

**LUCIANE YANASE HIRABARA**

**RECONHECIMENTO DE CARACTERES  
MANUSCRITOS BASEADO NA SELEÇÃO  
DINÂMICA DE MECANISMOS DE  
ZONEAMENTO**

Dissertação apresentada ao Programa de Pós-Graduação em Informática Aplicada da Pontifícia Universidade Católica do Paraná como requisito parcial para obtenção do título de Mestre em Informática.

**CURITIBA**

**2011**

**LUCIANE YANASE HIRABARA**

**RECONHECIMENTO DE CARACTERES  
MANUSCRITOS BASEADO NA SELEÇÃO  
DINÂMICA DE MECANISMOS DE  
ZONEAMENTO**

Dissertação apresentada ao Programa de Pós-Graduação em Informática Aplicada da Pontifícia Universidade Católica do Paraná como requisito parcial para obtenção do título de Mestre em Informática.

Área de Concentração: *Visão Computacional, Reconhecimento de padrões e Computação Forense*

Orientador: Prof. Dr. Cinthia Obladen de Almendra Freitas

Co-orientador: Prof. Dr. Alceu de Souza Britto Junior

**CURITIBA**

**2011**

Dados da Catalogação na Publicação  
Pontifícia Universidade Católica do Paraná  
Sistema Integrado de Bibliotecas – SIBI/PUCPR  
Biblioteca Central

H668r 2011	<p>Hirabara, Luciane Yanase</p> <p>Reconhecimento de caracteres manuscritos baseado na seleção dinâmica de mecanismos de zoneamento / Luciane Yanase Hirabara ; orientador, Cinthia Obladen de Almendra Freitas ; co-orientadores, Alice de Souza Britto Junior. – 2011. xiii, 88 f. : il. ; 30 cm</p> <p>Dissertação (mestrado) – Pontifícia Universidade Católica do Paraná, Curitiba, 2011 Bibliografia: f. 79-84</p> <p>1. Sistemas de reconhecimento de padrões. 2. Redes neurais. 3. Algoritmos. I. Freitas, Cinthia Obladen de Almendra. II. Britto Junior, Alceu de Souza. III. Pontifícia Universidade Católica do Paraná. Programa de Pós-Graduação em Informática. IV. Título.</p> <p>CDD 20. ed. – 006.4</p>
---------------	--

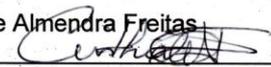
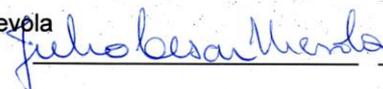


Pontifícia Universidade Católica do Paraná  
Centro de Ciências Exatas e de Tecnologia  
Programa de Pós-Graduação em Informática

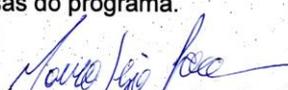
ATA DE DEFESA DE DISSERTAÇÃO DE MESTRADO  
PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

DEFESA DE DISSERTAÇÃO Nº 11/2011

Aos 23 dias do mês de Agosto de 2011 realizou-se a sessão pública de Defesa da Dissertação "**Reconhecimento de Carcateres Manuscritos Baseado na Seleção Dinâmica de Mecanismos de Zoneamento**" apresentada pela aluna **Luciane Yanase Hirabara**, como requisito parcial para a obtenção do título de Mestre em Informática, perante uma Banca Examinadora composta pelos seguintes membros:

Profª. Drª. Cinthia Obladen de Almeida Freitas  
PUCPR (Orientadora)  APROVADA  
(assinatura) (aprov/reprov.)  
Prof. Dr. Alceu de Souza Britto Junior  APROV.  
PUCPR  
Prof. Dr. Julio Cesar Nievoia  APROVADA  
PUCPR  
Prof. Dr. Sérgio Scheer  APROV  
UFPR

Conforme as normas regimentais do PPGIa e da PUCPR, o trabalho apresentado foi considerado APROVADO (aprovado/reprovado), segundo avaliação da maioria dos membros desta Banca Examinadora. Este resultado está condicionado ao cumprimento integral das solicitações da Banca Examinadora registradas no Livro de Defesas do programa.

  
Prof. Dr. Mauro Sérgio Pereira Fonseca  
Diretor do Programa de Pós-Graduação em Informática



Aos pais Orlando e Celina,  
as minhas irmãs Cris e Bel  
e ao meu marido Sakae,  
com amor...

## Agradecimentos

*“Existem pessoas cometa e pessoas estrela. Cometas são aquelas que só passam. As estrelas permanecem. ... estrelas são aquelas com os quais podemos contar, que estão sempre presentes quando precisamos, são coragem nos momentos difíceis, são luz nos momentos escuros, são segurança nos momentos de desânimo. ...”*(Autor desconhecido)

Agradeço por ter tantas pessoas estrela em minha vida.

Eternos agradecimentos aos meus pais, Celina e Orlando pelo amor e apoio incondicional, e as minhas irmãs, Cristiane e Elizabete pelo apoio e carinho, por estarem sempre presentes quando preciso, por incentivarem a prosseguir nesta jornada.

Agradeço a meu Marido Sakae, pela paciência e compreensão, por ser a segurança nos momentos de desânimo.

Agradeço a Prof. Dr. Cinthia, pelos ensinamentos e dedicação, por ser a estrela guia nesta jornada. Ao Prof. Dr. Alceu pelos ensinamentos, pela dedicação e atenção dispensadas.

Agradeço a meus familiares: tios, tias, primos, primas, em especial aos tios Nilo e Elaine, pelo apoio e por tornarem mais confortável esta jornada.

Agradeço a colega Simone e ao colega Jackson, pelo auxílio e atenção dispensados, e também aos colegas Marcelo e Fabio, que me auxiliaram no início desta jornada.

Aos professores e funcionários do PPGIA que direta ou indiretamente contribuíram para a execução desse trabalho.

# Sumário

<b>Agradecimentos</b> .....	i
<b>Sumário</b> .....	ii
<b>Lista de Figuras</b> .....	iv
<b>Lista de Tabelas</b> .....	vii
<b>Lista de Símbolos</b> .....	ix
<b>Lista de Abreviaturas</b> .....	x
<b>Resumo</b> .....	xii
<b>Abstract</b> .....	xiii
<b>Capítulo 1</b>	
<b>Introdução</b> .....	<b>1</b>
<b>1.1. Descrição do Problema</b> .....	1
<b>1.2. Objetivos</b> .....	3
<b>1.3. Desafios</b> .....	5
<b>1.4. Motivação</b> .....	5
<b>1.5. Organização da Dissertação</b> .....	6
<b>Capítulo 2</b>	
<b>Fundamentação Teórica</b> .....	<b>7</b>
<b>2.1. Reconhecimento de Manuscritos</b> .....	7
<b>2.2. Reconhecimento de Caracteres Manuscritos</b> .....	8
<b>2.3. Extração de Primitivas</b> .....	9
<b>2.4. Mecanismos de Zoneamento</b> .....	10
<b>2.5. Métodos de Classificação</b> .....	11
2.5.1. Redes Neurais Artificiais – RNA .....	12
2.5.1.1 Treinamento das Redes Neurais .....	13
2.5.1.2. MLP Convencional e Classe Modular .....	15

<b>2.6. Trabalhos Relacionados .....</b>	<b>18</b>
<b>2.7. Seleção Dinâmica.....</b>	<b>32</b>
<b>2.8. Considerações Finais .....</b>	<b>33</b>

### Capítulo 3

<b>Método Reconhecimento de Caracteres Manuscritos com Seleção Dinâmica de Mecanismos de Zoneamento .....</b>	<b>34</b>
<b>3.1. Visão Geral .....</b>	<b>34</b>
3.1.1. Pré-processamento .....	36
3.1.2. Extração de Primitivas .....	37
3.1.3. Classificação.....	40
<b>3.2. Seleção Dinâmica de Mecanismos de Zoneamento – Nível 01.....</b>	<b>41</b>
<b>3.3. Reconhecimento de Caracteres Manuscritos Minúsculos – Nível 02.....</b>	<b>46</b>
<b>3.4. Considerações Finais.....</b>	<b>48</b>

### Capítulo 4

<b>Resultados Experimentais.....</b>	<b>49</b>
<b>4.1. Experimentos com RNA-MLP Convencional.....</b>	<b>49</b>
<b>4.2. Experimentos com RNA MLP Convencional – Z4 – Nível 01.....</b>	<b>51</b>
<b>4.3. Experimentos com RNA MLP Classe - Modular – Nível 02.....</b>	<b>52</b>
<b>4.4. Reconhecimento de Caracteres Manuscritos com Seleção Dinâmica de Mecanismos de Zoneamento – Nível 01 + Nível 02.....</b>	<b>54</b>
<b>4.5. Análise dos Resultados.....</b>	<b>56</b>
4.5.1. Evolução dos Classificadores: RNA MLP Convencional e Classe-Modular.....	56
4.5.2. Análise das Matrizes de Confusão.....	58
4.5.3. Caracteres Minúsculos e Caracteres Maiúsculos.....	70
4.5.4. Abordagem pela Seleção Dinâmica de Zoneamento.....	78
<b>4.5. Considerações Finais.....</b>	<b>81</b>
<b>Conclusão.....</b>	<b>82</b>
<b>Referências Bibliográficas.....</b>	<b>85</b>

## Lista de Figuras

Figura 1.1	Diversidade de estilos de escrita – base de dados IRONOFF - caracteres minúsculos	2
Figura 1.2	Similaridade entre caracteres distintos - base de dados IRONOFF – caracteres minúsculos	2
Figura 1.3	Diferença entre caracteres maiúsculos e minúsculos - base de dados IRONOFF	6
Figura 2.1	Exemplo de formas manuscritas: a) caracteres Isolados e b) palavras	7
Figura 2.2	Esquema geral de um sistema de reconhecimento de manuscritos	9
Figura 2.3	Modelo de um Perceptron com D entradas e uma saída ([NIE04] – adaptado)	15
Figura 2.4	Arquitetura MLP com uma camada escondida ([NIE04] - adaptado)	15
Figura 2.5	Arquitetura de uma Rede MLP Classe-Modular: a) Subrede e b) Rede Classe-Modular completa	17
Figura 2.6	Diagramas esquemáticos de 2-classificador para uma classe $\omega_i$ ([OH01] - adaptado)	18
Figura 2.7	Representação das partições, os números 1, 2, 3, 4, 5, 6 representam o número de posições das partes [SUE94]	18
Figura 2.8	Os tipos de partições propostas para uma letra K [SUE94]	19
Figura 2.9	Exemplos de dois caracteres com um zoneamento 2LR	19
Figura 2.10	Exemplo de nomes de meses (base UFCG) [OLI02]	20
Figura 2.11	Exemplo de segmentação implícita [OLI02]	21
Figura 2.12	Zonas de Extração de primitivas [OLI02]	21
Figura 2.13	Pontos de fixação e suas respectivas janelas de atenção [COR02]	21
Figura 2.14	Template de codificação individual ([RAD03] - adaptado)	22
Figura 2.15	Estratégias de zoneamento ([RAD03] - adaptado)	23
Figura 2.16	Extração de primitivas: a) linha original, b) linha no arquivo binário, c) depois de distinguidas as direções e d) depois da normalização das direções	24

Figura 2.17	Amostras referentes aos valores por extenso (base PUCPR) [KAP04]	25
Figura 2.18	Mecanismo de zoneamento para um nome de mês [KAP04]	26
Figura 2.19	Esquema de zoneamento implícito pela combinação do modelo baseado em colunas e linhas [BRI04]	27
Figura 2.20	Zoneamento não simétrico Z5V, Z5H e Z7, respectivamente	28
Figura 2.21	Zoneamento baseado em regiões perceptivas: a) Z5H e b) Z7 [FRE07]	28
Figura 2.22	Meta-classes geradas a partir do conjunto de validação [AIR05]	29
Figura 2.23	a) Esquema de seleção estática b) e seleção dinâmica ([KO08] – adaptado)	33
Figura 3.1	Visão geral do método proposto	35
Figura 3.2	Amostras de Imagem Original (base IRONOFF)	36
Figura 3.3	Limiarização: a) imagem original e b) imagem resultante - método de Otsu	36
Figura 3.4	Exemplo de Bounding Box ao redor do caractere	37
Figura 3.5	Direções de busca (NSLO) e a sondagem para verificação de escape abaixo e direita	37
Figura 3.6	Exemplo de extração local de primitivas para um caractere “n”, e o vetor de características extraído para a primeira região do zoneamento 5H	39
Figura 3.7	Amostras de caracteres distintos e suas particularidades	40
Figura 3.8	Caracteres e suas particularidades representadas através da extração de primitivas e dos mecanismos de zoneamento	40
Figura 3.9	Esquema de seleção dinâmica de mecanismo de zoneamento	42
Figura 3.10	Zoneamento Global, 4, 5H, 5V e 7 para um caractere “e”	44
Figura 3.11	Criação das Matrizes de confusão	44
Figura 3.12	Obtenção dos 3 maiores valores de saída	45
Figura 3.13	Obtenção dos 6 valores de confusão a partir da matriz de confusão	45
Figura 3.14	Indicação do mecanismo de zoneamento	46
Figura 3.15	Esquema de subrede para o caractere “o”	47
Figura 4.1	Esquema de indicação do melhor zoneamento para um caractere desconhecido	52
Figura 4.2	Exemplo de caracteres que se confundiram	59

Figura 4.3	Exemplo de diversidade de estilos de escrita para caracteres “r” e “n”	60
Figura 4.4	Similaridade entre caracteres distintos – Z4	62
Figura 4.5	Amostras de caracteres “c”, “d” e “q”	63
Figura 4.6	Confusão entre caracteres distintos – Z5H	65
Figura 4.7	Exemplo de caracteres sem simetria horizontal	66
Figura 4.8	Similaridade e confusão entre caracteres distintos - Z5V	67
Figura 4.9	Confusão entre caracteres – Z7	69
Figura 4.10	Confusão entre caracteres: a) “i” minúsculo e “T” maiúsculo; b) caracteres “b” minúsculo e “D” maiúsculo; c) caracteres “t” minúsculo e “E” maiúsculo	72
Figura 4.11	Exemplos de estilos de escrita e simetria de caracteres.	73
Figura 4.12	Extração de Primitivas - Caracteres “H” e “h”	73
Figura 4.13	Caracteres incorretamente reconhecidos	81

## Lista de Tabelas

Tabela 2.1	Características de RNAs em função do tipo de aprendizagem [JAI96]	14
Tabela 2.2	Taxas de erro ([RAD03] - adaptado)	23
Tabela 2.3	Taxas de reconhecimento para cada método de seleção dinâmica de conjunto de classificadores ([KO08] - adaptado)	30
Tabela 2.4	Resumo de autores e trabalhos	31
Tabela 3.1	Rótulos dos pixels do fundo da imagem	38
Tabela 4.1	Classe e reconhecimento por zoneamento – RNA MLP Convencional	51
Tabela 4.2	Classe e reconhecimento por zoneamento – RNA MLP Classe-Modular	54
Tabela 4.3	Taxa de Reconhecimento - Zoneamento dinâmico	55
Tabela 4.4	Reconhecimento MLP Convencional e Classe-Modular	57
Tabela 4.5	Matriz de Confusão Z=9 MLP Classe Modular	59
Tabela 4.6	Matriz de Confusão Z=4 MLP Classe Modular	62
Tabela 4.7	Matriz de Confusão Z=5H MLP Classe Modular	65
Tabela 4.8	Matriz de Confusão Z=5V MLP Classe Modular	67
Tabela 4.9	Matriz de Confusão Z=7 MLP Classe Modular	69
Tabela 4.10	Quantidade de Amostras Utilizadas nos Experimentos	70
Tabela 4.11	Reconhecimento para caracteres maiúsculos e minúsculos	71
Tabela 4.12	Melhor Zoneamento para cada Classe de Caractere	74
Tabela 4.13	Mecanismos de Zoneamento e Confusões entre Classes	75
Tabela 4.14	Matriz de Confusão Z4 – MLP Classe-Modular: Maiúsculas e Minúsculas	75
Tabela 4.15	Matriz de Confusão Z5H – MLP Classe-Modular: Maiúsculas e Minúsculas	76

Tabela 4.16	Matriz de Confusão Z5V – MLP Classe-Modular: Maiúsculas e Minúsculas	76
Tabela 4.17	Matriz de Confusão Z7 – MLP Classe-Modular: Maiúsculas e Minúsculas	77
Tabela 4.18	Trabalhos Relacionados (base de dados: NIST e IRONOFF)	78
Tabela 4.19	Tabela comparativa: Zoneamentos individuais e seleção dinâmica	79
Tabela 4.20	Matriz de Confusão - Zoneamento selecionado dinamicamente	80

## Lista de Símbolos

$2-C(\omega_i)$	<i>2-classificador para uma classe <math>\omega_i</math>,</i>
$\omega_i$	<i>Classe de um problema</i>
$C_i$	<i>Classificador</i>
$\Omega_1$	<i>Conjunto de dados da classe <math>\omega_1</math></i>
$\Omega_0$	<i>Conjunto de dados da classe restante</i>
$x_i$	<i>Entrada de RNA</i>
$M_{\omega_i}$	<i>Modulo de classe <math>\omega_i</math></i>
$\underline{X}$	<i>Padrão de entrada</i>
$R_i$	<i>Primitiva</i>
$y_i$	<i>Saída de RNA</i>
$O_i$	<i>Saída de RNA</i>
$D$	<i>Vetor de probabilidade de uma amostra pertencer à classe <math>\Omega_0</math> ou <math>\Omega_1</math></i>

## Lista de Abreviaturas

CAS	<i>Conjunto de dados Automaticamente Segmentados</i>
CEDAR	<i>Center of Excellence for Document Analysis and Recognition</i>
DbD	<i>Distance-based Disagreement</i>
EoC	<i>Ensemble of Classifier</i>
EoZ	<i>Ensemble of Zonning</i>
KNN	<i>K –Nearest Neighbor</i>
KNORA	<i>K Nearest Oracle</i>
K-E	<i>KNORA Eliminate</i>
K-E-W	<i>KNORA Eliminate-W</i>
K-U	<i>KNORA Union</i>
K-U-W	<i>KNORA Union - W</i>
LCA	<i>Local Class Acuracy</i>
LR	<i>Left-Rigth</i>
MEM	<i>Modelos Escondidos de Markov</i>
MLP	<i>Multilayer Perceptron</i>
MCS	<i>Sistemas Multi-Classificadores (Multiple Classifiers System)</i>
MOEA	<i>Multi-Objective Evolutionary Algorithm</i>
NIST	<i>National Institute of Standards and Technology</i>
NSLO	<i>Norte-Sul-Leste-Oeste</i>
OLA	<i>Overall local accuracy,</i>
PUCPR	<i>Pontificia Universidade Católica do Paraná</i>
REC	<i>Reconhecimento</i>
UFMG	<i>Universidade Federal de Campina Grande</i>
RNA	<i>Redes Neurais Artificiais</i>
UD	<i>Up-Down</i>
Z	<i>Zoneamento</i>
Z4	<i>Zoneamento simétrico com 4 partições</i>
Z5H	<i>Zoneamento não simétrico com 5 partições Horizontais</i>

Z5V	<i>Zoneamento não simétrico com 5 partições Verticais</i>
Z7	<i>Zoneamento não simétrico com 7 partições</i>
Z9	<i>Zoneamento não simétrico com 9 partições</i>

## Resumo

Os estudos e pesquisas na área de reconhecimento de caracteres manuscritos vêm crescendo muito nos últimos tempos. Muitos métodos de extração de primitivas e diferentes métodos de zoneamento (*zoning*), bem como, de classificação têm sido avaliados permitindo a obtenção de resultados animadores. O presente trabalho buscou por meio da investigação e combinação de métodos existentes, incorporar aspectos da percepção humana no processo de reconhecimento de caracteres manuscritos, visando contribuir com a área de reconhecimento para caracteres manuscritos. Assim, é utilizada uma solução híbrida para o problema de reconhecimento de caracteres manuscritos minúsculo, utilizando extração Global de primitivas (baseada em concavidades e convexidades) combinada com uma extração Local (baseada em mecanismo de zoneamento perceptivo, não-simétrico). Tal método de extração de primitivas já foi avaliado e comprovadamente obteve bom desempenho para o reconhecimento de caracteres manuscritos maiúsculo. Investigou-se neste trabalho um grupo de 04 diferentes mecanismos de zoneamento com  $Z$  regiões, sendo  $Z = 4$ ,  $Z = 5H$ ,  $Z = 5V$  e  $Z = 7$ . A partir do estudo destes mecanismos foi proposto um método de seleção dinâmica do mecanismo mais adequado a cada uma das classes de caracteres a serem reconhecidas. Como método de classificação utilizou-se Redes Neurais MLP Convencional e Classe-Modular. Os experimentos realizados com cada mecanismo de zoneamento permitiram alcançar as seguintes taxas de reconhecimento:  $Z4$ : 81,10%,  $Z5H$ : 80,81%,  $Z5V$ : 81,63%,  $Z7$ : 83,60%. A partir dos resultados obtidos, propôs-se um método de reconhecimento de caracteres manuscritos minúsculos composto por dois níveis. No Nível 01, é realizada a seleção dinâmica do mecanismo de zoneamento mais adequado a ser utilizado para o reconhecimento de cada amostra de caractere manuscrito minúsculo, utilizando uma Rede Neural Artificial com arquitetura Convencional, visando obter uma classificação *a priori* do melhor mecanismo de zoneamento para cada classe de caractere. O Nível 02 corresponde à fase de reconhecimento propriamente dito, ou seja, depois de corretamente identificado o melhor zoneamento para o caractere em questão por meio da aplicação do Nível 01 é utilizado então, uma RNA MLP Classe-Modular. Os experimentos assim realizados permitiram alcançar taxas de reconhecimento de 88,13%, mostrando ser um enfoque promissor. A partir dos experimentos pode-se verificar que cada classe de caractere pode ter maior taxa de reconhecimento por zoneamentos distintos, sendo que não há predominância de um único mecanismo de zoneamento para o problema de reconhecimento de caracteres manuscritos minúsculo. Conclui-se assim que a utilização de um único mecanismo de zoneamento, mas sim, de um conjunto deles, pode ser fundamental para a melhoria das taxas de reconhecimento, visto ser possível escolher dinamicamente e *a priori* o melhor mecanismo de zoneamento para cada classe de caractere a ser tratada pelo sistema.

**Palavras-Chave:** Reconhecimento de Caracteres Manuscritos, Mecanismos de Zoneamento, Seleção Dinâmica, Redes Neurais.

## Abstract

The studies and research in handwritten characters recognition area has been increasing in recent times. Many feature extraction methods and different methods of zoning, as well as the classification have been assessed capable of producing encouraging results. This work sought through research and combination of existing methods, incorporating aspects of human perception in the process of handwritten characters recognition in order to contribute for handwritten characters recognition area. Thus, a hybrid solution is used to the problem of lowercase handwritten characters recognition, using global feature extraction (based on concavity and convexity) combined with local extraction (based on non-symmetrical perceptual zoning mechanism). This feature extraction method has already been evaluated and demonstrated good performance obtained for uppercase handwritten characters recognition. In this work is Investigated a group of 4 different zoning mechanisms with Z regions: Z = 4, Z = 5H, 5V and Z = Z = 7. From the study of these mechanisms has been proposed a dynamic selection method of the most appropriate mechanism for each class of characters to be recognized. A Conventional MLP Neural Networks and Class-Modular MLP Neural Networks are used for classification. The experiments carried out with each zoning mechanism had achieved the following recognition rates: Z4: 81.10% Z5H: 80.81% Z5V: 81.63% Z7: 83.60%. From the results obtained, a two-level method for handwritten characters recognition was proposed. At Level 01, Conventional Artificial Neural Network architecture is used to obtain a priori classification of the best zoning mechanism for each character class. In this first level is done a dynamic selection of the most appropriate zoning mechanism to be used for recognition of each handwritten character sample. The Level 02 corresponds to the recognition phase itself, i.e., after the Level 01 having correctly identified the best zoning for the character, a RNA-Class Modular MLP is used. The experiments made possible to achieve recognition rates of 88.13%, showing to be a promising approach. From the experiments we can observe that each character class has higher recognition rate for different zoning, and there is no predominance of a single zoning mechanism to the problem of lowercase handwritten characters recognition. Concluded that the use of a single zoning mechanism, but rather a ensemble of them, may be a key to improving rates of recognition, since it is possible to dynamically and a priori choose the best zoning mechanism for each character class to be evaluate by the system.

**Keywords:** Handwritten Character Recognition, Zoning Mechanism, Dynamic Selection, Neural Network.

# Capítulo 1

## Introdução

Devido ao alto grau de dificuldade em reproduzir a capacidade humana de ler palavras e caracteres manuscritos, muitos pesquisadores têm se dedicado ao estudo de métodos e técnicas visando o desenvolvimento de sistemas para o reconhecimento automático de manuscritos.

Pesquisadores se dedicam a estudos nessa área, utilizando diferentes análises, abordagens e combinações de técnicas envolvendo diferentes estratégias. Isto devido ao fato de que documentos manuscritos apresentam componentes de grande complexidade, tais como: diferentes estilos de escrita, contexto da escrita, contexto da aquisição do documento (*on-line* - dinâmico ou *off-line* - estático) e ainda, o tamanho do léxico a ser reconhecido e o número de escritores [FRE01].

Assim, o presente trabalho tem como área de pesquisa os documentos manuscritos, visando contribuir no desenvolvimento de sistemas automáticos de reconhecimento de caracteres manuscritos, explorando as potencialidades que métodos já propostos oferecem, apresentando assim uma nova abordagem de reconhecimento de caracteres utilizando seleção dinâmica de mecanismos de zoneamento como elementos complementares da extração de primitivas.

### 1.1. Descrição do problema

O fato do caractere se apresentar na forma manuscrita torna o problema de reconhecimento de caracteres ainda mais complexo devido aos fatores já citados anteriormente (estilos de escrita, contexto e aquisição do documento, etc.). A Figura 1.1

apresenta amostras de caracteres, retiradas na base de dados IRONOFF [GAU99]. Apesar de cada linha apresentar um mesmo caractere (“a”, “b” ou “c”), cada coluna pertence a um escritor em particular, observa-se, portanto, a conseqüente diversidade de estilos de escrita. Por diversidade de estilos entende-se o estilo pessoal de cada escritor, ou seja, a maneira como cada escritor representa cada caractere, pois sabe-se que cada escritor contém um conjunto de características que o diferencia dos demais escritores.

Os caracteres possuem diferentes particularidades, como por exemplo: laços, traços verticais e horizontais, entre outras, e, assim, o método de extração de primitivas deve ser capaz de representar eficientemente as características particulares de cada classe de caractere para que então o método de classificação selecionado possa reconhecê-los.

Além das diferenças entre os caracteres, a similaridade entre caracteres diferentes é outro grande problema na área de reconhecimento de caracteres. Deste modo, identificar em quais partes dos caracteres se encontra essa similaridade é outro fator importante. A Figura 1.2 ilustra a similaridade entre caracteres distintos. Estes exemplos ilustram similaridades já esperadas, pois uma maneira de melhorar os sistemas de reconhecimento automático é entender as similaridades entre as formas a reconhecer.

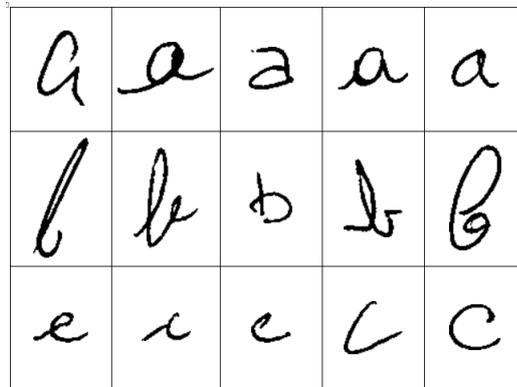


Figura 1.1: Diversidade de estilos de escrita – base de dados IRONOFF - caracteres minúsculos

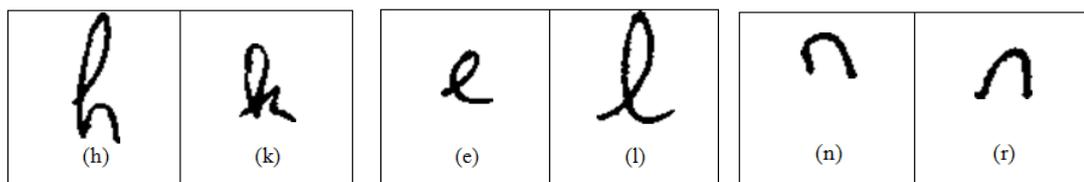


Figura 1.2: Similaridade entre caracteres distintos - base de dados IRONOFF – caracteres minúsculos

## 1.2. Objetivos

Existe um abismo entre as capacidades de leitura humana e os sistemas de reconhecimento automático da escrita humana. De acordo com a literatura, é necessário explorar e capturar informações da percepção humana para projetar novos sistemas [FRE01].

Neste trabalho, explora-se o uso de um conjunto de mecanismos de zoneamento perceptivos, aqui denominado de EoZ (*Ensemble of Zoning*). Uma abordagem analítica é utilizada considerando os 26 caracteres minúsculos do alfabeto latino, investigando informações das regiões perceptivas para o reconhecimento do caractere com base na percepção humana das formas. Este trabalho aborda os trabalhos já realizados por Freitas (2001) [FRE01], Aires (2005) [AIR05] e Freitas et al. (2007) [FRE07].

É então proposto e analisado um método de seleção dinâmica de mecanismos de zoneamento, utilizando o conceito de *Ensemble of Zoning* - EoZ, para predizer qual o melhor mecanismo de zoneamento para cada amostra de caractere. O sistema considera um método híbrido de extração de primitivas, associando um método Global (extração de concavidades e convexidades) a um método Local (baseado em diferentes mecanismos de zoneamento). Tais mecanismos de zoneamento são do tipo simétrico e não simétricos (perceptivo). Além disto, são utilizadas Redes Neurais Artificiais (RNA) com arquitetura MLP (*Multilayer Perceptron*) Convencional e Classe-Modular para o reconhecimento dos caracteres manuscritos.

Busca-se assim através da investigação e combinação de métodos existentes, que já foram avaliados e comprovadamente tiveram bons resultados, incorporar aspectos da percepção humana no processo de reconhecimento com o objetivo de se alcançar taxas de reconhecimento para caracteres manuscritos, iguais ou muito próximas das taxas alcançadas na literatura.

Nas fases do projeto, que engloba as etapas de pré-processamento, extração de primitivas, zoneamento, classificação e reconhecimento os seguintes objetivos específicos são:

- Utilizar os zoneamentos propostos por Aires (2005) [AIR05] e Freitas et al.(2007) [FRE07] para a extrair as características dos caracteres.
- Utilizar os arquivos gerados com vetores de características para caracteres manuscrito minúsculo com os zoneamentos  $Z=4$ ,  $Z=5H$ ,  $Z=5V$  e  $Z=7$  para as seguintes etapas:

- Configurar e combinar os modelos de Redes Neurais MLP com arquitetura convencional e Classe-Modular e utilizá-las para classificação e reconhecimento dos caracteres.
  - Utilizar as Redes Neurais MLP com arquitetura Classe-Modular, treinadas com os zoneamentos Z=4, Z5H, Z5V e Z7 e construir as matrizes de confusão a serem usadas no sistema proposto, composto por 2 níveis.
  - Utilizar a seleção dinâmica de zoneamento usando as Redes Neurais MLP com arquitetura convencional treinada com zoneamento Z=4 e indicar qual dos mecanismos de zoneamento reconhece uma determinada amostra de caractere minúsculo em uma base de teste – Nível 01.
  - Utilizar os modelos já configurados de Redes Neurais MLP com arquitetura Classe-Modular e os zoneamentos dinamicamente indicados na etapa anterior e então classificar e reconhecer os caracteres manuscrito minúsculo – Nível 02.
- Analisar os resultados:
    - Comparar os resultados obtidos neste trabalho com caracteres minúsculos com os obtidos por Aires (2005) [AIR05] de modo a confirmar se os zoneamentos propostos são aplicáveis tanto para caracteres maiúsculos quanto para minúsculos, para tal comparação foram realizados experimentos também com caracteres maiúsculos + minúsculos.
    - Comparar todos os resultados obtidos para as duas arquiteturas de Rede Neural, bem como, para os diferentes mecanismos de zoneamento propostos por Aires (2005) [AIR05].
    - Verificar quais classes tem melhor reconhecimento em função dos mecanismos de zoneamentos avaliados.
    - Verificar se a abordagem pela seleção dinâmica de zoneamento tem melhor desempenho que para seleção estática.

### 1.3. Desafios

Diante da complexidade da tarefa de reconhecimento de caracteres manuscritos, um grande desafio é utilizar o reconhecimento com base na percepção humana de formas e identificar quais partes dos caracteres geram confusão. Outro desafio na área de reconhecimento é a escolha de um conjunto de primitivas capaz de representar com eficiência um determinado padrão e a decisão de qual o melhor método a ser utilizado para a classificação adequada de uma determinada classe. O presente trabalho teve como desafio dar continuidade aos trabalhos de Freitas (2001) [FRE01], Aires (2005) [AIR05] e Freitas et al. (2007) [FRE07] e ao mesmo tempo incorporar o conceito de conjunto (*ensemble*) e seleção dinâmica proposta de Ko et al. (2008) [KO08].

### 1.4. Motivação

Uma das motivações para o desenvolvimento desse trabalho é contribuir para a resolução dos problemas na área de reconhecimento de caracteres. Outra motivação é investigar se a seleção dinâmica de mecanismos de zoneamento pode ter comportamento análogo à seleção dinâmica de classificadores e se tais mecanismos quando dispostos em conjunto (EoZ) podem afetar de forma positiva o desempenho no reconhecimento de caracteres manuscritos.

Outra motivação está em aplicar o método já desenvolvido e testado para caracteres manuscritos maiúsculos em caracteres minúsculos, de modo a investigar se as regiões perceptivas e os mecanismos de zoneamento já avaliados para caracteres maiúsculos também são válidos para caracteres minúsculos. Isto posto, devido ao fato de que a escrita de caracteres maiúsculos e minúsculos diferem no traçado e formação dos caracteres propriamente ditos, independentemente do escritor. A Figura 1.3 mostra exemplos de caracteres maiúsculos e minúsculos ressaltando as diferenças de traçado.

Assim, a grande motivação é pesquisar para auxiliar no desenvolvimento de um método de reconhecimento que se aproxime ao máximo do ser humano em sua capacidade de ler um texto, uma palavra ou caractere manuscrito qualquer, sem restrição [FRE01], pois mesmo depois de muitos esforços de pesquisa ainda há uma lacuna entre as capacidades de leitura humana e os sistemas de reconhecimento.

## 1.5. Organização da Dissertação

Nas seções seguintes são apresentados os estudos referentes ao problema em questão, bem como, a abordagem utilizada e as técnicas aplicadas para o desenvolvimento do referido sistema. Este trabalho está organizado como se segue. No Capítulo 2 são apresentados os conceitos relacionados ao escopo da proposta e os trabalhos já realizados na área. No Capítulo 3 apresenta-se a descrição do sistema proposto e analisado, bem como, os métodos adotados. No Capítulo 4 são apresentados os experimentos realizados com os zoneamentos simétricos e não simétricos a partir de diferentes arquiteturas de Redes Neurais Artificiais. Além disto, apresenta-se os experimentos utilizando de uma RNA com arquitetura Convencional para seleção do melhor mecanismo de zoneamento. Este capítulo resume também a análise dos resultados experimentais. No Capítulo 5 é apresentada a conclusão e as possíveis contribuições do trabalho para a área de reconhecimento de caracteres manuscritos, bem como, os trabalhos futuros.

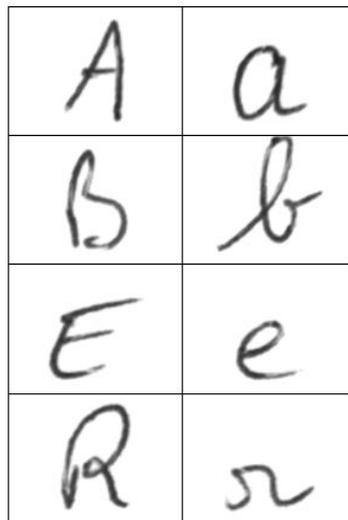


Figura1. 3: Diferença entre caracteres maiúsculos e minúsculos (base de dados IRONOFF)

## Capítulo 2

### Fundamentação Teórica

Este capítulo descreve os conceitos relacionados ao escopo da proposta e os trabalhos já realizados na área.

#### 2.1. Reconhecimento de Manuscrito

A escrita manuscrita pode se apresentar de duas formas: caracteres isolados (caracteres estão dispostos de forma não conectada) e palavras - caracteres dispostos com o objetivo de formar palavras, frases e textos com estilo completamente sem restrições, ou seja, podem ser conectados (estilo cursivo) ou desconectados (estilo letra de fôrma). A Figura 2.1 ilustra exemplos de caracteres isolados (a) e palavras(b).



Figura 2.1. Exemplo de formas manuscritas: a) caracteres Isolados e b) palavras

De um modo geral, um sistema de reconhecimento de manuscritos inclui tipicamente os estágios de extração de primitivas e classificação [HEU98], sendo compostas ainda de estágios de pré-processamento e segmentação.

A seguir é apresentada uma breve descrição dos estágios que compõem um sistema de reconhecimento de manuscrito:

- a) Pré-processamento é necessário para corrigir ou reduzir possíveis distorções na imagem;
- b) Segmentação explícita (a palavra é segmentada em caracteres) ou segmentação implícita (as palavras são tratadas como um todo);
- c) Extração de primitivas para representar as características particulares de cada palavra ou caractere, salientando as diferenças em relação a outras palavras ou caracteres;
- d) Classificação que atribui o caractere a uma das várias classes existentes.

De acordo com [FRE01a] e [FRE07] o reconhecimento de manuscritos pode ser tratado por uma Abordagem Local (*analytical approach* – em nível de caracteres isolados) ou por uma Abordagem Global (*global approach* – em nível das palavras, não sendo necessária a segmentação das palavras em caracteres).

Pode-se ainda classificar o campo de Reconhecimento de Manuscritos de várias maneiras, sendo que a mais simples é a distinção entre reconhecimento *on-line* (dinâmica) e *off-line* (estática) e esta classificação considera a natureza das aplicações [PLA00]. No reconhecimento *on-line* o usuário é conectado a um sistema usando uma caneta eletrônica ou *touch screen* e o reconhecimento é feito em tempo real; no reconhecimento *off-line* a escrita é capturada usando-se um processo de digitalização, sendo o texto, palavra, caractere ou dígito tratado como uma imagem. No reconhecimento *on-line* pode-se lançar mão de características dinâmicas da escrita, tais como: velocidade, pressão, progressão, sequência de traços, entre outras características. A Figura 2.2 ilustra um esquema básico para o reconhecimento de manuscritos.

## 2.2. Reconhecimento de Caracteres Manuscritos

No reconhecimento de caracteres manuscritos, os caracteres são tratados de forma isolada, podendo ser composto de letras, dígitos ou símbolos especiais. Considerando o alfabeto Latino as classes estão divididas em caracteres maiúsculos, minúsculos e dígitos.

As dificuldades relatadas no reconhecimento de caracteres manuscritos vêm do fato de que há caracteres que pertencem a diferentes classes, mas que apresentam formas muito similares. Se forem comparadas ao reconhecimento de caracteres manuscritos, o problema se torna maior, devido ao tamanho do léxico (26 para caracteres minúsculos e 26 para caracteres

maiúsculos). Muitas abordagens voltadas a resolver a questão do reconhecimento de caracteres manuscritos vêm sendo utilizadas pelos pesquisadores [KOE03] [AIR05] [FRE07] [XIA05].

Os caracteres manuscritos possuem diferentes particularidades entre si, e um mesmo caractere pode apresentar uma grande variabilidade na forma dependendo do estilo como se encontram escritos e por quem são escritos. Portanto, o fato de estarem na forma manuscrita torna o reconhecimento ainda mais difícil. O problema básico de reconhecimento de caractere é associar cada caractere a uma classe [PLA00]. Em sendo o problema em questão, caracteres isolados, a etapa de pré-processamento inclui somente a binarização (transformação da representação de 256 níveis de cinza para a representação binária, ou seja, branco/preto). Assim, não serão abordados e apresentados os métodos de pré-processamento necessários, por exemplo, no tratamento de palavras manuscritas, a saber: correção da inclinação da linha de base (*skew*), correção da inclinação vertical (*slant*), normalização do corpo de letras minúsculas, entre outros.

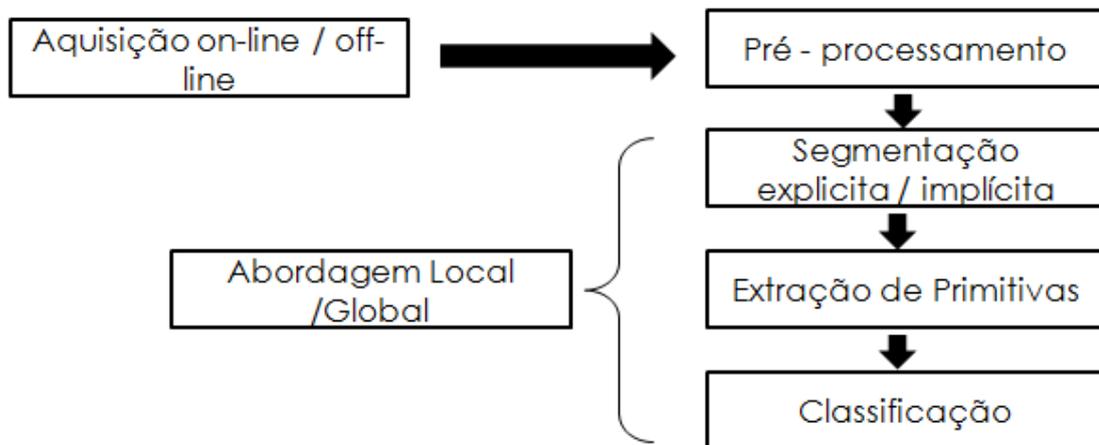


Figura 2.2: Esquema geral de um sistema de reconhecimento de manuscritos

### 2.3. Extração de Primitivas

A seleção de um método de extração de primitivas é provavelmente o fator mais importante para se obter alto desempenho nos sistemas de reconhecimento de caracteres [TRI96]. Assim, existem diversos tipos de conjuntos de primitivas utilizados em reconhecimento de manuscritos e a escolha de um conjunto de primitivas capaz de representar um caractere é um fator extremamente importante para o desempenho de um sistema de

reconhecimento de manuscrito. Muitas técnicas de extração de primitivas têm sido utilizadas na área de reconhecimento de manuscrito, podendo ser classificadas em primitivas estatísticas e estruturais [HEU98] [BRI04]. As primitivas estatísticas são baseadas na distribuição estatística dos pontos, tais como zoneamento, momentos, projeção de histogramas. As primitivas estruturais levam em consideração os aspectos topológicos e geométricos, como por exemplo, segmentos de linha, direções, intersecções, curvatura.

De acordo com Trier et al. (1996) [TRI96] é desejável que as primitivas a serem extraídas tenham as seguintes propriedades:

- **Invariância:** esta propriedade considera que as primitivas devem possuir aproximadamente os mesmos valores para amostras do mesmo caractere (parâmetros como translação, escala, rotação, estiramento, inclinação e espelhamento).
- **Reconstrução:** o princípio dessa propriedade é a de que caracteres possam ser reconstruídos a partir das primitivas extraídas, assim é extremamente desejável que as informações sobre os caracteres originais estejam presente nas primitivas extraídas.

Quanto à forma de representação, as primitivas podem ser representadas por [MAD01]:

- **Vetores de características ou matrizes:** a imagem do caractere é dividida em partes fixas ou variáveis e em cada segmento (zoneamento) é extraída então os vetores de primitivas.
- **Seqüência de símbolos:** onde uma palavra é representada por uma seqüência de símbolos, sendo levada em consideração a seqüência em que estes símbolos ocorrem na palavra.
- **Estruturas de grafos:** aqui a imagem toda é representada por um grafo, sendo as primitivas representadas por nós e as arestas representam as relações entre as primitivas.

## 2.4. Mecanismos de Zoneamento

Zoneamento (*zoning*) é um método para análise de informação local em partições de um dado padrão. De um modo geral, a quantidade de partições e sua distribuição podem ser

estabelecidas por um processo empírico [LI95] [SUE94] ou como resultado de um algoritmo de busca [RAD03]. Outros autores definem mecanismos de zoneamento baseados em conceitos perceptivos [AIR05] [FRE07].

Assim, o zoneamento é uma partição do padrão a partir de uma *bounding box* (ou seja, o menor retângulo que contém o padrão), sendo que cada partição é usada para auxiliar na identificação da posição em que as características do padrão são detectadas [DIL00]. O objetivo do zoneamento é obter características locais ao invés de características globais. De um modo geral, o projeto de zoneamento pode ser considerado de dois modos diferentes [LEC00] [FRE07]:

- **Fixo ou simétrico:** A *bounding box* é dividida em zonas regulares de igual tamanho;
- **Variável ou não simétrico:** a *bounding box* é não uniformemente dividida, seguindo um padrão específico pré-estabelecido ou ajustado.

A estratégia de zoneamento é geralmente definida por especialistas humanos usando o domínio de conhecimento [RAD03]. No presente trabalho utilizam-se os zoneamentos perceptivos baseados em partes específicas da imagem do caractere, como proposto por [FRE07].

## 2.5. Métodos de Classificação

O objetivo da classificação é atribuir a cada amostra desconhecida uma classe pré definida. Para o reconhecimento de caracteres os seguintes métodos de classificação de padrões são amplamente utilizados para o reconhecimento de caracteres [CHE07]:

- a) **Estatísticos:** são baseados na teoria da decisão de Bayes, sendo instanciados por classificadores baseados em estimativas de densidade paramétrica e não paramétrica. Na abordagem teórica da decisão estatística, as fronteiras de decisão são determinados pelas distribuições de probabilidade dos padrões pertencentes a cada classe.
- b) **Estruturais:** Interpretam a estrutura dos caracteres, armazenam menos parâmetros, que no métodos estatísticos e às vezes são mais precisos. Este paradigma tem sido utilizado em situações onde os padrões têm uma nítida estrutura que pode ser

capturada em termos de um conjunto de regras [JAI00]. Geralmente são utilizados em reconhecimento Online.

- c) **Redes Neurais Artificiais (RNA):** composta por um número de neurônios interligados, sendo que o modo como os neurônios estão interligados diferencia os modelos de rede: redes *feedforward*, redes recorrentes e assim por diante. As RNAs serão abordadas com mais detalhes na seção 2.5.1.
- d) **Máquina de Vetores de Suporte (*Supported Vetor Machine - SVM*):** é um novo tipo de classificador hiperplano, desenvolvido com base na teoria de aprendizagem estatística, com o objetivo de maximizar a margem geométrica do hiperplano, que está relacionada com o limite de erro da generalização. Geralmente, um classificador SVM é um classificador linear binário (duas classes), que utiliza uma função de Kernel, responsável por encontrar um hiperplano ótimo. A função de Kernel possui um conjunto de parâmetros que podem ser manipulados para obtenção de um melhor desempenho do classificador. [CHE07]
- e) **Combinação de múltiplos classificadores (*ensemble of classifiers*):** pode ser dividida em combinação paralela (horizontal) e seqüencial (vertical, cascata). A vantagem na combinação de múltiplos classificadores [OLI02], [XIA05], [KO08], é que os pontos fortes de classificadores diferentes podem ser combinados para alcançar um maior desempenho.

### 2.5.1. Redes Neurais Artificiais (RNA)

As redes neurais podem ser vistas como sistemas de computação massivamente paralelas que consistem em um número extremamente grande de processadores simples com muitas interconexões. Os Modelos de Redes Neurais tentam usar alguns princípios organizacionais, tais como aprendizado, generalização, adaptabilidade, tolerância a falhas, representação distribuída e modelos de computação de reconhecimento de padrões. Em uma rede representada por grafos ponderados, os nós são os neurônios artificiais, as arestas dirigidas (com pesos) são as conexões entre neurônios de entradas e saídas [JAI00], como pode ser verificado nas Figuras 2.3 e 2.4.

De acordo com Nievola (2004) [NIE04] para que uma RNA possa ser específica e única, devem-se indicar os seguintes elementos:

- Topologia da rede: Conjunto de interligações dos neurônios;
- Algoritmo de aprendizagem: Define como os pesos da rede neural, e eventualmente suas conexões, são alterados durante o treinamento;
- Codificação dos dados de entrada: Em muitos casos os dados de entrada devem ser convertidos para uma representação diferente daquela original;
- Decodificação dos valores de saída: As saídas correspondem a valores numéricos contínuos. Se esta não for a representação desejada, ela deve ser convertida para o formato adequado.

As RNAs são muito utilizadas na área de reconhecimento de padrões desempenhando tarefas de agrupamento, classificação, regressão, aproximação de funções, predição, entre outros. [FRE07].

#### **2.5.1.1. Treinamento das Redes Neurais**

Redes Neurais Artificiais possuem a capacidade de aprender por exemplos e fazer interpolações e extrapolações do que aprenderam. As principais características das redes neurais são a capacidade de aprender complexas relações de entrada-saída não-linear, uso sequencial de processos de treinamento e capacidade de adaptar-se aos dados [JAI00]. Os métodos para treinamento de redes podem ser agrupados em três paradigmas [JAI96], a saber:

- aprendizado supervisionado: No aprendizado supervisionado, conhecido também como aprendizado com um professor, a entrada e saída desejadas para a rede são fornecidas por um supervisor externo, sendo fornecida uma resposta correta (saída) para cada padrão de entrada. O objetivo é ajustar os parâmetros da rede para encontrar uma ligação entre os pares de entrada e saída fornecidos. Um caso especial de aprendizado supervisionado é o aprendizado por reforço, no qual a única informação de realimentação fornecida à rede é se uma determinada saída está correta ou não, isto é, não é fornecida à rede a resposta correta para o padrão de entrada, sendo assim não se sabe qual o valor de saída correto para cada exemplo de entrada; somente pode-se identificar se uma saída está correta ou não [BRA00] [JAI00] [NIE04];

- **aprendizado não-supervisionado:** No aprendizado não-supervisionado ou aprendizado sem um professor, não é preciso uma resposta correta associada a cada padrão de entrada no conjunto de treinamento. Para estes algoritmos, somente os padrões de entrada estão disponíveis para a rede; um caso especial de aprendizado não supervisionado é o aprendizado por competição. A idéia deste aprendizado parte do princípio de que dado um padrão de entrada as unidades de saída disputam entre si para serem ativadas, existindo assim uma competição entre as unidades de saída para decidir qual delas será a vencedora e, conseqüentemente, terá sua saída ativada e seus pesos atualizados no treinamento [BRA00] [JAI96] [NIE04];

Deste modo, pode-se agrupar as redes neurais de acordo com os paradigmas de aprendizagem, a arquitetura da rede e com o algoritmo de aprendizagem, tal qual mostrado na Tabela 2.1.

Tabela 2.1 - Características de RNAs em função do tipo de aprendizagem [JAI96]- adaptado

Paradigma	Regra de aprendizagem	Arquitetura	Algoritmo de aprendizagem	Tarefa
<b>Supervisionada</b>	Correção de erro	Perceptron com uma camada	Algoritmos de aprendizagem do perceptron	Classificação de padrões
		Perceptron com várias camadas	Retro-propagação; Adaline e Madaline	Aproximação de funções, predição e controle
	Boltzmann	Recorrente	Algoritmo de aprendizagem de Boltzmann	Classificação de padrões
	Correção de erro e Competitiva	Rede RBF	Algoritmo de aprendizagem RBF	Classificação de padrões, aproximação de funções, predição de controle
	Competitiva	Competitiva	Quantização do vetor de aprendizagem	Categorização em classes internas, compressão de dados
		Rede ART	ARTMAP	Classificação de padrões, categorização em classes internas
<b>Não supervisionada</b>	Correção de erro	Multicamadas em avanço	Projeção de Sammon	Análise de dados
	Hebb	Em avanço ou competitiva	Análise da componente principal	Análise de dados, compressão de dados
		Rede Hopfield	Aprendizagem de memória associativa	Memória associativa
	Competitiva	Competitiva	Quantização de vetores	Categorização, compressão de dados
		SOM (Kohonen)	SOM (Kohonen)	Categorização, análise de dados
		Rede ART	ART1, ART2	Categorização

### 2.5.1.2. MLP (Multilayer Perceptron) Convencional e Classe Modular

Uma rede MLP (*Multi-Layer Perceptron*) são RNAs que apresentam pelo menos uma camada intermediária ou escondida, sendo redes compostas por uma só camada resolvem apenas problemas linearmente separáveis. A solução de problemas não linearmente separáveis passa pelo uso de redes com mais camadas intermediárias ou escondidas. Em uma rede multicamadas, o processamento realizado por cada nó é definido pela combinação dos processamentos realizados pelos nós da camada anterior que estão conectados a ele [BRA00]. A rede MLP é uma extensão natural do *Perceptron* (primeira RNA utilizada, é composto basicamente por duas camadas: a de entrada e a de saída, conforme apresentado na Figura 2.3), sendo que os neurônios da camada escondida não têm ligações nem com os neurônios da entrada nem com a saída [NIE04]. A Figura 2.4 representa a arquitetura para uma RNA MLP convencional com uma camada escondida.

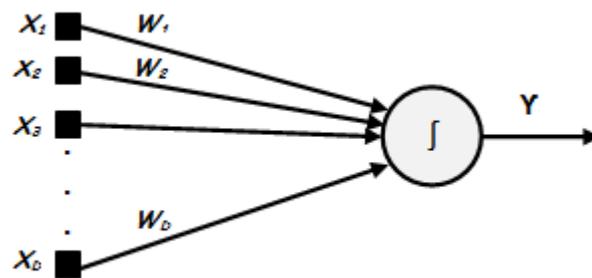


Figura 2.3: Modelo de um *Perceptron* com  $D$  entradas e uma saída [NIE04] – adaptado

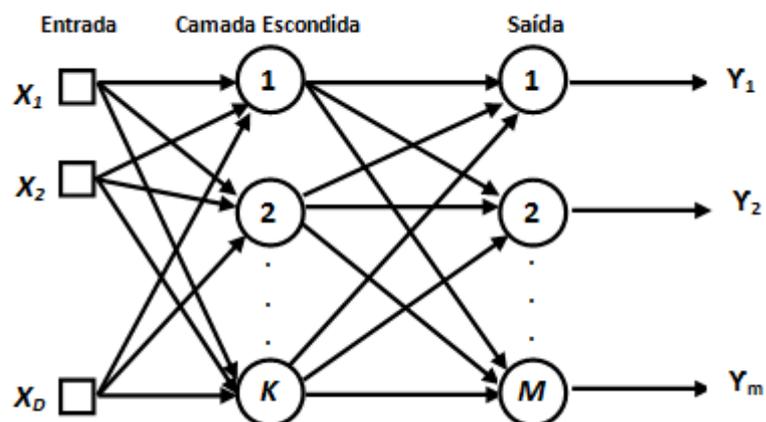


Figura 2.4.: Arquitetura MLP com uma camada escondida [NIE04]- adaptado

O número adequado de nós na camada intermediária depende de vários fatores, como o número de exemplos de treinamento, a quantidade de ruído presente nos exemplos, a complexidade da função a ser aprendida e a distribuição estatística dos dados de treinamento.

O algoritmo de aprendizado mais conhecido para treinamento das redes MLP é o algoritmo *backpropagation*, que foi o responsável pelo renascimento das RNAs, permitindo que redes neurais de múltiplas camadas apresentassem capacidade de aprendizado [JAI96].

O algoritmo *backpropagation* utiliza pares (entrada e saída desejadas) para ajustarem os pesos da rede, por meio de um mecanismo de correção de erros. O treinamento ocorre em duas fases: *forward* e *backward*, em que cada fase percorre a rede em um sentido. A fase *forward* é utilizada para definir a saída da rede para um dado padrão de entrada, enquanto que a fase *backward* utiliza a saída desejada e a saída fornecida pela rede para atualizar os pesos de suas conexões.[BRA00]

Um dos problemas enfrentados no treinamento de redes MLP diz respeito à definição de seus parâmetros. Pequenas diferenças nestes parâmetros podem levar a grandes diferenças tanto no tempo de treinamento quanto na generalização obtida. A frequência de ajuste dos pesos influencia o desempenho obtido durante o treinamento, sendo então um aspecto que precisa ser observado. Duas abordagens diferentes têm sido utilizadas quanto à frequência (periodicidade) para o ajuste de pesos pelo algoritmo *backpropagation*, a abordagem por padrão (*on-line*) e a abordagem por ciclo (*batch*). Na abordagem por padrão, os pesos são atualizados após a apresentação de cada padrão de treinamento. Esta abordagem é estável se a taxa de aprendizado for pequena (é aconselhável reduzir progressivamente esta taxa). Quando taxas elevadas são utilizadas, a rede geralmente se torna instável. A abordagem por padrão é geralmente mais rápida, sobretudo se o conjunto de treinamento for grande e redundante. Na abordagem por ciclo, por sua vez, os pesos são atualizados após todos os padrões terem sido apresentados. Esta técnica é geralmente mais estável, mas pode ser lenta se o conjunto de treinamento for grande e redundante. A escolha da abordagem a ser utilizada depende da aplicação e da distribuição estatística dos dados [BRA00].

Na arquitetura MLP Convencional as classes são treinadas juntas e a que apresentar um valor máximo de saída é considerada reconhecida. No conceito Classe-Modular (Class-Modular), o problema original  $K$  da classificação decompõe-se em  $K$ -subproblemas. Uma arquitetura modular consiste em  $K$  sub-redes, cada uma responsável para discriminar uma classe das outras  $K-1$  classes, no modelo MLP Classe-Modular cada módulo é treinado

independente dos outros, utilizando conjuntos de treinamento e validação [AIR05] [KAP03] [OH01]. A Figura 2.5 ilustra uma arquitetura Classe-Modular.

Um classificador Classe-Modular pode ser caracterizado por duas características principais:

- **Baixa complexidade do classificador:** Cada um dos 2-classificadores tem um número muito menor de parâmetros a serem estimados pelo processo de formação. Desde que 2-classificador utiliza todo o conjunto de treinamento para a sua formação, o processo de estimação de parâmetros pode ser realizada de uma maneira mais precisa e estável;
- **Independência das classes:** Cada um dos 2-classificadores pode ser projetado e treinado de forma independente das outras classes. Na forma mais geral, o módulo 2-classificador para uma classe específica pode ser visto como uma caixa preta das outras  $K-1$  classes, e o único objetivo do 2-classificador é discriminar as amostras de entrada com uma precisão muito elevada.

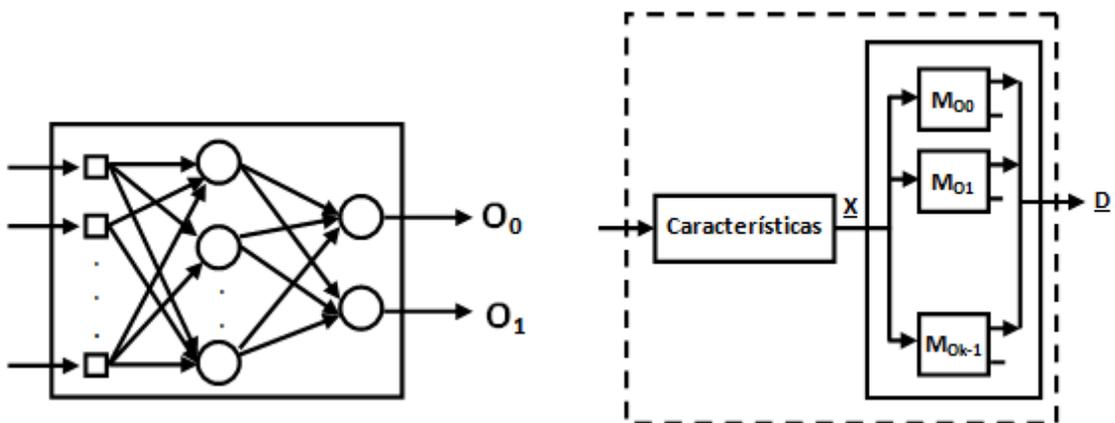


Figura 2.5: Arquitetura de uma Rede MLP Classe-Modular: a) Subrede e b) Rede Classe-Modular completa

Desde que as saídas de cada 2-classificadores são binárias, uma alternativa é a utilização de apenas uma saída. Pode se escolher então, uma das duas opções apropriadas para a aplicação [OH01]. A Figura 2.6 ilustra as duas opções possíveis, sendo que  $\underline{X}$  representa o padrão de entrada,  $2-C(\omega_i)$  denota o 2-classificador para uma classe  $\omega_i$ , e  $d$  denota o vetor que representa a probabilidade da amostra pertencer a uma classe  $\Omega_0$  ou  $\Omega_1$ :



Figura 2.6. Diagramas esquemáticos de 2-classificador para uma classe  $\omega_i$  ([OH01] - adaptado)

## 2.6. Trabalhos relacionados

Existem na literatura vários trabalhos na área de reconhecimento de manuscritos e, portanto, nesta seção são abordadas algumas dessas experiências, as quais foram consideradas relevantes para o contexto deste trabalho.

Suen et al. (1994) [SUE94] baseado na probabilidade de ocorrência de padrões e tomando por princípio quais partes são cruciais ao reconhecimento de caracteres, levando em consideração também o grau de confusão e a similaridade entre pares de caracteres, recomendam um reconhecimento baseado em um mecanismo de zoneamento simétrico, subdividindo os caracteres em 2, 4 e 6 regiões, como mostrado na Figura 2.7.

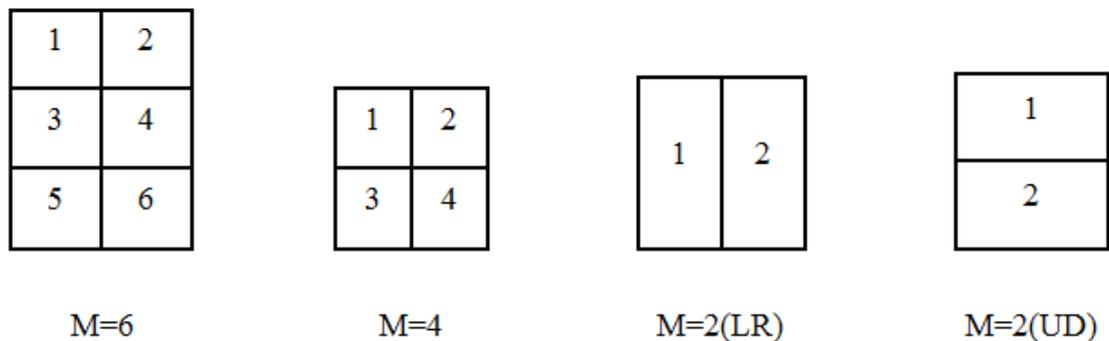


Figura 2.7: Representação das partições, os números 1, 2, 3, 4, 5, 6 representam o número de posições das partes [SUE94]

Para os experimentos foram utilizados 26 caracteres do alfabeto inglês e 10 numerais, num total de 89 padrões, com os estilos mais comuns de caracteres alfanuméricos impressos, escolhidos através de experimentos computacionais e humanos. Através de seus experimentos, Suen et al. (1994) [SUE94] constataram que o caractere “D” sempre alcança taxas altas de reconhecimento (100%), e que os caracteres “A”, “K” e “G” também obtêm taxas de reconhecimento altas. As taxas de reconhecimento para os zoneamentos 2UD, 2LR, 4

e 6 foram 86,12%, 85,88%, 61,73% e 42,91%, respectivamente. Observa-se ainda nos experimentos realizados que as partições podem influenciar na taxa de reconhecimento dos caracteres. Ao se analisar a Figura 2.8 nota-se que a metade superior no zoneamento 2UD, a parte esquerda no zoneamento 2LR e a posição 3 no zoneamento 6, notavelmente tem uma alta taxa de reconhecimento. Ao se analisar as duas partes do zoneamento 2LR para um caractere “Y”, ambas são perfeitas para o reconhecimento, porém para um caractere “B”, a parte esquerda pode ser confundida com a letra “E” e a parte direita com o numeral “8” ou “3”, como ilustrado na Figura 2.9. Assim, os autores definem como partes cruciais as partes que são facilmente reconhecidas por serem absolutamente distintas em um caractere. Definem, ainda, as partes de confusão são aquelas partes similares entre os caracteres distintos.

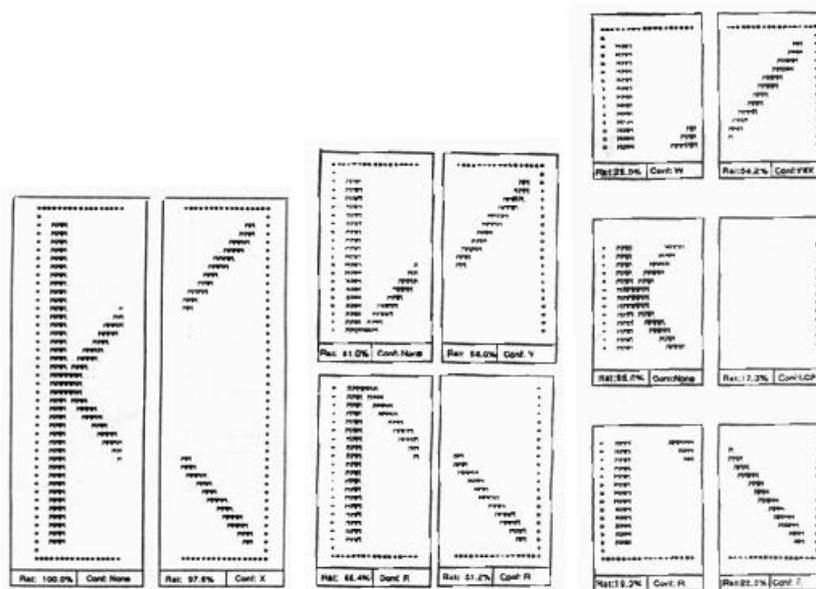


Figura 2.8: Os tipos de partições propostas para uma letra K [SUE94]



Figura 2.9. Exemplos de dois caracteres com um zoneamento 2LR

Tay et al. (2001) [TAY01] combina Modelos Escondidos de Markov (HMM – *Hidden Markov Model*) e Redes Neurais no reconhecimento de palavras manuscritas. Utiliza um

método de segmentação da esquerda para direita para extrair características geométricas para cada *frame* da imagem segmentada, ou seja, gerando assim uma segmentação gráfica que fornece todos os modos possíveis de se segmentar uma palavra em caracteres isolados. As características geométricas consideradas são as seguintes: dimensão e proporção da *bounding box* dos pixels nos *frames*, centro de gravidade, distâncias centro-zona, perfil em 8 direções, números de transições verticais e horizontais na direção diagonal de 45° e -45°. Os experimentos utilizam a base de dados IRONOFF [TAY01].

Oliveira et al.(2002) [OLI02] avalia o desempenho da combinação de Redes Neurais e Modelos Escondidos de Markov (MEM), propondo e analisando o desempenho das RNAs e dos MEMs aplicados ao reconhecimento de palavras manuscritas. A aplicação envolve o reconhecimento de manuscritos referentes aos nomes de meses do ano em língua portuguesa do Brasil, como exemplificado na Figura 2.10. Para os experimentos com as RNAs os autores utilizam a segmentação implícita, dividindo cada imagem em 8 sub-regiões, como mostrado na Figura 2.11. Para cada uma das 8 regiões é então definido um conjunto de 10 padrões, totalizando assim um vetor de primitivas com 80 padrões. Após a segmentação, o processo de extração das primitivas é realizada considerando dois tipos de primitiva: perceptivas e direcionais. Um exemplo de procedimento para a extração de primitivas perceptivas para uma palavra é ilustrada na Figura 2.12. Para os experimentos com o HMM, 3 zonas de extração de primitivas são definidas baseadas no histograma de projeção horizontal das transições (HT) branco-preto da palavra: zona ascendente, zona corpo da palavra e zona descendente, como mostrado na Figura 2.12. As taxas de reconhecimento obtidas foram de 75,9%, utilizando apenas o classificador HMM e de 90,4% considerando a combinação de RNAs e MEM [OLI02].

Handwritten names of months in Portuguese, arranged in a grid-like fashion:

Jan	Fevereiro	Março
Abril	dezenbra	Outubro
Junho	Junho	agosto
Agosto	dezembro	setembro
Julho	Janerio	fevereiro
Outubro	maio	setembro

Figura 2.10. Exemplo de nomes de meses [OLI02]



Figura 2.11: Exemplo de segmentação implícita [OLI02]



Figura 2.12. Zonas de Extração de primitivas [OLI02]

Correia et al. (2002) [COR02] propuseram uma abordagem de reconhecimento de manuscritos baseado na percepção humana, utilizam transformadas *Wavelet* para simular a capacidade de visão e para extração das primitivas baseadas nos pontos de fixação e detalhes da imagem, considerando três posições: vertical, horizontal e diagonal. Assim, o mecanismo de zoneamento baseado em janelas de "atenção" em cada ponto de fixação simula o movimento dos olhos e permite a extração seletiva e local de primitivas direcionais. A Figura 2.13 representa os pontos de fixação e suas respectivas janelas de atenção [COR02].

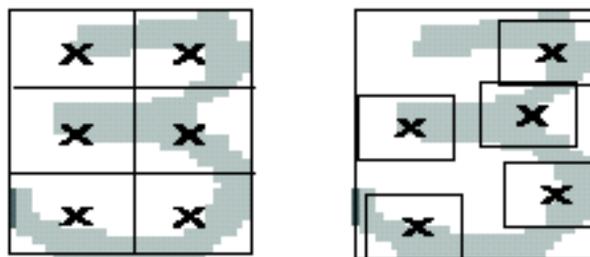


Figura 2.13: Pontos de fixação e suas respectivas janelas de atenção [COR02]

Radtke et al.(2003) [RAD03] apresenta uma abordagem automática para definir o zoneamento para o reconhecimento *off-line* de caracteres, usando *Multi-Objective Evolutionary Algorithm* (MOEA). O MOEA é um algoritmo baseado no método de pesquisa Darwiniana, no qual uma população de soluções candidatas evolui através de gerações de

operadores genéticos, tais como seleção, mutação e *crossover*. Os autores propõem uma metodologia adaptativa para definir a estratégia de zoneamento com  $m$  zonas de não sobreposição e uma taxa de erro aceitável, sem necessidade de intervenção humana durante a fase de pesquisa. Em seus experimentos as zonas foram definidas baseadas em divisões de posição fixa, que podem estar nos estados *on* ou *off*. A Figura 2.14 mostra o *template* utilizado como base para os experimentos. Dado que cada divisão tem dois estados (*on* ou *off*), tem-se que uma seqüência de 10 bits pode ser utilizada para codificar um indivíduo, sendo que cada bit indica se a divisão está *on* ou *off*. Assim, a seqüência descreve 1024 possíveis diferentes configurações de zoneamento. O conjunto de primitivas utilizados nos experimentos é composto de uma mistura de informações de concavidade e de contorno. Treze medidas de concavidades, um histograma de direções de contorno (8-direções Freeman) e o número de pixels pretos são extraídos de cada zona da imagem.

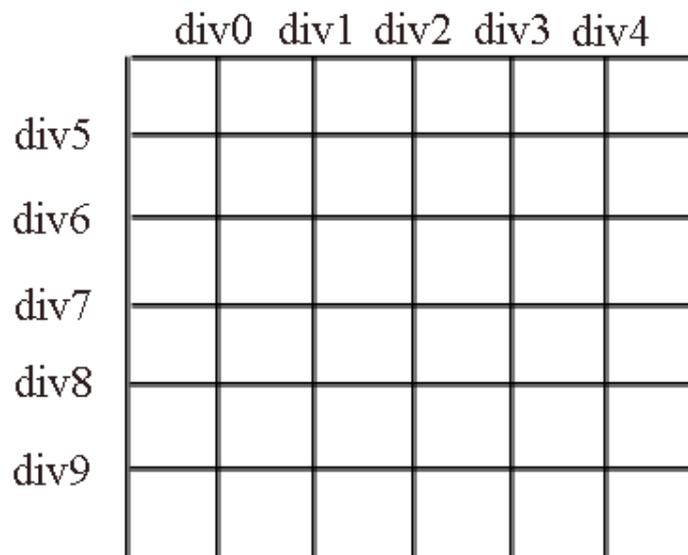


Figura 2.14: *Template* de codificação individual (Adaptado [RAD03])

Desta forma, uma estratégia de zoneamento como apresentado na Figura 2.15e irá produzir um vetor de primitivas com 132 elementos (22 x 6). As Figuras 2.15a até 2.15e mostram as estratégias de zoneamento que apresentaram o melhor *trade-off* entre número de zonas e taxa de erro, encontradas pela metodologia proposta por Radtke et al.(2003) [RAD03]. A Figura 2.15f representa um zoneamento concebido por um especialista humano, para fins de comparação com a metodologia proposta pelos autores.

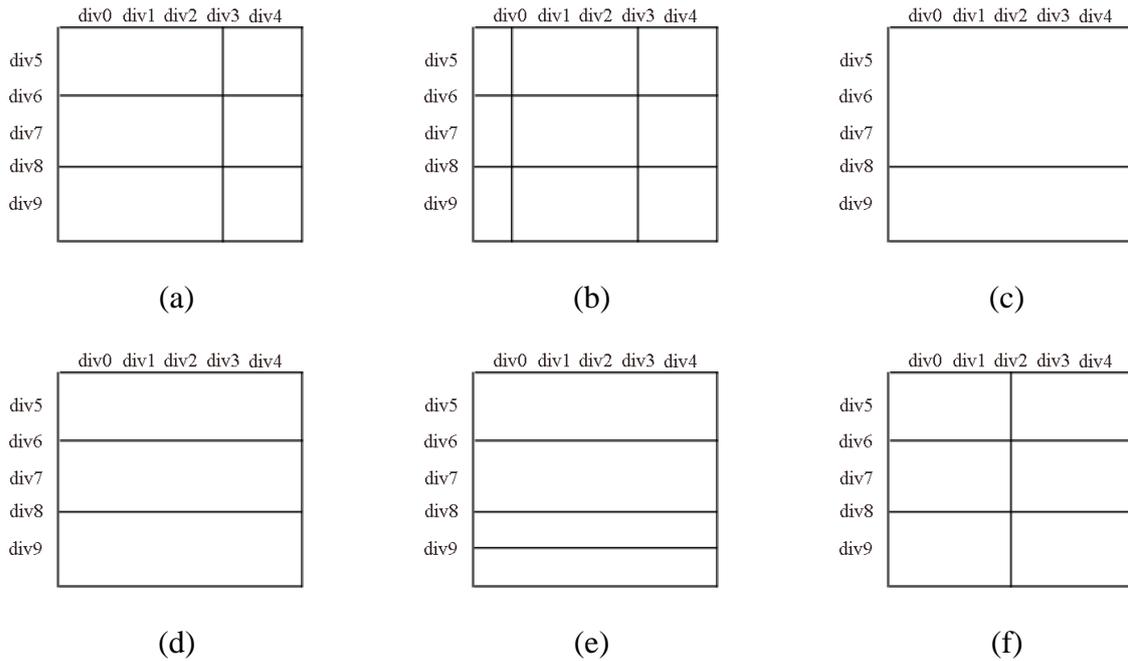


Figura 2.15: Estratégias de zoneamento (Adaptado [RAD03])

Para avaliar a taxa de erro individual, foi utilizado um classificador NN (*Nearest Neighbor*), sendo que uma comparação direta entre a taxa de erro em um classificador NN e uma RNA foi feita a fim de validar o poder de generalização da metodologia proposta, como mostrado na Tabela 2.2. Observa-se que o melhor resultado alcançado foi para o zoneamento com 9 regiões representada pela Figura 2.15b e indicada na Tabela 2.2.

Tabela 2.2: Taxas de erro ([RAD03] – adaptado)

Estratégia	NN (Nearest Neighbor)	Rede Neural
Figura 2.15a	5%	1.88%
→ <b>Figura 2.15b</b>	<b>5%</b>	<b>1.85%</b>
Figura 2.15c	6.92%	3.25%
Figura 2.15d	5.51%	2%
Figura 2.15e	5.32%	1.82%
Figura 2.15f	5.32%	1.97%

Blumenstein et al. (2003) [BLU03] apresenta uma nova técnica de extração de primitivas (primitivas de direção) para o reconhecimento de caracteres manuscritos. Seus

experimentos utilizam também classificadores neurais e zoneamento simples. Tal técnica de extração procura simplificar os limites ou representações diluídas de cada caractere através da identificação de cada direção ou segmentos de linha na imagem. Os segmentos de linha que podem ser determinados em cada caractere são de quatro tipos: linhas verticais, linhas horizontais, diagonal direita e diagonal esquerda. Ao lado dessas quatro representações de linha, a técnica também localiza pontos de intersecção entre cada tipo de linha, como mostrado na Figura 2.16.

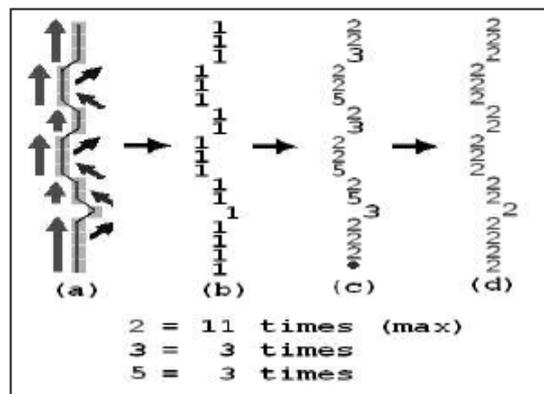


Figura 2.16. Extração de primitivas: a) linha original, b) linha no arquivo binário, c) depois de distinguidas as direções e d) depois da normalização das direções

Deste modo, Blumenstein et al (2003) [BLU03] trabalhando com caracteres minúsculos e maiúsculos organizados da base de dados do CEDAR, obtiveram taxas de reconhecimento de 83,65% utilizando as redes neurais com algoritmo *backpropagation* (utilizando a base de imagens pré-segmentadas - *Binary Alphanumeric Characters*). Para o conjunto de dados automaticamente segmentados (CAS) utilizando redes neurais e algoritmo *backpropagation* a taxa de reconhecimento para caracteres minúsculos foi de 69,78% sendo que foi usado primitivas de direção (*resized boundary*). Por outro lado, a taxa para caracteres maiúsculos foi de 80.62%, tendo sido utilizado o conjunto direcional de primitivas (*non resized thinned*).

Em Kapp et al.(2003) [KAP03] e Kapp (2004) [KAP04] é avaliada a arquitetura das RNAs Convencional e Classe-Modular, baseando no estudo de Oh e Suen [OH01], em que RNAs-MLP são utilizadas como classificadores organizados seguindo uma arquitetura Classe-Modular, para o reconhecimento de palavras manuscritas. Os autores utilizaram um conjunto de primitivas e aplicação de técnicas globais para a extração de padrões. Exploram

também primitivas perceptivas e características baseadas em concavidades e convexidades para o reconhecimento dos nomes dos meses do ano e dos valores por extenso, em formato manuscrito, em cheques bancários na língua portuguesa, sendo que em seus experimentos não há restrição quanto ao estilo de escrita, nem quanto ao número de escritores. Os autores propõem um método de reconhecimento automático *off-line* de dois conjuntos de palavras manuscritas encontrados no contexto de cheques bancários brasileiros. O primeiro conjunto corresponde aos meses do ano, formando um léxico de doze palavras, a base de dados utilizada foi a da Universidade Federal de Campina Grande (UFCG). O segundo conjunto é formado pelas palavras que compõem os valores por extenso, constituindo assim um léxico de trinta e nove palavras, sendo que a base de dados utilizada foi da Pontifícia Universidade Católica do Paraná (PUCPR). A Figura 2.17 mostra exemplo de palavras com valores por extenso utilizada pelos autores.

um dois seis sete dois um  
 três sete um três  
 oito sete  
 nove quatro nove  
 sete oito um nove  
 um dois oito nove oito  
 seis um sete quatro  
 três cinco seis três  
 quatro cinco  
 bom cinco dois um

Figura 2.17: Amostras referentes aos valores por extenso – base PUCPR [KAP04]

O método desenvolvido extrai primitivas globais das palavras, tais como, número de laços, ascendentes e descendentes, concavidades, convexidades, traços horizontais e verticais, gerando assim um vetor de 24 dimensões. O método utiliza o zoneamento das palavras em somente duas regiões, separadas pelo centro de gravidade da palavra (região da esquerda e região da direita da palavra), como mostrado na Figura 2.18. A justificativa para o uso de duas regiões no zoneamento é que separando as ocorrências das primitivas se obtém a informação sobre o posicionamento das mesmas dentro da palavra e isto oferece mais precisão na

classificação das formas. Além disto, os autores puderam estudar as confusões entre prefixos e sufixos que se repetem nas palavras dos léxicos em questão, por exemplo: “janeiro – fevereiro”, “quatro – quatrocentos”. Para os meses do ano, a taxa de reconhecimento obtida para a arquitetura RNA MLP Convencional foi de 77,08%, enquanto que a taxa de reconhecimento obtida para a arquitetura RNA MLP Classe - Modular foi de 81,75%. Para os valores por extenso, a taxa de reconhecimento alcançada para a arquitetura RNA MLP Convencional foi 48,93% , enquanto a arquitetura RNA MLP Classe - Modular alcançou taxa de reconhecimento de 52,34%.

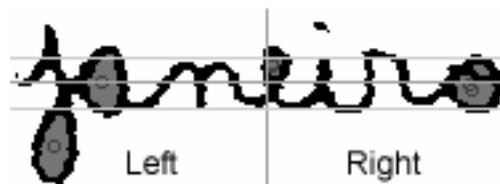


Figura 2.18. Mecanismo de zoneamento para um nome de mês - [KAP04]

Britto et al.(2004) [BRI04] combinam extração de primitivas do fundo da imagem de caracteres manuscritos e do primeiro plano da imagem do caractere em um sistema baseado em Modelos Escondidos de Markov (MEMs). As primitivas extraídas do fundo da imagem são baseadas em informações de concavidade e são utilizadas para realçar as propriedades topológicas e geométricas de cada classe do caractere. As primitivas do primeiro plano da imagem são baseadas em primitivas locais e globais, sendo estas calculadas de modo a levar em consideração os pixels do primeiro plano das colunas ou linhas da imagem do caractere. As características locais são baseadas em transições de pixel do fundo da imagem para o primeiro plano da imagem e vice-versa. Cada classe de caractere foi então representada por dois Modelos Escondidos de Markov (MEMs), um baseado em linhas e outro baseado em colunas da imagem. Deste modo um zoneamento implícito é criado, como mostrado na Figura 2.19. O sistema proposto trabalha com dígitos e caracteres isolados e cadeias (*strings*) de dígitos de diferentes tamanhos. Para os experimentos foi utilizada a base de dados NIST SD19. A taxa de reconhecimento para caracteres maiúsculos foi de 90.0% e para caracteres minúsculos foi de 84.0%.

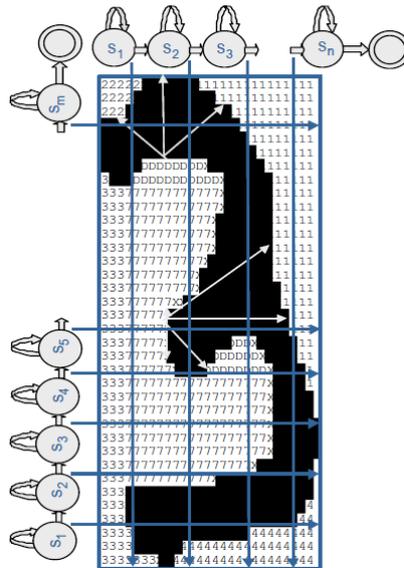


Figura 2.19. Esquema de zoneamento implícito pela combinação do modelo baseado em colunas e linhas [BRI04]

Aires (2005) [AIR05] e Freitas et al. (2007) [FRE07] propõem um zoneamento perceptivo para o reconhecimento de caracteres manuscritos baseado na confusão entre regiões específicas dos caracteres. Para isto, os autores utilizam informações contidas na matriz de confusão para propor uma estratégia perceptiva e não simétrica como mecanismo de zoneamento. Assim, os autores propõem um método híbrido de extração de primitivas, composto por uma rotulação global dos pixels do fundo da imagem (*background*) com posterior aplicação de mecanismos de zoneamento. A rotulação dos pixels do fundo da imagem tem por base a análise de deficiências de convexidades e concavidades (*concavity/convexity deficiencies*). A idéia básica de deficiência de concavidade e convexidade é que para cada pixel branco na imagem do caractere, busca-se nas quatro direções (norte, sul, leste e oeste) o número de pixels preto que se pode alcançar, assim como as direções em que os pixels pretos não são alcançados [AIR05]. A Figura 2.20 apresenta os mecanismos de zoneamento perceptivos e não simétricos propostos por Aires (2005) [AIR05] e Freitas et al. (2007) [FRE07], ilustrando a partição de uma *bounding box* (caixa delimitadora) dividida em 5 partes verticais (Z5V), 5 partes horizontais (Z5H) e 7 partições (Z7).

Para seus experimentos, Aires (2005) [AIR05] utilizou a base de dados IRONOFF, com caracteres manuscritos maiúsculos do alfabeto. As taxas médias de reconhecimento obtidas, para os quatro zoneamentos avaliados, foram as seguintes: Z4 = 82,89%, Z5H =

81,75%, Z5V = 80,94% e Z7 = 84,73%. Através dos experimentos a autora demonstrou que o mecanismo de zoneamento Z5H contribui para solucionar problemas em classes de caracteres que não são horizontalmente simétricas, apresentando melhores resultados para as seguintes confusões: “G”, “O”, e “Y”. Porém o zoneamento Z5V não contribuiu para o aumento da taxa de reconhecimento do classificador individual, não auxiliando na solução das confusões identificadas. Por outro lado, o zoneamento Z7 alcançou melhor resultado para caracteres não simétricos, extraíndo e representando a parte central dos caracteres diferentemente das partes superior e inferior, tal qual nos seguintes caracteres: “D” e “C”; “N” e “W”; “Y” e “X”. O zoneamento Z7 resultou melhor para as seguintes classes de caracteres: “B”, “C”, “D”, “E”, “K”, “N”, “P”, “R”, “U”, “W” e “X”. Alguns exemplos de mecanismos de zoneamentos perceptivos para caracteres manuscritos maiúsculos utilizados nos experimentos de Aires (2005) para Z5H e Z7 são ilustrados na Figura 2.21.

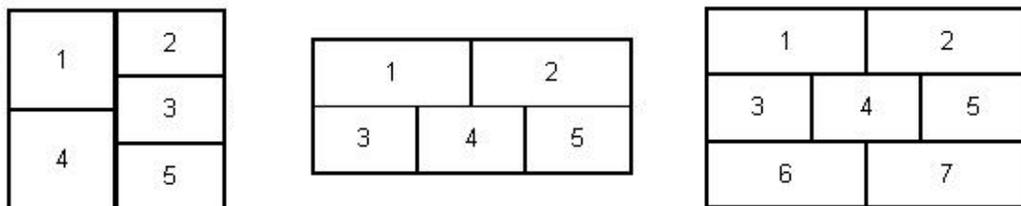


Figura 2.20: Zoneamento não simétrico Z5V, Z5H e Z7, respectivamente

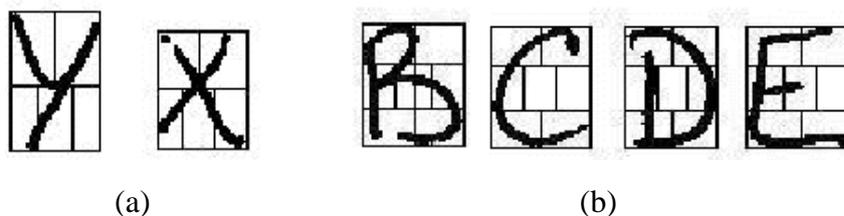


Figura 2.21: Zoneamento baseado em regiões perceptivas: a) Z5H e b) Z7 [FRE07]

No processo de reconhecimento dos caracteres manuscritos maiúsculo Aires (2005) [AIR05] utilizou um comitê de Rede Neurais Artificiais MLP com arquitetura Classe - Modular [FRE07]. Ainda em seu trabalho Aires (2005) [AIR05] aborda a criação de metaclassa por meio de uma abordagem baseada no cálculo de distâncias denominadas como *Distance-based Disagreement* (DbD). A abordagem DbD utiliza a informação contida nas matrizes de confusão, sendo que para cada classificador individual são computadas as distâncias que representam as discordâncias entre classificadores. O objetivo deste estudo estava em analisar

Sistemas Multi-Classificadores (*Multiple Classifiers System - MCS*). A distância pode ser obtida considerando que todas as matrizes de confusão possuem o mesmo tamanho (linhas e colunas). As distâncias foram medidas considerando os classificadores individuais, dois a dois e três a três, computando-se desta forma 10 medidas de distância, com as seguintes combinações: 4-5H, 4-5V, 4-7, 5H-5V, 5H-7, 5V-7, 4-5H-5V, 4-5H-7, 4-5V-7, 5H-5V-7. A partir da abordagem proposta, a autora identificou as seguintes meta-classes de caracteres manuscrito maiúsculo: *CDQRSZ*, *EJM*, *FPT*, *GX*, *HLNO*, *IY*, *UVW*. A Figura 2.22 ilustra de forma esquemática as meta-classes que foram identificadas nos experimentos realizados [AIR05].

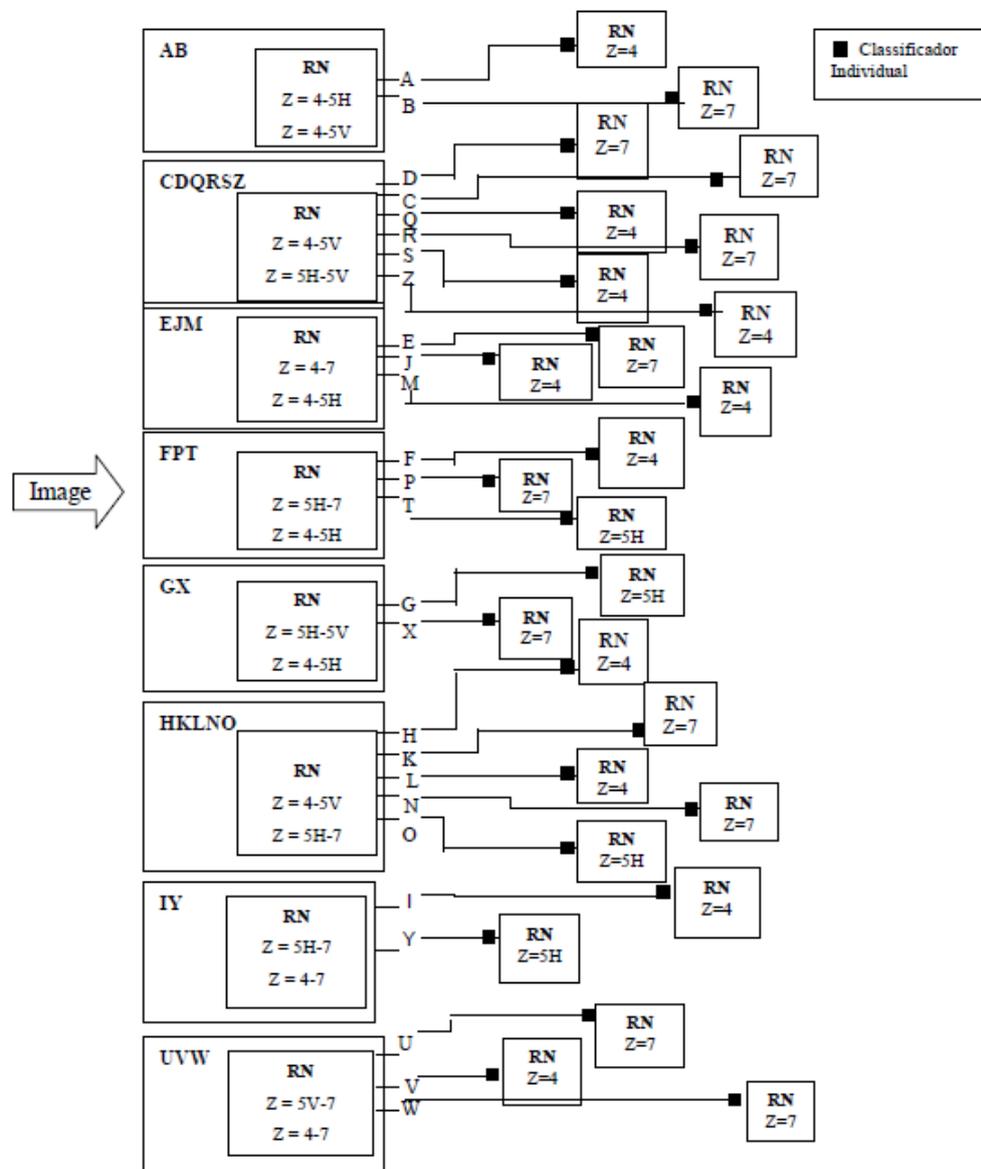


Figura 2.22: Meta-classes geradas a partir do conjunto de validação [AIR05]

Ko et al. (2008) [KO08] fez um estudo sobre seleção dinâmica de classificadores propondo 4 novos esquemas de seleção dinâmica para um conjunto de classificadores baseados no conceito de um oráculo. O método proposto é chamado de Seleção Dinâmica de Conjuntos de Classificadores (*EoC – Ensemble of Classifiers*), sendo abordado com mais detalhes na seção 2.7. Em um de seus experimentos, os autores realizaram experimentos com numerais manuscritos, considerando 10 classes de números e a utilização de um conjunto de 100 classificadores KNN gerados pelo método de Random Subspaces. Cada classificador KNN é composto por um conjunto de 32 primitivas. A base de dados utilizada pelos autores para os experimentos com os numerais foi o NIST SD19. Para os experimentos realizados com números, além dos quatro novos esquemas propostos pelos autores: KNORA Eliminate (KN-E), KNORA Eliminate - W (KN-E-W), KNORA Union (KN-U) e KNORA Union - W (KN-U-W), foram utilizados quatro outros métodos de seleção dinâmica de classificadores: Overall Local Accuracy (OLA), Local Class Accuracy (LCA), seleção a Priori e a seleção a Posteriori para fins de comparação. As taxas médias de reconhecimento alcançadas para KNORA Eliminate (KN-E) e KNORA Eliminate - W (KN-E-W), foram de 97,52 %, mostrando ser um método promissor.

A Tabela 2.3 mostra todos os resultados obtidos pelos autores nos experimentos realizados com números para os quatro novos esquemas propostos (KN-E, KN-E-W, KN-U e KN-U-W) e para os outros quatro métodos de seleção dinâmica (OLA, LCA, Seleção a Priori e Seleção a Posteriori).

Tabela 2.3. Taxas de reconhecimento para cada método de seleção dinâmica de conjunto de classificadores ([KO08]-adaptado)

	<b>KN-E</b>	<b>KN-E-W</b>	<b>KN-U</b>	<b>KN-U-W</b>	<b>OLA</b>	<b>LCA</b>	<b>A Priori</b>	<b>A Posteriori</b>
<b>Rec (%)</b>	97,52	97,52	97,25	97,25	94,11	97,40	94,12	97,40
<b>k</b>	7,8	7,8	1	1	30	1	30	1

Exposto os trabalhos considerados relevantes para o contexto do presente trabalho, pode-se então fazer um resumo relacionado os autores e seus trabalhos.

A Tabela 2.4 mostra o resumo com a relação dos trabalhos, seus autores, e os resultados obtidos em seus experimentos.

Tabela 2.4. Resumo de autores e trabalhos

<b>Autor</b>	<b>Resultados</b>	<b>Classificador</b>	<b>Proposta</b>	<b>Tipo manuscrito</b>
[SUE94] IRONOFF	2UD: 86,12% 2LR: 85,88% 4: 61,73% 6: 42,91%	MEM e RNA	Zoneamento simétrico com subdivisão dos caracteres em 2, 4 e 6 regiões. Considera probabilidade de ocorrência de padrões, partes cruciais, grau de confusão e similaridade entre caracteres.	Palavras
[TAY01] IRONOFF	MEM + RNA: 96,10%	MEM e RNA	Combinação de classificadores, com utilização de método de segmentação da esquerda para direita para extrair características geométricas para cada frame da imagem segmentada.	Palavras
[OLI02] PUCPR	MEM: 75,9% RNA + MEM: 90,4%	RNA e MEM	Combinação de classificadores	Palavras: meses do ano
[COR02] NIST	98,25%	RNA	Mecanismo de zoneamento baseado na percepção humana (janelas de "atenção"), utilizando Wavelet	Numerais
[RAD03] NIST SD -19	NN : 5,00% (taxa de erro) RNA: 1,85% (taxa de erro)	NN e RNA	Abordagem automática para definir zoneamento utilizando MOEA	Dígitos
[BLU03] CEDAR	RNA( <i>backpropagation</i> ) BAC: 83,65% CAS: minúsculo: 69,78% maiúsculo: 80,62%	RNA	Técnica de extração de primitivas (primitivas de direção)	Caracteres minúsculos, maiúsculos
[KAP03] e [KAP04]	<u>Nomes de meses</u> RNA Conv: 77,08% Classe-Modular: 81,75% <u>Valores por extenso</u> RNA Conv: 48,93% Classe-Modular: 52,34%.	RNA Convencional e Classe-Modular	Explora o uso de RNA Classe-Modular e uso de primitivas perceptivas e características baseadas em concavidades e convexidades para o reconhecimento.	Palavras: meses do ano e valores por extenso
[BRI04] NIST SD19.	Maiúsculos: 90.0% Minúsculos: 84.0%. Maiusculos + Minúsculos : 87.0% Dígitos: 89.6%	MEM	Combina extração de primitivas do fundo da imagem de caracteres manuscritos e do primeiro plano da imagem do caractere em um sistema baseado em MEM, e criação de zoneamento implícito.	Caracteres maiúsculos e minúsculos e dígitos.
[AIR05] e [FRE07] IRONOFF	Z4 : 82,89% Z5H: 81,75% Z5V: 80,94% e Z7: 84,73%.	RNA Classe-Modular	Zoneamento perceptivo com método híbrido de extração de primitivas, composto por uma rotulação global dos pixels do fundo da imagem com posterior aplicação de mecanismos de zoneamento.	Caracteres maiúsculos
[KO08] NIST SD19	KN-E e KN-E-W: 97,52. KN-U e KN-U-W: 97,25	KNN	Seleção Dinâmica de Conjuntos de Classificadores ( <i>EoC – Ensemble of Classifiers</i> )	Dígitos

## 2.7. Seleção Dinâmica

De acordo com Ko et al. (2008) [KO08] um modo para se definir os limites de desempenho de um Conjuntos de Classificadores (*EoC – Ensemble of Classifiers*) é através do conceito de oráculo. Um oráculo tem o poder de fazer predições. O conceito de oráculo aqui, é: se um classificador de um *EoC* pode classificar um dado padrão, então este *EoC* é considerado capaz de classificar este padrão em um oráculo.

Assim, partindo da natureza dos oráculos, são selecionados aqueles classificadores capazes de classificar corretamente um dado padrão. Isso é feito de um modo dinâmico, pois diferentes padrões requerem diferentes conjuntos de classificadores. Esse método proposto por Ko et al. (2008) [KO08] é chamado de Seleção Dinâmica de Conjuntos de Classificadores (*Dynamic Ensemble Selection of Classifiers*). Tal método investiga as técnicas de seleção dinâmica e propõe a criação de um Conjunto de Classificadores (*EoC – Ensemble of Classifiers*) a partir de um grupo de classificadores. O conjunto é escolhido de acordo com um método proposto pelos autores, denominado de KNORA (*K-Nearest Oracle*) *dynamic ensemble selection*.

Em seu trabalho os autores ressaltam ainda que uma das mais importantes questões para a criação de um Conjunto de Classificadores (*EoC – Ensemble of Classifiers*) é a seleção de um conjunto adequado de classificadores a partir de um grupo de diferentes classificadores. Diferentes classificadores apresentam dificuldades e possibilidades diversas para diferentes padrões de teste. Assim sendo, é razoável assumir que caso os conjuntos de classificadores possam se adaptar, então pode-se obter um conjunto resultante de classificadores melhor do que um simples *EoC* estático.

A diferença entre esquemas dinâmicos e estáticos está em que na seleção estática, um conjunto de *EoC* é selecionado para toda uma base de padrões de teste. Por outro lado, na seleção dinâmica, diferentes *EoC*'s podem ser selecionados para diferentes padrões de teste.

Esta diferença é representada de forma esquemática na Figura 2.23 que ilustra um esquema de seleção estática de um conjunto de classificadores e um esquema de seleção dinâmica de um conjunto de classificadores.

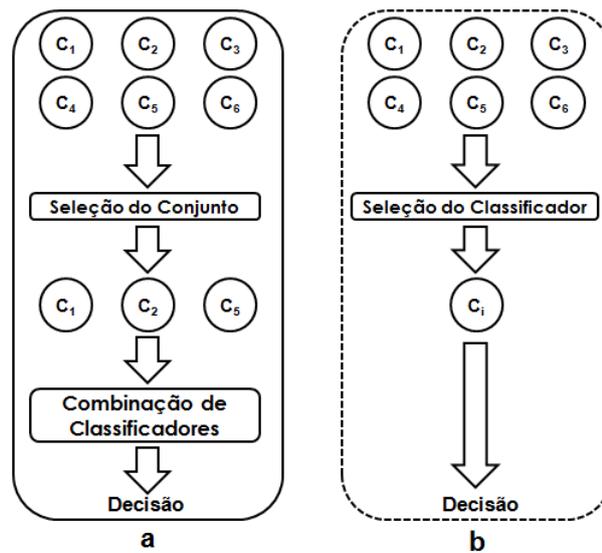


Figura 2.23. a) Esquema de seleção estática b) e seleção dinâmica [KO08] - adaptado

## 2.8. Considerações Finais

Neste capítulo foram apresentados os trabalhos já realizados na área de reconhecimento de caracteres manuscritos abordando algumas técnicas e resultados já obtidos na área, bem como a fundamentação teórica para a execução da metodologia proposta por Aires (2005) [AIR05], Freitas et al. (2007) [FRE07] e Ko et al (2008) [KO08]. O Trabalho de Ko et al (2008) [KO08] permitiu compreender a complexidade e importância na escolha de classificadores para compor um EoC robusto, contribuindo com o aprofundamento necessário para execução do trabalho, permitindo que se trabalhasse de forma análoga ao EoC, porém utilizando um conjunto de mecanismo de zoneamentos. Os experimentos realizados por Aires (2005) [AIR05] para caracteres manuscrito maiúsculo e os mecanismos de zoneamento perceptivos não simétricos propostos por Freitas et al. (2007) [FRE07] permitiram compreender a complexidade do problema de reconhecimento de caracteres manuscritos e a importância de se incorporar aspectos perceptivos no reconhecimento de caracteres, possibilitando assim a formação de um conjunto de mecanismos de zoneamento robusto. No capítulo seguinte serão descritos e apresentados os resultados dos experimentos já realizados, fazendo uma abordagem mais detalhada do método proposto.

## Capítulo 3

# Método de Reconhecimento de Caracteres Manuscritos com Seleção Dinâmica de Mecanismos de Zoneamento

Este capítulo especifica os detalhes do método proposto e analisado, a abordagem utilizada e as técnicas aplicadas para o desenvolvimento do referido sistema.

### 3.1. Visão Geral

A seleção de um método de extração de primitivas é provavelmente um dos fatores mais importante para se obter um alto desempenho em sistemas de reconhecimento. Utiliza-se neste trabalho, para a extração do conjunto de primitivas, uma abordagem híbrida, combinando extração global e local. São combinadas técnicas de rotulação dos pixels do fundo da imagem com mecanismos de zoneamento não simétricos e perceptivos. Esta combinação estabelece um conjunto de mecanismos de zoneamentos (EoZ) capaz de representar adequadamente um determinado padrão. A partir do EoZ, uma abordagem de classificação utilizando uma Rede Neural Artificial com arquitetura Convencional é utilizada para determinar o zoneamento mais indicado ao reconhecimento de cada amostra de caracteres manuscritos minúsculos. A predição do melhor mecanismo de zoneamento é feita de forma dinâmica, correspondendo ao Nível 01 do método como um todo. Assim, o Nível 02 diz respeito à fase de reconhecimento do padrão propriamente dito, na qual foram avaliadas Redes Neurais Artificiais com arquitetura Classe-Modular. É nesta fase que é realizado o

reconhecimento de cada amostra de caractere manuscrito minúsculo. A Figura 3.1 representa uma visão geral do método proposto.

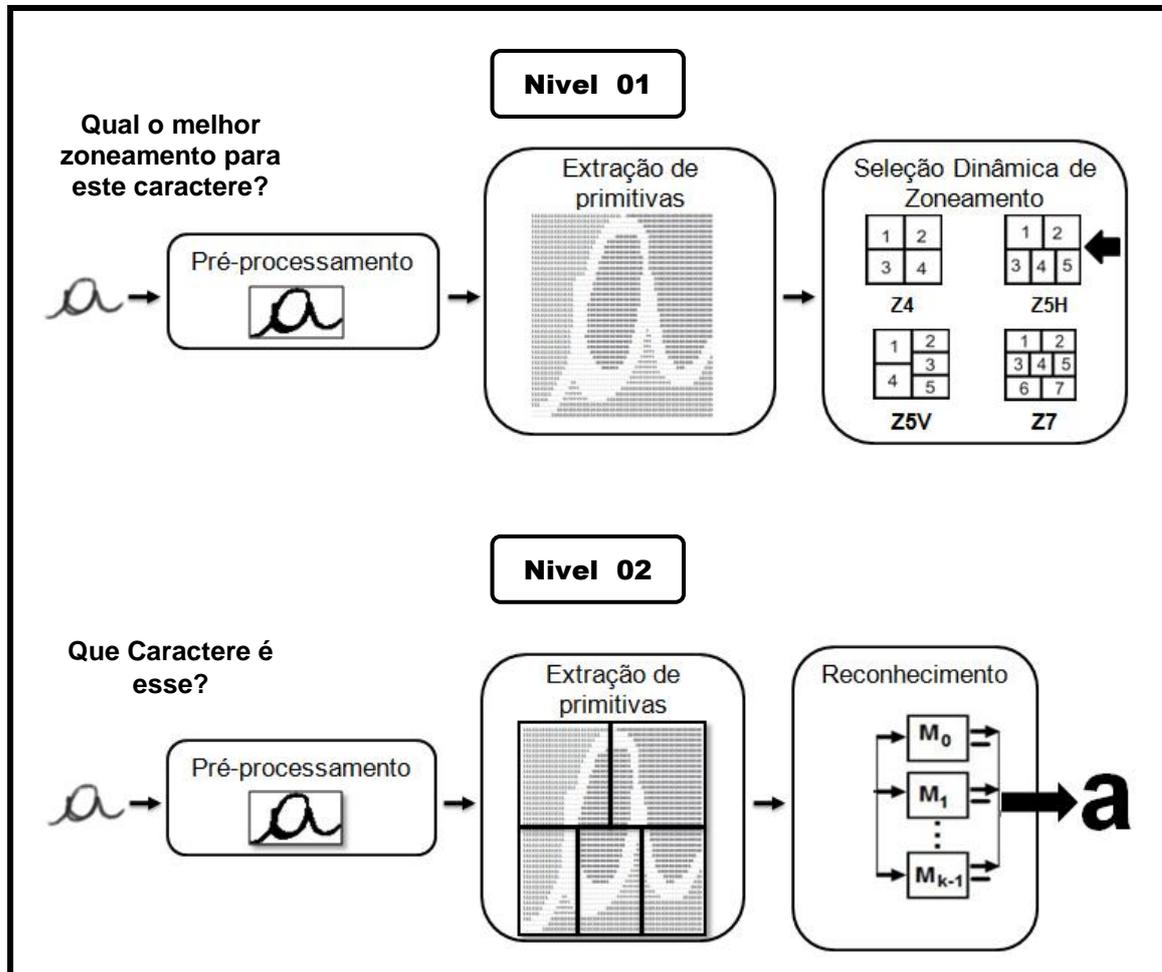


Figura 3.1: Visão geral do método proposto

Os algoritmos foram desenvolvidos em linguagem C++, sendo utilizada a biblioteca ALLEGRO para facilitar o processo de manipulação das imagens. A base de dados de caracteres manuscritos que se utilizou foi a IRONOFF, gerada pela IRESTE (*University of Nantes - France*) em colaboração com VISION OBJECTS [GAU99]. A base de dados em questão possui duas bases de manuscritos, contendo um grande número de caracteres isolados (10.685 caracteres minúsculos, 10.679 caracteres maiúsculos), dígitos (4086) e palavras cursivas (31.346 palavras de um léxico de 197 palavras). Esta base foi coletada a partir de 700 autores diferentes, principalmente de nacionalidade francesa. Cerca de dois terços dos escritores são do sexo masculino, 90% são destros, o escritor mais jovem tem 8 anos, e o mais velho 77 anos. Para este estudo foram utilizados arquivos da base *off-line*. Esta base de dados

está dividida em um léxico de 26 caracteres manuscritos minúsculos, sendo que as imagens dos caracteres já se encontram isoladas, como exemplificado na Figura 3.2. Utilizou-se ainda o JavaNNS, um simulador de redes neurais desenvolvido pela Universidade de Stuttgart para o treinamento das RNAs MLP Convencional e Classe-Modular.

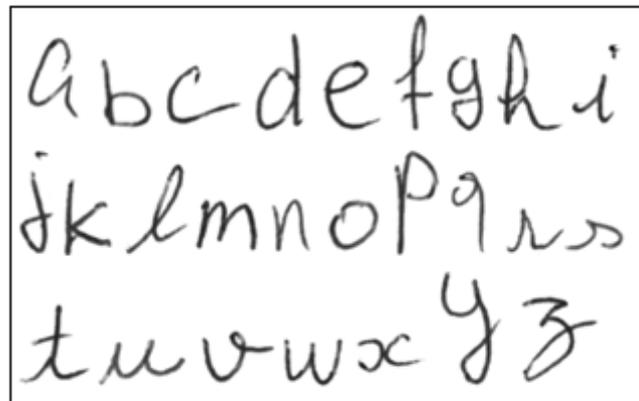


Figura 3.2: Amostras de Imagem Original - base IRONOFF

A seguir descreve-se as etapas de pré-processamento, extração e representação de primitivas e classificação utilizadas no presente trabalho.

### 3.1.1 Pré-processamento

A etapa de pré-processamento consiste basicamente de 2 tarefas:

- **Limiarização** (*thresholding*): converte a imagem original de 256 níveis de cinza para uma imagem binária (branco/preto). Para a limiarização das imagens utilizadas neste trabalho foi aplicado o algoritmo de limiarização de Otsu [OTS79], como ilustrado na Figura 3.3.

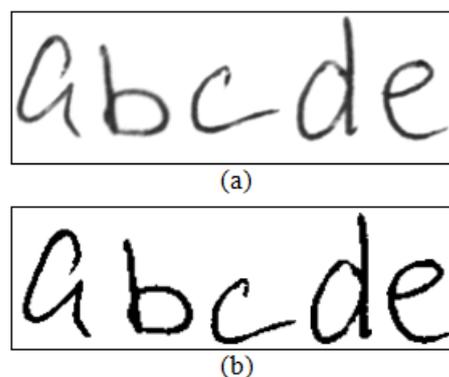


Figura 3.3. Limiarização: a) imagem original e b) imagem resultante - método de Otsu.

- **Bounding box:** Para auxiliar no processo de extração de primitivas, bem como, de zoneamento, é criada uma *bounding box* ao redor do caractere na imagem, como exemplificado na Figura 3.4.

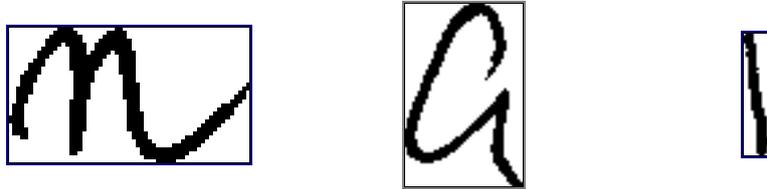


Figura 3.4: Exemplo de *Bounding Box* ao redor do caractere

### 3.1.2. Extração e Representação de Primitivas

A escolha por conjuntos de primitivas é provavelmente o fator mais importante para se obter alto desempenho nos sistemas de reconhecimento [TRI96]. Utiliza-se neste trabalho uma abordagem híbrida para o problema de extração das primitivas, utilizando uma extração global combinada com uma extração local. O método Global é composto pela rotulação dos pixels do fundo da imagem. Depois de realizada essa rotulação, a imagem do caractere é então dividida em regiões, ou seja, é aplicado um dos mecanismos de zoneamento e, assim, para cada região são obtidos os vetores de características correspondentes.

O conjunto de primitivas é baseado em convexidade e concavidade (*Concavity/Convexity Deficiencies*), tal qual proposto por Parker (1997) [PAR97] e adaptado por Aires (2005) [AIR05]. A idéia básica deste conjunto é que para cada pixel branco do fundo da imagem do caractere, procura-se em 4 direções (norte, sul, leste e oeste) até alcançar um pixel preto, correspondente ao traçado do caractere. Ainda, são realizadas sondagens para analisar os escapes a partir de cada pixel do fundo e, então, constatar a existência de laços fechados ou de concavidades/convexidades abertas, como destacado na Figura 3.5.

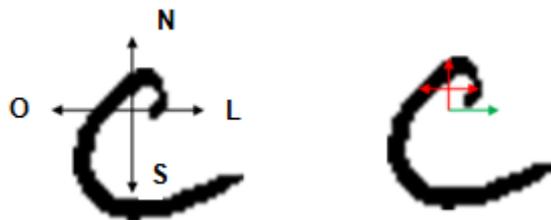


Figura 3.5. Direções de busca (NSLO) e a sondagem para verificação de escape abaixo e direita

A Tabela 3.1 resume os 23 rótulos utilizados e um exemplo de imagem rotulada é apresentado na Figura 3.6.

Tabela 3.1: Rótulos dos pixels do fundo da imagem

<b>Rótulo</b>	<b>Tipo</b>	<b>Busca</b>
0	Aberto à direita e acima	(NSLO)
1	Aberto à esquerda e acima	(NSLO)
2	Aberto à esquerda e abaixo	(NSLO)
3	Aberto à direita e abaixo	(NSLO)
4	Aberto à direita	(NSLO)
5	Aberto acima	(NSLO)
6	Aberto à esquerda	(NSLO)
7	Aberto abaixo	(NSLO)
8	Laço fechado	(NSLO)
A	Escape acima e direita	Sondagem
B	Escape acima e esquerda	Sondagem
C	Escape abaixo e direita	Sondagem
D	Escape abaixo e esquerda	Sondagem
E	Escape direita e acima	Sondagem
F	Escape esquerda e acima	Sondagem
G	Escape a direita e abaixo	Sondagem
H	Escape a esquerda e abaixo	Sondagem
J	Aberto esquerda/direita	(NSLO)
K	Aberto acima/abaixo	(NSLO)
L	Fechado abaixo	(NSLO)
M	Fechado acima	(NSLO)
N	Fechado direita	(NSLO)
O	Fechado esquerda	(NSLO)

Na sequência a extração Local de primitivas é realizada por meio do uso de mecanismos de zoneamento não simétricos e perceptivos, assim definidos por Aires (2005) [AIR05] : 5 regiões horizontais (5H), 5 regiões verticais (5V) e 7 regiões não simétricas. Além destes esquemas, foi utilizado um zoneamento composto por 4 regiões simétricas, a fim de validar os experimentos.

Para cada uma das  $Z$  regiões dos mecanismos de zoneamento é extraído um vetor de características, sendo tal vetor composto por 23 valores, cada qual representando um dos

rótulos. Todos os valores são normalizados pelo tamanho da região, sendo assim os valores no vetor de característica são representados em ponto flutuante. Para valores não encontrados, assume-se o valor igual a 0,001. A utilização deste valor justifica-se pelo fato das Redes Neurais serem aversivas ao valor zero, pois as ligações entre os neurônios são basicamente multiplicações e o uso de valores nulos retardaria o processo de aprendizagem [AIR05].

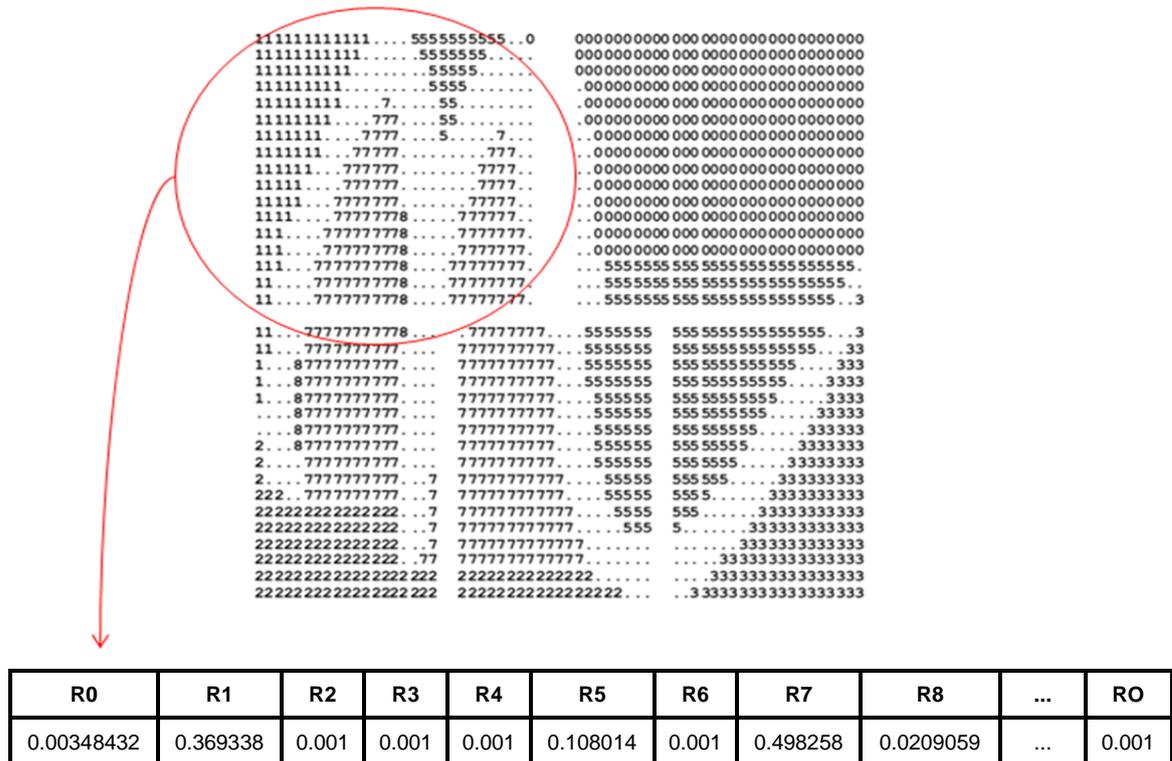


Figura 3.6. Exemplo de extração local de primitivas para um caractere “n”, e o vetor de características extraído para a primeira região do zoneamento 5H

Considerando-se o exposto, sabe-se que o problema de reconhecimento de caracteres manuscritos apresenta diferentes particularidades, tais como: linhas de intersecção, contornos e laços fechados, sendo que o método de extração de primitivas deve ser capaz de identificar estas particularidades. A Figura 3.7 representa imagens de caracteres diversos, cada qual com suas particularidades, como por exemplo, predomínio de laço fechado em um caractere “o”; laços fechados na parte superior e inferior em um caractere “g”; laços fechados e traços como em um caractere “f”; predomínio de região aberta para cima, em um caractere “u”, o que o diferencia de um caractere “a” ou “o”, é a presença de regiões convexas na parte inferior, como em um caractere “h” o que o diferencia de uma letra “l”, por exemplo.

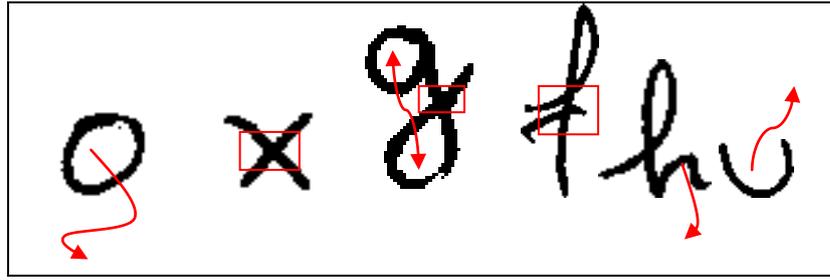


Figura 3.7. Amostras de caracteres distintos e suas particularidades

A Figura 3.8 representa as particularidades de alguns caracteres, mostrando-se os diferentes mecanismos de zoneamento para a extração de primitivas,  $Z = 4, 5H, 5V$  e  $7$ .

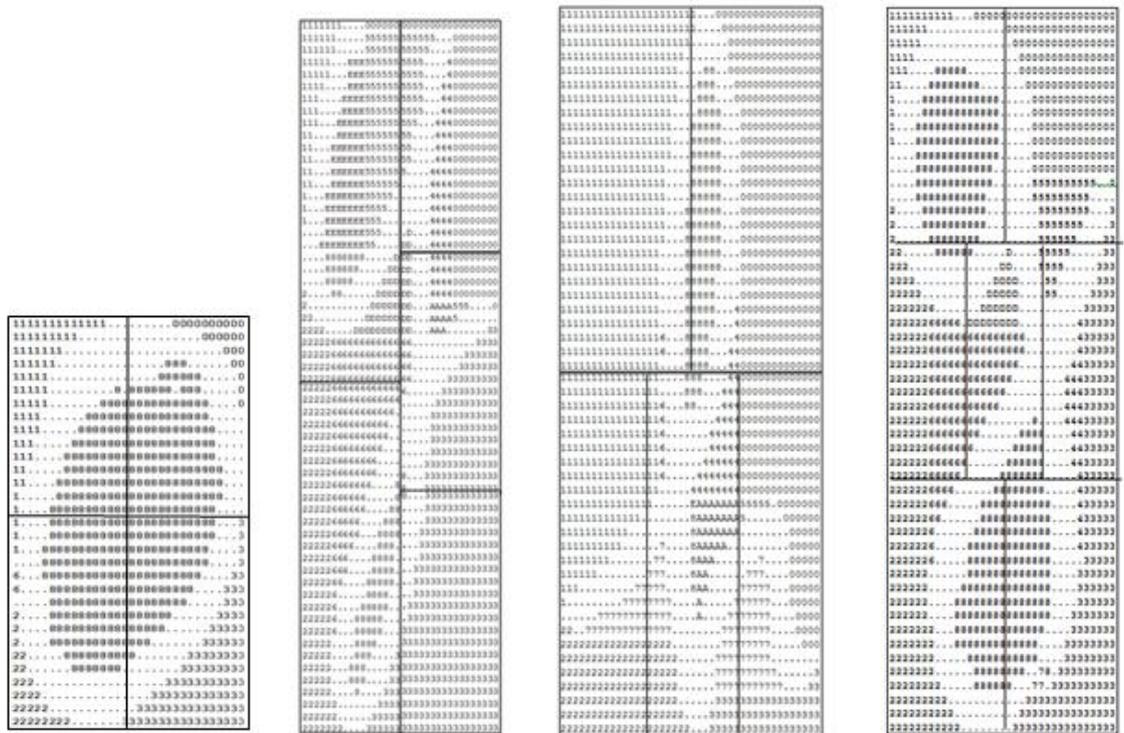


Figura 3.8. Caracteres e suas particularidades representadas através da extração de primitivas e dos mecanismos de zoneamento

### 3.1.3. Classificação

A classificação é uma das mais frequentes tarefas de tomada de decisão da atividade humana. O problema de classificação ocorre quando um objeto precisa ser atribuído a um grupo pré-definido ou classe com base em um número de atributos observados em relação a esse objeto [ZHA00].

Um problema de classificação é considerado fácil, quando as instâncias de diferentes classes são linearmente separáveis no espaço de características. Um problema de classificação se torna crítico quando instâncias de diferentes classes se sobrepõem no espaço de características. Neste caso um classificador simples não é capaz de resolver o problema de classificação [KOE03b].

Um problema relacionado ao reconhecimento de caracteres manuscritos é que existem muitas classes de caracteres que apresentam similaridades entre si, como por exemplo “p” e “q”, “m” e “n”, e que acabam se sobrepondo, ou seja, se confundindo.

No presente trabalho, como método de classificação foram utilizadas as Redes Neurais Artificiais MLP (*Multi-Layer Perceptron*) Classe-Modular, sendo feito também um estudo comparativo com as Redes Neurais Artificiais MLP Convencional. O método de classificação utilizado no presente trabalho é explicado com mais detalhes na seção 3.3.

### **3.2. Seleção Dinâmica de Mecanismos de Zoneamento – Nível 01**

Na área de reconhecimento de manuscrito a aplicação de múltiplos classificadores tem se apresentado muito útil, visto que os pesquisadores tem alcançado melhoras nas taxas de reconhecimento. Muitos métodos têm sido desenvolvidos na tentativa de melhorar as taxas de reconhecimento, e recentemente a combinação de classificadores, e/ou de diferentes métodos de extração de primitivas também vem sendo muito explorada. Um exemplo, é o trabalho de Xiang et al. (2005) [XIA05] que considera a combinação de sistemas utilizando métodos estruturais e estatísticos. Outros trabalhos consideram diferentes conjuntos de primitivas [KOE04], [AIR05] [FRE07]. O trabalho de Aires (2005) [AIR05] explora também a combinação de classificadores, tal qual proposto por Oh e Suen (2001) [OH01]. Apesar dos resultados promissores que os pesquisadores vêm alcançando em seus experimentos e dos grandes esforços empreendidos na área de reconhecimento de manuscrito, há ainda muitas questões remanescentes:

- Qual o melhor zoneamento para todas as classes de caracteres?
- Quais primitivas utilizar para o classificador selecionado?
- Utilizar seleção estática ou dinâmica de mecanismos de zoneamento?
- Criar novos classificadores ou utilizar os já existentes e explorar suas potencialidades?

À luz dessas questões investiga-se então a utilização de métodos de extração de primitivas perceptivas e a utilização de esquemas de seleção dinâmica de mecanismos de zoneamento, explorando as potencialidades de métodos já existentes, abordando não somente os mesmos problemas, mas utilizando soluções análogas para a resolução de problemas diferentes. Para o presente trabalho foram considerados os estudos e abordagens propostas por diversos pesquisadores, já mencionados neste trabalho, tais como [KO08] [FRE07] [AIR05] [KAP04] [OH01].

Estes pesquisadores acreditam que não exista um melhor ou pior método, pois para cada problema há uma solução adequada e diferente

De acordo com Ko et al. (2008) [KO08] para ter classificadores com diferentes erros, é aconselhável criar diversos classificadores e agrupá-los. De forma análoga, agrupou-se neste trabalho diferentes mecanismos de zoneamento, levando em consideração o fato que mecanismos de zoneamentos com diferentes características podem ser complementares, sem ter que necessariamente escolher somente um deles para o problema do reconhecimento de caracteres manuscritos como um todo.

Assim, no presente trabalho, propõe-se uma abordagem utilizando seleção dinâmica de mecanismos de zoneamento. Para tanto, selecionou-se aqueles mecanismos de zoneamentos que são capazes de corretamente reconhecer um dado padrão e agrupou-os em um conjunto de mecanismos de zoneamento, tal conjunto é aqui denominado de EoZ (*Ensemble of Zoning*), como mostrado na Figura 3.9.

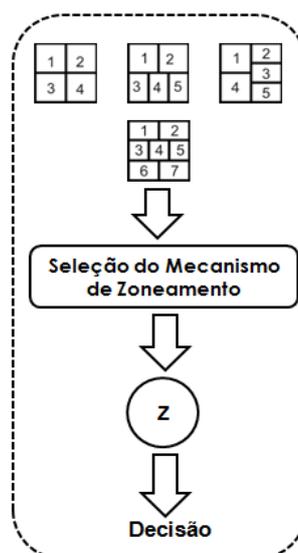


Figura 3.9: Esquema de seleção dinâmica de mecanismo de zoneamento

A escolha de um grupo adequado de mecanismo de zoneamento se deu levando em consideração que tais mecanismos de zoneamento já foram abordados por Aires (2005) [AIR05] e Freitas et al.(2007) [FRE07], para um problema semelhante: o de reconhecimento de caracteres manuscritos maiúsculos. Os experimentos realizados por Aires (2005) [AIR05] apresentaram desempenho promissor.

Estes mecanismos de zoneamento também foram investigados em esquemas estáticos no presente trabalho, e alcançaram resultados comparáveis aos apresentados na literatura. Outros mecanismos de zoneamento também foram investigados para compor o *EoZ* (*Ensemble of Zoning*), como o zoneamento com 9 partições não simétricas proposta por Radtke et al. (2003) [RAD03] e o zoneamento global (sem considerar qualquer tipo de partição), mas foram descartados por apresentarem taxas médias de reconhecimento muito abaixo do esperado. Assim, os componentes do *EoZ* são os mecanismos de zoneamento perceptivos simétrico: Z4 e não simétricos: Z5H, Z5V e Z7; propostos por Aires (2005) [AIR05] e Freitas et al. (2007) [FRE07].

Levando em consideração que o conjunto de primitivas aqui utilizado é um conjunto que já foi validado e que comprovadamente obteve resultados promissores, foi utilizado também nos experimentos um zoneamento global para comprovar a acurácia do método. O zoneamento global nada mais é que a extração global de primitivas, sem considerar qualquer tipo de zoneamento. A Figura 3.10 representa um exemplo de zoneamento global e dos mecanismos de zoneamento, com  $Z = 4$ ,  $Z = 5H$ ,  $Z = 5V$  e  $Z = 7$  partições para uma amostra de caractere “e”.

Assim, considerando o *EoZ*, para a seleção dinâmica de zoneamento, utilizou-se uma Rede Neural Artificial com arquitetura Convencional a partir da extração de primitiva combinada com o zoneamento perceptivo simétrico,  $Z = 4$  regiões. O objetivo da utilização da RNA MLP Convencional treinada com o  $Z = 4$  regiões é então prever qual o mecanismo de zoneamento mais adequado para uma determinada amostra de classe de caractere em uma base de teste. Assim, obteve-se a indicação de maneira dinâmica do mecanismo de zoneamento mais adequado a ser utilizado para o reconhecimento de cada classe de caractere manuscrito minúsculo, como apresentado no Figura 3.1 - Nível 01.

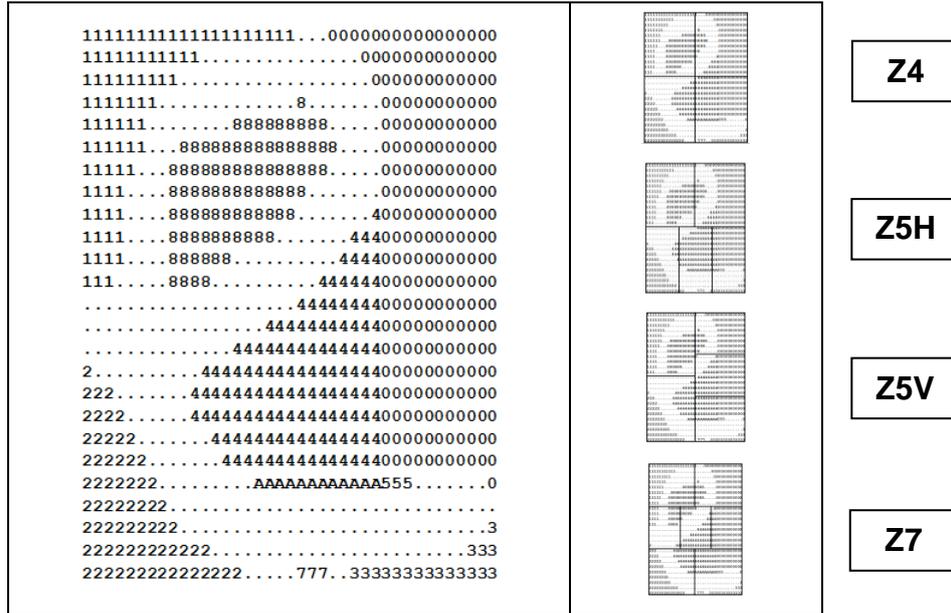


Figura 3.10: Zoneamento Global, 4, 5H, 5V e 7 para um caractere “e”

A indicação do melhor mecanismo de zoneamento é obtido realizando os passos:

1. A partir das RNA MLP Classe-Modular, configuradas e treinadas com os zoneamentos Z=4, Z=5H, Z=5V e Z=7 constrói-se as matrizes de confusão a serem usadas nesse primeiro nível do sistema proposto. Assim, são construídas 4 matrizes de confusão, uma para cada zoneamento do EoZ. A Figura 3.11 mostra de forma ilustrativa a criação das 4 matrizes a partir das RNA MLP Classe-Modular.

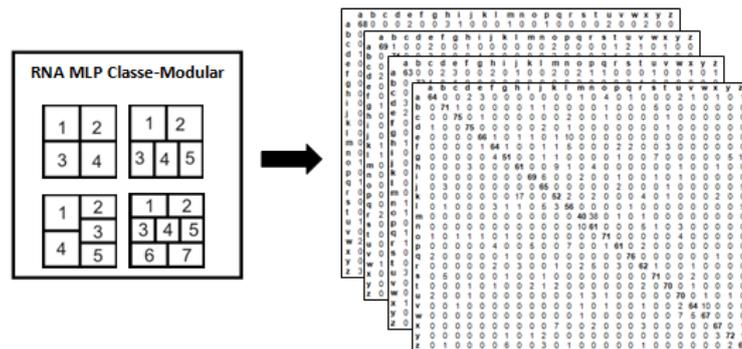


Figura 3.11. Criação das Matrizes de confusão

2. Utilizando a RNA MLP Convencional treinada com o zoneamento Z=4, cada amostra na base de teste é então submetida a entrada da rede. A RNA MLP Convencional é configurada com 92 entradas e 26 saídas, cada saída representando uma classe de

caractere minúsculo. Considera-se então as três saídas que obtiverem maior valor. A Figura 3.12 ilustra essa situação.

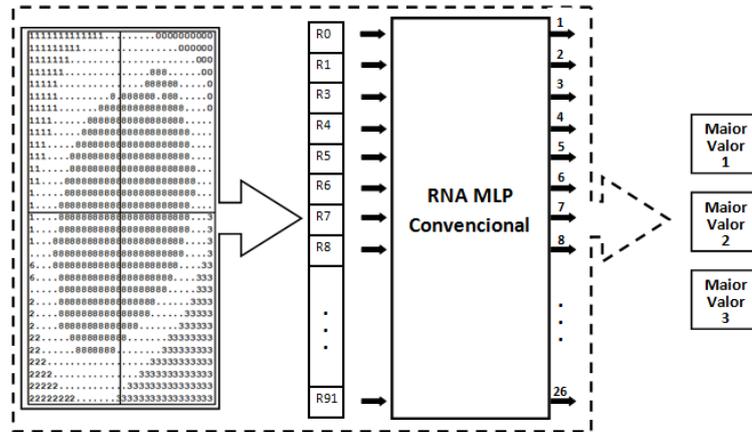


Figura 3.12. Obtenção dos 3 maiores valores de saída

3. A partir desses três valores, cada qual representando uma classe de caractere, é então feita uma análise, entre as confusões dessas três classes nas matrizes de confusão construídas a partir das RNA MLP Classe-Modular, cada valor representando uma classe é comparada 2 a 2. Dessa análise são extraídos 6 valores de confusão, como mostrado na Figura 3.13. Isto é realizado para cada uma das 4 matrizes de confusão obtidas com as RNA MLP Classe-Modular. Sendo assim cada Matriz de confusão possui 6 valores de confusão.

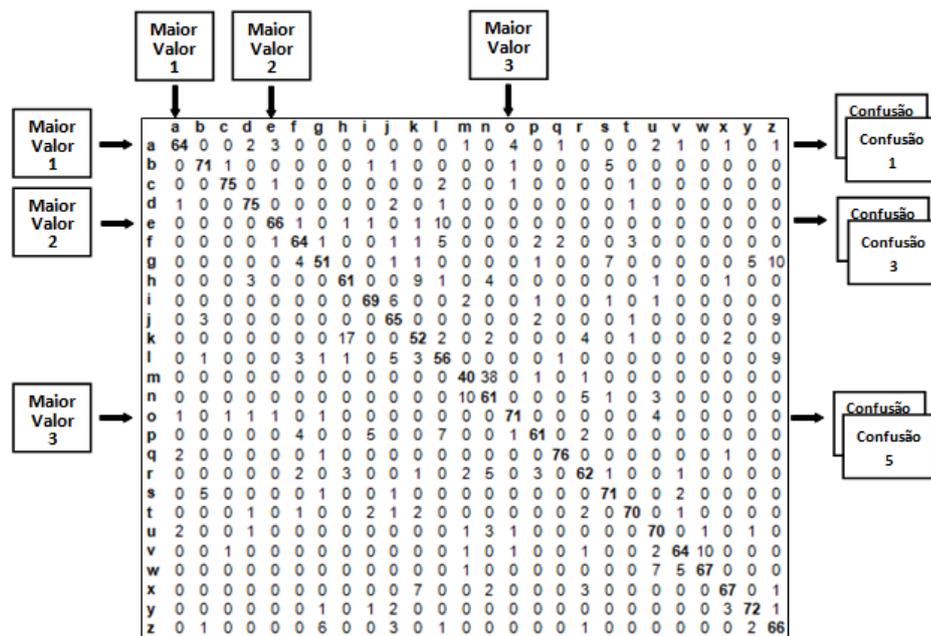


Figura 3.13. Obtenção dos 6 valores de confusão a partir da matriz de confusão

4. A seguir é realizada então uma comparação entre os valores das confusões encontradas, matriz a matriz, deste modo o zoneamento que obteve as menores confusões, é o zoneamento mais indicado para a amostra do caractere desconhecido. Em caso de empate entre os zoneamentos, é considerada então a maior taxa média de reconhecimento, isto posto que tem-se a classe de caractere que obteve o maior valor de saída. A Figura 3.14 ilustra a indicação do zoneamento que obteve as menores confusões, no caso do exemplo o zoneamento  $Z=5H$ .

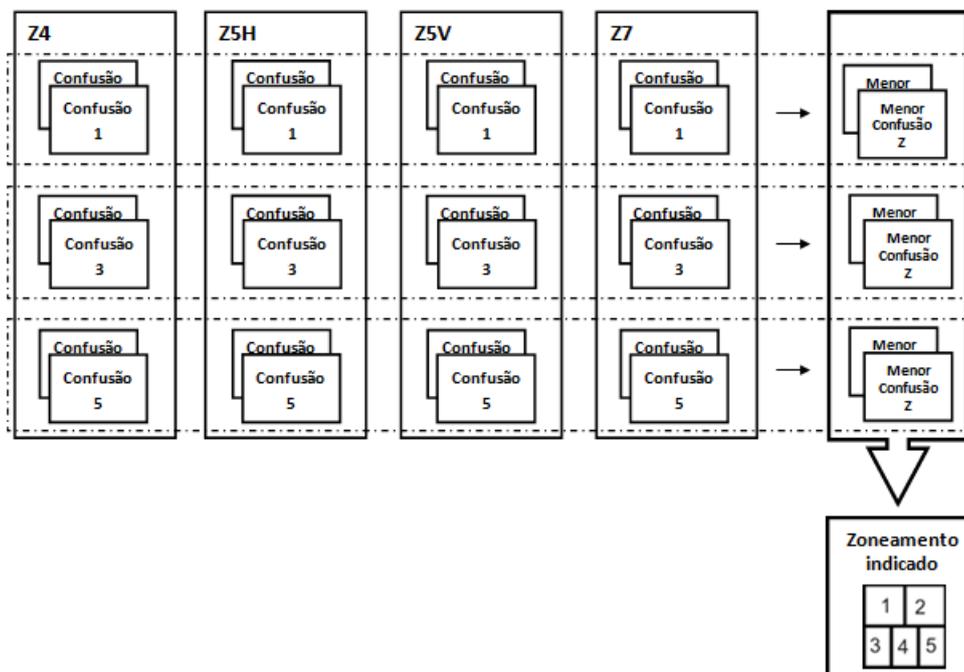


Figura 3.14. Indicação do mecanismo de zoneamento

### 3.3. Reconhecimento de Caracteres Manuscritos Minúsculos – Nível 02

Depois de realizada a predição do zoneamento mais adequado a cada amostra de caractere, submeteu-se cada amostra de caractere a uma RNA MLP-Classe-Modular já treinada com o zoneamento indicado pela Rede Neural Artificial do Nível 01. Para cada mecanismo de zoneamento, foram treinadas 26 subredes, uma para cada classe de caracteres manuscritos minúsculos, usando-se o conceito de modularidade, como proposto por Oh & Suen (2001) [OH01].

Utilizando o conceito Classe-Modular (*Class-Modular*), o problema original da classificação decompõe-se em subproblemas. Utilizando a arquitetura Classe-Modular, para cada classe de caractere manuscrito minúsculo desenvolveu-se uma subrede. Assim foram treinadas e validadas 26 subredes para cada mecanismo de zoneamento pertencente ao EoZ. Cada módulo (subrede) é treinado independentemente um do outro, utilizando conjuntos de treinamento e validação. A base de treinamento utilizada nos experimentos é composta de 240 amostras de cada classe de caractere e mais 9 amostras distintas das demais classes de caracteres, totalizando assim 465 amostras. Para a base de validação, considerou-se 80 amostras de cada classe de caractere e mais 3 amostras distintas para cada uma das demais classes de caractere, totalizando assim 155 amostras. Observa-se que essa configuração para as bases de treinamento e validação é necessária devido ao fato de se trabalhar com RNA MLP Classe-Modular, na qual o treinamento necessita conhecer o que é a classe e o que não é a classe a ser treinada.

O número de entradas das subredes é calculado a partir do mecanismo de zoneamento e do número de rótulos. Considerando o número de regiões e os 23 rótulos propostos, por exemplo, para o zoneamento simétrico composto por 4 regiões, tem-se 92 entradas (23 rótulos  $\times$  4 regiões) e uma saída. A Figura 3.15 ilustra um exemplo esquemático de uma subrede para o caractere “o”. Os resultados experimentais e as análises dos experimentos com as RNA MLP Classe-Modular são apresentados no Capítulo 4.

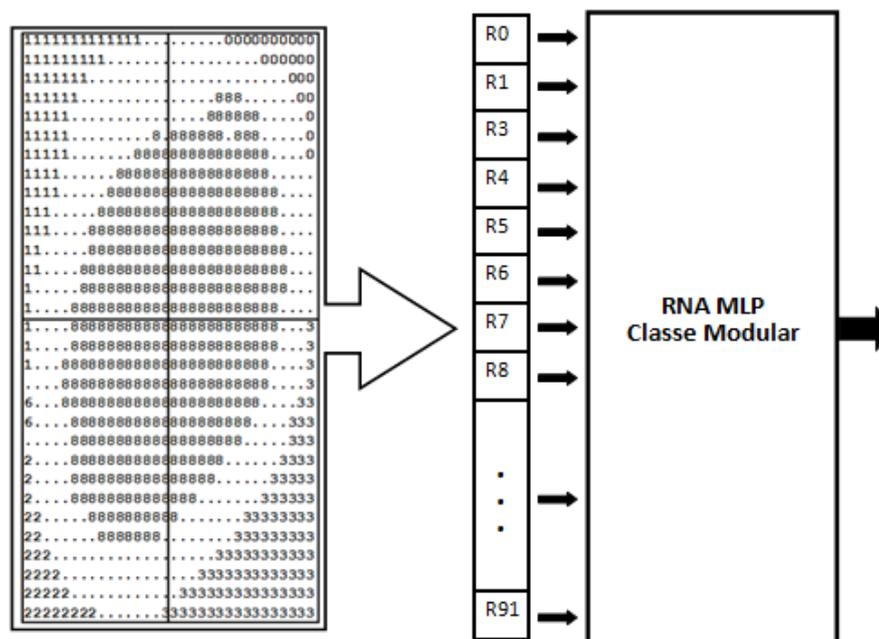


Figura 3.15 – esquema de subrede para o caractere “o”

### 3.4. Considerações Finais

Neste capítulo foi apresentada uma visão geral do método, mostrando os 2 níveis que compõem a método proposto:

- Nível 01: Utilização de um esquema de seleção dinâmica, a partir de um conjunto de mecanismos de zoneamento (EoZ), que faz a predição do melhor mecanismo de zoneamento para cada classe de caractere manuscrito minúsculo. Tal conjunto é composto por quatro mecanismos de zoneamento,  $Z = 4, 5H, 5V$  e  $7e$
- Nível 02: Que representa a fase de reconhecimento propriamente dita, ou seja, é neste nível que é realizado o reconhecimento por meio do mecanismo de zoneamento indicado no Nível 01, com a utilização das RNA MLP Classe-Modular.

Abordou-se também neste capítulo o método de extração de primitivas e os mecanismos de zoneamentos.

## Capítulo 4

### Resultados Experimentais

Este capítulo detalha os experimentos realizados tanto para o Nível 01 quanto para o Nível 02 do método proposto. Além disto, apresenta-se também uma análise mais aprofundada dos resultados experimentais. Finalmente é mostrada também uma análise comparativa sobre as diferenças entre caracteres manuscrito maiúsculos e minúsculos em termos de características perceptivas.

#### 4.1. Experimentos com RNA MLP Convencional

Os Experimentos com as Redes Neurais MLP Convencional, foram realizados utilizando mecanismos de zoneamento não simétricos ( $Z = 5H$ ,  $Z = 5V$  e  $Z = 7$ ) propostos por Aires (2005) [AIR05], um mecanismo de zoneamento simétrico ( $Z = 4$ ) e ainda um mecanismo não simétrico ( $Z = 9$ ) proposto por Radtke (2003) [RAD03], para fins de comparação.

Para todos os mecanismo de zoneamentos investigados utilizou-se o algoritmo de aprendizagem *Backpropagation* padrão, com parâmetro de aprendizagem também padrão 0,2. O número de neurônios na camada de entrada foi definido multiplicando o número de rótulos pelo número de partições do zoneamento. Foi considerado uma saída para cada classe de caractere manuscrito minúsculo, sendo assim cada RNA MLP Convencional foi configurada com 26 neurônios na camada de saída. Em relação ao número de neurônios na camada intermediária, este é em geral definido empiricamente [BRA00], e foi definido inicialmente em função da média aritmética entre o número de unidades de entrada e de saída. Em torno do

valor médio, foram testados outros valores com o intuito de se obter uma configuração de rede que obtivesse a maior taxa de reconhecimento.

Assim, para o mecanismo de zoneamento não simétrico  $Z = 9$ , foi utilizada uma Rede Neural MLP Convencional com 63 neurônios na camada escondida. Para o mecanismo de zoneamento simétrico  $Z = 4$ , foi utilizada uma Rede Neural MLP Convencional com 59 neurônios na camada escondida. Para os mecanismos de zoneamento não simétricos  $Z = 5H$  e  $Z = 5V$  utilizou-se uma Rede Neural MLP Convencional com 50 neurônios na camada escondida e, finalmente, para o zoneamento não simétrico  $Z = 7$  utilizou-se uma Rede Neural MLP Convencional com 73 neurônios na camada escondida.

O conjunto de dados utilizado nos experimentos (testes) foi composto por 400 imagens para cada classe de caractere manuscrito minúsculo, totalizando assim 10.400 amostras de imagens de caracteres manuscritos minúsculos. As bases de treinamento, validação e testes foram retiradas da base de dados IRONOFF e dividido em conjuntos que respeitam as seguintes porcentagens: treinamento (60%), validação (20%) e teste (20%). O conjunto de treinamento foi utilizado na modificação dos pesos, e o conjunto de validação foi utilizado para estimar a capacidade de generalização da rede durante o processo de aprendizagem. O treinamento das redes foi interrompido quando o erro do conjunto de validação começou a aumentar [NIE04]. Estimar o erro de generalização durante o processo de treinamento é uma forma de se evitar o *overfitting* (após certo ciclo do treinamento, a rede em vez de melhorar, começa a piorar a taxa de acertos para padrões diferentes daqueles utilizados para o ajuste dos pesos). O conjunto de teste foi utilizado para testar a capacidade de generalização da rede e, assim, obter-se as taxas de reconhecimento, visto que uma rede apresenta capacidade de generalização quando classifica corretamente padrões não-utilizados no treinamento [BRA00].

As taxas médias de reconhecimento obtidas nos experimentos realizados com os mecanismos de zoneamento simétrico  $Z = 4$  e não simétricos  $Z = 9$ ,  $Z = 5H$ ,  $Z = 5V$  e  $Z = 7$  foram de 79,52%, 77,88%, 80,48%, 79,81% e 80,82%, respectivamente.

A Tabela 4.1 detalha a relação entre as classes de caracteres manuscritos minúsculos e as taxas médias de reconhecimento obtidas nos experimentos com as RNA MLP Convencional.

Tabela 4.1. Classe e reconhecimento por zoneamento – RNA MLP Convencional

Classe	Mecanismos de Zoneamento (Z)				
	9	4	5H	5V	7
a	72,50	83,75	82,50	80,00	83,75
b	85,00	82,50	86,25	83,75	86,25
c	95,00	95,00	97,50	93,75	95,00
d	87,50	92,50	90,00	91,25	86,25
e	65,00	83,75	83,75	77,50	63,75
f	73,75	38,75	67,50	67,50	65,00
g	50,00	65,00	71,25	67,50	61,25
h	48,75	76,25	72,50	67,50	68,75
i	80,00	81,25	82,50	82,50	77,50
j	78,75	77,50	78,75	73,75	81,25
k	82,50	66,25	73,75	62,50	76,25
l	60,00	70,00	58,75	66,25	68,75
m	57,50	48,75	71,25	60,00	61,25
n	71,25	81,25	65,00	63,75	77,50
o	90,00	88,75	90,00	93,75	91,25
p	92,50	81,25	87,50	82,50	87,50
q	96,25	95,00	93,75	98,75	97,50
r	71,25	71,25	67,50	72,50	73,75
s	87,50	85,00	77,50	81,25	86,25
t	86,25	91,25	78,75	91,25	87,50
u	73,75	83,75	83,75	78,75	82,50
v	66,25	71,25	81,25	83,75	83,75
w	87,50	88,75	92,50	87,50	90,00
x	81,25	88,75	82,50	86,25	86,25
y	88,75	92,50	90,00	92,50	87,50
z	96,25	87,50	86,25	88,75	95,00
<b>Média(%)</b>	<b>77,88</b>	<b>79,52</b>	<b>80,48</b>	<b>79,81</b>	<b>80,82</b>

## 4.2. Experimentos com RNA MLP Convencional – Z4 – Nível 01

Para compor o Nível 01 do método proposto, foi utilizado a RNA MLP Convencional considerando 4 regiões simétricas ( $Z = 4$ ). Foram utilizados para a base de treinamento 240 amostras para cada classe de caractere, 80 amostras para a base de validação e para a base de teste 80 amostras para cada classe de caractere.

Para os experimentos considerou-se o zoneamento com 4 regiões simétricas por ser o zoneamento mais tradicional e por apresentar melhor resultado se comparado com os experimentos aqui realizados com um zoneamento global (sem considerar nenhum tipo de divisão).

Considerou-se nesse Nível os três maiores valores, ou seja as três classes de caracteres que obtiveram o maior valor de saída. Assim sendo, com base nas confusões entre as classes, utilizando as matrizes de confusão obtidas nos experimentos com RNA MLP Classe-Modular, determinou-se o zoneamento mais indicado para cada amostra de caractere na base de teste. A indicação do zoneamento mais indicado leva em consideração a menor confusão entre as classes com os três maiores valores de saída. Um exemplo das três classes de caracteres que obtiveram os maiores valores de saída pode ser visualizado na Figura 4.1. Assim, para uma amostra de caractere desconhecido, os maiores valores de saída foram para as três classes de caracteres: “a”, “e”, “o”. Fazendo uma comparação entre as confusões, utilizando as matrizes de confusão dos zoneamentos Z4, Z5H, Z5V e Z7, obtidos do treinamento com as RNA MLP Classe-Modular, encontra-se as menores confusões entre os pares (“a” e “e”; “a” e “o”; “e” e “a”; “e” e “o”; “o” e “a”; “o” e “e”). Deste modo o zoneamento que obteve as menores confusões, é o zoneamento mais indicado para a amostra do caractere desconhecido. Em caso de empate entre os zoneamentos, é considerada então a maior taxa média de reconhecimento, isto posto que tem-se a classe de caractere que obteve o maior valor de saída.

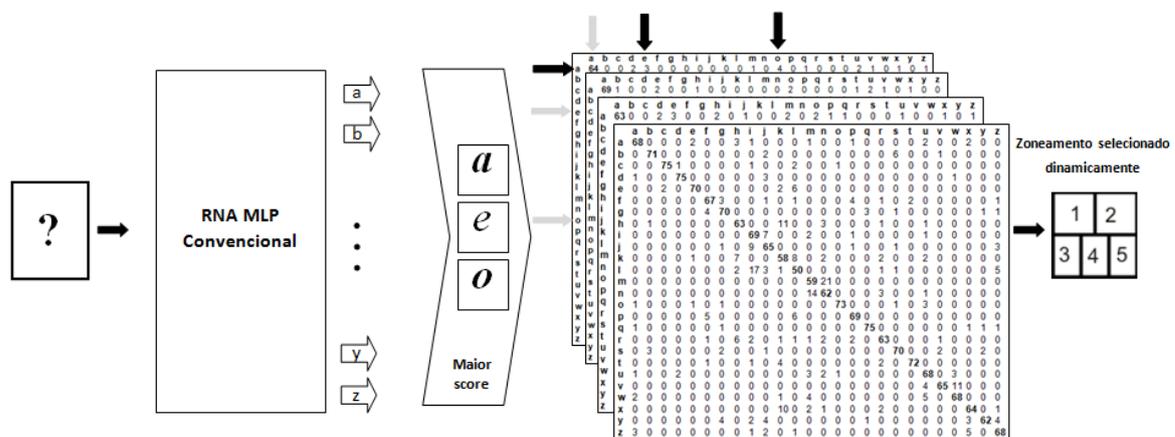


Figura 4.1 – esquema de indicação do melhor zoneamento para um caractere desconhecido

### 4.3. Experimentos com RNA MLP- Classe Modular – Nível 02

Para os experimentos com as RNAs MLP Classe-Modular foram utilizadas 400 imagens para cada classe de caractere, divididas em 3 conjuntos: Treinamento (60%), Validação (20%) e Teste (20%). Foram avaliados os zoneamento: Z = 4, 5H, 5V e 7. E,

ainda, um zoneamento não simétrico com 9 regiões proposto por Radtke (2003) [RAD03]. Nos experimentos utilizando a arquitetura RNA MLP Classe-Modular é necessário uma reorganização nas bases de treinamento e validação. Assim para a base de treinamento foram adicionados mais 9 amostras de cada classe de caractere e para a base de validação foram adicionados mais 3 amostras de cada classe de caractere.

Nos experimentos com as RNA MLP Classe-Modular, o zoneamento não simétrico Z9 utilizou 63 neurônios na camada escondida, sendo que os demais mecanismos de zoneamento utilizaram: 40 (Z4), 50 (Z5H e Z5V) e 63 (Z7). Todos os zoneamentos citados utilizaram o algoritmo de aprendizagem *Backpropagation* padrão, com parâmetro de aprendizagem 0,2. Do mesmo modo como nos experimentos com as RNA MLP Convencional, o número de neurônios foi definido empiricamente (inicialmente em função da média aritmética entre o número de unidades de entrada e de saída).

Nos experimentos realizados, o mecanismo de zoneamento não simétrico Z9, obteve taxa de reconhecimento de 79,62%, mostrando que este mecanismo treinado e validado na base de dados NIST não pode ser generalizado para outras bases de dados. Os outros mecanismos de zoneamento obtiveram 81,10% (Z4), 80,81% (Z5H), 81,63% (Z5V) e 83,61% (Z7) como taxa de reconhecimento.

A Tabela 4.2 ilustra os resultados obtidos em experimentos com as RNA MLP Classe-Modular, mostrando as taxas de reconhecimento para cada classe de caractere nos zoneamentos analisados (Z9, Z4, Z5H, Z5V e Z7), os valores em negrito representam o maior reconhecimento de uma classe por um determinado mecanismo de zoneamento. Ao se comparar estes valores com os resultados obtidos por Aires (2005) [AIR05] para caracteres manuscritos maiúsculos observa-se que as taxas de reconhecimento para os zoneamentos Z4, Z5H e Z7 foram maiores, todavia a taxa média de reconhecimento para caracteres manuscritos minúsculos obtida com o zoneamentos não simétrico Z5V foi mais alta. As taxas médias de reconhecimento obtidas por Aires (2005) [AIR05] em seus experimentos utilizando RNA MLP Classe-Modular no reconhecimento de caracteres manuscritos maiúsculos foram: Z4: 82,89%, Z5H: 81,75%, Z5V: 80,94% e Z7: 84,73% .

Comparando-se ainda os valores dos experimentos das RNA MLP Classe-Modular e os resultados obtidos nos experimentos com as RNA MLP Convencional (Seção 4.1) para o reconhecimento de caracteres manuscritos minúsculos, percebe-se o aumento nas taxas

médias de reconhecimento de todos os zoneamentos investigados. As matrizes de confusão com os experimentos RNA MLP Classe-Modular são mostradas e analisadas na Seção 4.4.

Tabela 4.2. Classe e reconhecimento por zoneamento – RNA MLP Classe-Modular

Classe	Mecanismo de Zoneamento (Z)				
	9	4	5H	5V	7
a	76,25	80,00	<b>86,25</b>	78,75	85,00
b	86,25	88,75	88,75	<b>90,00</b>	88,75
c	90,00	93,75	<b>95,00</b>	88,75	93,75
d	80,00	<b>93,75</b>	91,25	81,25	<b>93,75</b>
e	<b>87,50</b>	82,50	83,75	77,50	<b>87,50</b>
f	85,00	80,00	68,75	80,00	<b>83,75</b>
g	81,25	63,75	65,00	73,75	<b>87,50</b>
h	73,75	76,25	<b>80,00</b>	70,00	78,75
i	83,75	86,25	<b>88,75</b>	86,25	86,25
j	81,25	<b>81,25</b>	77,50	<b>81,25</b>	<b>81,25</b>
k	66,25	65,00	<b>72,50</b>	<b>72,50</b>	<b>72,50</b>
l	72,50	70,00	62,50	<b>72,50</b>	62,50
m	77,50	50,00	53,75	<b>85,00</b>	73,75
n	68,75	76,25	65,00	<b>82,50</b>	77,50
o	86,25	88,75	87,50	90,00	<b>91,25</b>
p	86,25	76,25	87,50	<b>88,75</b>	86,25
q	<b>95,00</b>	<b>95,00</b>	92,50	88,75	93,75
r	73,75	77,50	72,50	66,25	<b>78,75</b>
s	55,00	<b>88,75</b>	78,75	81,25	87,50
t	81,25	87,50	88,75	87,50	<b>90,00</b>
u	<b>90,00</b>	87,50	83,75	68,75	85,00
v	87,50	80,00	<b>82,50</b>	80,00	81,25
w	73,75	83,75	87,50	<b>90,00</b>	85,00
x	78,75	83,75	<b>86,25</b>	85,00	80,00
y	73,75	<b>90,00</b>	<b>90,00</b>	87,50	77,50
z	78,75	82,50	85,00	<b>88,75</b>	85,00
Média(%)	<b>79,62</b>	<b>81,10</b>	<b>80,81</b>	<b>81,63</b>	<b>83,61</b>

#### 4.4. Reconhecimento de Caracteres Manuscritos com Seleção Dinâmica de Mecanismos de Zoneamento - Nível 01 + Nível 02

O experimento descrito na Seção 4.2 com a RNA MLP Convencional e o mecanismo de zoneamento simétrico  $Z = 4$  faz parte do Nível 01 do método proposto. Deste modo, o Nível 01 seleciona dinamicamente o zoneamento mais indicado para cada amostra de

caractere a partir da RNA MLP Convencional e do zoneamento Z= 4. Esta predição é feita levando em consideração as 3 saídas que obtiveram os maiores valores na RNA MLP Convencional. O Nível 02 utiliza o zoneamento previamente indicado no Nível 01 para o reconhecimento através das RNA MLP Classe-Modular.

A taxa média de reconhecimento alcançada considerando os dois níveis (Nível 01 + Nível 02) do método proposto foi de 88,13%, sendo os resultados do reconhecimento com os zoneamentos dinamicamente selecionados utilizando as RNA MLP Classe-Modular mostrados na Tabela 4.3.

Tabela 4.3. Taxa de Reconhecimento - Zoneamento dinâmico

<b>Classe</b>	<b>Reconhecimento (%)</b>
a	85,00
b	91,25
c	96,25
d	93,75
e	87,50
f	86,25
g	88,75
h	83,75
i	90,00
j	78,75
k	80,00
l	80,00
m	82,50
n	82,50
o	92,50
p	92,50
q	95,00
r	82,50
s	95,00
t	92,50
u	91,25
v	86,25
w	90,00
x	87,50
y	91,25
z	88,75
<b>Média (%)</b>	<b>88,13</b>

## 4.5. Análise dos Resultados

Nas subseções que se seguem são apresentados os resultados e a análise das matrizes de confusão obtidas nos experimentos com caracteres manuscritos minúsculos. Também é apresentada uma análise sobre caracteres manuscritos maiúsculos e minúsculos, com o objetivo de se entender as diferenças e similaridades existentes entre eles em função das características perceptivas de cada um.

### 4.5.1. Evolução dos Classificadores: RNA MLP Convencional e Classe-Modular

Na arquitetura RNA MLP Convencional as classes são treinadas juntas e a que apresentar maior valor entre as saídas, é considerada a classe reconhecida. Sendo assim, tem-se uma única rede para todas as 26 classes representadas pelos caracteres manuscritos minúsculos. Já na RNA MLP Classe-Modular, há uma rede para cada classe de caractere, totalizando assim 26 subredes independentes, o que torna mais fácil modificar as configurações individuais de cada rede. Na RNA MLP Convencional tem-se então 26 saídas, uma saída para cada classe de caractere e na RNA MLP Classe-Modular apenas 2 saídas são necessárias.

Na RNA MLP Convencional a rede demora mais a convergir, e não é possível otimizar ou modificar o desempenho do reconhecimento de um único caractere especificamente sem que o desempenho de reconhecimento dos outros caracteres seja afetado. Além disto, não se tem garantia ou controle de que o desempenho dessas outras classes de caracteres seja afetado de forma positiva. Já na MLP Classe-Modular a configuração para cada sub-rede pode ser modificada e otimizada separadamente sem afetar o desempenho das outras sub-redes, pois cada módulo de rede é treinado independentemente [OH01].

Na Tabela 4.4 apresenta-se as taxas de reconhecimento para os mecanismos de zoneamentos utilizando as RNAs MLP Convencional e Classe-Modular. Analisando os resultados a partir das taxas médias de reconhecimento para cada mecanismo de zoneamento, tem-se que os experimentos com a arquitetura RNA MLP Classe-Modular obteve desempenho superior se comparado com os experimentos com as RNA MLP Convencional, comprovando assim que a modularidade das redes, combinadas com os mecanismos de zoneamentos se mostram um classificador promissor. Esta tabela apresenta também os resultados obtidos considerando um zoneamento global, sem partições, ou seja, somente com a aplicação da extração de primitivas e obtenção de um único vetor de características.

Observa-se o baixo desempenho obtido e, novamente, confirma-se que a aplicação dos mecanismos de zoneamento auxilia no reconhecimento de caracteres manuscritos.

Tabela 4.4. Reconhecimento RNA MLP Convencional e Classe-Modular

Classe	MLP- Convencional					MLP Classe Modular			
	ZGlobal	Z4	Z5H	Z5V	Z7	Z4	Z5H	Z5V	Z7
a	68.75	83.75	82.50	80.00	83.75	80.00	86.25	78.75	85.00
b	61.25	82.50	86.25	83.75	86.25	88.75	88.75	90.00	88.75
c	71.25	95.00	97.50	93.75	95.00	93.75	95.00	88.75	93.75
d	42.50	92.50	90.00	91.25	86.25	93.75	91.25	81.25	93.75
e	76.25	83.75	83.75	77.50	63.75	82.50	83.75	77.50	87.50
f	26.25	38.75	67.50	67.50	65.00	80.00	68.75	80.00	83.75
g	40.00	65.00	71.25	67.50	61.25	63.75	65.00	73.75	87.50
h	37.50	76.25	72.50	67.50	68.75	76.25	80.00	70.00	78.75
i	80.00	81.25	82.50	82.50	77.50	86.25	88.75	86.25	86.25
j	70.00	77.50	78.75	73.75	81.25	81.25	77.50	81.25	81.25
k	61.25	66.25	73.75	62.50	76.25	65.00	72.50	72.50	72.50
l	53.75	70.00	58.75	66.25	68.75	70.00	62.50	72.50	62.50
m	13.75	48.75	71.25	60.00	61.25	50.00	53.75	85.00	73.75
n	82.50	81.25	65.00	63.75	77.50	76.25	65.00	82.50	77.50
o	82.50	88.75	90.00	93.75	91.25	88.75	87.50	90.00	91.25
p	76.25	81.25	87.50	82.50	87.50	76.25	87.50	88.75	86.25
q	87.50	95.00	93.75	98.75	97.50	95.00	92.50	88.75	93.75
r	65.00	71.25	67.50	72.50	73.75	77.50	72.50	66.25	78.75
s	76.25	85.00	77.50	81.25	86.25	88.75	78.75	81.25	87.50
t	86.25	91.25	78.75	91.25	87.50	87.50	88.75	87.50	90.00
u	50.00	83.75	83.75	78.75	82.50	87.50	83.75	68.75	85.00
v	62.50	71.25	81.25	83.75	83.75	80.00	82.50	80.00	81.25
w	83.75	88.75	92.50	87.50	90.00	83.75	87.50	90.00	85.00
x	82.50	88.75	82.50	86.25	86.25	83.75	86.25	85.00	80.00
y	90.00	92.50	90.00	92.50	87.50	90.00	90.00	87.50	77.50
z	88.75	87.50	86.25	88.75	95.00	82.50	85.00	88.75	85.00
<b>Média</b>	66.01	79.52	80.48	79.81	<b>80.82</b>	81.11	80.82	81.63	<b>83.61</b>

#### 4.5.2. Análise das Matrizes de Confusão

A análise das matrizes de confusão é feita com o objetivo de investigar e entender a relação entre as regiões e as confusões geradas entre as classes de caracteres. Nos experimentos para o mecanismo de zoneamento não simétrico com  $Z = 9$  regiões, apesar do aumento na taxa média de reconhecimento se comparado com os experimentos realizados com a RNA MLP Convencional, os resultados obtidos não foram muito animadores. Considerando as confusões e as taxas médias de reconhecimento de caracteres manuscritos minúsculos comparando com os outros zoneamentos analisados, o desempenho foi menor.

Algumas confusões encontradas entre classes distintas para os experimento com o zoneamento não simétrico  $Z = 9$  foram:

- “a” e “b”;
- “b”, “f” e “s”;
- “e” e “l”;
- “h”, “i”, “k”, “m” e “n”;
- “i”, “j” e “l”;
- “k”, “h” e “l”;
- “m” e “n”;
- “p” e “l”;
- “r”, “e”, “i” e “n”;
- “s”, “n” e “r”;
- “t”, “e” e “i”;
- “w”, “m”, “u” e “v”;
- “x”, “k”, “m” e “n”;
- “y”, “g”, “i” e “j”;
- “z”, “a”, “g”, “j” e “x”.

A Tabela 4.5 mostra os resultados obtidos nos experimentos e as confusões no reconhecimento entre as classes de caracteres, ressaltando algumas das confusões identificadas. Analisando a matriz de confusão, pode-se perceber que o excesso de regiões no zoneamento não contribuiu para o reconhecimento, ou seja, as partições do zoneamento não auxiliaram na distinção de partes similares.

Tabela 4.5. Matriz de Confusão Z=9 Classe Modular

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	%Rec
a	61	5	1	3	1	1	1	0	1	0	1	1	1	1	1	0	0	0	0	0	1	0	0	0	0	0	76,25
b	2	69	0	0	1	4	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	1	0	0	0	0	86,25
c	0	1	72	0	2	0	0	0	1	0	0	2	0	1	0	0	0	0	0	0	1	0	0	0	0	0	90,00
d	1	0	0	64	1	0	0	0	2	3	0	1	1	1	0	0	0	0	0	0	2	0	1	0	1	2	80,00
e	0	0	0	0	70	0	0	0	1	0	0	8	0	0	0	0	0	0	0	0	1	0	0	0	0	0	87,50
f	0	0	0	0	0	68	1	0	0	0	1	2	0	0	0	2	0	3	2	1	0	0	0	0	0	0	85,00
g	0	1	0	0	0	7	65	0	0	3	0	1	0	1	0	0	0	0	0	0	0	0	0	0	2	0	81,25
h	0	0	0	0	0	0	1	59	3	0	5	1	6	4	0	0	0	0	0	0	1	0	0	0	0	0	73,75
i	0	0	0	1	2	0	0	0	67	4	1	3	1	1	0	0	0	0	0	0	0	0	0	0	0	0	83,75
j	0	0	0	1	0	1	3	0	3	65	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	81,25
k	0	1	1	1	1	0	0	10	2	1	53	4	1	1	0	0	0	3	0	0	1	0	0	0	0	0	66,25
l	0	0	0	2	2	0	0	4	1	1	58	0	0	0	2	0	1	1	1	1	0	0	0	0	0	0	72,50
m	0	0	0	0	0	0	0	0	0	0	0	62	13	0	0	0	0	0	1	0	0	0	0	4	0	0	77,50
n	0	0	0	0	0	0	0	0	0	0	0	25	55	0	0	0	0	0	0	0	0	0	0	0	0	0	68,75
o	3	0	0	0	1	1	3	0	0	0	0	0	0	0	69	0	0	0	1	0	1	1	0	0	0	0	86,25
p	0	0	0	0	0	2	0	0	0	0	0	6	1	0	0	69	0	2	0	0	0	0	0	0	0	0	86,25
q	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	76	0	0	0	0	0	0	0	1	1	0	95,00
r	0	0	0	0	3	0	0	1	3	0	0	1	1	7	0	0	0	59	1	0	1	1	0	1	0	1	73,75
s	0	0	0	0	0	0	0	0	0	0	0	1	0	7	0	0	0	25	44	0	1	1	0	1	0	0	55,00
t	0	0	0	0	2	1	0	0	5	1	0	0	0	0	0	0	2	0	65	0	0	0	0	1	0	3	81,25
u	0	0	0	2	0	0	0	0	0	0	1	0	1	2	1	0	0	0	0	0	72	0	1	0	0	0	90,00
v	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	70	7	0	0	0	87,50
w	0	0	0	0	0	0	0	1	0	0	0	8	0	0	0	0	0	0	0	0	6	6	59	0	0	0	73,75
x	0	0	0	0	0	0	0	0	1	0	4	0	4	3	0	0	0	2	0	0	0	0	0	63	2	1	78,75
y	0	1	0	1	0	1	3	0	3	3	0	1	1	0	1	0	1	0	0	0	0	1	0	2	59	2	73,75
z	3	0	0	0	0	1	3	0	0	3	0	0	0	0	0	1	0	0	0	0	0	0	0	4	2	63	78,75
<b>Taxa média de Reconhecimento</b>																									<b>79,62</b>		

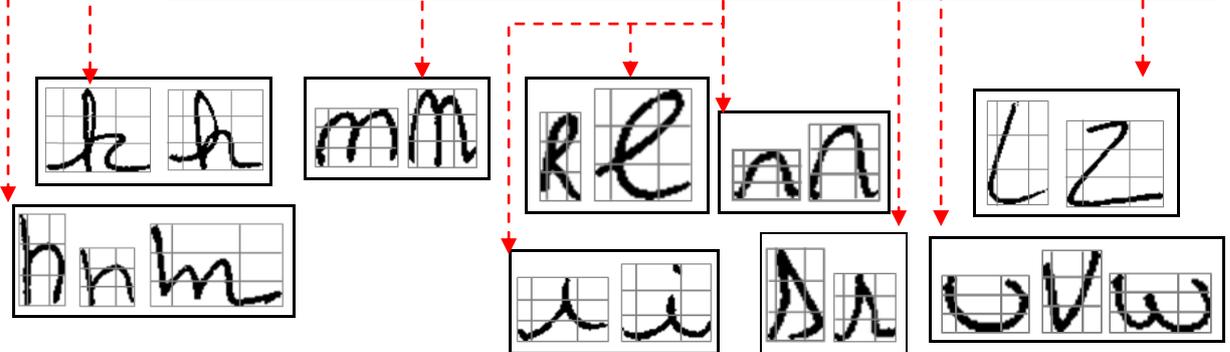


Figura 4.2. Exemplo de caracteres que se confundiram

A Figura 4.2 ilustra exemplos de caracteres distintos que se confundiram com a aplicação do mecanismo de zoneamento não simétrico  $Z = 9$ . Ao se analisar a Figura 4.2 pode-se observar que as partes que compõem o mecanismo de zoneamento  $Z = 9$  confundiu ainda mais as partes similares dos caracteres “l” e “z”, no caso, a parte esquerda e direita dos caracteres.

O mecanismo de zoneamento  $Z = 9$  também não ajudou a identificar as partes distintas dos caracteres, gerando mais confusões entre classes distintas, como se pode observar nas partes esquerdas dos caracteres “*r*” e “*s*”, como mostrado na Figura 4.2. Observou-se também que o zoneamento  $Z=9$  ressaltou ainda mais as similaridades de alguns caracteres, como se pode observar nos caracteres “*h*”, “*n*” e “*m*”, ilustrados também na Figura 4.2.

Outra observação, esta relacionada com a variabilidade de estilos, por exemplo, de caracteres minúsculos “*r*” e “*n*”, como mostrado na Figura 4.3. Percebeu-se então, que estes caracteres apresentaram taxas de reconhecimento baixas, visto que esta variabilidade não pode ser absorvida pelas técnicas implementadas.

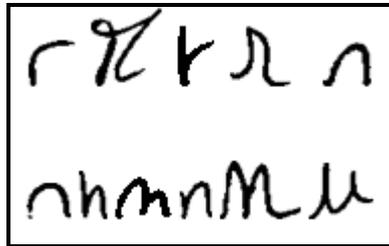


Figura 4.3. Exemplo de diversidade de estilos de escrita para caracteres “*r*” e “*n*”

O mecanismo de zoneamento simétrico  $Z = 4$  representou um aumento significativo nas taxas médias de reconhecimento comparando-se os experimentos realizados com as RNA MLP Convencional.

Apesar das várias confusões identificadas entre as classes, o mecanismo de zoneamento simétrico  $Z = 4$  teve desempenho melhor, em termos de taxa de reconhecimento, se comparado com o zoneamento não simétrico  $Z = 9$ , e uma análise mais aprofundada das confusões baseadas nas partes similares dos caracteres foi realizada.

A Tabela 4.6 mostra as confusões identificadas entre as classes, destacando alguns resultados. As confusões ressaltadas foram ilustradas, conforme pode se observar na Figura 4.4. Assim, encontram-se listadas a seguir as confusões identificadas entre as classes de caracteres para o experimento com o mecanismo de zoneamento simétrico  $Z = 4$ :

- “*a*” e “*o*”;
- “*b*” e “*s*”;
- “*e*” e “*l*”;
- “*f*” e “*l*”;

- “g”, “f” e “s”;
- “f”, “l” e “t”;
- “h”, “k” e “n”;
- “i” e “j”;
- “k”, “h” e “r”;
- “l”, “f”;
- “j” e “k”;
- “m” e “n”;
- “n”, “m”, “r” e “u”;
- “p”, “f”, “i” e “l”;
- “r”, “n”;
- “w”, “u”, e “v”;
- “x” e “k”;
- “y” e “x”;
- “z” e “g”

Analisando a matriz de confusão, encontram-se similaridades entre classes de caracteres distintos. Os caracteres “h” e “k” e “l” apresentam muita similaridade (Figura 4.4(a)), sendo que a diferença está na parte inferior dos caracteres. A parte superior de “k” é muito similar ao caractere “l”, sendo que as duas partes superiores no mecanismo de zoneamento  $Z = 4$  se confundem. As duas partes direitas dos caracteres “p” e “f” são bastante similares (Figura 4.4(b)), gerando confusão, sendo que a diferença entre os caracteres está na metade superior esquerda dos mesmos.

O caractere “r” teve dois estilos de escrita predominantes, gerando confusão com o caractere “n” similaridade (Figura 4.4(e)), sendo que uma sutil diferença está na região esquerda superior do caractere. Nota-se que a região esquerda do caractere “r” é também similar a do caractere “h” e, portanto, estes podem se confundir, principalmente quando se trata da região inferior esquerda, visto que a diferença está na parte central do caractere (onde as linhas se cruzam) e também na parte inferior direita.

Os caracteres “u”, “v” e “w”, devido as similaridades muito aparente, como se pode perceber na Figura 4.4(c) também se confundiram. Os caracteres “g” e “y” e “z” apresentam muitas similaridades (Figura 4.4(d)), e o zoneamento  $Z = 4$  não foi capaz de ressaltar as partes

distintas destes caracteres. Outros caracteres que a primeira vista parecem ser totalmente diferentes mas apresentam similaridades, é o caso dos caracteres “b” e “s” que apresentam similaridade na parte inferior.

Tabela 4.6. Matriz de Confusão Z=4 MLP Classe Modular

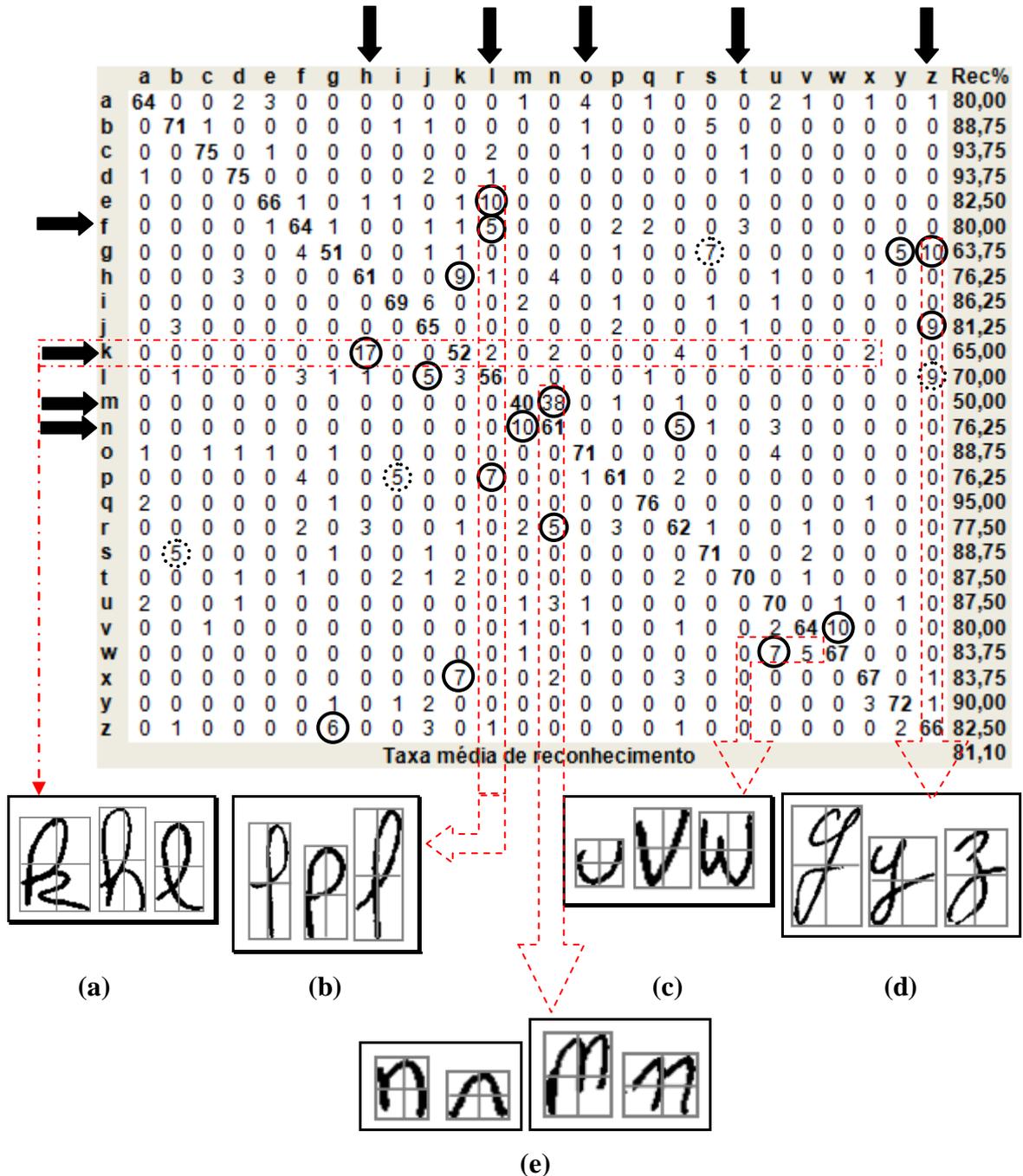


Figura 4.4. Similaridade entre caracteres distintos – Z4

Já para os caracteres manuscritos minúsculos “c”, “d” e “q” o zoneamento simétrico  $Z = 4$  obteve um excelente reconhecimento, chegando a altas taxas de reconhecimento (acima de 93%). Os caracteres em questão não apresentaram diversidades de estilos de escrita ou de forma contrastantes nas bases experimentais. Os caracteres manuscritos minúsculos “c”, “d” e “q” estão exemplificados na Figura 4.5.

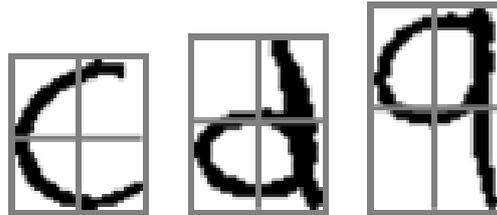


Figura 4.5. Amostras de caracteres “c”, “d” e “q”

Foram feitos em seguida experimentos com o mecanismo de zoneamento não simétrico  $Z = 5H$  procurando extrair características baseadas em caracteres que não são horizontalmente simétricos.

As seguintes confusões significativas foram identificadas entre as classes:

- “f”, “g”, “l” e “p”;
- “g” e “f”, “q”, “y” e “z”;
- “h”, “k”, “n”;
- “i” e “j”;
- “l”, “b” e “i”;
- “m” e “n”;
- “r”, “b”, “c” e “h”;
- “s”, “t” e “r”;
- “v”, “w”;
- “x”, “k” e “r”;
- “y” e “z”;
- “z”, “g”, “x” e “y”

A matriz de confusão, representada pela Tabela 4.7, mostra as confusões entre as classes, ressaltando alguns valores. Analisando a matriz de confusão representada pela Tabela 4.7, verifica-se que os caracteres “*h*” e “*r*” apresentaram muitas partes similares e se confundiram também na parte inferior central. A Figura 4.6 (a) ilustra esses caracteres.

Nos caracteres “*g*” e “*f*”, ilustrados na Figura 4.6 (b), a confusão aconteceu também na parte inferior, devido ao estilo de escrita do caractere “*f*”, nesse caso o mecanismo de zoneamento não simétrico  $Z = 5H$  não foi capaz de diferenciar a parte central inferior, ou seja, a parte do zoneamento inferior central não ajudou na distinção entre os dois caracteres.

Os caracteres “*p*” e “*f*” mostrados na Figura 4.6 (b) também se confundiram, pela similaridade em muitas partes dos caracteres, principalmente nas duas partes superiores dos caracteres, nesse caso ressaltar a parte inferior central dos caracteres também não ajudou na distinção entre esses dois caracteres.

Os caracteres “*i*”, “*j*” e “*l*” mostraram bastante similaridade relacionados à diversidade do estilo de escrita. Nesse caso a diferença visual aparente é muito pequena. A Figura 4.6 (c) mostra exemplos desses caracteres.

Outros caracteres como “*x*”, “*k*”, “*h*” e “*k*”, também apresentaram confusão relacionados aos estilos de escritas. Os caracteres “*x*”, “*k*”, “*h*” e “*k*” estão ilustrados na Figura 4.6 (d).

No caso dos caracteres “*u*”, “*v*” e “*w*”, o zoneamento não simétrico  $Z = 5H$  também não foi muito eficiente, causando confusão no reconhecimento, como se pode observar na Figura 4.6 (e). Observa-se ainda, que os caracteres manuscritos minúsculos “*u*” e “*v*” e “*w*” são muito similares visualmente, pois apresentam similaridade de traçado (forma).

O zoneamento não simétrico  $Z = 5H$  obteve melhor resultado, ou seja maior taxa de reconhecimento para as classes: “*d*”, “*q*” e “*y*”, que usam informações contidas na parte inferior do caractere para ser diferenciado de outros caracteres, pois estas classes contém caracteres que não são simétricos horizontalmente. O zoneamento não simétrico  $Z = 5H$  obteve alta taxa de reconhecimento também para a classe do caractere “*c*”, pois esta classe não apresentou similaridade de estilo de escrita com outras classes. A Figura 4.7 ilustra algumas amostras de caracteres “*d*”, “*q*” e “*y*”.

Tabela 4.7. Matriz de Confusão Z=5H MLP Classe Modular

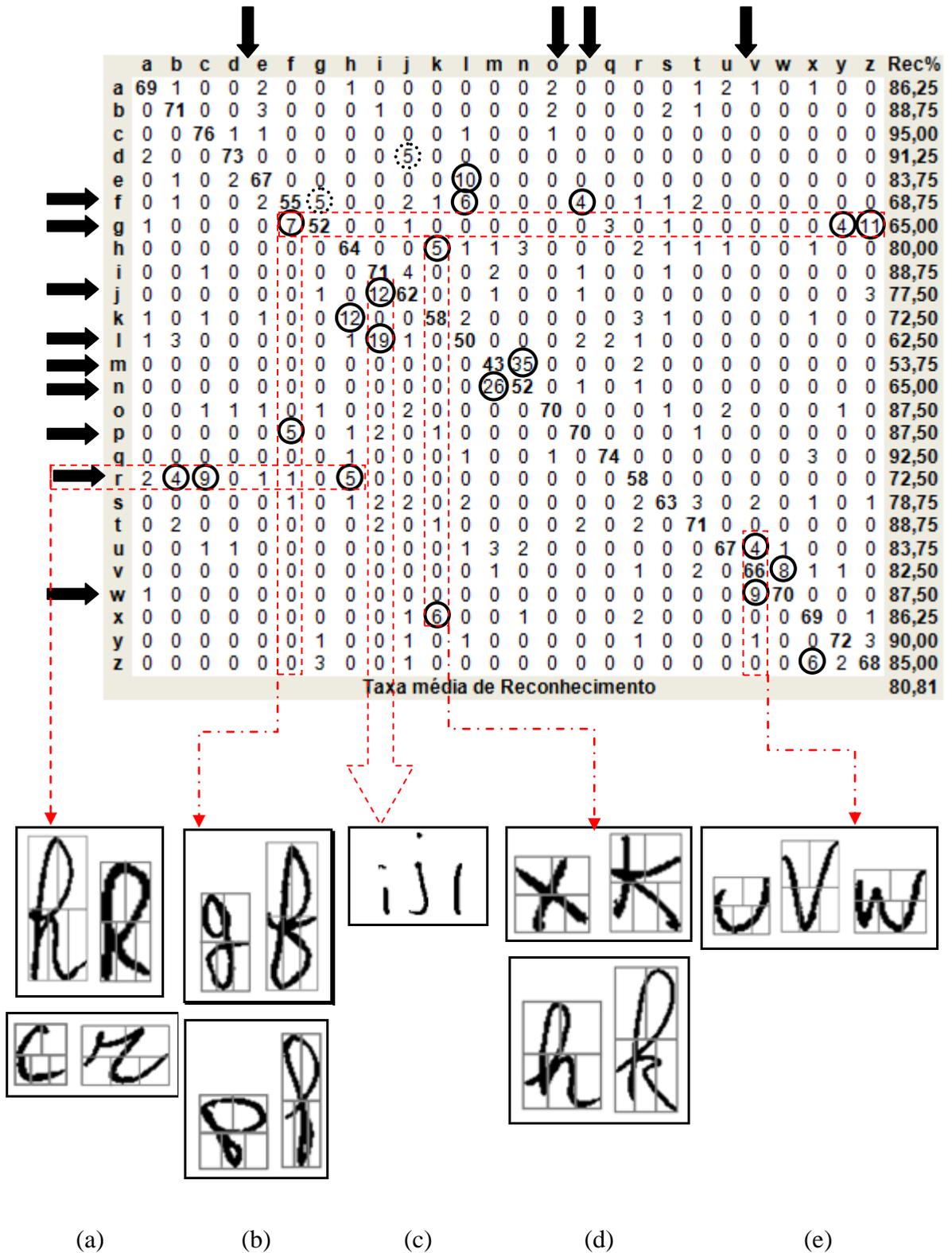


Figura 4.6. Confusão entre caracteres distintos –Z5H

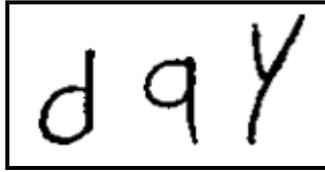


Figura 4.7. Exemplo de caracteres sem simetria horizontal

Para caracteres sem simetria vertical, o zoneamento  $Z = 5V$  foi o mais indicado, sendo que algumas confusões significativas podem ser citadas:

- “d” e “j”;
- “e” e “l”;
- “f”, “g” e “t”;
- “g”, “q” e “y”;
- “h”, “c” e “k”;
- “i” e “j”;
- “k”, “h” e “l”
- “l”, “e”, “i” e “z”;
- “m”, “n”;
- “n”, “u”;
- “r”, “c”, “h”;
- “u”, “v” e “w”;
- “x”, “t”;

O zoneamento  $Z = 5V$  obteve bom desempenho para os caracteres: “b”, “c”, “o”, “p”, “q”, “w” e “z”. A Tabela 4.8, mostra a relação entre as classes de caracteres e as confusões encontradas, sendo ressaltadas algumas confusões. A Figura 4.8(a) ilustra a similaridade entre caracteres: “t”, “k” e “x”; “h” e “k”; “f” e “t”. Na Figura 4.8(b) a parte direita central do zoneamento não ajudou na distinção entre os caracteres “c” e “h”. Já para os caracteres “z”, “g” e “y” o zoneamento  $Z = 5V$  evidenciou ainda mais as partes similares dos caracteres. O zoneamento  $Z = 5V$  não foi capaz de diferenciar esses caracteres e ressaltou ainda mais a intersecção dos traços que se encontram na parte central direita do caractere. A Figura 4.8(d) mostra alguns exemplos de caracteres “z”, “g” e “y” que se confundiram.

Tabela 4.8. Matriz de Confusão 5V MLP Classe Modular

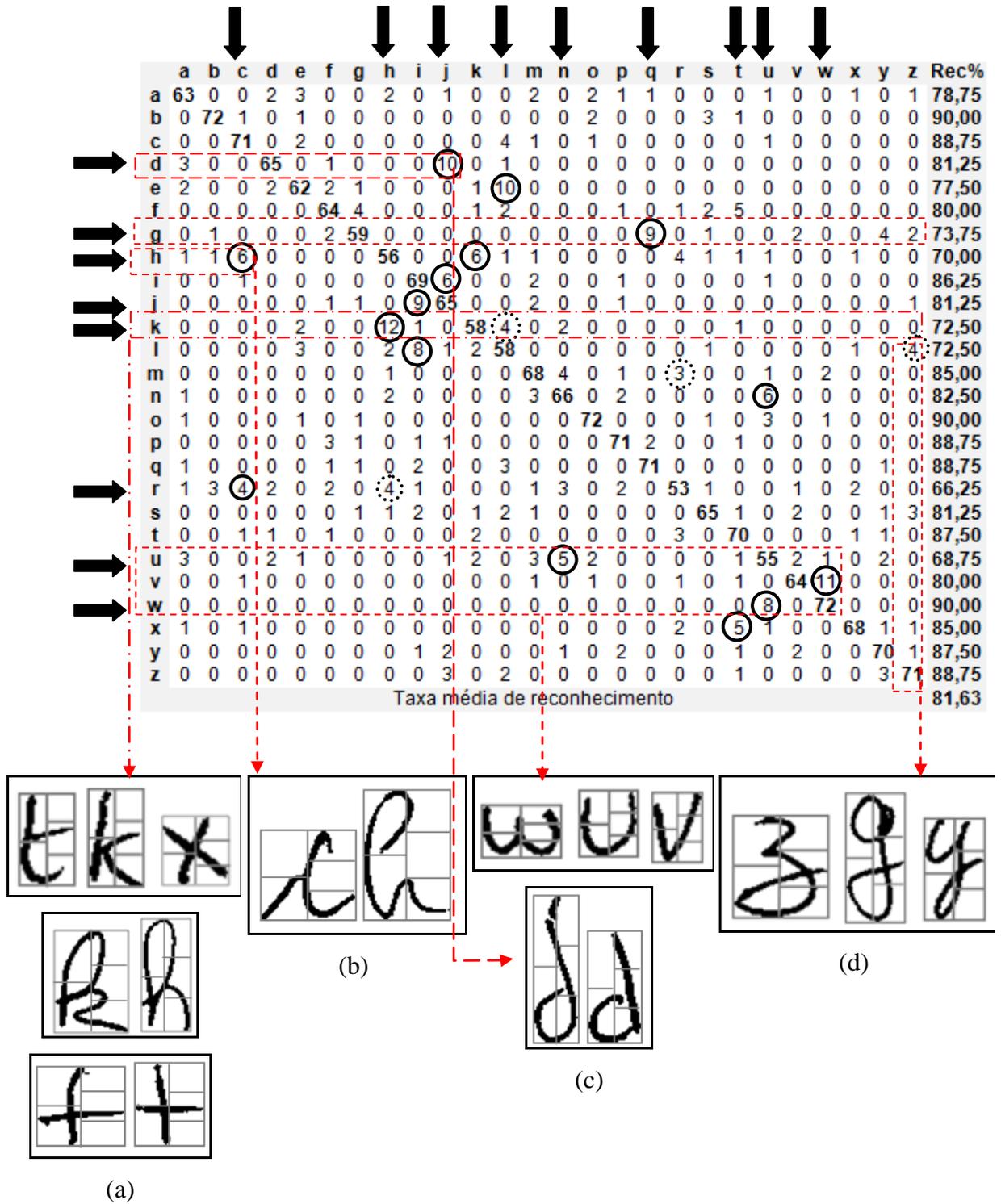


Figura 4.8. Similaridade e confusão entre caracteres distintos - Z5V

O mecanismo zoneamento não simétrico  $Z = 7$  apresentou bom desempenho para a maioria dos caracteres. O objetivo do zoneamento  $Z = 7$  é ressaltar a parte central dos caracteres, sendo que as confusões encontradas entre classes distintas são apresentadas na matriz de confusão representada na Tabela 4.9. Houve confusões significativas para as classes, podendo-se citar:

- “b” e “s”;
- “e” e “l”;
- “g” e “f”;
- “h” e “k”;
- “i” e “j”;
- “l” e “i”;
- “m” e “n”;
- “p”, “f” e “l”;
- “r” e “h”
- “u”, “v” e “w”;
- “x” e “k”;
- “y” e “z”;
- “z” e “x”;

Como mencionado anteriormente, o zoneamento  $Z = 7$  é baseado na representação da parte central do caractere, e ao se analisar as confusões, verifica-se que diversidades de estilo foram um dos principais fatores que ocasionou problemas no reconhecimento, com pode ser visto na Figura 4.9(a) para os caracteres que representam “f”, “p” e “l”. Já para os caracteres “h”, “k” e “l” (Figura 4.9(a) - superior) a confusão ocorreu devido as similaridades dos caracteres. Verificando-se as confusões entre os caracteres e analisando as similaridades e diferenças existentes, verificou-se que, de um modo geral houve melhora nas taxas de reconhecimento para caracteres não simétricos, onde a diferença está localizada na parte central do caractere. Já para os caracteres que apresentam similaridades na parte central dos caracteres, como “g”, “j”, “z” e “y”, estes foram confundidos, como ilustrado na Figura 4.9(c). Os melhores resultados foram obtidos para as classes: “c”, “d” e “q”.

Tabela 4.9. Matriz de Confusão Z=7 MLP Classe Modular

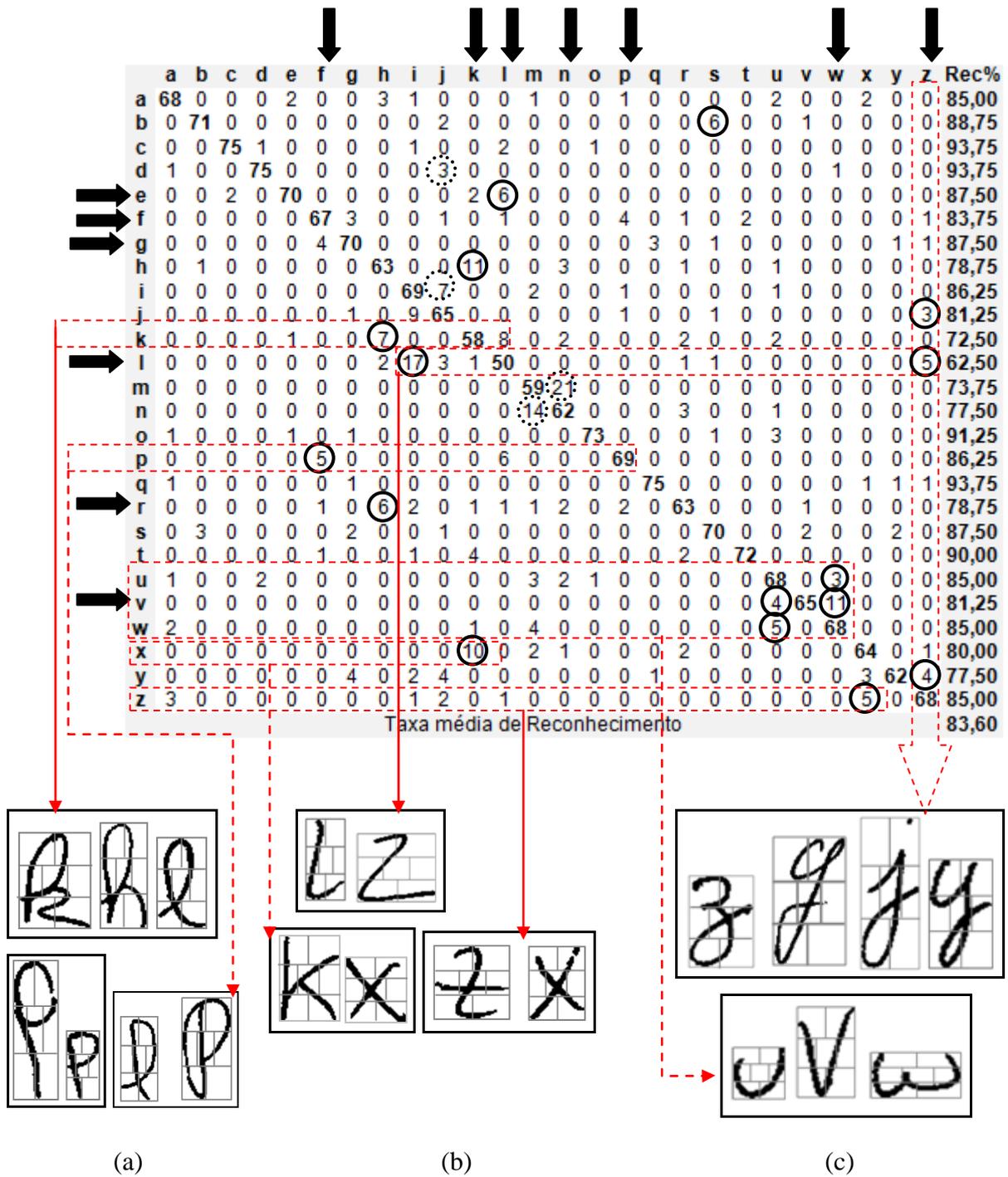


Figura 4.9. Confusão entre caracteres – Z7

Fazendo uma comparação entre os zoneamentos investigados percebe-se que para caracteres simétricos, tais como “o” e “c” obteve-se taxas de reconhecimentos altas em todos os mecanismos de zoneamento investigados. Caracteres sem diversidade de estilos

contrastantes, como “*q*”, também obteve altas taxas em todos os mecanismos analisados. Percebeu-se também que os caracteres “*n*” e “*m*” apresentaram muita confusão em todos os mecanismos analisados.

#### 4.5.3. Caracteres Minúsculos e Caracteres Maiúsculos

Um dos focos deste trabalho é investigar e comparar se os zoneamentos perceptivos e não simétricos que já foram avaliados para caracteres maiúsculos são válidos também para caracteres minúsculos, devido às diferenças de forma existentes entre eles. Deste modo foi realizada uma análise comparativa entre reconhecimento de caracteres manuscritos maiúsculos e minúsculos.

Para que a análise comparativa pudesse ser feita foram também realizados experimentos com RNA MLP Classe-Modular treinadas com caracteres manuscritos maiúsculos e minúsculos juntos, levando em consideração que em algumas situações não é necessário a distinção entre caracteres maiúsculos e minúsculos. Assim nesses experimentos as classes de caracteres maiúsculos e minúsculas foram reunidas em uma única classe, considerando por exemplo que um caractere maiúsculo “*A*” e um caractere minúsculo “*a*”, é um caractere da classe *A-a*. Em todos os experimentos foram utilizados a base de dados IRONOFF.

A Tabela 4.10 descreve as quantidades de amostras usadas para cada classe de caractere nos experimentos com os caracteres maiúsculos, minúsculos e para os experimentos realizados com caracteres maiúsculos e minúsculos.

Tabela 4.10. Quantidade de Amostras Utilizadas nos Experimentos

<b>Caracteres</b>	<b>Treinamento</b>	<b>Validação</b>	<b>Teste</b>
<b>maiúsculos</b>	(200 + 8 de cada classe) 400	(67 + 3 de cada classe) 142	(67 de cada classe)
<b>minúsculos</b>	(240 + 9 de cada classe) 465	(80 + 3 de cada classe) 155	(80 de cada classe)
<b>maiúsculos</b> + <b>minúsculos</b>	(200 + 8 de cada classe) + (200 + 8 de cada classe) 800	(67 + 3 de cada classe) + (67 + 3 de cada classe) 284	(67 de cada classe) + (67 de cada classe) 134

A Tabela 4.11 mostra as taxas de reconhecimento obtidas por caracteres manuscritos maiúsculos nos experimentos realizados por [FRE07], as taxas de reconhecimento dos experimentos com caracteres manuscritos minúsculos e as taxas médias de reconhecimento para caracteres maiúsculos e minúsculos treinadas em conjunto.

Tabela 4.11. Reconhecimento para caracteres maiúsculos e minúsculos

Caracteres Maiúsculos	Mecanismos de Zoneamento				Caracteres Minúsculos	Mecanismos de Zoneamento				Caracteres Maiúsculos e Minúsculos	Mecanismos de Zoneamento			
	Z4	Z5H	Z5V	Z7		Z4	Z5H	Z5V	Z7		Z4	Z5H	Z5V	Z7
A	<b>92.50</b>	86.60	89.60	91.00	a	80,00	<b>86,25</b>	78,75	85,00	A-a	76,12	82,84	81,34	<b>88,81</b>
B	65.70	64.20	74.60	<b>79.10</b>	b	88,75	88,75	<b>90,00</b>	88,75	B-b	<b>83,58</b>	74,63	82,09	84,33
C	82.10	79.10	68.70	<b>88.10</b>	c	93,75	<b>95,00</b>	88,75	93,75	C-c	<b>90,30</b>	79,85	83,58	88,81
D	73.10	65.70	68.70	<b>82.10</b>	d	<b>93,75</b>	91,25	81,25	<b>93,75</b>	D-d	80,60	76,12	67,16	<b>82,09</b>
E	83.60	85.10	89.60	<b>95.50</b>	e	82,50	83,75	77,50	<b>87,50</b>	E-e	35,82	36,57	30,60	<b>58,96</b>
F	<b>92.50</b>	91.00	89.60	<b>92.50</b>	f	80,00	68,75	80,00	<b>83,75</b>	F-f	83,58	83,58	<b>86,57</b>	<b>86,57</b>
G	82.10	<b>86.60</b>	80.60	80.60	g	63,75	65,00	73,75	<b>87,50</b>	G-g	76,12	76,12	65,67	<b>83,58</b>
H	<b>88.10</b>	85.10	70.10	76.10	h	76,25	<b>80,00</b>	70,00	78,75	H-h	69,40	77,61	<b>84,33</b>	73,13
I	<b>76.10</b>	71.60	<b>76.10</b>	71.60	i	86,25	<b>88,75</b>	86,25	86,25	I-i	79,85	<b>88,81</b>	86,57	73,88
J	<b>83.60</b>	79.10	79.10	82.10	j	<b>81,25</b>	77,50	<b>81,25</b>	<b>81,25</b>	J-j	76,87	86,57	<b>87,31</b>	79,85
K	77.60	76.10	77.60	<b>80.60</b>	k	65,00	<b>72,50</b>	<b>72,50</b>	<b>72,50</b>	K-k	68,66	67,91	70,90	<b>74,63</b>
L	<b>92.50</b>	89.60	86.60	91.00	l	70,00	62,50	<b>72,50</b>	62,50	L-l	67,16	74,63	<b>82,84</b>	79,85
M	<b>92.50</b>	82.10	85.10	88.10	m	50,00	53,75	<b>85,00</b>	73,75	M-m	58,96	69,40	76,87	<b>79,85</b>
N	68.70	77.60	70.10	<b>86.60</b>	n	76,25	65,00	<b>82,50</b>	77,50	N-n	<b>81,34</b>	67,16	70,15	68,66
O	86.60	<b>89.60</b>	88.10	83.60	o	88,75	87,50	90,00	<b>91,25</b>	O-o	85,07	88,06	88,81	<b>89,55</b>
P	86.60	92.50	91.00	<b>94.00</b>	p	76,25	87,50	<b>88,75</b>	86,25	P-p	88,06	<b>88,81</b>	88,06	83,58
Q	<b>82.10</b>	64.20	76.10	80.60	q	<b>95,00</b>	92,50	88,75	93,75	Q-q	88,06	<b>89,55</b>	88,06	87,31
R	86.60	89.60	88.10	<b>91.00</b>	r	77,50	72,50	66,25	<b>78,75</b>	R-r	78,36	82,84	77,61	<b>86,57</b>
S	<b>79.10</b>	<b>79.10</b>	<b>79.10</b>	76.10	s	<b>88,75</b>	78,75	81,25	87,50	S-s	81,34	88,06	<b>88,81</b>	<b>88,81</b>
T	95.50	<b>97.00</b>	<b>97.00</b>	<b>97.00</b>	t	87,50	88,75	87,50	<b>90,00</b>	T-t	<b>90,30</b>	82,09	88,81	88,06
U	80.60	85.10	82.10	<b>86.60</b>	u	<b>87,50</b>	83,75	68,75	85,00	U-u	83,58	81,34	83,58	<b>87,31</b>
V	<b>95.50</b>	82.10	88.10	82.10	v	80,00	<b>82,50</b>	80,00	81,25	V-v	77,61	83,58	86,57	<b>87,31</b>
W	70.10	74.60	65.70	<b>79.10</b>	w	83,75	87,50	<b>90,00</b>	85,00	W-w	73,13	<b>83,58</b>	74,63	76,12
X	76.10	74.60	70.10	<b>79.10</b>	x	83,75	<b>86,25</b>	85,00	80,00	X-x	85,82	85,07	86,57	<b>89,55</b>
Y	77.60	<b>89.60</b>	85.10	82.10	y	<b>90,00</b>	<b>90,00</b>	87,50	77,50	Y-y	91,04	90,30	85,82	<b>91,79</b>
Z	<b>89.60</b>	88.10	88.10	86.60	z	82,50	85,00	<b>88,75</b>	85,00	Z-z	69,40	76,87	<b>79,10</b>	<b>79,10</b>
Média	<b>83.00</b>	<b>81.70</b>	<b>80.90</b>	<b>84.70</b>	Média	<b>81,10</b>	<b>80,82</b>	<b>81,63</b>	<b>83,61</b>	Média	<b>77,70</b>	<b>79,31</b>	<b>79,71</b>	<b>82,23</b>

Tendo por base a percepção visual e, portanto, que caracteres maiúsculos e minúsculos são diferentes na forma, tamanho, traços, concavidades e convexidades, observa-se que o tratamento separado de caracteres minúsculos e maiúsculos é recomendável. Isso enfatiza a necessidade de incorporar a percepção humana para contribuir no processo de reconhecimento de caracteres. É necessário capturar evidências da percepção humana das formas para contribuir com o desenvolvimento de sistemas automático de reconhecimento.

Os caracteres minúsculos apresentam muita assimetria e as diversidades de estilos são mais marcantes que em caracteres maiúsculos. A partir dos experimentos com as RNAs MLP com arquitetura Classe-Modular para caracteres maiúsculos e minúsculos juntos pode-se verificar que caracteres minúsculos em geral apresentaram mais confusão não só entre caracteres minúsculos mas também com caracteres maiúsculos.

A Figura 4.10 ilustra exemplos de amostras de caracteres manuscritos minúsculos que se confundiram com caracteres maiúsculos.

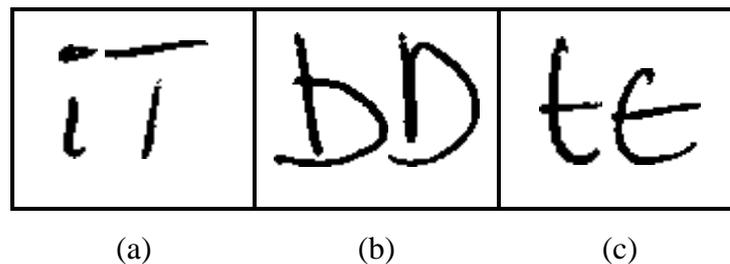


Figura 4.10: Confusão entre caracteres: a) “i” minúsculo e “T” maiúsculo; b) caracteres “b” minúsculo e “D” maiúsculo; c) caracteres “t” minúsculo e “E” maiúsculo

Já a Figura 4.11 mostra exemplos de escrita para os caracteres maiúsculo “R” e minúsculo “r”. Pode-se observar nesses caracteres que a diversidade de estilos de escrita para o caractere minúsculo “r” é visualmente mais evidente que para o caractere maiúsculo “R”. Ainda na Figura 4.11, pode-se observar exemplos de simetria de caractere maiúsculo “H” e do caractere minúsculo “h”. Em estudos da percepção humana das formas, simetria é considerada como lados opostos, que, embora não sejam exatamente iguais, mantém uma forte semelhança.

A Figura 4.12 apresenta um exemplo de extração de características aplicada aos caracteres maiúsculo “H” e minúsculo “h”, destacando as diferenças entre caracteres



Analisando os resultados experimentais, observa-se que os mecanismos de zoneamento  $Z = 4$  e  $Z = 7$  obtiveram melhores taxas de reconhecimento para caracteres maiúsculos. Já para caracteres minúsculos não há como estabelecer um melhor mecanismo de zoneamento. A Tabela 4.12 mostra a relação entre o melhor mecanismo de zoneamento para cada classe de caractere, sendo possível verificar que não existe predominância de reconhecimento por mecanismo de zoneamento para os caracteres manuscritos minúsculos.

Quando compara-se os resultados aplicando o mecanismo de zoneamento  $Z = 7$  para os experimentos com caracteres maiúsculos e minúsculos junto, os caracteres "*D - d*", "*E - e*", "*F - f*", "*K - k*", "*R - r*", alcançam as melhores taxas de reconhecimento, sejam estes maiúsculos ou minúsculos. Observa-se que os caracteres minúsculos, por apresentarem mais assimetria que caracteres maiúsculos, obtiveram taxas mais altas de reconhecimento para a maioria dos caracteres com os mecanismos de zoneamento não simétricos: 5H, 5V e 7.

Tabela 4.12. Melhor Zoneamento para cada Classe de Caractere

<b>Mecanismos de Zoneamento</b>	<b>Classes Maiúsculas</b>	<b>Classes Minúsculas</b>
<b>Z = 4</b>	A, F, H, I, J, L, M, Q, S, V, Z	d, j, q, s, u, y
<b>Z = 5H</b>	G, O, Y	a, c, h, i, k, v, x
<b>Z = 5V</b>	I, S, T	b, l, m, n, p, w, z
<b>Z = 7</b>	B, C, D, E, K, N, P, R, U, W, X	e, f, g, o, r, t

Algumas confusões significativas foram observadas entre algumas classes de caracteres nos experimentos com caracteres maiúsculos e minúsculos juntos. A Tabela 4.13 mostra a relação entre os mecanismos de zoneamentos e classes de caracteres que foram confundidos.

Observa-se ainda nos experimentos com os caracteres maiúsculos e minúsculos treinados juntos, que a classe "*E-e*" obteve a taxa de reconhecimento mais baixa entre todos os mecanismos de zoneamento investigados.

As confusões encontradas nos experimentos com caracteres manuscritos maiúsculos e minúsculos juntos utilizando os mecanismos de zoneamento  $Z4$ ,  $Z5H$ ,  $Z5V$  e  $Z7$  são ilustrados nas matrizes de confusão representadas nas Tabelas 4.14, 4.15, 4.16 e 4.17, respectivamente.

Tabela 4.13. Mecanismos de Zoneamento e Confusões entre Classes

Z = 4	Z = 5H	Z = 5V	Z = 7
A-a: D-d, H-h e Q-q	B-b: D-d	B-b: D-d	B-b: D-d
B-b: D-d	C-c: I-i e L-l	D-d: A-a e B-b	D-d: A-a e B-b
E -e: I-i, K-k, R-r, Z-z	D-d: A-a, B-b e O-o	E-e: R-r e T-t	E-e: T-t
J-j: I-i	E-e: I-i, J-j e K-k	G-g: B-b	H-h: M-m
K-k: H-h, R-r e X-x	G-g: B-b	K-k: L-l e X-x	I-i: J-j
L-l: I-i e T-t	H-h: M-m	M-m: H-h e N-n	K-k: H-h
M-m: N-n	M-m: N-n	R-r: H-h	M-m: H-h
S-s: J-j	U-u: W-w	W-w: N-n, U- u, V-v	N-n: U-u
V-v: W-w	Z-z: J-j	Z-z: X-x	W-w: U-u, V-v
Z-z: J-j			

Tabela 4.14 – Matriz de Confusão Z4 – MLP Classe-Modular : Maiúsculas e Minúsculas

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	%Rec
a	102	2	0	5	0	0	1	7	0	0	0	0	3	1	1	0	6	4	0	0	1	0	0	1	0	0	76,12
b	0	112	0	9	0	0	5	0	0	0	0	0	0	0	2	0	0	1	2	0	0	0	0	1	0	2	83,58
c	0	0	121	0	2	0	0	0	3	0	0	3	0	0	2	0	1	0	0	1	1	0	0	0	0	0	90,30
d	0	11	0	108	0	0	1	0	0	5	0	0	0	0	3	0	1	0	2	0	1	0	0	0	0	2	80,60
e	0	0	0	0	48	4	2	5	16	2	7	0	0	0	1	0	9	3	3	0	4	1	2	0	2	35,82	
f	0	0	1	0	5	112	2	0	0	0	1	4	0	0	0	1	1	1	2	3	0	0	0	1	0	0	83,58
g	0	5	0	2	1	8	102	0	0	2	0	0	0	0	0	0	3	0	4	0	0	0	0	0	3	4	76,12
h	1	4	1	5	2	1	5	93	6	1	1	0	3	3	1	0	0	1	0	1	3	1	0	0	0	1	69,40
i	0	0	0	1	0	1	0	0	107	9	0	4	2	0	2	1	0	0	1	1	1	0	0	3	0	1	79,85
j	0	0	0	0	0	1	0	0	25	103	0	0	2	0	0	0	0	0	0	0	1	0	0	0	0	2	76,87
k	0	0	0	3	0	0	1	7	0	0	92	4	1	0	1	0	2	10	0	1	0	0	0	8	1	3	68,66
l	0	1	1	1	1	0	1	1	18	1	1	90	0	0	0	1	0	1	1	7	3	0	0	0	0	5	67,16
m	0	0	0	0	0	0	0	2	0	0	1	0	79	39	0	1	0	4	0	0	8	0	1	2	3	0	58,96
n	0	1	0	2	0	0	0	1	0	0	0	0	7	109	0	0	0	4	0	0	5	3	2	0	0	0	81,34
o	1	2	0	9	0	0	1	0	0	0	0	0	0	114	0	1	0	1	1	3	0	0	0	0	1	0	85,07
p	0	0	0	0	0	4	0	0	2	0	0	4	0	1	0	118	5	0	0	0	0	0	0	0	0	0	88,06
q	6	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	118	1	0	0	3	0	0	0	3	0	88,06
r	4	0	0	0	1	1	0	6	5	0	1	3	1	5	0	0	0	105	0	0	0	0	0	2	0	0	78,36
s	1	7	0	2	0	0	0	0	0	14	0	0	0	0	0	0	0	0	109	0	0	1	0	0	0	0	81,34
t	0	0	0	0	1	0	0	0	7	0	1	0	0	0	0	0	0	4	0	121	0	0	0	0	0	0	90,30
u	0	0	0	1	0	0	0	1	2	2	1	0	0	3	2	0	1	0	0	0	112	0	8	0	1	0	83,58
v	0	0	1	1	0	0	1	0	0	0	1	0	2	0	2	0	0	0	1	2	2	104	15	1	0	1	77,61
w	0	0	0	0	0	0	0	0	0	0	0	0	0	17	1	0	0	0	0	0	13	5	98	0	0	0	73,13
x	0	0	0	0	0	0	0	2	0	0	4	0	1	0	0	0	0	3	0	1	2	0	2	115	3	1	85,82
y	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	1	0	0	1	0	1	0	1	2	122	1	91,04
z	1	4	0	7	0	1	2	1	1	7	0	0	0	0	0	0	0	1	0	0	1	1	0	10	4	93	69,40
Taxa Média de Reconhecimento																									77,70		

Tabela 4.15 – Matriz de Confusão Z5H – MLP Classe-Modular : Maiúsculas e Minúsculas

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	%Rec.
a	111	0	0	4	0	1	1	3	2	0	1	0	4	1	2	0	2	0	1	0	0	0	0	1	0	0	82,84
b	0	100	1	14	2	0	4	0	0	1	1	1	0	1	2	0	0	0	3	0	1	1	0	1	0	1	74,63
c	0	1	107	1	2	1	0	0	7	2	2	4	1	1	1	0	0	0	1	2	0	0	0	0	0	1	79,85
d	9	9	0	102	0	0	0	0	2	1	0	1	0	0	7	0	0	0	2	0	0	0	0	0	0	0	76,12
e	0	1	1	6	49	2	3	8	13	11	7	0	5	3	5	2	0	1	3	0	5	3	2	1	0	3	36,57
f	1	0	0	0	5	112	1	0	3	0	1	4	1	0	0	0	1	0	1	2	0	0	0	1	0	1	83,58
g	0	8	0	3	0	3	102	0	0	3	0	0	1	1	2	0	2	0	3	0	0	1	0	1	3	1	76,12
h	0	1	1	1	1	1	1	104	0	0	5	1	10	2	0	0	0	1	1	1	2	0	0	1	0	1	77,61
i	0	0	0	1	1	0	0	0	119	7	2	2	0	0	2	0	0	0	0	0	0	0	0	0	0	0	88,81
j	0	0	0	0	0	0	1	0	7	116	1	0	0	0	0	0	0	2	0	0	0	0	0	1	5	1	86,57
k	1	5	0	4	1	3	1	2	2	1	91	4	1	1	2	0	0	3	0	4	3	0	0	4	0	1	67,91
l	0	0	1	2	5	3	1	2	5	0	1	100	0	0	1	3	0	2	2	1	2	0	0	1	0	2	74,63
m	0	1	0	4	0	0	0	2	0	0	0	0	93	32	0	1	0	1	0	0	0	0	0	0	0	0	69,40
n	0	1	0	0	0	0	0	0	0	0	0	0	33	90	1	4	0	2	0	0	0	2	1	0	0	0	67,16
o	1	1	1	6	0	0	2	0	0	0	0	0	0	118	0	1	0	1	0	1	2	0	0	0	1	0	88,06
p	0	0	0	1	0	2	0	0	3	0	0	4	0	1	0	119	1	2	0	0	0	0	0	0	0	1	88,81
q	3	0	0	0	0	0	3	0	0	0	0	0	0	0	5	0	120	0	1	0	1	0	0	1	0	0	89,55
r	1	0	0	0	4	0	0	7	1	0	1	3	0	3	0	0	0	111	0	0	1	0	0	1	0	1	82,84
s	0	1	1	2	0	0	1	0	0	5	0	1	0	0	2	1	0	0	118	0	0	1	0	0	1	0	88,06
t	0	0	0	0	4	0	1	0	6	2	0	2	2	0	1	0	0	2	1	110	0	1	0	1	0	1	82,09
u	4	0	0	5	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	109	2	8	0	0	0	81,34
v	0	0	1	0	1	0	0	0	3	0	0	2	0	3	3	0	0	0	0	2	2	112	4	1	0	0	83,58
w	1	0	0	2	0	0	0	0	3	0	2	0	0	1	4	0	0	0	1	1	3	3	112	1	0	0	83,58
x	0	0	0	1	1	0	0	3	2	0	4	1	0	2	0	0	0	0	1	0	0	0	0	114	2	3	85,07
y	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	1	1	0	0	2	3	1	121	2	90,30
z	0	2	0	5	0	1	1	1	0	8	1	0	0	0	2	0	0	0	0	0	1	0	1	6	2	103	76,87
Taxa Média de Reconhecimento																											79,31

Tabela 4.16 – Matriz de Confusão Z5V – MLP Classe-Modular : Maiúsculas e Minúsculas

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	%Rec.	
a	109	1	0	4	0	0	1	2	6	1	0	0	1	1	2	0	4	0	0	0	0	0	0	1	0	1	81,34	
b	0	110	0	11	1	0	4	0	0	0	1	0	0	0	1	0	0	0	6	0	0	0	0	0	0	0	82,09	
c	0	1	112	1	2	1	0	0	5	2	0	2	1	1	2	0	0	0	1	1	1	0	0	0	0	1	83,58	
d	19	7	0	90	1	0	0	0	3	1	0	4	0	1	3	0	1	0	1	0	1	0	0	0	0	2	67,16	
e	1	2	1	2	41	2	3	2	3	6	8	4	0	2	3	1	1	17	0	22	1	1	1	4	0	6	30,60	
f	0	0	0	0	1	116	0	1	5	1	2	2	0	0	0	1	1	2	0	0	0	0	0	0	0	2	86,57	
g	0	10	0	3	1	5	88	0	0	4	1	1	1	1	3	0	2	0	3	1	0	0	0	1	8	1	65,67	
h	0	0	0	0	1	1	0	113	1	0	4	2	7	2	0	0	0	1	0	1	1	0	0	0	0	0	84,33	
i	0	0	0	1	0	0	0	0	116	4	0	10	0	0	1	0	0	0	0	1	0	0	0	0	0	1	86,57	
j	0	0	0	0	0	1	0	0	7	117	2	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1	4	87,31
k	0	1	0	3	1	1	1	4	2	0	95	9	0	2	0	0	0	1	1	2	0	0	0	10	0	1	70,90	
l	0	0	1	2	1	2	0	0	9	0	0	111	0	0	0	2	0	2	1	1	1	0	0	0	0	1	82,84	
m	0	0	0	0	0	2	0	14	0	0	0	1	103	9	0	1	0	0	0	1	1	0	0	0	1	1	76,87	
n	0	0	0	0	0	0	0	3	0	0	0	1	21	94	3	0	0	5	0	3	2	2	0	0	0	0	70,15	
o	0	1	0	8	0	0	3	0	0	0	0	0	0	119	0	1	0	1	0	1	0	0	0	0	0	0	88,81	
p	0	0	0	3	0	2	0	0	2	1	0	6	0	0	1	118	0	0	0	1	0	0	0	0	0	0	88,06	
q	5	0	0	0	0	0	2	1	0	0	1	0	0	0	2	0	118	0	0	0	0	0	0	1	3	1	88,06	
r	0	0	0	0	1	1	0	10	0	0	2	5	2	5	0	2	0	104	1	0	0	0	0	0	0	1	77,61	
s	0	1	0	3	0	0	2	0	2	4	0	0	0	0	1	0	0	119	0	0	1	0	0	1	0	1	88,81	
t	0	1	0	0	4	0	0	0	4	0	0	1	0	0	0	0	2	1	119	0	0	0	1	0	1	0	88,81	
u	2	0	0	2	0	0	0	1	0	1	0	0	0	1	7	0	0	0	0	0	112	2	5	0	1	0	83,58	
v	0	0	0	0	1	1	0	0	5	0	0	0	0	0	5	0	0	0	0	2	0	116	4	0	0	0	86,57	
w	1	0	0	0	0	0	0	2	0	0	0	0	15	1	0	0	0	0	0	9	6	100	0	0	0	0	74,63	
x	0	0	0	1	1	0	0	2	1	0	4	0	1	2	0	0	0	2	1	1	0	0	0	116	1	1	86,57	
y	0	0	0	0	0	1	0	0	3	6	0	0	1	0	1	1	1	0	0	0	2	1	0	115	2	3	85,82	
z	2	1	0	4	0	1	1	1	0	7	0	0	0	0	0	0	0	0	0	0	0	0	8	3	106	79,10		
Taxa Média de Reconhecimento																											79,71	

Tabela 4.17 – Matriz de Confusão Z7 – MLP Classe-Modular : Maiúsculas e Minúsculas

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	%Rec
a	119	1	0	2	1	0	1	2	0	0	0	0	3	1	0	0	1	2	0	0	0	0	0	1	0	0	88,81
b	1	113	0	10	0	0	3	0	0	1	0	0	0	0	1	0	0	0	5	0	0	0	0	0	0	0	84,33
c	0	0	119	0	1	1	0	0	3	0	1	1	1	0	2	0	0	0	1	1	2	0	0	1	0	0	88,81
d	10	8	0	110	0	0	0	0	0	2	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	1	82,09
e	0	0	0	0	79	2	2	0	8	3	3	1	0	0	0	0	0	1	1	26	0	1	0	1	2	2	58,96
f	1	0	0	0	0	116	2	0	1	2	4	1	0	0	0	2	1	0	1	3	0	0	0	0	0	0	86,57
g	0	2	0	1	0	6	112	1	0	2	0	0	0	1	0	1	0	3	0	0	0	0	0	0	3	2	83,58
h	2	0	0	0	0	0	0	98	0	0	7	0	13	5	3	0	0	1	0	0	5	0	0	0	0	0	73,13
i	0	0	0	3	0	1	0	0	99	12	0	7	2	0	2	1	1	0	1	0	1	0	1	1	0	2	73,88
j	0	0	0	1	0	1	0	0	14	107	0	0	2	0	1	1	0	0	0	0	1	0	0	0	0	6	79,85
k	0	0	0	1	2	1	0	16	0	0	100	0	2	0	0	1	7	0	1	0	0	0	3	0	0	74,63	
l	0	0	1	0	1	2	0	0	17	0	1	107	0	0	0	2	0	0	0	0	0	0	0	0	0	3	79,85
m	0	0	0	4	0	0	0	10	0	0	0	0	107	9	1	0	0	0	0	0	2	0	0	0	1	0	79,85
n	0	2	0	0	0	0	0	1	0	0	0	0	7	92	0	0	0	0	0	0	21	5	5	0	1	0	68,66
o	1	0	0	4	0	0	3	0	0	0	0	0	0	0	120	1	0	1	0	3	0	0	0	1	0	0	89,55
p	0	0	0	0	0	6	0	0	2	0	0	6	0	0	0	112	3	3	2	0	0	0	0	0	0	0	83,58
q	6	0	0	0	1	0	2	0	4	0	0	1	0	0	0	0	117	0	0	0	1	0	0	1	1	0	87,31
r	0	0	0	0	3	1	0	3	2	0	2	2	0	3	0	2	0	116	0	0	0	0	0	0	0	0	86,57
s	0	1	0	2	0	0	0	0	7	0	1	0	1	0	0	0	0	0	119	1	1	0	0	0	0	1	88,81
t	0	0	0	0	2	1	0	0	1	0	4	0	0	0	2	1	0	3	0	118	0	1	0	1	0	1	88,06
u	0	0	0	2	0	0	0	1	0	1	3	0	1	2	0	0	0	0	0	0	117	0	6	0	0	1	87,31
v	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	6	117	9	0	0	0	87,31
w	0	0	0	0	0	0	0	0	0	0	1	0	0	5	0	0	0	0	0	0	17	9	102	0	0	0	76,12
x	0	0	0	0	0	0	0	0	0	5	0	1	0	0	0	0	6	0	0	1	0	0	120	0	1	89,55	
y	0	0	0	0	0	0	2	0	0	2	0	0	0	0	0	0	2	1	1	1	1	0	0	0	123	1	91,79
z	0	0	0	5	0	1	1	1	0	9	0	0	0	0	0	0	0	0	0	0	0	1	0	7	3	106	79,10
Taxa Média de Reconhecimento																									82,23		

Em seu trabalho Koerich (2003) [KOE03b] e Koerich et. al. (2005) [KOE05], também faz experimentos utilizando redes neurais para o reconhecimento de caracteres maiúsculos e minúsculos. A base de dados utilizada foi a NIST, assim sendo não se pode fazer uma comparação direta com seus resultados, pois as bases de dados são distintas. Os autores também consideram em um de seus experimentos, que em determinada situação, dependendo do contexto não há necessidade de distinção entre caracteres manuscritos maiúsculos e minúsculos. Assim, considerou o seguinte: um problema de classificação com 52 classes, no qual caracteres minúsculos e maiúsculos são considerados diferentes; Um problema de classificação com 26 + 26 classes: caracteres minúsculos e maiúsculos são considerados diferentes e treinados em redes separadas e um problema de classificação com 26 classes, no qual caracteres minúsculos e maiúsculos são unidos, formando meta-classes.

Britto et al (2004) [BRI04] também faz experimentos com caracteres manuscritos, utilizando HMM e uma combinação de primitivas globais e locais. O HMM cria um zoneamento implícito baseado na combinação de colunas e linhas. A base de dados utilizada para os experimentos foi a NIST. O trabalho realizado por Britto et al. (2004) [BRI04] foi descrito resumidamente no Capítulo 2.

A literatura ainda é repleta de pesquisas com caracteres manuscritos, sendo possível uma comparação entre os experimentos realizados no presente trabalho com importantes contribuições disponíveis na literatura. A Tabela 4.18 mostra alguns resultados obtidos por experimentos realizados por diferentes autores, para caracteres manuscritos maiúsculos e minúsculos, utilizando as base de dados NIST e IRONOFF.

Tabela 4.18. Trabalhos Relacionados ( base de dados: NIST e IRONOFF )

	<b>Base de dados</b>	<b>Classificador</b>	<b>% Rec</b>
Britto [BRI04] Maiúscula Minúscula Maiúscula + Minúscula	NIST	HMM	90,00 84,00 87,00
Koerich [KOE05] Maiúscula Minúscula Maiúscula + Minúscula	NIST	MLP Convencional	92,49 86,73 87,79
Aires [AIR05] Maiúscula	IRONOFF	MLP-Classe Modular	84,70
Método Proposto Minúscula Maiúscula + Minúscula	IRONOFF	MLP-Classe Modular	83,60 82,23
Seleção Dinâmica Minúscula	IRONOFF	MLP Convencional + MLP-Classe Modular	88,13

#### 4.5.4. Abordagem pela Seleção Dinâmica de Zoneamento

A partir dos experimentos para caracteres manuscritos minúsculos e considerando as comparações realizadas entre caracteres manuscritos maiúsculos e minúsculos e, ainda, observando as análises das matrizes de confusão para todos os mecanismos de zoneamentos investigados, pode-se concluir quais classes distintas tiveram melhor reconhecimento em função dos mecanismos de zoneamentos avaliados. Tal fato é ainda mais evidente para caracteres manuscritos minúsculos, como se pode observar na Tabela 4.12, na qual é visível que não há um zoneamento que seja predominante no reconhecimento das classes de caracteres minúsculos. Como exemplo, tem se a classe do caractere “a”, que obteve taxa de reconhecimento mais alta no zoneamento  $Z = 5H$ , por sua vez o caractere “h” obteve taxa de reconhecimento mais baixa neste zoneamento e taxa mais alta no mecanismo de zoneamento  $Z = 7$ .

Percebeu-se ainda por meio dos experimentos, que variações nos estilos de escrita dos caracteres, mesmo que aparentemente pequenas, podem causar confusão e, conseqüentemente, resultar em reconhecimento incorreto, dependendo do mecanismo de zoneamento que se aplica. Assim, a seleção *a priori* de qual mecanismo de zoneamento utilizar para um determinado caractere torna-se um fator crucial para auxiliar no aumento das taxas de reconhecimento. A Tabela 4.19 mostra as taxas média de reconhecimento obtidas para cada classe de caractere manuscrito minúsculo, considerando os experimentos com os mecanismos de zoneamento Z4, Z5H, Z5V e Z7. Esta tabela mostra também os resultados obtidos pela seleção dinâmica de zoneamento para cada classe de caractere.

Tabela 4.19. Tabela comparativa: Zoneamentos individuais e seleção dinâmica

Classe	Z4	Z5H	Z5V	Z7	Seleção Dinâmica
a	80,00	<b>86,25</b>	78,75	85,00	85,00
b	88,75	88,75	90,00	88,75	<b>91,25</b>
c	93,75	95,00	88,75	93,75	<b>96,25</b>
d	<b>93,75</b>	91,25	81,25	<b>93,75</b>	<b>93,75</b>
e	82,50	83,75	77,50	<b>87,50</b>	<b>87,50</b>
f	80,00	68,75	80,00	83,75	<b>86,25</b>
g	63,75	65,00	73,75	87,50	<b>88,75</b>
h	76,25	80,00	70,00	78,75	<b>83,75</b>
i	86,25	88,75	86,25	86,25	<b>90,00</b>
j	<b>81,25</b>	77,50	<b>81,25</b>	<b>81,25</b>	78,75
k	65,00	72,50	72,50	72,50	<b>80,00</b>
l	70,00	62,50	72,50	62,50	<b>80,00</b>
m	50,00	53,75	<b>85,00</b>	73,75	82,50
n	76,25	65,00	<b>82,50</b>	77,50	<b>82,50</b>
o	88,75	87,50	90,00	91,25	<b>92,50</b>
p	76,25	87,50	88,75	86,25	<b>92,50</b>
q	<b>95,00</b>	92,50	88,75	93,75	<b>95,00</b>
r	77,50	72,50	66,25	78,75	<b>82,50</b>
s	88,75	78,75	81,25	87,50	<b>95,00</b>
t	87,50	88,75	87,50	90,00	<b>92,50</b>
u	87,50	83,75	68,75	85,00	<b>91,25</b>
v	80,00	82,50	80,00	81,25	<b>86,25</b>
w	83,75	87,50	<b>90,00</b>	85,00	<b>90,00</b>
x	83,75	86,25	85,00	80,00	<b>87,50</b>
y	90,00	90,00	87,50	77,50	<b>91,25</b>
z	82,50	85,00	<b>88,75</b>	85,00	<b>88,75</b>
Média (%)	<b>81,10</b>	<b>80,81</b>	<b>81,63</b>	<b>83,61</b>	<b>88,13</b>

A abordagem pela seleção dinâmica de zoneamento teve melhor resultado para a maioria das classes de caracteres, com exceção das classes “a”, “j” e “m”.

Analisando as confusões a partir da matriz de confusão com o zoneamento indicado dinamicamente pode-se perceber que vários caracteres que foram confundidos são praticamente irreconhecíveis até mesmo por humanos (no caso, por meio de inspeção visual desta pesquisadora). A Figura 4.13 ilustra alguns exemplos de caracteres que são visivelmente irreconhecíveis e que, conseqüentemente, foram confundidos com caracteres distintos:

- caractere “j” que foi confundido com “i”;
- caractere “f” que foi confundido com “r”;
- caractere “k” que foi confundido com “h”;
- caractere “r” que foi confundido com “i”;
- caractere “f” que foi confundido com “t”;

A Tabela 4.20 ilustra a matriz de confusão obtida por meio dos experimentos com a abordagem pela seleção dinâmica de zoneamento.

Tabela 4.20. Matriz de Confusão - Zoneamento selecionado dinamicamente

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	Rec%
a	68	0	0	0	2	0	0	4	0	0	0	0	0	0	2	0	0	0	0	1	1	0	0	2	0	0	85,00
b	0	73	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	6	0	0	0	0	0	0	0	91,25
c	0	0	77	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	96,25
d	0	0	0	75	0	0	0	0	0	3	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	93,75
e	0	0	0	0	70	1	0	1	0	0	1	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	87,50
f	0	0	0	0	0	69	2	0	0	0	1	2	0	0	0	3	0	1	0	2	0	0	0	0	0	0	86,25
g	0	0	0	0	0	4	71	0	0	0	0	0	0	0	0	3	0	1	0	0	0	0	0	0	1	0	88,75
h	0	0	0	1	0	0	0	67	0	0	7	0	0	1	0	0	0	2	0	1	1	0	0	0	0	0	83,75
i	0	0	0	0	0	0	0	0	72	4	0	2	0	0	0	2	0	0	0	0	0	0	0	0	0	0	90,00
j	0	1	0	0	0	1	0	0	6	63	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	7	78,75
k	0	1	0	0	1	0	0	6	0	0	64	5	0	1	0	0	0	0	0	2	0	0	0	0	0	0	80,00
l	0	0	0	0	1	0	1	0	8	3	1	64	0	0	0	0	0	0	0	0	0	0	0	0	0	2	80,00
m	0	0	0	0	0	0	0	1	0	0	0	0	66	10	0	1	0	2	0	0	0	0	0	0	0	0	82,50
n	0	0	0	0	0	0	0	2	0	0	0	0	9	66	0	0	0	0	0	0	3	0	0	0	0	0	82,50
o	0	0	1	0	0	0	2	0	0	0	0	0	0	0	74	0	0	0	0	0	3	0	0	0	0	0	92,50
p	0	0	0	0	0	1	0	0	1	0	0	2	0	0	0	74	2	0	0	0	0	0	0	0	0	0	92,50
q	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	76	0	0	0	0	0	0	1	0	0	95,00
r	0	0	0	0	0	1	0	4	1	0	1	0	1	2	0	2	0	66	0	0	0	2	0	0	0	0	82,50
s	0	2	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	76	0	0	0	0	0	0	0	95,00
t	0	0	0	0	0	0	0	0	1	0	3	0	0	0	0	0	0	2	0	74	0	0	0	0	0	0	92,50
u	0	0	0	1	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	73	0	2	0	1	0	91,25
v	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	1	69	6	0	1	0	86,25
w	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	6	2	72	0	0	0	90,00
x	0	0	0	0	0	0	0	0	0	0	5	0	1	1	0	0	0	2	0	0	0	0	0	70	1	0	87,50
y	0	0	0	0	0	0	1	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3	73	1	91,25
z	0	0	0	0	0	0	3	0	0	2	0	1	0	0	0	0	0	0	0	0	0	0	0	1	2	71	88,75
Taxa média de reconhecimento																										88,13	

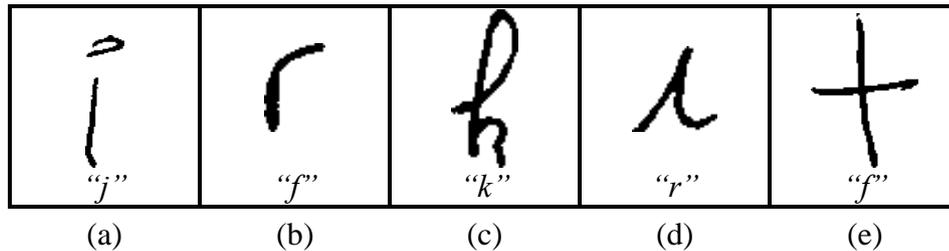


Figura 4.13. Caracteres incorretamente reconhecidos

#### 4.6. Considerações Finais

Neste capítulo foram apresentados os resultados dos experimentos realizados com as RNA MLP Convencional e Classe-Modular. Ao se comparar as duas arquiteturas verificou-se o aumento nas taxas de reconhecimento para todos os mecanismos de zoneamentos investigados quando se utiliza a arquitetura Classe-Modular. O reconhecimento com a seleção dinâmica de zoneamento considerando a utilização da arquitetura Convencional treinada com Z4 (Nível 1) e a arquitetura Classe-Modular (Nível 2) ocasionou aumento nas taxas de reconhecimento, permitindo alcançar taxas de reconhecimento de 88,13%, mostrando ser um enfoque promissor.

Foram ainda analisados as matrizes de confusão dos mecanismos de zoneamento  $Z=4$ ,  $Z=5H$ ,  $Z=5V$ ,  $Z=7$  e  $Z=9$ . A análise das matrizes de confusão permitiu entender a relação entre as regiões e as confusões geradas entre as classes de caracteres.

A partir dos experimentos realizados com os caracteres maiúsculos e minúsculos foi realizada uma análise em termos das características perceptivas dos caracteres. Dessa análise pode-se verificar que caracteres minúsculos apresentam mais assimetria que caracteres maiúsculos e também que a diversidade de estilos da escrita de caracteres minúsculos é mais contrastante que em caracteres maiúsculos. As classes maiúsculas tiveram melhor reconhecimento nos experimentos para o Z4 e Z7, já as classes minúsculas não tiveram predomínio de reconhecimento por nenhum dos zoneamentos investigados. Quando tratados separadamente, os caracteres apresentaram taxas de reconhecimento maiores, sendo assim o tratamento separado de caracteres maiúsculos e minúsculos é recomendável.

## Conclusão

O presente trabalho teve por objetivo contribuir para a área de reconhecimento de caracteres, dando continuidade as pesquisas realizadas por Freitas (2007) [FRE07] e Aires (2005) [AIR05], buscando incorporar aspectos da percepção humana no processo de reconhecimento. Assim, propôs e utilizou uma metodologia de seleção dinâmica de mecanismos de zoneamento para auxiliar no processo de reconhecimento de caracteres manuscritos minúsculos, considerando um léxico de 26 classes. Foram discutidos e comparados os resultados obtidos nos experimentos para caracteres minúsculos e maiúsculos com o objetivo de se compreender como os caracteres são diferentes em termos de formação de traçado e de características perceptivas.

As taxas de reconhecimento e as análises dos zoneamentos investigados indicam que o uso de aspectos perceptivos, em saber em quais partes do caractere está acontecendo à confusão, é um fator muito importante para que um mecanismo de zoneamento seja eficiente. Assim, foi constatado neste trabalho que a utilização de não somente um mecanismo de zoneamento perceptivo, mas sim um conjunto deles, pode ser fundamental para o aumento das taxas de reconhecimento. Isto, por meio de um classificador *a priori* que selecione o melhor mecanismo de zoneamento para cada amostra de caractere. Sabe-se que é difícil comparar os resultados deste trabalho com trabalhos de outros autores, os quais utilizaram técnicas diversas e base de dados diferentes, contudo acredita-se que os resultados alcançados sejam promissores. Acredita-se assim que o método proposto é promissor, tem grandes atrativos, podendo ainda ser melhorado.

Foram apresentados neste trabalho os experimentos utilizando uma abordagem por seleção dinâmica de zoneamento para o reconhecimento de caracteres manuscrito minúsculo, composta por 2 níveis: Nível 01 + Nível 02.

No Nível 01 foi utilizado uma RNA MLP convencional para predizer qual o zoneamento mais indicado para cada amostra de caractere numa base de teste, utilizou-se também o zoneamento Z4 por apresentar melhor resultado que o zoneamento global.

No Nível 2 o reconhecimento propriamente dito foi feito utilizando as RNA MLP com arquitetura Classe Modular treinadas com os 4 zoneamentos perceptivos: Z4, Z5H, Z5V, Z7.

Os 4 zoneamentos utilizados para compor o EoZ não foram escolhidos aleatoriamente, considerou-se os estudos preliminares feitos por Aires(2005)[AIR05] e Freitas et al. (2007) [FRE07] para caracteres manuscritos maiúsculos, e assim, partindo dos resultados promissores que estes autores tiveram em seus experimentos é que se passou a investigar e avaliar se os zoneamentos perceptivos e a classe modularidade das RNAs também eram válidos para os caracteres manuscritos minúsculos. Os experimentos realizados no presente trabalho com as RNA MLP Classe Modular para caracteres manuscritos minúsculos permitiram alcançar as seguintes taxas de reconhecimento: Z4: 81,10%, Z5H: 80,81%, Z5V: 81,63%, Z7: 83,60% . Partindo desses resultados animadores, pode-se então formar o conjunto de EoZ necessário para a seleção dinâmica de zoneamento para o reconhecimento de caracteres manuscritos minúsculos. Associou-se assim o que de melhor cada mecanismo de zoneamento tem, ou seja, cada um deles pode reconhecer melhor uma ou outra classe de caractere, devido ao fato de cada mecanismo de zoneamento individualmente e diferentemente ressaltar as características perceptivas de cada classe. Os resultados alcançados nos experimentos utilizando a abordagem de seleção dinâmica de zoneamento permitiu alcançar taxa de reconhecimento de 88,13%.

Finalmente foi mostrada também uma análise comparativa sobre as diferenças entre caracteres manuscrito maiúsculos e minúsculos em termos de características perceptivas. A análise leva em consideração que caracteres manuscrito maiúsculo e minúsculo tem características diferentes, diferindo no traçado das letras, e estilos de escrita. A análise considerou também que muitas vezes não existe a necessidade de diferenciar caracteres manuscritos maiúsculo de minúsculo, nesse caso, um “a” e um “A”, formam uma única classe. Os experimentos com caracteres minúsculas + maiúsculas alcançaram as seguintes taxas de reconhecimento: Z4: 77,70%, Z5H: 79,31%, Z5V: 79,71%, Z7: 82,23%. A partir da análise dos resultados e das matrizes de confusão e verificando imagem dos caracteres percebeu-se que caracteres maiúsculos tem reconhecimento mais alto pelos zoneamentos Z4 e Z7, e em geral apresentaram estilos de escritas menos contrastantes. Para caracteres minúsculos observou-se que não há predomínio aparente de mecanismo de zoneamento entre as classes. Tendo por base a percepção visual e, portanto, que caracteres maiúsculos e

minúsculos são diferentes na forma, tamanho, traços, concavidades e convexidades, observa-se que o tratamento separado de caracteres minúsculos e maiúsculos é recomendável.

Os resultados do presente trabalho mostraram-se promissores com a predição de um mecanismo de zoneamento, contudo o problema do reconhecimento de caracteres manuscrito é muito complexo, e se é possível prever qual o melhor mecanismo de zoneamento para determinado caractere, não há como prever o estilo de escrita que será utilizado por determinado autor. Seres humanos têm características distintas e a escrita de cada indivíduo conseqüentemente terá características diversas.

Assim sendo, ainda há muito a ser feito para que os sistemas de reconhecimento de manuscritos sejam infalíveis, e algumas investigações podem ser consideradas pra desenvolvimento futuro:

- Utilização de classificadores diferentes no Nível 01
- Utilização de bases de dados distintas que tenham características diferentes, como por exemplo a base NIST
- Investigação do uso de diferentes classificadores, como por exemplo as MEMs no Nível 02 que utilizem os mesmos mecanismos de zoneamento aqui analisados

Enfim, investigar se a abordagem aqui apresentada é válida também para outros problemas no domínio de Reconhecimento de manuscrito.

## Referências

- [AIR05] AIRES, S.B. K., *Reconhecimento de caracteres manuscritos baseado em regiões perceptivas*. Dissertação de Mestrado, PUCPR/PPGIa, Curitiba-PR, 2005, 82p.
- [BLU03] BLUMENSTEIN, M. VERMA, B., BASLI, H. *A Novel Feature Extraction Technique for the Recognition of Segmented Handwritten Characters*. Proceedings of the Seventh International Conference on Document Analysis and Recognition, p.137, August 03-06, 2003
- [BRA00] BRAGA, A.P. CARVALHO, A.P.L., LUDERMIR, T.B. *Redes neurais artificiais – Teoria e aplicações*. Rio de Janeiro: LTC – Livros Técnicos e Científicos Editora S.A. 2000
- [BRI04] BRITTO Jr.. A. S.; SABOURIN. R.; BORTOLOZZI. F.; SUEN C. Y.. *Foreground and background information in an HMM-Based method for recognition of isolated characters and numeral strings*. 9th Int. Workshop on Frontiers in Handwriting Recognition (IWFHR-9). Tokio Japan. pp. 371-37 (2004)
- [CHE07] CHERIET, M. KHARMA, N., LIU,C.L.,SUEN,C.Y. *Character Recognition System –A guide for students and practioners*. USA: Wiley – Interscience 2007
- [COR02] CORREIA, S.E.N., CARVALHO, J.M., SABOURIN, R. *Human-Perception Handwritten Character Recognition using Wavelets*. In Proceedings of XVI Brazilian Symposium on Computer Graphics and Image Processing, IEEE Computer Society, 2002, p.404.
- [FRE01a] FREITAS, C.O.A. *Uso de modelos escondidos de Markov para reconhecimento de palavras manuscritas*. Tese de Doutorado, PUCPR/PPGIA, Curitiba-PR, 2001, 188p.
- [FRE07] FREITAS, C.O.A., OLIVEIRA, L.E.S., BORTOLOZZI, F., AIRES, S.B.K...: *Handwritten Character Recognition using Non-Symmetrical Percpetual Zoning*,

International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI), Vol. 21, 1, pp. 1–21 (2007).

- [GAU99] GAUDIN, C.V., LALLICAN, P. M., BINTER, P., KNERR, S.: *The IRESTE On/Off (IRONOFF) Dual Handwriting Database*, Fifth International Conference on Document Analysis and Recognition. ICDAR, pp.455, 1999.
- [HEU98] HEUTTE, L., PAQUET, T., MOREAU, J.V., LECOURTIER Y., OLIVIER,C.: *A structural/statistical feature based vector for handwritten character recognition*, Pattern Recognition Letters, vol. 19, no. 7, pp. 629-641 (1998).
- [JAI96] A.K. JAIN, J. MAO, AND K.M. MOHIUDDIN, *Artificial Neural Networks: A Tutorial*, Computer. Mar. 1996, pp. 31-44.
- [JAI00] JAIN, A.K.; DUIN, R.P.W. ; & MAO, J. *Statistical Pattern Recognition: A Review*, IEEE Trans. Pattern Analysis and Machine Intelligence,Jan. 2000, vol. 22, no. 1, pp. 4-37.
- [KAP03] KAPP, M.N., FREITAS, C.O.D.A., NIEVOLA, J.C., SABOURIN, R.: *Evaluating the Conventional and Class-Modular Architectures Feedforward Neural Network for Handwritten Word Recognition*. In SIBGRAPI, 2003, 315-322
- [KAP04] KAPP, M.N., *Reconhecimento de Palavras Manuscritas Utilizando Redes Neurais Artificiais*. Dissertação de Mestrado, PUCPR/PPGIa, Curitiba-PR, 2004, 98p.
- [KOE03] KOERICH, A. L.; SABOURIN, R. and SUEN, C. Y. *Large vocabulary off-line handwriting recognition: A survey*. Pattern Analysis and Applications, 2003 6(2):97–121.
- [KOE03b] KOERICH, A.L.; *Improving Classification Performance Using Metaclasses*. IEEE International Conference on Systems, Man, and Cybernetics (SMC2003), Vol.1, pp.717-722, Washington DC, USA, October 2003. ISBN 0-7803-7952-7
- [KOE05] KOERICH, A.L& KALVA, P. R.. *Unconstrained Handwritten Character Recognition Using Metaclasses of Characters* IEEE International Conference on Image

Processing (ICIP2005), Vol.2, pp.542-545, Genova, Italy, September 2005. ISBN 0-7803-9134-9

- [KO08] KO, A. H. R.; SABOURIN, R.; BRITTO, A. S. Jr. *From dynamic classifier selection to dynamic ensemble selection*. Pattern Recognition, v.41 n.5, p.1718-1731, May, 2008
- [LEC00] LECCE, V.; DIMAURO, G.; GUERRIERO, A.; IMPEDOVO, S.; PIRLO G.; SALZO, A.; *Zoning design for handwritten numerical recognition*, 7th Int. Workshop on Frontiers in Handwriting Recognition (2000), pp. 583–588.
- [LI95] LI, Z.C., SUEN, C.Y., GUO, J. *A Regional Decomposition Method for Recognizing Handprinted Characters*. IEEE, Transactions on Systems, Man, and Cybernetics, Vol.25, junho 1995, p. 998-1010.
- [MAD01] MADHVANATH, S., & GOVINDARAJU, V.. *The role of holistic paradigms in handwritten word recognition*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(2), pp.149–164, 2001.
- [NIE04] NIEVOLA, J. C. *Redes Neurais Artificiais*. In: Sociedade Brasileira de Computação. (Org.). Inteligência Artificial. 1 ed. Porto Alegre: Editora da Sociedade Brasileira de Computação - ESBC, 2004, v. 1, p. 1-50.
- [OH01] OH, S.; LEE, J.S., SUEN, C.Y. "A Class-Modularity for Character Recognition," *icdar*, pp.0064, Sixth International Conference on Document Analysis and Recognition (ICDAR'01), 2001
- [OLI02] OLIVEIRA JR, J. J.; CARVALHO, J.M. DE C.; FREITAS, C. O.A.; SABOURIN R. *Evaluating NN and HMM classifiers for handwritten word recognition*. 15th Brazilian Symposium on Computer Graphics and Image Processing, 2002, p. 210-217.
- [OTS79] OTSU, N., *A threshold selection method from gray-level histograms*, IEEE Transactions Systems, Man. and Cybernetics, SMC 9, Vol.1, 1979, p.63-66

- [PAR97] Parker, J. R.: *Algorithms for Image Processing and Computer Vision*, John Wiley (1997)
- [PLA00] PLAMONDON, R.; SRIHARI, S. N. *On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey*. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22, NO.1, jan.2000, p.63-84.
- [RAD03] RADTKE, P.V.W.; OLIVEIRA, L.S.; SABOURIN, R. WONG, T.; *Intelligent zoning design using multi-objective evolutionary algorithms*, 7th Int. Conf. Document Analysis and Recognition (ICDAR2003), 2003, p. 824–828.
- [SUE94] SUEN, C.Y., GUO, J., LI, Z.C. *Analysis and Recognition of Alphanumeric Handprints by parts*. IEEE, Transactions on Systems, Man, and Cybernetics, Vol.24, abril 1994, p. 614-631.
- [TAY01] TAY, Y.H, LALLICAN, P.M, KHALID, M., GAUDIN, C.V, KNERR, S.. *An offline Cursive Handwritten Word Recognition System*. IEEE Region 10 Conference, TENCON 2001, Singapore, 2001, p.19-22.
- [TRI96] Ø. D. TRIER, A.K. JAIN, T. T. *Feature extraction methods for character recognition—A survey*, Pattern Recognition, 29:4(1996), 641–662.
- [XIA05] XIANG, P., XIUZI Y., SANYUAN Z., *A hybrid method for robust car plate character recognition*. Engineering Applications of Artificial Intelligence, V.18 n.8, December, 2005, p.963-972.
- [ZHA00] ZHANG, G. P. *Neural networks for classification: A survey*, IEEE Transactions Systems, Man, and Cybernetics-Part C: Applications and Reviews, Vol. 30, No. 4. 2000, p. 451–462.