertos são apresentados graficamente desta forma:

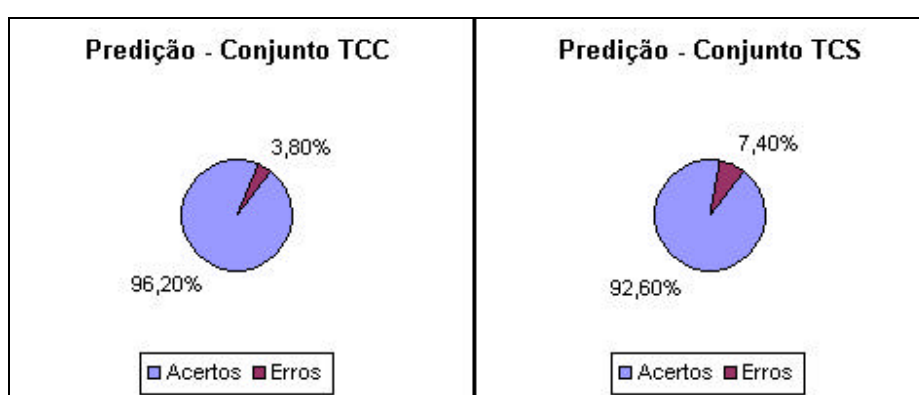


Figura 6.8: Gráfico da predição – conjuntos TCC e TCS.

Nas simulações que consideram a existência da faixa nebulosa obteve-se os resultados apresentados na “matriz de confusão” a seguir.

Predição	EXEMPLOS		Predição Total	% Erro Predição
	TCC	TCS		
Com Placa	478	36	514	7,00
Sem Placa	18	458	476	3,78
Não identificado	4	6	10	1,00
Total de Exemplos	500	500	1000	6,40
% Erro	3,60	7,20	5,40	
% Não identificado	0,80	1,20	1,00	
% Erro total	4,40	8,40	6,40	

Tabela 6.8: Resultados obtidos na simulação da rede treinada para segmentar as linhas (zona nebulosa)

Na análise incidente sobre os 500 exemplos de colunas apresentados com placa do veículo, observa-se que o percentual de erro total é de 4,4%, incluindo-se 3,6% referentes a predições erradas e 0,8% de predições não identificadas. Nos outros 500 exemplos sem placa do veículo, o percentual de erro total é de 8,4% incluindo-se 7,2% de predições erradas e 1,2% de predições não identificadas. O erro de predição total neste caso é de 6,4% incidente sobre os 1000 exemplos apresentados à rede, incluindo nesse percentual as predições erradas e as não identificadas. Do total de 514 exemplos preditos como 1 (sim), constata-se que o número de predições realmente corretas é 478 e o número de predições erradas é 36, que demonstra um erro de predição de 7%. No tocante ao total de 476 predições como 0 (não), constata-se que o número de predições corretas é 458 e o número de predições erradas é 18, que denota um erro de predição de 3,78%.

Os percentuais de erros, acertos e não identificados são apresentados graficamente desta forma:

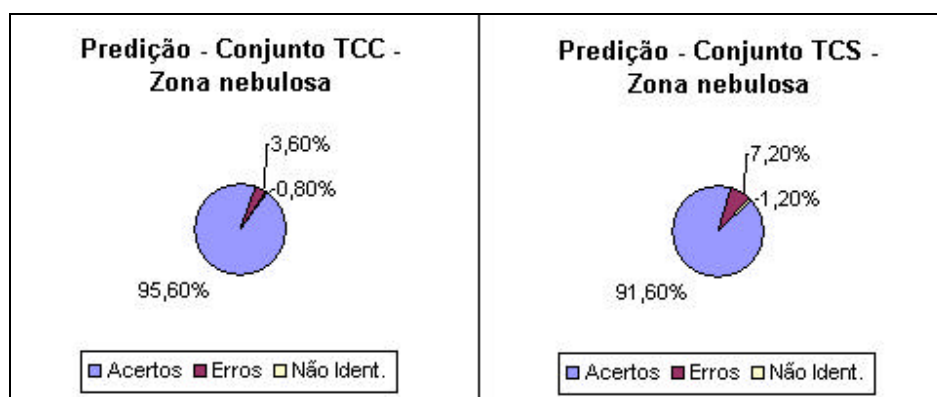


Figura 6.9: Gráfico da predição – conjuntos TCC e TCS (zona nebulosa).

6.5 Conclusão

Neste capítulo foram feitas considerações a respeito do treinamento e dos resultados obtidos pelas duas redes neurais apresentadas no capítulo anterior.

Foram descritos os processos de treinamento das duas redes neurais artificiais utilizadas para segmentar as linhas e colunas da imagem. O treinamento da primeira rede, deu-se em três (03) etapas e envolveu o total de 8000 exemplos de treinamento, e em todas elas, o erro médio desejado (0,001) foi obtido em tempo e número de épocas relativamente curtos. O treinamento da segunda rede foi executado em uma única etapa e exigiu o número de épocas consideravelmente maior para alcançar o erro médio especificado (0,001).

Este capítulo apresentou também considerações referentes às limitações do método utilizado para binarizar as linhas da imagem ocorridas naquelas que não apresentam boas condições de iluminação, sendo que todas as imagens tratadas neste trabalho foram previamente testadas quanto a esse problema.

Esclarecemos que na simulação do método, utilizou-se uma zona nebulosa no intervalo de respostas da rede definida empiricamente, sendo que a sua determinação precisa pode ser obtida por meio da análise mais detalhada dos erros obtidos durante os testes.

Após a simulação das duas redes neurais com os conjuntos de testes, verificou-se que é possível realizar a segmentação de imagens por meio do uso de tais técnicas. Os testes, considerando um *threshold* na saída das redes, apresentaram um erro geral de predição de 5% (cinco por cento) para as linhas da imagem e um erro geral de predição de 5,6% (cinco vírgula seis) para as colunas da imagem.

Capítulo 7: CONCLUSÕES

7.1 Introdução

Este trabalho relatou o desenvolvimento e teste de um método que, por meio do uso de técnicas de tratamento de imagens e principalmente de redes neurais artificiais, é capaz de segmentar a área correspondente a placa de um veículo em imagens digitais, sendo que o seu principal propósito é avaliar o desempenho de tais técnicas na execução dessa tarefa, contribuindo assim para solução do problema.

Foram abordados inicialmente assuntos referentes ao funcionamento dos controladores eletrônicos de velocidade, às imagens digitais obtidas por estes equipamentos e às informações relevantes nelas contidas, sendo apresentado também, um breve resumo das principais abordagens atualmente existentes a respeito do assunto.

É feita uma revisão conceitual das principais áreas aqui tratadas, começando pelos conceitos referentes à visão computacional e às imagens digitais; e em seguida, são apresentados os principais conceitos referentes às redes neurais artificiais, que correspondem à principal técnica utilizada neste trabalho. Todos os conceitos vistos são importantes para que se possa compreender o método proposto.

Neste capítulo verificamos os principais pontos a serem destacados nesta dissertação e os futuros trabalhos que poderão advir.

A seguir, é apresentado um sumário da dissertação, seguido das conclusões gerais e das perspectivas de trabalhos futuros.

7.2 Sumário da Dissertação

Esta dissertação contemplou várias etapas, desde o seu projeto inicial, passando pelas seguintes fases:

- **Revisão Bibliográfica:** etapa que se destinou à pesquisa e à leitura de textos pertinentes ao assunto, incluindo livros e artigos sobre redes neurais artificiais, Inteligência Artificial e visão computacional. O material revisto e consultado consta nas referências bibliográficas.
- **Aprendizagem sobre a plataforma escolhida:** A ferramenta escolhida para o desenvolvimento e testes do método ora proposto foi o MATLAB®, o qual possui *Toolboxes* específicas sobre diversas áreas, inclusive sobre visão computacional e redes neurais artificiais, estudadas nesta etapa do trabalho.
- **Implementação:** Nesta fase, foram aplicados os conhecimentos adquiridos no desenvolvimento do método proposto, bem como foram realizados testes com relação à escolha da arquitetura de rede a ser utilizada, preparação dos dados e definição de conjuntos de treinamento e testes para as redes ora propostas.
- **Execução de testes e avaliações:** Desenvolvido o método, foram executados testes no sentido de avaliar o desempenho do mesmo.

Realizadas todas essas etapas, foi desenvolvida esta dissertação de mestrado que trata de vários assuntos relacionados às imagens digitais e às redes neurais artificiais com a finalidade de documentar o processo de pesquisa envolvido no desenvolvimento e avaliação do método aqui proposto.

Para isto, desenvolve-se primeiramente um levantamento no Capítulo 2, que abrange aspectos referentes ao controle eletrônico de velocidade, o funcionamento de seus equipamentos, as imagens por eles geradas e as principais informações nelas contidas. Consta ainda nesse capítulo, uma breve

abordagem referente às técnicas atualmente existentes em relação ao problema, buscando-se introduzir uma visão geral sobre o mesmo

O Capítulo 3 aborda o tema da visão computacional, englobando aspectos referentes à fisiologia da visão e às imagens digitais, buscando dar um embasamento teórico a respeito de tais conceitos e possibilitando a necessária compreensão acerca da forma de captura e representação de imagens digitais.

O Capítulo 4 desta dissertação, dedica-se exclusivamente às Redes Neurais Artificiais, no qual se observam que tais técnicas tem inspiração em modelos neurais biológicos, possuindo estrutura distribuída e paralela que realiza processamento não linear. As redes neurais artificiais possuem topologias variadas e passam por um processo de aprendizado que pode ser supervisionado ou não supervisionado, neste processo os pesos referentes as conexões sinápticas são atualizados visando alterar a saída da rede.

O Capítulo 5, descreve o método para segmentação da área correspondente à placa de um veículo em uma imagem digital, abordando-se aspectos relativos à elaboração de conjuntos de dados utilizados na realização do treinamento e testes de duas redes neurais, do tipo *perceptron* de múltiplas camadas, capazes de identificar linhas e colunas que possuem características específicas de existência de placa na imagem.

Por fim, o Capítulo 6 complementa o anterior, demonstrando os resultados de algumas avaliações realizadas utilizando o método proposto.

7.3 Conclusões Gerais

A expectativa de contribuição prática desta dissertação é a de apresentar as redes neurais artificiais como uma ferramenta importante no que se refere à segmentação de imagens digitais e proporcionar um ponto de partida a pesquisas que se dedicam à solução do problema.

Podemos observar que a utilização de redes neurais artificiais nas tarefas de segmentação de imagens, é uma alternativa viável. Os testes realizados apresentaram erros 5,0% (cinco por cento) para classificação das linhas e 5,6% (cinco vírgula seis por cento) para a classificação das colunas da imagem. Tais resultados são satisfatórios e demonstram a efetiva aplicabilidade das redes neurais na solução do problema de segmentação da área da placa em imagens de veículos obtidas por controladores eletrônicos de velocidade.

Foram realizados também testes considerando a existência de uma zona nebulosa no intervalo de respostas das redes. Dessa forma, os valores de resposta situados dentro da zona nebulosa definida empiricamente como sendo o intervalo entre 0,4 e 0,6 são considerados como não identificados. Os testes realizados na classificação de linhas da imagem apresentaram erros de 3,9% (três vírgula nove por cento) e 2,2% (dois vírgula dois por cento) de não identificados. Os testes realizados na classificação de colunas da imagem apresentaram erros de 5,4% (cinco vírgula quatro por cento) e 1,0% (um por cento) de não identificados. A definição mais precisa da zona nebulosa pode ser obtida por meio de uma análise mais detalhada dos erros aferidos com a aplicação do conjunto de testes. Essa definição pode tornar mais confiável a classificação da rede.

É importante lembrar que a qualidade da imagem a ser tratada é de grande influência na tarefa de extrair informações da mesma, e técnicas de tratamento de imagens devem ser aplicadas no sentido de melhorar o desempenho das redes neurais que se proponham a realizar esta tarefa, haja vista que um sistema de obtenção de informações de uma imagem digital engloba vários aspectos, desde a captura da imagem; passando pelo seu armazenamento; pré-processamento e processamento propriamente dito, e sua eficiência depende da qualidade de cada uma destas etapas.

O método ora proposto serve como parâmetro a quem se propõe a desenvolver metodologias para obtenção de informação de imagens digitais ressaltando que para se tornar uma solução geral é necessário realizar estudos com a finalidade de melhorar seu desempenho, principalmente em imagens que

não apresentam boas condições de iluminação ou apresentam algum outro fator de ruído, como é o caso das imagens obtidas em dias de chuva e em veículos com a placa ofuscada por sujeira ou pintura gasta.

7.4 Trabalhos Futuros

É possível prever alguns estudos futuros como forma de continuação deste trabalho, tendo como tarefa pertinente a ser executada os testes de desempenho de outras topologias de redes e outros métodos de treinamento em comparação aos resultados obtidos pelas redes propostas neste trabalho. Outra tarefa bastante importante, no que se refere a buscar melhorar os resultados ora obtidos, é a utilização de técnicas de pré-processamento das imagens, como por exemplo as técnicas de realce, clareamento, identificação de bordas e contornos, afim de melhorar a qualidade da informação submetida à rede, possibilitando desta forma, avaliar os efeitos de tais técnicas sobre o resultado final na solução do problema. Uma análise mais detalhada dos resultados obtidos com a aplicação do conjunto de testes, pode ser realizada possibilitando uma definição mais precisa da zona nebulosa no intervalo de respostas das redes.

Podem-se, ainda, realizar pesquisas no sentido de utilizar as redes neurais na tarefa de obtenção de outras informações relevantes, tais como a identificação do modelo e do fabricante do veículo em uma imagem digital.

De forma geral, espera-se que este trabalho contribua na solução do problema de obtenção de informação de imagens digitais originadas por controladores eletrônicos de trânsito e velocidade, proporcionando elementos para futuras pesquisas nesta área.

Capítulo 8: REFERÊNCIAS BIBLIOGRÁFICAS

- [1] BRUGGE, M. H. T., NIJHUIS J. A. G., SPAANENBURG L., SDEVENS J. H.; *License Plate Recognition, Knowledge-based Intelligent Techniques in Character Recognition*, Edited by Lakhmi C. Jain and Beatrice Lazzerini, pp 263-295, 1999.
- [2] CUI, Y., HUANG, Q.; *Automatic License Extration from Moving Vehicles*. Siemens Corporate Research, Inc, ICIP'97.
- [3] BUSSADOR, A., FACON, J.; *Método Granulométrico para Segmentação de Placas de Veículos* - UCP, Curitiba, 2000.
- [4] BALLARD, D. H., BROWN, C. M.; *Computer Vision*. Prentice-Hall, Englew2ood Cliffs, 1982.
- [5] MARR, D.; *Vision*. W. H. Freeman & Company, New York, 1982.
- [6] MORGAN, C. T.; *Psicología Fisiológica*. Editora Pedagógica e Universitária, São Paulo, 1973.
- [7] KRECH, D., CRUTCHFIELD, R. S.; *Elementos de Psicologia*. Livraria Pioneira Editora, São Paulo, 1963
- [8] SAPARINA, Y.; *A Cibernética está em Nós*. Editora Saga, Rio de Janeiro, 1967.
- [9] DAVALO, E., NAÏ M, P.; *Dês Reseaux de Neurones*. Eyroles, Paris, 1990.
- [10] WILLRICH, R.; *Sistemas Multimídia Distribuídos*. Florianópolis, 2001.
- [11] SHEPHERD, G. M., KOCH, C.; *Introduction to synaptic circuits*. New York, 1994.

- [12] CHURCHLAND, P. S., SEJNOWSKI, T. J.; *The Computational Brain*. Cambridge, 1992.
- [13] HAYKIN, S.; *Redes Neurais Princípios e Prática*. Bookman, Porto Alegre, 2001.
- [14] SUGA, N.; *Computations of velocity and range in the bat auditory system for echo location*. Cambridge, 1990.
- [15] RAMÓN y CAJÁL.; *Histologie du Systems Nerveux de L'homme et des vertébrés*. Paris, 1911.
- [16] FAGGIN, F.; *VLSI Implementation of neural networks*. Seattle, 1991.
- [17] ALMEIDA, M. A. F.; *Introdução ao estudo de redes neurais artificiais*. UFSC, Florianópolis.
- [18] HECHT-NIELSEN, R.; *Applications of Counterpropagation Networks , Neural Networks*, Vol. 1, pgs 131-140. 1988.
- [19] McCULLOCH, W. S., PITTS, W.; *A logical calculus of the ideas immanent in nervous activity*. 1943
- [20] HINTON, G. E.; *Connectionist Learning Procedures, Artificial Intelligence*, Vol. 40, pgs. 185-234. 1989.
- [21] SIMPSON, P.; *Artificial Neural Systems: Foundations, Paradigms, Applications and Implementations*, Pergamon Press Inc. 1990.
- [22] CARBONELL, J. G.; *Introduction: Paradigms for Machine Learning, Artificial Intelligence*, Vol. 40, pgs. 1-9. 1989.
- [23] PESSOA; *Aprendizado Não-Supervisionado em Redes Neurais*, Tese de Mestrado, COPPE Sistemas, UFRJ. 1990.

- [24] LIPPMAN, R. P.; *An Introduction to Computing with Nerual Nets*, IEEE ASSP Magazine, pgs. 4-22. 1987.
- [25] KIM, K. I., PARK, S. H., JUNG, K., PARK, M. H., KIM, H. J.; *The neural network based method for locating car license plate*. Proc. ITC-CSCC'99. 1999.
- [26] RUMELHART, D. E., McCLELLAND, J.; *Parallel Distributed Processing*, vol. 1, MIT Press. 1986.
- [27] HERTZ, J., KROGH, A., PALMER, R. G.; *Introduction to the theory of neural computation*. Redwood City, Adison-Wiley. 1991.
- [28] BEALE, R., JACKSON, T.; *Neural Computing – An Introduction*, Adam Hilger Publishers. 1990.
- [29] RUMELHART, D. E., WEIGEND, S. A.; *Predicting the Future: A Connectionist Approach*, Stanford PDP-90-01, PARC-SSL-90-20. 1990.
- [30] SOH, Y. S., CHUN, B. T., YOON, H. S.; *Design of Real Time Vehicle Identification System*, Proceedings of the 1994 IEE International Conference on Systems, Man and Cybernetics, San Antonio. 1994.
- [31] RIBEIRO, F. F.S.; *Detecção de Volume de Tráfego de Veículos Proporcionada por Visão Computacional via Redes Neurais*, Dissertação de mestrado submetida ao Depto. de Eng. Elétrica da UFBA, Salvador, BA. 1999.
- [32] DUARTE, A. A., LIMA, A. C.; *Wavelets e Redes Neurais Aplicadas na Medição do Fluxo de Veículos Proporcionado por Visão Computacional*, Depto. de Eng. Elétrica – Escola Politécnica – UFBA, Salvador, BA. 2001.
- [33] GOSE, E, JOHNSONBAUGH, R., JOST, S.; *Pattern recognition & image analysis*, Prentice Hall PTR, Upper Saddle River, NJ. 1996.

- [34] McCULLOCH, W. S., PITTS, W.; *How we know universals; the perception of auditory and visual forms*, Bulletin of Mathematical Biophysics. 1947.
- [35] HEBB, D. O.; *The Organization of Behavior*, Wiley. 1949.
- [36] ROSENBLATT, F.; *Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms*. Spartan Books, Washington, DC, 1962.
- [37] MINSKY, M., PAPER, S.; *Perceptrons; an introduction to computational geometry*. MIT Press, Cambridge, MA, 1969.
- [38] WERBOS, P. J.; *Backpropagation: Basics and New Development*. *The Handbook of Brain Theory and Neural Networks*, MIT Press, Cambridge, 1995.
- [39] RUMELHART, D. E., HINTON, G. E., WILLIAMS, R. J.; *Learning internal representations by error propagation*, in: D.E. Rumelhart, J.L. McClelland (Eds.), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, MIT Press, Cambridge, MA, 1986.

Bibliografia Retirada da Internet

- [40] <http://www.abrancet.com.br> - Associação brasileira de monitoramento e controle eletrônico de trânsito, 2001.
- [41] <http://www.iee.ufrgs.br/iee/siav> – Sistema Automático de Identificação de Veículos, 2000.
- [42] <http://www.inf.ufsc.br/~visao/> - Seminário Introdução a Visão Computacional. CPGCC - INE – UFSC.
- [43] <http://www.utad.pt/~jbarroso/html/isie97.html> - Number plate reading using computer vision, 1999.

Outras Bibliografias Utilizadas

- [44] NEURAL NETWORK TOOLBOX USER´S GUIDE, Copyright 1995-1998
by The Math Works, Inc.