

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMATICA
CURSO DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

UM ESTUDO SOBRE RECONHECIMENTO
VISUAL DE CARACTERES ATRAVÉS DE
REDES NEURAIS

por

FERNANDO SANTOS OSORIO

Dissertação submetida como requisito parcial para
a obtenção do grau de Mestre em
Ciência da Computação

Prof. Anatólio Laschuk
Orientador

Porto Alegre, outubro de 1991

CATALOGAÇÃO DA PUBLICAÇÃO

Osório, Fernando Santos

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

302p.

Dissertação (mestrado), Universidade Federal do Rio Grande do Sul, Curso de Pós-Graduação em Ciências da Computação, Porto Alegre, BR-RS, 1991.
Orientador: Laschuk, Anatólio.

Dissertação: Processamento de Imagens: OCR
Reconhecimento de Caracteres: Redes Neurais

"O mundo está dividido entre aqueles que fazem as coisas e aqueles que levam as glórias. Tente ficar no primeiro grupo, porque ali há menos competição."

Lord Chesterfield

"O assunto mais importante do mundo pode ser simplificado até o ponto em que todos possam apreciá-lo e compreendê-lo. Isso é, ou deveria ser, a mais elevada forma de arte."

Charles Chaplin

"Considero a ciência simplesmente como uma maneira de entender como o mundo é e por que ele é assim. Em qualquer época, nosso conhecimento científico é apenas a modernização do nosso entendimento. Não acredito em verdades absolutas. Tenho medo delas porque bloqueiam a busca de um melhor entendimento. Sempre que pensamos ter chegado à resposta final, o progresso, a ciência e o entendimento acabam. No entanto, o conhecimento que nos rodeia não é um objetivo a ser alcançado por si só. Ele deve ser perseguido, acredito, para tornar o nosso mundo um lugar melhor e a vida mais gratificante."

Eliyahu M. Goldratt

Aos meus pais,
à minha amada esposa.

AGRADECIMENTOS

Agradeço ao meu orientador, professor Anatólio Laschuk, pela paciência, apoio e confiança em mim e no meu trabalho;

A todo corpo docente do CPGCC, em especial aos professores Philippe Navaux e Dante Barone, os quais acompanharam e incentivaram meus trabalhos junto ao CPGCC.

Aos funcionários do II, pela ajuda prestada e o "apoio técnico" necessário para o bom andamento de meu trabalho. Agradeço em especial aos funcionários da secretaria, dos laboratórios e em destaque aos da biblioteca, indispensáveis a um bom trabalho científico;

Aos colegas da pós-graduação pelo seu companheirismo, incentivo, e pelas inúmeras discussões e sugestões dadas durante a evolução deste trabalho. Agradeço especialmente a: Alceu Frigeri, Gilberto Marchioro, Fernando Moraes, Carlos Eduardo Pereira, Marcelo Walter, Gladimir Baranovsky, Alexandre Casacurta, Alexandre Agustini e Alex Guazzelli.

Em especial a Adelmo Cechin e Remis Balaniuk, meus verdadeiros "mestres", que me apresentaram as redes neurais e me incentivaram na escolha do tema abordado em minha dissertação de mestrado. O estudo das redes neurais junto ao CPGCC-UFRGS é fruto do trabalho e dedicação destas duas pessoas, aliadas ao incentivo dado pelo prof. Navaux;

A CAPES e CNPQ pelo auxílio financeiro, na forma de bolsa de estudos;

A UFRGS e ao CPGCC pelas condições de estudo e trabalho oferecidas;

A minha família e meus amigos, os quais por mais difícil que tenha sido esta caminhada, nunca me deixaram sem o apoio tão necessário para que eu pudesse atingir meus objetivos;

Obrigado a todos vocês !

SUMÁRIO

GLOSSARIO	9
LISTA DE ABREVIATURAS	10
LISTA DE FIGURAS	11
LISTA DE TABELAS	15
RESUMO	16
ABSTRACT	18
1 INTRODUÇÃO	20
2 PROCESSAMENTO DE IMAGENS	24
2.1 Introdução	24
2.2 Imagem Digital	26
2.2.1 Digitalização	27
2.2.1.1 Digitalizadores	28
2.2.1.2 Amostragem	30
2.2.1.3 Quantização	33
2.2.1.4 Classes de Imagens	35
2.2.2 Armazenamento	38
2.2.2.1 Compactação	39
2.2.2.2 Formato de Arquivos	40
2.2.3 Exibição	42
2.2.3.1 <i>Halftoning</i>	42
2.2.3.2 <i>Dithering</i>	46
2.3 Tratamento de Imagens	47
2.3.1 Histograma	48
2.3.2 Modificação da Escala de Cinza	50
2.3.2.1 Equalização de Histograma	50
2.3.2.2 Limiarização	51
2.3.3 Filtros	53
2.3.3.1 Operações no Domínio Espaço	53
2.3.3.1.1 Convolução	54
2.3.3.2 Operações no Domínio Freqüência	57

2.4 Reconhecimento de Padrões	57
2.4.1 Obtenção de Descrições	59
2.4.2 Obtenção de Classes	59
2.4.2.1 Análise de Padrões	63
2.4.2.2 Análise de Elementos	63
2.4.2.3 Análise de Contornos	63
2.4.2.4 Casamento de Padrões	64
2.4.2.5 Abordagem do Discriminante	65
2.4.2.6 Abordagem Sintática	65
2.4.2.7 Avaliação do Processo de Classificação	65
 3 SISTEMAS DE RECONHECIMENTO DE CARACTERES	67
3.1 Origens e Evolução	67
3.2 Tipos de Sistemas de Reconhecimento de Caracteres	69
3.3 Descrição dos Sistemas OCR	76
3.3.1 Etapas de Processamento	77
3.3.2 Técnicas de Reconhecimento	80
3.3.2.1 Atributos Globais	80
3.3.2.1.1 Casamento de Padrões e Correlações	80
3.3.2.1.2 Transformações e Expansões em Séries	80
3.3.2.2 Distribuição de Pontos	81
3.3.2.2.1 Zoneamento	81
3.3.2.2.2 Momentos	81
3.3.2.2.3 N-Uplas	81
3.3.2.2.4 <i>Characteristic Loci</i>	81
3.3.2.5 Cruzamentos e Distâncias	82
3.3.2.6 Atributos Geométricos e Topológicos	82
3.3.3 Avaliação do Reconhecimento	83
3.3.4 Principais Dificuldades	87
3.4 Análise de Algoritmos	94
3.4.1 Segmentação	94
3.4.1.1 Separação de Textos e Gráficos	94
3.4.1.2 Separação de caracteres	98
3.4.2 Extração de Atributos	104
3.4.2.1 Vizinhança Global	105
3.4.2.2 Vizinhança Local	105
3.4.2.3 Projeção	106

3.4.2.4 Alternância	106
3.4.2.5 Simetria	107
3.4.2.6 Densidade	107
3.4.3 Alteração de Escala e Posição	108
3.5 Perspectivas	111
3.5.1 Adaptação	111
3.5.2 Processamento Paralelo	112
3.5.3 Uso de Contexto	113
4 REDES NEURAIS	115
4.1 Conceitos Básicos	115
4.2 Origem e Evolução	121
4.3 Características e Aplicações das Redes Neurais	123
4.4 Modelos de Redes Neurais	125
4.4.1 Classificação dos Modelos	126
4.4.2 Modelos Adaptativos	130
4.4.2.1 Perceptron	131
4.4.2.2 Adaline	133
4.4.2.3 Madaline	137
4.4.2.4 Redes Multinível	140
4.4.2.5 ART	142
4.4.3 Modelos Competitivos	145
4.4.3.1 Modelo de Hopfield	145
4.4.3.2 Máquina de Boltzmann	149
4.4.3.3 BAM	149
4.4.3.4 Modelo de Kohonen	152
4.5 Aplicação das Redes Neurais Junto aos Sistemas OCR	155
5 SISTEMA N ² OCR	157
5.1 Organização Geral do Sistema	161
5.2 Aquisição de Imagens	164
5.3 Tratamento da Imagem	172
5.4 Preparação dos Dados para a Rede Neural	180
5.4.1 Segmentação	184
5.4.2 Rotulação	189
5.4.3 Ajuste de Escala e Posição	195

5.5 Rede Neural	195
5.5.1 Modelo Adotado	200
5.5.1.1 Escolha do Modelo	200
5.5.1.2 Descrição do Modelo	204
5.5.2 Aprendizado	204
5.5.3 Reconhecimento	212
5.5.4 Extensão do Modelo	213
5.5.5 Implementação em Hardware	217
5.6 Arquivo de Configuração	219
5.7 Avaliação do Desempenho do Sistema	223
5.8 Evolução do Sistema	229
 6 CONCLUSÃO	233
 ANEXO A.1 - Padrões de Descrição de Imagens	236
ANEXO A.2 - Padrões de Descrição de Caracteres	244
ANEXO A.3 - Padrões de Teste Empregados	247
ANEXO A.4 - Avaliação do Desempenho do Sistema N ² OCR	265
ANEXO A.5 - Listagens de Arquivos do Sistema N ² OCR	283
 BIBLIOGRAFIA CITADA	291
BIBLIOGRAFIA CONSULTADA	301

GLOSSÁRIO

Amostragem	- Discretização da imagem no espaço
Artefatos	- Pontos de "sujeira" em uma imagem. São elementos da imagem que não fazem parte do objeto alvo da análise
<i>Bit-Map</i>	- Conjunto de pontos de uma imagem digital
Desempenho	- Avaliação do funcionamento de um sistema em relação a sua eficácia e eficiência
Eficácia	- Desempenho em termos de taxa de acertos
Eficiência	- Desempenho em termos de rapidez do sistema
Performance	- Desempenho de um sistema
Pixel ou Pel	- Elemento de imagem, um ponto de um padrão bit-map ("picture element")
Quantização	- Discretização da imagem na amplitude

LISTA DE ABREVIATURAS

A/D	- Conversor de sinais analógicos para digitais
ADALINE	- Adaptive Linear Element
ADANLIENE	- Adaptive Non-Lienar Expandible Network
ANSI	- American National Standards Institute
ART	- Adaptive Resonance Theory
ASCII	- American Standard Code for Information Interchange
B&W	- Branco e Preto (dispositivo de 2 cores apenas)
BRASCII	- Código ASCII brasileiro, inclui acentuação
BTC	- Block Truncation Encoding
CCD	- Charged Coupled Devices
CPGCC	- Curso de Pós-Graduação em Ciência da Computação
D/A	- Conversor de sinais digitais para analógicos
DPI	- Dots per Inch
FFT	- Fast Fourier Transform
HDTV	- High Definition Television
IEEE	- Institute of Electrical and Electronic Engineers
II	- Instituto de Informática
LUT	- Look Up Table
LZW	- Lempel-Ziv & Welch Algorithm
MADALINE	- Many ADALINE (Adaptive Linear Element)
N ² OCR	- Neural Network for Optical Character Recognition
P&B	- Preto e Branco, o mesmo que B&W
OCR	- Optical Character Recognition
OHCR	- Optical Handwritten Character Recognition
OPCR	- Optical Printed Character Recognition
RAC	- Relative Address Coding
RGB	- Red, Green, Blue
RLE	- Run Length Encoding
RLSA	- Run Length Smoothing Algorithm
SCR	- Single Character Recognition
SIRENE	- Projeto "Ambiente para SIMulação de REdes NEurais"
UFRGS	- Universidade Federal do Rio Grande do Sul
XOR	- Função Ou Exclusivo (Exclusive OR)

LISTA DE FIGURAS

Figura 2.1 - Relações entre as áreas constituintes do processamento Gráfico	25
Figura 2.2 - Esquema de um digitalizador com câmera de vídeo	29
Figura 2.3 - Classificação dos dispositivos de digitalização	30
Figura 2.4 - Exemplos de matrizes de caracteres	32
Figura 2.5 - Exemplo de compactação RLE	40
Figura 2.6 - Exemplos de aproximação de intensi- dades através do uso de halftoning	45
Figura 2.7 - Matrizes de halftoning usadas na geração de cinco tons	45
Figura 2.8 - Histograma da figura 2.6 (b) com o halftoning desfeito (matriz 6x6)	49
Figura 2.9 - Binarização da imagem do texto da figura 2.6 (a)	52
Figura 2.10 - Imagem original e imagem após convolução para a detecção de contornos	56
Figura 2.11 - Esquema geral de um processo de reconhecimento e classificação de imagens	61
Figura 3.1 - Classificação dos sistemas de reconhecimento de caracteres	70
Figura 3.2 - Exemplos de padronização de conjuntos de caracteres para sistemas de OCR	78
Figura 3.3 - Esquema de etapas de processamento em um sistema OCR	77
Figura 3.4 - Complexidade dos sistemas OCR	91
Figura 3.5 - Etapas do processamento do algoritmo RLSA	96
Figura 3.6 - Separação dos textos do documento da figura 3.5	98
Figura 3.7 - Algoritmo geral de separação de caracteres	101

Figura 3.8 - Localização do inicio de uma linha	101
Figura 3.9 - Algoritmo para isolar um caractere	103
Figura 3.10 - Extração de um caractere desconexo	104
Figura 3.11 - Exemplo de vizinhança global	105
Figura 3.12 - Exemplo de padrões de vizinhança local	105
Figura 3.13 - Exemplo de cálculo de projeções	106
Figura 3.14 - Exemplo de cálculo de simetria	107
Figura 3.15 - Exemplo de cálculo da densidade de sub-regiões	107
Figura 3.16 - Separação e posicionamento de carac- teres	109
Figura 3.17 - Redução de um padrão 3x3 para 2x2	110
Figura 3.18 - Algoritmo de alteração de escala	110
Figura 4.1 - Aparência externa de um neurônio	116
Figura 4.2 - Modelo do <i>Psychon</i> de McCulloch & Pitts	119
Figura 4.3 - Topologias de interconexões em redes neurais	121
Figura 4.4 - Estrutura de um neurônio artificial	126
Figura 4.5 - Taxonomia das redes artificiais de neurônios	129
Figura 4.6 - Funções de transferência	130
Figura 4.7 - Exemplo de uso do Perceptron	132
Figura 4.8 - Esquema do neurônio tipo Adaline	134
Figura 4.9 - Algoritmo de aprendizado Delta Rule	135
Figura 4.10 - Níveis de uma rede versus regiões de decisão	138
Figura 4.11 - Solução do problema do XOR	139
Figura 4.12 - Funcionamento do modelo do ART	144
Figura 4.13 - Organização de uma rede de Hopfield	146
Figura 4.14 - Utilização da rede de Hopfield	146
Figura 4.15 - Interação entre os neurônios do modelo de Hopfield	147
Figura 4.16 - Algoritmos da rede de Hopfield	148
Figura 4.17 - Processo de codificação de uma BAM	151

Figura 4.18 - Organização de uma rede neural de Kohonen	153
Figura 4.19 - Agrupamento da atividade das saídas de uma matriz bidimensional de neurônios	154
Figura 4.20 - Exemplos de topologias de vizinhança	155
Figura 5.1 - Esquema de funcionamento do sistema N ² OCR	161
Figura 5.2 - Protótipo do sistema N ² OCR	162
Figura 5.3 - Tela de ajuda do módulo gerador de imagens	165
Figura 5.4 - Configuração do módulo gerador de imagens	166
Figura 5.5 - Padrão do caractere 'a' gerado sem distorção	168
Figura 5.6 - Padrão do caractere 'a' gerado com a inversão aleatória de 5% dos pixels	169
Figura 5.7 - Padrão do caractere 'a' gerado com a inversão aleatória de 10% dos pixels	170
Figura 5.8 - Padrão do caractere 'a' gerado com a rotação de 10 graus	171
Figura 5.9 - Tela de ajuda do módulo de tratamento de imagens	173
Figura 5.10 - Conversão dos formatos de arquivos	174
Figura 5.11 - Visualização de uma imagem de um texto ..	175
Figura 5.12 - Visualização de uma imagem com halftoning ..	176
Figura 5.13 - Histograma da imagem da fig. 5.12 com halftoning desfeito	177
Figura 5.14 - Binarização da imagem cujo histograma aparece na fig. 5.13	178
Figura 5.15 - Convolução para detecção de bordas na imagem da fig. 5.14	179
Figura 5.16 - Tela de ajuda do módulo de preparação de dados	181
Figura 5.17 - Imagem exibida pelo módulo de preparação de dados	182
Figura 5.18 - Exibição de outra área da imagem da fig. 5.17 (deslocamento interativo)	183
Figura 5.19 - Parâmetros de configuração da segmentação	185

Figura 5.20 - Primeiro caractere isolado da imagem	186
Figura 5.21 - Segundo caractere isolado da imagem	187
Figura 5.22 - Terceiro caractere isolado da imagem	188
Figura 5.23 - Rotulação de caracteres	190
Figura 5.24 - Configuração do processo de rotulação ...	192
Figura 5.25 - Caractere isolado com código de identificação	193
Figura 5.26 - Caractere isolado com rótulo de desconhecido	194
Figura 5.27 - Ajuste de escala e posição	196
Figura 5.28 - Padrão com ajuste de escala 10 x 10	197
Figura 5.29 - Padrão com ajuste de escala 16 x 16	198
Figura 5.30 - Padrão com ajuste de escala 20 x 20	199
Figura 5.31 - Organização da rede neural do modelo ADAn-LIENE	203
Figura 5.32 - Expansão da rede neural	206
Figura 5.33 - Tabela de controle de aprendizado	210
Figura 5.34 - Proposta simplificada de uma arquitetura ..	218
Figura 5.35 - Avaliação automática do reconhecimento ..	226
Figura 5.36 - Recodificação de um padrão	230

LISTA DE TABELAS

Tabela 3.1 - Taxas de reconhecimento obtidas por diferentes sistemas OHCR	86
Tabela 3.2 - Comportamento das técnicas de reconhecimento	90
Tabela 3.3 - Erros de substituição mais comuns	92
Tabela 5.1 - Funções previstas no sistema N ² OCR e funções implementadas no protótipo	158

RESUMO

Este trabalho apresenta um estudo sobre o reconhecimento visual de caracteres através da utilização das redes neurais. São abordados os assuntos referentes ao Processamento Digital de Imagens, aos sistemas de reconhecimento de caracteres, e às redes neurais. Ao final é apresentada uma proposta de implementação de um sistema OCR orientado ao reconhecimento de caracteres impressos, que utiliza uma rede neural desenvolvida especificamente para esta aplicação. O sistema proposto, que é denominado de sistema N²OCR, possui um protótipo implementado, que também é descrito neste trabalho.

Em relação ao Processamento Digital de Imagens são apresentados diversos temas, abrangendo os assuntos referentes à aquisição de imagens, ao tratamento das imagens e ao reconhecimento de padrões. A respeito da aquisição de imagens são destacados os aspectos referentes aos dispositivos de aquisição e os tipos de imagens obtidas através destes. Sobre o tratamento de imagens são abordados os aspectos referentes a imagens textuais, incluindo: halftoning, geração e modificação de histograma, limiarização e operações de filtragem. Quanto ao reconhecimento de padrões é feita uma breve análise das técnicas relacionadas a este tema.

Os diversos tipos de sistemas de reconhecimento de caracteres são abordados, assim como as técnicas e algoritmos empregados por estes. Além destes tópicos é apresentada uma discussão a respeito da avaliação dos resultados obtidos por estes sistemas, assim como é feita uma análise das principais dificuldades enfrentadas por estas aplicações.

Neste trabalho é feita uma apresentação a respeito das redes neurais, suas características, histórico e evolução das pesquisas nesta área. É feita uma descrição dos principais modelos de redes neurais em destaque na atualidade: Perceptron, Adaline, Madaline, redes multinível, ART, modelo de Hopfield, máquina de Boltzmann, BAM e modelo de Kohonen.

A partir da análise dos diferentes modelos de redes neurais empregados na atualidade, chega-se a proposta de um novo modelo de rede a ser utilizado pelo sistema N²OCR. São descritos os itens referentes ao aprendizado, ao reconhecimento e as possíveis extensões deste novo modelo. Também é abordada a possibilidade de implementação de um hardware dedicado para este modelo.

No final deste trabalho é fornecida uma visão global do sistema N²OCR, descrevendo cada um de seus módulos. Também é feita uma descrição do protótipo implementado e de suas funções.

Palavras-chave: Processamento de Imagens, Sistemas OCR, Reconhecimento de Caracteres, Redes Neurais.

ABSTRACT

This work presents a study of visual character recognition using neural networks. It describes some aspects related to Digital Image Processing, character recognition systems and neural networks. The implementation proposal of one OCR system, for printed character recognition, is also presented. This system uses one neural network specifically developed for this purpose. The OCR system, named N²OCR, has a prototype implementation, which is also described.

Several topics related to Digital Image Processing are presented, including some referent to image acquisition, image processing and pattern recognition. Some aspects on image acquisition are treated, like acquisition equipments and kinds of image data obtained from those equipments. The following items about text image processing are mentioned: halftoning, histogram generation and alteration, thresholding and filtering operations. A brief analysis about pattern recognition related to this theme is done.

Different kinds of character recognition systems are described, as the techniques and algorithms used by them. Besides, a discussion about performance estimation of this OCR systems is done, including typical OCR problems description and analysis.

In this work, neural networks are presented, describing their characteristics, historical aspects and research evolution in this field. Different famous neural network models are described: Perceptron, Adaline, Madaline, multi-level networks, ART, Hopfield's model, Boltzmann machine, BAM and Kohonen's model.

From the analysis of such different neural network models, we arrive to a proposal of a new neural network model, where are described items related to learning, recognition and possible model extensions. A possible hardware implementation of this model is also presented.

A global vision of N²OCR system is presented at the end of this work, describing each of its modules. A description of the prototype implementation and functions is also provided.

Key-words: Image Processing, OCR Systems, Character Recognition, Neural Networks.