

UNIVERSIDADE FEDERAL DE SÃO CARLOS

CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA

PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA
COMPUTAÇÃO

**"Arquitetura de Redes Neurais para o Reconhecimento
Facial Baseado no Neocognitron**

ORIENTADOR: Prof. Dr. José Hiroki Saito

CO-ORIENTADOR: Prof. Dr. Ednaldo Brigante Pizzolato

ALUNO: Ângelo Rodrigo Bianchini

**São Carlos
Julho/2001**

**Ficha catalográfica elaborada pelo DePT da
Biblioteca Comunitária da UFSCar**

B577ar

Bianchini, Ângelo Rodrigo.

Arquitetura de redes neurais para o reconhecimento facial baseado no neocognitron / Ângelo Rodrigo Bianchini. - São Carlos : UFSCar, 2004.

141 p.

Dissertação (Mestrado) -- Universidade Federal de São Carlos, 2001.

1. Processamento de imagens. 2. Redes neurais (computação). 3. Reconhecimento facial. 4. Rede neural neocognitron. I. Título.

CDD: 006.42 (20^a)

Dedicat3rio

A minha v3, Dizolina Coleone Bianchini,
pela dedica33o e carinhos concedidos.

(In Memoriam)

Agradecimentos

A Deus, porque através d'Ele tudo é possível.

Ao Prof. José Hiroki Saito, pela orientação, amizade, paciência, compreensão e principalmente pela oportunidade oferecida para o desenvolvimento deste trabalho.

Ao Prof. Ednaldo Brigante Pizzolato pela orientação, ajuda e constante incentivo para o desenvolvimento do trabalho.

À minha família pelo incentivo.

Aos meus tios Erminia e Pedro e a minha prima Adriana pelo carinho e apoio.

À minha irmã Silvana Aparecida Bianchini e ao meu cunhado Ciro Rosa pelo carinho, ajuda e compreensão concedidos.

Aos professores e funcionários do programa de pós-graduação.

Aos colegas de pós-graduação Nassim, Luciano, Jacqueline, Mateus, Márcia, Reinaldo, Luciana, Milene e demais pela ajuda nos momentos necessários.

Aos amigos Fernando Salina, Fernando Teixeira, Miriam, Cláudia e Jefferson pelo constante apoio.

A CAPES pelo auxílio financeiro concedido.

A todas as pessoas que contribuíram para que esse trabalho pudesse ser realizado.

Lista de Figuras

2.1- Diversidade de imagens faciais de um mesmo indivíduo.....	27
2.2- Neurônio biológico.....	29
2.3- Neurônio artificial.....	30
2.4- Analogia entre neurônio biológico e artificial.....	30
2.5- Exemplo do neurônio do Perceptron.....	33
2.6- Exemplo da arquitetura do Perceptron Multicamada.....	34
2.7- Tipos de regiões de decisão que podem ser formadas pelo (a) Perceptron e (b) Perceptron Multicamadas	35
2.8- Redes Neurais Recorrentes	36
2.9- Exemplo de uma rede probabilística para a detecção de faces.....	38
2.10- Exemplo de uma rede convolucional.....	40
2.11- Exemplo simplificado do neocognitron para reconhecimento de caracteres.	40
2.12- Exemplo detalhado do neocognitron para reconhecimento de caracteres.....	41
3.1- Módulos do sistema para detecção facial.....	43
3.2- Exemplo da aplicação de máscara.....	44
3.3- Modelo de face genérica baseada na distribuição – (a) Distribuição da amostra facial; (b) Aproximação com clusters gaussianos; (c) Distribuição da amostra não-facial, sobreposta aos contornos dos clusters faciais; e (d) Aproximação com clusters gaussianos	46
3.4- Medidas de diferença entre o padrão de teste e os centróides.....	47
3.5- Módulos do sistema para detecção facial.....	50
3.6- Exemplo de imagem não facial e falsas detecções.....	51
3.7- Exemplo de sobreposições e falsas detecções.....	52
3.8- Heurística para detecções sobrepostas – (a) Pirâmide de entrada com detecções sobrepostas; (b) Detecção dos centros; (c) “Spreading Out” das detecções; (d) Detecção do centróide; (e) Avaliação dos centróides; e (f) Resultado Final	53
3.9- Arbitração entre múltiplas redes para eliminar detecções sobrepostas.....	53
3.10- Módulos do sistema para reconhecimento facial.....	58

3.11- PDBNN para a tarefa de reconhecimento facial.....	58
3.12- Diferentes tipos de imagens utilizadas no teste.....	59
3.13- Módulos “B” do sistema para reconhecimento facial.....	60
4.1- Estrutura hierárquica da rede neocognitron.....	62
4.2- Diagrama de uma célula-S	62
4.3- Modelo de abrangência das conexões entre as células dos diferentes planos celulares.....	63
4.4- Seqüência de reconhecimento do neocognitron.....	64
4.5- Algoritmo para computar um estágio.....	64
4.6- Seqüência de treinamento do neocognitron	65
4.7- Treinamento de um estágio no neocognitron.....	66
4.8- Etapas para o treinamento não-supervisionado no neocognitron.....	68
4.9- Exemplo de erro na distinção dos padrões dos números 5 e 6, em decorrência da similaridade dos padrões extraídos na camada Uc2.....	70
4.10- Exemplo de distinção correta dos padrões dos números 5 e 6	70
4.11- Arquitetura de rede convolucional para reconhecimento de padrões de imagens médicas	72
4.12- Exemplo de saída utilizando-se a associação de incerteza (fuzzy) construídas pela distribuição gaussiana	73
5.1- Etapas da NEOPC: (a) Pré-Processamento, (b) RNDPC: obtenção dos pontos de controle, (c) RNRFF: treinamento e (d) RNRFF: reconhecimento...	77
5.2- Exemplo de segmentação por <i>threshold</i> : (a) imagem original e (b) histograma em níveis de cinza da (a) imagem “A”	78
5.3- Segmentação da imagem facial utilizando-se valores diferentes de <i>threshold</i> : (a) imagem original, (b) <i>threshold</i> = 0.70 e (c) <i>threshold</i> = 0.90..	79
5.4- Exemplo de segmentação da imagem facial: (a) imagem original, (b) sombreamento da imagem original, (c) histograma em níveis de cinza da imagem original, (d) histograma em níveis de cinza do sombreamento da imagem original, (e) resultado da segmentação imagem original com <i>threshold</i> = 0.90, e (f) resultado da segmentação do sombreamento da imagem original com <i>threshold</i> de 0.90.....	80
5.5- Diagrama das etapas do pré-processamento	81

5.6-	Etapas do Pré-Processamento: o pré-processamento é realizado na ordem em que é demonstrado na Figura, iniciando da parte superior, da esquerda para direita: (a) imagem original, (b) sombreamento da imagem, (c) segmentação do sombreamento da imagem original com o uso de <i>threshold</i> , e (d) redimensionamento da imagem (57x57)	82
5.7-	Ilustração dos estágios da RNDPC	83
5.8-	Diferenças em três neurônios biológicos	85
5.9-	RNDPC – Obtenção dos pontos de controle considerando um limiar para detecção de bordas e de linhas de 0.84	86
5.10-	RNDPC – Obtenção dos pontos de controle considerando um limiar para detecção de bordas de 0.60 e de linhas de 0.84	87
5.11-	Exemplo do treinamento não supervisionado utilizando os pontos de controle	88
5.12-	RNRF no reconhecimento de amostras	90
5.13-	Base de Imagens 1: classes e amostras subfaciais – olho esquerdo	92
5.14-	Base de Imagens 1: classes e amostras faciais completas	94
5.15-	Base de Imagens 1: classes e amostras segmentadas	95
5.16-	Espaço físico elaborado para aquisição das imagens.....	104
5.17-	Base de Imagens 2: exemplos de amostras	105
5.18-	Base de Imagens 2: taxas de erros apresentadas durante o reconhecimento, considerando-se 3 classes e diferentes quantidades de amostras	109
5.19-	Base de Imagens 2 - Taxas de erros apresentadas durante o reconhecimento, considerando-se 5 classes e diferentes quantidades de amostras	112
5.20-	Base de Imagens 2 - Analogia entre as taxas de erros obtidas variando o número de classes e amostras contidas na base de imagens 2	113
5.21-	Base de Imagens 3: exemplos de amostras	114
5.22-	Base de Imagens 4: exemplos de amostras	116
5.23-	Base de Imagens 4: Taxas de erros apresentadas durante o reconhecimento, considerando-se 3 classes e diferentes quantidades de amostras	125
5.24-	Base de Imagens 4: Taxas de erros apresentadas durante o reconhecimento, considerando-se 6 classes e diferentes quantidades de amostras	125
5.25-	Base de Imagens 4: Taxas de erros apresentadas durante o reconhecimento, considerando-se 10 classes e diferentes quantidades de amostras	129

5.26- Base de Imagens 4: Taxas de erros apresentadas durante o reconhecimento, considerando-se 17 classes e diferentes quantidades de amostras	133
5.27- Base de Imagens 4: Analogia entre as taxas de erros obtidas variando o número de classes e amostras	133

Lista de Tabelas

3.1- Proporção de Detecção x Distância Métricas.....	49
3.2- Proporção de Detecção x Composição do Protótipo.....	49
3.3- Número de Falsas Detecções para o Conjunto de Teste I.....	55
3.4- Número de Falsas Detecções para o Conjunto de Teste II.....	56
3.5- Resultados obtidos através do módulo “A” (A Falsa Aceitação foi computada usando padrões de uma base de dados desconhecida).....	59
3.6- Resultados obtidos através do módulo “B” (A Falsa Aceitação foi computada usando padrões de uma base de dados desconhecida).....	60
5.1- Base de Imagens 1 – Amostras Subfaciais – Olho Esquerdo: <i>Thresholds</i> e taxa de reconhecimento considerando-se as mesmas amostras para o treinamento e reconhecimento (5 classes contendo 5 amostras cada uma)....	93
5.2- Base de Imagens 1 – Amostras Faciais Completas: <i>Thresholds</i> e taxa de reconhecimento considerando-se as mesmas amostras para o treinamento e reconhecimento (5 classes contendo 5 amostras cada uma).....	94
5.3- Base de Imagens 1 – Amostras Faciais Segmentadas: <i>Thresholds</i> e taxa de reconhecimento considerando-se as mesmas amostras para o treinamento e reconhecimento (5 classes contendo 5 amostras cada uma).....	97
5.4- Base de Imagens 1 – Amostras Faciais Segmentadas - Teste Aumentando o Número de Classes: <i>Thresholds</i> e taxa de reconhecimento considerando-se as mesmas amostras para o treinamento e reconhecimento (10 classes contendo 10 amostras cada uma).....	98
5.5- Base de Imagens 1 – Amostras Faciais Segmentadas - Teste usando Padrões de Treinamento distintos dos Padrões de Reconhecimento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 5 amostras cada uma para o treinamento e cinco amostras para o reconhecimento).....	99

5.6-	Base de Imagens 1 – Amostras Faciais Segmentadas - Teste usando Padrões de Reconhecimento de Classes distintas das Classes dos Padrões de Treinamento: <i>Thresholds</i> e taxa de erro considerando-se classes diferentes para o treinamento e reconhecimento (5 classes contendo 5 amostras cada uma).....	100
5.7-	Base de Imagens 1 – Amostras Faciais Segmentadas - Teste usando Padrões de Imagens Faciais contendo Acessórios: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 5 amostras cada uma para o treinamento e cinco amostras para o reconhecimento).....	101
5.8-	Base de Imagens 1 – Amostras Faciais Segmentadas - Número de Fatores Treinados x Quantidade de Amostras sem Acessórios: número de fatores treinados pelas camadas Us3, Us4 e Us5, de acordo com o número de classes e amostras utilizadas.....	102
5.9-	Base de Imagens 1 – Amostras Faciais Segmentadas - Número de Fatores Treinados x Quantidade de Amostras com Acessórios: número de fatores treinados pelas camadas Us3, Us4 e Us5, de acordo com o número de classes e amostras utilizadas.....	103
5.10-	Base de Imagens 2 – 3 Classes e 5 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 5 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	106
5.11-	Base de Imagens 2 – 3 Classes e 10 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 10 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	107
5.12-	Base de Imagens 2 – 3 Classes e 15 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 15 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	107
5.13-	Base de Imagens 2 – 3 Classes e 20 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 20 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	108

5.14-	Base de Imagens 2 – 3 Classes e 25 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 25 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	108
5.15-	Base de Imagens 2 – 5 Classes e 5 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 5 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	110
5.16-	Base de Imagens 2 – 5 Classes e 10 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 10 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	110
5.17-	Base de Imagens 2 – 5 Classes e 15 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 15 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	111
5.18-	Base de Imagens 2 – 5 Classes e 20 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 20 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	111
5.19-	Base de Imagens 2 – 5 Classes e 25 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 25 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	112
5.20-	Base de Imagens 3 – 3 Classes e 5 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 5 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	112
5.21-	Base de Imagens 4 – 3 Classes e 5 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 5 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	117

5.22-	Base de Imagens 4 – 3 Classes e 10 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 10 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	118
5.23-	Base de Imagens 4 – 3 Classes e 10 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 10 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	118
5.24-	Base de Imagens 4 – 3 Classes e 20 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 20 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	119
5.25-	Base de Imagens 4 – 3 Classes e 30 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 30 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	119
5.26-	Base de Imagens 4 – 3 Classes e 40 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 40 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	120
5.27-	Base de Imagens 4 – 3 Classes e 50 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 50 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	120
5.28-	Base de Imagens 4 – 6 Classes e 5 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (6 classes contendo 5 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	121
5.29-	Base de Imagens 4 – 6 Classes e 10 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (6 classes contendo 10 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	122

5.30- Base de Imagens 4 – 6 Classes e 15 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (6 classes contendo 15 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	122
5.31- Base de Imagens 4 – 6 Classes e 20 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (6 classes contendo 20 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	123
5.32- Base de Imagens 4 – 6 Classes e 30 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (6 classes contendo 30 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	123
5.33- Base de Imagens 4 – 6 Classes e 40 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (6 classes contendo 40 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	124
5.34- Base de Imagens 4 – 6 Classes e 50 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (6 classes contendo 50 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	124
5.35- Base de Imagens 4 – 10 Classes e 5 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (10 classes contendo 5 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	125
5.36- Base de Imagens 4 – 10 Classes e 10 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (10 classes contendo 10 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	126
5.37- Base de Imagens 4 – 10 Classes e 15 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (10 classes contendo 15 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	126

5.38- Base de Imagens 4 – 10 Classes e 20 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (10 classes contendo 20 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	127
5.39- Base de Imagens 4 – 10 Classes e 30 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (10 classes contendo 30 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	127
5.40- Base de Imagens 4 – 10 Classes e 40 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (10 classes contendo 40 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	128
5.41- Base de Imagens 4 – 10 Classes e 50 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (10 classes contendo 50 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	128
5.42- Base de Imagens 4 – 17 Classes e 5 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (17 classes contendo 5 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	129
5.43- Base de Imagens 4 – 17 Classes e 10 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (17 classes contendo 10 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	130
5.44- Base de Imagens 4 – 17 Classes e 15 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (17 classes contendo 15 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	130
5.45- Base de Imagens 4 – 17 Classes e 20 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (17 classes contendo 20 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	131

5.46- Base de Imagens 4 – 17 Classes e 30 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (17 classes contendo 30 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	131
5.47- Base de Imagens 4 – 17 Classes e 40 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (17 classes contendo 40 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	132
5.48- Base de Imagens 4 – 17 Classes e 50 Amostras para o Treinamento: <i>Thresholds</i> e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (17 classes contendo 50 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).....	132

Resumo

Nesta dissertação de mestrado é apresentado um sistema de reconhecimento facial baseado no neocognitron (NEOPC). Este sistema é dividido em três importantes etapas: (a) pré-processamento: que normaliza as amostras faciais utilizadas nas arquiteturas de redes neurais, (b) Rede Neural para Detecção dos Pontos de Controle (RNDPC): que obtém os pontos de controles que são adicionados na estrutura de treinamento não-supervisionado do neocognitron e (c) Rede Neural para o Reconhecimento Facial (RNRF): que realiza o reconhecimento das amostras faciais.

A principal característica da NEOPC é a utilização dos pontos de controle para a extração de fatores localizados em posições estratégicas, tais como olhos, nariz e lábios, utilizados no treinamento não-supervisionado da RNRF.

Foram realizados os testes, da NEOPC, considerando amostras faciais com variações de ângulos e expressões, e diferentes quantidades de classes e amostras para a fase de treinamento e reconhecimento.

Foram utilizadas, para a obtenção dos resultados, uma base de imagens disponibilizada pela Universidade de Cambridge e três bases de imagens desenvolvidas como parte deste trabalho.

Para comparar os vários resultados obtidos com as quatro bases de imagens utilizadas nesta dissertação, foram apresentados, em todas as bases, informações relevantes, tais como: os *thresholds* utilizados para a excitação dos neurônios, a taxa de acerto e a taxa de erro.

Os resultados demonstraram que o desempenho do NEOPC aumenta com a quantidade de amostras utilizadas para o treinamento, até um ponto ótimo, e posteriormente passa a decrescer. Isso significa que novos atributos de novas amostras aumentam o número de planos celulares dificultando o reconhecimento.

Abstract

In this dissertation it is presented a face recognition system based in the neocognitron (NEOPC). This system is divided into three major stages: (a) pré-processing: which normalizes the facial samples utilized in the neural network architectures; (b) Neural Network for Detection of Control Points (NNDCP): which obtains the control points to be used in the non supervised training structure of the neocognitron and (c) Neural Network for the Facial Recognition (NNFR), which carried out the recognition of face samples.

The main characteristic of NEOPC is the use of control points for the extraction of patterns strategically located, such as eyes, noses and lips, used on non supervised training of the NNFR.

The tests of the NEOPC were carried out, considering facial samples having variations of angles and expressions, and different amounts of classes and samples to the training and recognition stages.

For the attainment of the results, a base of images released by the University of Cambridge and three bases of images developed as part of this work.

In order to compare the several results obtained with the four bases of images investigated in this dissertation we present in all of them relevant information, such as, the thresholds used for the excitation of neurons, rightness and error rates.

The results shows that the NEOPC performance increases with the number of samples used for the training, until an optimal point, and then decreases. It is explained by the increase of the cell numbers with the features from new samples, difficulty the recognition.

Sumário

1- Introdução.....	20
1.1 – Organização do Trabalho.....	22
2- Reconhecimento Facial e Redes Neurais.....	23
2.1 – Reconhecimento Facial.....	23
2.2 – Redes Neurais.....	28
2.2.1 – Modelos de Redes Neurais.....	32
2.2.1.1 – Perceptron.....	32
2.2.1.2 – Perceptron Multicamada.....	33
2.2.1.3 – Rede Neural Recorrente.....	36
2.2.1.4 – Rede Neural Baseada na Decisão Probabilística.....	36
2.2.1.5 – Rede Neural Convolutacional.....	39
3- Redes Neurais Aplicadas a Detecção e Reconhecimento Facial.....	42
3.1 – Detecção de Faces baseada no Aprendizado.....	42
3.1.1 – Pré-Processamento.....	43
3.1.2 – Modelo Baseado na Distribuição.....	44
3.1.3 – Medidas de Diferença – Padrões x Modelo.....	45
3.1.4 – Classificador.....	46
3.2 – Detecção de Faces baseada em Redes Neurais.....	50
3.3 – Aplicação da PDBNN para o Reconhecimento Facial.....	57
4- Neocognitron e suas Aplicações.....	61
4.1 – Neocognitron.....	61
4.2 – Aplicações com o Neocognitron.....	67
4.2.1 – Reconhecimento de Padrões no Neocognitron provido por uma Adaptação Neural	67

4.2.2 – Uma Arquitetura de Rede Neural Convolutacional baseada no Neocognitron para o Reconhecimento de Padrões de Imagens Médicas.....	71
5- Especificação do Projeto e Resultados.....	76
5.1 – Projeto.....	76
5.1.1 – Pré-Processamento.....	78
5.1.2 – Rede Neural para a Detecção dos Pontos de Controle (RNDPC)	83
5.1.3 – Rede Neural para o Reconhecimento Facial (RNRF).....	87
5.2 – Resultados.....	90
5.2.1 – Base de Imagens 1.....	91
5.2.1.1 – Resultados para Amostras Subfaciais – Olho Esquerdo	91
5.2.1.2 – Resultados para Amostras Faciais Completas.....	93
5.2.1.3 – Resultados para Amostras Faciais Segmentadas.....	95
5.2.1.3.1 – Teste Aumentando o Número de Classes....	98
5.2.1.3.2 – Teste usando Padrões de Treinamento distintos dos Padrões de Reconhecimento.....	99
5.2.1.3.3 – Teste usando Padrões de Reconhecimento de classes distintos dos Padrões de Treinamento	100
5.2.1.3.4 – Teste usando Padrões de Imagens Faciais contendo Acessórios	101
5.2.1.3.5 – Número de Fatores Treinados x Amostras Utilizadas	102
5.2.2 – Base de Imagens 2.....	103
5.2.2.1 – Teste Considerando 3 Classes e Diferentes Quantidades de Amostras	106
5.2.2.2 – Teste Considerando 5 Classes e Diferentes Quantidades de Amostras	109
5.2.3 – Base de Imagens 3.....	113
5.2.4 – Base de Imagens 4.....	115
5.2.4.1 – Teste Considerando 3 Classes e Diferentes Quantidades de Amostras Distintas para o Treinamento e Reconhecimento.....	117

5.2.4.2 – Teste Considerando 6 Classes e Diferentes Quantidades de Amostras Distintas para o Treinamento e Reconhecimento.....	121
5.2.4.3 – Teste Considerando 10 Classes e Diferentes Quantidades de Amostras Distintas para o Treinamento e Reconhecimento.....	125
5.2.4.4 – Teste Considerando 17 Classes e Diferentes Quantidades de Amostras Distintas para o Treinamento e Reconhecimento.....	129
6- Conclusão.....	135
7- Bibliografias	137

Capítulo I

Introdução

Há vários anos um grande número de pesquisas vem sendo realizadas para prover sistemas de reconhecimento automático, devido à grande demanda de aplicações necessárias no mercado. Um exemplo clássico é o da área bancária. Os métodos convencionais baseados no uso de cartões e senhas não são mais altamente confiáveis. Os cartões podem ser perdidos ou extraviados, enquanto que as senhas podem ser fraudadas [Bol00]. Assim, instituições bancárias perdem milhões a cada ano em decorrência das fraudes existentes [Cell99].

Uma das primeiras técnicas “não-convencionais” utilizada em sistemas de reconhecimento automático foi a impressão digital. Esses sistemas têm sido comercialmente utilizados desde 1960, embora tenham sido aplicados, quase que exclusivamente, em áreas forenses para investigações criminais.

Atualmente, novas técnicas de reconhecimento por meio de faces, íris, retina e voz, além de outras, têm sido abordadas para aplicações em sistemas de reconhecimento automático [Bol00].

O reconhecimento facial pode ser considerado como uma das mais importantes funções do ser humano, por permitir identificar uma grande quantidade de faces e aspectos psicológicos demonstrados pela fisionomia.

Diversas áreas de pesquisas relacionadas ao reconhecimento facial, como detecção, representação e processo, têm despertado um interesse ainda maior, em decorrência da vasta possibilidade existente de aplicações: sistemas de controle automático de acesso, interface homem-máquina, sistemas de vigilância, etc [Fre92].

Embora o reconhecimento facial venha de encontro a várias necessidades atuais de aplicações, esta ainda é uma área de pesquisa que pode ser considerada como um problema clássico da visão artificial. Isto se deve à complexidade existente na detecção e reconhecimento de características e padrões.

No entanto, em virtude das novas exigências do mercado, o avanço das técnicas de reconhecimento facial vem ocorrendo de maneira surpreendente.

Aproximadamente há 20 anos atrás, o reconhecimento facial era uma área totalmente em estudo entre os grandes problemas da inteligência artificial e visão computacional. Hoje, apesar de ainda ser considerado como um problema clássico, há vários sistemas para reconhecimento facial capazes de atuarem em grandes bases de dados com resultados consideráveis.

O reconhecimento facial vem se desenvolvendo junto à “quarta geração” de computadores e tecnologia da informação através de sua aplicação na nova geração de interfaces. A próxima geração de interfaces detectará a presença de pessoas e as reconhecerão.

Porém, ainda há muitas linhas de pesquisas para serem abordadas a ponto de tornar o reconhecimento facial como parte integrante da nova geração de interfaces.

Neste trabalho são apresentados alguns resultados para uma nova arquitetura de redes neurais baseado no neocognitron. Optou-se pela arquitetura do neocognitron em decorrência dos resultados obtidos por Fukushima[Fuk95] para o reconhecimento de caracteres manuscritos. Espera-se, que o avanço de uma pesquisa estruturada na proposta de uma arquitetura de redes neurais para o reconhecimento facial baseada na visão biológica, como o neocognitron, possa apresentar as soluções necessárias para os problemas atuais, devido principalmente à estrutura convolucional dessas redes.

1.1 – Organização do Trabalho

Esta dissertação está organizada nos seguintes capítulos, além do presente capítulo I, que tem caráter introdutório:

- Capítulo II – Reconhecimento Facial e Redes Neurais: é apresentada uma abordagem conceitual sobre o reconhecimento facial e redes neurais;
- Capítulo III – Redes Neurais Aplicadas a Detecção e Reconhecimento Facial: são apresentadas algumas pesquisas desenvolvidas para a detecção e o reconhecimento facial. Este capítulo visa a interação com a área, bem como análise dos diferentes tipos de redes neurais;
- Capítulo IV – Neocognitron e suas Aplicações: são apresentadas a arquitetura do neocognitron e algumas pesquisas relacionadas às suas aplicações;
- Capítulo V – Especificação do Projeto e Resultados: são apresentados a arquitetura desenvolvida e os resultados obtidos;
- Capítulo VI – Conclusões; e
- Capítulo VII – Referências Bibliográficas.

Capítulo II

Reconhecimento Facial e Redes Neurais

Neste capítulo será apresentada uma abordagem conceitual sobre o reconhecimento facial e redes neurais, uma vez que estes tópicos consistem no alicerce sobre o qual está embasada a pesquisa. Os conceitos apresentados sobre redes neurais consistem apenas em definições gerais, já que o próximo capítulo estará abordando suas aplicações no reconhecimento facial.

2.1 - Reconhecimento Facial

Apesar do constante avanço tecnológico, a interação humano-computador ainda continua fortemente baseada sobre sinais e comportamentos que não são naturais no cotidiano.

Não se pode negar que o teclado e o mouse sejam provas desse avanço, mas, no entanto, o desenvolvimento de interfaces mais naturais e intuitivas que não requerem treinamento especializado ainda é motivo de constantes pesquisas [Dau97].

Dentre as áreas que têm se dedicado a estabelecer um novo tipo de interação humano-computador está a “visão computacional”. Embora normalmente conhecida por esta titulação, esta área possui uma vasta variedade de denominações, como: visão cibernética, visão de máquina, visão robótica e visão por computador.

Pertencente a área do Reconhecimento e Análise de Padrões, o reconhecimento facial é um dos ramos de pesquisa abordado pela visão computacional.

No início dos anos 70, o reconhecimento facial foi baseado principalmente sobre atributos faciais mensuráveis como os olhos, nariz, sobrancelhas, etc. Devido à escassez dos recursos computacionais e à ineficiência dos algoritmos de extração de características, somente um pequeno número de testes foram conduzidos e a performance dos sistemas não atingiu o esperado. Então, as pesquisas na área ressurgiram por volta dos anos 90, inovando os métodos existentes [Hon98].

Atualmente, ainda pode-se considerar o reconhecimento facial como um campo quase que exclusivamente em estudos [LeD95], e, conseqüentemente, com poucas aplicações no mercado.

Um exemplo bastante promissor da aplicação do reconhecimento facial é o reconhecimento da “identidade pessoal”. Esta aplicação pode ser realizada em automação de escritórios, personalização automática de ambiente, aplicações em segurança de sistemas, criminologias, controle de acesso físico e prevenção de fraude. Hoje, os meios mais utilizados para se verificar a “identidade pessoal” estão centrados no uso de cartões magnéticos e senhas, que requerem sua utilização durante uma transação, mas que não verificam sua idoneidade [Dau97].

Outro motivo que certamente contribuirá para o reconhecimento facial, são as vantagens que o mesmo possui em relação ao reconhecimento por impressão digital e por íris. No caso do reconhecimento por impressão digital a desvantagem encontra-se no fato de que nem todas as pessoas possuem uma impressão digital “boa” que possa ser reconhecida por um sistema. Quanto ao reconhecimento pela íris, já ficou comprovado que a mesma apresenta uma textura de alta confiabilidade e de larga variação, sendo estável pela vida toda. Neste caso o problema está relacionado ao modo de captura da íris. Isso ocorre devido ao alinhamento que deve haver entre a câmera e os olhos da pessoa. Apesar de algumas soluções já terem sido apresentadas para este problema, como um sistema de três câmeras que fazem a procura pela íris em uma face sem que haja um perfeito alinhamento, elas ainda são muito caras e impraticáveis [Dau97] [Cell99].

O reconhecimento facial não pode ser visto de maneira isolada, pois consiste em uma técnica que faz uso de diversas tarefas que compõem o processamento facial.

Dependendo do âmbito de aplicação, um sistema de reconhecimento totalmente automatizado pode ser composto por tarefas preliminares, como a detecção e segmentação [Sun98].

Deste modo, o processamento facial é composto por diferentes tarefas que variam entre as aplicações, de acordo com o objetivo almejado. Dentre as principais tarefas encontram-se:

1. Classificação: é a designação de uma face visualmente baseada em categorias como sexo, pele e olhos [Abd95];
2. Identificação: é a categorização de um nome ou outra semântica chave (por exemplo, meu médico) [Abd95]; é a verificação se uma face pertence ao conjunto de faces “conhecidas” [Cel99];
3. Reconhecimento: é a categorização de uma face como familiar ou não [Abd95];
4. Detecção: dada uma imagem qualquer, identifica o local onde aparece(m) a(s) face(s) [Cel99];
5. Segmentação: identifica as partes constituintes de uma face [Cel99];
6. Representação: seleciona as informações da face que serão utilizadas para representá-la [Cel99].

Como pode ser verificado, a definição da tarefa *identificação* possui uma variação entre os conceitos apresentados.

Pode-se considerar que cada uma destas tarefas usa diferentes características das faces e diferentes abordagens computacionais [LeD95].

Uma das dificuldades dos modelos computacionais é tratar a complexidade dos padrões visuais. Embora todas as faces sejam compostas por padrões reconhecidos universalmente (boca, olhos e nariz), elas possuem variações únicas. Assim, para o reconhecimento facial torna-se de grande importância a utilização destas variações sutis na determinação das características relevantes [Abd95].

Outro importante aspecto a ser considerado em relação às características a serem utilizadas para a tarefa de reconhecimento, é que elas possuam uma larga variação estatística para serem consideradas únicas para cada indivíduo. O ponto ideal é que a variância inter-classe seja grande e a intra-classe pequena, para que faces diferentes gerem códigos os mais distintos possíveis, enquanto diferentes imagens de um mesma face gerem códigos bastantes similares [Dau97].

Portanto, estabelecer uma *representação* que capture as características ideais é um difícil problema. Uma solução que vem sendo utilizada é a aplicação das redes neurais diretamente sobre as imagens faciais, como uma poderosa ferramenta para selecionar as informações mais relevantes para cada tarefa [Abd95].

Pode-se identificar pelo menos dois tipos de categorias de sistemas para o reconhecimento facial [Law97]:

- sistemas onde o objetivo é reconhecer a face de uma pessoa dentro de uma base de imagens, como no caso da área policial. Estes sistemas geralmente retornam um lista das possíveis faces correspondentes aos padrões de entradas e não necessitam usualmente ser em tempo real; e
- sistemas onde o objetivo é reconhecer a face de uma pessoa em particular e em tempo real, como os sistemas de segurança, ou permitir o acesso para um grupo de pessoas e negar para outras, como no caso de acesso ao computador.

Dentre os principais problemas encontrados para a tarefa de reconhecimento facial, encontram-se:

- iluminação;

-
- ângulos e poses;
 - expressões;
 - cosméticos e acessórios;
 - extração da face do contexto ou do fundo.

Na Figura 2.1 são mostradas diversas imagens faciais de um mesmo indivíduo, onde existem variações dos itens mencionados.

Outra questão bastante polêmica é em relação à melhor maneira de representação das faces: métodos baseados na aparência 2D ou métodos baseados no modelo 3D. Dentro deste contexto, surgem alguns questionamentos. Por exemplo, em ambos os casos

- qual seria a variância ótima das características para extração dos dados?
- como decidir a melhor maneira durante uma incerteza?

Devido a falta de evidência científica que comprove uma determinada técnica, o problema de representação das faces ainda está aberto.

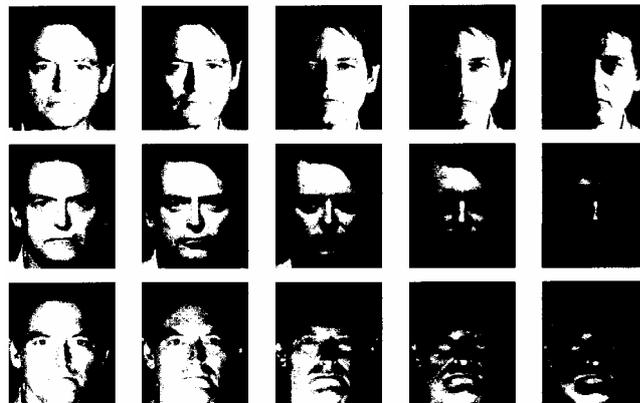


Figura 2.1 – Diversidade de imagens faciais de um mesmo indivíduo.

2.2 – Redes Neurais

Existem diversas definições utilizadas para redes neurais. A seguir serão apresentadas duas delas:

- uma rede neural artificial é um sistema de processamento de informação que tem como princípio alguma inspiração numa rede neural biológica. As redes neurais artificiais têm sido desenvolvidas como generalizações de modelos matemáticos da cognição humana ou biologia neural [Fau94];
- uma rede neural artificial é um processador distribuído paralelamente para armazenar conhecimentos experimentais e torná-los disponíveis para uso. O conhecimento é adquirido de forma semelhante ao cérebro, através de um processo de aprendizado [Hay94].

Uma vez que as redes neurais artificiais tiveram sua origem inspirada no modelo do sistema nervoso biológico, será apresentado, a seguir, o modelo biológico do neurônio, seguido de um modelo computacional do mesmo.

Modelo Biológico: conforme pode ser verificado na Figura 2.2, o neurônio é composto pelos seguintes componentes:

- dentrites: responsáveis pela coleta dos impulsos provenientes dos outros neurônios;
- corpo celular: também conhecido por soma, é responsável pelo processamento dos sinais recebidos;
- axônio: responsável pela propagação dos sinais;

Os dentrites e o corpo celular constituem a principal parte da superfície de entrada do neurônio. A conexão entre o axônio de um neurônio e o dentrite de outro é chamado de sinapse [Arb95].

Sendo o neurônio a unidade básica de processamento, a união de bilhões deles constituem o sistema nervoso.

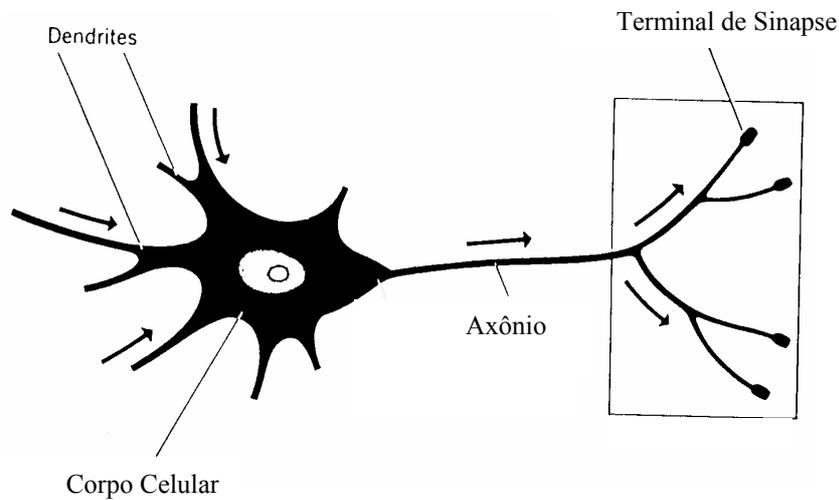


Figura 2.2 – Neurônio biológico

Modelo Computacional: o neurônio artificial (Figura 2.3) apresenta características semelhantes ao do modelo biológico, podendo ser identificados os seguintes elementos:

- um conjunto de sinapses: conforme especificado no modelo biológico, a sinapse consiste na entrada dos sinais para o neurônio. No caso do neurônio computacional, para cada sinal de entrada existe um peso associado ao mesmo no qual será realizada a operação de multiplicação. Assim, cada neurônio pode tratar o impulso recebido de maneira diferente, atribuindo-lhe um valor. Outro fator importante a ser ressaltado, é que os pesos atribuídos serão positivos se a sinapse associada é excitatória e negativos se a sinapse associada é inibitória;
- somador: que realiza o somatório de todos os resultados provenientes da multiplicação dos sinais pelos pesos; e
- função de ativação: responsável por limitar a amplitude de saída de um neurônio.

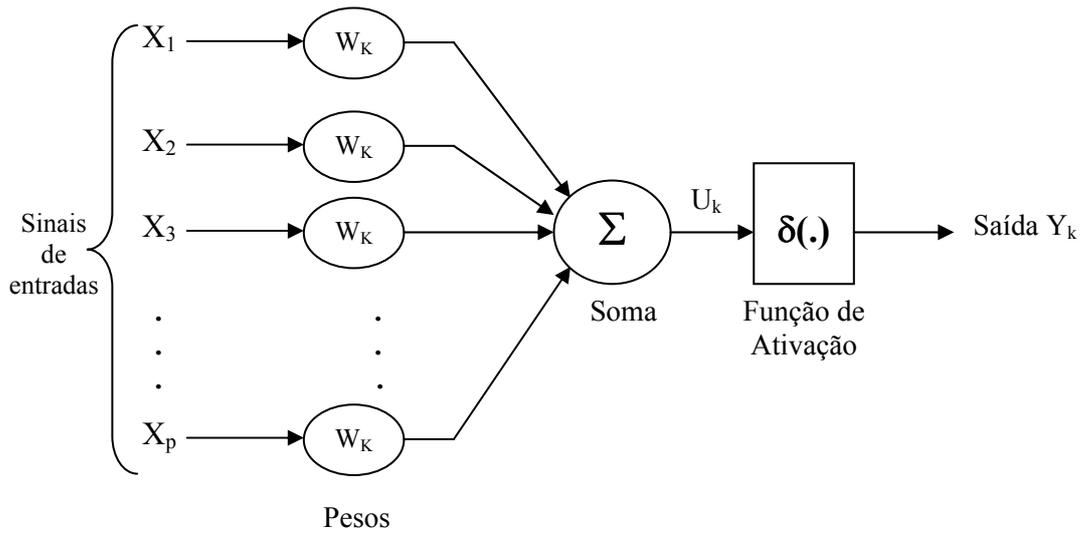


Figura 2.3 – Neurônio Computacional

A Figura 2.4 representa, de modo geral, a analogia entre ambos os modelos apresentados:

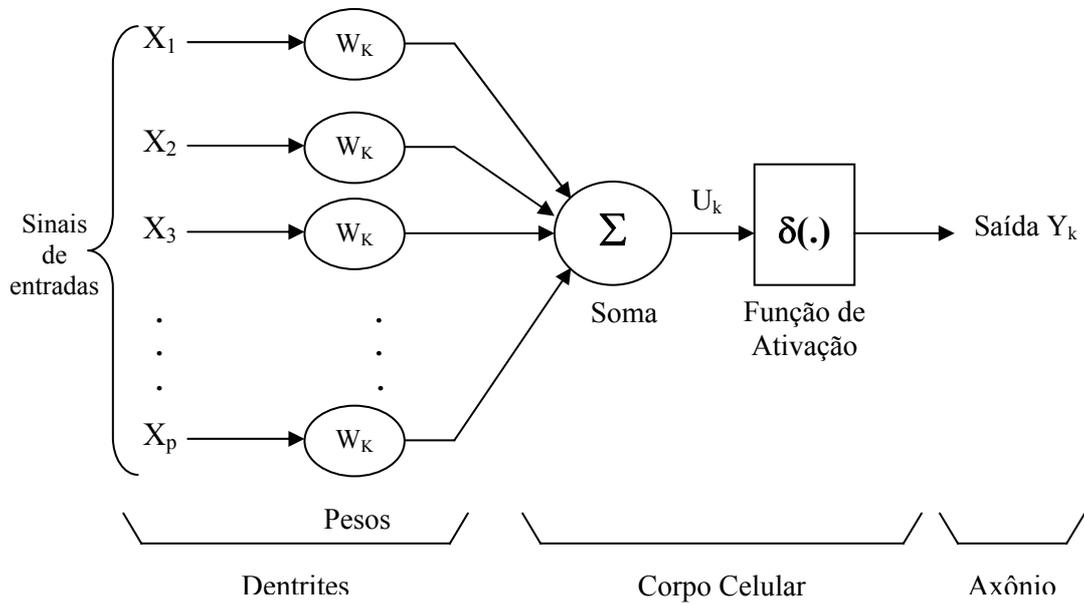


Figura 2.4 – Analogia entre o neurônio biológico e artificial

Soulié [Sou95] apresenta quatro categorias principais onde as redes neurais são aplicadas:

- classificação e diagnóstico: um padrão é apresentado para a rede, que decide por qual conjunto de classes pré-definidas o mesmo pertence. Consiste em uma situação típica de reconhecimento de padrões;
- função de aproximação: a rede tem que associar ao padrão de entrada um valor, representando o fenômeno real. Este problema pode ser encontrado em predição de séries temporais, identificação e controle de processos;
- compressão, extração de características e quantização: em muitas aplicações, dados não processados, como os recebidos de sensores, possuem alta dimensão e ruído. As redes neurais podem ser usadas para reduzir a dimensionalidade e ruído, ambos por compressão numérica ou através de características relevantes. A quantização, supervisionada ou não-supervisionada, pode também ser utilizada para representar os dados através de um conjunto limitado de referência; e
- otimização: as redes neurais podem ser um conjunto, de modo que seus pesos incluem tanto a função alvo a ser otimizada como o confinamento sobre as variáveis. A solução é então obtida como o estado limite das redes dinâmicas.

Já as principais áreas onde as redes neurais têm sido utilizadas destacam-se:

- aplicações militares: processamento de sinais (por exemplo, radar) para identificação de alvos ou fontes, fusão de dados, análise de imagem para localização e identificação de alvos;
- processamento de imagem: reconhecimento de faces, objetos, caracteres e análise de imagens de satélites;
- voz: reconhecimento de voz e identificação da fala;

-
- manufatura: prognóstico do consumo de energia e controle de qualidade;
 - controle de processo: identificação e controle de processos industriais;
 - robótica;
 - biologia e medicina: análise de imagens, processamento de sinais e diagnósticos automáticos;
 - petróleo: análise de imagem sísmicos;
 - física de alta energia: detecção de eventos;
 - setor terciário: autenticação de assinaturas, detecção de fraude e marketing;
 - finanças: índices econômicos e análise de risco;

As redes neurais artificiais, do mesmo modo que a rede neural biológica, são constituídas por diversos neurônios. A maneira em que os neurônios da rede estão estruturados e são treinados estabelecem diferentes tipos de modelos, demonstrados no próximo sub-item.

2.2.1 – Modelos de Redes Neurais

2.2.1.1 – Perceptron: o perceptron é um modelo de rede neural utilizado para a classificação de padrões linearmente separáveis [Hay94], ou seja, ele aprende uma função de decisão para classificar dois conjuntos linearmente separáveis [Gon92]. Este modelo consiste apenas de um neurônio e utiliza o método de aprendizado supervisionado.

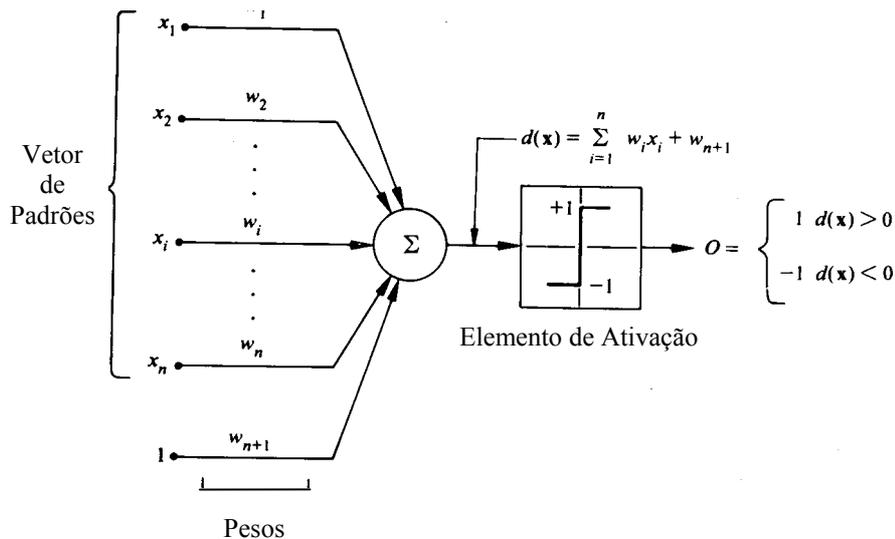


Figura 2.5 – Exemplo do neurônio do Perceptron

A Figura 2.5 exemplifica o modelo de um perceptron. Sua resposta é baseada na soma dos pesos de entrada;

$$d(x) = \sum_{i=1}^n w_i x_i + w_{n+1} \quad (\text{I})$$

os coeficientes $w_i, i = 1, 2, 3, \dots, n, n+1$, chamados pesos, modificam as entradas antes delas serem somadas e enviadas para o elemento de ativação. A função de decisão de limiar responsável pela ativação é chamada por função de ativação (threshold). No caso do exemplo apresentado, o perceptron responderá 1 quando $d(\mathbf{x}) > 0$, indicando que o padrão x foi reconhecido como pertencente à classe w , e -1 caso contrário.

2.2.1.2 – Perceptron Multicamadas (Multilayer Perceptron – MLP): o perceptron multicamada representa uma generalização do perceptron, sendo também conhecido como rede multicamada progressiva (*feedforward*). O MLP é denominado como sendo progressiva (*feedforward*) pelo fato dos neurônios não possuírem suas saídas (axônios) conectadas (sinapses) aos neurônios localizados nas camadas anteriores. O MLP consiste em uma importante arquitetura de rede neural, pelo fato de ser

aplicável a problemas de reconhecimento de padrões, independente das classes serem separáveis linearmente [Gon92] [Hay94]. A Figura 2.7 apresenta o tipo de região de decisão que podem ser formadas por redes “*feedforward*”.

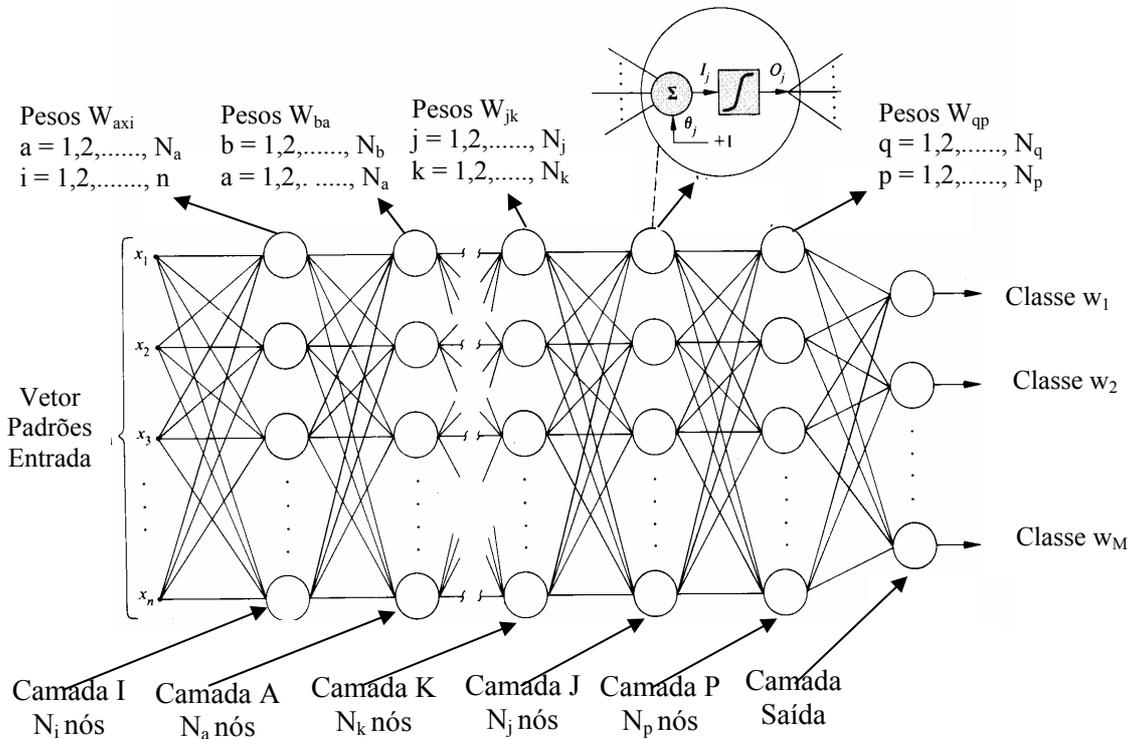


Figura 2.6 – Exemplo da arquitetura do Perceptron Multicamada

O MLP possui três características básicas [Hay94]:

- cada neurônio na rede inclui uma função de decisão linear, como por exemplo uma função de ativação sigmoide. No caso da função sigmoide, os valores de saída poderão oscilar entre -1 e 1 , prevalecendo como resposta a classe relacionada ao neurônio da camada de saída que apresentar o maior valor;
- a rede contém uma ou mais camadas ocultas; e
- a rede apresenta um alto grau de conectividade entre as camadas.

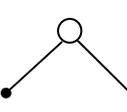
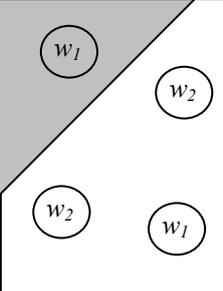
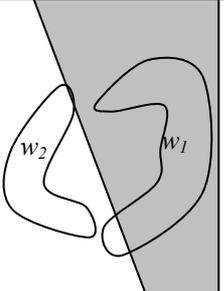
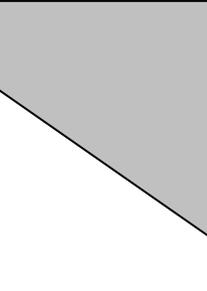
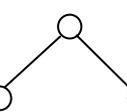
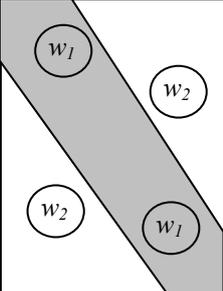
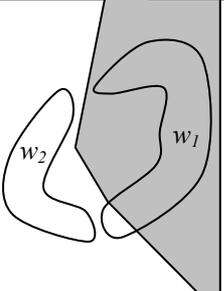
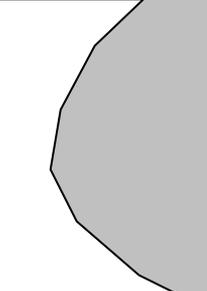
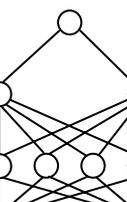
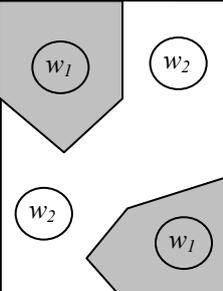
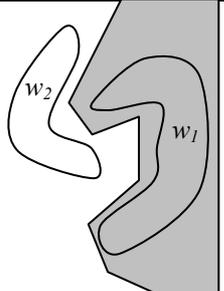
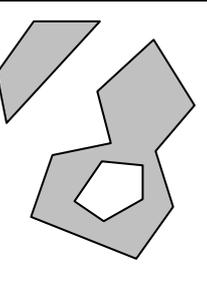
	Estrutura da rede	Tipo de região de decisão	Solução para o problema do OU-exclusivo	Classes com regiões entrelaçadas	Formas mais gerais das superfícies de decisão
(a) Perceptron	Camada única 	Único hiperplano			
(b) Perceptron Multicamada	Duas camadas 	Regiões convexas abertas ou fechadas			
	Três camadas 	Arbitrária (complexidade limitada pelo número de nós)			

Figura 2.7 – Tipos de regiões de decisão que podem ser formadas pelo (a) Perceptron e (b) Perceptron Multicamadas [Gon92].

O MLP é uma rede supervisionada, tendo o seu algoritmo de treinamento composto por dois passos:

- feedforward: um vetor de padrões é apresentado para a camada de entrada, sendo propagado pela rede, camada por camada. Assim, cada neurônio ao receber um sinal, calcula sua função de ativação e envia o valor para os neurônios da próxima camada, até alcançar os neurônios de saída. Os neurônios de saída calculam suas ativações resultando no valor de saída da rede neural.
- backpropagation: como se trata de uma rede supervisionada que requer uma resposta desejada para ser treinada, o erro para cada neurônio de saída é

calculado comparando o valor obtido para a saída com o valor alvo. Esse erro é então propagado para as camadas inferiores da rede, obtendo-se a ponderação dos erros da camada superior.

Esses passos são repetidos até os erros dos neurônios de saída serem considerados satisfatórios.

2.2.1.3 – Rede Neural Recorrente: as redes neurais recorrentes distinguem-se das MLP por possuírem pelo menos um *loop* de realimentação (feedback). Por exemplo, uma rede neural recorrente pode consistir de uma única camada de neurônios, com cada neurônio alimentando as entradas dos demais neurônios (Figura 2.8).

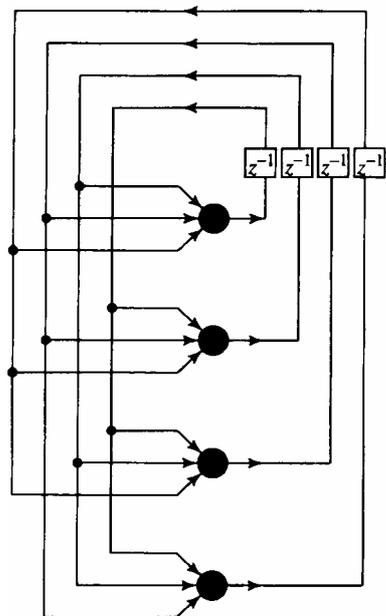


Figura 2.8 – Redes Neurais Recorrentes

2.2.1.4 – Rede Neural Baseada na Decisão Probabilística (PDBNN): consistem de uma variação das DBNN (redes neurais baseadas em decisão), sendo relevante para a sua especificação os seguintes aspectos:

- regras de aprendizado: sendo uma extensão das DBNN, as PDBNN utilizam-se da regra de aprendizado não-supervisionado localmente (neurônios ocultos) e

supervisionado globalmente (saída), o qual consiste em uma das possíveis regras utilizadas no modelo DBNN.

Também conhecido por LUGS (*Locally Unsupervised Globally Supervised*), este esquema de treinamento é composto por duas fases:

- LU (*Locally Unsupervised*): cada subrede é treinada individualmente. Os valores dos parâmetros na rede são inicializados nesta fase, que pode adotar um dos esquemas de aprendizado não supervisionado, como por exemplo, o K-médias [Dud73]
 - GS (*Globally Supervised*): nesta fase é utilizada a informação do professor para melhor estabelecer os limites de decisão (*threshold*). Quando um padrão de treinamento é mal-classificado, utilizam-se as técnicas de aprendizado reforço ou anti-reforço para adequar os limites de decisão (estas técnicas são exemplificadas no sub-item “Atualização do *Threshold*”).
- função discriminante: a subrede de função discriminante é designada para modelar as funções probabilísticas. A partir do resultado apresentado pela função discriminante, é realizada a adequação dos pesos para o grau de possíveis envolvimento (mensurados pela probabilidade) para cada *cluster*.

O exemplo de uma função discriminante probabilística seria: dado um conjunto de “i” padrões $X = \{ x(t); t=1,2,3,4... N \}$, e assumindo-se que a função probabilística $p(x(t)/w)$ para a classe w (por exemplo, classe de faces e classe de não faces) é uma mistura de distribuição gaussiana, pode-se definir $p(x(t)/w, \Theta r)$ para ser uma das distribuições gaussianas que compreende em:

$$p(x(t) / w) = \sum_{r=1}^R P(\Theta r / w) p(x(t) / w, \Theta r) \quad (\text{II})$$

onde Θr representa o $r^{\text{ésimo}}$ *cluster* na subrede e

$P(\Theta r/w)$, a probabilidade a priori do *cluster* r .

Em outras palavras, a probabilidade de um dado vetor de entrada $x(t)$ pertencer a uma classe w é obtido através do somatório das probabilidades em relação ao número de *cluster* existente, calculando-se para cada *cluster* a probabilidade a priori do mesmo pertencer à classe w ($P(\Theta_r/w)$) e a probabilidade do vetor de entrada $x(t)$ pertencer a uma classe w juntamente com a probabilidade de ocorrer o *cluster* em questão ($p(x(t)/w, \Theta_r)$).

Um modelo de função discriminante probabilística pode ser dado como:

$$\log\left[\sum_r P(\Theta_r/w) p(x(t)/\Theta_r, w)\right] \quad (III)$$

A Figura 2.9 demonstra um exemplo da aplicação da função discriminante apresentada, para a tarefa de detecção de faces.

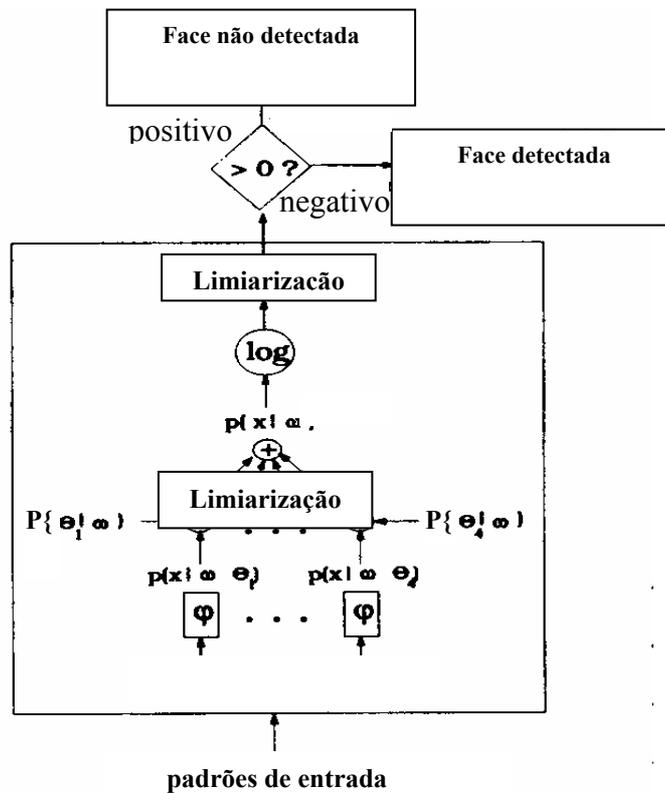


Figura 2.9 – Exemplo de uma rede probabilística para a Detecção de Faces

Assim, pode-se estabelecer, por exemplo, o vetor de entrada $x(t)$ pertencente a classe “face” caso o resultado da função logarítmica esteja acima do *threshold*, e pertencente a classe não-face caso contrário.

- atualização do *threshold*: o valor do *threshold* (T) também pode ser aprendido através de regras de aprendizado de reforço e anti-reforço. Por exemplo, dada uma entrada $x(t)$, se $x(t) \in w$, mas $p(x(t)/w) < T$, então o valor de T deverá ser reduzido (técnica anti-reforço). No entanto, se $x(t) \notin w$ mas $p(x(t)/w) > T$, então T deverá ser incrementado (técnica reforço).

2.2.1.5 – Rede Neural Convolutacional: um problema clássico encontrado nas redes neurais é a habilidade para lidar com situações onde ocorram deformações e variações nos dados de entrada. As redes convolucionais visam tratar estes tipos de situações.

Pode-se apresentar como um exemplo típico de rede neural convolutacional a Figura 2.10. Ela consiste de um conjunto de camadas contendo um ou mais planos. A imagem é apresentada para a primeira camada, conceituada como camada de entrada. Cada unidade em um plano recebe entrada de uma pequena vizinhança de um plano da camada prévia. Os campos receptivos locais em um plano possuem pesos estabelecidos que são forçados para serem iguais, no intuito de extrair as mesmas características. Cada plano pode ser considerado como um mapeamento de características, onde são detectadas características pré-estabelecidas. Assim, múltiplos planos são usualmente utilizados em cada camada para que múltiplas características possam ser detectadas. Estas camadas são denominadas de camadas convolucionais. Após cada camada convolutacional encontram-se as camadas que realizam as operações de sub-amostragem, responsáveis pelo fator de tolerância a distorções da rede.

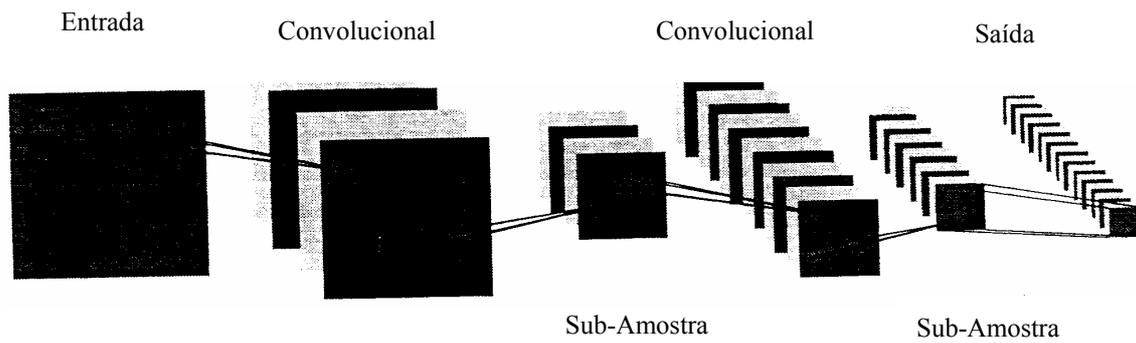


Figura 2.10 – Exemplo de uma rede convolucional

Um exemplo de arquitetura que apresenta os princípios de redes neurais convolucionais e pode ser considerada como sua primeira realização é o neocognitron¹ [Neu98]. A Figura 2.11 ilustra, de maneira simplificada, o funcionamento do neocognitron para o reconhecimento de caracteres [Fuk95] e a Figura 2.12 mostra de maneira mais detalhada as suas camadas. Cada retângulo da Figura 2.12 representa os resultados obtidos por uma camada convolucional e de sub-amostragem. Os padrões detectados por cada camada estão relacionados a um plano, isto é, cada plano de uma camada é treinado para reconhecer um tipo diferente de fator.

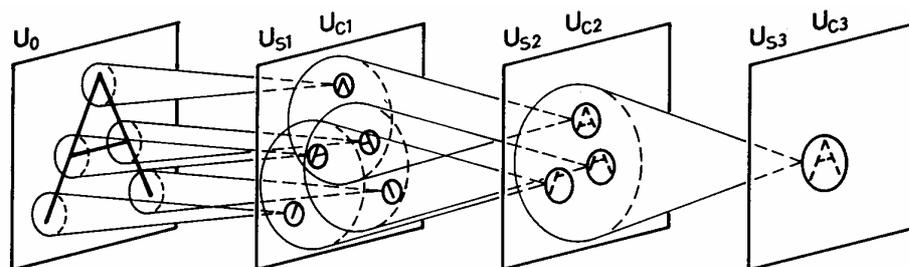


Figura 2.11 – Exemplo simplificado do neocognitron para reconhecimento de caracteres.

¹ A arquitetura do neocognitron será especificada detalhadamente no capítulo V.

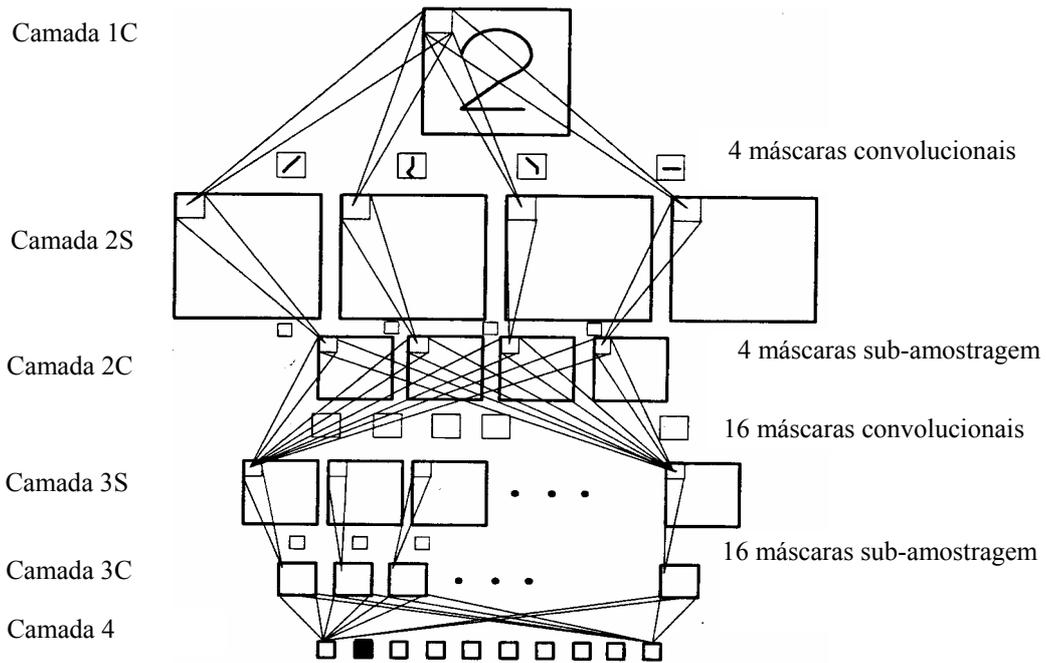


Figura 2.12 – Exemplo detalhado do neocognitron para reconhecimento de caracteres.

Pode-se concluir este capítulo com alguns conceitos. No tópico de reconhecimento facial pode-se observar relatos importantes sobre a viabilidade da tarefa de reconhecimento através do uso de faces ao invés da impressão digital ou íris, a viabilidade do uso de redes neurais e os problemas encontrados para as aplicações. Quanto às redes neurais, além da breve definição dos conceitos e aplicações apresentados, pode-se verificar diferentes arquiteturas e maneiras de propagação dos sinais caracterizando diferentes modelos.

No próximo capítulo serão apresentados alguns exemplos de aplicações de redes neurais no processamento facial.

Capítulo III

Redes Neurais Aplicadas ao Reconhecimento Facial

Conforme foi apresentado no capítulo anterior, a aplicação das redes neurais vem sendo, também, utilizada como ferramenta para as aplicações em processamento facial [Abd95]. Neste capítulo serão apresentados alguns exemplos de aplicações voltadas para a tarefa de detecção e reconhecimento facial.

3.1 - Detecção de Face baseada no Aprendizado²

Neste trabalho, Sung [Sun98] considerou a detecção facial como um problema de aprendizado para reconhecer padrões faciais de exemplos. Assim, uma possível solução encontrada foi a construção de um modelo de face genérico baseado na distribuição com todas as possíveis variações de padrões. Um procedimento de decisão foi treinado para, empiricamente, descobrir um conjunto de parâmetros e limiares que diferenciem padrões “faciais” e “não-faciais” sobre uma seqüência de exemplos.

Uma vez definido um modelo de face genérica baseada na distribuição dos padrões, modelo baseado na distribuição, o sistema realiza a detecção facial da seguinte maneira (Figura 3.1):

- Realiza-se um pré-processamento da janela de padrões a ser analisada;
- Calcula-se um conjunto de “medidas de diferença” entre o modelo de face canônica e a janela de padrões analisada; e

² Esta técnica consiste no trabalho apresentado por Sung [Sun98]

- Apresenta-se o conjunto de “medidas de diferença” para um classificador (desenvolvido com redes neurais multicamadas) que determina se a janela de padrões contém ou não uma face.

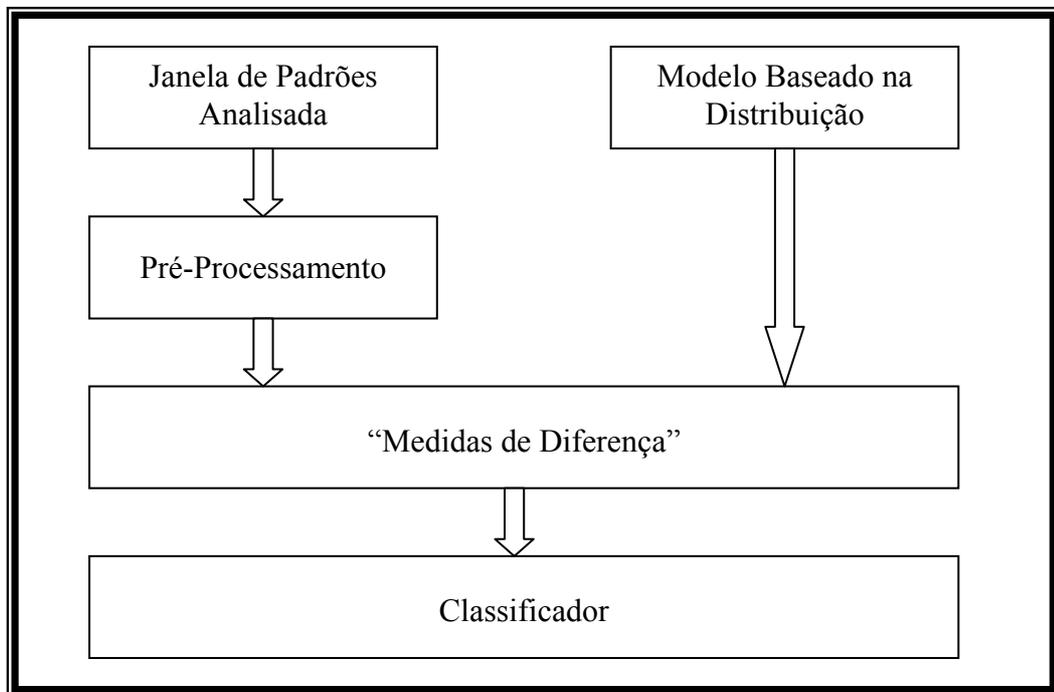


Figura 3.1 – Módulos do sistema para a detecção facial

A seguir serão apresentadas cada uma das fases acima mencionadas:

3.1.1 - Pré-Processamento: é um procedimento utilizado para normalizar a base de imagens faciais e não-faciais utilizada para definir o modelo de face canônica. Este procedimento pode ser descrito pelos seguintes passos:

- redimensionamento da imagem: que altera o tamanho da imagem (redimensiona) para 19x19 pixels;
- máscara: onde uma máscara binária é utilizada para assegurar que não seja inserida nenhuma estrutura incorreta para a representação da face (Figura

3.2), como por exemplo, uma estrutura pertencente ao “background” da imagem;

- correção do gradiente de iluminação: que consiste em subtrair um plano de iluminação da janela de padrões, antes de se aplicar a máscara; e
- equalização do histograma: utilizado para melhorar o contraste da imagem.

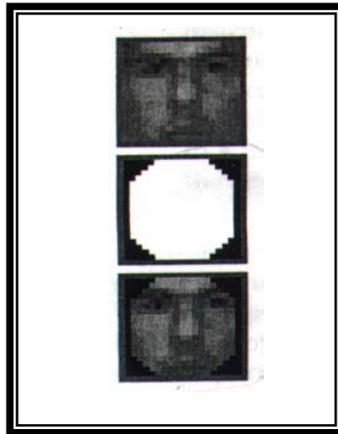


Figura 3.2 – Exemplo da aplicação da máscara

3.1.2 - Modelo Baseado na Distribuição: a detecção de faces é realizada comparando-se cada janela de padrões analisada a um modelo canônico. Como uma das dificuldades está em prestar conta da vasta variação possível em uma imagem, o trabalho se dirige a este problema usando um esquema de modelagem baseada na distribuição para representar o conjunto de todos os padrões que são observados na face canônica. Este esquema trata de todos os pixels da imagem como um vetor de espaço que é dimensionalmente igual ao número de pixels da imagem.

Como não é possível obter todos os padrões *faciais* através de uma única face canônica, o autor utilizou uma base de dados contendo padrões de face frontal (19x19 pixels) para se obter uma representação segura de uma face genérica. Também utilizou uma base de dados com padrões *não-faciais* com o intuito de se refinar o limite entre padrões *faciais* e *não-faciais*.

Sung[Sun98] distribuiu os padrões *faciais* e *não-faciais* através de 6 clusters cada (6 *faciais* e 6 *não-faciais*).

Devido ao alto número de padrões de faces frontais normalizados contidos na base de dados (4.150) para sintetizar 6 clusters de padrões de faces referentes à modelagem da distribuição dos padrões faciais, foi utilizada uma versão modificada do algoritmo K-médias [Sun98] para calcular o centro dos clusters e sua matriz de covariância.

O algoritmo K-médias modificado [Sun98] usa uma variável adaptativa da distância métrica de *Mahalanobis* [Dud73] normalizada para particionar a amostra de dados em clusters.

De um modo geral, pode-se dizer que o algoritmo do K-médias modificado difere do original, devido a um *loop* adicional que tem o objetivo de refinar e recuperar a forma do cluster para se ter melhor representação da face.

A modelagem da distribuição dos padrões *não-faciais* foi realizada a fim de reduzir o erro de classificação. Isto se mostrou necessário pelo fato da face ter sido representada com um pequeno número de clusters Gaussianos, tornando possível que alguns padrões *pseudo-faces* estivessem localizados mais próximos aos centróides dos clusters faciais do que alguns padrões de faces reais. Portanto, os clusters *não-faciais* esculpem a região negativa em volta dos clusters da face que não correspondem aos padrões da face.

De maneira análoga ao anterior, nesta modelagem também foi utilizado o algoritmo de agrupamento de clusters K-médias modificado para se obter as seis posições centrais e suas matrizes de covariância em uma base de dados de 6.189 padrões *pseudo-faces* normalizados. A diferença existente é que a base de dados foi incrementalmente gerada no modelo *bootstrap*, que realiza o seguinte procedimento: inicia um pequeno conjunto de exemplos *não-faciais* no treinamento da base de dados; treina o classificador com a corrente base de dados de exemplo; percorre sobre uma seqüência randômica de imagens e coleta todos os padrões *não-face* que o sistema classificar erroneamente como *face* e adiciona-os para o treinamento da base de dados como um novo exemplo negativo.

Uma vez realizadas ambas as modelagens, *facial* e *não-facial*, o resultado será um modelo de face genérico baseado na distribuição, conforme demonstrado na Figura 3.3.

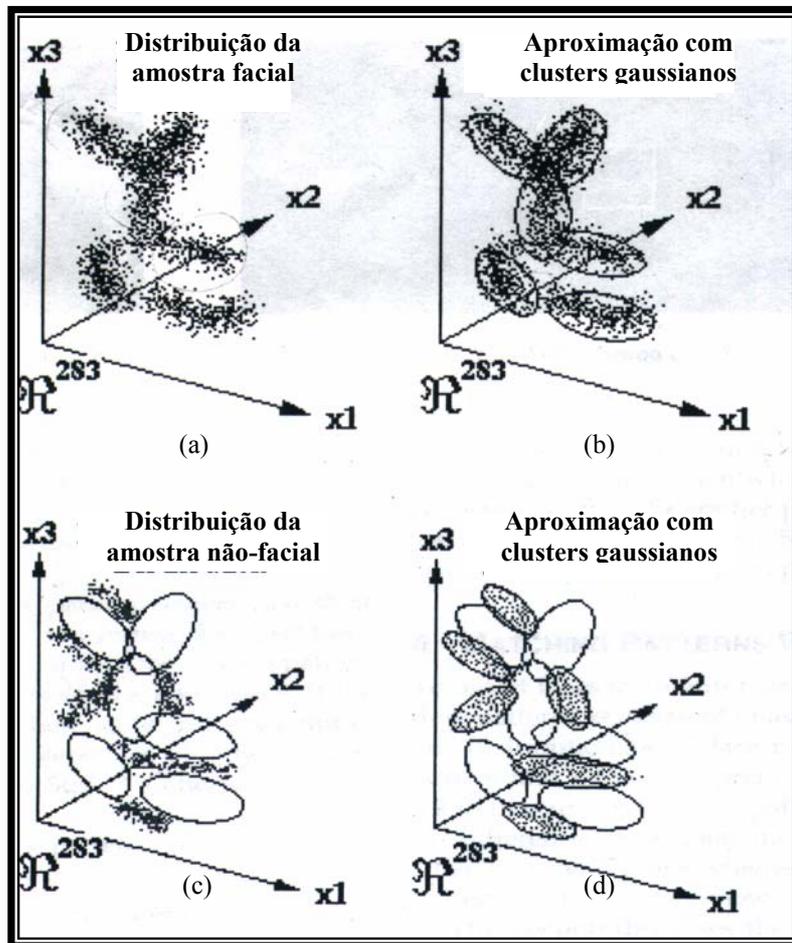


Figura 3.3 – Modelo de face genérica baseada na distribuição – (a) Distribuição da amostra *facial*; (b) Aproximação com clusters gaussianos; (c) Distribuição da amostra *não-facial*, sobreposta aos contornos dos clusters *faciais*; e (d) Aproximação com clusters gaussianos.

3.1.3 - “Medidas de Diferença” – Padrões x Modelo: após ter sido desenvolvido o modelo genérico de face baseado na distribuição, para detectar faces em uma imagem o sistema de Sung[Sun98] compara janelas de padrões da imagem contra o modelo de face baseado na distribuição. Cada comparação retorna o conjunto de “medidas de diferença” que são utilizadas para treinar o classificador.

Cada conjunto de medidas é um vetor que armazena as 12 distâncias calculadas entre padrão de teste e os 12 centróides (Figura 3.4).

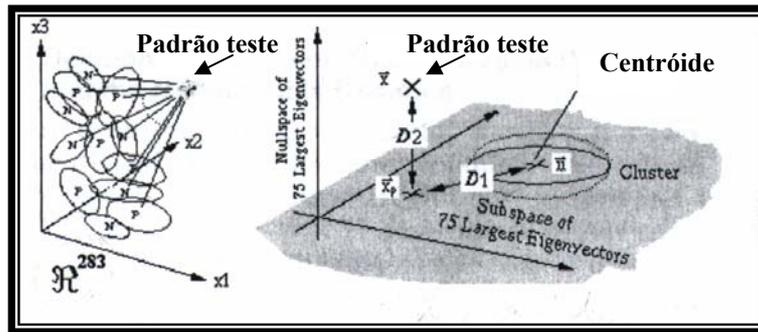


Figura 3.4 – Medidas de diferença entre o padrão de teste e os centróides.

O conjunto de todas as 12 distâncias pode ser visualizado como uma “diferença” bruta entre o padrão de teste e a completa classe de padrões da “face canônica” (modelo baseado na distribuição).

A “medida de diferença” consiste de duas componentes denominadas de “distância de 2-valores”, conforme demonstrado na Figura 3.4. Como pode ser verificado, a primeira componente, $D1$, é a distância de *Mahalanobis* normalizada entre o padrão de teste e o centróide do cluster, em um subespaço baixo dimensional explicado pelos maiores 75 autovetores do cluster. Esta componente de “distância” está direcionada para refletir a localização dos padrões de teste relativo para principal direção de alongamento da distribuição local dos dados. A segunda componente é uma distância Euclidiana normalizada entre o padrão de teste e sua projeção no subespaço baixo dimensional. Esta é uma componente de distância que está direcionada para as diferenças dos padrões não inclusos na primeira componente devido à possível inexatidão da modelagem.

3.1.4 - O Classificador: desenvolvido através de um perceptron multicamada (MLP), ele classifica as janelas de padrões como *faciais* e *não-faciais* baseado nas “medidas de diferença”. A rede tem 12 pares de unidades de entrada, 1 unidade de saída e 24 unidades ocultas. Cada par de entrada recebe a “distância de 2-valores” para um centróide específico. A saída retorna 1 se o vetor de distâncias fornecido para a entrada pertencer a um padrão de face e 0 caso contrário.

Sung[Sun98] treinou a rede sobre as características dos vetores de distância utilizando-se o algoritmo de aprendizado *backpropagation*.

Para obtenção dos resultados foram adotados os seguintes critérios:

- 2 bases de dados, sendo a 1ª composta por imagens digitalizadas de alta qualidade e a 2ª composta por imagens de alta e baixa qualidade;
- 2 arquiteturas para o classificador:
 - perceptron multicamada (MLP) e o
 - perceptron linear (composto por uma única camada);
- 4 modelos de distância métrica
 - Distância Métrica de 2-Valores,
 - Distância de Mahalanobis Normalizada (D1 na Distância Métrica de 2-Valores),
 - Distância Euclidiana (D2 na Distância Métrica de 2-Valores),
 - Distância de Mahalanobis Normalizada Padrão;
- 2 diferentes composições de protótipos para o modelo baseado na distribuição
 - um sistema com 12 clusters de *face* e *não face*,
 - um sistema com 12 clusters de *face*;
 - um sistema com somente 6 clusters de *face*.

Os resultados obtidos estão apresentados nas Tabelas 3.1 e 3.2.

Tabela 3.1 - Proporção de Detecção x Distância Métricas

Arquitetura do Classificador		Distância Métrica			
		2-Valores	D1	D2	Mahalanobis
MLP	1ª Bas	96.3%	91.6%	91.4%	84.1%
	2ª Bas	79.9%	85.1%	65.1%	42.6%
Perceptron	1ª Base	96.7%	93.3%	92.3%	93.0%
	2ª Base	84.6%	85.1%	68.2%	58.6%

Tabela 3.2 - Proporção de Detecção x Composição do Protótipo

Arquitetura do Classificador	Composição do Protótipo		
	6 F & 6 NF	12 Face	6 Face
Multicamada	96.3%	85.3%	59.7%
	79.9%	69.6%	60.9%
Uma camada	96.7%	52.1%	66.6%
	84.6%	49.7%	55.4%

Conforme pode ser verificado, a rede desenvolvida por Sung[Sun98] apresenta melhor desempenho quando é usada uma base de imagens de alta qualidade (1ª base), a distância de 2-valores como padrões de entrada e a composição do protótipo considerando clusters faciais e não-faciais. Os resultados apresentados na Tabela 3.2 foram obtidos utilizando a 1ª base de imagens e a distância de 2-valores como padrões de entrada.

3.2 - Detecção de Faces baseada em Redes Neurais³

Pode-se dizer que este algoritmo trabalha aplicando um conjunto de filtros baseados em redes neurais sobre as janelas de padrões. Como o resultado pode ser composto por um conjunto de detecções sobrepostas e de falsas detecções, foram desenvolvidas duas técnicas direcionadas a estes problemas.

A primeira componente é um filtro baseado em redes neurais que recebe uma imagem de entrada e gera uma saída variando de 1 a -1 , significando a presença ou ausência de uma face. Conforme pode ser visualizado na Figura 3.5, uma fase de pré-processamento antecede a entrada da janela de padrões na rede, sendo composta pela correção da iluminação e equalização do histograma. A imagem original é decrementada em um fator x , para que as janelas de padrões extraídas para alimentar a rede possam ser compostas por combinações diferentes, ou seja, no exemplo apresentado abaixo, os padrões dos componentes olhos, nariz e boca da face maior serão extraídos em uma única janela quando a imagem original estiver totalmente decrementada.

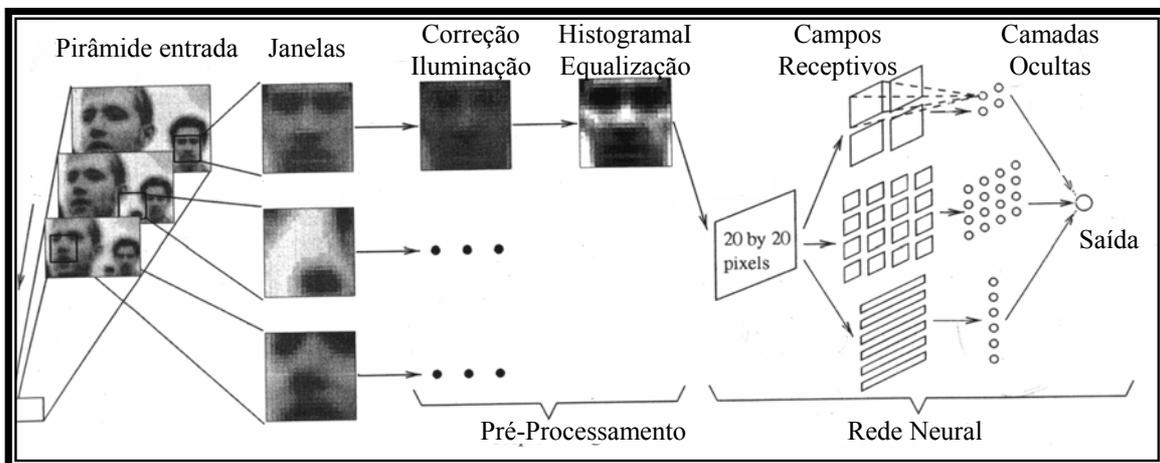


Figura 3.5 – Módulos do sistema para detecção facial

Após realizado o pré-processamento, a janela é apresentada para a camada de entrada da rede neural e propagada por três tipos de camadas ocultas:

³ Esta técnica consiste no trabalho apresentado por Rowley [Row98].

-
- 04 que “olham” para as subregiões 10x10 pixels;
 - 16 que “olham” para as subregiões 5x5 pixels; e
 - 06 que “olham” para linhas horizontais 20x5 pixels.

Cada um destes tipos foram escolhidos para permitir que as unidades ocultas detectem as características locais que deva ser importantes para a detecção facial. Em particular, as linhas horizontais permitem que as unidades ocultas detectem características como bocas ou par de olhos, enquanto que as demais unidades permitem a detecção de características como olhos individuais, nariz ou o contorno da boca. Embora a Figura 3.5 mostre uma única unidade oculta, pode-se utilizar um número maior de unidades.

Para treinar a rede neural para atuar como um filtro, foram utilizados por volta de 1050 exemplos de faces em diferentes tamanhos, orientações, posições e intensidades, sendo os olhos, nariz, contorno e centro da boca de cada face rotulados manualmente. Utilizou-se ainda um procedimento para normalizar cada face, utilizada para o treinamento, para a mesma escala, orientação e posição. Com o objetivo de completar o número de faces utilizadas, foram gerados 15 exemplos de cada imagem original, aplicando-se a translação, rotação, escala e borramento.

Já para os exemplos *não-faciais*, devido à dificuldade de se reunir um conjunto de padrões *não-faciais* significantes, optou-se pelo método *bootstrap*, apresentado no estudo de caso anterior.

A Figura 3.6 demonstra um exemplo de imagem *não-facial* apresentada para a rede e as falsas detecções utilizadas como novos exemplos de imagens *não-faciais* no treinamento (conforme o método *bootstrap*).

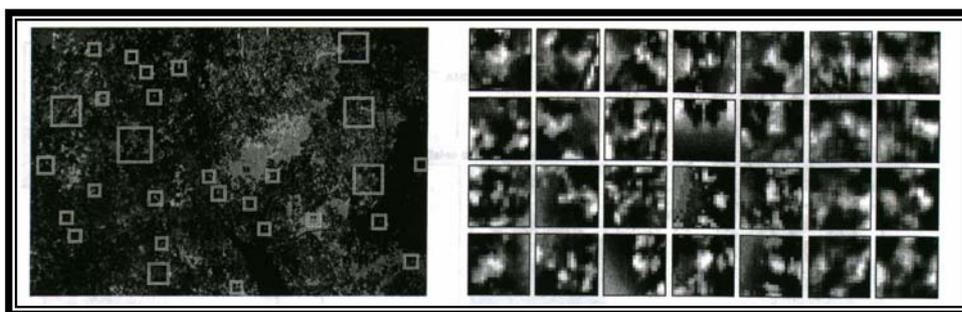


Figura 3.6 – Esquerda: exemplo de imagem “não-facial”; Direita: exemplo de falsas detecções.

Outro tipo de exemplo é apresentado pela figura abaixo, onde pode-se notar nitidamente as sobreposições e as falsas detecções.

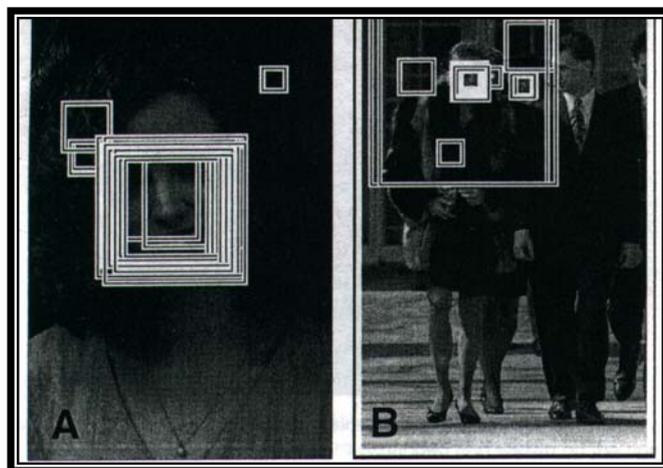


Figura 3.7 – Exemplo de sobreposições e falsas detecções.

Através da observação realizada nos resultados apresentados pela rede, pode-se notar que as faces são detectadas através de múltiplas posições ou escalas vizinhas, enquanto que as falsas detecções ocorrem com menor consistência. Esta observação conduz para uma heurística que consiste em unir as detecções sobrepostas podendo eliminar muitas detecções falsas. A implementação desta heurística pode ser visualizada pela Figura 3.8.

Cada detecção em uma particular localização e escala é registrada em uma pirâmide de saída, calculando-se o ponto central de cada detecção. Então cada ponto passa por um processo de *spreading out*, que consiste na expansão dos pontos. Um limiar que determina um número mínimo de detecções a ser considerado como face é aplicado para estes valores, calculando-se em seguida, os centróides de todas as regiões consideradas como face pela limiarização. Cada centróide é então examinado, iniciando-se daquele que tiver o maior número de detecções dentro de uma vizinhança especificada. Se nenhuma outra localização de uma face representada por um centróide sobrepuser o centróide corrente, então o centróide corrente é removido da pirâmide de saída e todos os demais centróides de localização da face restante constituem o resultado final de detecção. Caso contrário, o centróide corrente representa o resultado final.

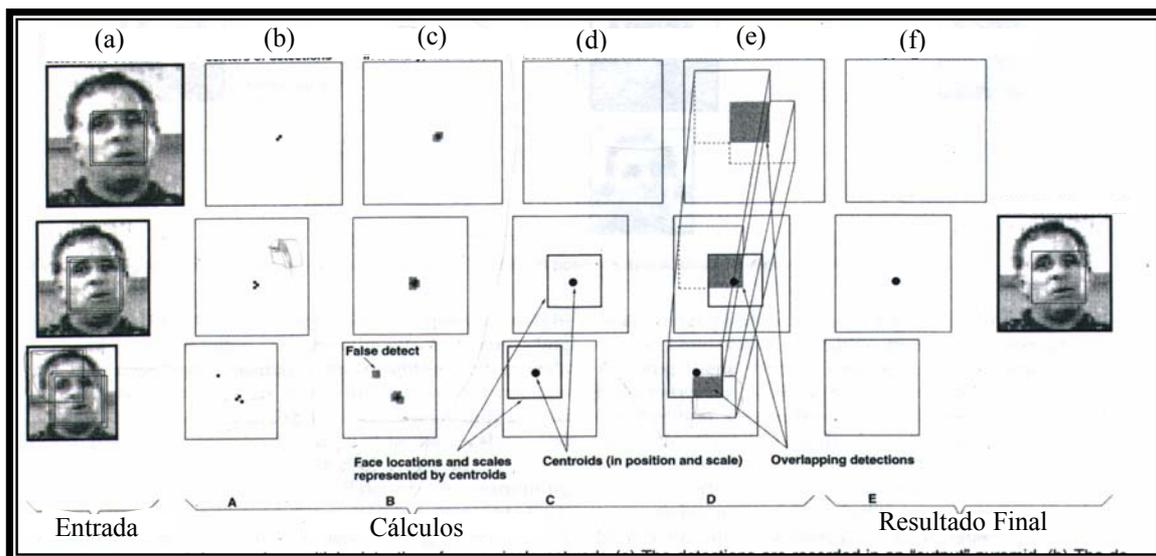


Figura 3.8 – Heurística para detecções sobrepostas – (a) Pirâmide de entrada com detecções sobrepostas; (b) Detecção dos centros; (c) “Spreading Out” das detecções; (d) Detecção do centróide; (e) Avaliação dos centróides; (f) Resultado final.

Outra técnica de reduzir o número de falsas detecções é a aplicação de múltiplas redes, com a arbitração de suas saídas para produzir uma decisão final.

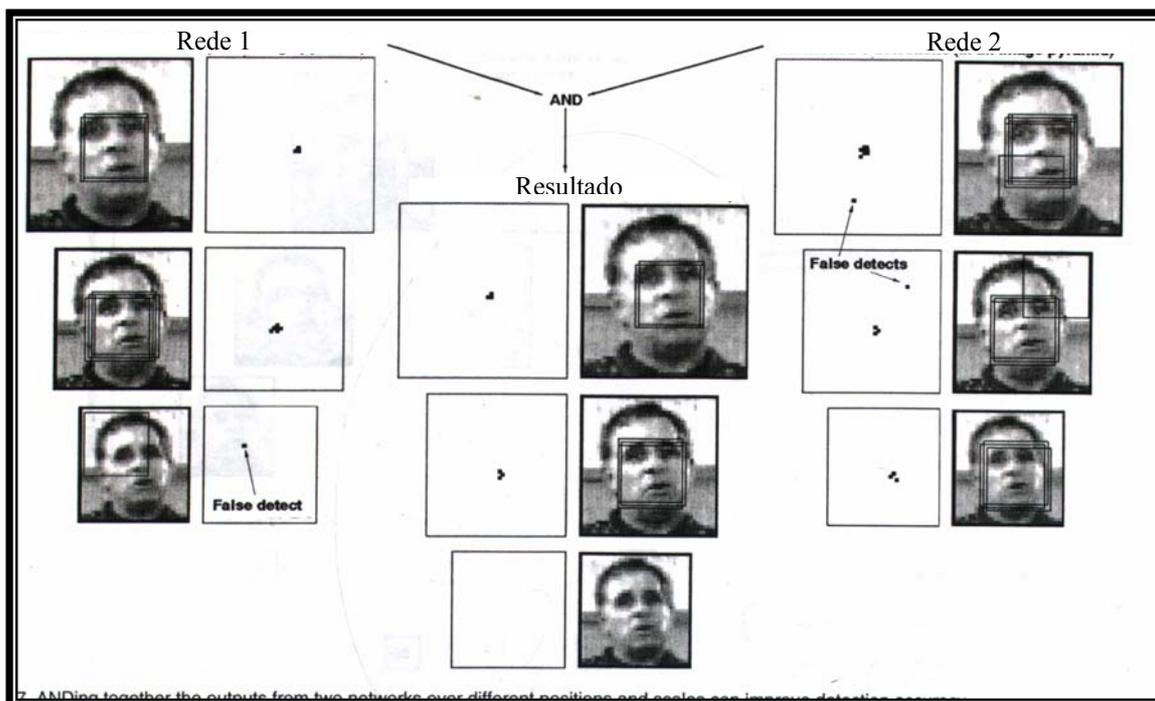


Figura 3.9 – Arbitração entre múltiplas redes para eliminar detecções sobrepostas

Na arbitração, cada detecção em uma particular localização e escala é registrada em uma pirâmide de saída (como foi feito na heurística anterior). No caso do modelo apresentado na Figura 3.9, o caminho utilizado para combinar as duas pirâmides foi através do operador *AND*. Portanto, esta estratégia transmite uma detecção somente se ambas as redes detectarem uma face com exatamente a mesma escala e posição. Isto permite eliminar muitas detecções falsas, já que as redes individuais raramente combinam sobre uma falsa detecção de face.

Esta estratégia de arbitração pode ser realizada com outros tipos de operadores como o *OR* e o *voting* (votação), podendo ser aplicada antes ou depois das heurística anterior.

Para a realização dos testes são considerados dois grandes conjuntos de imagem. O primeiro conjunto é constituído de 130 imagens obtidas de diversas maneiras, por exemplo de jornais e fotografias, compostas por um total de 507 faces frontais. O segundo conjunto é composto por imagens que possuem boa iluminação e uma única face.

Em outras palavras, o primeiro conjunto foi adotado para analisar a proporção de análises falsas do sistema, enquanto que o segundo conjunto foi adotado para analisar a sensibilidade angular do sistema.

As Tabelas apresentadas a seguir demonstram os resultados obtidos com a rede. Os resultados estão organizados em quatro categorias: (a) única rede sem heurística; (b) única rede com heurística (exemplo apresentado na Figura 3.8); (c) arbitrando entre duas redes (exemplo apresentado na Figura 3.9); e (d) arbitrando entre três redes (exemplo apresentado na Figura 3.9 substituindo os operadores). Conforme pode ser verificado, para a 1ª base de imagens (Tabela 3.3) o uso da arbitração entre duas redes e três redes apresentou um número menor de falsas detecções. No entanto, para a 2ª base de imagens (Tabela 3.4) obteve-se um resultado melhor através da arbitração entre duas redes.

Tabela 3.3 – Proporção de Erros e Detecções para o Conjunto de Teste I

Tipo	Sistema	Falsas Detecções	Proporção de Falsas Detecções
Única Rede sem Heurística	Rede 1	945	1/87935
	Rede 2	862	1/96402
	Rede 3	738	1/112600
	Rede 4	819	1/101464
Única Rede com Heurística	Rede 1 -> limiar (2,1)	570	1/145788
	Rede 2 -> limiar (2,1)	506	1/164227
	Rede 3 -> limiar (2,1)	440	1/188861
	Rede 4 -> limiar (2,1)	484	1/171692
Arbitrando entre 2 redes	Redes 1 e 2 -> AND (0)	79	1/1051888
	Redes 1 e 2 -> AND (0) -> limiar (2,3)	2	1/41549605
	Redes 1 e 2 -> limiar (2,2) -> AND (2)	23	1/3613009
	Redes 1 e 2 -> limiar (2,2) -> OR (2) -> limiar (2,1)	185	1/449184
Arbitrando entre 3 redes	Redes 1,2 e 3 -> voting (0)	99	1/839385
	Redes 1,2 e 3 -> arbitração da rede (5 unidades ocultas) -> limiar(2,1)	16	1/5193700
	Redes 1,2 e 3 -> arbitração da rede (10 unidades ocultas) -> limiar (2,1)	10	1/8309921
	Redes 1,2 e 3 -> arbitração da rede (perceptron) -> limiar (2,1)	2	1/6924934

Tabela 3.4 – Número de Falsas Detecções para o Conjunto de Teste II

		Faces Frontais	Ângulo 15°	Ângulo 22,5°
Número de Imagens:		1001	241	378
Número de Faces:		1001	241	378
Número de Janelas:		255129875	61424875	96342750
Tipo	Sistema	1	1	1
Única Rede sem Heurística	Rede 1	1747	447	819
	Rede 2	1457	481	592
	Rede 3	1242	374	605
	Rede 4	1665	458	709
Única Rede com Heurística	Rede 1 -> limiar (2,1)	643	136	263
	Rede 2 -> limiar (2,1)	458	118	146
	Rede 3 -> limiar (2,1)	421	85	133
	Rede 4 -> limiar (2,1)	563	120	207
Arbitrando entre 2 redes	Redes 1 e 2 -> AND (0)	141	46	85
	Redes 1 e 2 -> AND (0) -> limiar (2,3)	0	0	1
	Redes 1 e 2 -> limiar (2,2) -> AND(2)	12	3	3
	Redes 1 e 2 -> limiar (2,2) -> OR (2) -> limiar (2,1)	137	35	53
Arbitrando entre 3 redes	Redes 1,2 e 3 -> voting (0)	74	23	38
	Redes 1,2 e 3 -> arbitração da rede (5 unidades ocultas) -> limiar(2,1)	5	1	3
	Redes 1,2 e 3 -> arbitração da rede (10 unidades ocultas) -> limiar (2,1)	4	1	2
	Redes 1,2 e 3 -> arbitração da rede (perceptron) -> limiar (2,1)	3	1	2

3.3 – Aplicação da PDBNN para o Reconhecimento Facial⁴

Tratando-se de uma rede probabilística, o reconhecimento é realizado com base sobre as técnicas de reconhecimento biométrico. A principal justificativa para o uso da técnica biométrica encontra-se no avanço tecnológico da microeletrônica e sistemas de visão, tornando-os economicamente praticáveis.

O PDBNN apresentada por Shang[Sha97] para o reconhecimento facial consiste de três módulos principais:

- detector de face: localiza uma face humana na imagem inserida;
- localizador de olhos: determina a posição de ambos os olhos para gerar um vetor significativo de características; e
- reconhecedor de face.

Como se trata de um modelo hierárquico, os módulos de processamento são executados sequencialmente. Pode-se definir, de modo geral, o funcionamento da rede da seguinte maneira: depois de se obter uma imagem através de um sistema de aquisição, é aplicada uma PDBNN para realizar a detecção de uma face. Caso positivo, uma outra PDBNN é ativada para localizar ambos os olhos. Em seguida, uma sub-imagem correspondente para a região facial (composta pelas sobrancelhas, olhos e nariz) é extraída. Finalmente, o vetor de característica, referente a região facial obtida, é apresentada a uma PDBNN para o reconhecimento facial.

A PDBNN para a tarefa de detecção da face e localização dos olhos foi apresentada no capítulo II. Já a PDBNN para a tarefa de reconhecimento facial pode ser considerada como uma extensão da PDBNN para a detecção de faces (Figura 2.9), onde para K -pessoas existem K diferentes subredes (Figura 3.11).

⁴ Esta técnica consiste no trabalho apresentado por Shang [Sha97].

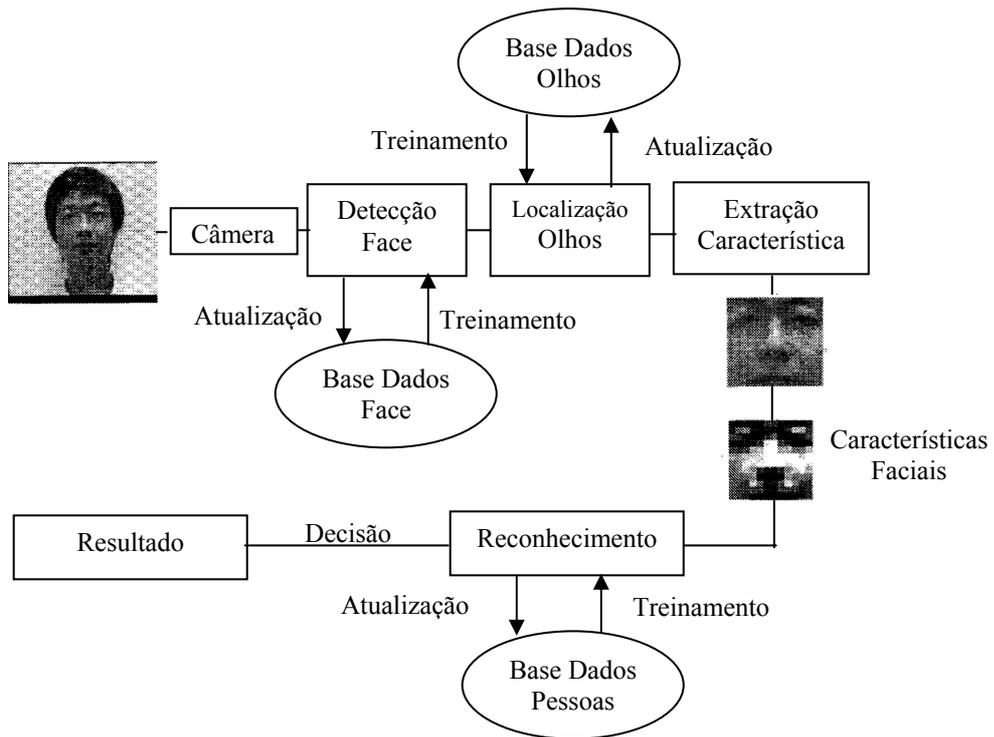


Figura 3.10 – Módulo “A” do sistema para reconhecimento facial.

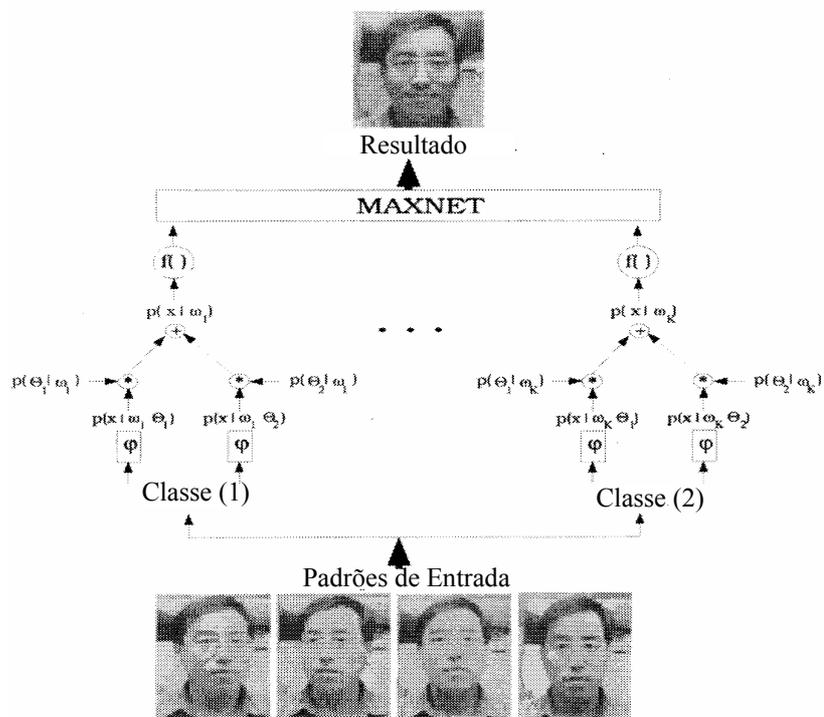


Figura 3.11 – PDBNN para a tarefa de reconhecimento facial

Os testes foram realizados sobre uma base de dados contendo imagens de 66 pessoas, com 5 diferentes tipos de orientações, 4 condições de iluminação e 3 tipos de expressões (sorriso, olhos carrancudos e surpresa; cada qual contendo 2 imagens), conforme exemplificado na Figura 3.12.

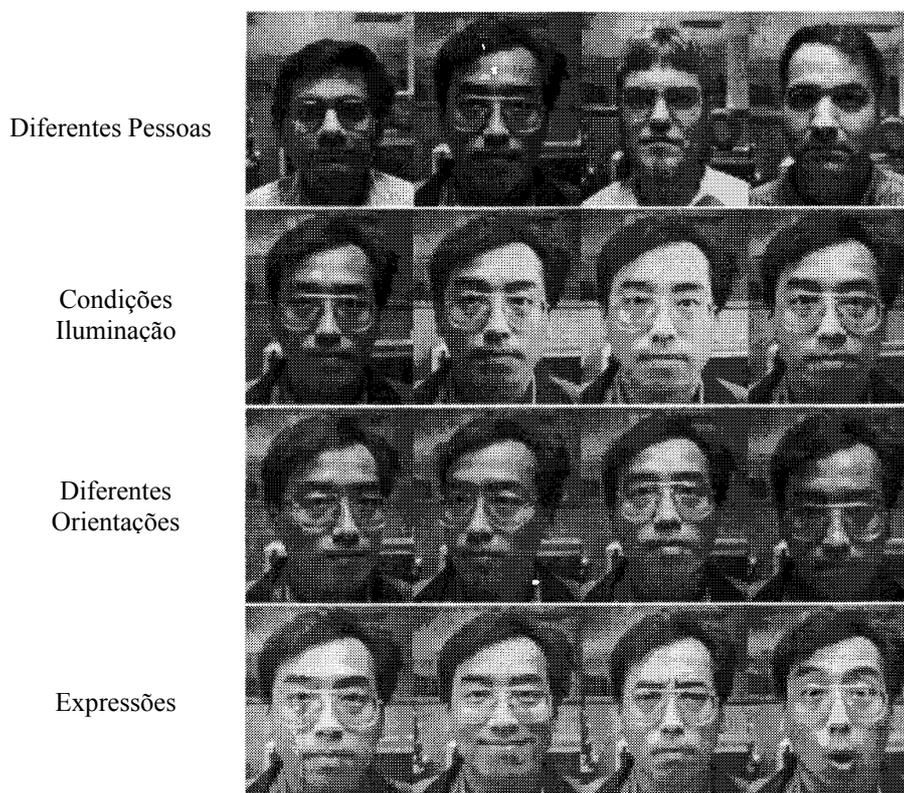


Figura 3.12 – Diferentes tipos de imagens utilizadas nos testes.

Aplicando-se o PDBNN para o reconhecimento facial sobre a base de dados apresentada, notou-se um bom resultado para as variações de orientações e expressões, mas não para as variações de iluminações. Os resultados obtidos são mostrados na Tabela 3.5.

Tabela 3.5 – Resultados obtidos através do módulo “A” (A Falsa Aceitação foi computada usando padrões de uma base de dados desconhecida).

Resultados	
Reconhecimento	92,65%
Falsa Rejeição	7,29%
Má-Classificação	0,06%
Falsa Aceitação	9,35%

Uma solução apresentada pelo mesmo autor para o problema de iluminação, foi a inserção de mais um módulo na estrutura do sistema: o verificador de faces (Figura 3.13). Sua função é aceitar ou rejeitar a decisão do reconhecimento “primário”, tendo como entrada uma sub-imagem pertencente a região da “testa”. Isto se deve ao fato desta região ser mais homogênea do que a região dos olhos e nariz, permitindo a normalização dos efeitos de iluminação sobre esta área de maneira mais fácil e eficaz. Os resultados obtidos são mostrados na Tabela 3.6.

Tabela 3.6 – Resultados obtidos através do módulo “B”(A Falsa Aceitação foi computada usando padrões de uma base de dados desconhecida).

	Resultados
Reconhecimento	97,75%
Falsa Rejeição	2,25%
Má-Classificação	0%
Falsa Aceitação	0%

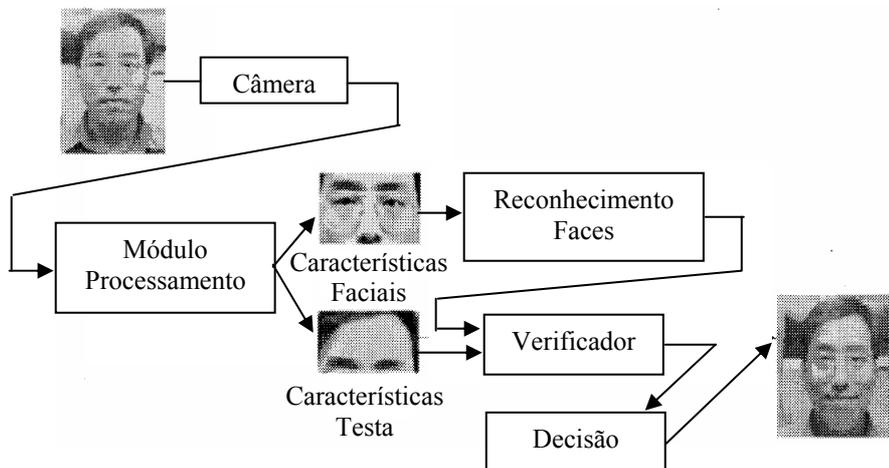


Figura 3.13 – Módulo “B” do sistema para reconhecimento facial.

No próximo capítulo serão apresentadas a arquitetura do neocognitron e suas aplicações.

Capítulo IV

Neocognitron e suas Aplicações

Neste capítulo será apresentada a arquitetura de redes neurais desenvolvida por Fukushima [Fuk95] originalmente para o reconhecimento de caracteres manuscritos: o neocognitron. Trata-se de uma arquitetura de redes neurais baseada no modelo biológico de visão. Serão também apresentadas arquiteturas de redes neurais que são alterações do neocognitron, e suas aplicações.

4.1 - Neocognitron

O neocognitron é uma rede neural composta de várias camadas de células neurais, organizadas matricialmente, inspirada no modelo biológico de visão de Hubel e Wiesel [Fuk95]. Como em outros modelos de redes neurais, no neocognitron é possível identificar uma fase de treinamento não-supervisionado, e uma fase de reconhecimento.

A estrutura da rede, consiste de vários estágios sucessivos, cada qual com dois tipos principais de camadas de células: U_s de células-S, responsáveis pela extração dos fatores, e U_c de células-C que fazem a generalização dos sinais. O estágio inicial da rede é a camada de entrada, chamada de U_0 (Figura 4.1).

As células-S são variáveis e necessitam de treinamento (aprendizado) para obterem as características dos padrões de entrada necessárias para o reconhecimento, ou seja, as células-S são ativadas somente se as características forem apresentadas na posição correta da camada de entrada, sendo que as características a serem obtidas são definidas durante o processo de treinamento (aprendizado). A Figura 4.2, apresenta abaixo, demonstra um exemplo típico das células-S, onde os elementos $u(1)$, $u(2)$..., $u(N)$, são

denotados de “*entradas excitatórias*”. O elemento “*v*” denota as células inibitórias, utilizadas para auxiliar na obtenção das características.

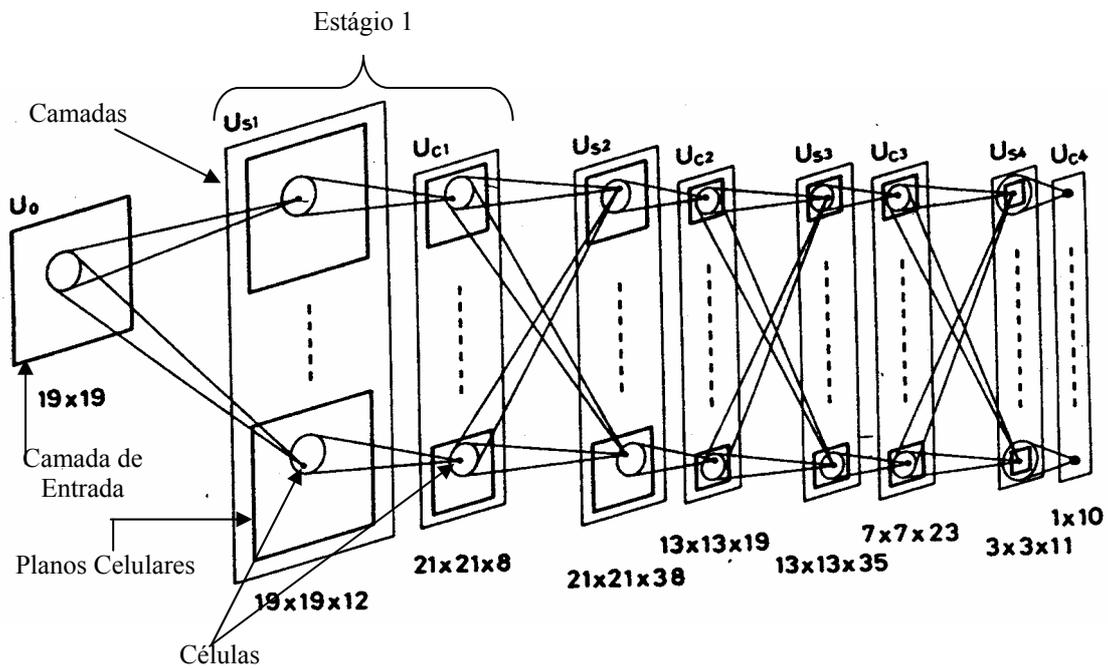


Figura 4.1 – Estrutura hierárquica da Rede Neocognitron.

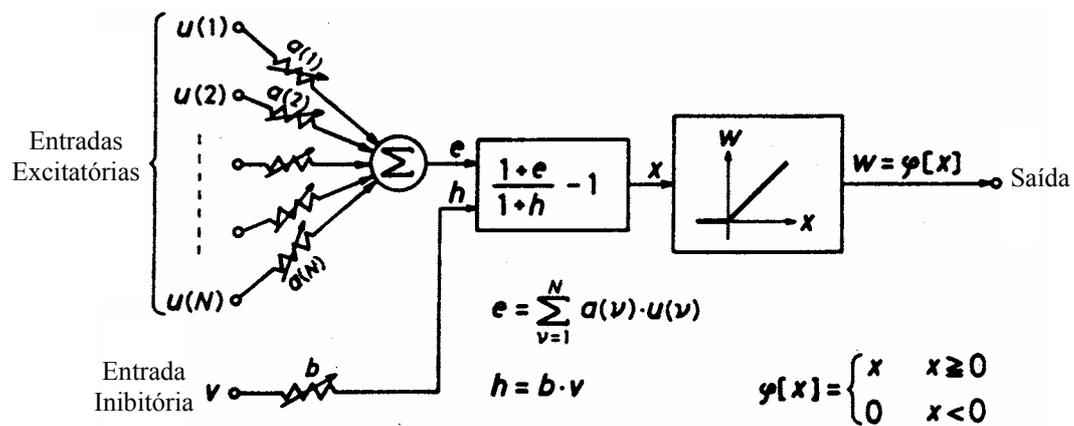


Figura 4.2 – Diagrama de uma célula-S.

Diferente das células-S, as células-C são fixas e invariáveis e servem para se obter valores correspondentes a uma operação de borrimento baseado num grupo de células-S. Essas células são responsáveis pela tolerância à deformação do neocognitron.

Assim, cada um dos i -ésimos estágios da rede é composta de camadas U_{si} que consistem de um número de estruturas matriciais de células- S , seguidas de camadas U_{ci} que consistem de um número de matrizes de células- C . Cada estrutura matricial de células, denominada plano celular, é associada a um único fator extraído dos padrões de teste, durante a fase de treinamento.

Cada célula de um plano-celular recebe conexões de entrada dos planos celulares da camada precedente, conforme apresentado no diagrama da Figura 4.3.

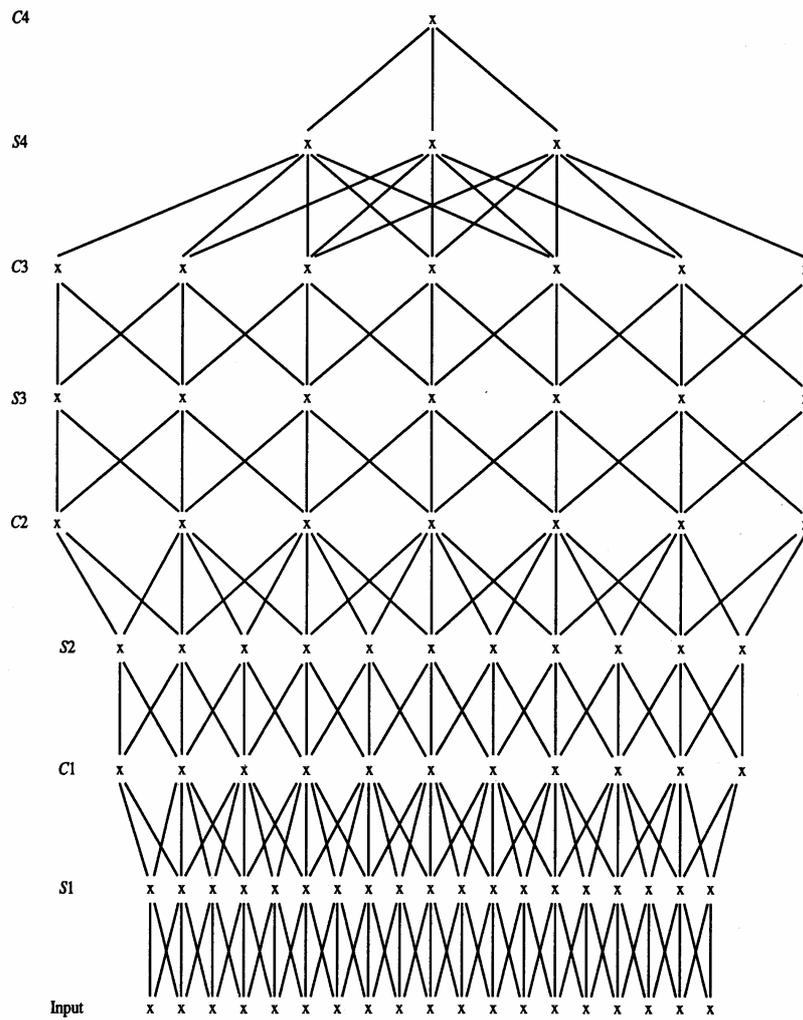


Figura 4.3 – Modelo de abrangência das conexões entre as células dos diferentes planos celulares, através de uma secção da rede neocognitron.

A fase de reconhecimento do neocognitron, segue a sequência mostrada no algoritmo da Figura 4.4, da camada 1 a L , onde L é o número total de estágios.

Program neocognitron ;

```
begin  
  for  $l = 1$  to  $L$  do   computar_estagio ( $l$ );  
end;
```

Figura 4.4 - Sequência de reconhecimento do neocognitron.

Procedure computar_estagio (l) ;

```
begin  
  for  $k = 1$  to  $K_l$  do begin  
    for  $n = 1$  to  $N$  do begin  
      for  $\kappa = 1$  to  $K_{l-1}$  do  
        for all  $v \in S_v$  do begin  
           $e(n,k) := e(n,k) + a(v, \kappa, k) \cdot u_{Cl-1}(n+v, \kappa)$  ;  
           $h(n,k) := h(n,k) + c(v) \cdot \{u_{Cl-1}(\kappa, n+v)\}^2$  ;  
        end;  
           $u_{Sl}(n,k) := (\theta / (1-\theta)) \cdot \varphi((1+e(n,k)) /$   
               $(1+\theta \cdot b(k) \cdot \text{sqrt}(h(n,k))) - 1)$  ;  
        for all  $v \in S_v$  do  
           $u_{Cl}(n,k) := u_{Cl}(n,k) + d(v) \cdot u_{Sl}(n+v, k)$  ;  
           $u_{Cl}(n,k) := \Psi(u_{Cl}(n,k))$   
        end  
      end  
    end  
  end;
```

Figura 4.5 - Algoritmo para computar um estágio.

A Figura 4.5 mostra o algoritmo que computa os valores $u_{Sl}(n, k)$ de células-S, e os valores $u_{Cl}(n, k)$ de células-C, dentro de um estágio l . São computados esses valores nos K_l planos-celulares e em todas as N posições dentro de um plano-celular.

Para o cálculo do valor $u_{SI}(n, k)$ são computados os somatórios $e(n, k)$ e $h(n, k)$ de todas as entradas conectadas aos K_{l-1} planos-celulares da camada precedente, numa dada área de conexão S_v , que circunda a posição da célula n da camada de células- C , do estágio anterior ou da camada de entrada, pela iteração sobre os comandos:

$$\begin{aligned} e(n, k) &:= e(n, k) + a(v, \kappa, k) \cdot u_{Cl-1}(n + v, \kappa), \quad e \\ h(n, k) &:= h(n, k) + c(v) \cdot \{u_{Cl-1}(\kappa, n + v)\}^2 \quad . \end{aligned} \quad (IV)$$

Então o valor $u_{SI}(n, k)$ é obtido pelo comando:

$$u_{SI}(n, k) := (\theta / (1 - \theta)) \cdot \varphi((1 + e(n, k)) / (1 + \theta \cdot b(k) \cdot \text{sqrt}(h(n, k)))) - 1 \quad (V)$$

onde $\varphi(x) = x$, quando $x > 0$, e $\varphi(x) = 0$, caso contrário. A variável θ representa o limiar da função, cujos valores ficam entre 0 e 1, e $b(k)$ representa o coeficiente de inibição.

Para se obter o valor $u_{Cl}(n, k)$, é computado primeiramente o somatório de todas as entradas correspondentes aos valores $u_{SI}(n, k)$, previamente obtidos em uma área de conexão S_v que circunda a posição n da camada de células- S precedente, pela iteração sobre o seguinte comando:

$$u_{Cl}(n, k) := u_{Cl}(n, k) + d(v) \cdot u_{SI}(n + v, k) \quad (VI)$$

seguida do cálculo da função de transferência $\Psi(x) = \varphi(x) / (1 + \varphi(x))$, que limita a saída das células- C , no intervalo $[0, 1)$.

```

Program treinar_neocognitron();
Begin
  For  $l = 1$  to  $L$  do treinar_estagio( $l$ );
End

```

Figura 4.6 - Sequência de treinamento do neocognitron.

```

procedure treinar_estagio (l);
begin
  repeat
    for k = 1 to  $K_l + 1$  do computar_estagio(l);
    seleccionado := false;
    repeat
      if proximo_vencedor > 0 then begin
        vencedor := proximo_vencedor;
        seleccionado := true;
        for k = 1 to  $K_l$  do
          if  $u_S(\textit{vencedor}, k) > 0$  then
            seleccionado := false;
          end;
        until (seleccionado or proximo_vencedor = 0);
        if seleccionado then begin
          for  $\kappa = 1$  to  $K_{l-1}$  do
            for all  $v \in S$  do
               $a(v, \kappa, k) := a(v, \kappa, k) + q \cdot c(v) \cdot u_{CI-1}(\textit{vencedor} + v, \kappa);$ 
               $b(k) := b(k) + q \cdot \textit{sqrt}(h(\textit{vencedor}, k));$ 
             $K_l := K_l + 1$ 
          end
        until not (seleccionado);
      end
    until not (seleccionado);
end

```

Figura 4.7 - Treinamento de um estágio do neocognitron.

A Figura 4.6 mostra a sequência de treinamento dos estágios do neocognitron, enquanto que a Figura 4.7 mostra o algoritmo de treinamento dentro de um estágio. O treinamento do neocognitron procede computando os valores de células-S para os planos celulares $k = 1$ a $k = K_l$, relacionados aos fatores já existentes, e $k = K_l + 1$, que corresponde ao plano de seleção de fator (*seed selecting plane*). Então, é verificado nesse último plano, plano de seleção, a célula vencedora (*winner*), ou seja, aquela em que o valor de resposta é máximo. Se, na posição da célula vencedora, existe alguma resposta maior que zero nos K_l planos-celulares, o algoritmo segue na procura de um novo vencedor; caso contrário,

cada conexão de entrada da célula vencedora é reforçada proporcionalmente à intensidade da conexão de entrada, através dos comandos:

$$\begin{aligned} a(v, \kappa, k) &:= a(v, \kappa, k) + q \cdot c(v) \cdot u_{CI-1}(vencedor+v, \kappa), \text{ e} \\ b(k) &:= b(k) + q \cdot \text{sqrt}(h(vencedor, k)). \end{aligned} \quad (\text{VII})$$

Dessa forma um novo fator é obtido e um novo plano-celular é adicionado à camada, incrementando K_I de um. Uma vez feito o reforço das conexões de entrada na célula vencedora, essa célula passa a ser a célula-semente, pois todas as demais células do mesmo plano-celular terão os mesmos pesos nas conexões de entrada. Daí surge o nome *Seed-Selecting-Plane*, ou Plano de Seleção de Fatores (PSF), ao plano usado para se obter a célula vencedora, que vem a ser a célula-semente.

O procedimento de treinamento da camada-S, descrito acima, é repetido até que todos os novos fatores sejam detectados, com a apresentação dos padrões de treinamento na camada de entrada.

A Figura 4.8, mostra em diagrama de blocos o treinamento do neocognitron.

4.2 – Aplicações com o Neocognitron

4.2.1 - Reconhecimento de Padrões no Neocognitron provido de uma Adaptação Neural⁵

O propósito do trabalho desenvolvido por Ooyen [Ooy93] consiste na adaptação neural das células-C para melhorar o desempenho no reconhecimento de padrões.

Em consequência do algoritmo de treinamento utilizado no neocognitron, as células-S que respondem para as características alvo têm suas conexões normalmente mais reforçadas que as demais. Isso também ocorrerá para as demais camadas, sendo que as

⁵ Esta técnica consiste no trabalho apresentado por Ooyem [Ooy93].

células nos níveis sucessivos terão suas conexões reforçadas mediante as células precedentes que são frequentemente ativadas e possui um alto nível de atividade.

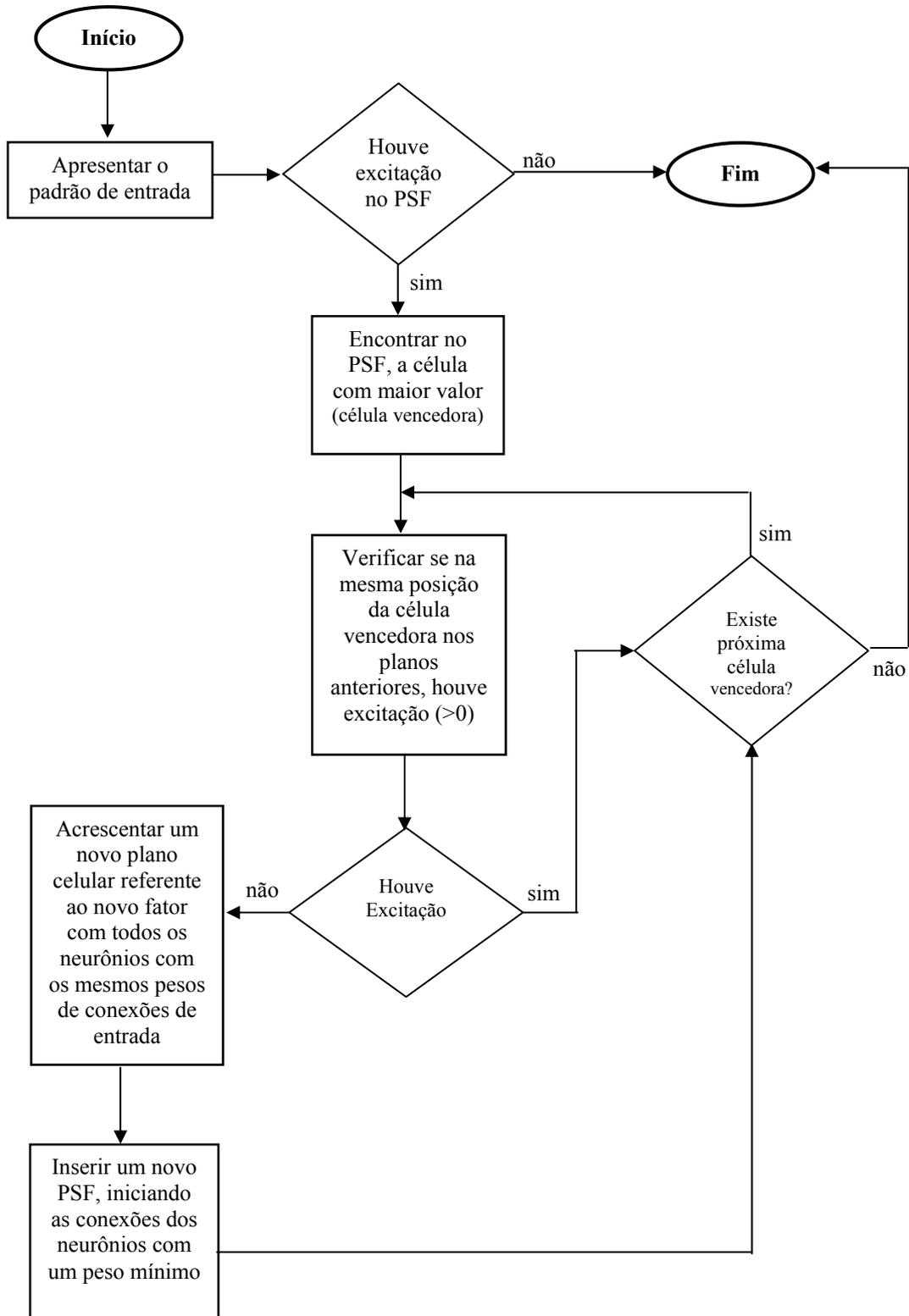


Figura 4.8 – Etapas para o treinamento não-supervisionado no neocognitron

No entanto, pode ocorrer das informações chegarem ao nível mais alto das células-C, camada de saída, sem estarem especificadas o suficiente para o reconhecimento. Por exemplo, o treinamento da rede com conjuntos de padrões semelhantes pode acarretar em um mau aprendizado, uma vez que a propagação pelas camadas ocorrerá quase que de maneira similar.

Neste sentido, este trabalho propõe o uso de neurônios adaptativos. Esse termo é utilizado para descrever o fenômeno onde o neurônio torna-se gradativamente menos sensível quando estimulado frequentemente.

Assim, as respostas das células-C para as características frequentemente ativadas serão mais fracas, devido a adaptação, e a conexão destas células para as células-S sucessivas não serão reforçadas. Já as conexões das células-C que respondem para as características particulares de cada padrão tornam-se reforçadas, em consequência dos valores das saídas destas células serem relativamente altas em decorrência do baixo grau de adaptação.

Em outras palavras, pode-se concluir que ocorrerá o processo inverso do neocognitron convencional. Ao invés de reforçar as conexões que forem altamente ativadas, realizar-se-á o decremento. Como resultado tem-se as conexões que possuem um número baixo de ativação com alto valor de resposta, caracterizando características únicas de cada padrão.

A análise foi realizada sobre ambos os modelos, testando-se os padrões de entrada 8,5,6 e 9 apresentados de maneira sincronizada (por exemplo, 85698569...).

A arquitetura do neocognitron modificado é composta por uma camada de entrada, quatro camadas de células-S e quatro camadas de células-C. Já a arquitetura do neocognitron original consiste de uma camada de entrada, três camadas de células-S e três camadas de células-C.

Conforme pode ser verificado na Figura 4.2, o neocognitron original não conseguiu distinguir entre os padrões 5 e 6. A Figura apresenta o resultado das células das camadas

C, com a U_{C3} sendo a camada de saída da rede. O problema ocorreu na segunda camada onde foram extraídas características similares.

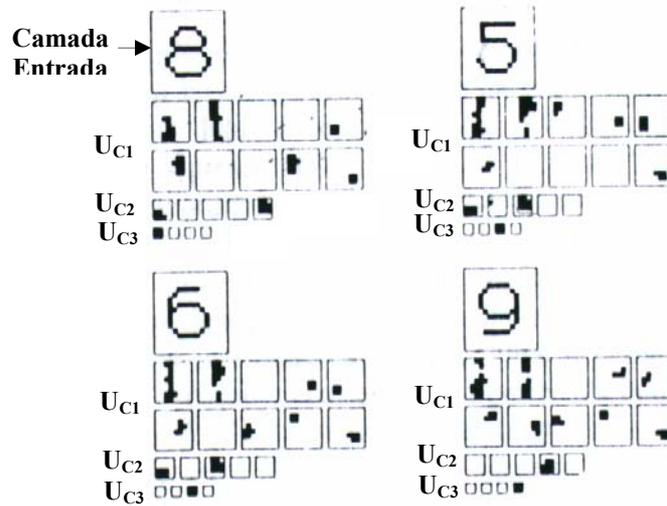


Figura 4.9 – Exemplo de erro na distinção dos padrões dos números 5 e 6, em decorrência da similaridade dos padrões extraídos na camada U_{C2} .

Já o neocognitron modificado apresentou todos os resultados satisfatórios (Figura 4.10).

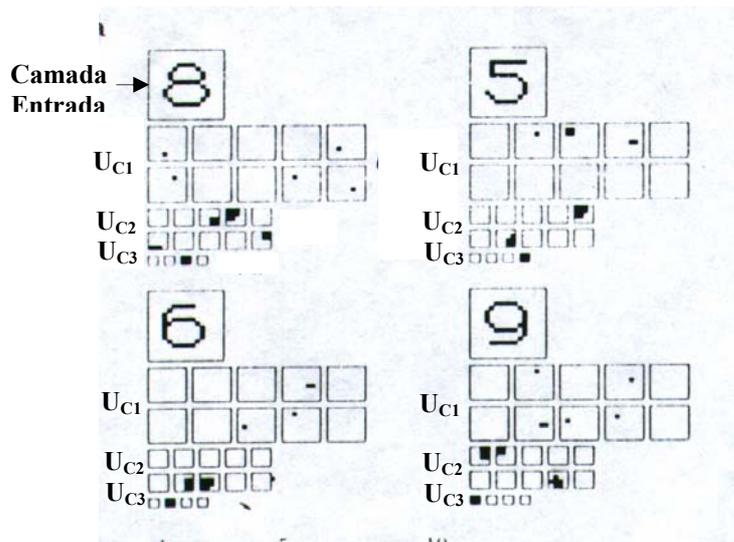


Figura 4.10 – Exemplo de distinção correta dos padrões dos números 5 e 6.

4.2.2 - Uma arquitetura de Rede Neural Convolutacional baseada no Neocognitron para o Reconhecimento de Padrões de Imagens Médicas⁶

Este trabalho consiste da aplicação de uma rede neural convolutacional e vários métodos de treinamento voltados para a detecção de pequenos nódulos pulmonares e microcalcificações, sobre a projeção digital de imagens em raios-X.

Diversas técnicas têm sido propostas para a detecção de nódulos pulmonares (como por exemplo, técnicas de operações morfológicas). No entanto, nota-se que há um compromisso, “*trade-off*”, entre o aumento da sensibilidade e o decremento da especificação, isto é, se a sensibilidade do programa de especificação for relativamente alta, o número de falsas especificações aumentará, caso contrário, acarretará em várias perdas de especificações verdadeiras. Por causa dessas ocorrências é que decidiu-se pela aplicação de um modelo de rede neural.

A estrutura da rede neural convolutacional utilizada no trabalho em questão é uma estrutura simplificada do neocognitron.

De acordo com o autor [Shi95], a motivação pelo uso do neocognitron para o reconhecimento de padrões na área médica se deu pelo fato do mesmo parecer “... ser capaz de incorporar a informação de uma área circundante à área suspeita sobre o seu processo e tem o potencial para tratar das ambigüidades no conjunto de informações” [Shi95].

A arquitetura é composta por duas camadas ocultas. O algoritmo de treinamento é composto por uma fase supervisionada e outra não-supervisionada. Apesar do treinamento não-supervisionado utilizar o mesmo algoritmo do neocognitron para a propagação progressiva, o treinamento supervisionado de uma camada para a próxima foi realizado com o método backpropagation e não com o proposto por Fukushima [Fuk95].

⁶ Esta técnica consiste no trabalho apresentado por Shing [Shi95]

Os pesos utilizados no processamento da rede foram divididos em dois grupos: centro e região circundante. O propósito de utilizar dois grupos é o de treinar a rede para aprender diferentes padrões de imagem.

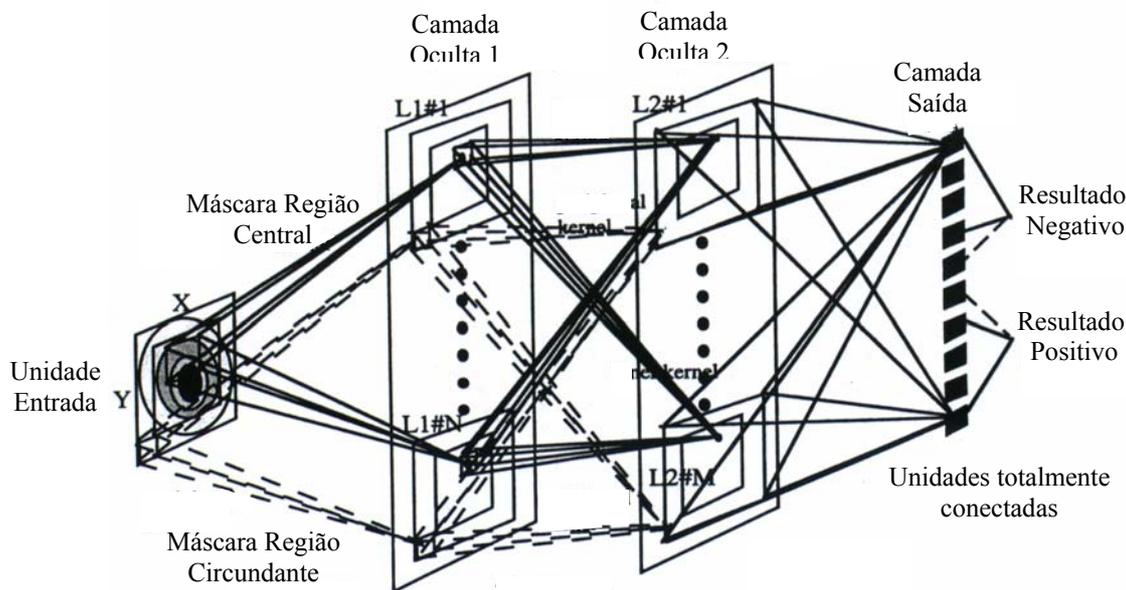


Figura 4.11 – Arquitetura da rede convolucional para reconhecimento de padrões de imagens médicas.

No entanto, a arquitetura com dois grupos de pesos torna-se viável quando o reconhecimento envolve formas onde se pode identificar um limite entre ambas as regiões. Assim, a arquitetura contendo os dois grupos foi utilizada para a detecção de nódulos pulmonares, enquanto que para a detecção das microcalcificações foi utilizada uma arquitetura com um único grupo de pesos.

O bloco de entrada consiste de 32x32 pixels extraídos de uma imagem como suspeitos. Com o objetivo de melhorar o desempenho da rede, optou-se por reduzir a imagem para um bloco de 16x16 pixels, encontrando-se a média para cada quatro pixels.

A primeira camada oculta consiste de doze grupos formatados em uma matriz quadrada de 12x12 pixels, com área central definida em 8x8 pixels. A Segunda consiste de doze grupos formatados em uma matriz quadrada de 8x8 pixels, com área central definida em 6x6 pixels. Já a camada de saída é composta por dez posições totalmente conectadas com a segunda camada oculta.

No caso do reconhecimento das microcalcificações a arquitetura prevalece a mesma, desconsiderando-se apenas as definições das áreas centrais, uma vez que para esta aplicação foi considerado um único grupo de pesos.

Uma observação interessante é em relação ao número de camadas necessárias. Quanto mais complexo forem os padrões, maior o número de camadas necessárias para distinguir a alta ordem de informação da estrutura da imagem.

Como o objetivo da rede é o de distinguir os padrões de doenças, foram realizados alguns acompanhamentos junto aos radiologistas. Notou-se alguns casos onde era possível o diagnóstico claro da doença e em outros se estimava a probabilidade. Neste último caso, a rede neural pode ser mais realística tratando saídas em termos de probabilidade. Em outras palavras, pode-se definir uma escala de pesos de incidência sobre as saídas.

No entanto, deparou-se com o problema em que as regras de decisões utilizadas pelos radiologistas não estavam relacionadas. Para isso, foi proposto o uso de uma distribuição gaussiana para estabelecer uma associação de incerteza, “*fuzzy*”⁷, entre as saídas adjacentes. Assim, o radiologista ao determinar uma probabilidade, para um padrão de doença em uma área da imagem, baseado em sua experiência, esta probabilidade será acompanhada por uma variação.

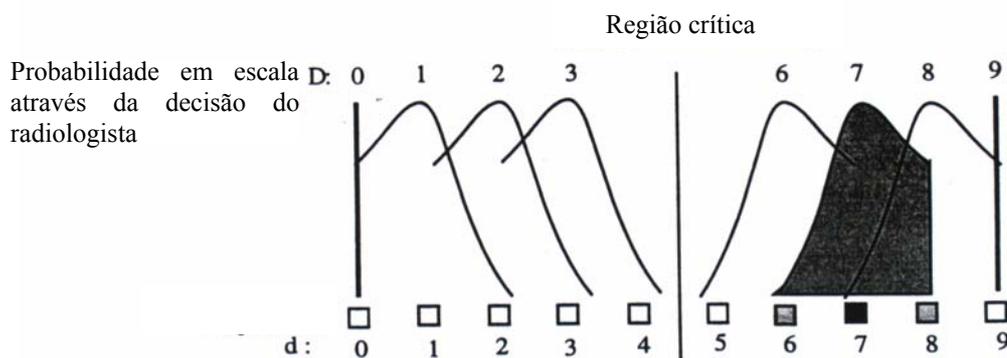


Figura 4.12 – Exemplo de saída utilizando-se a associação de incerteza (*fuzzy*) construídas pela distribuição gaussiana.

⁷ Indica o grau de pertinência em que um determinado padrão pertence a um conjunto.

Os blocos de imagens utilizados pela rede passaram por uma fase de pré-processamento, onde foram borrados e rotacionados em 90°, 180° e 270°.

Para o caso dos nódulos pulmonares, os experimentos foram realizados sobre radiografias de pacientes com câncer na fase primária e avançada, contendo um ou mais nódulos.

Os resultados obtidos através dos experimentos realizados por Shing[Shi95] são mostrados através do índice A_z baseado nas características ROC⁸:

- $A_z = 0,77$ de acerto usando-se a rede com uma função delta para determinar a saída;
- $A_z = 0,83$ de acerto usando-se a rede com a distribuição gaussiana para associação das saídas; e
- $A_z = 0,87$ de acerto usando-se a rede com a distribuição gaussiana para associação das saídas e considerando-se oito tipos de blocos de imagens rotacionadas (ao invés de três).

Durante a fase do pré-processamento para as microcalcificações, foi utilizada a transformada Wavelet (domínio do espaço -> domínio da frequência) e aplicado um filtro para eliminar as baixas frequências (filtro passa-alta) antes da reconstrução da imagem (domínio da frequência -> domínio do espaço). Os blocos resultantes foram utilizados como entrada para a rede.

Os resultados obtidos através dos experimentos realizados foram:

- $A_z = 0,89$ de acerto usando-se a rede com a determinação baseada em microcalcificações individuais;

⁸ *Receiver Operating Characteristic* (ROC) é um método analítico aplicado para avaliação do desempenho de sistemas. O índice A_z representa o desempenho indexado do sistema usando análise ROC.

-
- $A_z = 0,97$ de acerto usando-se a rede com a determinação baseada em clusters de microcalcificações.

Ao fim do trabalho o autor [Shi95] conclui que as redes neurais convolucionais e suas técnicas de treinamento podem ser estendidas para muitas aplicações.

No próximo capítulo serão demonstrados a especificação da arquitetura desenvolvida e os resultados obtidos para o reconhecimento facial.

Capítulo V

Especificação do Projeto e Resultados

Neste capítulo serão apresentados a rede neural NEOPC, desenvolvida como parte desta dissertação, os resultados e as conclusões obtidas. A arquitetura da NEOPC, rede neural baseada no Neocognitron com utilização de Pontos de Controle, é composta por uma fase de pré-processamento e duas redes neurais: RNDPC (rede neural para a detecção dos pontos de controle) e a RNRF (rede neural para o reconhecimento facial). O algoritmo para o pré-processamento foi implementado em Matlab 5.2 e as redes neurais em Delphi 5.0, sendo utilizado um computador Pentium 500 Mhz com 128 Mbytes de RAM.

5.1 – Projeto

A NEOPC é composta de um recurso adicional estabelecido à estrutura descrita no treinamento do neocognitron, que é o uso de pontos de controle para a obtenção das células-semente, além do padrão de entrada. Conforme visto na seção 4.1, os pontos de controle são usados pela rede para que as células vencedoras só sejam escolhidas como célula-semente, caso elas estejam posicionadas estrategicamente em relação ao padrão de entrada.

A justificativa para se acrescentar os pontos de controle sobre os padrões de entrada é que a aplicação exclusiva do padrão de entrada no algoritmo de treinamento, pode resultar na obtenção de células-semente que não sejam adequadas para garantir a

seletividade⁹ de rede., como por exemplo, aquelas correspondentes às regiões totalmente escuras pertencentes aos cabelos.

A Figura 5.1 apresenta o arquétipo da NEOPC, seguida pela explicação de cada etapa.

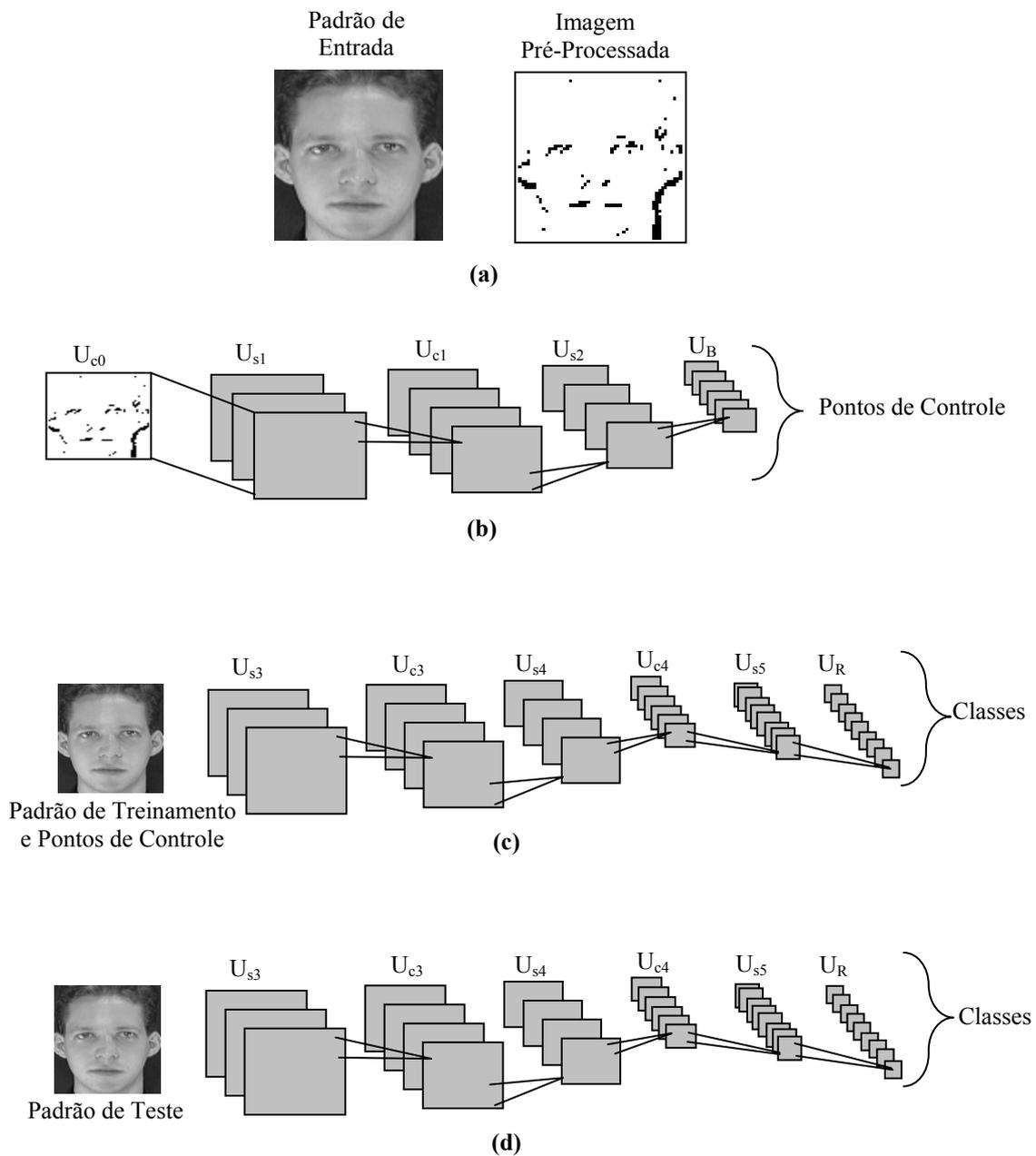


Figura 5.1 – Etapas da NEOPC: (a) Pré-Processamento, (b) RNDPC: obtenção dos pontos de controle, (c) RNRF: treinamento e (d) RNRF: reconhecimento.

⁹ Entende-se como seletividade, a capacidade de identificar corretamente a classe correspondente ao padrão de entrada.

5.1.1 – Pré-Processamento

A rede neural para detecção dos pontos de controle (RNDPC) foi treinada de modo a detectar os pontos de curvatura dos padrões apresentados. Portanto, o processamento da RNDPC é baseado na detecção de bordas, seguida da detecção de segmentos de retas (linhas) e de pontos de curvatura.

Neste sentido, a etapa de pré-processamento foi realizada afim de obter, como imagem resultante, os contornos dos olhos, nariz e lábios, que serão utilizados como padrões de entrada para a RNDPC (Figura 5.1);

O pré-processamento realizado consiste basicamente na segmentação da imagem considerando-se a similaridade dos pixels [Gon92]. A Figura 5.2 apresenta um histograma de níveis de cinza correspondente a imagem “A”. Como pode ser verificado, os pixels da imagem possuem seus níveis de cinza agrupados em dois grupos: o círculo (parte clara da imagem) e o fundo (parte escura). Assim, para segmentar a imagem ($f(x,y)$) basta considerar $f(x,y) > T$, onde T representar o valor de *threshold*, para obter o círculo ou $f(x,y) < T$ para obter o fundo.

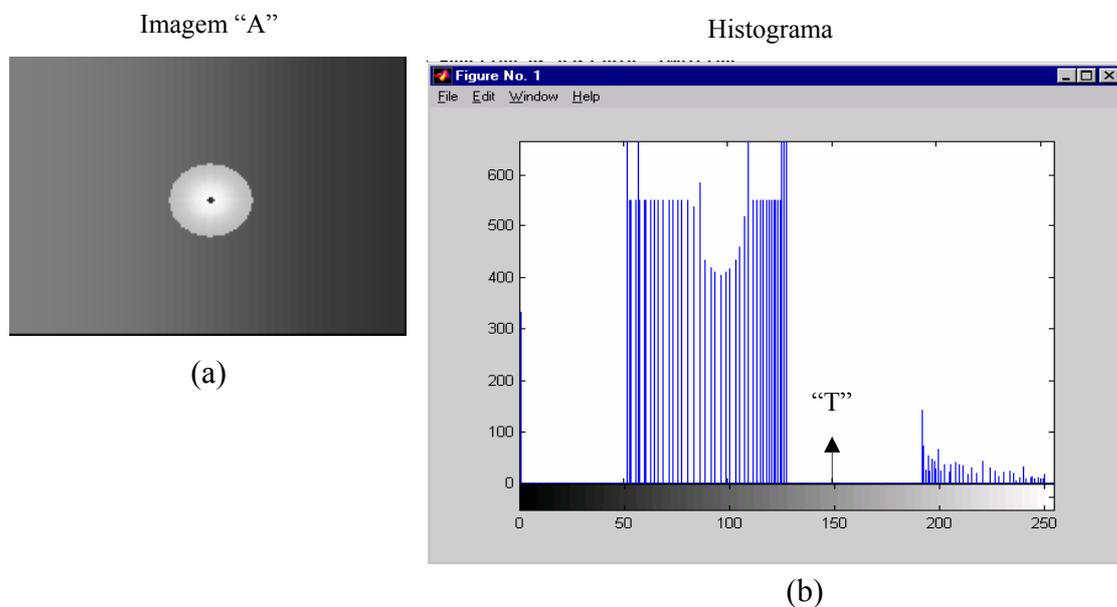


Figura 5.2 – Exemplo de segmentação por *threshold*: (a) imagem original e (b) histograma em níveis de cinza da imagem “A” (a).

No caso das imagens faciais não é possível identificar o agrupamento de níveis de cinza para cada estrutura da face (olhos, nariz, lábios). Porém, torna-se possível extrair o contorno destas estruturas uma vez que elas são compostas por pixels de menor intensidade. Deste modo, a segmentação torna-se cada vez maior em relação ao decremento do valor do *threshold*. A Figura 5.3 apresenta o resultado da segmentação para uma imagem facial, variando o valor do *threshold*.

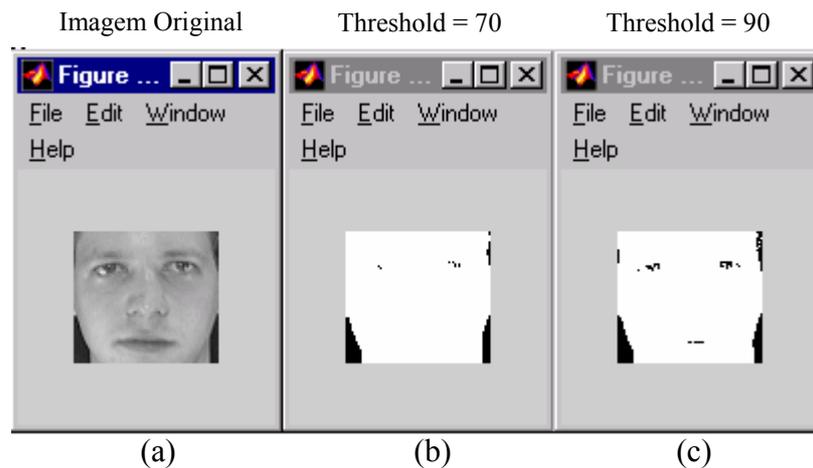


Figura 5.3 – Segmentação da imagem facial utilizando-se valores diferentes de *threshold*: (a) imagem original, (b) *threshold* = 0.70 e (c) *threshold* = 0.90.

Ao analisar os histogramas de níveis de cinza das imagens faciais (Figura 5.4), observou-se uma grande frequência de pixels nas intensidades que correspondem aos contornos das estruturas da face ocasionando dificuldades para estabelecer o valor de *threshold*.

Em outras palavras, considerando-se um valor de *threshold* muito baixo, perdem-se muitos dos padrões que compõem os contornos das estruturas faciais, enquanto que para um valor de *threshold* muito alto obtém-se diversos padrões que não correspondem aos contornos das estruturas faciais.

Neste sentido, foi realizado a obtenção do sombreamento da imagem facial, que resulta em uma imagem com descontinuidade dos pixels nas regiões onde estão localizadas as estruturas faciais. A Figura 5.4 apresenta o histograma de níveis de cinza para uma

imagem facial e para o sombreamento desta imagem, assim como, o resultado da segmentação por *threshold* para ambos os casos.

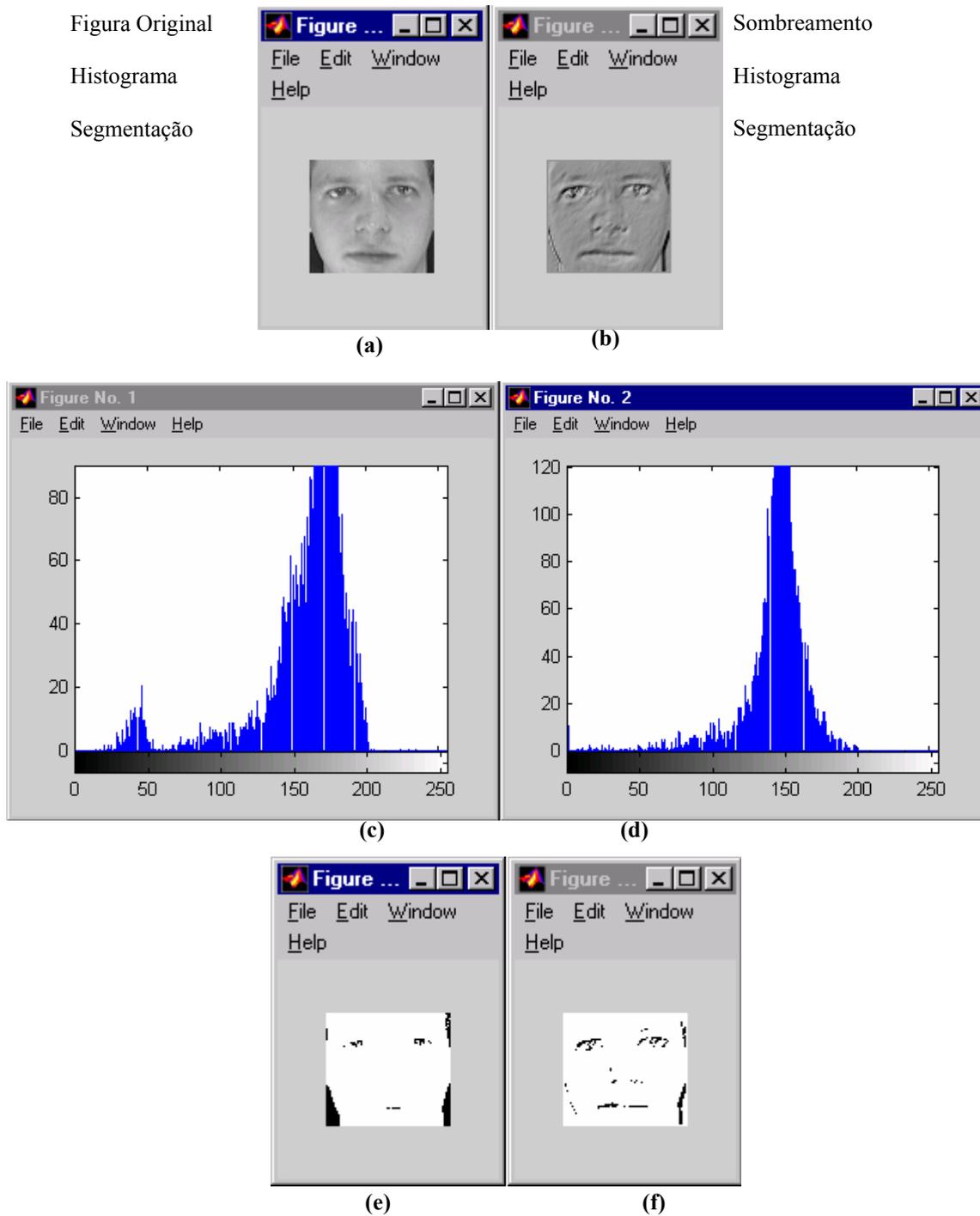


Figura 5.4 – Exemplo de segmentação da imagem facial: (a) a imagem original, (b) sombreamento da imagem original, (c) histograma em níveis de cinza da imagem original, (d) histograma em níveis de cinza do sombreamento da imagem original, (e) resultado da segmentação da imagem original com *threshold* de 0.90 e (f) resultado da segmentação do sombreamento da imagem original com *threshold* de 0.90.

Como pode ser verificado nos histogramas apresentados na Figura 5.4, torna-se possível estabelecer um *threshold* (“T”) para segmentar dois grupos: contornos e fundo.

Outros testes como, por exemplo, a aplicação do filtro de Sobel [Gon92] também foram realizados para detecção de bordas. Porém, como já era esperado, esses filtros não apresentaram bons resultados, uma vez que as imagens faciais apresentam, em regiões como os olhos, nariz e lábios, uma variação muito pequena na intensidade dos pixels.

A Figura 5.5 apresenta as etapas realizadas durante a fase de pré-processamento.

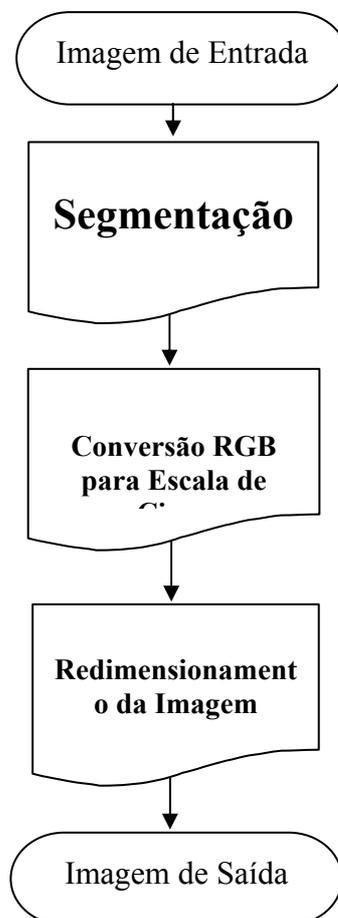


Figura 5.5 – Diagrama das etapas do pré-processamento

A Figura 5.5 apresenta três etapas que foram utilizadas de acordo com os testes realizados e a base de imagens adotada.

A etapa de “Segmentação” consiste na obtenção de algumas estruturas da face que foram utilizadas para a realização de determinados testes, como por exemplo, o olho esquerdo. A segmentação foi realizada de forma interativa neste trabalho.

Já a etapa “Conversão de RGB para Escala de Cinza” está ligada as bases de imagens utilizadas, ou seja, duas bases de imagens estão no formato RGB e a arquitetura de rede neural desenvolvida trabalha com imagens em níveis de cinzas.

A última etapa especificada no pré-processamento, “Redimensionamento da Imagem”, também está ligada a arquitetura da rede neural desenvolvida. A camada de entrada da rede neural, especificada no próximo item, foi projetada para imagens de dimensão 57x57. Assim, através do método de interpolação de ordem zero [Gon92] a imagem resultante é redimensionada para 57x57.

A Figura 5.6 demonstra o resultados das etapas do pré-processamento.

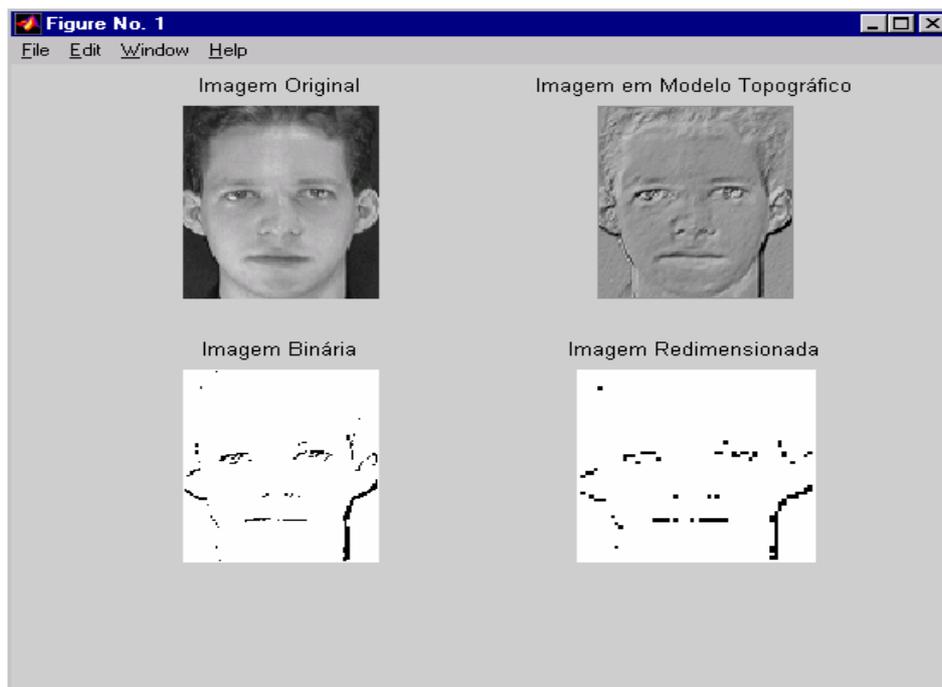


Figura 5.6 – Etapas do Pré-Processamento. O pré-processamento é realizado na ordem em que é demonstrado na Figura, iniciando da parte superior, da esquerda para direita: (a) imagem original, (b) sombreamento da imagem, (c) segmentação do sombreamento da imagem original com o uso de *threshold* e (d) redimensionamento da imagem (57x57).

5.1.2 – Rede Neural para a Detecção dos Pontos de Controle (RNDPC)

A Rede Neural para a Detecção de Pontos de Controle (RNDPC) é baseada na estrutura do neocognitron e, conforme já mencionado, é responsável pela obtenção dos pontos de controle a serem utilizados pela RNRN (rede neural para o reconhecimento facial) durante o treinamento não-supervisionado.

O padrão de entrada (camada U_0) utilizado para a RNDPC, corresponde a uma amostra pré-processada (Figura 5.1).

A RNDPC possui suas interconexões pré-definidas, que estabelece a detecção de bordas, seguida da detecção de segmentos de retas, de pontos de curvatura, e de pontos extremos, sobre os padrões de entrada.

A Figura 5.7 ilustra as etapas correspondentes à RNDPC, onde as bordas são detectadas sobre ambos os lados (A e B) de uma linha em 8 direções principais, que variam de 22.5 graus, a partir do ângulo 0. A existência de duas bordas (uma de cada lado) numa determinada direção, identifica um segmento de reta, ou linha. Pequenas distorções na espessura de uma linha é absorvida pela estrutura da rede, que permite essas variações, pelo uso da camada de células-C.

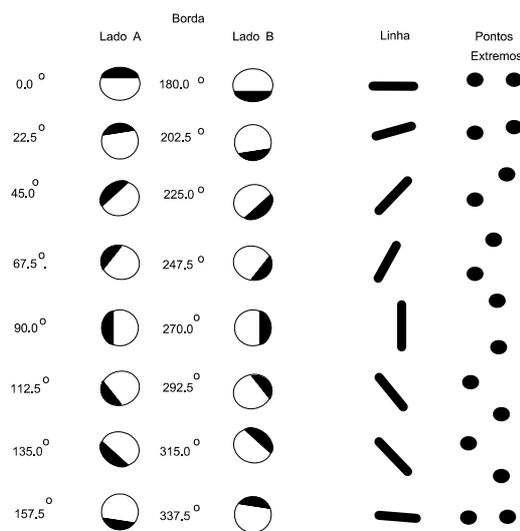


Figura 5.7 - Ilustração dos estágios da RNDPC.

As conexões das células- S das camadas Us_1 e Us_2 são semi-fixas¹⁰. No estado inicial, quando a rede é construída, as conexões são modificadas pelo método de treinamento supervisionado do neocognitron. Mais especificamente, uma célula-semente é arbitrariamente selecionada num plano-celular, e um padrão de treinamento, atribuído à célula, é apresentado na camada de entrada U_0 . Cada conexão de entrada da célula-semente é reforçada proporcionalmente à intensidade de resposta da célula, de onde provem a conexão. As conexões de entrada de todas as outras células do plano-celular são automaticamente reforçadas, na mesma proporção da célula-semente.

O padrão de treinamento para cada plano-celular de Us_1 é uma borda numa orientação particular, conforme se verifica na Figura 5.7, à esquerda. A borda, contudo, não é localizada no centro do campo receptivo. Existe um pequeno deslocamento, em relação ao centro, para que a célula responda a uma borda de uma linha muito estreita.

Uma linha estreita em uma particular orientação é usada para o treinamento de um plano-celular da camada Us_2 . Na Figura 5.7 é vista após os padrões de bordas, os padrões de segmentos de retas. Apenas uma simples apresentação de uma linha de certa espessura é necessária para treinar as células de extração de linhas, em diferentes espessuras.

A camada seguinte U_B é responsável pela detecção de pontos extremos das linhas retas. Como uma curva é composta de um conjunto de segmentos de linhas retas, os pontos extremos que pertencem à junção de dois segmentos de linhas representam um ponto de curvatura. Uma célula- B recebe conexões de entrada somente de um único plano-celular da camada Us_2 . Essas conexões de entrada são fixas e tem uma distribuição espacial que permite a detecção de um dos pontos extremos de uma linha numa determinada direção. Em contraste às células- S , uma célula- B (*bend points*), funciona de forma subtrativa, ao invés da forma multiplicativa, e portanto, não usa as células- V que estão embutidas na estrutura de cálculo das células- S .

A estrutura dessa rede é semelhante a usada numa rede de extração de curvaturas [Fuk71]. O arranjo espacial das conexões excitatórias e inibitórias coincide com a

¹⁰ As conexões são ditas semi-fixas porque, nenhuma alteração é feita sobre essas células, posteriormente.

orientação preferencial do plano-celular da camada precedente U_{S_2} . As células- B são ativadas numa das extremidades de uma linha obtida em U_{S_2} . Quando uma linha reta é apresentada à camada de entrada U_0 , as respostas das células- B são ativas numa das extremidades das linhas. Quando uma linha curva é apresentada, as células- B respondem aos pontos de curvatura da linha.

A rede neural RNDPC implementada permite variar o limiar de disparo na detecção de bordas, correspondente ao primeiro estágio, e na detecção de segmentos de reta, correspondente ao segundo. Com esse recurso, os pontos de controle obtidos podem variar usando limiares de disparo diferentes, fazendo com que as células-semente da rede neural secundária sejam distintas, para limiares diferentes. O limiar de disparo relaciona-se à excitabilidade dos neurônios envolvidos. Essa excitabilidade varia de neurônio a neurônio no cérebro biológico, como pode ser visto na ilustração da Figura 5.8 [Guy93].

Nessa Figura são vistos três neurônios cerebrais, e os respectivos pontos de disparo. É possível verificar também que além da posição de excitação (coordenada horizontal), os neurônios variam na quantidade de descargas por segundo (coordenada vertical), que correspondem ao peso das conexões do modelo proposto.

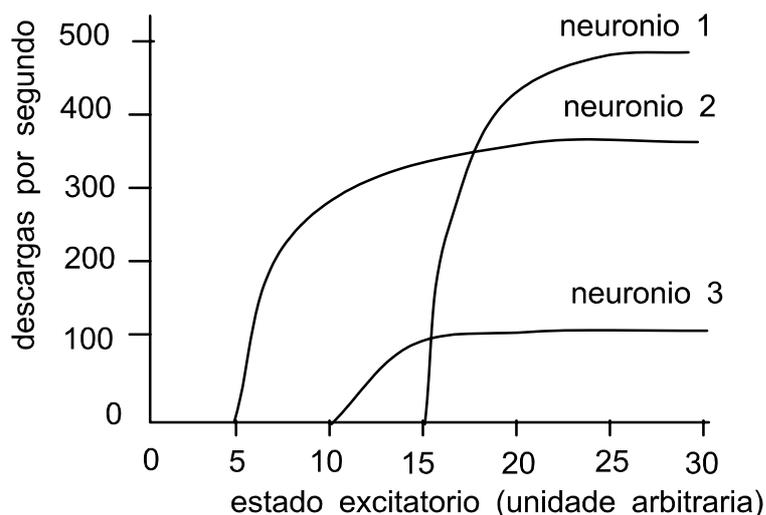


Figura 5.8 - Diferenças em três neurônios biológicos

A Figura 5.9 mostra o resultado da obtenção dos pontos de controle usando um limiar de disparo alto para a detecção de bordas (0.84) e de linhas (0.84). A primeira coluna à esquerda mostra a imagem do padrão de entrada. As duas colunas seguintes mostram respectivamente as bordas de ambos os lados (A e B) dos segmentos de retas detectados. A coluna situada no centro corresponde aos segmentos de retas, ou linhas, detectados, nas oito direções. As duas últimas colunas correspondem aos respectivos pontos extremos. As seleções vistas no canto direito correspondem respectivamente aos limiares de detecção de borda e linha.

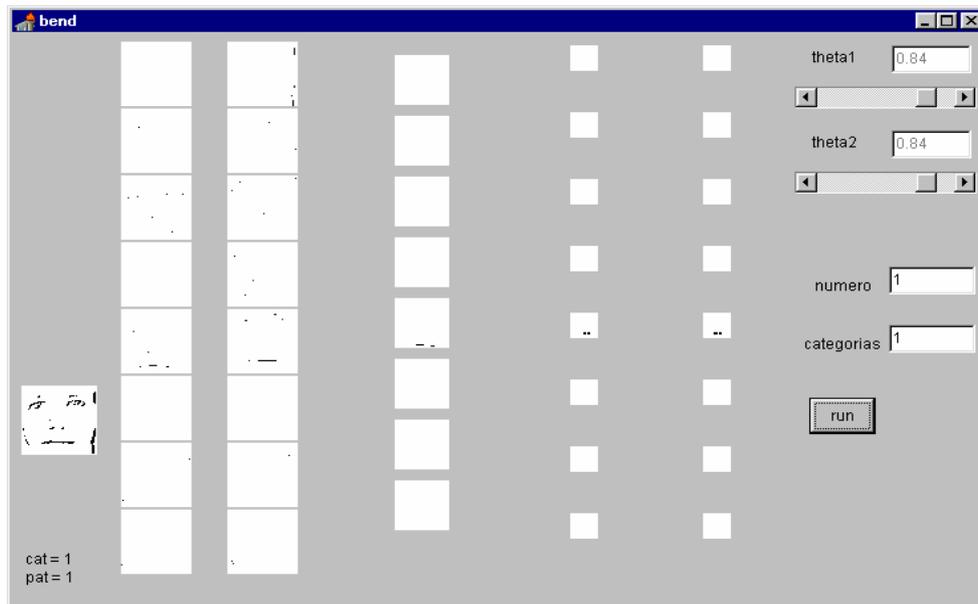


Figura 5.9 – RNDPC – Obtenção dos pontos de controle considerando um limiar para detecção de bordas e de linhas de 0.84.

A Figura 5.10 por outro lado, mostra o mesmo resultado para o mesmo padrão de entrada da Figura 5.9, porém, com limiar de disparo de detecção de bordas menor (0.6), e mesmo limiar de detecção de linhas (0.8). Com isso, é possível verificar que mais linhas foram detectadas e portanto, mais pontos extremos, ou pontos de controle, foram obtidos.

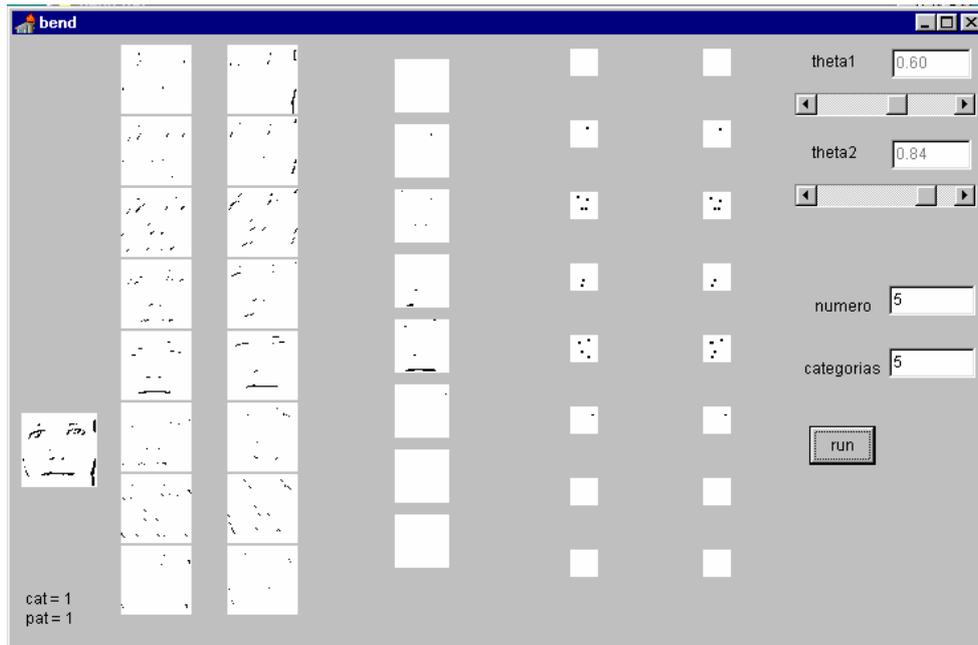


Figura 5.10 – RNDPT - Obtenção dos pontos de controle considerando um limiar para detecção de bordas de 0.60 e de linhas de 0.84.

5.1.3 – Rede Neural para o Reconhecimento Facial (RNRF)

A Rede Neural para o Reconhecimento Facial (RNRF) é uma rede baseada no neocognitron, com uma diferença no treinamento da sua primeira camada.

O treinamento da RNRF consiste de aprendizado não-supervisionado, que usa, além do padrão de entrada (Figura 5.1), os pontos de controle obtidos pela RNDPC, que consistem de pontos extremos e pontos de curvatura.

Assim, é apresentada à camada de entrada U_0 , uma amostra facial, cujas posições estratégicas de obtenção dos fatores são obtidos previamente RNDPC. O algoritmo de treinamento não-supervisionado é então aplicado, verificando-se que as células-semente, coincidem com os pontos de controle. Dessa forma assegura-se que os fatores específicos de um padrão de entrada são aprendidos, em posições estratégicas, definidas pelos pontos de controle. A Figura 5.11, ilustra o treinamento não supervisionado, utilizando os pontos de controle.

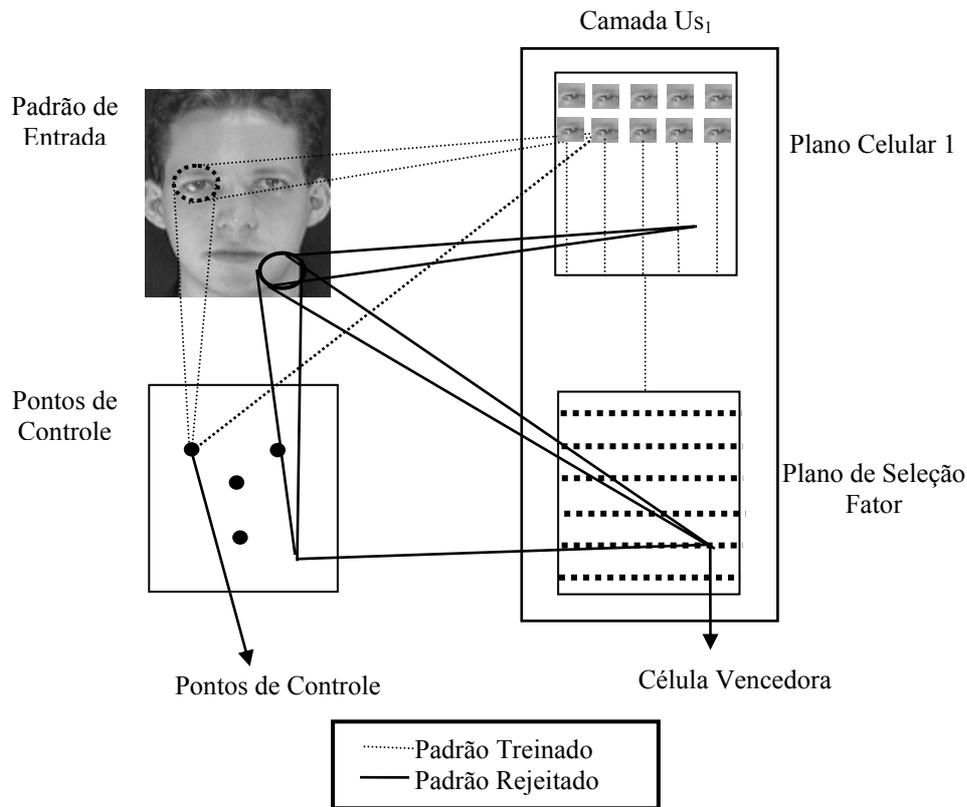


Figura 5.11 – Exemplo do treinamento não supervisionado utilizando os pontos de controle.

No caso da Figura 5.11, o plano celular 1 da camada U_s , foi treinado para um fator existente no olho esquerdo. Nota-se a existência de um ponto de controle na posição relativa ao olho esquerdo. Assim, caso o plano de seleção de fator seja excitado para o mesmo fator do olho esquerdo, para alguma posição do padrão de entrada, as células da camada 1 também serão excitadas naquela posição, notando-se que todas as células de um mesmo plano celular tem os mesmos pesos nas suas respectivas conexões de entrada. Neste caso, o algoritmo de treinamento parte em busca de outra célula vencedora. Porém, caso uma célula do plano de seleção seja excitada para fator inexistente nas camadas anteriores, o algoritmo de treinamento não supervisionado irá verificar se a mesma posição da célula vencedora há algum ponto de controle. Na Figura 5.11, para o fator localizado na região inferior da face, não há nenhum ponto de controle correspondente. Neste caso, o algoritmo parte em busca de uma nova célula vencedora.

A RNRF obtida é composta de três estágios, numerados de 3 a 5, cada um dos quais composto de uma camada de células-*S* e uma outra de células-*C*, com exceção do último estágio, que é composto de uma camada de células-*S*, seguida de uma camada de células-*R*. A seqüência de camadas existentes são portanto: $U_{s3} > U_{c3} > U_{s4} > U_{c4} > U_{s5} > U_R$, notando-se que as camadas de numerações 0,1 e 2, correspondem à rede RNDPC. O treinamento e o reconhecimento para essa rede é igual ao neocognitron, com exceção do último estágio, que corresponde de células-*S* com entradas que identificam as respectivas categorias. Durante o treinamento, a categoria do padrão de entrada é associada a uma das entradas da célula-*S* do último estágio. Assim, uma célula-*R* escolhe a resposta de maior intensidade da última camada de células-*S*. Durante a fase de reconhecimento, a rede adota a célula-*R* de maior intensidade de resposta, como a resposta correta.

A Figura 5.12 mostra o resultado do reconhecimento de padrão, pelas respectivas camadas, na RNRF. No exemplo, uma amostra correspondente a classe 1 é usado para teste de reconhecimento, apresentada no lado esquerdo inferior.

O conjunto de matrizes bidimensionais à esquerda mostra os sinais de excitação (em manchas escuras) das células-*S* da primeira camada U_{s1} , da RNRF. No caso existem 7 matrizes, ou planos celulares, nessa camada. Na Figura 5.12 não estão sendo mostrados os planos de células-*C*, para simplificação, que se situam posteriormente às células-*S*. Na parte central são vistos os planos celulares da camada U_{s2} , que no exemplo, são em número de 30, e que os primeiros planos, situados à esquerda, apresentam maiores sinais de excitação de suas células. Logo à direita, são vistos os planos-celulares da camada U_{s3} , e posteriormente, uma coluna correspondente às células-*R*, sendo uma célula para cada classe ser reconhecido, de 1 a 10. No exemplo, a amostra utilizada pertence a classe 1, sendo reconhecida corretamente. No canto direito são vistas as seleções para o ajuste dos limiares das células nas três camadas e a seleção de um novo padrão de teste.

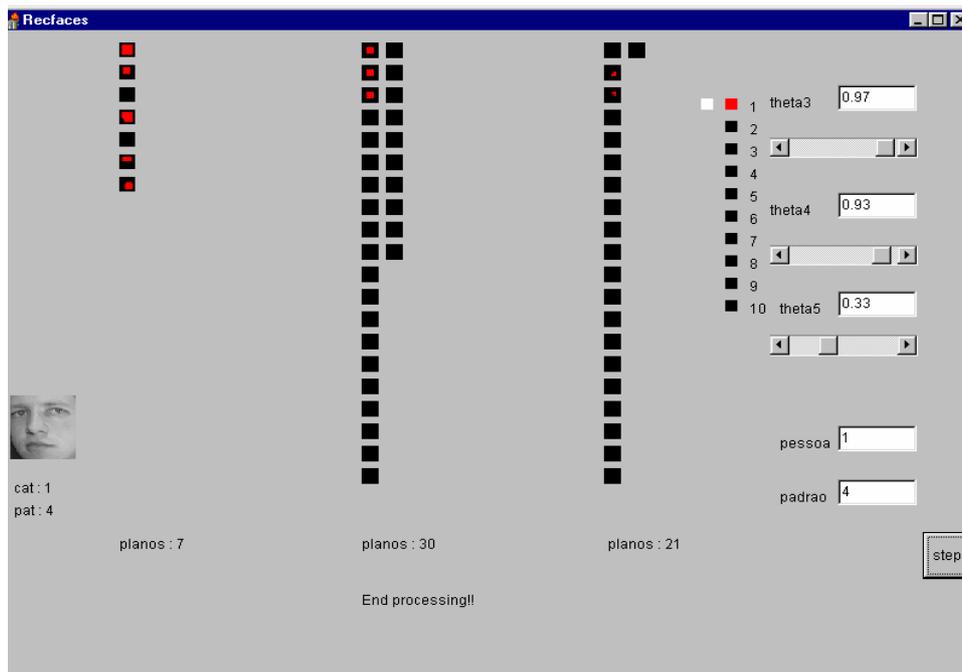


Figura 5.12 – RNRf no reconhecimento de amostras

A fase de reconhecimento usa somente a RNRf, uma vez que a RNDPC serve apenas para a obtenção dos pontos de controle.

5.2 – Resultados

Conforme mencionado no capítulo II, para a tarefa de reconhecimento facial é de grande importância a utilização das variações sutis apresentada por cada imagem para a determinação das características relevantes que as mesmas possuem [Abd95], bem como, estabelecer uma variação onde diferentes faces gerem códigos os mais distintos possíveis, enquanto que diferentes imagens de uma mesma face gerem códigos bastantes similares [Dau97]. Na seção 5.1, pode-se verificar que a NEOPC permite a manipulação da extração de características através do uso da variação de *thresholds*, que são limiares de excitação de neurônios. Assim, para os testes realizados, considerou-se como resultado as simulações realizadas com o melhor *threshold* encontrado, ou seja, o que apresentou o melhor resultado para as amostras utilizadas.

As Figuras apresentadas no texto exemplificando as imagens utilizadas para a realização do testes estão organizadas por classes de amostras. Cada classe representa o conjunto de amostras de uma mesma pessoa.

Os testes estão organizados através de quatro bases de imagens: base 1, base 2, base 3 e base 4. As características de cada base de imagens estão especificadas nos próximos sub-itens, onde são apresentados os resultados obtidos.

5.2.1 – Base de Imagens 1

As imagens faciais utilizadas foram obtidas pelo site www.cam.ac.uk/facesataglance.html, da Universidade de Cambridge, onde é disponível um banco de imagens faciais de 40 pessoas, com 10 imagens diferentes para a mesma pessoa. As imagens possuem diferentes poses, expressões faciais, luminosidade e acessórios.

Os testes com a NEOPC foram iniciados de modo à avaliar a operacionalidade para diferentes tipos de amostras. Assim, foram obtidos os resultados considerando amostras sub-faciais – olho esquerdo, amostras faciais completas e amostras faciais - central.

5.2.1.1 - Resultados para Amostras Subfaciais – Olho Esquerdo.

A análise da operacionalidade da NEOPC foi iniciada considerando apenas o olho esquerdo da face como amostra, conforme apresentado na Figura 5.13.

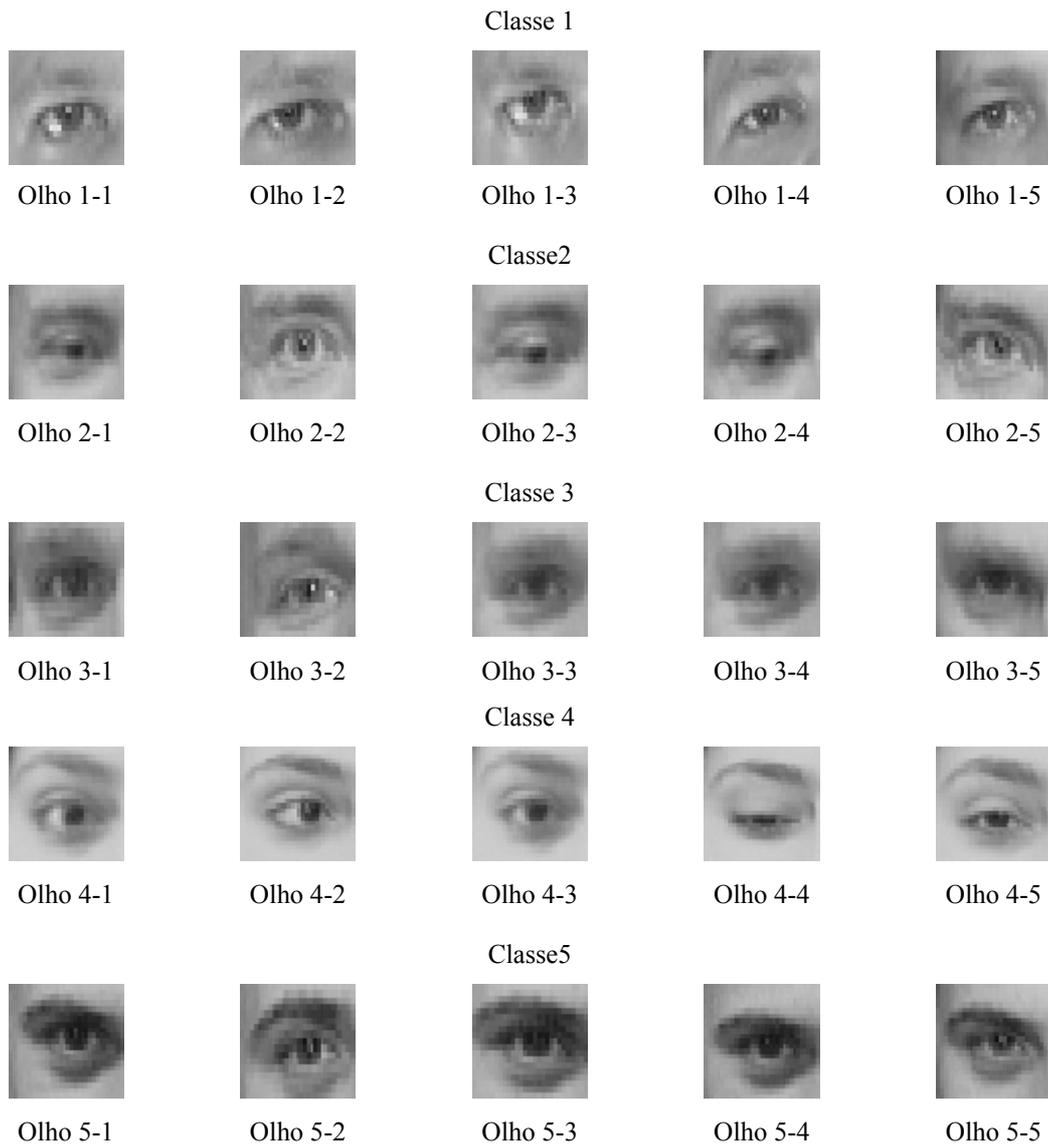


Figura 5.13 – Base de Imagens 1: classes e amostras subfaciais – olho esquerdo

O teste foi realizado considerando todas as amostras apresentadas para cada classe, tanto para a fase de treinamento como de reconhecimento. Os resultados obtidos estão apresentados na Tabela 5.1.

Tabela 5.1 – Base de Imagens 1 - Amostras Subfaciais – Olho Esquerdo: *Thresholds* e taxa de reconhecimento considerando-se as mesmas amostras para o treinamento e reconhecimento (5 classes contendo 5 amostras cada uma).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.84	0.84	0.97	0.97	0.97	0.97	0.93	0.33

Taxa de Acerto	72%
Taxa de Erro	28%

Como pode ser verificado, a rede apresentou uma taxa de acerto de 72% utilizando-se as mesmas amostras para o treinamento e reconhecimento. Mediante este fato, optou-se pela realização de um novo teste considerando a imagem facial completa como amostra.

5.2.1.2 – Resultados para Amostras Faciais Completa

Uma vez que o objetivo é verificar o desempenho da NEOPC em relação ao tipo de amostra utilizada, foi considerado o mesmo conjunto de amostras para o treinamento e reconhecimento, conforme o teste realizado anteriormente. A Figura 5.14 apresenta as amostras faciais utilizadas



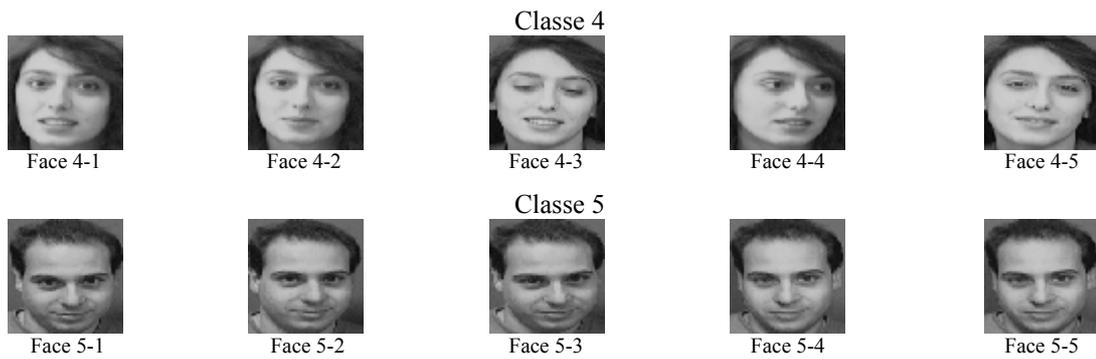


Figura 5.14 – Base de Imagens 1: classes e amostras faciais completas.

Os resultados obtidos estão apresentados na tabela abaixo.

Tabela 5.2 – Base de Imagens 1 – Amostras Faciais Completas: *Thresholds* e taxa de reconhecimento considerando-se as mesmas amostras para o treinamento e reconhecimento (5 classes contendo 5 amostras cada uma).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.84	0.84	0.93	0.99	0.97	0.97	0.93	0.33
Taxa de Acerto		89%						
Taxa de Erro		11%						

Comparando a Tabela 5.1 e 5.2, nota-se que a taxa de acerto é melhor quando se considera a amostra facial completa.

Notou-se, no entanto, que durante a fase de obtenção dos pontos de controle para o treinamento das amostras faciais completas, foram detectados um número excessivo de pontos de controle para determinadas amostras. Este fator gerou uma certa dificuldade para a fase de treinamento e reconhecimento.

Assim, para o próximo teste foram considerados a região central de cada amostra.

5.2.1.3 – Resultados para Amostras Faciais Segmentada

Com base nos resultados da etapa de detecção dos pontos de controle do teste anterior, decidiu-se em alterar mais uma vez a imagem de entrada considerando apenas a região central da face, composta pelos olhos, nariz e lábios. As amostras faciais – central foram obtidas através da etapa de segmentação apresentada na fase de pré-processamento. A Figura 5.15 apresenta o resultado da segmentação, tendo 15 classes em linhas, subsequentes, e 10 colunas para cada classe, numeradas da esquerda para a direita em ordem crescente. Assim, a amostra 3 da classe 6, é denotada Face 6-3. Essa notação será usada nas descrições posteriores.



Classe 6



Face 6-1 Face 6-2 Face 6-3 Face 6-4 Face 6-5 Face 6-6 Face 6-7 Face 6-8 Face 6-9 Face 6-10

Classe 7



Face 7-1 Face 7-2 Face 7-3 Face 7-4 Face 7-5 Face 7-6 Face 7-7 Face 7-8 Face 7-9 Face 7-10

Classe 8



Face 8-1 Face 8-2 Face 8-3 Face 8-4 Face 8-5 Face 8-6 Face 8-7 Face 8-8 Face 8-9 Face 8-10

Classe 9



Face 9-1 Face 9-2 Face 9-3 Face 9-4 Face 9-5 Face 9-6 Face 9-7 Face 9-8 Face 9-9 Face 9-10

Classe 10



Face 10-1 Face 10-2 Face 10-3 Face 10-4 Face 10-5 Face 10-6 Face 10-7 Face 10-8 Face 10-9 Face 10-10

Classe 11



Face 11-1 Face 11-2 Face 11-3 Face 11-4 Face 11-5 Face 11-6 Face 11-7 Face 11-8 Face 11-9 Face 11-10

Classe 12



Face 12-1 Face 12-2 Face 12-3 Face 12-4 Face 12-5 Face 12-6 Face 12-7 Face 12-8 Face 12-9 Face 12-10

Classe 13



Face 13-1 Face 13-2 Face 13-3 Face 13-4 Face 13-5 Face 13-6 Face 13-7 Face 13-8 Face 13-9 Face 13-10

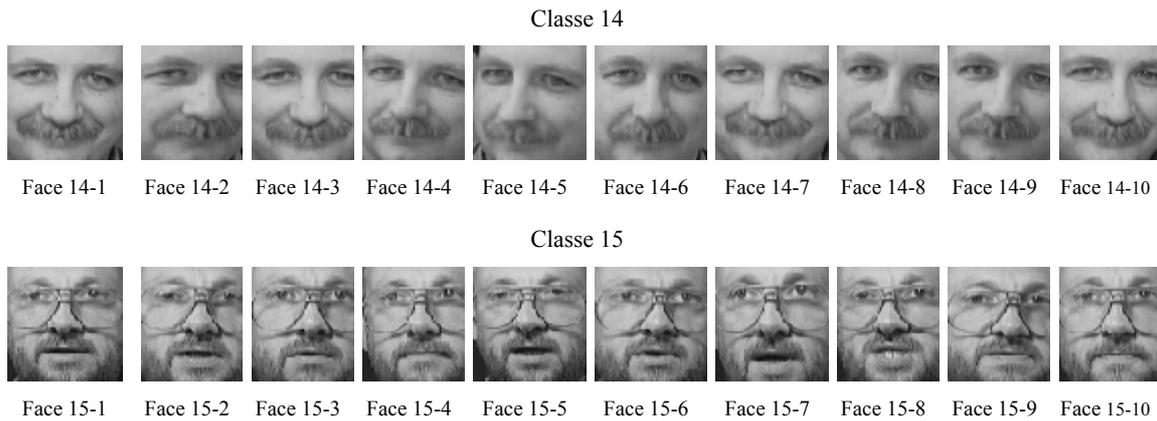


Figura 5.15 – Base de Imagens 1: classes e amostras segmentadas.

A Tabela 5.3 apresenta os resultados obtidos quando avaliada a operacionalidade da NEOPC adotando-se os mesmos critérios dos testes anteriores, mesmo conjunto de amostras para a fase de treinamento e reconhecimento. Este teste foi realizado para avaliar, a princípio, a taxa de acerto com o uso de amostras faciais segmentadas em relação aos demais tipos de amostras utilizadas (olho esquerdo e a face completa) e a quantidade de amostras utilizadas para o treinamento. Como as amostras utilizadas para o treinamento e reconhecimento são as mesmas, a taxa de acerto deve ser de 100%. Caso isso não ocorra, o número de amostras utilizadas para o treinamento deve ser considerada como insatisfatória, ou seja, os padrões treinados pela rede não são suficientes para estabelecer uma boa variação inter-classe. O teste foi realizado com as classes 1, 2, 3, 4 e 5 e com as 5 primeiras amostras (1 a 5) de cada classe, totalizando 25 amostras.

Tabela 5.3 – Base de Imagens 1 – Amostras Faciais Segmentada: *Thresholds* e taxa de reconhecimento considerando-se as mesmas amostras para o treinamento e reconhecimento (5 classes contendo 5 amostras cada uma).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.84	0.84	0.97	0.97	0.97	0.97	0.93	0.33
Taxa de Acerto		100%						
Taxa de Erro		0						

Conforme pode ser verificado, o uso de amostras faciais segmentadas apresentou uma taxa de acerto superior aos demais tipos de amostras (Tabela 5.1 e 5.2) considerando-se a mesma quantidade de amostras para o treinamento e reconhecimento. Portanto, são apresentados a seguir, outros testes conduzidos através do uso de amostras faciais segmentadas. Pode-se verificar também, que o número de amostras utilizadas apresentaram uma taxa de acerto de 100% para o reconhecimento, utilizando-se as mesmas amostras para ambas as etapas.

5.2.1.3.1 – Teste Aumentando o Número de Classes

A Tabela 5.4 apresenta os resultados considerando um número maior de classes e amostras. Foram utilizadas todas as amostras pertencentes às classes de 1 a 10. Para este teste também foram aplicadas as mesmas classes e amostras para a fase de treinamento e reconhecimento.

Tabela 5.4 – Base de Imagens 1 – Amostras Faciais Segmentadas – Teste Aumentando o Número de Classes: *Thresholds* e taxa de reconhecimento considerando-se as mesmas amostras para o treinamento e reconhecimento (10 classes contendo 10 amostras cada uma).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.84	0.84	0.97	0.97	0.97	0.97	0.97	0.97
Taxa de Acerto		59%						
Taxa de Erro		41%						

Neste caso, o número de amostras utilizadas são insuficientes para o treinamento resultando numa taxa de acerto de 59%.

5.2.1.3.2 – Teste usando Padrões de Treinamento distintos dos Padrões de Reconhecimento

Os testes apresentados até o momento foram realizados considerando as mesmas amostras para a fase de treinamento e reconhecimento da NEOPC.

A Tabela 5.5 apresenta os resultados considerando amostras diferentes de uma mesma classe para a fase de treinamento e reconhecimento. Neste teste foram considerados um total de cinco classes, de 1 a 5, sendo as cinco primeiras amostras (1 a 5) utilizadas para o treinamento e as cinco restantes (6 a 10) para o reconhecimento.

Tabela 5.5 – Base de Imagens 1 – Amostras Faciais Segmentadas - Teste usando Padrões de Treinamento distintos dos Padrões de Reconhecimento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 5 amostras cada uma para o treinamento e cinco amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.84	0.84	0.97	0.97	0.97	0.97	0.93	0.33
Taxa de Acerto		64%						
Taxa de Erro		36%						

Como pode ser verificado, a quantidade de amostras utilizadas, por classe, para a fase de treinamento dos testes apresentados nas Tabelas 5.3 e 5.5 são as mesmas. Porém, para os resultados apresentados na Tabela 5.3 foram utilizadas as mesmas amostras para o treinamento e reconhecimento, obtendo uma taxa de reconhecimento de 100%. Já os resultados apresentados na Tabela 5.5 foram obtidos utilizando amostras diferentes de uma mesma classe para a fase de treinamento e reconhecimento, apresentando uma taxa de reconhecimento de 64%. Entende-se que a taxa de acerto pode ser melhorada acrescentado um número maior de amostras para o treinamento.

5.2.1.3.3 – Teste usando Padrões de Reconhecimento de Classes distintas das Classes dos Padrões de Treinamento.

Os resultados apresentados abaixo, Tabela 5.6, demonstram a operacionalidade da NEOPC considerando, para o reconhecimento, amostras pertencentes a classes diferentes das utilizadas para o treinamento. Foram consideradas para a fase de treinamento as classes pertencentes ao intervalo de 1 a 5 com as cinco primeiras amostras (da esquerda para a direita) e para o reconhecimento as classes de 6 a 10 com as cinco primeiras amostras (1 a 5).

Tabela 5.6 – Base de Imagens 1 – Amostras Faciais Segmentadas - Teste usando Padrões de Reconhecimento de Classes distintas das Classes dos Padrões de Treinamento: *Thresholds* e taxa de erro considerando-se classes diferentes para o treinamento e reconhecimento (5 classes contendo 5 amostras cada uma).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.84	0.84	0.97	0.97	0.97	0.97	0.93	0.33
Taxa de Acerto		48%						
Taxa de Erro		52%						

A taxa de erro apresentada na Tabela acima, 52%, demonstra a semelhança dos padrões entre as amostras de classes distintas. Se a taxa de erro deve ser minimizada, implica-se na necessidade de uma quantidade maior de amostras para a realização do treinamento da rede.

Uma das principais dificuldades do reconhecimento facial é a diversidade de imagens faciais que um mesmo indivíduo pode ter. Nos testes apresentados até o momento, foram consideradas imagens faciais com diferentes orientações, tipos de expressões faciais e luminosidade.

5.2.1.3.4 – Teste usando Padrões de Imagens Faciais contendo Acessórios

A Tabela 5.7 apresenta os resultados com amostras que possuem, além de diferentes orientações, tipos de expressões faciais e luminosidade, diferentes tipos de acessórios, tais como, óculos, barba e bigode. As amostras utilizadas pertencem a classe de 11 a 15, sendo consideradas as cinco primeiras (1 a 5) tanto para a fase de treinamento como reconhecimento.

Tabela 5.7 – Base de Imagens 1 – Amostras Faciais Segmentadas - Teste usando Padrões de Imagens Faciais contendo Acessórios: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 5 amostras cada uma para o treinamento e cinco amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.84	0.84	0.97	0.97	0.97	0.97	0.93	0.33
Taxa de Acerto		60%						
Taxa de Erro		40%						

Os número de classes e amostras utilizadas nos resultados apresentados na Tabela 5.7, correspondem aos da Tabela 5.5. Porém, houve uma redução de 4% na taxa de reconhecimento para as amostras compostas por acessórios (Tabela 5.3). Durante o treinamento das amostradas compostas por acessórios, pode-se observar que muitos dos padrões treinados correspondem aos acessórios contidos nas amostras. Este fator ocasionou uma diminuição na variação dos padrões inter-classe. Por exemplo, um padrão treinado para uma determinada amostra em uma região como a barba, acaba se tornando característico a todas as amostras compostas por barba. Assim, a variação dos padrões detectados durante a fase de reconhecimento acaba se tornando pequena para a distinção das amostras. Neste caso, uma solução para melhorar a taxa de reconhecimento seria o aumento da quantidade de amostras por classe para a fase de treinamento e, talvez, a variação do *threshold*.

5.2.1.3.5 – Número de Fatores Treinados x Amostras Utilizadas

Os dados apresentados na Tabela 5.8 especificam a variação da quantidade de fatores treinados pelas camadas U_{S_3} , U_{S_4} e U_{S_5} , de acordo com o número de classes e amostras utilizadas. Em outras palavras, a Tabela abaixo especifica o número de planos celulares treinados em cada camada. Os resultados foram obtidos com as cinco primeiras amostras das classes 1,2,3,4 e 5.

Tabela 5.8 – Base de Imagens 1 – Amostras Faciais Segmentadas - Número de Fatores Treinados x Quantidade de Amostras sem Acessórios: número de fatores treinados pelas camadas U_{S_3} , U_{S_4} e U_{S_5} , de acordo com o número de classes e amostras utilizadas.

		Quantidade de Amostras por Classe											
		2			3			4			5		
		Camadas			Camadas			Camadas			Camadas		
		U_{S_3}	U_{S_4}	U_{S_5}	U_{S_3}	U_{S_4}	U_{S_5}	U_{S_3}	U_{S_4}	U_{S_5}	U_{S_3}	U_{S_4}	U_{S_5}
Quantidade de classes	2	2	4	3	3	6	4	4	7	6	4	10	8
	3	3	7	4	4	9	6	5	12	9	5	16	13
	4	4	7	7	5	13	9	6	19	14	8	38	32
	5	5	12	8	6	19	12	7	27	17	8	38	24

O número de fatores treinados por camadas apresentado na Tabela anterior, demonstra a operacionalidade da rede para amostras com diferentes orientações, tipos de expressões faciais e luminosidade. A Tabela 5.7 demonstra o número de fatores treinados para amostras com imagens contendo algum tipo de acessório (óculos, barba e bigode).

Tabela 5.9 – Base de Imagens 1 – Amostras Faciais Segmentadas - Número de Fatores Treinados x Quantidade de Amostras com Acessórios: número de fatores treinados pelas camadas U_{S_3} , U_{S_4} e U_{S_5} , de acordo com o número de classes e amostras utilizadas.

		Quantidade de Amostras por Classe											
		2			3			4			5		
		Camadas			Camadas			Camadas			Camadas		
		U_{S_3}	U_{S_4}	U_{S_5}	U_{S_3}	U_{S_4}	U_{S_5}	U_{S_3}	U_{S_4}	U_{S_5}	U_{S_3}	U_{S_4}	U_{S_5}
Quantidade de classes	2	4	5	3	4	7	6	4	9	8	4	11	9
	3	6	11	6	7	17	9	7	23	12	8	20	13
	4	7	14	8	9	21	12	9	26	16	11	30	18
	5	9	19	10	13	20	15	14	25	16	17	41	30

Através dos resultados apresentados nas Tabelas 5.7 e 5.8, pode-se concluir que as amostras contendo acessórios apresentam um número maior de fatores para treinamento. Isso pode ser verificado nas Tabelas apresentadas, através do incremento do número de fatores treinados por camadas. Por exemplo, o treinamento realizado considerando 2 classes com duas amostras cada uma, obteve para as amostras sem acessórios, 2 fatores treinados na camada U_{S_3} , 4 fatores na U_{S_4} e 3 fatores na U_{S_5} , enquanto que para as amostras contendo acessórios foram treinados 4 fatores na camada U_{S_3} , 5 fatores na U_{S_4} e 3 fatores na U_{S_5} .

5.2.2 – Base de Imagens 2

A base de imagens utilizada para os testes anteriores, base de imagens 1, possui um número limitado de amostras por classes (10 amostras para cada classe). Neste sentido, optou-se por adquirir uma base de imagens contendo um número maior de amostras para cada classe. Inicialmente foram analisadas algumas bases de imagens disponibilizadas via internet por diversas instituições, como por exemplo, a base de imagens da *University Purdue*. No entanto, deparou-se com o mesmo problema apresentado pela base de imagens utilizada anteriormente (Base de Imagens – 1): número restrito de amostras por classe. Assim, optou-se pelo desenvolvimento de uma nova base de imagens.

A base de imagens 2 foi desenvolvida em um espaço improvisado dentro do laboratório GAPIS (Grupo de Arquitetura e Processamento de Imagens e Sinais) do Departamento de Computação da Universidade Federal de São Carlos através de uma câmara fotográfica digital Kodak DC210A, adotando-se os seguintes critérios:

- luminosidade: todas as fotos foram realizadas considerando a mesma luminosidade;
- background: considerando-se a base utilizada para os primeiros experimentos (base 1), optou-se pelo uso do background preto afim de manter um padrão das bases de imagens utilizadas para os testes;
- Resolução: padrão;
- Qualidade das imagens: boa; e
- Formato das imagens: bmp.

A Figura 5.16 demonstra o espaço físico elaborado para aquisição das imagens.

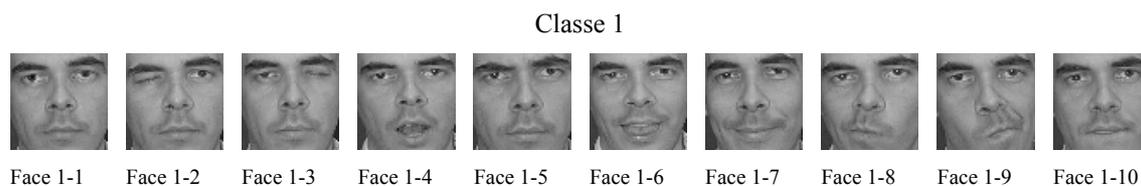


Figura 5.16 - Espaço físico elaborado para aquisição das imagens.

Na parede localizada de frente para os voluntários foram posicionados os pontos para as poses com diferentes ângulos. Os voluntários foram instruídos a permanecer com os olhos retos e rotacionar a cabeça em direção aos pontos especificados. Nesta base foram consideradas diferentes expressões faciais e ângulos, conforme a relação de poses apresentadas abaixo:

- expressão normal;
- olho direito fechado;
- olho esquerdo fechado;
- boca aberta;
- testa franzida;
- mostrando a língua;
- sorrindo;
- boca virada para o lado direito;
- boca virada para o lado esquerdo;
- mordendo os lábios;
- olho direito e esquerdo fechados;
- olho direito e esquerdo fechados e sorrindo;
- olho direito e esquerdo fechados e boca aberta;
- olho direito e esquerdo fechados e mordendo os lábios;
- olho direito e esquerdo arregalados;
- 5 poses em diferentes ângulos da direita para a esquerda;
- 5 poses em diferentes ângulos de cima para baixo; e
- 5 poses em diferentes ângulos movendo o pescoço da direita para a esquerda.

Foram consideradas 5 classes contendo 30 amostras cada uma, conforme exemplificado na Figura 5.17. De acordo com os resultados obtidos com a base de imagens 1, as amostras pertencentes à base de imagens 2 consistem em amostras faciais segmentadas.



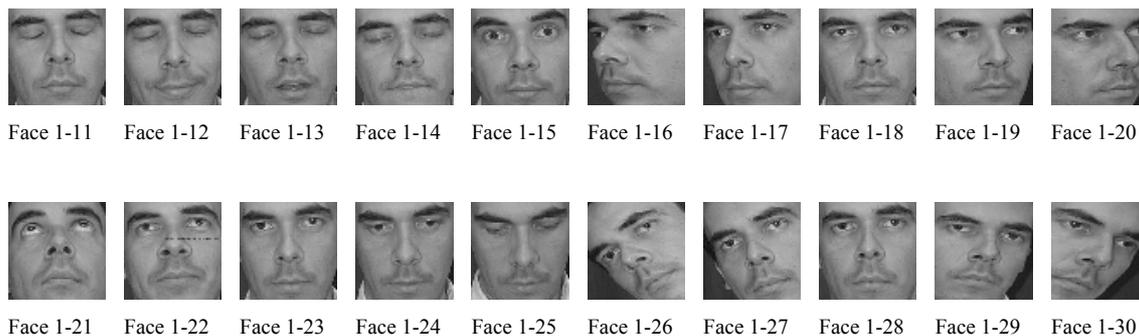


Figura 5.17 – Exemplo de amostras contidas na base de imagens 1.

Os testes foram realizados afim de verificar a operacionalidade da NEOPC considerando diferentes quantidades de classes e amostras para o treinamento. Os resultados obtidos são apresentados a seguir.

5.2.2.1- Testes Considerando 3 Classes e Diferentes Números de Amostras

Os testes foram realizados considerando as cinco primeiras amostras de cada classe para o reconhecimento e as demais amostras para o treinamento. Foram utilizadas as três primeiras classes apresentadas no Anexo 1 – Base de Imagens 2. As Tabelas apresentadas abaixo demonstram os resultados obtidos, variando o número de amostras utilizadas para o treinamento (Tabelas 5.10, 5.11, 5.12, 5.13 e 5.14).

→ 5 amostras

Tabela 5.10 – Base de Imagens 2 – 3 Classes e 5 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 5 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.85	0.33
Taxa de Acerto		66%						
Taxa de Erro		34%						

→ 10 amostras

Tabela 5.11 - Base de Imagens 2 – 3 Classes e 10 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 10 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.85	0.33
Taxa de Acerto		44%						
Taxa de Erro		56%						

→ 15 amostras

Tabela 5.12 - Base de Imagens 2 – 3 Classes e 15 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 15 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.80	0.33
Taxa de Acerto		44%						
Taxa de Erro		56%						

→ 20 amostras

Tabela 5.13 - Base de Imagens 2 – 3 Classes e 20 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 20 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.80	0.10
Taxa de Acerto		33%						
Taxa de Erro		67%						

→ 25 amostras

Tabela 5.14 - Base de Imagens 2 – 3 Classes e 25 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 25 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.80	0.10
Taxa de Acerto		33%						
Taxa de Erro		67%						

A Figura 5.18 apresenta a curva de erros em relação ao número de amostras utilizadas para o treinamento, obtida nos testes acima mencionados.

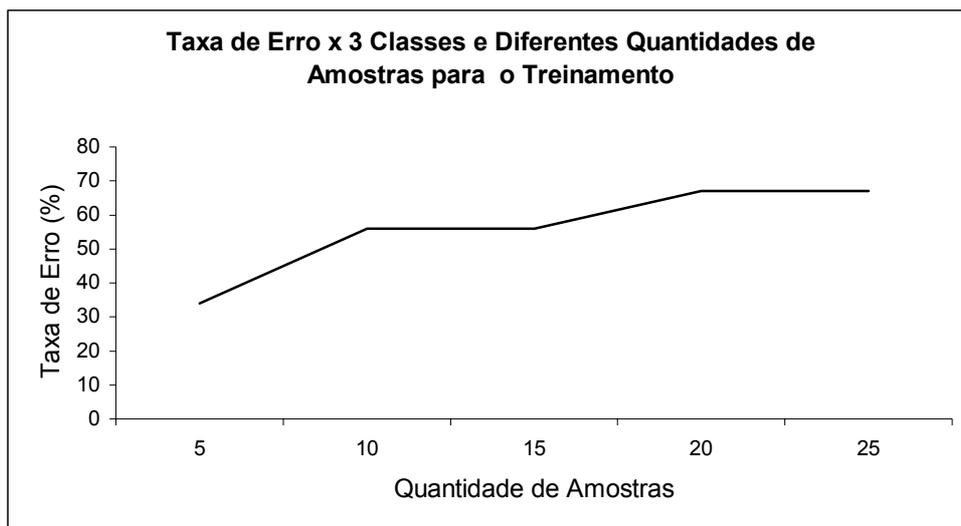


Figura 5.18 – Base de Imagens 2: taxas de erros apresentadas durante o reconhecimento, considerando-se 3 classes e diferentes quantidades de amostras.

5.2.2.2- Testes Considerando 5 Classes e Diferentes Números de Amostras

Neste item os testes foram realizados adotando-se o mesmo critério do teste anterior, ou seja, as cinco primeiras amostras de cada classe foram utilizadas para o reconhecimento e as demais amostras para o treinamento. Foram utilizadas todas as classes, cinco, apresentadas no anexo 1 – Base de Imagens 2. As Tabelas apresentadas abaixo demonstram os resultados obtidos (Tabelas 5.15, 5.16, 5.17 e 5.18).

→ 5 amostras

Tabela 5.15 - Base de Imagens 2 – 5 Classes e 5 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 5 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.85	0.33
Taxa de Acerto		44%						
Taxa de Erro		54%						

→ 10 amostras

Tabela 5.16 - Base de Imagens 2 – 5 Classes e 10 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 10 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.80	0.20
Taxa de Acerto		36%						
Taxa de Erro		64%						

→ 15 amostras

Tabela 5.17 - Base de Imagens 2 – 5 Classes e 15 Amostras para o Treinamento *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 15 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.95	0.80	0.20
Taxa de Acerto		24%						
Taxa de Erro		76%						

→ 20 amostras

Tabela 5.18 - Base de Imagens 2 – 5 Classes e 20 Amostras para o Treinamento *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 20 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.95	0.05	0.30
Taxa de Acerto		8%						
Taxa de Erro		92%						

→ 25 amostras

Tabela 5.19 - Base de Imagens 2 – 5 Classes e 25 Amostras para o Treinamento *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (5 classes contendo 25 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.95	0.05	0.30
Taxa de Acerto		8%						
Taxa de Erro		92%						

A Figura 5.19 apresenta a curva de erros relacionada aos testes apresentados acima. Já a Figura 5.20 demonstra as taxas de erros obtidas com os testes realizados com a base de imagens 2, variando o número de classes (3 e 5) e a quantidade de amostras para o treinamento.

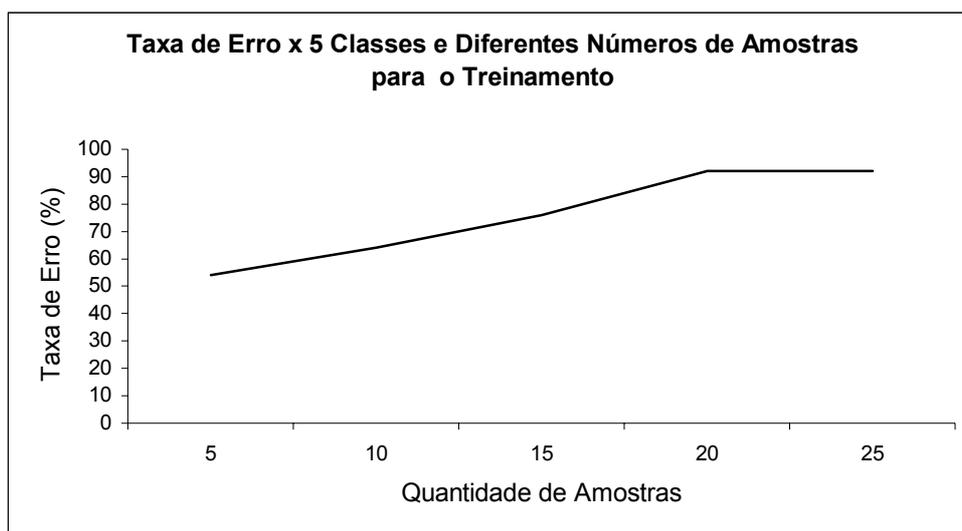


Figura 5.19 – Base de Imagens 2 – Taxas de erros apresentadas durante o reconhecimento, considerando-se 5 classes e diferentes quantidades de amostras.

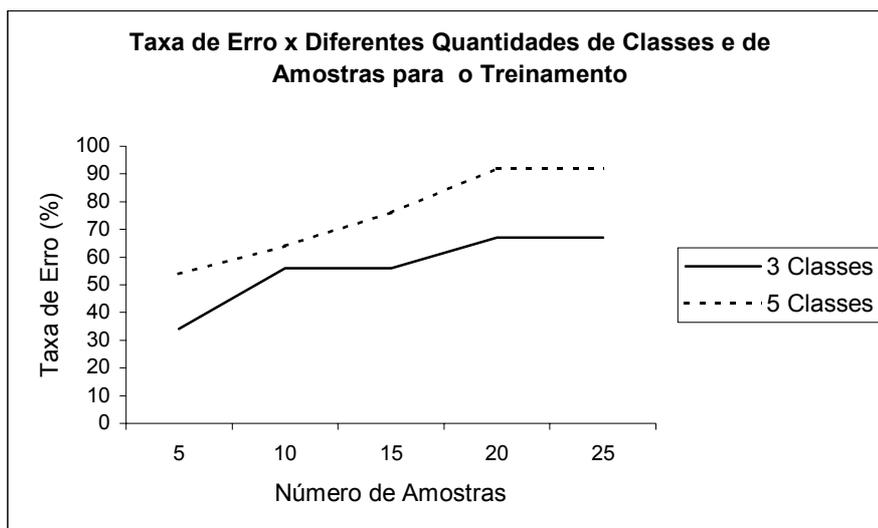


Figura 5.20 – Base de Imagens 2 - Analogia entre as taxas de erros obtidas variando o número de classes e amostras contidas na base de imagens 2

Conforme pode ser verificado, a taxa de acerto diminuiu em relação ao aumento do número de amostras utilizadas para o treinamento. Entende-se que os resultados obtidos são conseqüências da ampla variação dos padrões apresentadas pelas amostras.

5.2.3 – Base de Imagens 3

A base de imagens 3 foi desenvolvida considerando-se os mesmo critérios utilizados para o desenvolvimento da base de imagens 2, alterando apenas as poses das imagens. Conforme mencionado anteriormente, entende-se que as poses consideradas na base de imagem 2 dificultou o reconhecimento em decorrência da alta variação dos padrões contidos nas amostras. Assim, a base de imagens 3 foi desenvolvida considerando-se apenas diferentes variações de ângulos, conforme exemplificado na Figura 5.21. A base de imagens 3 é composta por 17 classes contendo 60 amostras cada uma, organizada através de 2 subconjuntos para cada classe: treinamento e reconhecimento. Assim, para cada classe foram consideradas 50 amostras para o subconjunto de treinamento (utilizadas para a fase de treinamento) e 10 amostras para o subconjunto de reconhecimento (utilizadas para a fase de reconhecimento), apresentadas no Anexo 2 – Base de Imagens 3.



Figura 5.21 – Base de Imagens 3 - Exemplos de amostras.

O primeiro teste foi realizado considerando as 3 primeiras classes e as 5 primeiras amostras de cada classe para o treinamento. Já para o reconhecimento foram consideradas as cinco primeiras amostras de cada classe apresentadas no subconjunto de amostras para reconhecimento no Anexo 2 – Base de Imagens 3. Os resultados estão apresentados na Tabela 5.19.

Tabela 5.20 – Base de Imagens 3 – 3 Classes e 5 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 5 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.80	0.33
Taxa de Acerto		47%						
Taxa de Erro		53%						

Analisando os resultados apresentados pelas Tabelas 5.5 (5.2.1.3.3 – Teste usando Padrões de Treinamento distintos dos Padrões de Reconhecimento) e 5.10 (5.2.2.1 – Testes Considerando 3 Classes e Diferentes Números de Amostras) verifica-se que ambas apresentaram uma taxa de acerto de 64% e 66%, respectivamente. Neste caso, acredita-se que a taxa de acerto da Tabela 5.19 tenha diminuído em decorrência da diferença dos ângulos das faces contidas nas amostras utilizadas para o treinamento e para o reconhecimento e em decorrência da variação de ângulos existentes nas faces utilizadas para o treinamento, ou seja, acredita-se que o uso de amostras com faces frontais possibilitam melhor desempenho da rede.

5.2.4 – Base de Imagens 4

Os resultados apresentados neste item correspondem as amostras contidas na Base de Imagens 3, organizadas em uma seqüência diferente. A nova seqüência foi realizada priorizando as amostras com faces pouco rotacionadas (pouca variação de ângulo), conforme exemplificado na Figura 5.22. Denominou-se esta base como sendo Base de Imagens 4.

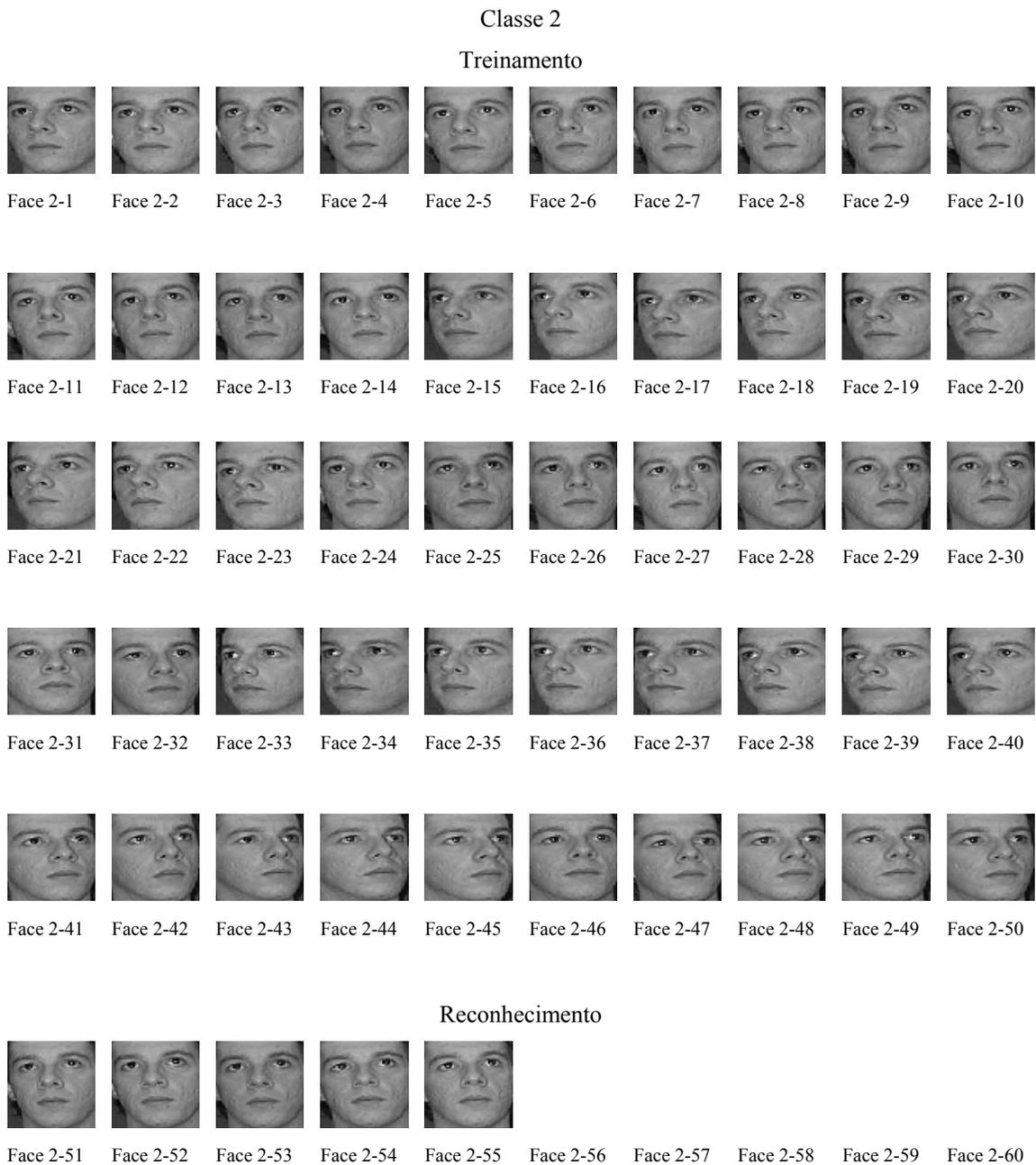


Figura 5.22 – Base de Imagens 4: Exemplos de amostras.

A seguir são apresentados os resultados obtidos variando a quantidade de classes (3, 6, 10 e 17 classes) e a quantidade de amostras utilizadas para o treinamento (5, 10, 15, 20, 30, 40 e 50 amostras). As classes e as amostras utilizadas para cada teste foram extraídas da base de imagens 4, considerando a forma em que estão organizadas. Por exemplo, para o teste com 3 classes e 5 amostras para o treinamento, foram consideradas as 3 primeiras classes contidas na base de imagens e as 5 primeiras amostras pertencentes ao subconjunto “treinamento” de cada classe, da esquerda para a

direita, em ordem crescente (Classe 1: Face 1-1, Face 1-2, Face 1-3....,Face 1-5; Classe 2: Face 2-1, Face 2-2, Face 2-3.... Face 2-5; Classe 3: Face 3-1, Face 3-2, Face 3-3....Face 3-5). Já para a fase de reconhecimento foram utilizadas as 5 primeiras amostras pertencentes ao subconjunto “reconhecimento” de cada classe. Os resultados obtidos estão apresentados nos próximos sub-itens.

5.2.4.1 – Testes Considerando 3 Classes e Diferentes Quantidades de Amostras Distintas para o Treinamento e Reconhecimento

→ 5 amostras

Tabela 5.21 – Base de Imagens 4 – 3 Classes e 5 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 5 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.95	0.20
Taxa de Acerto	80%							
Taxa de Erro	20%							

→ 10 amostras

Tabela 5.22 – Base de Imagens 4 – 3 Classes e 10 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 10 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.80	0.20
Taxa de Acerto		86%						
Taxa de Erro		14%						

→ 15 amostras

Tabela 5.23 – Base de Imagens 4 – 3 Classes e 10 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 10 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.86	0.05
Taxa de Acerto		73%						
Taxa de Erro		27%						

→ 20 amostras

Tabela 5.24 – Base de Imagens 4 – 3 Classes e 20 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 20 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.80	0.15
Taxa de Acerto		53%						
Taxa de Erro		47%						

→ 30 amostras

Tabela 5.25 – Base de Imagens 4 – 3 Classes e 30 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 30 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.75	0.20
Taxa de Acerto		46%						
Taxa de Erro		54%						

→ 40 amostras

Tabela 5.26 – Base de Imagens 4 – 3 Classes e 40 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 40 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.97	0.10	0.30
Taxa de Acerto		40%						
Taxa de Erro		60%						

→ 50 amostras

Tabela 5.27 – Base de Imagens 4 – 3 Classes e 50 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (3 classes contendo 50 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.97	0.05	0.30
Taxa de Acerto		33%						
Taxa de Erro		67%						

A Figura 5.23 apresenta a curva de erros em relação ao número de amostras utilizadas para o treinamento.

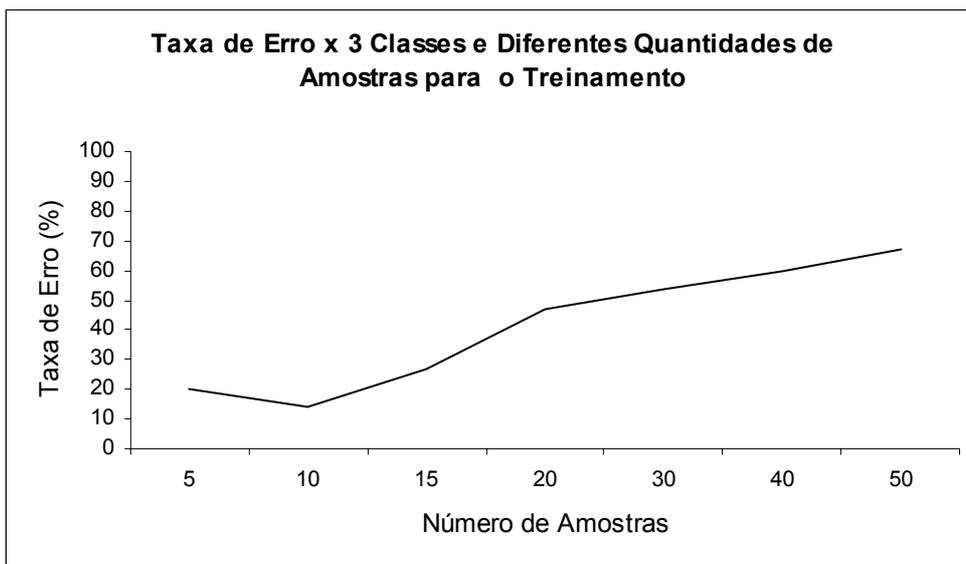


Figura 5.23 – Base de Imagens 4 – Taxas de erros apresentadas durante o reconhecimento, considerando-se 3 classes e diferentes quantidades de amostras.

5.2.4.2 – Testes Considerando 6 Classes e Diferentes Quantidades de Amostras Distintas para o Treinamento e Reconhecimento

→ 5 amostras

Tabela 5.28 – Base de Imagens 4 – 6 Classes e 5 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (6 classes contendo 5 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.90	0.20
Taxa de Acerto		53%						
Taxa de Erro		47%						

→ 10 amostras

Tabela 5.29 – Base de Imagens 4 – 6 Classes e 10 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (6 classes contendo 10 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.75	0.20
Taxa de Acerto		63%						
Taxa de Erro		37%						

→ 15 amostras

Tabela 5.30 – Base de Imagens 4 – 6 Classes e 15 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (6 classes contendo 15 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.45	0.30
Taxa de Acerto		30%						
Taxa de Erro		70%						

→ 20 amostras

Tabela 5.31 – Base de Imagens 4 – 6 Classes e 20 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (6 classes contendo 20 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.98	0.10	0.10
Taxa de Acerto		17%						
Taxa de Erro		83%						

→ 30 amostras

Tabela 5.32 – Base de Imagens 4 – 6 Classes e 30 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (6 classes contendo 30 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.97	0.98	0.98	0.97	0.15	0.33
Taxa de Acerto		17%						
Taxa de Erro		83%						

→ 40 amostras

Tabela 5.33 – Base de Imagens 4 – 6 Classes e 40 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (6 classes contendo 40 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.96	0.98	0.98	0.96	0.10	0.10
Taxa de Acerto		17%						
Taxa de Erro		83%						

→ 50 amostras

Tabela 5.34 – Base de Imagens 4 – 6 Classes e 50 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (6 classes contendo 50 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.95	0.98	0.98	0.95	0.80	0.15
Taxa de Acerto		7%						
Taxa de Erro		93%						

A Figura 5.24 apresenta a curva de erros em relação ao número de amostras utilizadas para o treinamento.

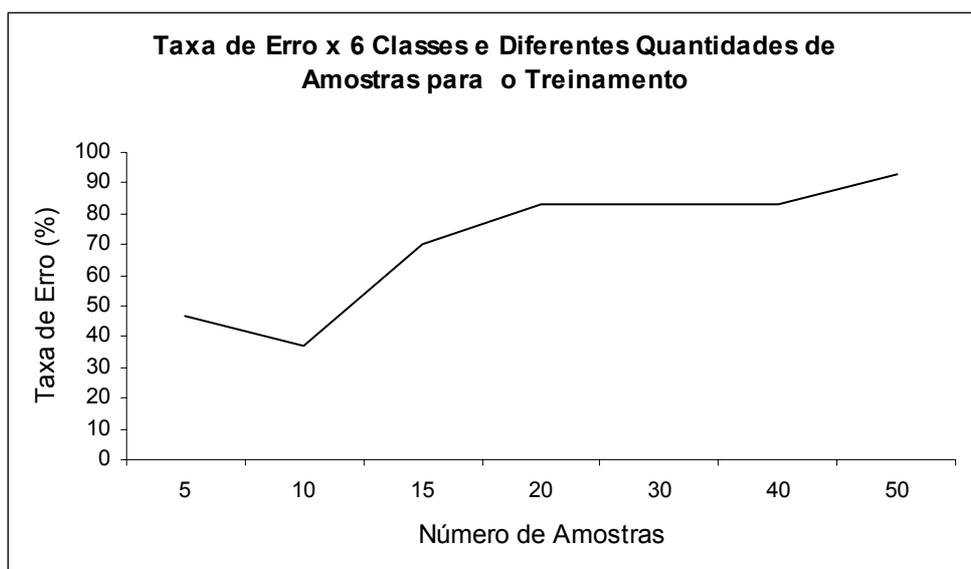


Figura 5.24 – Base de Imagens 4 – Taxas de erros apresentadas durante o reconhecimento, considerando-se 6 classes e diferentes quantidades de amostras.

5.2.4.3 – Testes Considerando 10 Classes e Diferentes Quantidades de Amostras Distintas para o Treinamento e Reconhecimento

→ 5 amostras

Tabela 5.35 – Base de Imagens 4 – 10 Classes e 5 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (10 classes contendo 5 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.98	0.98	0.98	0.97	0.10	0.33
Taxa de Acerto		10%						
Taxa de Erro		90%						

→ 10 amostras

Tabela 5.36 – Base de Imagens 4 – 10 Classes e 10 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (10 classes contendo 10 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.97	0.98	0.98	0.97	0.60	0.50
Taxa de Acerto		4%						
Taxa de Erro		96%						

→ 15 amostras

Tabela 5.37 – Base de Imagens 4 – 10 Classes e 15 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (10 classes contendo 15 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.96	0.98	0.98	0.95	0.70	0.35
Taxa de Acerto		12%						
Taxa de Erro		88%						

→ 20 amostras

Tabela 5.38 – Base de Imagens 4 – 10 Classes e 20 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (10 classes contendo 20 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.95	0.98	0.98	0.95	0.86	0.35
Taxa de Acerto		36%						
Taxa de Erro		64%						

→ 30 amostras

Tabela 5.39 – Base de Imagens 4 – 10 Classes e 30 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (10 classes contendo 30 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.95	0.98	0.98	0.95	0.86	0.35
Taxa de Acerto		10%						
Taxa de Erro		90%						

→ 40 amostras

Tabela 5.40 – Base de Imagens 4 – 10 Classes e 40 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (10 classes contendo 40 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.95	0.98	0.98	0.95	0.80	0.30
Taxa de Acerto		8%						
Taxa de Erro		92%						

→ 50 amostras

Tabela 5.41 – Base de Imagens 4 – 10 Classes e 50 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (10 classes contendo 50 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.95	0.98	0.98	0.95	0.80	0.30
Taxa de Acerto		8%						
Taxa de Erro		92%						

A Figura 5.25 apresenta a curva de erros em relação ao número de amostras utilizadas para o treinamento.

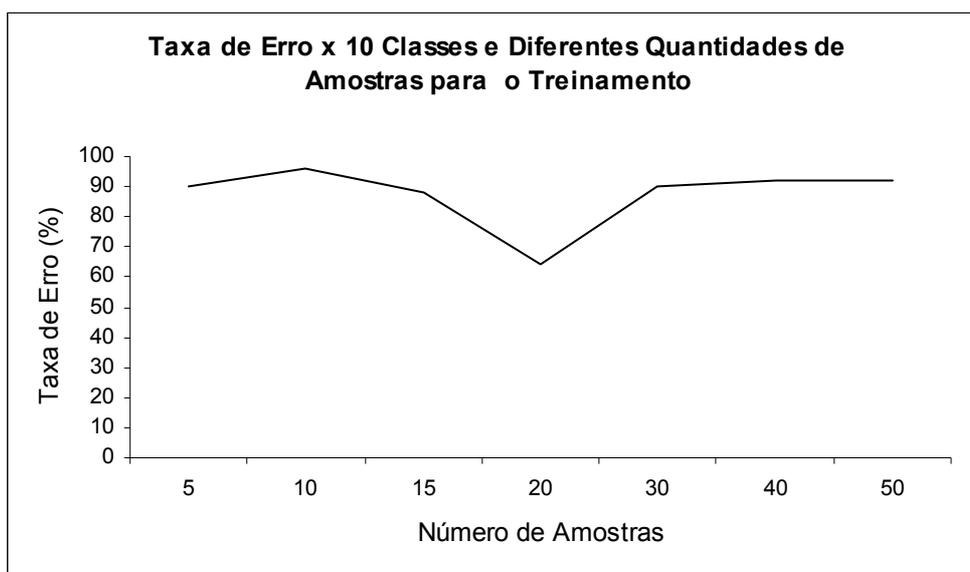


Figura 5.25 – Base de Imagens 4 – Taxas de erros apresentadas durante o reconhecimento, considerando-se 10 classes e diferentes quantidades de amostras.

5.2.4.4 – Testes Considerando 17 Classes e Diferentes Quantidades de Amostras Distintas para o Treinamento e Reconhecimento

→ 5 amostras

Tabela 5.42 – Base de Imagens 4 – 17 Classes e 5 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (17 classes contendo 5 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.97	0.98	0.98	0.94	0.10	0.20
Taxa de Acerto		6%						
Taxa de Erro		94%						

→ 10 amostras

Tabela 5.43 – Base de Imagens 4 – 17 Classes e 10 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (17 classes contendo 10 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.96	0.98	0.98	0.97	0.75	0.20
Taxa de Acerto		7%						
Taxa de Erro		93%						

→ 15 amostras

Tabela 5.44 – Base de Imagens 4 – 17 Classes e 15 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (17 classes contendo 15 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.95	0.95	0.95	0.95	0.90	0.35
Taxa de Acerto		8%						
Taxa de Erro		92%						

→ 20 amostras

Tabela 5.45 – Base de Imagens 4 – 17 Classes e 20 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (17 classes contendo 20 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.95	0.98	0.98	0.97	0.90	0.50
Taxa de Acerto		34%						
Taxa de Erro		66%						

→ 30 amostras

Tabela 5.46 – Base de Imagens 4 – 17 Classes e 30 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (17 classes contendo 30 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.94	0.94	0.98	0.97	0.90	0.50
Taxa de Acerto		4%						
Taxa de Erro		96%						

→ 40 amostras

Tabela 5.47 – Base de Imagens 4 – 17 Classes e 40 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (17 classes contendo 40 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.94	0.94	0.98	0.97	0.90	0.50
Taxa de Acerto		4%						
Taxa de Erro		96%						

→ 50 amostras

Tabela 5.48 – Base de Imagens 4 – 17 Classes e 50 Amostras para o Treinamento: *Thresholds* e taxa de reconhecimento considerando-se amostras diferentes para o treinamento e reconhecimento (17 classes contendo 50 amostras cada uma para o treinamento e 5 amostras para o reconhecimento).

<i>Thresholds</i>	Pontos de Controles		Treinamento			Reconhecimento		
	Borda	Linha	Us ₃	Us ₄	Us ₅	Us ₃	Us ₄	Us ₅
	0.82	0.82	0.94	0.94	0.98	0.97	0.90	0.50
Taxa de Acerto		4%						
Taxa de Erro		96%						

A Figura 5.26 apresenta a curva de erros em relação ao número de amostras utilizadas para o treinamento.

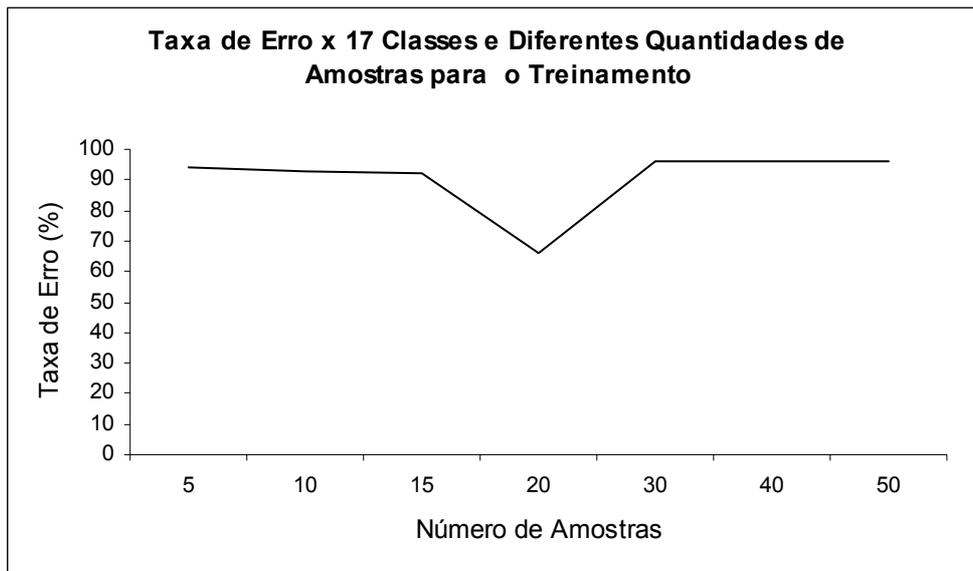


Figura 5.26 – Base de Imagens 4 – Taxas de erros apresentadas durante o reconhecimento, considerando-se 17 classes e diferentes quantidades de amostras.

A Figura 5.27 demonstra as taxas de erros obtidas com os testes realizados com a base de imagens 4, onde foram consideradas diferentes quantidades de classes e amostras para o treinamento.

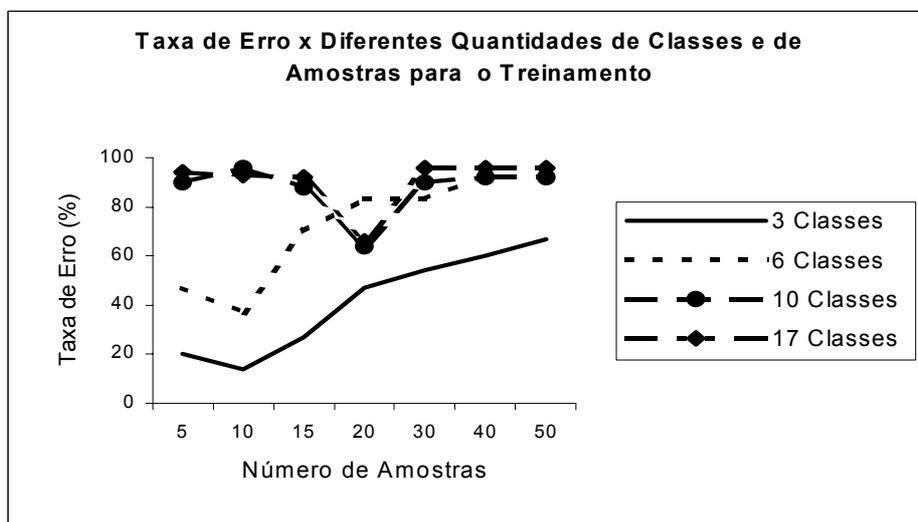


Figura 5.27 – Base de Imagens 4 - Analogia entre as taxas de erros obtidas variando o número de classes e amostras.

Conforme pode ser verificado no gráfico apresentado na Figura 5.27, independente da quantidade de classes consideradas, a NEOPC apresentou um comportamento onde se torna possível identificar uma quantidade “ótima” de amostra para o treinamento onde a

taxa de erro é menor. No entanto, o ideal seria que, de acordo com o acréscimo de amostras para o treinamento, a taxa de erro fosse diminuída. Isso não acontece, porque novos atributos de novas amostras aumentam o número de planos celulares dificultando o reconhecimento.

No próximo capítulo será apresentada a conclusão do trabalho realizado.

Capítulo VI

Conclusão

Neste trabalho foram apresentados alguns resultados para uma nova arquitetura de rede neural convolucional baseada no neocognitron: NEOPC.

Conforme pode ser verificado, o número de amostras contidas nas bases de imagens elaboradas são suficientes para verificar características do sistema implementado, como por exemplo, o aumento da taxa de erro em função do aumento do número de amostras usadas para o treinamento, sendo que é esperado exatamente o oposto. Uma das possíveis razões para tal é o uso de amostras com grandes variações de ângulo e de expressões faciais, dificultando a operacionalidade da NEOPC.

Com a grande variedade de ângulos e expressões o número de fatores (características) é elevado, e conseqüentemente, o número de planos celulares se torna grande, o que pode implicar na diminuição do desempenho da rede.

Uma possibilidade de melhoria no desempenho do sistema é a diminuição do threshold durante o treinamento e o reforço dos pesos das conexões de entradas em função das excitações a cada novo padrão de entrada durante o treinamento. Assim, com a diminuição do threshold é possível obter um número menor de novos fatores, enquanto que o reforço possibilita que os mesmos fatores sejam adaptados para diversos padrões de entrada. Na implementação realizada, os pesos das conexões foram mantidos fixos desde o primeiro treinamento do plano celular correspondente. Para cada novo fator extraído sempre eram acrescentados novos planos celulares.

Quanto à fase de reconhecimento, notou-se que o aumento do número de classes e amostras para o treinamento dificultou a determinação do valor de threshold, para

excitação dos neurônios, em decorrência da similaridade dos fatores treinados para cada classe.

Neste caso, entende-se que embora tenham sido utilizados os pontos de controle durante a fase de treinamento para a obtenção de fatores, o aumento do número de classes e amostras resultou em um conjunto de fatores similares para o treinamento.

Uma possível solução seria a utilização dos pontos de controle juntamente com a adaptação dos neurônios proposta por Oyen [Oye93], que torna o neurônio gradativamente menos sensível quando estimulado freqüentemente. Assim, torna-se possível a identificação de variações sutis apresentada por cada amostra para a determinação das características relevantes que as mesmas possuem [Abd95].

A seguir serão indicadas algumas sugestões para estudos futuros:

- reforço dos pesos para cada novo padrão de treinamento;
- adaptação dos neurônios para determinação dos fatores relevantes para cada classe; e
- especificação de um algoritmo para obtenção dos thresholds (limiar de excitação dos neurônios). Neste caso, o NEOPC seria capaz de auto-avaliar o melhor threshold a ser utilizado.

Capítulo VII

Bibliografias

- [Abd95] Abdi, Hervé; O'Toole, Alice J.; Valentim, Dominique, "Face Recognition", **The Handbook of Brain Theory and Neural Networks**, Michael A. Arbib, Massachusetts Institute of Technology, Cambridge, pp. 388-390, 1995.
- [Arb95] Arbib, Michael A., "Introducing the Neuron", **The Handbook of Brain Theory and Neural Networks**, Massachusetts Institute of Technology, Cambridge, pp. 4-11, 1995.
- [Adi97] Adini, Y; Moses, Y.; Ullman, S., "Face Recognition: The problem of Compensating for Changes in Illumination Direction", **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Vol. 19, N. 7, pp. 721-732, julho de 1997.
- [Bel97] Belhumeur, P.N.; Hespanha, J. P.; Kriegman, D.J., "Eigenfaces vs. Fisherfaces: Recognition Using Specific Linear Projection", **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Vol. 19, N. 7, pp. 711-720, julho de 1997.
- [Bia99] Bianchini, A. R.; Saito, J. H., **Uma Arquitetura de Redes Neurais Primária e Secundária baseada no Neocognitron**, Anais do WORKCOMP'99 – Workshop de Computação, pp. 63-68, outubro de 1999.

-
- [Bol00] Bolle, Ruud M.; Jain, Anil, “Biometrics: The Future of Identification”, **Computer – Innovate Technology for Computer Professionals**, IEEE Computer Society, pp. 46-49, fevereiro de 2000.
- [Cel99] Celli, Pedro Luis Fagá, **Um Sistema de Apoio à Identificação de Suspeitos com Reconhecimento Automático de Faces**, Dissertação de Mestrado, Universidade de São Paulo, abril de 1999.
- [Dau97] Daugman, John, “Face and Gesture Recognition: Overview”, **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Vol. 19, N. 7, pp. 675-676, julho de 1997.
- [Dou93] Dougherty, E.R., **Mathematical Morphology in Image Processing**, Marcel Dekker, 1993.
- [Dud73] Duda, R.O.; Hart, P.E., **Pattern Classification and Scene Analysis**, John Wiley & Sons, 1973.
- [Fau94] Fausett, Laurene, **Fundamentals of Neural Networks Architectures, Algorithms, and Applications**, Florida Institute of Technology, Prentice Hall International, 1994.
- [Fre92] Freiman, James A. and Skapura, David M., **Neural Networks Algorithms, Applications and Programming Techniques**, Addison-Wesley Publishing Company Inc, 1992.
- [Fri00] Frischholz, R.W.; Dieckmann U., BioID: A Multimodal Biometric Identification System, **IEEE – Computer Society**, pp. 64-68, fevereiro de 2000.
- [Fuk95] Fukushima, Kunihiko, “Neocognitron: A Model for Visual Pattern Recognition”, **The Handbook of Brain Theory and Neural Networks**, Michael A. Arbib, Massachusetts Institute of Technology, Cambridge, pp. 613-616, 1995.

-
- [Gon92] Gonzales, Rafael C.; Woods, Richard E., **Digital Image Processing**, Addison-Wesley Publishing Company, pp. 571-658, 1994.
- [Hay94] Haykin, Simon, **Neural Networks – A Comprehensive Foundation**, Prentice Hall International, 1994.
- [Hon98] Hong, Lin; Jain, Anil, “Integrating Faces and Fingerprints for Personal Identification”, **IEEE Transactions on Pattern and Machine Intelligence**, Vol. 20, N. 12, pp. 1295- 1307, dezembro de 1998.
- [Jai99] Jain, L. C.; Halici, U.; Hayashi, I.; Lee, S.B.; Tsutsui, S., **Intelligent Biometric Techiques in Fingerprint and Face Recognition**, CRC Press LLC, 1999.
- [Kua94] Kuan, C. M.; Hornik, K; White, H. A Convergence Result for Learning in Recurrent Neural Networks, **Neural Computation**, vol. 6, pp. 420-440, 1994.
- [Lan97] Lanitis, A.; Taylor, C.J.; Cootes, T.F., “Automatic Interpretation and Coding of Face Images Using Flexible Models”, **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Vol. 19, N. 7, pp. 743-756, julho de 1997.
- [Law97] Lawrence, Steve; Giles, C. Lee; Tsoi, Ah Chung; Back, Andrew D., “Face Recognition: A Convolutional Neural-Network Approach”, **IEEE Transactions on Neural Networks**, Vol. 8, N. 1, pp. 98-113, janeiro de 1997.
- [Led95] LeDoux, Joseph E.; Fellous, Jean-Marc, “Emotion and Computational Neuroscience”, **The Handbook of Brain Theory and Neural Networks**, Michael A. Arbib, Massachusetts Institute of Technology, Cambridge, pp. 356-359, 1995.

-
- [Lee97] Lee, Seong-Whan; Song, Hee-Heon, “A New Recurrent Neural-Network Architecture for Visual Pattern Recognition”, **IEEE Transactions on Neural Networks**, Vol. 8, N. 2, pp. 331- 339, março de 1997.
- [Mas89] Macarenhas, N.D.A.; Valesco, F.R.D., **Processamento Digital de Imagens**, 2ª Edição – IV Escola Brasileiro-Argentina de Informática, Universidade Católica de Santiago Del Estero, 1989.
- [Mor01] Morishima, Shigeo; “Face Analysis and Synthesis”, **IEEE Signal Processing Magazine**, Vol. 18, N. 3, pp. 26-34, maio de 2001.
- [Neu98] Neubauer, Claus, “Evaluation of Convolutional Neural Networks for Visual Recognition”, **IEEE Transactions on Neural Networks**, Vol. 9, N. 4, pp. 685-696, julho de 1998.
- [Ooy93] Ooyen, A.V.; Nienhuis, B., Pattern recognition in the neocognitron is improved by neuronal adaptation, **Biological Cybernetics**, vol. 70, Nº 1, pp. 47-56, junho de 1993.
- [Par97] Parker, J.R., **Algorithms For Image Processing and Computer Vision**, Jhon Wiley & Sons, 1997.
- [Pen00] Pentland, A.; Tanzeem, C., Face Recognition for Smart Environments, **IEEE – Computer Society**, pp. 50-55, fevereiro de 2000.
- [Phi00] Phillips, P.J.; Wilson, A.M.C.L.; Przybocki, M., An Introduction to Evaluating Biometric Systems, **IEEE – Computer Society**, pp. 56-63, fevereiro de 2000.
- [Row98] ROWLEY, Henry A.; BALUJA, Shumeet; TAKEO, Kanade. Neural Network-Based Face Detection, **IEEE – Transactions on Pattern Analysis and Machine Intelligence**, vol. 20, Nº 1, janeiro de 1998.

-
- [Sha97] Shang, Hung Lin; Sun, Yuan Kung, “Face Recognition/Detection by Probabilistic Decision-Based Neural Network ”, **IEEE Transactions on Neural Networks**, Vol. 8, N. 1, pp. 114-132, janeiro de 1997.
- [Sha00] Sharath, P.; Bolle, R.M.; Jain, A., Biometrics: The Future of Identification, **IEEE – Computer Society**, pp. 46-49, fevereiro de 2000.
- [Shi95] Shih, C.B.; Heang, P.C.; Jyh, S.L.; et al., Artificial Convolution Neural Network for Medical Image Pattern Recognition, **Special Issue Neural Networks**, vol. 8, N° 7/8, pp. 1201-1214, maio de 1995.
- [Sou95] Soulié, F. F., “Applications of Neural Networks”, **The Handbook of Brain Theory and Neural Networks**, Michael A. Arbib, Massachusetts Institute of Technology, Cambridge, pp. 94-98, 1995.
- [Sun98] Sung, K.K.; Poggio, T., Example-Based Learning for View-Based Human Face Detection, **IEEE – Transactions on Pattern Analysis and Machine Intelligence**, vol. 20, N° 1, janeiro de 1998.
- [Wec98] Wechsler, H.; Phillips, P.J.; Bruce, V.; Soulié, F.F.; Huang, T.S., **Face Recognition: From Theory to Applications**, Springer-Verlag Berlin Heidelberg New York, series F: Computer and Systems Sciences, vol. 163, 1998.