



UNIVERSIDADE FEDERAL DO AMAZONAS
FACULDADE DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**DETECÇÃO DE PONTOS FIDUCIAIS EM FACES
HUMANAS UTILIZANDO MÁQUINA DE
VETORES SUPORTE**

Luiz Eduardo Sales e Silva

Manaus – Amazonas

Maio de 2014

Luiz Eduardo Sales e Silva

**DETECÇÃO DE PONTOS FIDUCIAIS EM FACES
HUMANAS UTILIZANDO MÁQUINA DE
VETORES SUPORTE**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas.

Orientador: Prof. D.Sc. Waldir Sabino da Silva Júnior

Ficha Catalográfica
(Catalogação realizada pela Biblioteca Central da UFAM)

Silva, Luiz Eduardo Sales e
S586p Detecção de pontos fiduciais em faces humanas utilizando máquina
de vetores suporte / Luiz Eduardo Sales e Silva. - Manaus, 2014.
122f. il. color.
Dissertação (mestrado em Engenharia Elétrica) – Universidade
Federal do Amazonas.
Orientador: Dr. Prof. D.Sc. Waldir Sabino da Silva Junior
1. Engenharia elétrica 2. Estrutura de dados 3. Reconhecimento
de padrões I. Silva Júnior, Waldir Sabino (Orient.) II. Universidade
Federal do Amazonas III. Título

CDU 2007 621.3(043.3)

LUIZ EDUARDO SALES E SILVA

DETECÇÃO DE PONTOS FIDUCIAIS EM FACES HUMANAS
UTILIZANDO MÁQUINA DE VETORES SUPORTE

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Amazonas, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica na área de concentração Controle e Automação de Sistemas.

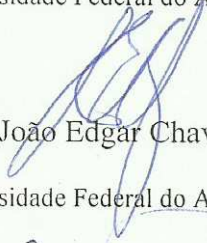
Aprovado em 14 de Maio de 2014.

BANCA EXAMINADORA



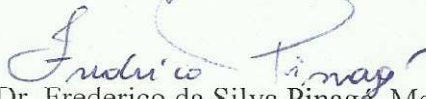
Prof. Dr. Waldir Sabino da Silva Junior, Presidente

Universidade Federal do Amazonas- UFAM



Prof. Dr. João Edgar Chaves Filho, Membro

Universidade Federal do Amazonas- UFAM



Prof. Dr. Frederico da Silva Pinagé, Membro

Fundação Centro de Análise, Pesquisa e Inovação Tecnológica- FUCAPI

Agradecimentos

- Primeiramente à minha mãe, Vanda Sales pelo apoio, esforço, dedicação e compreensão em todos os momentos e por ser a minha inspiração.
- À minha irmã Samara pelos momentos de distração.
- Aos meus tios, Bento, José e Messias por acreditarem no meu potencial.
- A Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM) pelo apoio financeiro.
- Ao meu orientador Waldir Sabino, pela amizade, compreensão, paciência, orientações e por sempre me ajudar nos momentos difíceis da execução desta dissertação, além de ser um profissional exemplar.
- Aos colegas de Laboratório de Visão Computacional pelos trabalhos conjuntos.
- Aos colegas de trabalho da Fundação Nokia pelo apoio.
- À Professora Ana Rita e aos Professores Edirley e Marden, por possibilitarem a conciliação entre os horários de trabalho e estudo, além de compreenderem a importância desta formação para o desenvolvimento profissional.
- Aos amigos Sandiego, Paulo César, Afonso e Fábio Lima pela amizade de longa data, pela compreensão, apoio e respeito.

Resumo da Dissertação apresentada à UFAM como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia Elétrica

DETECÇÃO DE PONTOS FIDUCIAIS EM FACES HUMANAS UTILIZANDO MÁQUINA DE VETORES SUPORTE

Luiz Eduardo Sales e Silva

Orientador: Waldir Sabino da Silva Júnior

Programa: Pós-Graduação em Engenharia Elétrica

Atualmente, o problema da detecção de pontos fiduciais em faces humanas vem recebendo crescente atenção da comunidade científica. Recentemente, podemos encontrar na literatura alguns sistemas desenvolvidos com o objetivo de realizar a detecção de pontos fiduciais. Alguns destes representantes utilizam classificadores SVM. Um dos problemas enfrentados por essa abordagem reside no fato de que o desempenho dos classificadores SVM é extremamente sensível à mudança de seus parâmetros. Esta é uma dificuldade que não é trivial de se explorar. Nesta dissertação, propomos um sistema de detecção de pontos fiduciais, que utiliza classificadores SVM, com o objetivo de investigar o desempenho dos classificadores para um conjunto de parâmetros pré-definidos. Com esta investigação, pretendemos descrever o comportamento do sistema para o conjunto de parâmetros e determinar qual é a combinação de parâmetros, e de seus valores, que produz o melhor desempenho possível para o sistema como um todo. O sistema proposto é complexo devido a grande quantidade de etapas envolvidas. A avaliação de desempenho do sistema proposto foi realizada para onze pontos fiduciais da face humana para duas bases de dados de imagens (*BioID* e *Feret*). Os resultados demonstram que o desempenho do sistema proposto é incrementado quando utilizamos a abordagem adotada. O desempenho do sistema é satisfatório quando comparado com os de sistemas similares.

Palavras-chave: reconhecimento de padrões, detecção de pontos fiduciais, máquina de vetores suporte, grid de parâmetros e C-SVC.

Abstract of Dissertation presented to UFAM as a partial fulfillment of the requirements for the degree of Master in Electrical Engineering

FIDUCIAL POINTS DETECTION IN HUMAN FACES USING SUPPORT
VECTORS MACHINE

Luiz Eduardo Sales e Silva

Advisor: Waldir Sabino da Silva Júnior

Department: Postgraduate in Electrical Engineering

Currently, the problem of facial fiducial points detection has received growing attention from the scientific community. Recently, we can find in the literature some systems developed aiming to perform the fiducial points detection. Some of these systems use SVM classifiers. One of the problems faced by this approach lies in the fact that the performance of SVM classifiers is extremely sensitive to changes in its parameters. This approach has a problem. The performance of SVM classifiers is extremely sensitive to changes in its parameters. This is a difficulty that is not trivial of exploring. In this dissertation, we propose a system for fiducial points detection, which uses SVM classifiers, with the aim of investigating the performance of the classifiers for a set of pre-defined parameters. With this research, we aim to describe the behavior of the system for the parameter set and determine which is the combination of parameters and their values, which produces the best possible performance for the system. The proposed system is complex due to the large number of stages involved. The performance evaluation of the proposed system was done for eleven fiducial points of the human face for two images database (*BioID* and *Feret*). The results show that the performance of the proposed system is increased when using the approach adopted. System performance is satisfactory when compared with similar systems.

Keywords: pattern recognition, fiducial points detection, support vectors machine, parameters grid and C-SVC.

Sumário

1	Introdução	1
1.1	Objetivos da Dissertação	3
1.2	Organização da Dissertação	4
2	Fundamentos Teóricos	6
2.1	Conceitos Iniciais	6
2.2	Ciclo de Projeto	7
2.3	Aprendizado de Máquina	9
2.4	Técnicas de Classificação e Classificadores	9
2.4.1	Discriminante Linear de Fisher	9
2.4.2	<i>Template Matching</i>	11
2.4.3	<i>Hidden Markov Models</i>	12
2.5	Trabalhos Relacionados a Detecção de Pontos Fiduciais	12
2.5.1	Detectores por Produto Interno	12
2.5.2	Detectores por SVM	14
2.5.3	<i>Facial Point Detection using Boosted Regression and Graph Models</i>	15
2.5.4	<i>Face Detection, Pose Estimation, and Landmark Localization in the Wild</i>	17
2.5.5	<i>Multiple Kernel Learning SVM and Statistical Validation for Facial Landmark Detection</i>	18
3	Máquina de Vetores Suporte (SVM)	21
3.1	Introdução	21
3.2	Tipos de SVM	23

3.2.1	Margens Rígidas	24
3.2.2	Margens Suaves	26
3.2.3	Formulações do SVM	28
3.2.4	SVM Não Linear	29
4	Sistema de Detecção de Pontos Fiduciais utilizando Classificadores C-SVC	31
4.1	Sistema de Detecção de Pontos Fiduciais utilizando Classificadores SVM	31
4.1.1	Pré-processamento	32
4.1.2	Procedimento de Treino	34
4.1.3	Procedimento de Teste	45
4.2	Investigações Propostas para o Sistema de Detecção de Pontos Fiduciais	47
4.2.1	Iterações do SVM	47
4.2.2	Definição das Classes dos Blocos de Treino	48
4.2.3	<i>Grid</i> de Parâmetros	51
5	Experimentos e Resultados	53
5.1	Base de Dados	53
5.2	Projeto dos Classificadores C-SVC	54
5.3	Métrica de Desempenho	57
5.4	Validação Cruzada	58
5.5	Experimentos	59
5.6	Resultados	61
5.6.1	Análise de Resultados: Conjunto de Imagens da Base de Dados <i>BioID</i>	62
5.6.2	Comparação entre Métodos: Conjunto de Imagens da Base de Dados <i>BioID</i>	77
5.6.3	Análise de Resultados: Conjunto de Imagens da Base de Dados <i>Feret</i>	79
5.6.4	Comparação entre Métodos: Conjunto de Imagens da Base de Dados <i>Feret</i>	86
5.6.5	Síntese de Resultados: Comparação <i>BioID</i> e <i>Feret</i>	88

6 Conclusão	92
6.1 Propostas para Trabalhos Futuros	93
Referências Bibliográficas	95
A Artigos Publicados	102
A.1 Artigos Diretamente Relacionados ao Tema	102
A.2 Outros Artigos	102
B Combinação de Parâmetros	104

Lista de Figuras

2.1	Diagrama em blocos de um sistema genérico de reconhecimento de padrões.	7
2.2	Diagrama em blocos que representa o ciclo de projeto de um sistema de reconhecimento de padrões.	8
3.1	Exemplo de conjunto de dados composto por duas classes, classe de quadrados e classe de triângulos.	22
3.2	Separação dos dados da classe de quadrados e da classe de triângulo.	23
3.3	Apresentação do SVM margem rígida com a separação da classe de quadrados e a classe de triângulos.	26
3.4	Apresentação do SVM margem suave com a separação da classe de quadrados e a classe de triângulos.	27
3.5	Exemplo do uso do <i>kernel</i> para a separação de dados não linearmente separáveis.	30
4.1	Diagrama em blocos para pré-processamento das imagens (treino e teste).	33
4.2	Diagrama em blocos para o procedimento de treino do sistema de detecção de pontos fiduciais.	35
4.3	Utilização da marcação manual fornecida pela base de dados como o centro do primeiro bloco da classe de positivos no ponto fiducial centro do olho esquerdo, sendo representado pelo símbolo ” + ”.	39
4.4	Exemplo da operação de mascaramento.	41
4.5	Representação da quantidade de blocos da classe de negativos antes da etapa de <i>Salt and Pepper</i>	41
4.6	Exemplificação da etapa <i>Salt and Pepper</i>	42

4.7	Exemplo de transformação matriz-vetor para o bloco da classe de positivos do ponto fiducial centro do olho esquerdo.	43
4.8	Diagrama em blocos para o procedimento de teste do sistema de detecção de pontos fiduciais.	46
4.9	Aplicação da janela deslizante B_z na ROI da etapa de teste.	47
5.1	Exemplos de imagens da base dados <i>BioID</i>	53
5.2	Exemplos de imagens da base dados <i>Feret</i>	54
5.3	Pontos fiduciais da face humana usada nos experimentos.	59
5.4	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>BioID</i> . Os resultados são apresentados para a região do olho direito são para os pontos fiduciais 0, 1 e 2. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,1.	63
5.5	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>BioID</i> . Os resultados são apresentados para a região do olho esquerdo são para os pontos fiduciais 3, 4 e 5. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,1.	64
5.6	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>BioID</i> . Os resultados são apresentados para a região do nariz são para os pontos fiduciais 6, 7 e 8. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,1.	65
5.7	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>BioID</i> . Os resultados são apresentados para a região da boca são para os pontos fiduciais 9 e 10. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,1.	66
5.8	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>BioID</i> . Os resultados são apresentados para a região do olho direito são para os pontos fiduciais 0, 1 e 2. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,01.	68
5.9	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>BioID</i> . Os resultados são apresentados para a região do olho esquerdo são para os pontos fiduciais 3, 4 e 5. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,01.	69

5.10	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>BioID</i> . Os resultados são apresentados para a região do nariz são para os pontos fiduciais 6, 7 e 8. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,01.	70
5.11	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>BioID</i> . Os resultados são apresentados para a região da boca são para os pontos fiduciais 9 e 10. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,01.	71
5.12	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>BioID</i> . Os resultados são apresentados para a região do olho direito são para os pontos fiduciais 0, 1 e 2. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,001.	73
5.13	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>BioID</i> . Os resultados são apresentados para a região do olho esquerdo são para os pontos fiduciais 3, 4 e 5. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,001.	74
5.14	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>BioID</i> . Os resultados são apresentados para a região do nariz são para os pontos fiduciais 6, 7 e 8. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,001.	75
5.15	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>BioID</i> . Os resultados são apresentados para a região da boca são para os pontos fiduciais 9 e 10. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,001.	76
5.16	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>Feret</i> . Os resultados são apresentados para a região do olho direito são para os pontos fiduciais 0, 1 e 2. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,1.	81
5.17	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>Feret</i> . Os resultados são apresentados para a região do olho esquerdo são para os pontos fiduciais 3, 4 e 5. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,1.	82

5.18	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>Feret</i> . Os resultados são apresentados para a região do nariz são para os pontos fiduciais 6, 7 e 8. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,1.	83
5.19	<i>Grid</i> de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados <i>Feret</i> . Os resultados são apresentados para a região da boca são para os pontos fiduciais 9 e 10. Neste <i>grid</i> de parâmetros, o valor de ξ_i igual à 0,1.	84

Lista de Tabelas

3.1	Problemas de otimização e restrições do método SVM.	29
5.1	Função para Treinamento do SVM linear.	55
5.2	Critérios de elaboração do classificador.	55
5.3	Numerações e descrições dos pontos fiduciais das bases de dados <i>BioID</i> e <i>Feret</i>	60
5.4	Desempenho do sistema proposto para 11 pontos usando base de dados <i>BioID</i>	80
5.5	Desempenho do sistema proposto para 11 pontos usando base de dados <i>Feret</i>	89
B.1	Combinações de parâmetros utilizados nos experimentos.	104

Capítulo 1

Introdução

Reconhecimento de Padrões é uma área de pesquisa que visa classificar dados em classes conhecidas por intermédio da extração de características, identificando-os como pertencentes a um conjunto [1–3]. Os métodos desenvolvidos na área de Reconhecimento de Padrões são utilizados em diversas áreas e linhas de pesquisa e desenvolvimento, como por exemplo: o processamento de sinais de voz, sensoria-mento remoto, visão computacional e identificação de assinaturas [4, 5].

Atualmente, os métodos desenvolvidos na área de Reconhecimento de Padrões para a detecção de pontos fiduciais da face humana vem recebendo grande atenção da comunidade científica. Os pontos fiduciais podem ser definidos como pontos de referência ou características salientes de um objeto [6–8]. Na face humana, os pontos fiduciais são caracterizados por representar o encontro de linhas ou centros de formas geométricas presentes na face. Por exemplo, podemos representar estas respectivas características pelos cantos da boca, centros dos olhos e a ponta do nariz. Geralmente, os sistemas de detecção de pontos fiduciais em faces humanas são aplicados no reconhecimento de expressões faciais e na estimação de pose [9].

Na literatura, é comum categorizarmos os métodos de detecção de pontos fiduciais de duas formas denominadas por *locais* e *globais* [10]. Em métodos locais, a principal característica reside no fato de que os pontos fiduciais são tratados e detectados de maneira individual, ou seja, informações complementares ao ponto fiducial não são utilizadas (por exemplo, os detectores são projetados somente para o canto esquerdo da boca). Os métodos globais são caracterizados por detectar os pontos fiduciais de modo conjunto utilizando modelos deformáveis, sendo menos

suscetíveis à variação de pose e iluminação em relação aos locais.

Recentemente, podemos encontrar na literatura alguns sistemas desenvolvidos com o objetivo de realizar a detecção de pontos fiduciais. Alguns destes representantes [11, 12] utilizam classificadores obtidos através da técnica por Máquina de Vetores Suporte (SVM) [2], [1], [13] e [14], que constitui um dos métodos estado-da-arte na área de Reconhecimento de Padrões, devido principalmente ao seu desempenho de classificação. Um dos problemas enfrentados por essa abordagem reside no fato de que o desempenho dos classificadores SVM é extremamente sensível à mudança de seus parâmetros [15–17]. Esta é uma dificuldade que não é trivial de se explorar.

Nesta dissertação, propomos um sistema de detecção de pontos fiduciais, que utiliza classificadores SVM, com o objetivo de investigar o desempenho dos classificadores para um conjunto de parâmetros pré-definidos. Com esta investigação, pretendemos descrever o comportamento do sistema para o conjunto de parâmetros e determinar qual é a combinação de parâmetros, e de seus valores, que produz o melhor desempenho possível para o sistema como um todo. O sistema proposto é dividido em um procedimento de treino e um de teste. Cada procedimento possui diversas etapas. Dessa forma, podemos comentar que a implementação em si do sistema é complexa. Em uma breve descrição, podemos comentar que o sistema possui: uma etapa de pré-processamento para os procedimentos de treino e teste, responsável por realizar a retirada do *background*, escalonamento e correção de iluminação; uma etapa definição de região de interesse que utiliza um modelo gaussiano à priori, responsável por definir a região com alta probabilidade de encontrarmos um ponto fiducial; para o procedimento de treino, temos uma etapa para seleção dos blocos de interesse formados por blocos cujo centro é o ponto fiducial (positivos) e por outros blocos da imagem (negativos). Por fim, ainda no procedimento de treino, os blocos (positivos e negativos) são utilizados para o projeto de um classificador SVM que gera o hiperplano responsável pela separação (ou categorização) entre blocos positivos e negativos. No procedimento de teste, as etapas são semelhantes, porém, vale ressaltar que o classificador é utilizado para categorizar os blocos.

Nesta dissertação, os parâmetros responsáveis pela separação dos blocos positivos e negativos estão associados ao próprio sistema de detecção de pontos fiduciais

e ao algoritmo de treino do classificador SVM utilizado no sistema. Nós utilizamos a formulação matemática SVM margem suave. Neste caso, precisamos definir quais valores de parâmetros, que combinados, produzirão o melhor desempenho na detecção do ponto fiducial. A definição destes valores é um processo custoso e que demanda elevada quantidade de experimentos, pois cada parâmetro apresenta uma faixa de valores que influencia sensivelmente o desempenho do sistema de detecção de pontos fiduciais. Para a estimação dos valores dos parâmetros do sistema e do algoritmo do SVM utilizamos a técnica de *grid* de parâmetros.

A avaliação de desempenho do sistema de detecção foi realizada para onze pontos fiduciais da face humana com o uso da validação cruzada. Em nossos experimentos, utilizamos duas bases de dados denominadas por *BioID* [18] e *Feret* [19], com respectivamente, 503 e 2004 imagens. Para comparação de desempenho, utilizamos os resultados apresentados em [20].

1.1 Objetivos da Dissertação

O objetivo principal desta dissertação é investigar o desempenho de um sistema de detecção de pontos fiduciais de faces humanas, que utiliza classificadores SVM, para um conjunto de parâmetros pré-definidos.

Objetivos Específicos da Dissertação

- (1) Implementar todos os procedimentos e etapas do sistema de detecção de pontos fiduciais de faces humanas na plataforma *OpenCV*;
- (2) Investigar o desempenho dos classificadores SVM para um conjunto de parâmetros pré-definidos.
- (3) Descrever o comportamento do sistema de detecção de pontos fiduciais de faces humanas para o conjunto de parâmetros e determinar qual é a combinação de parâmetros, e de seus valores, que produz o melhor desempenho.
- (4) Compreender os conceitos teóricos da formulação SVM;
- (5) Realizar os experimentos, utilizando validação cruzada, para 11 pontos fiduciais de faces humanas para as imagens das bases de dados *BioID* e *Feret*;

- (6) Explorar trabalhos diretamente relacionados com o tema proposto.

1.2 Organização da Dissertação

Esta dissertação está organizada conforme abaixo.

- No Capítulo 2, apresentamos os fundamentos teóricos utilizados na dissertação. Inicialmente, os conceitos fundamentais sobre a área de Reconhecimento de Padrões são abordados. Em seguida, técnicas de classificação em reconhecimento de padrões são discutidas. Por fim, alguns sistemas de detecção de pontos fiduciais são expostos. Este capítulo contempla o objetivo específico (6) além de fornecer uma breve fundamentação teórica a cerca das teorias envolvidas.
- No Capítulo 3, apresentamos a técnica de aprendizado denominada por Máquina de Vetores Suporte (SVM). Neste capítulo, discutimos os fundamentos da teoria do SVM. Em seguida, os tipos de SVM são apresentados. O SVM margem rígida, caracterizado por não permitir a existência de dados das classes de positivos e negativos entre as margens, é apresentado a seguir. Na sequência, apresentamos o SVM margem suave, onde é permitida a existência de padrões das classes de positivos e negativos entre as margens de separação. O tipo de SVM com *kernels* são apresentados com o objetivo de expor a separação dos dados de classes não linearmente separáveis através do mapeamento para um espaço de dimensão superior. Este capítulo contempla o objetivo específico (4).
- No Capítulo 4, expomos uma das contribuições desta dissertação. Neste, o sistema de detecção de pontos fiduciais é dividido nos procedimentos de treino e de teste. O capítulo é dedicado ao detalhamento destes procedimentos e de suas etapas. Por fim, apresentamos o conjunto de parâmetros investigados para avaliar o sistema de detecção de pontos fiduciais. Este capítulo contempla o objetivo específico (1) e parcialmente os objetivos específicos (2) e (3).
- O Capítulo 5 mostra os experimentos realizados. Neste capítulo, são abordadas as bases de dados, a métrica de desempenho e a validação cruzada, importantes

para a obtenção do desempenho. Por fim, os resultados são apresentados. Este capítulo contempla parcialmente o objetivo específico (5).

- No Capítulo 6, são apresentados as considerações finais e em seguida sugestões para trabalhos futuros.

Capítulo 2

Fundamentos Teóricos

2.1 Conceitos Iniciais

Reconhecimento de padrões é a área de pesquisa que visa a classificação de dados de entrada em classes conhecidas. Para classificarmos os dados em classes conhecidas é necessária a extração de características ou atribuições fundamentais e individuais do objeto em estudo [1,2]. Os dados de entrada, dependendo da aplicação, podem ser imagens, formas de onda, ou ainda quaisquer tipos de elementos que necessitem ser classificados, independente das medições obtidas para a classificação do objeto de estudo.

O termo *padrão* consiste em um conceito importante sendo definido por modelo composto por um conjunto de características. No conjunto de características que compõem o padrão, citamos as características extraídas do objeto e o rótulo que representa a que classe o padrão pertence. Os objetos são os elementos cujas características extraídas são classificadas e incluídas em determinada classe [1]. A classe representa o conjunto a que os objetos a serem classificados pertencem.

Com os padrões disponíveis é possível a elaboração de um classificador, alinhada com a possibilidade de indicação das classes, caracterizando os aprendizados de máquina supervisionado e não supervisionado, descrito posteriormente. Os padrões com a presença ou ausência das classes são os elementos responsáveis pela implementação da função de decisão. A função de decisão mapeará pontos do espaço de características distinguindo os objetos de cada classe em uma região definida pela própria função de decisão. Para a criação do classificador é executada uma

sequência de operações computacionais sobre o padrão, denominada algoritmo de reconhecimento.

No procedimento de classificação, objetos que serão classificados, ou ainda, objetos de entrada. A entrada do classificador utiliza como entrada as características para realizar a categorização do padrão.

Distintos sistemas de reconhecimento de padrões possuem em sua composição etapas semelhantes. Estas etapas são projetadas de acordo com os objetos de interesse. Semelhante ao apresentado em [2], por exemplo, um sistema genérico de reconhecimento de padrões pode ser dividido em seis etapas: aquisição, pré-processamento, segmentação, extração de características, classificação e pós-processamento. A Figura 2.1, apresenta o diagrama em blocos de um sistema de reconhecimento de padrões genérico.

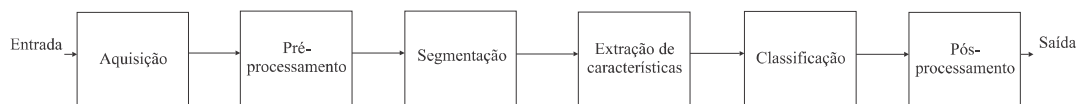


Figura 2.1: Diagrama em blocos de um sistema genérico de reconhecimento de padrões.

Na etapa de aquisição, um elemento sensor converte um tipo de energia em outro. Por exemplo, o áudio ou vídeo sendo convertido em um sinal elétrico. Entretanto, os sinais são compostos por informações indesejadas, como por exemplo, ruídos e interferências do ambiente de captação da informação. Deste modo, a eliminação de ruídos e interferências ocorre através do pré-processamento. Para as etapas posteriores, os sinais de interesse necessitam ser separados dos demais sinais, processo denominado de segmentação. Na etapa posterior, as características de interesse dos objetos segmentados são extraídos para elaboração do classificador. A partir de técnicas de classificação, as características são utilizadas para definir a que classe os objetos de interesse pertencem. Por fim, no pós-processamento os objetos classificados são analisados para a decisão final, informando a que classe pertence.

2.2 Ciclo de Projeto

O ciclo de projeto consiste nas etapas necessárias para o desenvolvimento de um sistema de reconhecimento de padrões [2]. A utilização do ciclo de projeto

facilita a organização do sistema de reconhecimento de padrões e possibilita o planejamento de cada especificidade das etapas. Para a elaboração de um sistema de reconhecimento de padrões são necessárias as seguintes etapas: coleta de dados, escolha de características, escolha do modelo, treinamento do classificador e avaliação. O diagrama que representa estas operações encontra-se na Figura 2.2.

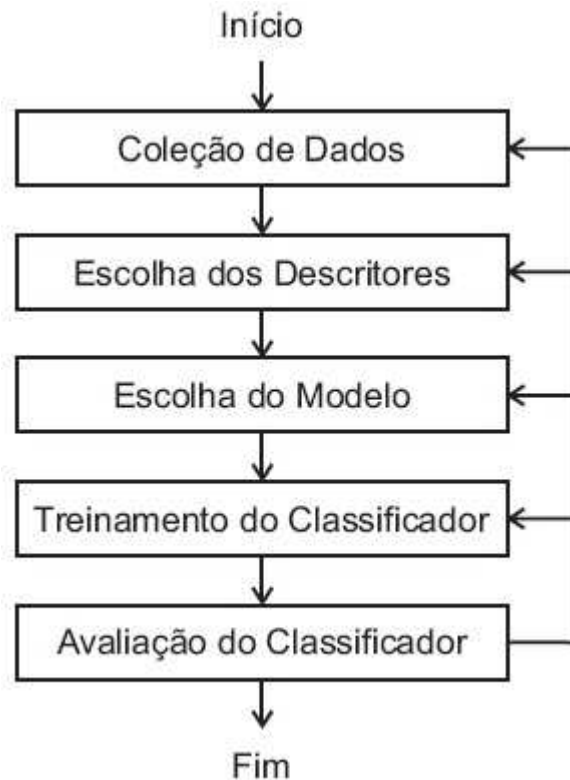


Figura 2.2: Diagrama em blocos que representa o ciclo de projeto de um sistema de reconhecimento de padrões [2].

Na Figura 2.2, diversos objetos podem ser selecionados, destes, inúmeras características podem ser extraídas, porém adota-se no projeto somente características que não sofram alterações conforme diversas modificações que possam existir nos padrões de interesse. A escolha do modelo estará relacionada ao classificador e a metodologia de classificação utilizada no sistema desenvolvido. Em seguida, o treinamento utiliza as características extraídas anteriormente e contribui para o ajuste do classificador para obter o melhor desempenho possível. Na avaliação do classificador, é necessária a utilização de uma base de dados de teste que deve ser diferente da utilizada na composição do classificador para verificar a eficiência do mesmo.

2.3 Aprendizado de Máquina

O aprendizado de máquina consiste na incorporação das características dos objetos para a elaboração do classificador. Existem três tipos de aprendizado de máquina, são eles: supervisionado, não supervisionado e semi-supervisionado [1]. No primeiro, os dados previamente segmentados pertencem a uma classe pré-definida, deste modo, informa-se ao classificador as características e os rótulos de classe para que os parâmetros de classificação sejam ajustados. Porém, nem sempre os dados estarão organizados, sendo desconhecidas as classes dos objetos. Neste caso, o processo de aprendizado denomina-se não supervisionado. No aprendizado não supervisionado, são informadas características extraídas dos objetos, sem indicar quaisquer informações a respeito das classes, com essas informações o algoritmo de treino busca por grupos representativos no conjunto de dados, ou seja, grupos que possuam características semelhantes. No procedimento de aprendizado semi-supervisionado são utilizados padrões com classes pré-definidas e classes que não são pré-definidas. Dessa forma, o algoritmo de treinamento realizará os procedimentos para o aprendizado supervisionado e o não-supervisionado para elaborar o classificador responsável pela separação das classes.

2.4 Técnicas de Classificação e Classificadores

Os classificadores possuem o objetivo de categorizar os objetos de interesse a uma determinada classe. Visando a solução dos problemas de classificação, diversas técnicas foram desenvolvidas para que os dados dos objetos de entrada sejam classificados corretamente. Cada técnica tem por objetivo obter o melhor desempenho durante o procedimento de classificação. A seguir, iremos abordar os seguintes classificadores: Discriminante Linear de Fisher [21, 22], *Template Matching* [23, 24] e *Hidden Markov Models* [25].

2.4.1 Discriminante Linear de Fisher

Desenvolvido por Fisher [1, 2], com o objetivo de dividir o espaço de características através de hiperplanos, o método Discriminante Linear de Fisher foi proposto para transformar sistemas multiclases em duas classes. Atualmente, atualmente o

mesmo é utilizado para redução da dimensionalidade ou no pós-processamento de classificadores mais complexos. O discriminante de Fisher pode ser obtido como segue: Sendo $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ um conjunto d-dimensional composto por N amostras. Estas amostras são divididas em dois conjuntos, \mathbf{X}_1 com n_1 amostras pertencentes à classe C_1 e \mathbf{X}_2 , com n_2 amostras pertencentes à classe C_2 . Seja ainda $\boldsymbol{\mu}_i$, a média, e $\boldsymbol{\Sigma}_i$, a matriz de covariância, das amostras da classe i , para $i = 1, 2$.

A separação do espaço de características pode ser feita através do hiperplano apresentado a seguir:

$$f(\mathbf{x}) = \mathbf{w}^t \mathbf{x} + b \quad (2.1)$$

onde \mathbf{w} é a direção normal ao hiperplano, o sobrescrito t representa o transposto do vetor e $|b|$ é a distância do hiperplano à origem ($\|\mathbf{w}\| = 1$). Projetando cada elemento \mathbf{x}_n no hiperplano definido por \mathbf{w} , um novo conjunto $\mathbf{Y} = \{y_1, \dots, y_N\}$ é gerado:

$$y_n = \mathbf{w}^t \mathbf{x}_n \quad (2.2)$$

O objetivo é encontrar a melhor classificação para \mathbf{Y} . Assim, dada a média μ_i do conjunto de dados, são definidas a matriz de espalhamento intra-classe,

$$\mathbf{S}_w = \sum_{i=1}^2 \frac{n_i}{n} \boldsymbol{\Sigma}_i, \quad (2.3)$$

e a matriz de espalhamento extra-classes,

$$\mathbf{S}_B = \sum_{i=1}^2 \frac{n_i}{n} (\boldsymbol{\mu}_i - \boldsymbol{\mu})(\boldsymbol{\mu}_i - \boldsymbol{\mu})^t, \quad (2.4)$$

Com as matrizes de espalhamento intra-classes e extra-classes é necessário relacioná-las para que o hiperplano que separe os dados das classes, seguindo o critério de menor separação intra-classes e maior separação extra-classes. A relação entre as matrizes de espalhamento pode ser apresentada pela função objetivo, definida a seguir:

$$\mathbf{J}(\mathbf{w}) = \frac{\mathbf{w}^t \mathbf{S}_B \mathbf{w}}{\mathbf{w}^t \mathbf{S}_w \mathbf{w}} \quad (2.5)$$

Se \mathbf{S}_w for inversível, a função objetivo tem solução na região em que \mathbf{Y}

apresenta menores distâncias intra-classe e o maiores distâncias entre classes:

$$\mathbf{w} = \mathbf{S}_w^{-1}(\mu_1 - \mu_2) \quad (2.6)$$

Por mais que tenhamos definido o hiperplano que separa as classes de acordo a Equação (2.6), o parâmetro b pode ser definido utilizando o critério de minimização do erro de classificação. A regra de classificação é dada por:

$$\begin{cases} \mathbf{x}_n \in C_1 & \text{se } \mathbf{w}^t \mathbf{x}_n + b > 0; \\ \mathbf{x}_n \in C_2 & \text{se } \mathbf{w}^t \mathbf{x}_n + b < 0. \end{cases} \quad (2.7)$$

2.4.2 *Template Matching*

O *template matching* é uma técnica de classificação utilizada para a identificação de partes de uma imagem, ou múltiplas imagens, a partir de um padrão de imagem determinado. Assim, a técnica *template matching* é considerada flexível por ser simples em diversas aplicações na localização objetos e por possuir baixo custo computacional [23, 24].

O *template matching* necessita que os objetos de classificação tenham características similares para que apresentem os melhores desempenhos. O funcionamento do *template matching* é definido conforme a seguir. Inicialmente, para a formação do objeto padrão no *template matching*, a dimensão do objeto padrão é definida e as suas respectivas características são extraídas e armazenadas em um vetor, sendo considerado o vetor padrão. Após a formação do padrão e a sua respectiva matriz, para que a classificação seja realizada, a matriz de características do objeto de interesse é extraída. Por fim, a matriz com características do candidato a classificação é classificada quando comparada a matriz de características padrão.

O *template matching* utiliza funções matemáticas para fazer a comparação entre a matriz analisada e a matriz padrão. Dentre as funções matemáticas, podemos citar o erro, o erro quadrático, correlação, covariância cruzada, a correlação cruzada e a correlação cruzada normalizada [23, 24]. Dentre estas funções matemáticas, a utilizada com frequência devido a eficiência dos resultados é a correlação cruzada [23, 24]. Nestes métodos, a classificação é indicada pela correlação entre as matrizes analisada e padrão. Esta correlação possui variação no intervalo de $[-1, 1]$. Onde

o valor 1 indicada máxima semelhança e o valor -1 indica máxima diferença entre a matriz padrão e a matriz analisada. A partir desta indicação de semelhança é definido se a matriz classificada pertence ou não a classe desejada.

2.4.3 *Hidden Markov Models*

O *Hidden Markov Models* (HMM) é um modelo estatístico utilizado para o reconhecimento padrões com características temporais, ou sejam, as classificações de objetos futuros dependem das classificações dos objetos atuais. Inicialmente desenvolvido para o reconhecimento de sinais de voz [1, 25].

O HMM possui seus princípios baseados na cadeia de *Markov*. Esta caracterização está relacionada ao fato de que os objetos classificados no estado futuro dependem da classificação do objeto no estado atual [1, 25]. Deste modo, o HMM possui alto custo computacional.

As aplicações que utilizam o HMM são modeladas utilizando-se o conceito de máquina de estados. Na máquina de estados, primeiramente, uma quantidade finita de estados é definida. Em seguida, para cada transição entre estados existe uma probabilidade associada. Por fim, a cada teste realizado, a probabilidade é definida, determinando a que estado o objeto de teste pertence.

2.5 Trabalhos Relacionados a Detecção de Pontos Fiduciais

2.5.1 Detectores por Produto Interno

No sistema de detecção de pontos fiduciais proposto por [26], na etapa de classificação foi utilizada a cascata de classificadores IPD (Detector por Produto Interno) baseado em filtros de correlação [1–3]. O IPD é caracterizado por incorporar as informações estatísticas do conjunto de treino no projeto do detector com baixo custo computacional. Além disso, o mesmo é caracterizado por ser tolerante a pequenas variações dos padrões. O projeto do classificador IPD é caracterizado por um vetor capaz de discriminar um padrão através de um produto interno com um sinal desconhecido. A formulação do IPD é obtida da seguinte forma: dado um

problema com N classes, em que as classes são: $\{A_1, \dots, A_n, \dots, A_N\}$, o objetivo é detectar amostras pertencentes à A_n e rejeitar as demais. Assim, é necessário obtermos um detector \mathbf{h}_{A_n} de tal forma que o produto interno com um sinal desconhecido \mathbf{x} tenha um valor elevado se $\mathbf{x} \in A_n$ e pequeno no caso contrário. Matematicamente, podemos escrever da seguinte forma:

$$\mathbf{h}_{A_n}^t \mathbf{x} = c, \quad (2.8)$$

onde idealmente $c = 1$ se $\mathbf{x} \in A_n$ e $c = 0$ caso contrário. Definindo o erro quadrático como

$$\|e\|^2 = (\mathbf{h}_{A_n}^t \mathbf{x} - c)^2, \quad (2.9)$$

a solução de mínimos quadrados é

$$\mathbf{h}_{A_n} = (E[\mathbf{x}\mathbf{x}^t])^{-1} E[\mathbf{x}c]. \quad (2.10)$$

Podemos escrever os termos $E[\mathbf{x}\mathbf{x}^t]$ and $E[\mathbf{x}c]$ em função do conjunto de treino conforme a seguir:

$$E[\mathbf{x}\mathbf{x}^t] = \sum_{i=1}^N p(A_i) \mathbf{R}_{A_i}, \quad (2.11)$$

$$E[\mathbf{x}c] = p(A_n) \boldsymbol{\mu}_{A_n}, \quad (2.12)$$

onde \mathbf{R}_{A_i} é a matriz de autocorrelação das amostras de treino de A_i , $\boldsymbol{\mu}_{A_i}$ é a respectiva média e $p(A_i)$ é a probabilidade da amostra de A_i . Substituindo as Equações (2.11) e (2.12) na Equação (2.10) temos:

$$\mathbf{h}_{A_n} = \left(\sum_{i=1}^N p(A_i) \mathbf{R}_{A_i} \right)^{-1} p(A_n) \boldsymbol{\mu}_{A_n}. \quad (2.13)$$

Na equação acima, a expressão $\sum p(A_i) \mathbf{R}_{A_i}$ é a soma ponderada das matrizes de correlação de todas as classes. Neste caso, os pesos são as probabilidades das classes. O termo deve ser invertido, implicando que o número de amostras diferentes deve ser maior que o tamanho dos vetores. Usado em casos particulares como este de detecção de pontos fiduciais. Na expressão $p(A_n) \boldsymbol{\mu}_{A_n}$, a média da classe desejada

é ponderada pela sua probabilidade.

Na cascata de classificadores IPD, as imagens detectadas tem seus dados inseridos na sequência de classificadores, a cada estágio, e na saída do primeiro estágio, os pontos considerados positivos (pontos classificados como pontos fiduciais) são encontrados e indicados aos blocos subsequentes e os pontos classificados como negativos são rejeitados. Ao final da cascata de classificadores IPD, a quantidade de pontos classificados como positivos será reduzida e para definir se o candidato é um ponto fiducial. Os pontos classificados como positivo tendem a se agrupar em uma região, o pós-processamento por intermédio de uma mediana vetorial faz com que o sistema indique somente um ponto fiducial na saída, representando a detecção do ponto fiducial.

Em [26], a detecção de 11 pontos fiduciais é implementada utilizando a base de dados *BioID* [18]. Para os procedimentos de treino e teste, a base de dados foi dividida em 7 subconjuntos de mesmo tamanho para a realização de 7 experimentos. Neste procedimento de treino e teste é aplicada a validação cruzada para que o desempenho do sistema de detecção de pontos fiduciais utilizando a cascata de classificadores IPD seja obtido. Assim, podemos sintetizar que o sistema proposto apresenta bons resultados.

2.5.2 Detectores por SVM

No trabalho desenvolvido em [11], um sistema de detecção de pontos fiduciais, semelhante ao que será apresentado nesta pesquisa, foi desenvolvido. Em [11], o sistema de detecção de pontos fiduciais é composto por etapas de segmentação da face, correção de iluminação, escalonamento das imagens e das marcações manuais fornecidas pelas base de dados, definição da região de interesse, detecção dos pontos fiduciais e por fim, uma etapa de pós-processamento.

As etapas de segmentação da face, correção de iluminação, escalonamento e definição da região de interesse em [11] são agrupadas no procedimento denominado pré-processamento. O procedimento de pré-processamento é definido conforme a seguir. Inicialmente, as imagens fornecidas pelas bases de dados são indicadas ao sistema para que ocorra o enquadramento da face e a eliminação do *background* utilizando o detector de faces *Viola Jones* [27–30]. A seguir, a imagem enquadrada é

submetida a correção de iluminação [31] para ajustar a não uniformidade na iluminação das imagens da base de dados. Em seguida, o escalonamento é aplicado para que as imagens sejam ajustadas as dimensões definidas pelos autores, o objetivo deste ajuste é que as imagens e as marcações manuais fornecidas para o procedimento de treino possuam dimensões iguais. Por fim, a região com alta probabilidade de conter os pontos fiduciais é definida, ou seja, a região utilizada nos procedimentos de treino e teste para cada ponto fiducial trabalhado é delimitada. Neste trabalho, os autores utilizam um SVM com *kernel* Gaussiano como técnica de classificação dos blocos extraídos da região de interesse.

O sistema de detecção é constituído para detectar 11 pontos fiduciais. O projeto do classificador utiliza os bancos de dados *BioID* [18] e *FERET* [19]. Para determinar os parâmetros do SVM com *kernel* gaussiano, é feito o *grid search* para que os melhores valores sejam encontrados. De maneira semelhante à aplicada no sistema de detecção de pontos fiduciais proposto por [26], o conjunto de treino e teste é dividido em 7 partes e a validação cruzada é empregada.

O sistema de detecção de pontos fiduciais em faces humanas proposto por [11] apresenta semelhanças quando comparado ao sistema proposto nesta dissertação. Na Seção 4.2, as diferenças entre os sistemas são apresentadas. Dentre as diferenças, podemos citar: o tipo de *kernel*, o procedimento de seleção dos blocos das classes de positivos e negativos e o procedimento de *grid* de parâmetros para validação cruzada.

2.5.3 Facial Point Detection using Boosted Regression and Graph Models

No trabalho desenvolvido em [32] é apresentado um sistema de detecção de pontos fiduciais usando *Boosted Regression* e *Graph Models*. Neste sistema, a técnica de regressão do algoritmo SVM, o SVR (*Support Vector Regression*) [15], é combinada com a técnica de campos aleatórios de *Markov* [1] para a detecção de pontos fiduciais. Este sistema utiliza três fontes de informação para detecção dos pontos fiduciais que são combinadas para modelar suas posições relativas. Deste modo, podemos dividir o sistema nas etapas de probabilidade *a priori*, na predição por regressão e por fim, na localização relativa dos pontos fiduciais.

Na etapa de definição da região de probabilidade *a priori*, o objetivo é delimitar a região de interesse para a detecção de pontos fiduciais. Inicialmente, as imagens são enquadradas utilizando o detector de faces *Viola Jones* [27–30] e são redimensionadas para as dimensões de 200×280 . Na sequência, uma normalização de iluminação global é empregada às imagens redimensionadas para que os efeitos da variação de iluminação sejam reduzidos. Por fim, os 22 pontos fiduciais de interesse são divididos em estáveis e instáveis. Esta divisão de pontos em estáveis e instáveis é associada a capacidade de mudança de localização na face entre alterações de expressões faciais e fala. Os primeiros pontos detectados são os estáveis. Eles auxiliam na detecção dos pontos fiduciais considerados instáveis. Os valores de média e desvio-padrão da base de dados são obtidas a partir dos pontos estáveis, resultando no modelo gaussiano *a priori*. A partir da região de interesse formada pelos pontos estáveis, a rotação das imagens é obtida e em seguida, este valor é utilizado para corrigir a rotação de cada imagem da base de dados. A definição da região de interesse para os pontos instáveis é realizada de modo semelhante a realizada nos pontos estáveis, utilizando como referência as coordenadas fornecidas pela base de dados.

Na etapa seguinte, a cada ponto fiducial é atribuído um vetor para detecção. A detecção de pontos fiduciais é definida a partir da decomposição em dois problemas de regressão separados. O primeiro problema de regressão é utilizado para encontrar o ângulo do vetor e o segundo problema é utilizado para encontrar o comprimento do vetor. Os autores utilizam o SVR com *kernel* Gaussiano devido ao bom desempenho. A obtenção dos parâmetros ótimos do SVR com *kernel* Gaussiano é realizada em um *loop* de validação cruzada independente da base de dados de teste. Os regressores fornecem uma localização precisa do ponto fiducial. Além disso, os erros de localização são reduzidos a partir do uso de recursivas iterações da validação cruzada.

Na etapa de localização relativa, as posições espaciais entre os pontos fiduciais são relacionadas levando em consideração as restrições atropométricas [32] como referência para determinar a localização do ponto fiducial. O uso das restrições espaciais é implantada para melhorar a estimativa de acerto dos pontos, impedindo a combinação de pontos fiduciais inviáveis, ou seja, combinação entre pontos estáveis

e pontos instáveis. Para cada ponto fiducial é indicado um vetor. A correlação entre dois pontos fiduciais é obtida pela relação entre os vetores de cada ponto fiducial é realizada considerando o ângulo e o comprimento dos vetores. Em seguida, os vetores de cada ponto fiducial são modelados utilizando probabilidade conjunta usando campos aleatórios de *Markov*. Assim, de acordo com as combinações das posições relativas de cada ponto fiducial para o ponto de referência (baseada na posição da cabeça) é identificado se o ponto fiducial desejado foi estimado corretamente ou erroneamente.

Após as três etapas de detecção de pontos fiduciais apresentadas anteriormente, os autores implementam um algoritmo que utiliza as localizações com alta probabilidade de conter o ponto fiducial para gerar as características para os filtros de *Haar* que em seguida serão utilizados pelos regressores para fazer a predição para cada ponto fiducial. Os pontos candidatos são analisados pela rede de *markov* e classificados pelos regressores. O ponto que possuir maior taxa de predições corretas por regressor é denominado ponto fiducial detectado.

2.5.4 Face Detection, Pose Estimation, and Landmark Localization in the Wild

No trabalho desenvolvido em [33] é apresentado um modelo unificado para a detecção de face, estimação de pose e localização de pontos fiduciais em imagens. Neste modelo unificado é aplicado o modelo de árvores de decisão. Os pontos fiduciais da face humana são mapeados e denominados partes, e em seguida são utilizadas combinações globais para capturar as mudanças topológicas de acordo com o ponto de vista da imagem. Deste modo, para a identificação das informações de deformação da imagem e da estrutura tridimensional (detecção de face, estimação de pose e localização de pontos fiduciais) é utilizado no sistema o modelo de árvores dividida em partes.

Inicialmente, todos os pontos fiduciais da face humanas são modelados como uma parte do sistema e as informações globais são utilizadas para capturar mudanças de face na imagem para identificar as deformações para um ponto de vista simples. O modelo de parte estruturada em árvore apresenta o somatório das características em evidência extraídas de acordo com o *template* da expressão da face por ponto fiducial,

este é associado a uma combinação global e a localização por pixel. Neste somatório de aparências, os vetores de características são extraídos pixel a pixel utilizando o descritor *Histogram of Oriented Gradients (HOG)* [34]. Em seguida, pontuações das combinações de regiões faciais são adquiridas pela diferença dos valores dos pixels adjacentes. Assim, as relações entre pontos fiduciais são encontradas e as mudanças nas expressões faciais são mapeadas com facilidade, obtendo o formato da face. O modelo de forma é comparado a um modelo espacial que utiliza probabilidade gaussiana e os autores identificam que o modelo proposto apresenta baixa variação nos valores dos autovetores, caracterizando a identificação das variações faciais.

Na etapa seguinte, as informações de aparência e de formato da face são combinadas a um termo escalar para identificar o ponto de vista que será analisado. Nesta configuração, todas as combinações possíveis de variação são enumeradas e para cada combinação são encontradas as melhores combinações dos pontos fiduciais, ou das partes. As melhores combinações são obtidas pela maximização interna das pontuações entre pontos fiduciais.

Por fim, para aplicar a aprendizagem ao modelo, os autores assumem um cenário supervisionado, onde são indicadas as imagens positivas com pontos fiduciais e a combinação de rótulos, assim como as imagens negativas são indicadas. No aprendizado, o formato da face e os parâmetros de aparência são discriminados usando um *framework* de predição. Neste *framework*, os exemplos positivos e negativos são indicados e uma função objetivo flexível a combinação flexível de partes é aplicada. As pontuações referentes aos exemplos positivos devem ser maior que 1, enquanto os exemplos negativos possuem pontuações menores que -1 . Com esta fase de aprendizagem, os autores detectam a face, estimam a pose e localizam o ponto fiducial em imagens.

2.5.5 Multiple Kernel Learning SVM and Statistical Validation for Facial Landmark Detection

No trabalho desenvolvido em [35] é apresentado um método robusto para a detecção de pontos fiduciais em imagens com faces com expressões. Neste método, é introduzido um *framework* com diversas resoluções baseado no algoritmo de múltiplos *kernels* para a identificação do ponto fiducial. Em seguida, o método combina

o algoritmo de múltiplos *kernels* ao procedimento de *bootstrap* [36] e identificamos se os pontos detectados são verdadeiros ou falso positivos com o intuito de fornecer robustez ao sistema.

Na etapa inicial, as faces são enquadradas utilizando o detector de faces *Viola Jones* [27–30]. Em seguida, o método proposto utiliza *templates* de trechos individuais de características para o procedimento de extração de características. A resolução de cada trecho é proporcional a distância intraocular, ou seja, pequenas resoluções caracterizam informações locais, e altas resoluções caracterizam as informações globais. As características de cada trecho são extraídas em escala de cinza independente da resolução. Para cada ponto fiducial trabalhado é desenvolvido um detector que pode apresentar 16 resoluções distintas.

Na etapa posterior, os autores [35] utilizam o classificador SVM para o treino das características e justificam a sua escolha devido a determinação do modelo de parâmetros a solução do problema de otimização. Na utilização do *multikernel* SVM, os *kernels* Gaussiano e polinomial são combinados para que a melhor separação entre classes seja obtida. Para a realização dos testes do classificador SVM, as características extraídas da região de interesse para cada candidato a ponto fiducial são indicadas ao classificador SVM e por fim, os pontos são classificados como pontos fiduciais ou não.

Na etapa de treino, em [35], o procedimento de *bootstrap* foi aplicado para tornar o detector de pontos fiduciais robusto a falsas classificações obtidas no procedimento de treino do classificador. No procedimento de *bootstrap*, o sistema é iterativamente treinado novamente com dois conjuntos que contenham falsas classificações produzidas após a detecção de pontos fiduciais realizada anteriormente.

Na etapa seguinte, os pontos fiduciais, que são flexíveis devido as variações de expressões da face são identificados. Para mapear a variação de cada ponto fiducial, o modelo *Gaussian Mixture* [37] é aplicado. Com esta aplicação, as expressões da face são identificadas. Por fim, os autores identificam se os pontos fiduciais flexíveis detectados em cada modelo gaussiano são verdadeiros ou falsos positivos utilizando a distância de *Mahalanobis*. O uso da distância de *Mahalanobis* é aplicada para a identificação da distância entre os pontos fiduciais classificados no SVM e o modelo gaussiano proposto, caso a distância seja menor que 95% o ponto fiducial é

considerado detectado corretamente.

Capítulo 3

Máquina de Vetores Suporte (SVM)

3.1 Introdução

A Máquina de Vetor Suporte (SVM, do inglês *Support Vector Machine*), proposta por Vapnik [1, 2, 13, 14] é um método de aprendizado de máquina supervisionado em crescente exploração. A elevação da quantidade de pesquisas é atribuída ao desempenho na solução de diversos problemas de classificação e regressão. A classificação consiste em separar categorias ou classes. A regressão por sua vez, consiste em generalizar um extenso conjunto de dados em uma função linear.

O SVM é considerado um contraste em relação aos métodos estatísticos clássicos. Quando comparado aos demais métodos estatísticos clássicos, o principal fator de comparação é o tratamento do espaço de características na solução dos problemas de classificação e regressão. Nos métodos estatísticos clássicos, por exemplo, a dimensão do espaço de características é reduzida para a melhoria do desempenho. Em contrapartida, o SVM incrementa a dimensão do espaço de características para obter melhores desempenhos.

O SVM foi desenvolvido inicialmente para trabalhar com a separação de duas classes, ou seja, com classificações binárias. Para a solução destes problemas, duas abordagens podem ser empregadas. Na primeira abordagem, o SVM transforma as múltiplas classes em diversos problemas de duas classes [12, 38]. Na segunda abordagem, o algoritmo de treino do SVM é modificado, onde versões multiclasse do algoritmo são criadas [12, 38]. Comparando as duas abordagens, a primeira abordagem é utilizada com mais frequência por não produzir algoritmos com alto

custo computacional quando comparado à segunda abordagem.

O objetivo do SVM é separar as classes por um hiperplano ótimo. O hiperplano ótimo separa os conjuntos de dados das classes de modo a maximizar a distância entre eles (denominada por margem) e as amostras mais próximas de cada conjunto. Observe na Figura 3.1. Note que são apresentadas duas classes, os triângulos e os quadrados, a finalidade é encontrar o melhor hiperplano de separação entre as classes.

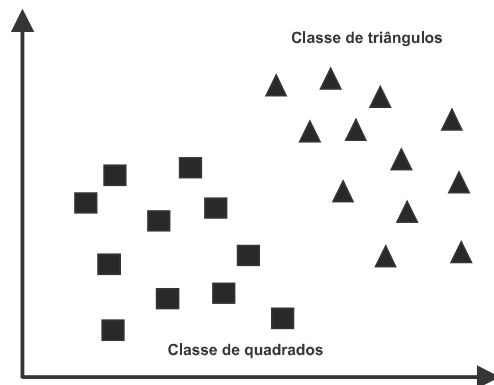


Figura 3.1: Exemplo de conjunto de dados composto por duas classes, classe de quadrados e classe de triângulos. Note que os dados são apresentados sem separação.

O hiperplano responsável pela separação da classe de triângulos e da classe de quadrados da Figura 3.1 é apresentado na Figura 3.2. Neste caso, tratamos de um problema de separação de duas classes, temos um hiperplano de dimensão igual a um, ou seja, temos uma reta. Caso, a quantidade de classes separadas fossem três, visualizaríamos um plano separando as classes.

Observe na Figura 3.2 que no espaço entre as classes, representados por linhas pontilhadas, diversos hiperplanos podem ser utilizados para a separação da classe de quadrados da classe de triângulos. Porém, somente um pode ser considerado o hiperplano ótimo. O hiperplano é considerado ótimo por encontrar a maior separação entre os triângulos e quadrados e por dispor a melhor separação entre os próprios elementos da mesma classe, neste caso o hiperplano ótimo é destacado na cor vermelha.

Finalizada a geração do hiperplano ótimo, aos elementos de treino situados na margem pré-estabelecida denominamos *vetores suporte*. Os vetores suporte são importantes para o processo de classificação. Os vetores suporte delimitam a região

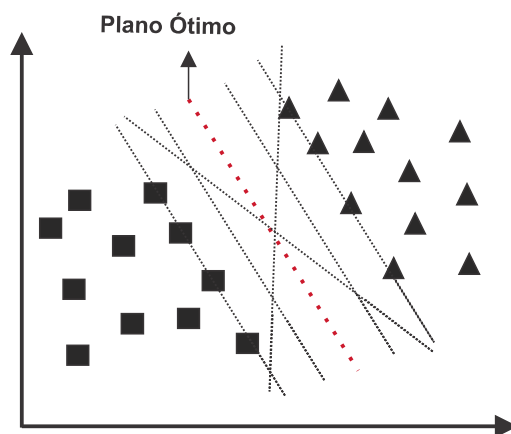


Figura 3.2: Separação dos dados da classe de quadrados e da classe de triângulo. Observe que diversos planos podem ser utilizados para esta separação. Entretanto, somente um é considerado ótimo, por encontrar a maior separação entre e intra classes. Nesta Figura, o plano ótimo é representado na cor vermelha.

pertencente as classes e em conjunto com o hiperplano ótimo delimitam a margem de separação. Deste modo, encontrados os vetores suporte, os demais vetores inseridos para treino são descartados pelo algoritmo.

3.2 Tipos de SVM

O SVM foi desenvolvido visando a análise de dados com um algoritmo de otimização matemática. Esta análise de dados do SVM é baseada na programação quadrática convexa [5]. A otimização matemática e a programação quadrática convexa mostram que a compreensão do algoritmo do SVM não é simples do ponto vista de sua construção e funcionamento.

A margem de separação das classes é um fator diretamente relacionado ao processo de otimização. De acordo com a presença ou ausência de dados das classes entre as margens, o procedimento de otimização da programação quadrática convexa pode se tornar custoso. Esta elevação do custo computacional ocorre devido a dificuldade em encontrar um hiperplano ótimo para a separação dos dados das classes.

A margem do SVM pode ser classificada em rígida e suave. Na margem rígida, o classificador SVM gerado após o procedimento de treino não possui qualquer elemento entre o hiperplano ótimo e os vetores de suporte, garantindo melhor

desempenho durante a classificação. Na margem suave é permitida a presença de elementos entre o hiperplano ótimo e os vetores suporte, a este conjunto de dados denominamos de não linearmente separáveis.

Quando tratamos de conjuntos de dados com elementos não linearmente separáveis, além do uso do SVM margem suave, podemos utilizar o SVM não linear com *kernels*. O SVM não linear com *kernels* mapeia o conjunto de dados de treino em um espaço de dimensão maior, de modo que os dados de treino tornem-se linearmente separáveis. Assim, ao encontrar o espaço de dimensão maior que separe o conjunto de dados de treino, o hiperplano ótimo de separação das classes e os vetores suporte são obtidos com o auxílio de *kernels*.

3.2.1 Margens Rígidas

Conforme mencionado, caso os dados das classes sejam linearmente separáveis, existe a possibilidade de separação das classes por um hiperplano ótimo. Para esta separação das classes, o método de margens rígidas pode ser adotado. O método de margens rígidas pode ser obtido conforme a seguir: seja T um conjunto de treinamento com n dados $\mathbf{x}_i \in X$ e seus respectivos rótulos $y_i \in Y$, em que X constitui o espaços de dados e $Y = \{-1, +1\}$.

O conjunto T será considerado linearmente separável se for possível separar os dados das classes $+1$ e -1 por um hiperplano. Descrito matematicamente por:

$$f(\mathbf{x}) = \mathbf{w}^t \cdot \mathbf{x} + b = 0 \quad (3.1)$$

Onde $\mathbf{w}^t \cdot \mathbf{x}$ é o produto escalar entre \mathbf{w}^t e \mathbf{x} , $\mathbf{w}^t \in X$ é o vetor normal ao hiperplano, $\frac{b}{\|\mathbf{w}^t\|}$ é a distância do hiperplano a origem, ou ainda, esta distância é igual a $|b|$ com ($\|\mathbf{w}^t\| = 1$).

A Equação (3.1) representa matematicamente a reta responsável por separar um conjunto considerado linearmente separável, logo uma função sinal $g(\mathbf{x}) = \text{sgn}(\mathbf{x})$ pode ser usada na obtenção das classificações, conforme a expressão a seguir:

$$g(\mathbf{x}) = \text{sgn}(\mathbf{x}) = \begin{cases} 1 & \text{se } \mathbf{w}^t \cdot \mathbf{x} + b > 0; \\ -1 & \text{se } \mathbf{w}^t \cdot \mathbf{x} + b < 0. \end{cases} \quad (3.2)$$

O hiperplano canônico em relação a T será aquele em que \mathbf{w}^t e \mathbf{x} for o mais próximo de $f(\mathbf{x}) = 0$, satisfazendo a expressão abaixo, definindo que qualquer objeto aplicado no classificador deve possuir módulo igual a 1.

$$|\mathbf{w}^t \cdot \mathbf{x}_i + b| = 1 \quad (3.3)$$

Ao utilizarmos o resultado da Equação (3.3), temos como implicação as inequações apresentadas a seguir, onde os rótulos de cada classe são definidos, representados pelos valores $+1$ e -1 .

$$\begin{cases} \mathbf{w}^t \cdot \mathbf{x}_i + b \geq +1 & \text{se } y_i = +1 \\ \mathbf{w}^t \cdot \mathbf{x}_i + b \leq -1 & \text{se } y_i = -1 \end{cases} \quad (3.4)$$

$$y_i(\mathbf{w}^t \cdot \mathbf{x}_i + b) - 1 \geq 0, \forall (\mathbf{x}_i, y_i) \in T \quad (3.5)$$

Na Figura 3.3, ilustramos a representação das inequações utilizando o exemplo de separação da classe de quadrados e de triângulos. Nesta figura, a Equação (3.1) denomina-se por hiperplano de separação ótima. As Equações (3.4) delimitam a margem de separação das classes. Em sequência à separação das classes, a distância entre o hiperplano ótimo e os vetores suporte (posicionados nos hiperplanos das margens rígidas) é calculada.

Com os hiperplanos definidos na Figura 3.3, a margem é definida pela distância entre as amostras mais próximas de cada classe e o hiperplano ótimo. As amostras mais próximas de cada classe são denominadas de vetores suporte e estão sobre os dois hiperplanos paralelos que delimitam as margens. Sendo a distância d dada por $\frac{2}{\|\mathbf{w}\|}$, para obtermos o hiperplano que maximiza a margem, é necessário minimizar $\|\mathbf{w}\|$ [39]. Neste caso, é necessário modelar esta função de modo a obter o ponto mínimo global, para isso é recorrido ao problema de otimização, conforme descrito a seguir:

$$\begin{aligned} & \text{Minimizar}_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 \\ & \text{Restrições : } y_i(\mathbf{w}^t \cdot \mathbf{x}_i + b) - 1 \geq 0, \forall_i = 1, \dots, n. \end{aligned} \quad (3.6)$$

A Equação (3.6) apresenta o problema de otimização e as respectivas restri-

ções, impostas para garantir a inexistência de dados de treinamento entre as margens de separação das classes.

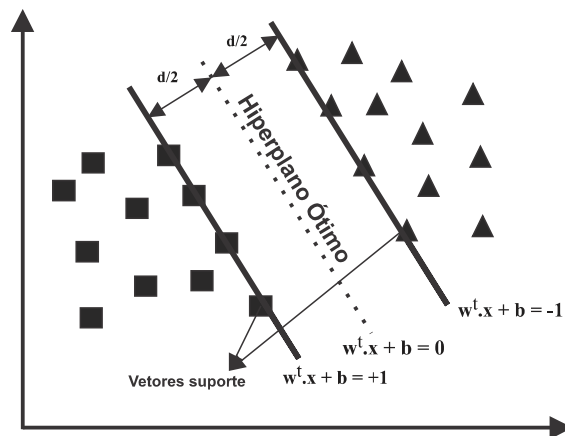


Figura 3.3: O SVM margem rígida é apresentado com a separação da classe de quadrados e a classe de triângulos. Note que os vetores suporte são os dados que delimitam cada classe. Cada vetor suporte corresponde a uma igualdade $w^t x_i + b = \pm 1$. A distância entre os vetores suporte e o hiperplano ótimo é chamado de margem.

O problema de otimização obtido é quadrático e devido as restrições possuirá um único mínimo global. Para este problema de otimização, a solução pode ser encontrada utilizando a função Lagrangeana [39,40]. A função Lagrangeana é utilizada por inserir as restrições impostas à função que deseja-se otimizar, facilitando a separação da classe de positivos da classe de negativos sem dados de treinamento entre as margens de separação.

3.2.2 Margens Suaves

Nas aplicações cotidianas que envolvem os problemas reais de reconhecimento de padrões é difícil encontrar dados linearmente separáveis devido a existência dos ruídos e *outliers* [3] que são dados de treinamento que ficam situados entre as margens de separação. Nestas situações, o SVM margem rígida não separa os dados de treinamento, pois as restrições apresentadas na Equação (3.6) foram inseridas para garantir que não existam dados de treinamento entre as margens de separação.

Para separar dados de treinamento que possuem ruídos e *outliers* [3] recorre-se ao uso do SVM margem suave. O SVM margem suave propõe suavizar o efeito das restrições impostas ao problema de otimização do SVM margem rígida. A

suavização das restrições do problema de otimização permitem a existência de dados de treinamento entre as margens de separação com o uso do classificador linear.

Na Figura 3.4 encontramos um exemplo com dados não linearmente separáveis. Note que encontramos dados da classe de quadrados e triângulos entre os hiperplanos separadores e os vetores suporte. Logo, a margem suave é aplicada para separação dos dados da classe de quadrados e da classe de triângulos.

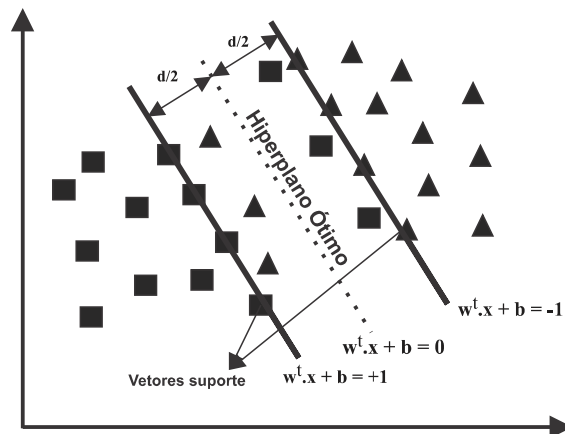


Figura 3.4: O SVM margem suave é apresentado com a separação da classe de quadrados e a classe de triângulos. Note que os vetores suporte delimitam cada classe. Cada vetor suporte corresponde a uma igualdade $\mathbf{w}^t \mathbf{x}_i + b = \pm 1$. A distância entre os vetores suporte e o hiperplano ótimo é chamado de margem. Diferente do SVM margem rígida, observe que encontramos dados da classe de triângulos próximo a região da classe de quadrados e vice-versa.

Na margem suave, as variáveis de folga (*slack variables*) [3] permitem que exista uma solução para os dados considerados não linearmente separáveis. Na margem suave, a variável de folga ξ_i é acrescentada como um compromisso entre a maximização da margem e o erro de classificação. O uso da margem suave implica na descrição a seguir:

$$y_i(\mathbf{w}^t \mathbf{x}_i + b) \geq 1 - \xi_i, \text{ para } i = \{1, \dots, n\}. \quad (3.7)$$

Na Equação (3.7) temos que: ξ_i , que assume valores ≥ 0 , é a variável de folga, y_i são os rótulos de classe, n representa o total de elementos de classe e o sobrescrito t representa o transposto do vetor. Para maximizar a margem de separação das classes e obtermos o hiperplano de separação ótima torna-se necessário encontrar a

solução do problema de otimização definido por:

$$\begin{aligned}
 \text{Minimizar} \quad & \frac{1}{2} \|\mathbf{w}\|^2 + C \left(\sum_{i=1}^n \xi_i \right). \\
 \text{Restrito à} \quad & y_i (\mathbf{w}^t \mathbf{x}_i + b) \geq 1 - \xi_i, \\
 & \text{para } i = \{1, \dots, n\}.
 \end{aligned} \tag{3.8}$$

Na Equação (3.8), C representa a ponderação imposta à margem suave e regula o compromisso entre a maximização da margem e o erro de classificação. A formulação deste problema de otimização é conhecida por C-SVC [15] e valores $\xi_i \in (0, 1]$, indicam dados entre as margens. Semelhante ao apresentado no SVM margem rígida, a solução do problema de otimização do SVM margem suave com restrições é encontrada com o uso da função lagrangeana [39, 40]. Na função lagrangeana os coeficientes de lagrange são utilizados para inserir as restrições impostas à função otimização da Equação (3.8), conseqüentemente, facilita a separação dos dados de treinamento através do hiperplano ótimo com dados entre as margens de separação.

3.2.3 Formulações do SVM

As formulações do SVM são elaboradas com o objetivo de encontrar o hiperplano ótimo para atender a solução de problemas reais. Dentre estes problemas, podemos citar a classificação, distribuição de estimação e a regressão [15]. Para cada formulação do SVM, parâmetros são acrescentados à formulação e às restrições com o objetivo de facilitar o controle do desempenho do problema de otimização.

Além da formulação C-SVC, utilizada neste trabalho, podemos encontrar na literatura várias formulações do algoritmo SVM [15]. As formulações do algoritmo SVM diferem-se entre si em vários aspectos: na formulação das equações, na função de otimização utilizada e por fim, os parâmetros utilizados para o controle do desempenho do problema de otimização. As principais formulações do SVM são apresentadas na Tabela 5.1. Note que o desempenho de cada algoritmo dependerá da escolha adequada dos parâmetros, justificando a procura pelo melhor conjunto de parâmetros do SVM.

A Tabela 5.1 apresenta as formulações do SVM, as restrições e os respectivos problemas de otimização. Nos problemas de otimização SVC (*Support Vector*

Classification), o parâmetro C (com variação $(0, \infty)$) é aplicado para redução do parâmetro de folga e o parâmetro ν (variando de $(0, 1]$) é aplicado para criar um limiar entre o valor máximo de erro de treinamento e o mínimo da quantidade de vetores suporte [41, 42]. Em seguida, a formulação *One-class SVM* é empregada em sistemas de classificação onde a classe de negativos não possui os dados definidos ou distribuídos de modo uniforme para a separação, neste caso, a classe de positivos é priorizada. Na formulação *One-class SVM*, ρ é o principal parâmetro para a solução deste problema de otimização [43, 44], permitindo a estimação de uma distribuição de dados de treinamento com dimensão elevada. Por fim, a SVM SVR (*Support Vector Regression*) é a formulação do SVM utilizada em problemas de regressão e os parâmetros C , ϵ e ν são os responsáveis pelo controle do desempenho do problema de otimização [45, 46].

Tabela 3.1: Problemas de otimização e restrições do método SVM.

Formulação	Problema de Otimização	Parâmetros
C-SVC	Minimizar: $\frac{1}{2} \ \mathbf{w}\ ^2 + C \left(\sum_{i=1}^n \xi_i \right)$ Restrito à: $y_i(\mathbf{w}^t \mathbf{x}_i + b) \geq 1 - \xi_i$, para: $\xi_i \geq 0 \quad i = \{1, \dots, n\}$.	C
ν -SVC	Minimizar: $\frac{1}{2} \ \mathbf{w}\ ^2 - \nu\rho + \frac{1}{n} \left(\sum_{i=1}^n \xi_i \right)$ Restrito à: $y_i(\mathbf{w}^t \mathbf{x}_i + b) \geq \rho - \xi_i$, para: $\xi_i \geq 0$ e $i = \{1, \dots, n\}$.	ν, ρ
<i>One Class</i>	Minimizar: $\frac{1}{2} \ \mathbf{w}\ ^2 - \rho + \frac{1}{\nu \cdot n} \left(\sum_{i=1}^n \xi_i \right)$ Restrito à: $\mathbf{w}^t \mathbf{x}_i \geq \rho - \xi_i$, para: $\xi_i \geq 0$ e $i = \{1, \dots, n\}$.	ν, ρ
ϵ -SVR	Minimizar: $\frac{1}{2} \ \mathbf{w}\ ^2 + C \left(\sum_{i=1}^n \xi_i \right) + C \left(\sum_{i=1}^n \xi_i^* \right)$ Restrito à: $\mathbf{w}^t \mathbf{x}_i + b - z_i \geq \epsilon + \xi_i$, $z_i - \mathbf{w}^t \mathbf{x}_i - b \geq \epsilon + \xi_i^*$, para: $\xi_i, \xi_i^* \geq 0$ e $i = \{1, \dots, n\}$.	C, ϵ
ν -SVR	Minimizar: $\frac{1}{2} \ \mathbf{w}\ ^2 + C \left(\nu\epsilon + \frac{1}{n} \sum_{i=1}^n (\xi_i + \xi_i^*) \right)$ Restrito à: $\mathbf{w}^t \mathbf{x}_i + b - z_i \geq \epsilon + \xi_i$, $z_i - \mathbf{w}^t \mathbf{x}_i - b \geq \epsilon + \xi_i^*$, para: $\xi_i, \xi_i^* \geq 0$ e $i = \{1, \dots, n\}$.	C, ν, ϵ

3.2.4 SVM Não Linear

O SVM é caracterizado por aumentar a dimensão do espaço que contem os dados que serão separados para melhoria do desempenho na separação dos dados. Quando os dados de treinamento não são linearmente separáveis, é necessário aumentar a dimensão do espaço, até tornar os dados linearmente separáveis [1–3]. Por

exemplo, dados não lineares são apresentados na Figura 3.5a, os retângulos e triângulos não podem ser separados de maneira linear, na Figura 3.5b, é criado um hiperplano que não mapeia de forma linear as classes, porém não separa as classe de maneira eficiente. Após o aumento da dimensão do espaço, encontra-se um hiperplano ótimo que separe de maneira linear conforme visto na Figura 3.5c.

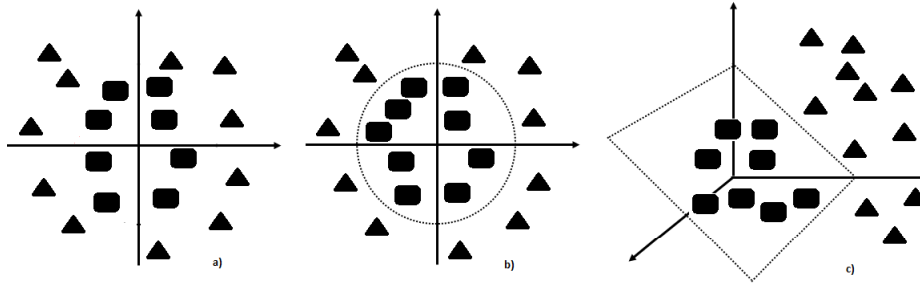


Figura 3.5: Exemplo do uso do *kernel* para a separação de dados não linearmente separáveis. a) Conjunto de dados não linear; b) Fronteira não linear para separação de classes; c) Fronteira Linear para separação de classes

O aumento de dimensão do espaço no SVM transforma o procedimento de otimização do algoritmo mais complexo quando comparado ao SVM de margem rígida e ao SVM de margem suave. Para garantir que o problema de otimização seja convexo são utilizados os *kernels*. Os *kernels* mapeiam os dados das classes em dimensões superiores e o que, na maioria dos casos, ajuda na separação das classes por intermédio do hiperplano de ótimo. Dentre os tipos de *kernels* utilizados nas aplicações em reconhecimento de padrões, citamos os polinomiais, os gaussianos e os sigmoidais [1–3].

Capítulo 4

Sistema de Detecção de Pontos Fiduciais utilizando Classificadores C-SVC

4.1 Sistema de Detecção de Pontos Fiduciais utilizando Classificadores SVM

Pontos fiduciais podem ser definidos como marcações de referência [6–8, 28]. Em algumas aplicações/problemas [33, 35] a detecção de pontos fiduciais é uma etapa importante. Por exemplo, nos sistemas apresentados em [33, 35] que abordam o problema do reconhecimento de expressões faciais e estimação de pose os autores utilizam a detecção de pontos fiduciais para compor o sistema completo. Nestes sistemas, a detecção de pontos fiduciais é feita para mapear pontos ao redor de uma região da face, como por exemplo, a boca e em seguida, o mapeamento é utilizado para integrar um modelo de forma culminando no reconhecimento da expressão da boca e estimação da pose da face.

Nesta dissertação, nosso objetivo é explorar parâmetros de classificadores SVM (do inglês, *Support Vector Machines*) [17] de um sistema de detecção de pontos fiduciais de faces humanas. Iremos explorar 4 parâmetros: a margem que regula o compromisso da separação de classes (C); o peso (W) atribuído as classes durante o treinamento do sistema; o número de iterações ($\#it$) e a tolerância (ε), necessários

para gerar o hiperplano para a separação dos elementos de classes. Com exploração dos parâmetros podemos visualizar com maior clareza como as variáveis influenciam o desempenho do sistema de detecção de pontos fiduciais.

Geralmente, os pontos fiduciais das faces humanas são disponibilizados pelas bases dados [18, 19]. Quando não fornecidos pela base de dados, os pontos fiduciais em faces humanas são identificados por representar pontos salientes, como por exemplo, ponta do nariz e boca, e o encontro de linhas e estruturas geométricas, representados pelos olhos e os cantos da boca, por exemplo. Dentre as aplicações mais comuns para a detecção de pontos fiduciais, citamos o reconhecimento de expressões [6], a extração de características [7], *tracking* [7] e a estimação de pose [33].

O sistema de detecção de pontos fiduciais utilizando classificadores SVM usado para a exploração de parâmetros é dividido em dois procedimentos distintos, designados por *procedimento de treino* e de *teste*. Antes de apresentarmos o treino e teste, mostraremos uma etapa comum a ambos denominada por *pré-processamento*.

4.1.1 Pré-processamento

O pré-processamento é uma etapa comum aos procedimentos de treino e de teste do sistema de detecção de pontos fiduciais. A aplicação do pré-processamento é destinada as imagens da base de dados para a redução de variações que influenciam no desempenho do sistema de detecção de pontos fiduciais. Dentre estas variações, podemos citar: pequenas rotações, variações de resoluções ou escala e variações na iluminação. Conforme apresentado na Figura 4.1, o pré-processamento é composto por três etapas distintas, são elas: um detector de faces *Viola Jones* [27–30]; um escalonamento e a correção de iluminação [31].

Em nosso trabalho, aplicamos um dos mais difundidos detectores de faces, conhecido popularmente como detector *Viola Jones* [27, 28]. Este detector é usado em diversos sistemas de detecção/reconhecimento [29, 30]. No diagrama em blocos para pré-processamento das imagens, o sub-bloco do detector *Viola Jones* é usado para detectar a face dos indivíduos em uma cena. Observe que a saída do sub-bloco consiste na face do indivíduo na mesma resolução da imagem de entrada.

As faces das imagens da base de dados não possuem a mesma resolução. Na sequência do detector *Viola Jones*, as imagens da base de dados e as marcações ma-

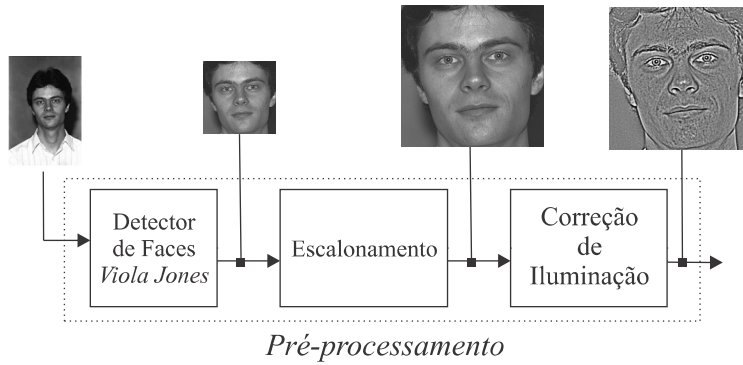


Figura 4.1: Diagrama em blocos para pré-processamento das imagens (treino e teste). Em particular, uma imagem com *background* é inserida na entrada do detector de faces *Viola Jones* para a detecção da face. Para fins didáticos e devido a restrição de espaço, esta imagem está representada com uma resolução menor. Na realidade, a resolução das imagens é maior do que a exposta nesta ilustração. Em seguida, a imagem é redimensionada. Por fim, a correção de iluminação é empregada com o objetivo de diminuir os efeitos de iluminação da cena.

nuais fornecidas são submetidas ao procedimento de escalonamento para ajustarmos a resolução para 220×220 . A adequação das imagens à resolução 220×220 tem por objetivo a facilidade na extração de características dos elementos de cada classe e a construção de um sistema de detecção para processar os pontos fiduciais das faces em uma resolução comum. Conforme apresentado no sub-bloco de escalonamento, as imagens de entrada e saída possuem resoluções diferentes, pois a imagem de saída está ajustada para a resolução 220×220 e conforme visualizamos, a imagem de entrada apresenta uma resolução diferente.

As marcações manuais são coordenadas cartesianas disponibilizadas pela base de dados e representam a localização do ponto fiducial na imagem. As marcações manuais são o centro dos blocos utilizados para compor as classes de positivos e de negativos. A resolução dos blocos, em que os centros são as marcações manuais, é disponibilizada na mesma resolução das imagens da base de dados. Devido ao escalonamento das imagens, aplicamos o escalonamento nas marcações manuais. Observe que as marcações manuais são modificadas quando realizamos o procedimento de escalonamento. Matematicamente, as marcações manuais, após o escalonamento, são obtidas da seguinte maneira:

$$(x_1, x_2) = \frac{220}{L}(x'_1, x'_2) \quad (4.1)$$

As variáveis (x_1, x_2) representam o par ordenado da marcação manual da imagem após o escalonamento para 220×220 . A variável L é o valor da largura da imagem, lembrando que a imagem é quadrada. Por fim, o par ordenado (x'_1, x'_2) representa a marcação manual original da base de dados.

Conforme descrito no início desta subseção, um dos problemas que reduzem o desempenho dos classificadores nos sistemas de detecção de pontos fiduciais é a não uniformidade da iluminação das imagens. A não uniformidade da iluminação das imagens da bases dados surge a partir de variações do ambiente e o modo de captação das imagens [47, 48]. Dentre estas variações, podemos citar: a distância entre o objeto e a câmera e as condições do ambiente de iluminação em locais abertos ou fechados, influenciando na intensidade de luz que incide nos objetos [47, 48].

Conforme [31], existem duas maneiras de tratar a correção de iluminação. Na primeira, as variações de iluminação das imagens são generalizadas em um modelo que posteriormente é aplicado nas novas imagens para a correção da iluminação. Na segunda, técnicas de processamento de imagens são aplicadas para correção de iluminação em cada imagem. Por fim, optamos pelo segundo método [31], pela complexidade de implementação e a correção de iluminação de individual por imagem.

4.1.2 Procedimento de Treino

O procedimento de treino do sistema de detecção de pontos fiduciais é dividido em 5 sub-etapas. Primeiramente, as imagens das bases dados e seus respectivos rótulos manuais são submetidos ao pré-processamento através da detecção da face, dos escalonamento e a correção de iluminação conforme comentado na Seção 4.1.1. A seguir, a região com alta possibilidade em conter o ponto fiducial é obtida através do probabilístico Gaussiano. Em seguida, blocos, que formam os elementos de classe, são selecionados. Após, os vetores de características são gerados para cada bloco selecionado. Finalmente, o classificador é obtido através do treinamento do classificador C-SVC. Na Figura 4.2, o diagrama em blocos da etapa de treinamento é apresentado.

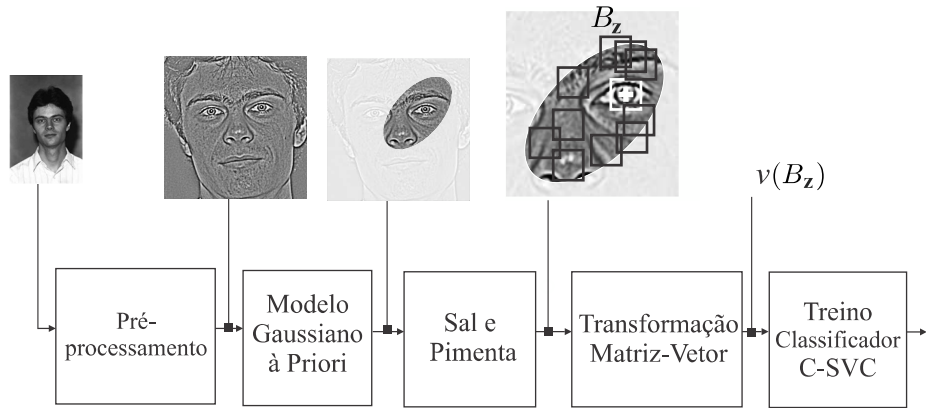


Figura 4.2: Diagrama em blocos para o procedimento de treino do sistema de detecção de pontos fiduciais. Em particular, uma imagem do conjunto de treino é submetida ao pré-processamento. Note que neste caso, conforme comentado anteriormente, adequamos a resolução da imagem. A imagem resultante do pré-processamento é inserida na entrada do Modelo Gaussiano para a definição da região de interesse. Na etapa seguinte, a definição das classes de positivos e negativos é realizada com o uso do *Salt and Pepper*. Em seguida, as classes são submetidas à etapa de transformação matriz-vetor para extração de características. Por fim, os vetores de características das classes de positivos e negativos são indicados à última etapa para que o classificador C-SVC seja projetado.

Definição da Região de Interesse

A restrição do espaço de interesse é uma tarefa importante em nosso sistema de detecção de pontos fiduciais de face humana. A utilização de uma região de específica é comum para agrupar as marcações manuais. Este agrupamento, facilita o projeto do classificador SVM durante o treino e a detecção do ponto fiducial no procedimento de teste, pois, ao invés do uso de todos os blocos pertencentes a imagem, processamos apenas os blocos que estão contidos nessa região (tanto para treino quanto para teste) o que obviamente reduz o tempo de resposta do sistema de detecção.

A esta restrição do espaço de interesse, nomeamos de Região de Interesse (ROI, do inglês *Region of Interest*). A ROI é construída utilizando-se as marcações manuais redimensionadas e cada ponto fiducial possui uma ROI distinta. Note que a mesma ROI é utilizada para os procedimentos de treino e teste.

Neste trabalho, a Região de Interesse é obtida utilizando o modelo Gaussiano à Priori [37], implementado com o auxílio da função densidade de probabilidade gaussiana, comum para as etapas de treino e teste. O modelo Gaussiano à Priori

é utilizado devido a sua capacidade de agrupar as marcações manuais em uma região específica, considerando a rotação existente nas faces das imagens das bases de dados e por fornecer a região de interesse com a maior probabilidade em possuir os pontos fiduciais. Matematicamente, este modelo é construído como segue: suponha um conjunto de coordenadas dos pontos fiduciais $\mathbf{x}_i = (x_1, x_2)$ do conjunto de treinamento (onde $i = \{1, 2, \dots, N\}$), eles são responsáveis por gerar a média $\boldsymbol{\mu}_{\mathcal{X}}$ e matriz de covariância $\boldsymbol{\Sigma}_{\mathcal{X}}$ da função de densidade de probabilidade gaussiana, onde \mathcal{X} é uma variável aleatória bidimensional com N realizações iguais as coordenadas $\mathbf{x}_i = (x_1, x_2)$. Dessa forma, um ponto na imagem será considerado um candidato a ponto fiducial **PF** com coordenadas $\mathbf{y} = (y_1, y_2)$ para a etapa de treino e teste, como segue:

$$\max_{\text{rótulo}(\mathbf{x}_i)=\mathbf{PF}} \{(1 + \rho)|\mathbf{x}_i - \boldsymbol{\mu}_{\mathcal{X}}|^2\} \geq (\mathbf{y} - \boldsymbol{\mu}_{\mathcal{X}})^t \boldsymbol{\Sigma}_{\mathcal{X}}^{-1} (\mathbf{y} - \boldsymbol{\mu}_{\mathcal{X}}). \quad (4.2)$$

No primeiro lado da desigualdade da Equação (4.2), a média aritmética das marcações manuais escalonadas $\boldsymbol{\mu}_{\mathcal{X}}$ é obtida com o auxílio das próprias marcações manuais \mathbf{x}_i . Os valores de $\boldsymbol{\mu}_{\mathcal{X}}$ e \mathbf{x}_i são rotacionados e em seguida é calculada a máxima distância euclidiana entre as marcações manuais rotacionadas e a média das marcações manuais rotacionadas. Esta máxima distância é calculada para definirmos o limite máximo que um ponto pode ser considerado candidato a ponto fiducial. Definido valor da distância máxima, o mesmo é acrescido do percentual ρ , sendo $\rho = 5\%$ para etapa de treino e $\rho = 8\%$ para etapa de teste. Este percentual indica a tolerância para um ponto da imagem ser considerado um candidato a ponto fiducial.

No lado direito da Equação (4.2), a distância de *Mahalanobis* é aplicada para identificar se os candidatos a pontos fiduciais $\mathbf{y} = (y_1, y_2)$ pertencem a região de interesse. Os atributos estatísticos média $\boldsymbol{\mu}_{\mathcal{X}}$ e matriz de covariância $\boldsymbol{\Sigma}_{\mathcal{X}}$ representam as marcações manuais agrupadas em uma região elíptica, considerando a sua distribuição e orientação na imagem. Deste modo, a distância de *Mahalanobis* identifica a distância entre os pontos com coordenadas $\mathbf{y} = (y_1, y_2)$ e as marcações manuais agrupadas na ROI. Por fim, a distância euclidiana não pode ser empregada, pois representa a distância entre dois pontos e não a distância entre um ponto e uma região de interesse, realizada com a distância de *Mahalanobis*.

Por fim, se o valor resultante da distância de *Mahalanobis* é menor que o va-

lor da distância máxima acrescida do percentual ρ , dizemos que $\mathbf{y} = (y_1, y_2)$ estará inserido na ROI. Este procedimento é realizado para todas as coordenadas cartesianas existentes numa imagem com dimensões 220×220 e os pontos que atenderem a condição, pertencerão a ROI.

Nesta dissertação, a ROI é definida pelos Algoritmos 1 e 2. Ao executarmos os algoritmos, a ROI será formada pelos candidatos a pontos fiduciais com valores de distância menores que a saída do Algoritmo 1. Vale ressaltar, que a formação da ROI no procedimento de treino é realizada através do acréscimo de 5% e a formação da ROI no procedimento de teste com o acréscimo de 8% através do uso do ρ .

Definição de Classes

No aprendizado de máquina supervisionado [1], os dados segmentados pertencem a uma classe pré-definida. A distinção dos elementos de classe é necessária para o projeto do classificador SVM. O sistema de detecção de pontos fiduciais utiliza um classificador que categoriza os elementos entre duas classes, ou seja, o classificador é binário [1]. O sistema de detecção de pontos fiduciais identifica se o ponto é ou não um ponto fiducial. Neste caso, é comum nomearmos as categorizações como *classe de positivos* e *classe de negativos*.

A classe de positivos representa os blocos que possuem os pontos fiduciais como centro. A localização do centro destes blocos por imagem é fornecida através das coordenadas cartesianas disponibilizadas pela base de dados. Assim, para cada ponto fiducial, o centro dos blocos da classe de positivos é fornecida.

Os blocos da classe de positivos possuem como centro as marcações manuais com as suas coordenadas cartesianas redimensionadas para a resolução de 220×220 . Note que na Figura 4.3, as marcações manuais são utilizadas como referência para que os pontos com coordenadas posteriores e anteriores nos sentidos horizontal e vertical da imagem sejam utilizados como centro dos blocos da classe de positivos. Observe que os pontos ao redor da marcação manual, ou seja, os vizinhos de um pixel nas direções horizontal e vertical, são considerados como ponto fiducial. Dessa forma, cada ponto fiducial irá gerar 5 blocos de dimensões 13×13 para a classe de positivos.

Em sequência, temos a etapa de separação de classes. Assim, os pontos

Algoritmo 1: Cálculo da distância máxima entre as coordenadas média das marcações e o ponto mais distante da média

Entrada: coordenadas \mathbf{x}_i do conjunto de treinamento com $i = \{1, 2, \dots, N\}$ e N sendo o total de imagens

Saída: distância máxima entre o ponto médio rotacionado σ_r e os pontos fiduciais, com o acréscimo do percentual ρ , sendo $\rho = 5\%$ para etapa de treino e $\rho = 8\%$ para etapa de teste.

início

1) Ler *marcações*:

$$\mathbf{x}_i = (x_1, x_2);$$

2) Calcular a matriz de covariância das *marcações*:

$$\Sigma_{\mathcal{X}} = E[(\mathbf{x}_i - \boldsymbol{\mu}_{\mathcal{X}})^t (\mathbf{x}_i - \boldsymbol{\mu}_{\mathcal{X}})];$$

3) Fazer a rotação das *marcações*:

$$\mathbf{x}_{ri} = (x_{r1}, x_{r2}) = \mathbf{x}_i \Sigma_{\mathcal{X}};$$

4) Calcular o vetor médio das *marcações* rotacionadas:

$$\boldsymbol{\mu}_r = \frac{1}{N} \sum_{i=0}^N \mathbf{x}_{ri};$$

5) Calcular os desvios padrões rotacionados:

$$\sigma_r = \sqrt{E(\mathbf{x}_{ri}^2) - (E(\mathbf{x}_{ri}))^2};$$

6) Calcular a distância das *marcações* rotacionadas para a média das *marcações* rotacionadas, normalizando por σ_r :

$$\mathbf{D}_{ri} = \frac{|\mathbf{x}_{ri} - \boldsymbol{\mu}_r|^2}{\sigma_r};$$

7) Determinar a maior distância:

$$\mathbf{D}_{max} = \max \frac{|\mathbf{x}_{ri} - \boldsymbol{\mu}_r|^2}{\sigma_r};$$

8) Acrescentar o percentual ρ a maior distância, sendo $\rho = 5\%$ para etapa de treino e $\rho = 8\%$ para etapa de teste:

$$D_{final} = (1 + \rho)|\mathbf{D}_{max}|;$$

fim

Algoritmo 2: Candidatos a pontos fiduciais

Entrada: desvio padrão rotacionado σ_r ; média dos pontos fiduciais rotacionados μ_r ; distância final entre média e ponto mais distante rotacionados D_{final} ; coordenadas y_i do conjunto de candidatos a pontos fiduciais com $i = \{1, 2, \dots, N\}$ e N sendo o total de imagens.

Saída: Pontos pertencentes a região de interesse.

início

1) Ler *marcações*:

$$\mathbf{y} = (y_1, y_2);$$

2) Fazer a rotação dos candidatos a pontos fiduciais:

$$\mathbf{y}_r = (y_{r1}, y_{r2}) = \mathbf{y}\Sigma\chi;$$

3) Calcular a distância dos candidatos a pontos fiduciais rotacionados para o vetor médio das *marcações* rotacionadas, normalizando-os σ_r :

$$D_{ry} = \frac{|\mathbf{y}_r - \mu_r|^2}{\sigma_r};$$

fim

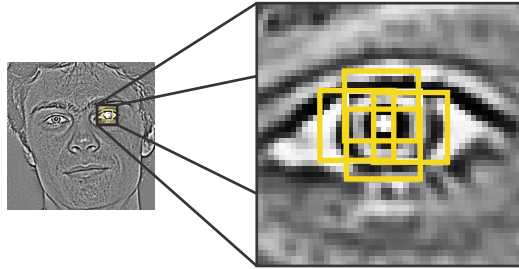


Figura 4.3: A marcação manual fornecida pela base de dados é utilizada como o centro do primeiro bloco da classe de positivos no ponto fiducial centro do olho esquerdo, sendo representado pelo símbolo " + ". Esta marcação manual serve de referência para definir os demais blocos da classe de positivos. Note que após o *zoom* na região do centro do olho, a visualização dos blocos da classe de positivos torna-se clara.

selecionados como centro dos blocos da classe de positivos são isolados. O isolamento destes pontos é realizado para evitar que um ponto seja utilizado como centro do bloco da classe de positivos e posteriormente como centro de um bloco da classe de negativos. Caso um ponto seja utilizado como centro dos blocos da classe de positivos e da classe de negativos, o SVM possuirá dificuldades na separação das classes, uma vez que um mesmo bloco fora apresentado como representante de classes distintas. Essa dificuldade, resultará em classificações incorretas na etapa de testes.

Este procedimento de isolar os pontos utilizados como centro dos blocos da classe de positivos denominamos de mascaramento. Realizado após a definição dos blocos da classe de positivos, o mascaramento é feito conforme a Figura 4.4. Inicialmente, aos pontos internos à ROI atribuímos o valor 1 e aos demais pixels da imagem o valor 0 é atribuído em sua representação matricial. A seguir, os pontos utilizados como centro dos blocos da classe de positivos são usados como o centro de sub-blocos de dimensões 3×3 e armazenados em uma imagem conforme a Figura 4.4. A criação desses sub-blocos isola a região ao redor da marcação manual para que as características próximas deste ponto não sejam utilizadas no vetor de características do blocos da classe de negativos. A estes sub-blocos atribuímos o valor 0 e para os demais pontos da imagem é atribuído o valor 1 em sua representação matricial. Por fim, realiza-se uma multiplicação pixel a pixel entre a imagem da ROI e a imagem com os sub-blocos, e os pontos resultantes a este procedimento são utilizados para ser o centro dos blocos da classe de negativos com dimensões 13×13 .

A imagem exibida na Figura 4.4, ao final da etapa de mascaramento, apresenta os pontos que serão utilizados como centro dos blocos da classe de negativos. Note que a quantidade restante de pontos da ROI fará com que a classe de negativos possua um elevado número de blocos de dimensões 13×13 quando comparada a quantidade de blocos da classe de positivos. Assim, a quantidade muito superior de blocos da classe de negativos em relação a classe de positivos torna a etapa de projeto do classificador SVM ineficaz.

A redução da quantidade de blocos da classe de negativos é feita através da seleção dos pontos pertencentes a ROI, após o mascaramento. A etapa de seleção destes pontos é realizado de maneira aleatória. O objetivo da seleção aleatória é não tornar o sistema de detecção de pontos fiduciais tendencioso, deste modo, não teremos o controle de quais blocos serão indicados ao classificador SVM.

Esta etapa de seleção aleatória de pontos, nomeamos de *Salt and Pepper*. A etapa *Salt and Pepper* consiste em gerar aleatoriamente pontos pretos e brancos distribuídos em quantidades iguais na ROI (região de interesse). Conforme apresentado na Figura 4.5, os pontos pretos gerados na ROI serão o centro dos blocos da classe de negativos. Obviamente, a quantidade de pontos pode ser controlada (esse detalhamento será feito no capítulo de experimentos). Dessa maneira, podemos reduzir

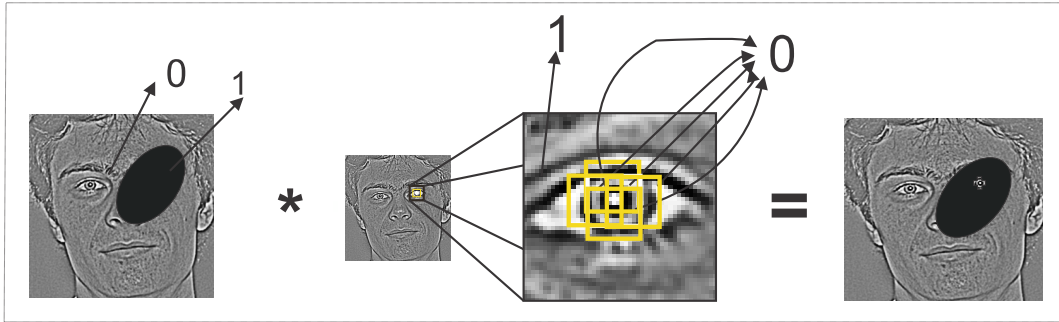


Figura 4.4: A operação de mascaramento é realizada com o uso da imagem da esquerda em que os pontos inseridos na ROI o valor 1 é atribuído e aos demais pontos da imagem o valor 0 é atribuído. Na imagem central, os sub-blocos com dimensões 3×3 são atribuídos o valor 0 e aos demais pontos desta imagem é atribuído o valor 1. A multiplicação pixel a pixel é realizada entre as duas imagens, e a imagem resultante (à esquerda) é apresentada com a ROI e os sub-blocos excluídos (na cor branca).

a quantidade de blocos negativos indicados ao classificador.

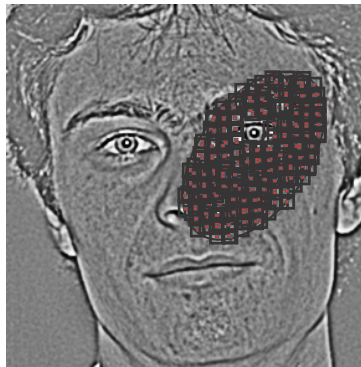


Figura 4.5: Representação da quantidade de blocos da classe de negativos antes da etapa de *Salt and Pepper*. Note que cada ponto da ROI, destacado na cor vermelha, é utilizado como centro do bloco, conseqüentemente a quantidade de blocos da classe de negativos é elevada. Observe, por exemplo que o resultado do mascaramento para o ponto fiducial centro do olho esquerdo é nítido, pois a região pertencente aos blocos da classe de positivos foi isolada.

A etapa *Salt and Pepper* é exemplificada na Figura 4.6. Primeiro, após o mascaramento, a cada ponto pertencente a imagem (o ponto interno a ROI é destacado na cor vermelha) é atribuído um valor, ou ainda uma *tag*, para ser utilizado como referência para a seleção dos pontos. Posteriormente, os pontos são selecionados aleatoriamente por intermédio das suas *tags*. Finalmente, os pontos selecionados são utilizados como centro dos blocos da classe de negativos com dimensões 13×13 .

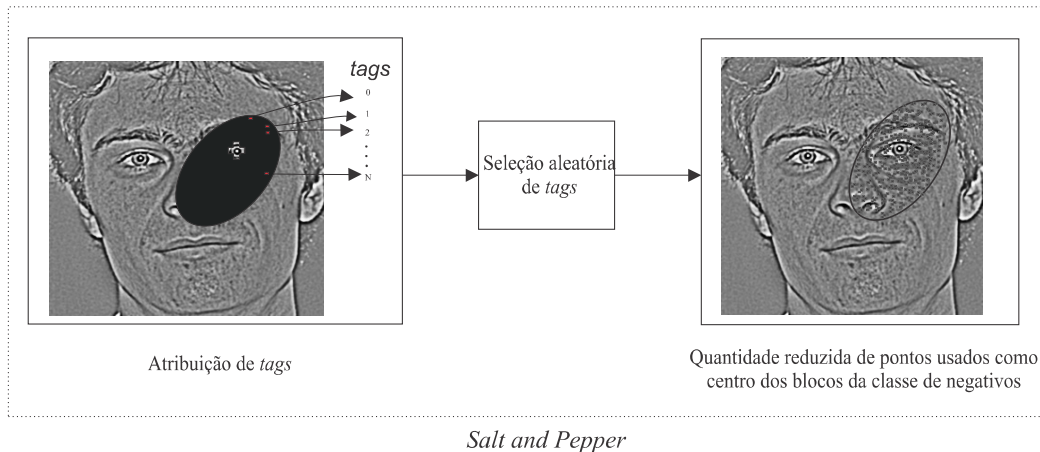


Figura 4.6: Exemplificação da etapa *Salt and Pepper*. Inicialmente, *tags* são atribuídas aos pontos internos à ROI (destacados na cor vermelha) da imagem resultante após a etapa de mascaramento. A quantidade total de *tags* atribuídas é igual a $N - 1$, onde N representa a quantidade de pontos internos à ROI. Em seguida, as *tags* são selecionadas aleatoriamente. Por fim, as *tags* selecionadas indicam os pontos utilizados como centro dos blocos que formam a classe de negativos. Note que a quantidade de pontos utilizados na imagem final como centro dos blocos da classe de negativos é inferior a quantidade de pontos no início da etapa de *Salt and Pepper*.

Transformação Matriz-Vetor

A extração do vetor de características é uma etapa comum para os procedimentos de treino e teste. Para o treino, o vetor de características é utilizado no projeto do classificador SVM para a separação das classes de positivos e negativos. Similarmente, no procedimento de teste, o vetor de características, associado a um candidato a ponto fiducial, é utilizado como entrada do classificador SVM (projetado no procedimento de treino) para que o mesmo decida se o candidato é ou não um ponto fiducial, ou seja, se ele pertence a classe de positivos ou negativos. Os blocos pertencentes à classe de positivos e negativos tem os vetores de características extraídos pixel a pixel a partir das imagens da base de dados. Os valores dos pixels são extraídos em escala de cinza. A resolução da escala de cinza é igual a 8 bits.

A etapa de transformação matriz-vetor extrai as características dos blocos da classe de positivos e negativos. A extração de características para o centro do olho direito é apresentada na Figura 4.7. Por exemplo, um bloco da classe de positivos é selecionado e a marcação manual é destacada com um retângulo. Na matriz a direita, os valores em escala de cinza extraídas as características do bloco e armazenados

em uma matriz de dimensões 13×13 . Na sequência, os valores correspondentes a esta matriz são usados para preencher um vetor com dimensão igual a 169. O vetor de características é preenchido com cada linha da matriz apresentada, colocadas sequencialmente no vetor até o preenchimento das 169 posições.

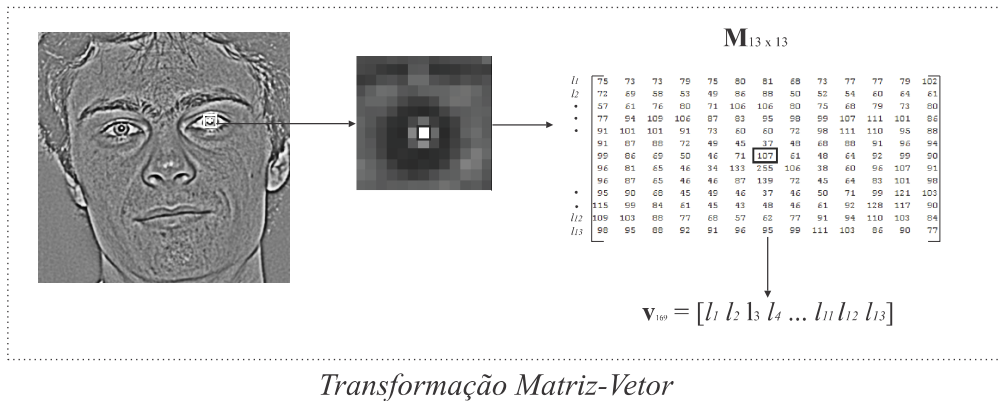


Figura 4.7: Exemplo de transformação matriz-vetor para o bloco da classe de positivos do ponto fiducial centro do olho esquerdo. Note que inicialmente o bloco e a sua respectiva marcação manual são demarcados. Na imagem seguinte, o bloco da classe de positivos com dimensões 13×13 é extraído da imagem e a sua marcação permanece em destaque através do retângulo. Posteriormente, os valores em escala de cinza são extraídos e armazenados em uma matriz de dimensões 13×13 . Note que os valores em escala de cinza do bloco é apresentado e o valor em escala de cinza da marcação manual é destacado. Por fim, cada linha da matriz \mathbf{M} preenche sequencialmente o vetor de características \mathbf{v} , no fim, indicado ao SVM para o projeto do classificador C-SVC.

Treino do Classificador C-SVC

No SVM, a margem de separação pode conter ou não elementos de classe. Quando a margem de separação não possui elementos de classe, chamamos o classificador de SVM margens rígidas. Entretanto, quando tratamos de classes não linearmente separáveis (o caso desta dissertação), o problema de otimização é insolúvel para margens rígidas. Assim, é necessário utilizarmos um classificador de margens suaves para a solução de problemas próximos da realidade.

O classificador C-SVC (*C-Support Vector for Classification*) é um classificador SVM de margens suaves. O classificador C-SVC é também chamado de Máquina de Vetores Suporte Regulada ou ainda de Algoritmo Padrão. A denominação de Máquina de Vetores Suporte Regulada é adotada devido ao controle realizado na

separação de classes durante o treino e a permissão da existência de elementos entre as margens de separação.

O principal parâmetro do classificador C-SVC é o C (Custo). O parâmetro C é responsável por regular o compromisso entre o hiperplano que separa os elementos de classe e gerar os vetores de suporte, considerando os limites de cada classe. Conseqüentemente, o custo regula o compromisso entre a maximização da margem e o erro de classificação, permitindo assim que o SVM possa lidar com problemas de separação de classes mais próximos da realidade.

Seguindo o procedimento de treino do sistema de detecção de pontos fiduciais, para cada bloco da classe positivos e da classe de negativos selecionados e extraídos das imagens das bases de dados, os vetores de características são indicados ao classificador em conjunto com um vetor contendo os rótulos dos blocos, obviamente, o comprimento deste vetor será igual à quantidade de vetores de características da classe de positivos e negativos somados. Os rótulos necessitam ser distintos para indicar ao classificador C-SVC que tratam-se de classes distintas. Deste modo, indicamos ao classificador dois números diferentes, um representando a classe de positivos e o outro representando a classe de negativos.

Além dos vetores de características e os rótulos dos blocos da classe de positivos e da classe de negativos, indicamos ao classificador C-SVC os valores do parâmetro de peso de cada classe com o objetivo de aumentar ou diminuir o tamanho da margem de separação individualmente. Com o uso do parâmetro peso, as margens de separação tornam-se mais restritas e deste modo podemos evitar a presença de elementos de classe entre as margens de separação, tornando o processo de separação de classes ágil.

O parâmetro de iterações é indicado ao classificador C-SVC e informa a quantidade de execuções do algoritmo para obter-se a separação de classes. Para cada valor de iteração indicado, o classificador C-SVC apresenta uma separação de classes distinta. Isto ocorre pelo fato da quantidade de iterações indicada ser insuficiente para separar todos os blocos da classe de positivos e negativos durante o treino.

A tolerância ao erro do classificador SVM é o último parâmetro da etapa de treino a ser indicado. A tolerância refere-se ao erro entre a resposta desejada e a

saída do SVM. Deste modo, durante os testes do classificador C-SVC, a quantidade de blocos classificados incorretamente não excedem o valor da tolerância.

Ao final do treino do classificador e da conseqüente separação de classes, um arquivo no formato XML é salvo e nele são apresentados os parâmetros indicados ao SVM. Além disso, o vetor de características que representa o hiperplano que separa a classe de positivos e a classe de negativos, e os vetores de suporte que são responsáveis por delimitar cada classe são armazenados neste arquivo de formato XML.

4.1.3 Procedimento de Teste

O procedimento de teste apresenta etapas semelhantes às descritas no procedimento de treino. A semelhança tem por objetivo prover as imagens do procedimento de teste o mesmo tratamento realizado nas imagens de treino nas etapas antecedentes ao projeto do classificador C-SVC. Com a igualdade nas etapas iniciais, as imagens das bases de dados de teste apresentarão vetores de características com valores em escala de cinza semelhantes aos do procedimento de treino, contribuindo para o desempenho do sistema de detecção de pontos fiduciais.

Na Figura 4.8 é apresentado o diagrama em blocos do procedimento de teste, subdividido em 4 etapas. Inicialmente, o pré-processamento é aplicado às imagens da base de dados, semelhante a etapa de treinamento. Em seguida, o modelo gaussiano à priori delimita a região da face a ser testada, utilizando a ROI acrescida de 8%. Depois, a transformação matriz-vetor é empregada para a mesma finalidade da etapa de treinamento. Por fim, o vetor de características da região delimitada é testado utilizando o classificador projetado na etapa de treinamento, definindo se o ponto testado é um ponto fiducial.

O pré-processamento aplicado ao procedimento de teste é similar ao utilizado na etapa de treino. Conforme apresentado na Subseção 4.1.1, o pré-processamento é aplicado as imagens das bases de dados de teste para a redução de pequenas rotações, variações de resoluções ou escala e variações na iluminação que influenciam no desempenho do sistema de detecção de pontos fiduciais. Deste modo, as imagens de teste são indicadas inicialmente a um detector de faces *Viola Jones* [27–30] para o enquadramento da face e a eliminação do *background*. Em seguida, a imagem

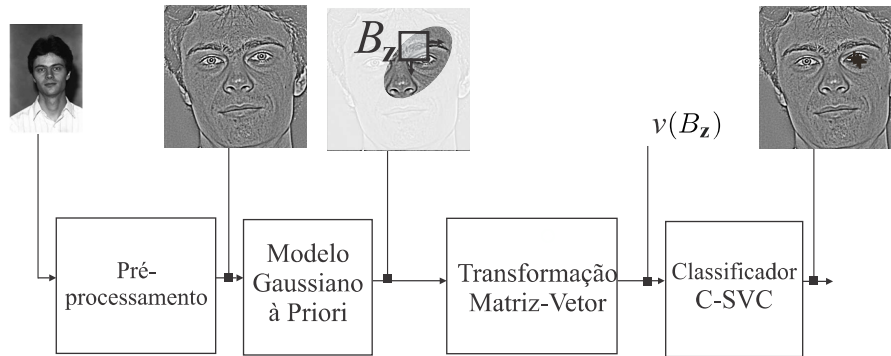


Figura 4.8: Diagrama em blocos para o procedimento de teste do sistema de detecção de pontos fiduciais. Em particular, uma imagem do conjunto de teste é inserida no sistema, após o pré-processamento, o Modelo Gaussiano é empregado para definir os blocos dos candidatos a pontos fiduciais a partir dos pontos inseridos na ROI de teste. Em seguida, os candidatos a pontos fiduciais são submetidos à etapa de transformação matriz-vetor para extração de características dos seus respectivos blocos. Por fim, os vetores de características dos candidatos a pontos fiduciais são testados usando o classificador projetado na etapa de treino, definindo-se a que classe o candidato a ponto fiducial pertence.

enquadrada é submetida ao escalonamento para o ajuste da imagem às dimensões de 220×220 . E por fim, a correção de iluminação [31] é empregada para ajustar a não uniformidade da iluminação das imagens das bases de testes.

A seguir no procedimento de teste, a ROI formada na etapa de treino, com o acréscimo de 8% é usada no modelo Gaussiano à Priori da etapa de teste para indicar os candidatos a pontos fiduciais. A ROI na etapa de teste delimita a região da face a ser testada nas imagens das bases de dados de teste. Na Figura 4.9, a ROI da etapa de teste é apresentada, e em seguida, cada bloco de dimensões 13×13 B_z é processado utilizando-se uma janela deslizante. Note que todos pontos pertencentes a esta ROI serão utilizados como centro dos blocos dos candidatos a pontos fiduciais.

Os vetores de características dos blocos de candidatos a pontos fiduciais são obtidos de forma semelhante à realizada no procedimento de treino. Por intermédio da etapa de transformação matriz-vetor, cada bloco B_z tem os seus valores em escala de cinza extraídos e em seguida, estes valores preenchem o vetor $v(B_z)$, composto de 169 posições.

Finalizando o procedimento de teste, cada vetor $v(B_z)$ é indicado ao classificador C-SVC projetado no procedimento de treino. Assim, cada vetor $v(B_z)$ é categorizado na classe de positivos ou na classe de negativos. O arquivo XML salvo

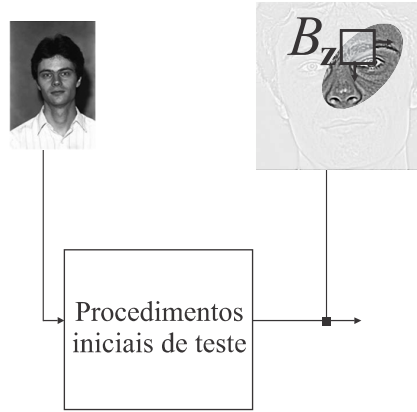


Figura 4.9: Aplicação da janela deslizante B_z na ROI da etapa de teste. Cada ponto pertencente a ROI de teste é utilizado como centro do bloco dos candidatos a pontos fiduciais.

no procedimento de treino é utilizado como referência para mapear a região pertencente a classe de positivos e a classe de negativos, deste modo, cada vetor $v(B_z)$ é mapeado no espaço dos vetores de características do SVM salvo no arquivo XML e de acordo com a sua localização em relação ao hiperplano de separação de classes é definido se $v(B_z)$ pertence a classe de positivos ou negativos.

4.2 Investigações Propostas para o Sistema de Detecção de Pontos Fiduciais

4.2.1 Iterações do SVM

O hiperplano ótimo produz a separação das classes quando utilizamos o método SVM. Proveniente do procedimento de treino do classificador, entre os vetores suporte e o hiperplano ótimo não existem blocos das classes de positivos e de negativos, e assim, cada classe que será separada pelo SVM possui uma região na superfície de separação. Deste modo, conforme citado em [15], o hiperplano ótimo visa o alcance do valor mínimo global da sua função $\|\mathbf{w}\|$ para a solução do problema de otimização da separação de classes. Ao procedimento de separação de classes realizado a cada iteração do algoritmo do SVM denominamos de convergência do SVM.

A convergência do SVM é uma consequência do procedimento de treino para o classificador SVM. Neste, decide-se em qual iteração do algoritmo será encerrado

o procedimento de treino para o classificador SVM. Dessa forma, o hiperplano que separa as classes é definido.

O critério de encerramento de treino para o classificador SVM depende de dois fatores: o número de iterações e a tolerância. O número de iterações apresenta ao SVM a quantidade de execuções do algoritmo para obter-se a separação de dados. A tolerância, refere-se ao erro adquirido durante o procedimento de classificação. Ao classificar um bloco, o classificador SVM pode indicar um resultado incorreto, ou seja, o resultado do classificador em relação ao bloco é diferente do resultado real. Deste modo, a tolerância é utilizada para regular este erro. Além disso, o critério de encerramento pode ser aplicado com prioridades. Dentre as prioridades, a aplicação pelo uso da tolerância, do número de iterações, ou ainda, a utilização de ambos.

A convergência do SVM e o respectivo critério de encerramento do algoritmo de treino do classificador SVM são responsáveis por definir o hiperplano ótimo para a separação dos dados de treino das classes de positivos e negativos. A cada hiperplano ótimo utilizado para a separação da classe de positivos da classe de negativos, o desempenho do sistema de detecção de pontos fiduciais pode ser influenciado. Além disso, conforme apresentado, existem três maneiras de encerrar o algoritmo de treino do classificador SVM e a investigação da influência de cada critério de encerramento no sistema de detecção de ponto fiduciais é relevante nesta pesquisa, pois podemos identificar qual a influência deste fator no desempenho do sistema proposto e na formação do hiperplano ótimo quanto a separação dos dados de treino da classe de positivos e da classe de negativos.

4.2.2 Definição das Classes dos Blocos de Treino

A seleção dos blocos no procedimento de treino do classificador do sistema de detecção de pontos fiduciais é realizada através da seleção de pontos a partir da Região de Interesse (ROI, do inglês *Region of Interest*). A ROI é caracterizada por representar a região com alta probabilidade de conter o ponto fiducial. Procedimentos matemáticos são aplicados às coordenadas das marcações manuais da base de dados para formar a ROI. Dentre estes procedimentos matemáticos, podemos citar o Modelo Gaussiano à Priori, explicado na Seção 4.1.2.

Finalizada a definição da ROI, indica-se ao classificador as características das

classes que serão utilizadas, e os seus respectivos rótulos. Para o caso de categorizar os blocos de treino e teste, denominamos as classes por classe de positivos e classe de negativos. Primeiramente, a classe de positivos possui essa denominação por representar os blocos com o ponto fiducial como centro, e a localização do centro destes blocos por imagem é fornecida através das coordenadas cartesianas disponibilizadas pela base de dados. Por último, os blocos da classe de negativos possuirão como centro a coordenada dos pontos que não sejam os pontos fiduciais e que estejam inclusos na região de interesse.

Os pontos fiduciais do sistema de detecção são obtidos a partir das marcações manuais das bases de dados. Em geral, as marcações manuais informam ao classificador as coordenadas cartesianas do ponto fiducial por imagem. As marcações manuais são relevantes para o sistema de detecção de pontos fiduciais por representar o centro do bloco da classe de positivos e servir de referência para o centro dos demais blocos da classe de positivos, neste trabalho, representado pelas coordenadas cartesianas posteriores e anteriores nos sentidos horizontais e verticais à coordenada fornecida pela base de dados. Por fim, o ponto fiducial corresponde ao objeto de detecção do sistema proposto e torna-se inviável projetar o classificador por SVM sem os blocos da classe de positivos, sendo necessário indicar ao classificador classes distintas para que ocorra a separação de classes.

Dessa forma, em nosso sistema, a seleção de blocos da classe de positivos é de fundamental importância para o projeto do sistema de detecção de pontos fiduciais. Conforme apresentado na Seção 4.1.2, durante a separação de classes do SVM na etapa de treino erros podem surgir. Estes erros são provenientes da classe de positivos e da classe de negativos que são indicados ao classificador para o treino com os valores dos blocos da classe de negativos e rótulos incorretos. Deste modo, quando os blocos e os rótulos da classe de negativos são apresentados incorretamente ao SVM para realizar o procedimento de treino, o SVM gera o hiperplano ótimo com a separação incorreta da classe de positivos e da classe de negativos.

A etapa de mascaramento apresentada na Seção 4.1.2 será aplicada para que os pontos que serão utilizados como centro dos blocos da classe de positivos não possam ser utilizados como centro para os blocos da classe de negativos. Ao ser empregado, o mascaramento busca evitar que um mesmo bloco da classe de negativos

na transformação matriz-vetor seja indicado ao classificador repetidamente, porém com rótulos distintos. A indicação de blocos da classe de negativos que não possuam os mesmos valores em escala de cinza dos pixels dos blocos da classe de positivos ao classificador é proporcionada ao classificador SVM e é caracterizado como um diferencial desta pesquisa.

Finalizado o mascaramento, os pontos usados como centro da classe de positivos são excluídos, logo, a ROI passa a apresentar uma região sem pontos. Os pontos que restam na ROI são utilizados como centro dos blocos da classe de negativos e podem apresentar elevada, conseqüentemente, a quantidade de blocos da classes de negativos também por ser elevada, desequilibrando a relação da quantidade de blocos da classe de positivos e da classes de negativos.

Caso o desequilíbrio entre a quantidade de blocos da classe de positivos e da classe de negativos durante os experimentos seja elevado, a redução de pontos utilizados como centro da classe de negativos é realizada por um algoritmo responsável por gerar aleatoriamente pontos pretos e brancos distribuídos em quantidades iguais na região de interesse, a este algoritmo, denominamos *salt and pepper* [49–51].

Em decorrência desta investigação, o uso do algoritmo *salt and pepper* influenciará na quantidade de blocos da classe de negativos. Assim, o sistema de detecção de pontos fiduciais poderá acertar mais, ou seja, o seu desempenho poderá ser melhorado.

A seleção dos blocos das classes de positivos e negativos possui papel relevante nesta pesquisa. Conforme apresentado nesta seção, a formação da ROI para a definição da região que contenha o ponto fiducial é apresentado como diferencial desta pesquisa por delimitar a região em que os blocos das classes de positivos e negativos serão extraídos em contrapartida a utilização de toda a região da face fornecida pela base de dados, pois neste caso dificulta a separação dos blocos da classe de positivos e da classe de negativos. O mascaramento dos pontos utilizados como centro dos blocos da classe de positivos com o intuito de não reutilizamos para o centro dos blocos da classe de negativos é outro fator a ser levado em consideração. Caso um mesmo ponto seja utilizado como centro do bloco da classe de positivos e do bloco da classe de negativos o algoritmo de treino do classificador do SVM possuirá dificuldades em separar os dados, pois o blocos serão apresentados com classes

distintas. Deste modo, o tamanho da máscara também influencia no desempenho do sistema de detecção de pontos fiduciais e sua investigação é relevante nesta pesquisa. Por fim, a influência do algoritmo *salt and pepper* no estudos do controle da quantidade de blocos da classe de negativos indicados ao classificador para o treino do SVM é uma contribuição notável por influenciar na separação da classe de positivos da classe de negativos e conseqüentemente no desempenho do sistema de detecção de pontos de pontos fiduciais. Deste modo, o conjunto de etapas desenvolvidas para a seleção dos blocos das classes de positivos e negativos para o treino do SVM são contribuições notáveis no sistema de detecção de pontos fiduciais e a investigação da influência destas etapas no desempenho do sistema proposto é importante para esta pesquisa.

4.2.3 *Grid* de Parâmetros

Definimos *Grid* de parâmetros pela análise da influência conjunta de parâmetros [16] num sistema genérico. Quando os parâmetros são variados em uma faixa de valores, o desempenho do sistema de detecção de pontos fiduciais pode ser analisado. A partir desta análise, os parâmetros investigados conjuntamente norteiam a identificação dos respectivos valores que aumentam ou reduzem o desempenho do sistema de detecção de pontos fiduciais.

Conforme apresentado na Seção 4.1, em cada etapa do procedimento de treino, podemos encontrar parâmetros que podem ser importantes para cada etapa. Os parâmetros estão distribuídos tanto no sistema de detecção de pontos fiduciais quanto no algoritmo de treino do classificador SVM, este último responsável pela separação das classes de positivos e negativos. Dentre os parâmetros do sistema de detecção de pontos fiduciais citamos o tamanho da máscara, a intensidade de pontos utilizada no algoritmo *salt and pepper* que influencia na quantidade de blocos da classe de negativos e o peso (W) atribuído a cada classe para ser levado em consideração durante o treinamento do algoritmo do SVM no sistema de detecção de pontos fiduciais. No treino do algoritmo do classificador SVM podemos citar a margem que regula o compromisso da separação das classes de positivos e negativos (C). Além deste, podemos citar o número de iterações ($\#it$) e a tolerância (ε) responsáveis pelo critério de encerramento do algoritmo do classificador SVM apresentado ante-

riormente. Cada parâmetro apresentado pode indicar um desempenho distinto no sistema de detecção de pontos fiduciais.

Os parâmetros investigados serão o peso (W), o parâmetro de ponderação (C), e o número de iterações ($\#it$) e a tolerância (ε) do critério de encerramento por representarem parâmetros que podem influenciar diretamente na formação do hiperplano ótimo de separação das classes. Para investigar o comportamento destes parâmetros e buscar os melhores valores que produzem os melhores desempenho do sistema de detecção de pontos fiduciais, o *grid* de parâmetros é aplicado. A cada parâmetro será definida uma faixa de valores. Assim, os valores do *grid* de parâmetros serão inseridos no sistema de detecção de pontos fiduciais e no algoritmo de treino do classificador SVM. Ao final, o desempenho do sistema de detecção de pontos fiduciais será avaliado e poderemos identificar quais as combinações de parâmetros que permitem os melhores desempenhos.

Capítulo 5

Experimentos e Resultados

5.1 Base de Dados

Nesta dissertação, iremos utilizar duas bases de dados denominadas por *BioID* e *Feret*. A *BioID* é uma base de dados caracterizada por apresentar variações da pose, condições de iluminação natural e *backgrounds* distintos. Estas características tornaram a *BioID* uma base de dados amplamente utilizada em pesquisas para a comparação de desempenho entre técnicas de classificação aplicadas à sistemas de detecção de pontos fiduciais. A base de dados *BioID* original é composta por 1521 imagens em escala de cinza com resolução de 384×286 . A *BioID* possui imagens que foram retiradas de 23 indivíduos em diferentes formas. Neste conjunto, os indivíduos podem conter óculos, barba e bigode. Além das imagens, a base de dados possui uma anotação de 20 pontos na face e outra contendo somente os olhos. Na Figura 5.1, podemos visualizar exemplos das imagens da base de dados *BioID*.



Figura 5.1: Exemplos de imagens da base dados *BioID* [18].

Nos experimentos desta dissertação, utilizaremos um subconjunto da base de

dados *BioID* composto por 503 imagens. Neste subconjunto, as imagens de indivíduos com óculos, barba e bigode, e as imagens com grandes rotações foram excluídas por inibirem a extração das características dos blocos das classes de positivos e negativos. Os experimentos utilizam 11 marcações manuais.

A base de dados *Feret* (*Facial Recognition Technology*) foi desenvolvida pelo Departamento de Defesa em Tecnologia Antidrogras dos Estados Unidos com o objetivo de desenvolver reconhecimento de rosto de forma automática auxiliando na segurança, inteligência e agentes da lei no exercício de suas funções [19, 52, 53]. A *Feret* é composta por 11338 (versão *color*), codificadas em 24 bits imagens de diferentes indivíduos em distintas expressões faciais e condições de iluminação [19, 52, 53]. Os indivíduos apresentam óculos, barba e/ou bigode, representando 856 indivíduos. Todas as imagens da base de dados encontram-se dispostas na escala de cinza, no formato PPM com resolução de 512×768 pixels. As faces possuem variação em escala e iluminação. As imagens possuem rótulos manuais para 4 pontos fiduciais. Na Figura 5.2 apresentamos alguns exemplos da base de dados *Feret*.



Figura 5.2: Exemplos de imagens da base dados *Feret* [19].

Nos experimentos desta dissertação, utilizaremos um subconjunto da base de dados *Feret* composto por 2004 imagens. Neste subconjunto, as imagens de indivíduos com óculos, barba e bigode, e as imagens com grandes rotações foram excluídas por inibirem a extração das características dos blocos das classes de positivos e negativos. Os experimentos utilizam 11 marcações manuais.

5.2 Projeto dos Classificadores C-SVC

Nesta dissertação, o projeto do classificador e as etapas dos procedimentos de treino e teste do sistema de detecção de pontos fiduciais são implementados

utilizando a ferramenta *OpenCV*¹ [57]. O *OpenCV* é utilizado devido aos recursos e ferramentas disponíveis no uso com imagens e vídeos, ao uso recente em pesquisas de visão computacional, além de utilizar a *LibSVM* como implementação do SVM. A *LibSVM* disponibiliza classificadores SVM com *kernel* linear e não linear.

A sintaxe responsável pelo treino do algoritmo SVM é apresentada na Tabela 5.1. Note que, a matriz de treino deve conter todos os vetores de características utilizados para o treino do classificador, a matriz de classes é responsável por informar os rótulos de cada vetor de características, as matrizes de variações e a de respostas a variações são utilizadas para construção de classificadores com *templates* deformáveis não utilizado neste trabalho, deste modo o valor 0 é inserido. Por fim, os parâmetros do SVM são responsáveis por informar o tipo de SVM utilizado (neste caso o C-SVC), o tipo de *kernel* e o valor do parâmetros de margem (*C*) representando o SVM margem suave.

Tabela 5.1: Função para Treinamento do SVM linear.

```
bool CvSVM::train
const CvMat* _train_data // matriz de treinamento;
const CvMat* _responses // matriz de classes;
const CvMat* _var_idx= 0 // matriz de variações;
const CvMat* _sample_idx= 0 //matriz de respostas a variações;
CvSVMParams _params=CvSVMParams() // parâmetros do SVM;
```

Além dos parâmetros descritos acima, para que o algoritmo de treino seja realizado é necessário inserir os critérios de encerramento do treino do classificador. No *OpenCV*, corresponde a função descrita na Tabela 5.2.

Tabela 5.2: Critérios de elaboração do classificador.

```
class TermCriteria
COUNT=1 // critério;
MAX_ITER=COUNT // número de iterações;
EPS=2 // tolerância;
```

Na Tabela 5.2, o critério de encerramento é apresentado de acordo com a abordagem da Seção 4.2.3. O treino do algoritmo do classificador SVM poderá ser encerrado ao atingir o número de iterações (*#it*), ao atingir a tolerância (ϵ), ou ainda

¹Na literatura, podemos encontrar outras implementações do SVM. Por exemplo, temos o *SVM-light* desenvolvido por Joachims [39,54], a *LibSVM* desenvolvido por Chih-Chung Chang e Chih-Jen Lin [55] e *SVMperf* desenvolvido por Teo [56].

ao atingir os valores de iterações e tolerância conjuntamente. Nesta dissertação, investigaremos o desempenho do sistema de detecção de pontos fiduciais utilizando o critério de encerramento com alcance dos valores de iterações e tolerância conjuntamente.

Além dos parâmetros, de ponderação de classes (C), do número de iterações ($\#it$) e da tolerância (ε) referentes ao algoritmo do SVM por C-SVC, nesta dissertação investigaremos o peso (W) atribuído a cada classe para ser levado em consideração durante o treino do algoritmo do SVM no sistema de detecção de pontos fiduciais.

Nesta investigação é necessário definirmos exatamente quais são os parâmetros utilizados e quais são os seus intervalos numéricos. Especificamente em nossa pesquisa, exploramos os parâmetros a seguir:

- Parâmetro de regularização C com valores iguais à 2^n com $n = \{-12, -11, \dots, 8\}$;
- Variáveis de folga ξ_i (veja Tabela 5.1) com valores iguais à 0, 001; 0, 01; 0, 1.
- Número máximo de iterações para convergência ($\#it$): 10000; 18000; 26000; 34000; 42000; 50000.
- Ponderação da quantidade de elementos de classe em proporção na ordem ($w_p : w_n$): (1:1), (1:5), (2:10), (10:500), (5:1), (10:2) e (500:10).

As faixas de variações investigadas dos parâmetros foram obtidas com experimentos iniciais. Em nossos estudos, identificamos por intermédio de experimentos quais os parâmetros do classificador C-SVC e do sistema de detecção de pontos fiduciais que mais influenciavam no desempenho do sistema. Nesta identificação de parâmetros, variamos os valores a uma extensa faixa, e para cada valor da faixa, um classificador foi gerado no procedimento de treino e em seguida realizamos o procedimento de teste para identificar o desempenho. Identificamos os limites inferiores e superiores de valores para cada parâmetro através a avaliação do desempenho dos classificadores gerados no procedimento de treino. Assim, notamos que acima dos valores do limite superior e abaixo dos valores do limite inferior de cada faixa de valores de parâmetros, o desempenho do sistema de detecção de pontos fiduciais não era influenciado sensivelmente. Neste procedimento, identificamos os parâmetros e

a respectiva faixa de valores apresentados acima. Verificamos que estas faixas influenciavam sensivelmente o desempenho do sistema de detecção de pontos fiduciais. Um aspecto importante é observar que cada um dos parâmetros e suas faixas de variações representam um experimento que deve ser explorado, ou seja, será um classificador gerado no procedimento de treino e em seguida o uso deste classificador para o procedimento de teste. Esta abordagem denomina-se *grid* de parâmetros abordado na Seção 4.2.3. Em nossa pesquisa, temos um extenso *grid* de parâmetros formado por 756 combinações por ponto fiducial pesquisado.

5.3 Métrica de Desempenho

Os resultados da classificação do sistema de detecção de pontos fiduciais são obtidos com o uso da métrica de desempenho do classificador. Neste caso, a métrica de desempenho será aplicada a cada experimento realizado no procedimento de teste para a detecção de 11 pontos fiduciais. Deste modo, identificaremos os parâmetros que apresentem os melhores desempenhos no sistema de detecção de pontos fiduciais.

O desempenho do classificador é medido utilizando-se uma métrica normalizada pela distância entre as coordenadas manuais (fornecida pela base de dados) dos centros dos olhos esquerdo e direito. Esta distância, denominada \tilde{d}_o , é obtida da seguinte forma:

$$\tilde{d}_o = \|\widetilde{\mathbf{OE}} - \widetilde{\mathbf{OD}}\| \quad (5.1)$$

onde: $\widetilde{\mathbf{OE}}$ e $\widetilde{\mathbf{OD}}$ corresponde respectivamente às coordenadas manuais do centro do olho esquerdo e do centro do olho direito fornecidas pela base de dados. Supondo um candidato \mathbf{x} e as coordenadas manuais do ponto fiducial igual a $\widetilde{\mathbf{PF}}$ então define-se $d_{\mathbf{PF}}$ como:

$$d_{\mathbf{PF}} = \|\mathbf{x} - \widetilde{\mathbf{PF}}\| \quad (5.2)$$

O desempenho do classificador é medido através das taxas de Verdadeiros Positivos (TP) e Falsos Positivos (FP) utilizando as expressões:

$$\text{TP} = 100 \frac{\text{Total de faces em que } d_{\mathbf{PF}} \leq 0,1 \tilde{d}_o}{\text{Total de faces}}, \quad (5.3)$$

$$FP = 100 \frac{\text{Total de candidatos a PF com } d_{\mathbf{PF}} > 0, 1 \tilde{d}_o}{\text{Total de negativos}}. \quad (5.4)$$

Com esta métrica de desempenho, avaliaremos o desempenho do sistema de detecção de pontos através da quantificação das taxas de TP e FP por imagem. O uso destas duas medidas de desempenho facilitam a identificação dos resultados por ponto fiducial classificado. Deste modo, o sistema de pontos fiduciais pode ter seus parâmetros ajustados para o melhor desempenho.

5.4 Validação Cruzada

A avaliação estatística do desempenho pode ser realizada utilizando validação cruzada [35]. A validação cruzada tem por objetivo prover o desempenho do sistema proposto para um novo conjunto de dados, neste caso, a base de testes. Em resumo, a validação cruzada é responsável por dividir a base de dados em dois subconjuntos mutuamente exclusivos. Destes subconjuntos, o primeiro será utilizado para o procedimento de treino e o segundo para o procedimento de teste.

Dentre os tipos validação cruzada, podemos citar o sorteio, o *leave-one-out* e o *k-fold*. O sorteio consiste na seleção de uma quantidade constante de dados para treino, e o restante das amostras é utilizado para teste. O *leave-one-out*, com N amostras, $N-1$ amostras são utilizadas para treinamento e a amostra restante para teste, processo realizado N vezes. Na validação por *k-fold*, a base de dados é dividida em k -subconjuntos, por sua vez, cada subconjunto é dividido em $k-1$ subconjuntos para treino e o restante destes subconjuntos para teste, ao final, é realizada a média aritmética do desempenho dos k subconjuntos, inferindo sobre o comportamento dos parâmetros. Para avaliação de desempenho do sistema, utilizamos o último tipo, devido a quantidade de *folds* e distribuição das imagens da base de dados em cada *fold*, os classificadores gerados no procedimento de treino são robustos às variações das imagens da base de dados utilizadas no procedimento de teste, tornando o desempenho do sistema de detecção de pontos fiduciais sem variações nos resultados por *fold*. Nesta pesquisa, a validação cruzada é feita por *k-fold*, onde o conjunto de dados foi dividido em 7 partes.

5.5 Experimentos

Os experimentos foram executados utilizando-se um cluster de computadores com 30 núcleos, cada um com memória *RAM* de 4GB e processador com 2,26GHz. O sistema operacional utilizado foi o Linux na versão Ubuntu 10.04. Em nossos experimentos, utilizamos dois subconjuntos divididos em 7 partes. As imagens, utilizadas nos experimentos, são agrupadas em conjuntos da seguinte maneira:

- Um conjunto formado por 503 imagens da base de dados *BioID* e;
- Um conjunto formado por 2004 imagens da base de dados *Feret*.

Foram utilizados 11 pontos fiduciais distintos, apresentados na Figura 5.3. Para cada ponto fiducial utilizamos o sistema descrito no Capítulo 4, ou seja, temos 11 sistemas de detecção distintos.



Figura 5.3: Pontos fiduciais da face humana usada nos experimentos.

Dentre os pontos fiduciais apresentados na Figura 5.3, a Tabela B apresenta as numerações e descrições dos 11 pontos fiduciais utilizados nos experimentos desta dissertação.

Nos experimentos, os parâmetros investigados para cada base de dados serão:

- Parâmetro de regularização C com valores iguais à 2^n com $n = \{-12, -11, \dots, 8\}$;
- Variáveis de folga ξ_i com valores iguais à 0,001; 0,01; 0,1.
- Número máximo de iterações para convergência ($\#it$): 10000; 18000; 26000; 34000; 42000; 50000.

Tabela 5.3: Numerações e descrições dos pontos fiduciais das bases de dados *BioID* e *Feret*

Pontos Fiduciais	
Número	Descrição
00	Canto Externo do Olho Direito
01	Centro do Olho Direito
02	Canto Interno do Olho Direito
03	Canto Interno do Olho Esquerdo
04	Centro do Olho Esquerdo
05	Canto Externo do Olho Esquerdo
06	Canto Direito do Nariz
07	Ponta do Nariz
08	Canto Esquerdo do Nariz
09	Canto Direito da Boca
10	Canto Esquerdo da Boca

- Ponderação da quantidade de elementos de classe em proporção na ordem $(w_p : w_n)$: (1:1), (1:5), (2:10), (10:500), (5:1), (10:2) e (500:10).

Para a análise de resultados de desempenho do sistema para o *grid* de parâmetros utilizaremos gráficos com o agrupamento dos pontos fiduciais por região da face humana, conforme as regiões a seguir:

- Região do olho direito (com os pontos fiduciais 0, 1 e 2);
- Região do olho esquerdo (com os pontos fiduciais 3, 4 e 5);
- Região do nariz (com os pontos fiduciais 6, 7 e 8);
- Região da boca (com os pontos fiduciais 9 e 10).

Cada gráfico será responsável por apresentar as características necessárias para a análise de desempenho para o *grid* de parâmetros, dentre as características, citamos:

- A taxa de verdadeiros positivos (TP);
- A taxa de falsos positivos (FP);
- O desvio padrão para cada taxa (TP e FP);
- Cada gráfico terá o parâmetro de folga (ξ_i) mantido fixo, assim teremos resultados para ξ_i com valores fixos em 0,001; 0,01; 0,1 .

Considerando todos os parâmetros para projeto do classificador C-SVC e os 11 pontos fiduciais da face serão realizados um total de 11088 experimentos:

- 8316 experimentos para a base de dados *BioID*;
- 2772 experimentos para a base de dados *Feret*.

O tempo de processamento é um aspecto relevante neste *setup* de experimentos. O período de processamento dos experimentos é avaliado sob a análise dos tempos de geração e validação do classificador, ou ainda, tempos de execução dos procedimentos de treino e de teste para os 7 *folds* de cada ponto fiducial. Deste modo, o tempo médio para a realização do experimento para um ponto fiducial é de 780 horas. O experimento mais custoso para esta dissertação foi o realizado no ponto fiducial 10, com o tempo de 1080 horas. Assim, atribuímos ao alto tempo de processamento a realização do procedimento de separação das classes do algoritmo SVM.

5.6 Resultados

Nesta seção, apresentamos os resultados obtidos nos experimentos do sistema de detecção de pontos fiduciais. O desempenho do sistema para o *grid* de parâmetros para cada ponto fiducial é apresentado a seguir. Dividimos a apresentação dos resultados em duas subseções, sendo uma para cada base de dados. Em cada subseção será apresentada uma análise dos parâmetros e em seguida a comparação com outros métodos. Os gráficos com os desempenhos para os 11 pontos fiduciais das bases de dados *BioID* e *Feret* são divididos em regiões. A divisão é feita em região do olho direito (com os pontos fiduciais 0, 1 e 2), região do olho esquerdo (com os pontos fiduciais 3, 4 e 5), região do nariz (com os pontos fiduciais 6, 7 e 8) e por fim, a região da boca (com os pontos fiduciais 9 e 10). Nos gráficos, apresentamos o desempenho do *grid* de parâmetros para as taxas de Verdadeiros Positivos (TP) e Falsos Positivos (FP) com os respectivos desvios padrões. Os parâmetros C , número máximo de iterações para convergência ($\#it$) e ponderação da quantidade de elementos de classe (W) são variados na sequência de valores apresentada neste capítulo, e o parâmetro de folga (ξ_i) é mantido com o valor fixo. Deste modo, cada gráfico apresentará 252 resultados distintos.

5.6.1 Análise de Resultados: Conjunto de Imagens da Base de Dados *BioID*

A seguir, apresentamos o desempenho do sistema de detecção de pontos fiduciais para o *grid* de parâmetros com 11 pontos utilizando as bases de dados *BioID*. Os resultados por base de dados são divididos em 4 regiões. A primeira região é a do olho direito, em seguida o olho esquerdo, a seguir o nariz e por fim a região da boca. Nas Figuras 5.4, 5.5, 5.6 e 5.7, apresentamos os resultados para a base de dados *BioID*, com valor de ξ_i igual à 0, 1.

Em seguida, nas Figuras 5.8, 5.9, 5.10 e 5.11 apresentamos os resultados para a base de dados *BioID*, com valor de ξ_i igual à 0, 01. Por fim, nas Figuras 5.12, 5.13, 5.14 e 5.15 apresentamos os resultados para a base de dados *BioID*, com valor de ξ_i igual à 0, 001.

Nas Figuras 5.4 a 5.15 apresentamos os resultados obtidos utilizando a base de dados *BioID*. Após análise das figuras podemos verificar que os resultados para ξ_i igual à 0, 1:

- Na Figura 5.4, os resultados para a região do olho direito são apresentados. Observamos que os melhores desempenhos são aqueles que a taxa TP apresenta resultados acima de 90% e a taxa de FP é inferior a 2,5%. Nos gráficos identificamos combinações de parâmetros que apresentem estes bons resultados nas 84 primeiras combinações. Nas demais partes dos gráficos as taxas de TP e FP são elevadas. Dessa forma, podemos inferir que o sistema não possui desempenho satisfatório para este grupo de combinações.
- Na Figura 5.5 os resultados da região do olho esquerdo são apresentados, e de maneira semelhante apresentado na região do olho direito, os melhores desempenhos são apresentados para as 84 primeiras combinações dos valores dos parâmetros. Note que os pontos 3 e 4 apresentam desempenhos de TP próximos a 100%. Assim, inferimos que neste grupo de combinações o sistema apresenta ótimo desempenho.
- Na Figura 5.6 os resultados para a região do nariz são apresentados e observamos que os desempenhos são oscilantes, e os melhores resultados são notados nas 42 primeiras combinações. Note que para o ponto fiducial 8 o desempenho

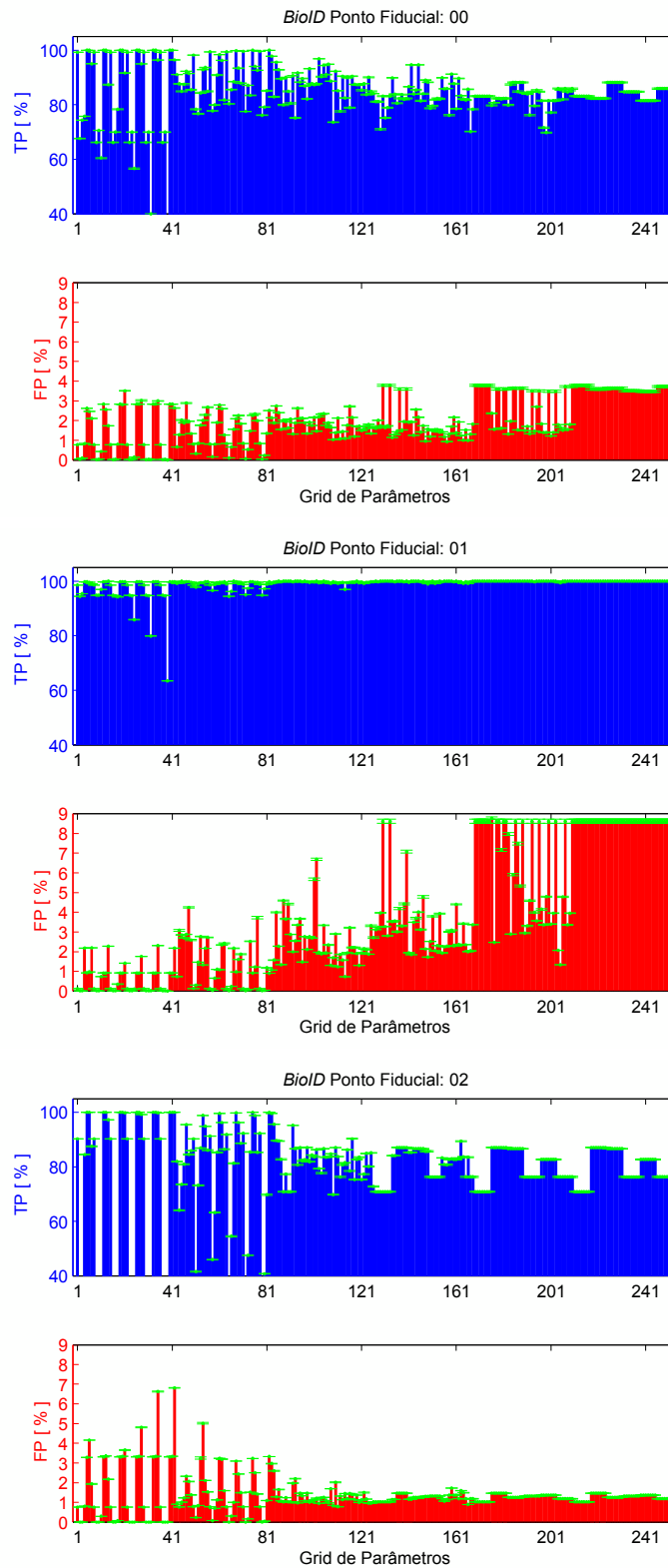
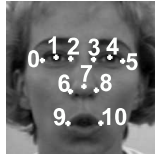


Figura 5.4: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *BioID*. Os resultados são apresentados para a região do olho direito são para os pontos fiduciais 0, 1 e 2. Neste *grid* de parâmetros, o valor de ξ_i igual à 0, 1.

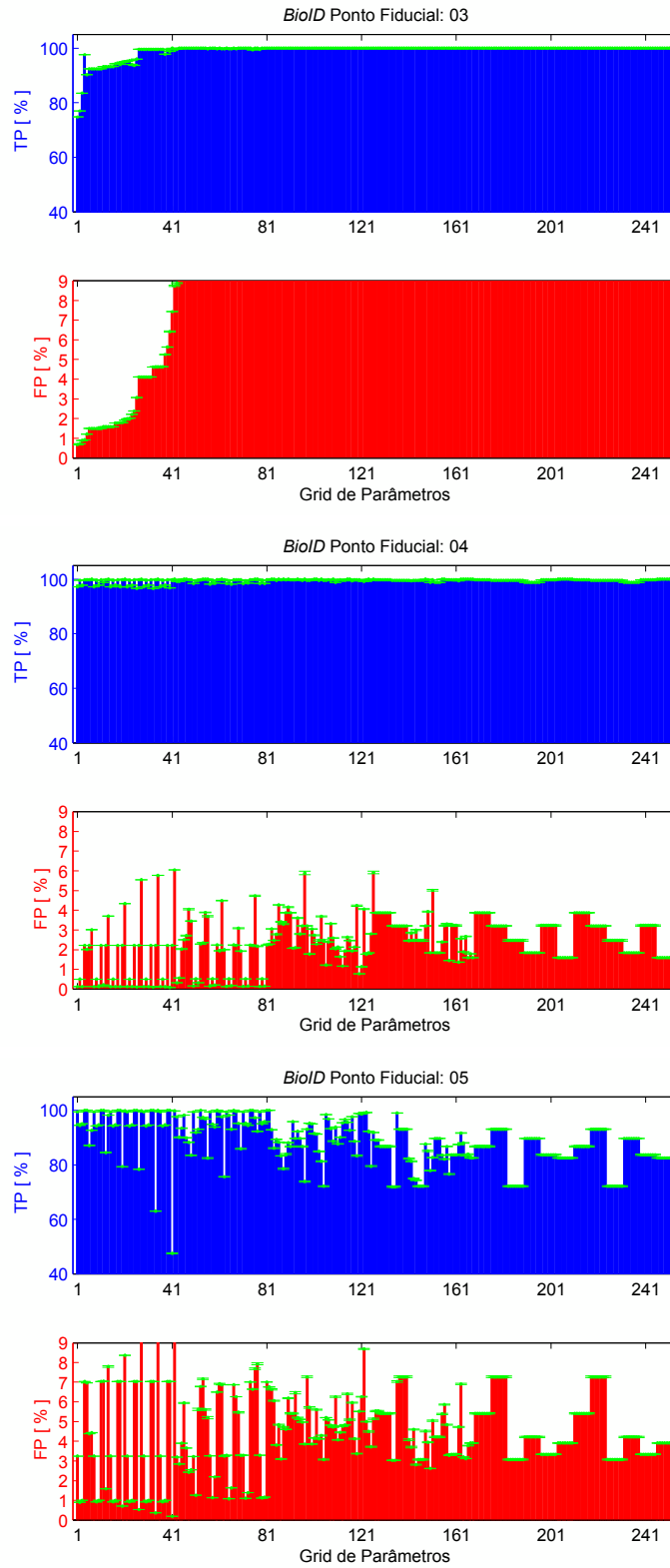
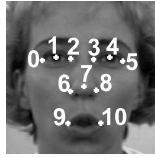


Figura 5.5: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *BioID*. Os resultados são apresentados para a região do olho esquerdo são para os pontos fiduciais 3, 4 e 5. Neste *grid* de parâmetros, o valor de ξ_i igual à 0, 1.

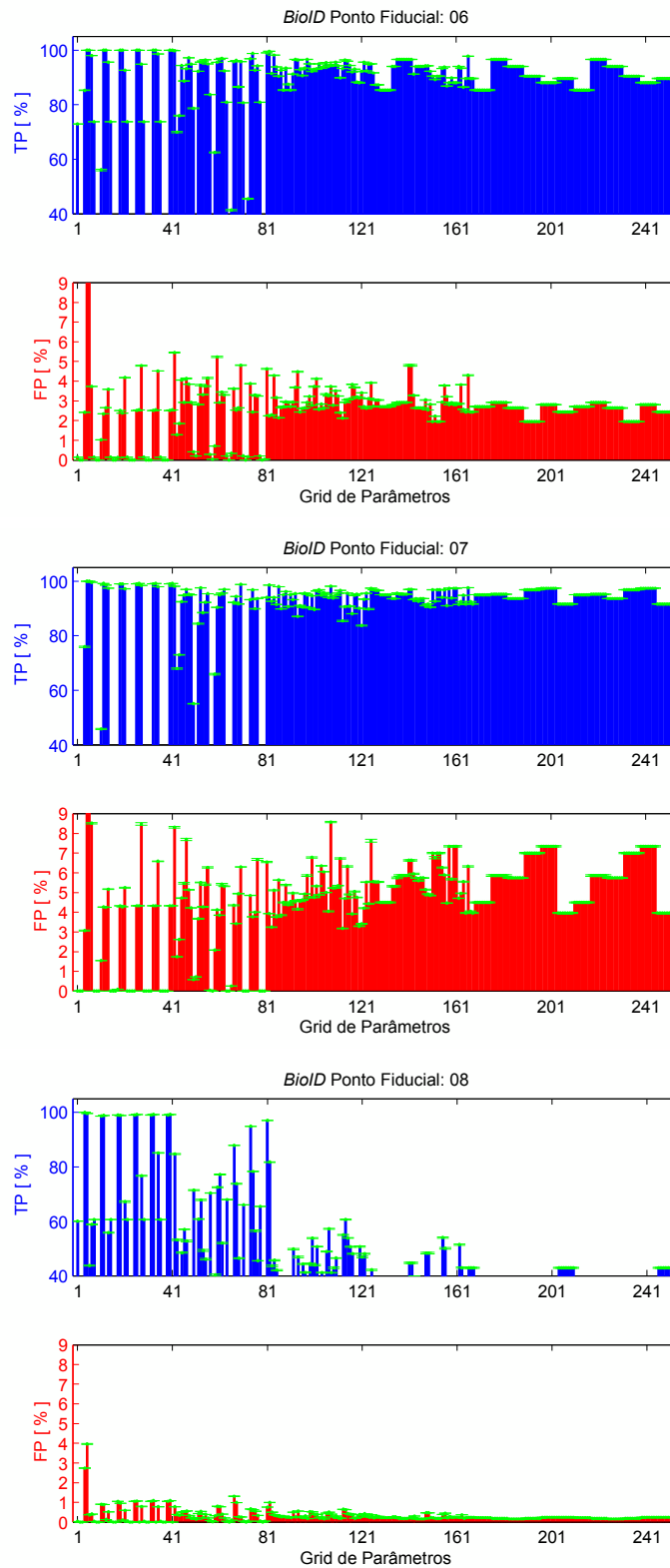
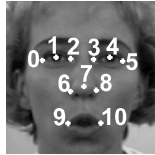


Figura 5.6: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *BioID*. Os resultados são apresentados para a região do nariz são para os pontos fiduciais 6, 7 e 8. Neste *grid* de parâmetros, o valor de ξ_i igual à 0, 1.

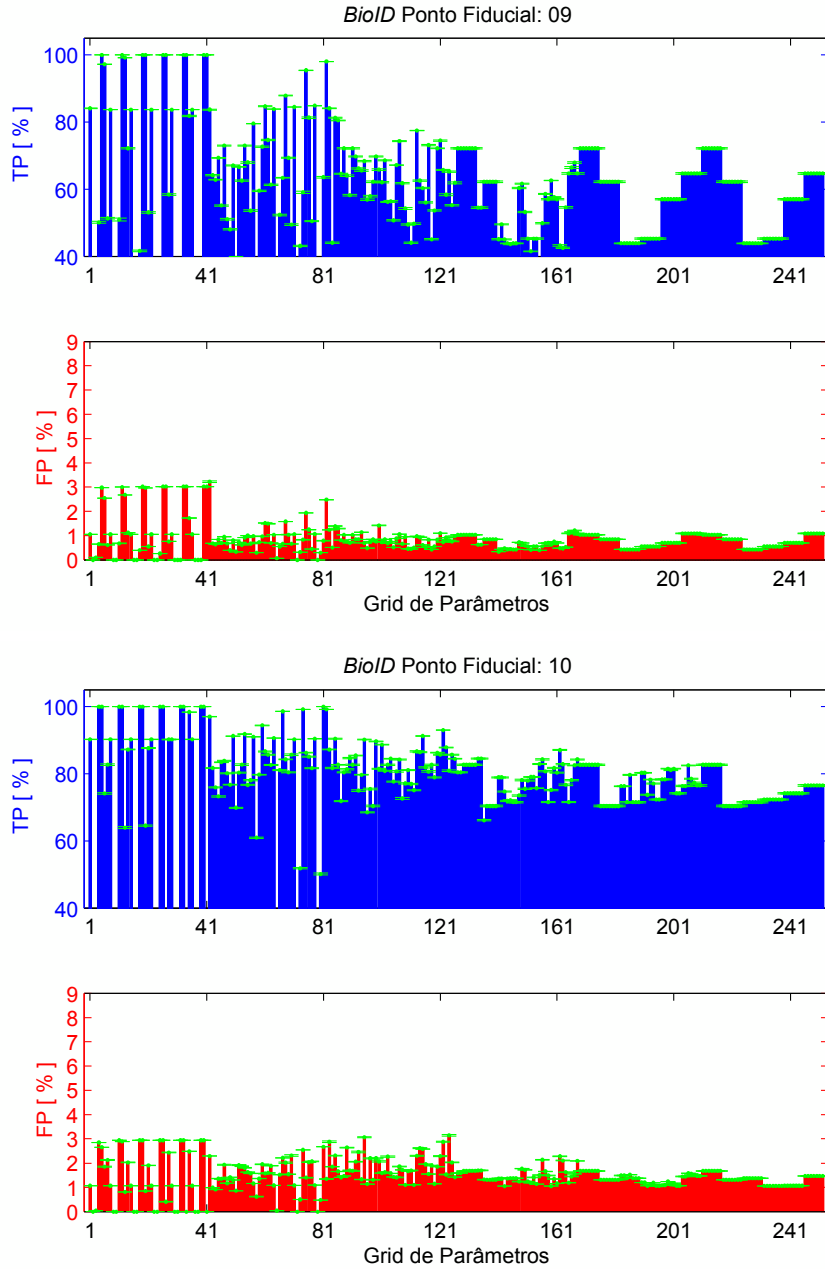
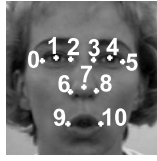


Figura 5.7: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *BioID*. Os resultados são apresentados para a região da boca são para os pontos fiduciais 9 e 10. Neste *grid* de parâmetros, o valor de ξ_i igual à 0, 1.

em geral dos classificadores possuem baixas taxas de TP e altas taxas de FP. Deste modo, podemos inferir que o sistema possui desempenho satisfatório somente para os grupo de combinações iniciais.

- Na região da boca o gráfico é apresentado na Figura 5.7. Note que os melhores desempenhos estão localizados nas 84 primeiras combinações do *grid* de parâmetros. Note que a taxa de FP possui valores pequenos em toda a extensão do *grid*, entretanto a taxa de TP não atende as especificações após as 84 primeiras combinações do *grid*. Dessa forma, concluímos que no grupo inicial de combinações o sistema apresenta ótimo desempenho.

Em seguida, analisamos os resultados para ξ_i igual à 0,01, apresentados nas Figuras 5.8 até 5.11:

- Na Figura 5.8 observamos que os melhores resultados para a região do olho direito estão concentrados nas 84 primeiras combinações. É notável que o desempenho do ponto fiducial 01 apresenta valores da taxa de TP próximos a 100% no decorrer do *grid*, entretanto a taxa de FP apresenta valores elevados. Dessa forma, podemos inferir que o sistema neste grupo de combinações apresenta ótimo desempenho.
- Na Figura 5.9 os melhores desempenhos são apresentados para as 84 primeiras combinações dos valores dos parâmetros. Note que o desempenho para todos os pontos desta região apresentam valores próximos a 100%, entretanto o desempenho da taxa de FP possui valores elevados, principalmente quando analisamos o ponto fiducial 03. Deste modo, podemos inferir que o sistema para este grupo de das primeiras 84 combinações apresenta desempenho satisfatório.
- Os resultados para a região do nariz na Figura 5.10 mostram que os melhores resultados são notados nas 41 primeiras combinações. Nesta região, as taxas de TP e FP são elevadas para os pontos 06 e 07. No ponto 08, o desempenho para a taxa de TP são baixos a partir das 84 primeiras combinações do *grid*, enquanto que a taxa de FP são constantemente abaixo de 2,5%.

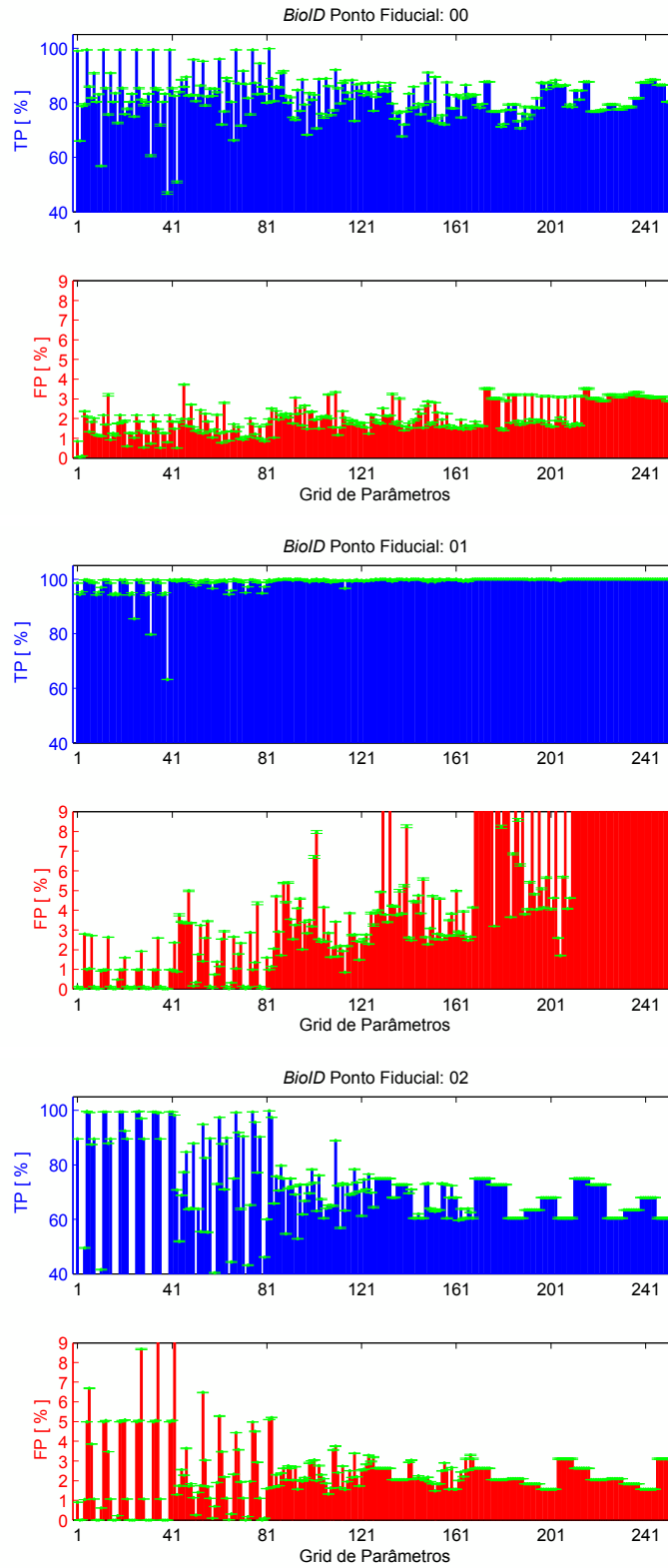
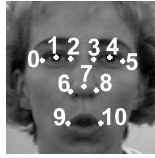


Figura 5.8: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *BioID*. Os resultados são apresentados para a região do olho direito são para os pontos fiduciais 0, 1 e 2. Neste *grid* de parâmetros, o valor de ξ_i igual à 0,01.

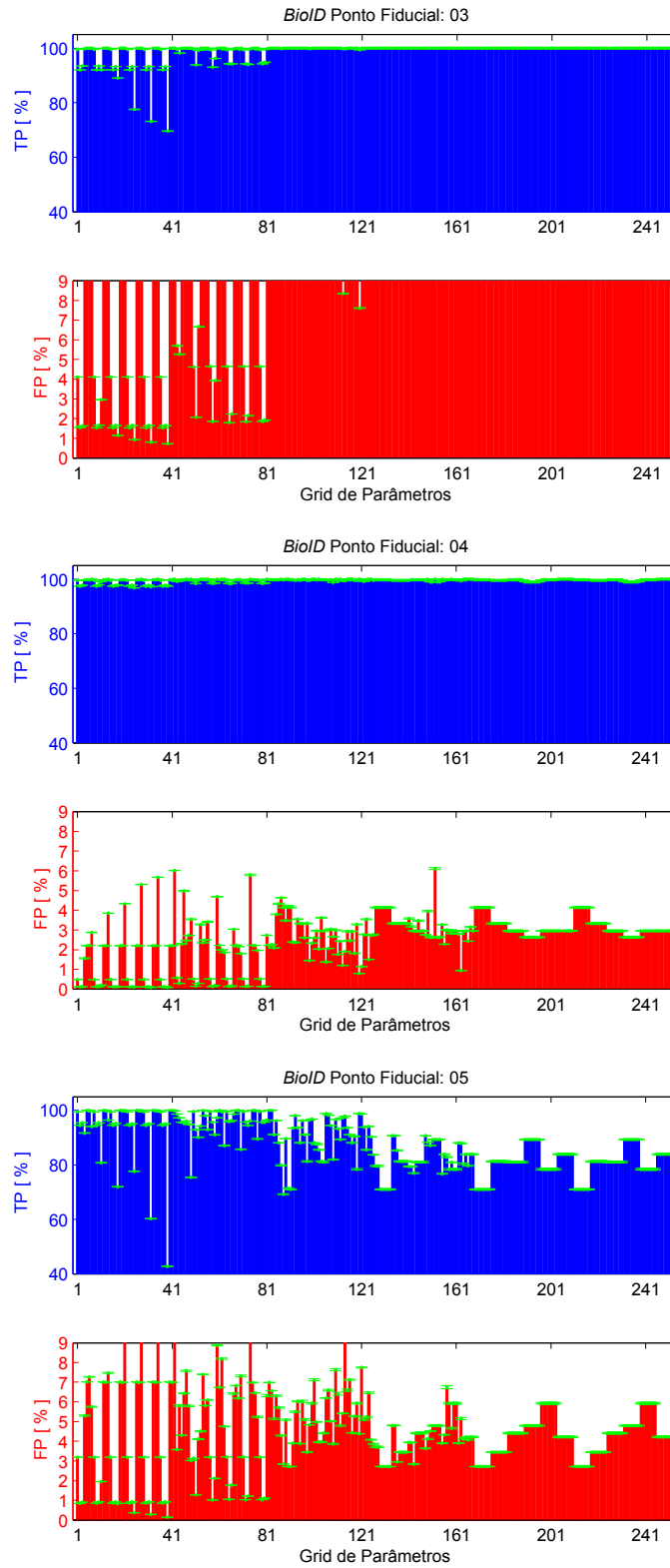
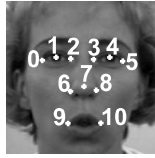


Figura 5.9: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *BioID*. Os resultados são apresentados para a região do olho esquerdo são para os pontos fiduciais 3, 4 e 5. Neste *grid* de parâmetros, o valor de ξ_i igual à 0,01.

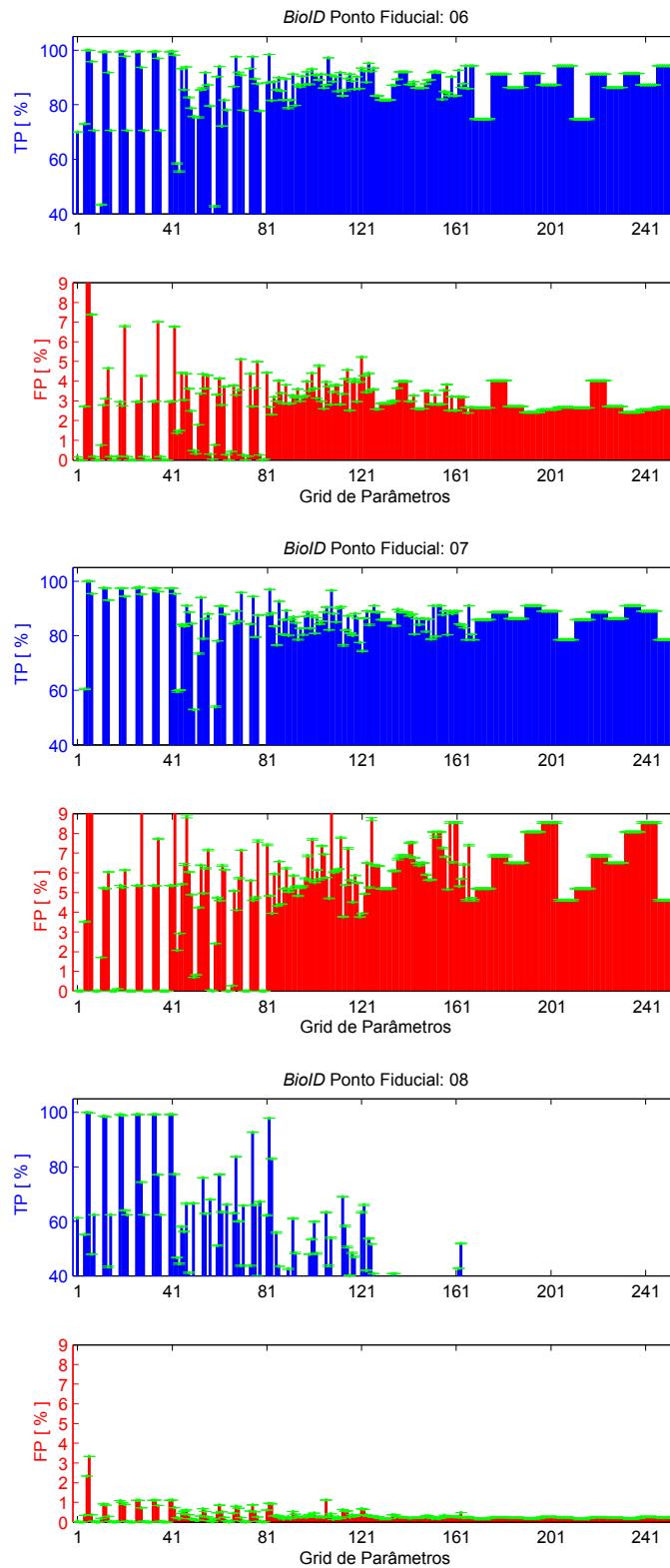
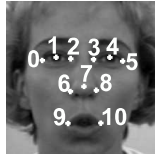


Figura 5.10: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *BioID*. Os resultados são apresentados para a região do nariz são para os pontos fiduciais 6, 7 e 8. Neste *grid* de parâmetros, o valor de ξ_i igual à 0,01.

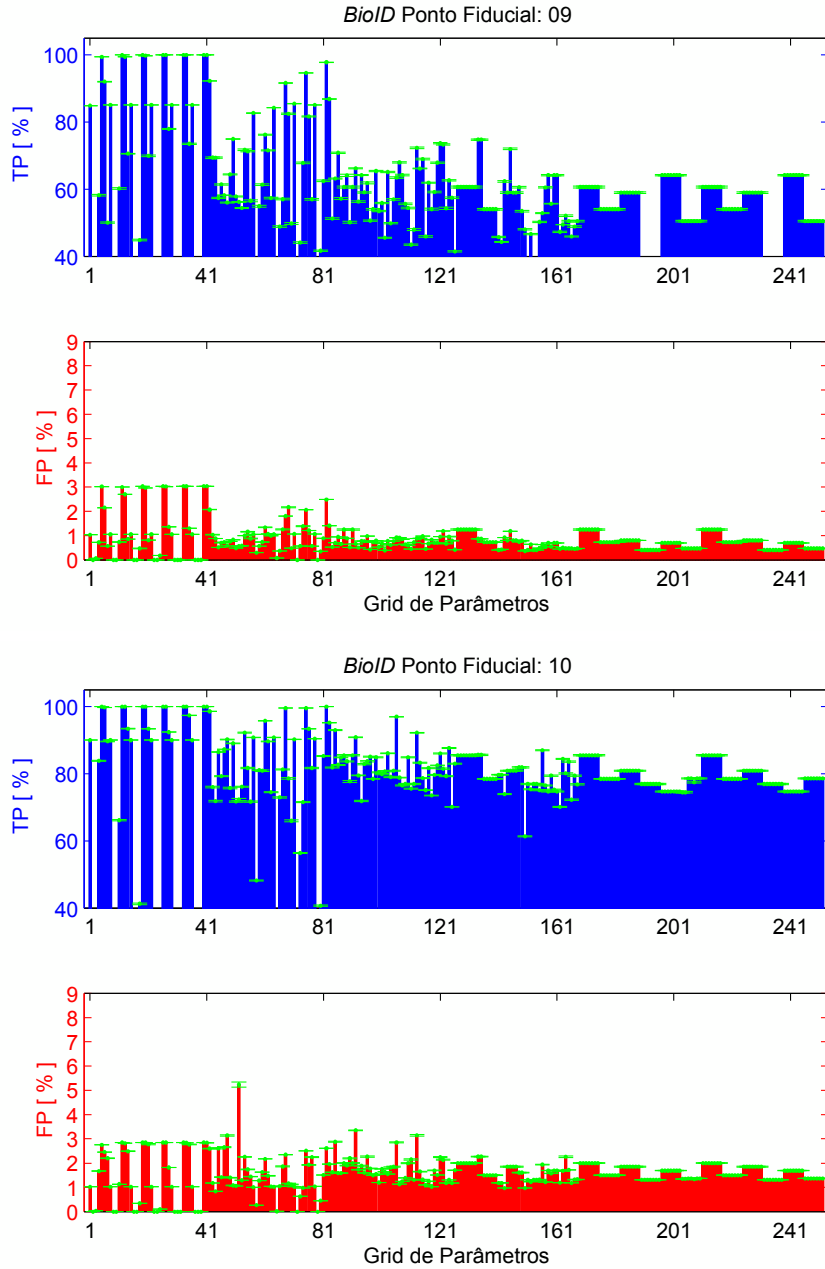
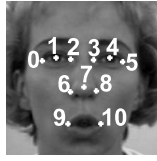


Figura 5.11: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *BioID*. Os resultados são apresentados para a região da boca são para os pontos fiduciais 9 e 10. Neste *grid* de parâmetros, o valor de ξ_i igual à 0,01.

- Na região da boca o gráfico é apresentado na Figura 5.11. Note que os melhores desempenhos estão localizados nas 42 primeiras combinações do *grid* de parâmetros. As taxas de FP para esta região é considerada aceitável na extensão do *grid* de parâmetros. Assim, podemos inferir que para este grupo de combinações o desempenho do sistema é satisfatório.

Por fim, analisamos os resultados para ξ_i igual à 0,001, apresentados nas Figuras 5.12 até 5.15:

- Na Figura 5.12, nas 84 primeiras combinações observamos os melhores desempenhos. Os resultados apresentados em geral apresentam bastantes classificadores com valores elevados para a taxa de TP, assim como alguns resultados de FP maiores que 2%. Assim, podemos inferir que o desempenho do sistema para este grupo de combinações é considerado bom.
- Os desempenhos da região do olho esquerdo são apresentados na Figura 5.13 e as 42 primeiras combinações dos valores dos parâmetros são notadas como os melhores resultados. Dessa forma, inferimos que o sistema para esta combinação de parâmetros apresenta desempenho satisfatório.
- Na Figura 5.14 apresentamos os melhores resultados nas 84 primeiras combinações do *grid* de parâmetros. Notamos que a taxa de FP é acima de 3% na extensão do *grid* para os pontos fiduciais 06 e 07, enquanto que no ponto fiducial 08 identificamos que o maior percentual de classificadores apresentam taxas de FP abaixo de 2%. Dessa forma, podemos inferir que o sistema para esta combinação de parâmetros não apresenta desempenho satisfatório.
- Na Figura 5.15 note que os melhores desempenhos estão localizados nas 84 primeiras combinações do *grid* de parâmetros. As taxas de TP apresentam muita oscilação para o ponto fiducial 09, enquanto que as taxas de FP para esta região é considerada aceitável na extensão do *grid* de parâmetros.

Em resumo, nestes experimentos propomos observar o desempenho do sistema de detecção de pontos fiduciais para a base de dados *BioID* na extensão do *grid* de parâmetros. Apresentamos os gráficos com os pontos fiduciais divididos em 4 regiões (olho direito, olho esquerdo, boca e nariz). Além disso, dividimos para cada

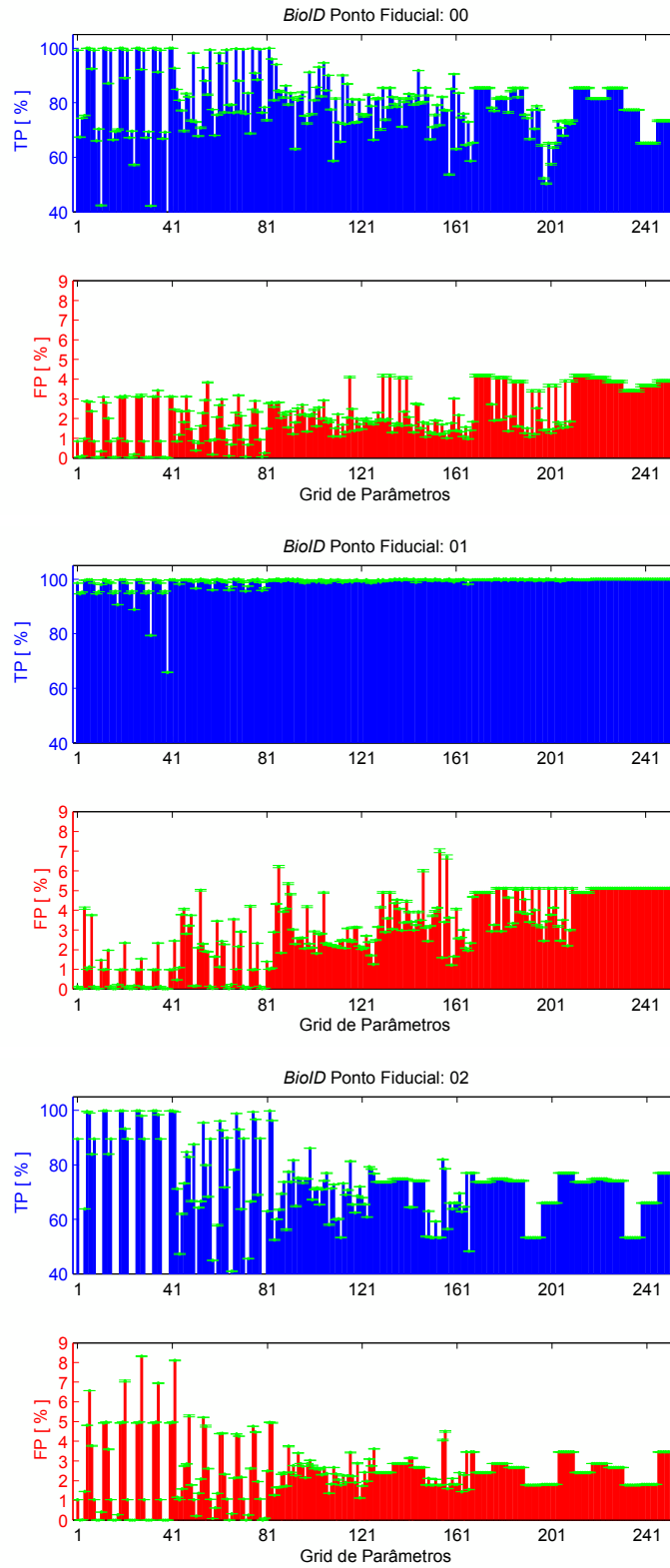
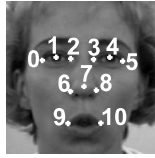


Figura 5.12: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *BioID*. Os resultados são apresentados para a região do olho direito são para os pontos fiduciais 0, 1 e 2. Neste *grid* de parâmetros, o valor de ξ_i igual à 0,001.

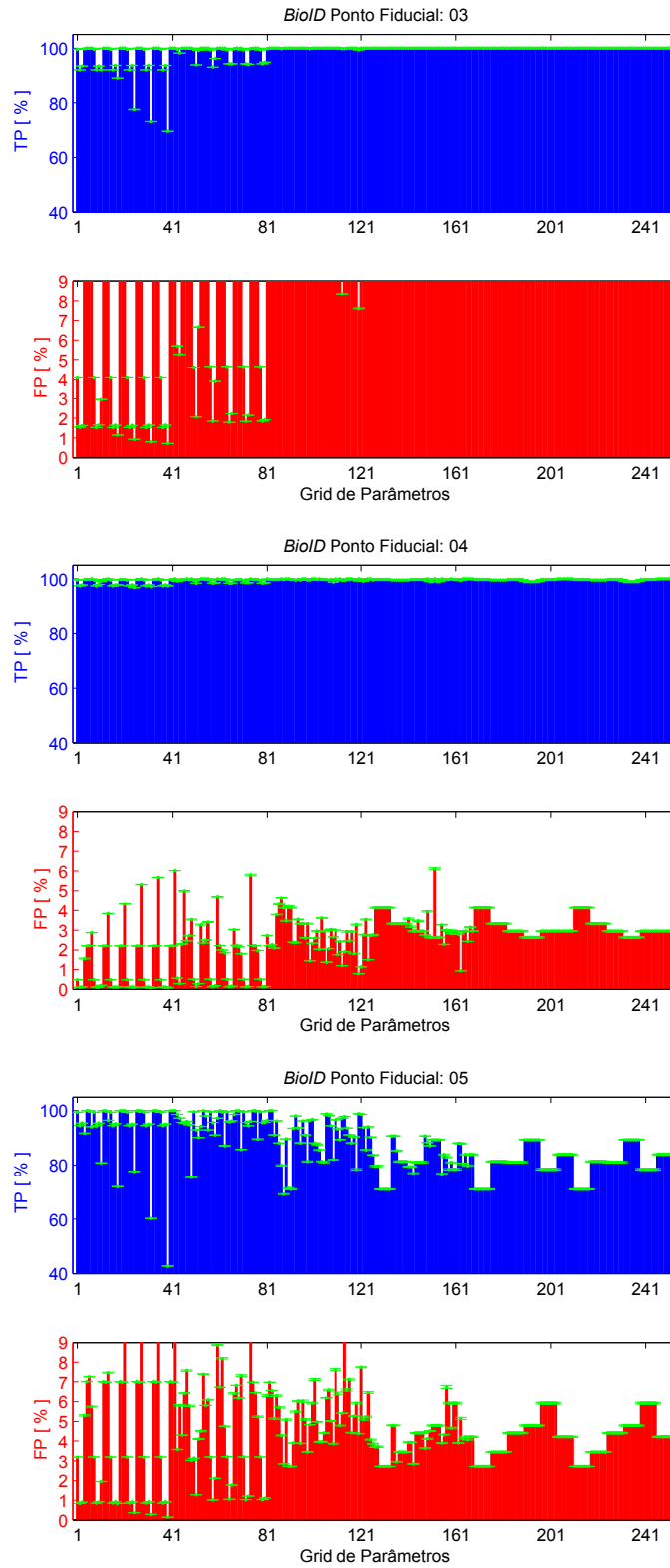
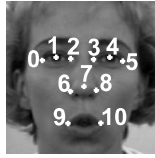


Figura 5.13: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *BioID*. Os resultados são apresentados para a região do olho esquerdo são para os pontos fiduciais 3, 4 e 5. Neste *grid* de parâmetros, o valor de ξ_i igual à 0,001.

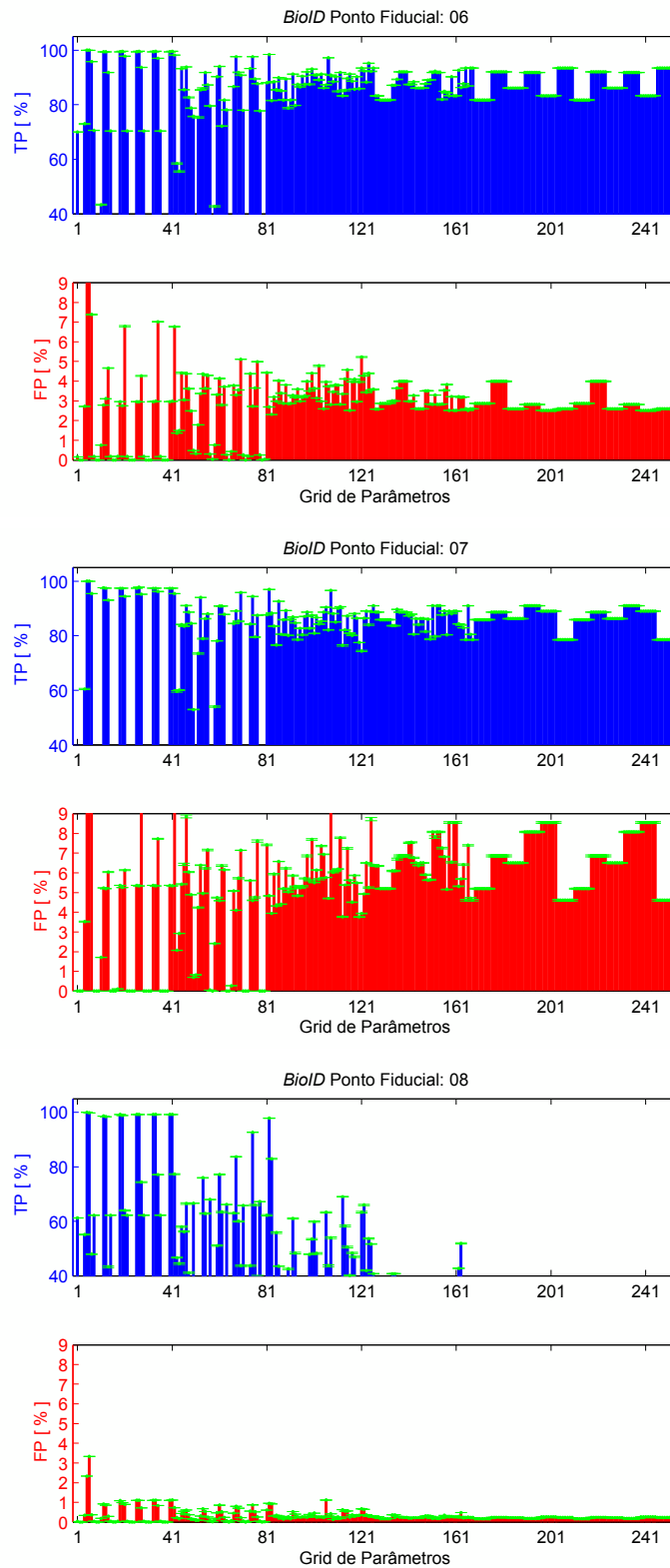
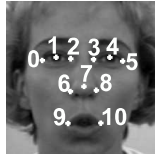


Figura 5.14: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *BioID*. Os resultados são apresentados para a região do nariz são para os pontos fiduciais 6, 7 e 8. Neste *grid* de parâmetros, o valor de ξ_i igual à 0,001.

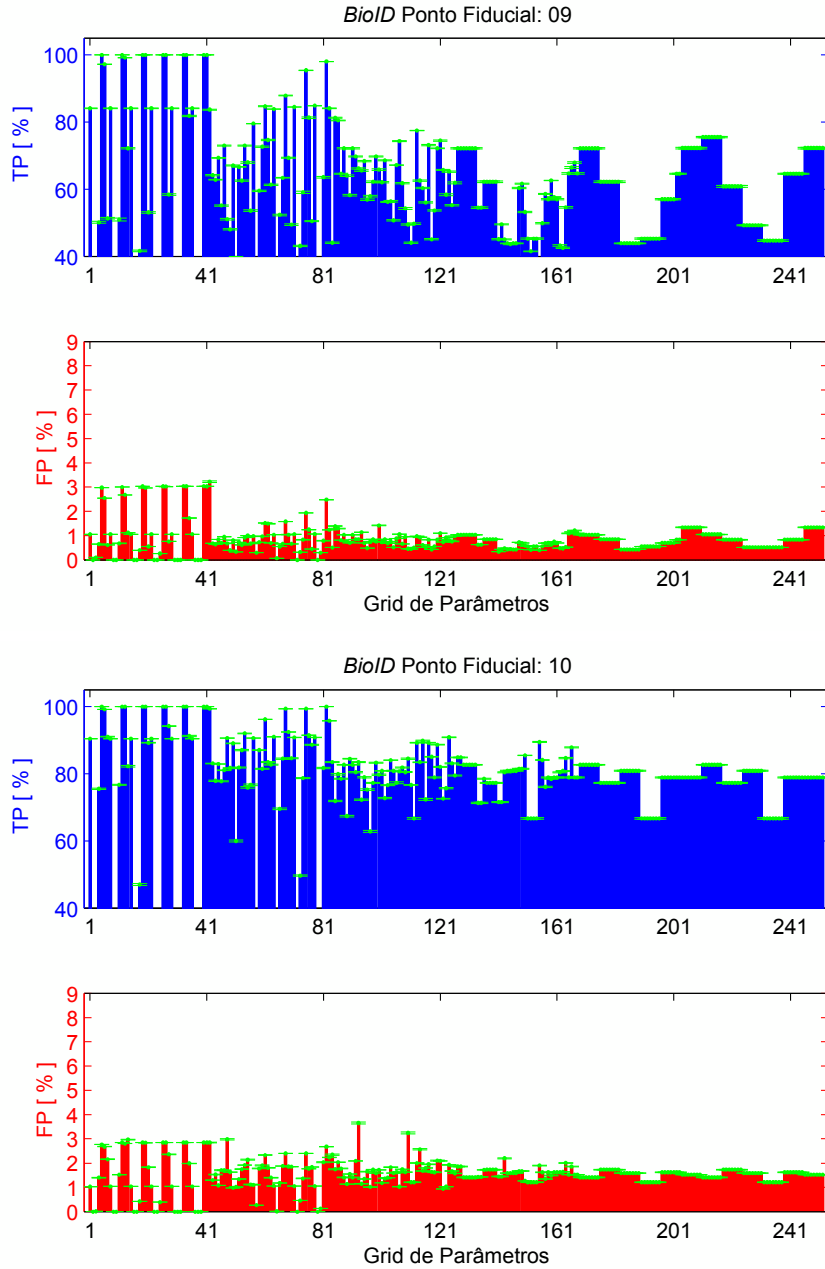
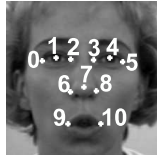


Figura 5.15: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *BioID*. Os resultados são apresentados para a região da boca são para os pontos fiduciais 9 e 10. Neste *grid* de parâmetros, o valor de ξ_i igual à 0,001.

ponto três análises, cada uma utilizando a variável de folga fixa em um valor (0, 01, 0, 001 e 0, 1). Assim, identificamos no desenvolvimento dos experimentos que o desempenho do sistema para todas as combinações dos parâmetros apresenta variações nas taxas de TP e FP. Entretanto, observamos que os classificadores com desempenho satisfatório estão inseridos dentre as 84 primeiras combinações e no grupo restante de combinações podemos inferir que o desempenho não é satisfatório, principalmente por apresentar valores elevados da taxa FP. Dentre as combinações de parâmetros que apresentam desempenho satisfatório do sistema, identificamos os valores de C com valores iguais à 2^{-12} e 2^{-9} , e as demais variações dos parâmetros de ponderação e do número de iterações.

5.6.2 Comparação entre Métodos: Conjunto de Imagens da Base de Dados *BioID*

Com o objetivo de comparar o desempenho do sistema de detecção de pontos fiduciais proposto com outros sistemas, identificamos os melhores desempenhos do *grid* de parâmetros e tabulamos. Os resultados são apresentados na Tabela 5.4 com a aplicação dos melhores parâmetros do *grid* para 11 pontos fiduciais da base de dados *BioID*. Os valores apresentados na tabela são as médias das taxas de TP e FP (segundo as Equações (5.3) e (5.4) e os respectivos desvios padrões para identificarmos a variação do valor do desempenho para os 7 *folds*. De modo geral, a situação ideal é a que os resultados apresentem desempenhos elevados na taxa de TP e os menores desempenhos possíveis quando tratamos da taxa de FP.

Na Tabela 5.4, note que para comparação de desempenho utilizamos os resultados dos métodos por filtragem discriminativa com e sem análise de componentes principais e detecção por produto interno apresentados em [20] e de um sistema que utiliza classificador SVM com *kernel* linear. Nesta dissertação, o pós-processamento não é empregado, entretanto o cálculo das taxas de positivos e negativos (TP e FP) são calculadas considerando 10% ao redor do ponto fiducial.

Analizamos os resultados da Tabela 5.4 e agrupamos os comentários por região, semelhante ao apresentado na Seção de análise dos resultados gráficos. Assim, identificamos que:

- Para a região do olho direito, note que desempenho do sistema proposto é competitivo quando comparado aos demais métodos, pois apresentamos taxas de TP superiores e as taxas de FP são abaixo de 2,0%. Além disso, podemos observar que os resultados apresentados são superiores ao dos métodos DF e SVM L. Por fim, os resultados são equivalentes quando comparados aos métodos DF-PCA e IPD, os dois com taxas de FP baixas;
- O desempenho do sistema proposto para a região do olho esquerdo apresenta melhores resultados para a taxa de TP (valores acima de 97%) quando comparados aos demais métodos. As taxas de FP do método proposto é abaixo de 2,0%.
- Na região do nariz, as taxas de TP do método proposto apresentam valores acima de 90%, exceto para o ponto fiducial 07. As taxas de FP são consideradas são consideradas razoáveis. Para os pontos fiduciais 06 e 08 identificamos bons resultados. Os pontos 06 e 07 apresentaram um alto desvio padrão. Observamos que os métodos DF-PCA e IPD apresentam desempenhos elevados. E por fim, os métodos DF e SVM L apresentam desempenhos fracos quando comparado ao método proposto;
- As taxas de TP do método proposto para a região da boca apresenta valores acima de 99% e com a taxa de FP abaixo de 2,0% tornando o método proposto competitivo. Note que o desempenho do método proposto supera o desempenho dos métodos SVM L e DF, enquanto que apresenta desempenho equivalente aos métodos DF-PCA e IPD.

Em resumo, podemos inferir os comentários abaixo para a comparação com os demais métodos:

- DF: o método proposto supera o desempenho do sistema utilizando o método DF em todos os pontos fiduciais quando comparamos os valores das taxas de TP, com exceção ao ponto 07. Note que os resultados apresentados para os desvios padrões são competitivos. Nesta comparação, inferimos que o desempenho do sistema de detecção de pontos é ótimo.

- SVM linear: o método SVM L possui o desempenho do sistema superado para as taxas de TP e FP nos pontos fiduciais 00, 01, 02, 04, 05, 09 e 10. Nos pontos fiduciais 03, 06, e 08, note que o SVM L apresenta desempenho superior nas taxas de TP e FP. Além disso, podemos inferir que o desempenho do ponto fiducial de 07 para ambos os sistemas não são satisfatórios.
- DF-PCA: no método proposto, os pontos 00, 01, 03, 04, 05, 08, 09 e 10 apresentam o desempenho de TP superior ao método DF-PCA. Quando comparamos o desempenho de FP, identificamos que o método DF-PCA apresenta ótimos valores, abaixo de 0,3%, sendo superior ao método C-SVC. Assim, podemos inferir que o desempenho dos métodos são competitivos, uma vez que os desvios padrões do método proposto são menores que os demais em comparação ao DF-PCA, com exceção para os pontos fiduciais 02, 06 e 07.
- IPD: Identificamos que o método proposto apresenta melhor desempenho na taxa de TP para os pontos fiduciais 00, 01, 03, 04, 05, 09 e 10 em comparação ao método IPD. Este método por sua vez apresenta ótimo desempenho quando tratamos da taxa de FP, pois os valores são abaixo de 0,10%. Valores de FP abaixo de 2,0% são considerados satisfatórios, esse fato também pode ser observado nos resultados do método proposto, exceto no ponto fiducial 9. Assim, podemos inferir que o desempenho do método IPD é ligeiramente superior ao método proposto.

5.6.3 Análise de Resultados: Conjunto de Imagens da Base de Dados *Feret*

A seguir, apresentamos o desempenho do *grid* de parâmetros para a base de dados *Feret*. Note que os resultados apresentados são somente para a variável de folga fixa em 0,1, pois de acordo com os experimentos realizados com a base de dados *BioID*, identificamos que os resultados para as variações deste parâmetro são equivalentes. Nas Figuras 5.16, 5.17, 5.18 e 5.19 apresentamos os resultados para as quatro principais principais da face humana.

Analizamos os resultados para a base de dados *Feret*, apresentados nas Figuras 5.16 até 5.19:

Tabela 5.4: Desempenho do sistema proposto para 11 pontos usando base de dados *BioID*.

PF	DF-PCA [58]			DF [58]			IPD [58]			SVM L [58]			C-SVC (Proposto)		
	FP	TP	σ	FP	TP	σ	FP	TP	σ	FP	TP	σ	FP	TP	σ
00	0,05	0,03	2,2	0,22	0,03	4,7	0,05	0,03	1,3	1,09	0,74	9,8	0,76	0,07	0,7
01	0,10	0,05	2,8	0,30	0,04	6,8	0,05	0,03	1,3	0,32	0,09	1,0	0,40	0,03	0,5
02	0,24	0,14	2,1	0,39	0,03	6,1	0,05	0,02	1,0	1,58	0,59	14,3	0,76	0,17	4,2
03	0,08	0,03	2,5	0,30	0,03	13,7	0,05	0,02	1,9	0,46	0,21	1,7	0,92	0,40	1,7
04	0,11	0,04	2,8	0,40	0,02	9,6	0,03	0,02	0,9	2,49	4,76	4,2	0,50	0,11	0,7
05	0,05	0,02	2,5	0,23	0,04	12,4	0,03	0,02	2,1	2,27	0,62	3,1	0,31	0,14	0,7
06	0,14	0,03	1,6	0,25	0,04	9,3	0,06	0,02	1,5	1,40	1,44	1,0	1,95	0,79	11,7
07	0,18	0,02	3,7	0,20	0,01	8,1	0,10	0,02	1,8	0,06	0,03	9,6	0,58	0,28	12,7
08	0,17	0,02	2,6	0,23	0,01	5,0	0,03	0,01	1,7	0,40	0,09	3,9	0,75	0,15	2,3
09	0,18	0,04	3,0	0,22	0,01	5,5	0,16	0,04	2,3	2,91	5,12	5,5	2,08	0,33	1,1
10	0,18	0,02	2,9	0,11	0,04	6,2	0,20	0,05	3,0	1,47	0,49	1,9	1,34	0,50	0,7

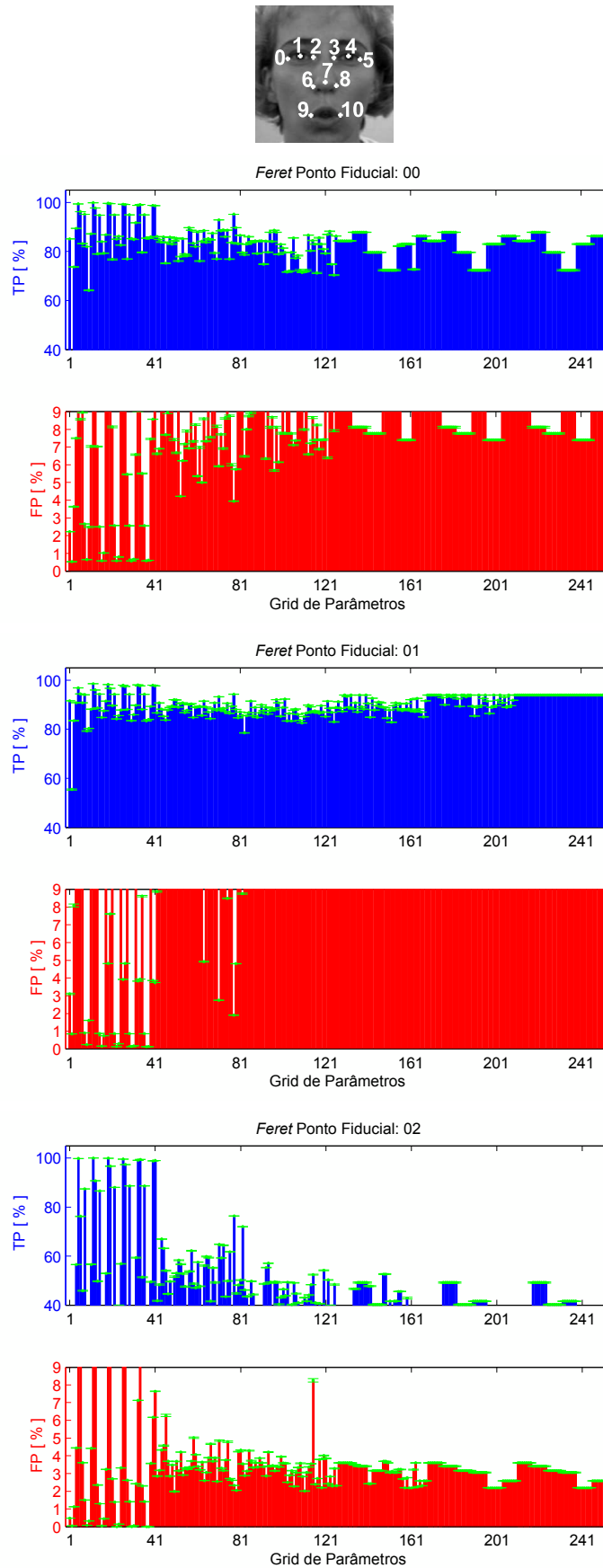


Figura 5.16: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *Feret*. Os resultados são apresentados para a região do olho direito são para os pontos fiduciais 0, 1 e 2. Neste *grid* de parâmetros, o valor de ξ_i igual à 0, 1.

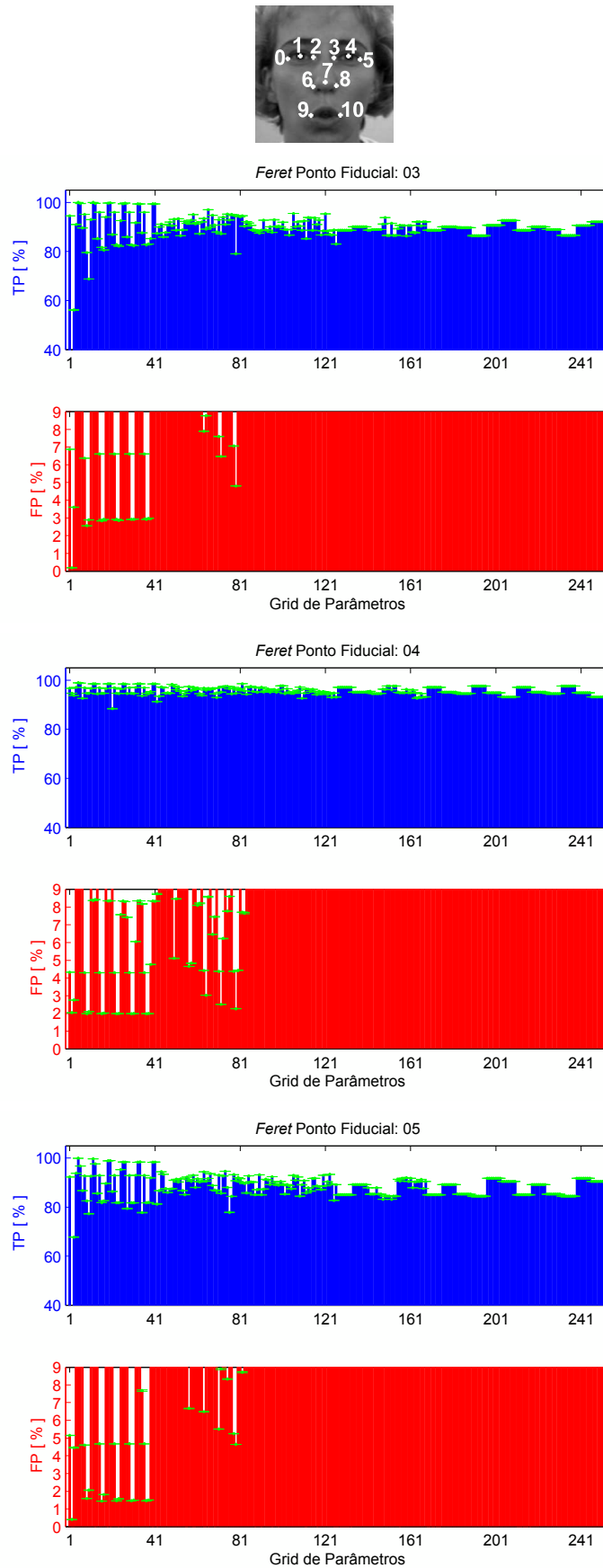


Figura 5.17: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *Feret*. Os resultados são apresentados para a região do olho esquerdo são para os pontos fiduciais 3, 4 e 5. Neste *grid* de parâmetros, o valor de ξ_i igual à 0, 1.

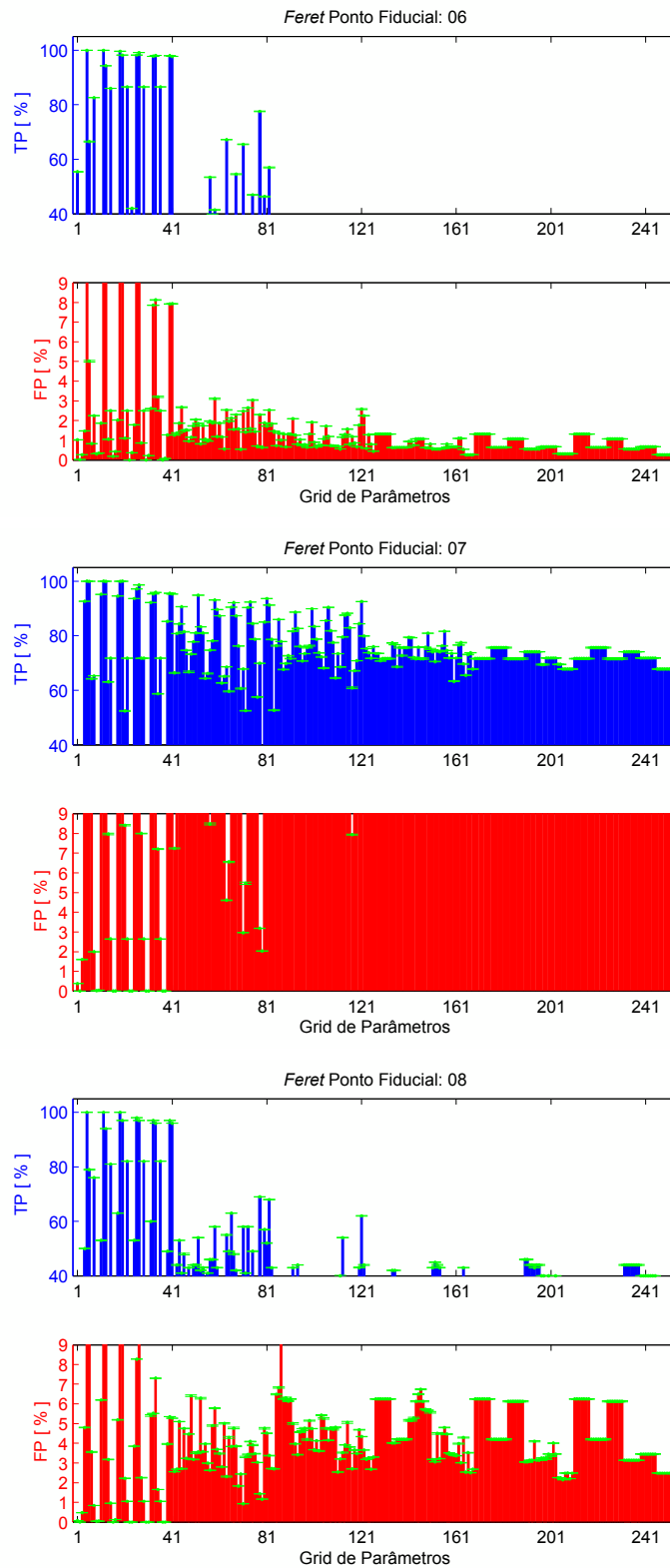
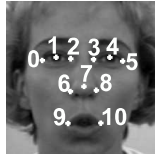
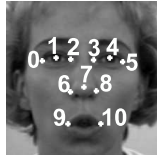
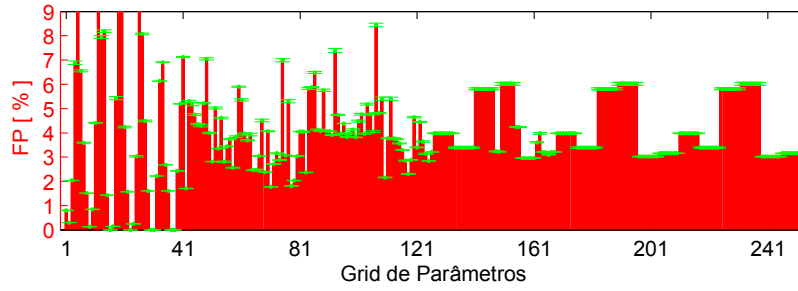
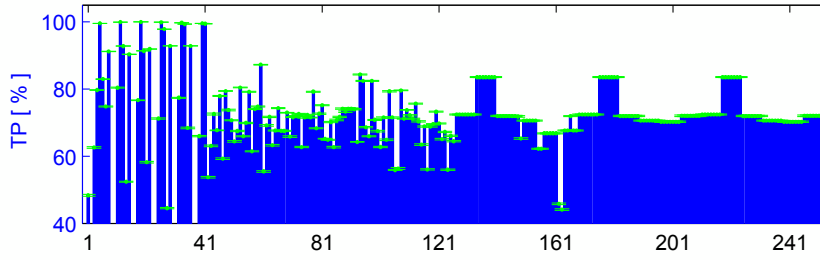


Figura 5.18: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *Ferret*. Os resultados são apresentados para a região do nariz são para os pontos fiduciais 6, 7 e 8. Neste *grid* de parâmetros, o valor de ξ_i igual à 0, 1.



Feret Ponto Fiducial: 09



Feret Ponto Fiducial: 10

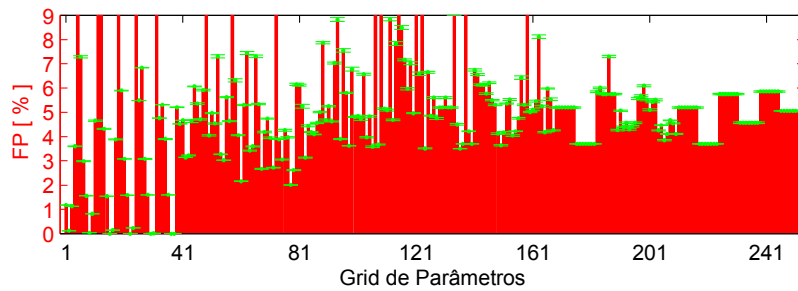
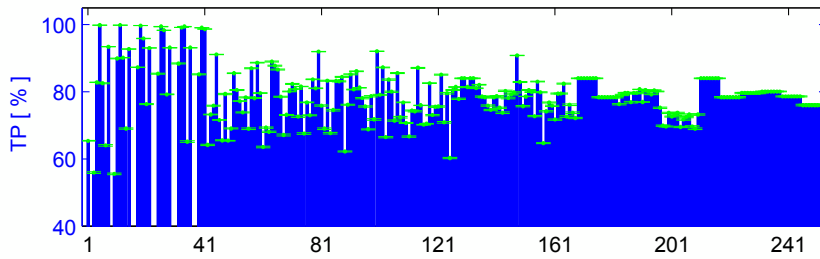


Figura 5.19: *Grid* de parâmetros do desempenho do sistema de detecção de pontos fiduciais para a base de dados *Feret*. Os resultados são apresentados para a região da boca são para os pontos fiduciais 9 e 10. Neste *grid* de parâmetros, o valor de ξ_i igual à 0, 1.

- Na Figura 5.16, os resultados para a região do olho direito são apresentados. Nos gráficos, identificamos os melhores resultados dentre as 42 primeiras combinações. Nas demais partes dos gráficos, as taxas de TP e FP são elevadas para os pontos fiduciais 0 e 1, tornando o desempenho dos classificadores não competitivos. No ponto fiducial 2, a taxa de TP apresenta resultados elevados, tornando estes classificadores eficazes para a detecção deste ponto fiducial. Dessa forma, podemos inferir que o sistema não apresenta desempenho satisfatório para as combinação de parâmetros apresentada.
- O desempenho do sistema de detecção de pontos fiduciais para o *grid* de parâmetros são apresentados na 5.17 para a região do olho esquerdo. Observe que na maior parte da extensão do *grid* de parâmetro o desempenho da taxa de FP é elevada. Os melhores desempenhos de TP e FP são encontrados dentre as 42 primeiras combinações, pois notamos algumas combinações com o desempenho da taxa de TP elevada e a taxa de FP abaixo de 2,0%. Assim, podemos inferir que o sistema não apresenta desempenho satisfatório para as combinações de parâmetros.
- Na Figura 5.18, os resultados para a região do nariz são apresentados e observamos que os melhores desempenhos são notados nas 42 primeiras combinações do *grid* de parâmetros. Note que para os pontos fiduciais 6 e 8 as taxas de TP apresentam desempenho abaixo de 40% a partir das 84 primeiras combinações. O ponto fiducial 07 apresenta taxas de FP elevadas, tornando os classificadores não competitivos.
- Na Figura 5.19, note que os melhores desempenhos estão localizados nas 42 primeiras combinações do *grid* de parâmetros. Após as 42 primeiras combinações do *grid*, note que as taxas de TP e FP apresentam valores elevados. Assim podemos inferir que o sistema para esta combinação de parâmetros não apresenta desempenho satisfatório.

Em resumo, nestes experimentos propomos observar o desempenho do sistema de detecção de pontos fiduciais para a base de dados *Feret* na extensão do *grid* de parâmetros. Apresentamos os gráficos com os pontos fiduciais divididos em 4 regiões (olho direito, olho esquerdo, boca e nariz). Além disso, com a variável de folga

fixa em 0, 1. Assim, identificamos no desenvolvimento dos experimentos que o desempenho do sistema para todas as combinações dos parâmetros apresenta variações nas taxas de TP e FP. Entretanto, observamos que os classificadores com desempenho satisfatório estão inseridos dentre as 42 primeiras combinações e no grupo restante de combinações podemos inferir que o desempenho não é satisfatório, principalmente por apresentar valores elevados da taxa FP. Dentre as combinações de parâmetros que apresentam desempenho satisfatório do sistema, identificamos os valores de C com valores iguais à 2^{-12} .

5.6.4 Comparação entre Métodos: Conjunto de Imagens da Base de Dados *Feret*

De maneira semelhante a realizada com os resultados da base de dados *BioID* identificamos os melhores desempenhos do *grid* de parâmetros e tabulamos para a base de dados *Feret*. Os resultados são apresentados na Tabela 5.5 com a aplicação dos melhores parâmetros do *grid* para 11 pontos fiduciais da base de dados *Feret*. De modo geral, a situação ideal é a que os resultados apresentem desempenhos elevados na taxa de TP e os menores desempenhos possíveis quando tratamos da taxa de FP.

Analizamos os resultados da Tabela 5.5. Assim, identificamos que:

- Para os pontos fiduciais para a região do olho direito, note que os resultados são competitivos. Note que as taxas de TP do método proposto são acima de 90,0% e as taxas de FP são abaixo de 2,0% com exceção ao ponto fiducial 00. Além disso, podemos observar que os resultados apresentados são superiores aos dos métodos DF e SVM L, inclusive quando levamos em consideração o desvio padrão para as taxas de TP e FP. Por fim, os resultados são equivalentes quando comparados aos métodos DF-PCA e IPD, os dois com taxas de FP e desvios padrões baixas.
- O desempenho do sistema proposto para a região do olho esquerdo apresenta melhores resultados para a taxa de TP valores acima de 95,0% para os pontos 03 e 04. Note que as taxas de FP apresentam valores abaixo de 2,0%, com exceção do ponto fiducial 03. Assim observamos que os valores apresentados pelo método proposto são superiores em comparação aos métodos DF e SVM

L, e apresentam baixa competitividade quando comparamos aos métodos DF-PCA e IPD.

- Na região do nariz, as taxas de TP do método proposto apresentam valores abaixo de 90,0%, exceto para o ponto fiducial 08. A taxa de FP é aceitável para o ponto fiducial 08, pois para os pontos 06 e 07 encontramos taxas de FP acima de 2,0%. Note que o método proposto apresenta desempenho equivalente quando comparado DF e SVM L, onde encontramos desvios padrões melhores. Além disso, observe que os resultados dos métodos DF-PCA e IPD apresentam resultados com maior competitividade.
- As taxas de TP do método proposto para a região da boca apresenta valores acima de 95,0% e com a taxa de FP abaixo de 2,0% tornando o método proposto competitivo. Note que o desempenho do método proposto supera os desempenho dos métodos SVM L e DF, com baixos valores de desvio padrão. Além disso, note que o desempenho do método proposto é competitivo quando comparado aos métodos DF-PCA e IPD, pois a taxa de TP do método C-SVC é superior e as taxas de FP dos pontos 09 e 10 estão abaixo de 2,0%.

Em resumo, podemos inferir os comentários abaixo para a comparação com os demais sistemas:

- DF: o método proposto supera o desempenho do sistema utilizando o método DF em todos os pontos fiduciais quando comparamos os valores das taxas de TP. Note que os resultados apresentados para a taxa de FP são melhores por estarem abaixo de 0,6%, enquanto que o do método proposto é considerado aceitável por estarem em média próximo ou abaixo de 2,5%. Nesta comparação, inferimos que o desempenho do sistema de detecção de pontos proposto é ótimo.
- SVM linear: o método SVM L possui o desempenho do sistema superado para as taxas de TP em todos os pontos fiduciais com exceção ao ponto fiducial 07. Com relação ao parâmetro FP, o método SVM L supera os resultados do método proposto nos pontos fiduciais 04, 05, 06 e 07. Podemos inferir que o desempenho do sistema proposto é superior quando comparado ao SVM linear.

- DF-PCA: no método proposto, os pontos 01, 03, 04, 09 e 10 apresentam o desempenho de TP superior ao método DF-PCA. Quando comparamos o desempenho de FP, identificamos que o método DF-PCA apresenta ótimos valores, abaixo de 0,4%, sendo superior ao método C-SVC. Assim, podemos inferir que o desempenho do método proposto é competitivo, porém o método DF-PCA apresenta desempenho ótimo.
- IPD: identificamos que o método proposto apresenta melhor desempenho na taxa de TP para os pontos fiduciais 01 e 05 em comparação ao método IPD. Note que este método apresenta ótimo desempenho quando tratamos da taxa de FP, pois os valores são abaixo de 0,30%. Valores de FP abaixo de 2,0% são considerados satisfatórios, esse fato também pode ser observado nos resultados do método proposto, exceto no ponto fiducial 00, 03, 06 e 07. Assim, podemos inferir que o desempenho do método IPD é superior ao método proposto.

5.6.5 Síntese de Resultados: Comparação *BioID* e *Feret*

Nesta seção apresentamos a análise de resultados geral para os experimentos propostos pelo sistema de detecção de pontos fiduciais em faces humanas utilizando máquina de vetores suporte. Vale ressaltar que os pontos fiduciais foram divididos em 4 regiões (olho direito, olho esquerdo, boca e nariz). Deste modo, analisamos os resultados para os subconjuntos das bases de dados *BioID* e *Feret*, ambas compostas por 503 e 2004 imagens e sob o desenvolvimento de experimentos com 252 combinações do *grid* de parâmetros.

Nos experimentos, o desempenho do sistema de detecção de pontos fiduciais foi avaliado utilizando a métrica de TP e FP. Independente da base de dados avaliada, observamos que os pontos fiduciais apresentem taxas de TP e FP diferenciadas. Além disso, notamos que os classificadores com desempenho satisfatório estão inseridos dentre as 42 primeiras combinações e no grupo restante de combinações podemos inferir que o desempenho não é satisfatório, principalmente por apresentar valores elevados valores da taxa FP.

A seguir, apresentamos a síntese de resultados segmentada por região da face:

- Os resultados para a região do olho direito para as bases de dados *BioID* e

Tabela 5.5: Desempenho do sistema proposto para 11 pontos usando base de dados *Feret*.

PF	DF-PCA [58]			DF [58]			IPD [58]			SVM L [58]			C-SVC (Proposto)					
	FP	σ	TP	FP	σ	TP	FP	σ	TP	FP	σ	TP	FP	σ	TP	FP	σ	TP
00	0,22	0,05	95,4	0,18	0,04	35,1	0,18	0,02	95,6	1,2	3,00	2,04	91,1	12,4	2,57	0,76	95,0	1,5
01	0,27	0,05	93,1	0,31	0,06	15,8	0,08	0,01	88,6	4,5	1,85	3,30	80,0	15,7	0,86	0,06	94,3	1,4
02	0,32	0,16	96,2	0,22	0,09	32,2	0,15	0,02	98,1	0,9	19,51	15,60	89,8	25,3	1,39	0,20	90,2	1,1
03	0,29	0,03	92,8	0,52	0,01	64,9	0,12	0,01	95,4	1,8	4,82	5,47	93,6	9,7	2,80	0,33	95,2	1,4
04	0,16	0,01	94,7	0,40	0,01	25,7	0,07	0,01	91,0	1,8	0,43	0,36	63,7	39,6	1,98	0,13	95,0	1,2
05	0,25	0,01	95,9	0,26	0,00	33,0	0,15	0,02	97,1	1,1	0,84	0,61	73,3	27,6	1,44	0,10	84,2	1,7
06	0,28	0,07	96,3	0,14	0,04	74,1	0,20	0,04	96,7	1,5	2,06	3,32	73,9	33,7	2,25	0,54	87,6	3,2
07	0,30	0,02	89,9	0,22	0,01	42,6	0,27	0,03	96,6	1,3	1,16	0,62	80,7	13,0	2,64	0,16	71,8	3,7
08	0,36	0,03	97,2	0,15	0,01	77,8	0,18	0,02	96,2	2,1	2,13	2,42	86,2	16,3	1,05	0,04	90,1	1,9
09	0,16	0,01	94,5	0,12	0,00	18,4	0,19	0,01	99,2	0,8	1,57	1,43	84,7	17,4	1,61	0,09	97,4	2,3
10	0,14	0,04	93,1	0,07	0,02	9,8	0,22	0,02	98,5	0,8	0,76	1,33	54,4	38,0	0,69	0,19	95,5	2,5

Feret apresentam desempenhos da taxa de TP satisfatórios para o decorrer do *grid* de parâmetros. Note que o desempenho da taxa de FP para a base de dados *Feret* apresenta valores elevados no decorrer do *grid*. Deste modo, os classificadores com desempenhos competitivos estão localizados na região do *grid* em que C apresenta valores iguais à 2^{-12} e 2^{-9} para a bases de dados *BioID* e 2^{-12} para a base de dados *Feret*.

- O desempenho do sistema de detecção de pontos fiduciais para o *grid* de parâmetros da região do olho esquerdo apresentam resultados próximos a 100% quando analisamos a taxa TP para ambas as bases. O desempenho da taxa de FP para o ponto fiducial 03 é elevado no decorrer do *grid* para as duas bases de dados. Em geral, a taxa de FP é elevada para os experimentos realizados com a base de dados *Feret*. Os melhores desempenhos no *grid* de parâmetros são aqueles cujo o parâmetro C apresenta valores iguais à 2^{-12} e 2^{-9} para a bases de dados *BioID* e 2^{-12} para a base de dados *Feret*.
- Na análise de resultados, identificamos que o desempenho do sistema de detecção de pontos fiduciais para as duas bases de dados na região nariz são semelhantes, pois encontramos poucas regiões do *grid* que apresentam desempenhos do limitante inferior dos gráficos para as taxas de TP e nas taxas de FP observamos o desempenho abaixo de 2,5% no decorrer dos gráficos. Assim, o classificadores com bons desempenhos estão localizados na região cujo o parâmetro C apresenta valores iguais à 2^{-12} .
- Nos resultados apresentados para a região da boca, observamos que a taxa de TP apresenta desempenho satisfatório ao desejado para as duas bases de dados. Quando analisamos a taxa de FP, identificamos que o desempenho da base de dados *BioID* é superior (abaixo de 2,5%) ao apresentado pela base de dados *Feret*. Os melhores desempenhos no *grid* de parâmetros são aqueles cujo o parâmetro C apresenta valores iguais à 2^{-12} e 2^{-9} para a bases de dados *BioID* e 2^{-12} para a base de dados *Feret*.

Em resumo, a diferença de resultados é justificada devido ao subconjunto da base de dados *Feret* apresentar maior quantidade de imagens e variações (rotação e não uniformidade de iluminação) quando comparado a base de dados *BioID*, o que

torna o procedimento da separação de classes mais custoso e com restrições para apresentar bons desempenhos de TP e FP para o *grid* de parâmetros completo.

Capítulo 6

Conclusão

Nesta dissertação, investigamos os parâmetros que influenciam na geração do hiperplano ótimo responsável pela separação dos dados das classes de positivos e negativos da técnica SVM. Os parâmetros investigados nesta dissertação foram o parâmetro de ponderação (C), o número de iterações ($\#it$) e a tolerância (ϵ) utilizados no algoritmo do SVM e o peso (W) atribuído a cada classe para ser levado em consideração durante o treino do algoritmo do SVM. Os parâmetros investigados foram submetidos ao *grid* de parâmetros para a identificação da combinação que apresente o melhor desempenho, sendo uma das principais contribuições desta dissertação.

O desempenho do sistema de detecção de pontos fiduciais para o *grid* de parâmetros foi avaliado utilizando um sistema de detecção de pontos fiduciais, dividido em procedimentos de treino e teste. No procedimento de treino as imagens da base de dados de treino são submetidas a etapa de pré-processamento, a definição da região de interesse, a extração de características e por fim, ao treino do algoritmo SVM. Outra contribuição importante dessa dissertação é observada na definição dos blocos das classes de treino, onde aplicamos uma série de procedimentos para a seleção dos pontos utilizados como centro dos blocos da classe de positivos e da classes de negativos. No procedimento de teste, as imagens tem os blocos da região de interesse extraídos e em seguida os blocos são classificados utilizando o classificador obtido na etapa de treino.

Os experimentos para os 11 pontos fiduciais foram realizados utilizando a base de dados *BioID*, com 503 imagens e a base de dados *Feret* com 2004 imagens

através do procedimento de validação cruzada. Na validação cruzada foram criados dois subconjuntos, o de treino e o de teste. O subconjunto de treino é composto por 6/7 das imagens da *bioID* e o restante das imagens pertencem ao subconjunto de teste.

O desempenho do sistema para o *grid* de parâmetros é agrupado em 4 regiões. Região do olho direito, do olho esquerdo, do nariz e da boca. Em cada gráfico, são apresentadas as taxas de TP e FP e os seus respectivos desvios padrões. Na análise destes resultados, identificamos no *grid* as combinações que apresentam os melhores desempenhos. Identificamos que o parâmetro ponderação (C) com valores iguais à 2^{-12} e 2^{-9} combinados as variações dos parâmetros de iteração e ponderação apresentam os melhores resultados para a base de dados *BioID*. E para a base de dados *Feret*, o parâmetro ponderação (C) com valor igual à 2^{-12} apresenta os melhores resultados, neste caso, combinados as variações dos parâmetros de iteração e ponderação.

Os resultados obtidos nos 11088 experimentos foram comparados aos resultados de métodos por filtragem discriminativa e detecção por produto interno apresentados em [20] e de um sistema de detecção de pontos fiduciais que utiliza classificador SVM com *kernel* linear. Avaliando o desempenho em termo das taxas de TP e FP podemos sintetizar que o sistema proposto apresenta bons resultados. Com a identificação dos melhores resultados, os parâmetros ótimos para cada ponto fiducial foi obtido. Por fim, comentamos a importância da pesquisa de métodos que proporcionam a detecção de pontos fiduciais em faces humanas devido ao grande número de aplicações, como por exemplo, citamos aplicações em sistemas de segurança, automação residencial e industrial.

6.1 Propostas para Trabalhos Futuros

Nesta seção, apresentamos algumas propostas para trabalhos futuros a partir desta dissertação:

- Investigar a formulação do SVM ν -SVC e avaliar o desempenho comparando com o desempenho obtido nesta pesquisa.
- Investigar formas de pré-processamento para estimar o ponto fiducial.

- Projetar um filtro levando em consideração a relação de aspecto, cada filtro vai detectar um ponto fiducial levando em consideração várias relações de aspecto.
- Inserir o procedimento de pós-processamento para identificar o desempenho do sistema de detecção de pontos fiduciais utilizando a formulação do C-SVC do SVM.
- Investigar os parâmetros de *kernels* não lineares do SVM, por exemplo, o gaussiano.
- Investigar a variação da quantidade de blocos das classes no desempenho do sistema de detecção de pontos fiduciais utilizando a formulação do C-SVC do SVM.
- Investigar e propor métodos para definição da região de interesse, investigando se o método Gaussiano à Priori é o método ideal para todos os pontos fiduciais trabalhados. Além disso, identificar a possibilidade de um método de definição da região de interesse para cada ponto fiducial, identificando em quais pontos fiduciais o uso do método multimodal é necessário.

Referências Bibliográficas

- [1] THEODORIDIS, S., KOUTROUMBAS, K., *Pattern Recognition*. 4th ed. Elsevier, 2009.
- [2] DUDA, R., HART, P., STORK, D., *Pattern Classification*. 2nd ed. Wiley-Interscience: Secaucus, NJ, USA, 2001.
- [3] BISHOP, C. M., *Pattern Recognition and Machine Learning*. 1st ed. Springer: New York, NY, USA, 2006.
- [4] YUJIAN, W., JIAZHENG, Y., LILI, F., et al., “Application Research of Support Vector Machine in Multi-Spectra Remote Sensing Image Classification”. In: *Proceedings of the Conference on Image and Signal Processing*, pp. 1–5, Oct 2009.
- [5] PRADHAN, A., “Support Vector Machine - A Survey”, *International Journal of Emerging technology and Advanced Engineering*, v. 2, pp. 82–85, August 2012.
- [6] YUN, T., GUAN, L., “Automatic Fiducial Points Detection for Facial Expressions using Scale Invariant Feature”. In: *Proceedings of the International Workshop on Multimedia Signal Processing*, pp. 1–6, Oct 2009.
- [7] WEIWEI, Y., NANNAN, Y., “Facial Feature Extraction on Fiducial Points and used in Face Recognition”, *IEEE Transactions of Workshop on Intelligent Information Technology Applications*, v. 3, pp. 274–277, 2009.
- [8] YUN, T., GUAN, L., “Fiducial Point Tracking for Facial Expression using Multiple Particle Filters with Kernel Correlation Analysis”. In: *Proceedings of Conference on Image Processing*, pp. 373–376, Sept 2010.

- [9] MARTINEZ, B., VALSTAR, M., BINEFA, X., et al., “Local Evidence Aggregation for Regression-Based Facial Point Detection”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 35, n. 5, pp. 1149–1163, May 2013.
- [10] AMBERG, B., VETTER, T., “Optimal Landmark Detection using Shape Models and Branch and Bound”. In: *Proceedings of the Conference on Computer Vision*, pp. 455–462, Nov 2011.
- [11] ARAUJO, G. A., “Detecção de Landmarks Faciais usando SVM”. In: *Simpósio Brasileiro de Telecomunicações*, pp. 1–5, 2011.
- [12] ZHANG, X., MAHOOR, M., VOYLES, R., “Facial Expression Recognition using HessianMKL Based Multiclass-SVM”. In: *Proceedings of the Conference and Workshop on Automatic Face and Gesture Recognition*, pp. 1–6, April 2013.
- [13] OSUNA, E. E., FREUND, R., GIROSI, F., *Support Vector Machines: Training and Applications*, Tech. rep., 1997.
- [14] CORTES, C., VAPNIK, V., “Support-Vector Networks”, *Machine Learning*, v. 20, n. 3, pp. 273–297, September 1995.
- [15] CHANG, C.-C., LIN, C.-J., “LIBSVM: A Library for Support Vector Machines”, *ACM Transactions on Intelligent System Technology*, v. 2, n. 3, pp. 27:1–27:27, May 2011.
- [16] WANG, S., “Support Vector Machines Classification for High-Dimensional Dataset”. In: *Proceedings of the Conference on Multimedia Information Networking and Security*, pp. 315–318, Nov 2012.
- [17] WU, K. P., WANG, S. D., “Choosing the Kernel Parameters for Support Vector Machines by the Inter-cluster Distance in the Feature Space”, *Elsevier Transactions on Pattern Recognition*, v. 42, n. 5, pp. 710–717, May 2009.
- [18] Disponível em: <http://www.bioid.com/support/downloads/software/bioid-face-database.html>. Acesso em: Dezembro de 2013.

- [19] Disponível em: <http://www.itl.nist.gov/iad/humanid/feret/>. Acesso em: Dezembro de 2013.
- [20] WALDIR, S., ARAUJO, G., DA SILVA, E., et al., “Facial Fiducial Points Detection using Discriminative Filtering on Principal Components”. In: *Proceedings of the Conference on Image Processing*, pp. 2681–2684, Sept 2010.
- [21] VOS, F., CAAN, M. W. A., VERMEER, K., et al., “Linear and Kernel Fisher Discriminant Analysis for Studying Diffusion Tensor Images in Schizophrenia”. In: *Proceedings of the International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 764–767, April 2007.
- [22] HUANG, H., LIU, J., PAN, Y., “Local Fisher Discriminant Analysis with Maximum Margin Criterion for Image Recognition”. In: *Proceedings of the International Conference on Computer Graphics, Imaging and Visualization*, pp. 92–97, Aug 2011.
- [23] DUFOUR, R., MILLER, E., GALATSANOS, N., “Template Matching based Object Recognition with Unknown Geometric Parameters”, *IEEE Transactions on Image Processing*, v. 11, n. 12, pp. 1385–1396, Dec 2002.
- [24] JOGLEKAR, J., GEDAM, S. S., “Area Based Image Matching Methods - A Survey”, v. 2, pp. 130–136, Jan.
- [25] LI-SHA, X., HUA-JING, F., LUO, Z., “Survey on Hidden Markov Model based Fault Diagnosis and Prognosis”. In: *Proceedings of the Conference on Control and Decision*, pp. 4880–4884, May 2013.
- [26] ARAUJO, G., WALDIR, S., DA SILVA, E., et al., “Facial Landmarks Detection Based on Correlation Filters”. In: *Proceedings of the International Symposium on Telecommunications*, pp. 1–5, October 2010.
- [27] VIOLA, P., JONES, M., “Robust Real-Time Object Detection”, *International Journal of Computer Vision*, v. 57, n. 2, pp. 137–154, July 2001.
- [28] VIOLA, P., JONES, M., “Rapid Object Detection using a Boosted Cascade of Simple Features”. In: *Proceedings of the Conference on Computer Vision*

and Pattern Recognition, pp. 511–518, Kauai Marriott, Hawaii, December 2001.

- [29] EPHRAIM, T., HIMMELMAN, T., SIDDIQI, K., “Real-Time Viola-Jones Face Detection in a Web Browser”. In: *Proceedings of the Conference on Computer and Robot Vision*, pp. 321–328, May 2009.
- [30] REN, J., KEHTARNAVAZ, N., ESTEVEZ, L., “Real-Time Optimization of Viola -Jones Face Detection for Mobile Platforms”. In: *Proceedings of the Conference on Circuits and Systems Workshop: System-on-Chip - Design, Applications, Integration, and Software*, pp. 1–4, Oct 2008.
- [31] TAN, X., TRIGGS, B., “Enhanced Local Texture Feature Sets for Face Recognition under Difficult Lighting Conditions”. In: *Proceedings of the International Workshop on Analysis and Modeling of Faces and Gestures*, pp. 168–182, Rio de Janeiro, Brazil, October 2007.
- [32] VALSTAR, M., MARTINEZ, B., BINEFA, X., et al., “Facial Point Detection using Boosted Regression and Graph Models”. In: *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 2729–2736, June 2010.
- [33] ZHU, X., RAMANAN, D., “Face Detection, Pose Estimation, and Landmark Localization in the Wild.” In: *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 2879–2886, June 2012.
- [34] LI, Y., QI, C., “Face Recognition using Hog Feature and Group Sparse Coding”. In: *Proceedings of the Conference on Image Processing*, pp. 3350–3353, Sept 2013.
- [35] RAPP, V., SENECHAL, T., BAILLY, K., et al., “Multiple Kernel Learning SVM and Statistical Validation for Facial Landmark Detection”. In: *Proceedings of the Conference on Automatic Face Gesture Recognition and Workshops*, pp. 265–271, March 2011.

- [36] ARIFFIN, S., MIDI, H., “Robust Bootstrap Methods in Logistic Regression Model”. In: *Proceedings of the Conference on Statistics in Science, Business, and Engineering*, pp. 1–6, Sept 2012.
- [37] ZHANG, H., WU, Q., NGUYEN, T. M., “Image Segmentation by a Robust Modified Gaussian Mixture Model”. In: *Proceedings of the Conference on Acoustics, Speech and Signal Processing*, pp. 1478–1482, 2013.
- [38] RAVIKUMAR, B., THUKARAM, D., KHINCHA, H. P., “Comparison of Multiclass SVM Classification Methods to use in a Supportive System for Distance Relay Coordination”, *IEEE Transactions on Power Delivery*, v. 25, n. 3, pp. 1296–1305, July 2010.
- [39] JOACHIMS, T., BURGESS, C., SCHOLKOPF, B., et al., (eds), “Advances in Kernel Methods: Support Vector Learning”, chap. Making Large-scale Support Vector Machine Learning Practical, pp. 169–184, MIT press: Cambridge, MA, 1999.
- [40] JAIN, A. K., DUIN, R. P. W., MAO, J., “Statistical Pattern Recognition: A Review”, v. 22, pp. 4–37, 2000.
- [41] YUJIAN, W., JIAZHENG, Y., LILI, F., et al., “Application Research of Support Vector Machine in Multi-Spectra Remote Sensing Image Classification”. In: *Proceedings of the Conference on Image and Signal Processing*, pp. 1–5, Oct 2009.
- [42] NOVAKOVIC, J., VELJOVIC, A., “C-Support Vector Classification: Selection of Kernel and Parameters in Medical Diagnosis”. In: *Proceedings of the International Symposium on Intelligent Systems and Informatics*, pp. 465–470, Sept 2011.
- [43] XU, Y., “Predicting Customer Churn with Extended One-Class Support Vector Machine”. In: *Proceedings of the Conference on Natural Computation*, pp. 97–100, May 2012.
- [44] WANG, J., ZHOU, Y., KOU, Z., “One Class Support Vector Machine for Predicting Avian-to-Human Transmission of Avian Influenza A Virus”. In:

Proceedings of the Conference on Computer Science and Automation Engineering, v. 3, pp. 184–188, May 2012.

- [45] LI-YANG, W., ZHI, L., ZHI-GUANG, Z., et al., “A Mixed-Kernel-based SVR Controller for Biped Robots”. In: *Proceedings of the Conference on Control Conference*, pp. 3925–3930, July 2011.
- [46] XIAOFEN, J., BAITING, Z., ZHAOQUAN, C., “Genetic Algorithm Optimization based on Support Vector Machine Image Interpolation”. In: *Proceedings of the Conference on Cross Strait Quad-Regional Radio Science and Wireless Technology*, v. 2, pp. 1319–1322, July 2011.
- [47] BUCIU, I., “Efficiency Analysis of Illumination Correction Methods for Face Recognition Performance”, *IEEE Transactions on Computational Photography*, v. 0, pp. 211–216, 2010.
- [48] WU, Y., LIU, Z., HAN, Y., et al., “An Image Illumination Correction Algorithm Based on Tone Mapping”. In: *Proceedings of the Conference on Image and Signal Processing*, v. 2, pp. 645–648, 2010.
- [49] ESAKKIRAJAN, S., VEERAKUMAR, T., SUBRAMANYAM, A., et al., “Removal of High Density Salt and Pepper Noise Through Modified Decision Based Unsymmetric Trimmed Median Filter”, *IEEE Transactions on Signal Processing Letters*, v. 18, n. 5, pp. 287–290, May 2011.
- [50] DEIVALAKSHMI, S., SARATH, S., PALANISAMY, P., “Detection and Removal of Salt and Pepper Noise in Images by Improved Median Filter”. In: *Proceedings of the Conference on Recent Advances in Intelligent Computational Systems*, pp. 363–368, Sept 2011.
- [51] RAHMAN, M., ABDULLAH-AL-WADUD, M., PREZA, C., “A Decision-based Filter for Removing Salt-and-Pepper Noise”. In: *Proceedings of the Conference on Informatics, Electronics Vision*, pp. 1064–1068, May 2012.
- [52] PHILLIPS, P., MOON, H., RAUSS, P., et al., “The FERET Evaluation Methodology for Face-Recognition Algorithms”. In: *Proceedings of Conference on Computer Vision and Pattern Recognition*, pp. 137–143, Jun 1997.

- [53] PHILLIPS, P., MOON, H., RIZVI, S., et al., “The FERET Evaluation Methodology for Face-Recognition Algorithms”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 22, n. 10, pp. 1090–1104, Oct 2000.
- [54] Disponível em: <http://svmlight.joachims.org/>. Acesso em: Maio de 2013.
- [55] Disponível em: <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>. Acesso em: Novembro de 2013.
- [56] Disponível em: http://svmlight.joachims.org/svm_perf.html. Acesso em: Julho de 2013.
- [57] Disponível em: <http://en.wikipedia.org/wiki/OpenCV>. Acesso em: Julho de 2013.
- [58] DA SILVA JÚNIOR, W. S., *Reconhecimento de Padrões Utilizando Filtros de Correlação com Análise de Componentes Principais*, Tese de Doutorado, COPPE - Universidade Federal do Rio de Janeiro, 2010.
- [59] ZHIQIAN WANG; RAO, K. R.; BEN-ARIE, J., “Optimal Ramp Edge Detection using Expansion Matching”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 18, n. 11, pp. 1092–1097, November 1996.
- [60] RAO, K. R.; BEN-ARIE, J., “Optimal Edge Detection using Expansion Matching and Restoration”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 16, n. 12, pp. 1169–1182, December 1994.
- [61] TAN, X., TRIGGS, B., “Fusing Gabor and LBP Feature Sets for Kernel-Based Face Recognition”. In: *Proceedings of the International Workshop on Analysis and Modeling of Faces and Gestures*, pp. 235–249, Rio de Janeiro, BR, October 2007.
- [62] DUAN, H., LIU, X., “Lower C Limits in Support Vector Machines with Radial Basis Function Kernels”. In: *Proceedings of the International Symposium on Information Technology in Medicine and Education*, v. 2, pp. 768–771, Aug 2012.
- [63] Disponível em: <http://www.mmfacedb.com/>. Acesso em: Janeiro de 2013.

Apêndice A

Artigos Publicados

Neste apêndice, apresentamos os artigos desenvolvidos nesta Dissertação.

A.1 Artigos Diretamente Relacionados ao Tema

1. SILVA, L. E. S.; SANTOS, K. V.; SILVA JUNIOR, W. S. "Sistema de Detecção de Pontos Fiduciais utilizando Classificadores Lineares C-SVC". In: *Anais do XXXI Simpósio Brasileiro de Telecomunicações*, Setembro de 2013, Fortaleza. SBrT'13, 2013.
1. SILVA, L. E. S.; SANTOS, K. V.; TOMAZ JUNIOR, P. D.; SILVA JUNIOR, W. S. "Fiducial Points Detection Using SVM Linear Classifiers". In: *Second International Conference on Computational Science and Engineering*, Dubai, 2014.

A.2 Outros Artigos

1. SANTOS, K. V.; SILVA, L. E. S.; SILVA JUNIOR, W. S. "Avaliação do Desempenho dos Filtros Discriminativos em um Sistema de Detecção de Pontos Fiduciais". In: *Anais do XXXI Simpósio Brasileiro de Telecomunicações*, Setembro de 2013, Fortaleza. SBrT'13, 2013.
1. SILVA, L. E. S.; SANTOS, K. V.; SILVA JUNIOR, W. S. "Fiducial Points Influence of Quantity of Principal Component In Discriminative Filtering". In:

Second International Conference on Computational Science and Engineering,
Dubai, 2014.

Apêndice B

Combinação de Parâmetros

Neste apêndice, apresentamos as combinações dos parâmetros utilizados nos experimentos desta dissertação. Levar em consideração que o valor do parâmetro (ξ_i) é fixo em 0,001; 0,01; 0,1 de acordo com o experimento realizado.

Tabela B.1: Combinações de parâmetros utilizados nos experimentos.

Combinação	Parâmetros			
	Custo	Iterações	Peso (Wp)	Peso (Wn)
1	0,000244	10000	1	5
2	0,000244	42000	1	5
3	0,000244	50000	1	5
4	0,000244	34000	1	5
5	0,000244	26000	1	5
6	0,000244	50000	2	10
7	0,000244	18000	1	5
8	0,000244	42000	2	10
9	0,000244	34000	2	10
10	0,000244	26000	2	10
11	0,000244	10000	1	1
12	0,000244	18000	2	10
13	0,000244	26000	1	1
14	0,000244	42000	1	1
15	0,000244	50000	1	1
16	0,000244	34000	1	1
17	0,000244	18000	1	1
18	0,000244	10000	2	10
19	0,001953	50000	1	1
20	0,001953	18000	10	500
21	0,001953	26000	500	10
22	0,001953	26000	5	1
23	0,000244	34000	500	10
24	0,000244	42000	500	10
25	0,015625	18000	10	2
26	0,001953	50000	1	5
27	0,001953	42000	1	1
28	0,001953	42000	500	10
29	0,015625	26000	1	1
30	0,001953	18000	5	1
31	0,015625	18000	1	5
32	0,015625	50000	2	10
33	0,001953	50000	10	2

34	0,000244	42000	10	500
35	0,015625	42000	1	1
36	0,000244	50000	500	10
37	0,001953	18000	1	5
38	0,015625	42000	10	2
39	0,001953	10000	1	1
40	0,001953	42000	2	10
41	0,001953	26000	1	1
42	0,001953	26000	10	2
43	0,000244	18000	500	10
44	0,000244	18000	10	500
45	0,015625	34000	1	1
46	0,001953	18000	10	2
47	0,015625	42000	1	5
48	0,001953	26000	2	10
49	0,001953	34000	1	5
50	0,015625	34000	1	5
51	0,015625	42000	10	500
52	0,015625	42000	500	10
53	0,125	42000	2	10
54	0,125	42000	10	500
55	0,125	42000	5	1
56	0,125	42000	10	2
57	0,125	42000	500	10
58	1	42000	1	1
59	1	42000	1	5
60	1	42000	2	10
61	1	42000	10	500
62	1	42000	5	1
63	1	42000	10	2
64	1	42000	500	10
65	8	42000	1	1
66	8	42000	1	5
67	8	42000	2	10
68	8	42000	10	500
69	8	42000	5	1
70	8	42000	10	2
71	8	42000	500	10
72	0,001953	18000	1	1
73	0,015625	50000	1	1
74	0,000244	50000	10	500
75	0,000244	10000	10	500
76	0,001953	34000	10	500
77	0,001953	10000	10	500
78	0,001953	42000	1	5
79	0,015625	26000	10	500
80	0,015625	26000	500	10
81	0,125	26000	1	1
82	0,125	26000	1	5
83	0,125	26000	2	10
84	0,125	26000	10	500
85	0,125	26000	5	1
86	0,125	26000	10	2
87	0,125	26000	500	10
88	1	26000	1	1
89	1	26000	1	5
90	1	26000	2	10
91	1	26000	10	500
92	1	26000	5	1
93	1	26000	10	2
94	1	26000	500	10
95	8	26000	1	1
96	8	26000	1	5

97	8	26000	2	10
98	8	26000	10	500
99	8	26000	5	1
100	8	26000	10	2
101	8	26000	500	10
102	0,015625	26000	1	5
103	0,001953	18000	500	10
104	0,015625	50000	10	2
105	0,015625	34000	10	2
106	0,001953	50000	500	10
107	0,015625	18000	10	500
108	0,015625	18000	500	10
109	0,125	18000	1	1
110	0,125	18000	1	5
111	0,125	18000	2	10
112	0,125	18000	10	500
113	0,125	18000	5	1
114	0,125	18000	10	2
115	0,125	18000	500	10
116	1	18000	1	1
117	1	18000	1	5
118	1	18000	2	10
119	1	18000	10	500
120	1	18000	5	1
121	1	18000	10	2
122	1	18000	500	10
123	8	18000	1	1
124	8	18000	1	5
125	8	18000	2	10
126	8	18000	10	500
127	8	18000	5	1
128	8	18000	10	2
129	8	18000	500	10
130	0,000244	26000	500	10
131	0,001953	34000	500	10
132	0,015625	42000	5	1
133	0,001953	26000	10	500
134	0,000244	10000	10	2
135	0,000244	50000	5	1
136	0,001953	34000	1	1
137	0,001953	42000	10	500
138	0,015625	42000	2	10
139	0,015625	18000	5	1
140	0,015625	10000	1	1
141	0,001953	42000	10	2
142	0,015625	10000	1	5
143	0,001953	10000	5	1
144	0,000244	10000	500	10
145	0,000244	42000	5	1
146	0,125	42000	1	1
147	0,125	42000	1	5
148	0,001953	34000	10	2
149	0,001953	34000	2	10
150	0,001953	10000	10	2
151	0,015625	18000	2	10
152	0,001953	42000	5	1
153	0,000244	50000	10	2
154	0,015625	18000	1	1
155	0,015625	26000	2	10
156	0,015625	34000	2	10
157	0,015625	50000	500	10
158	0,125	50000	2	10
159	0,125	50000	10	500

160	0,125	50000	5	1
161	0,125	50000	10	2
162	0,125	50000	500	10
163	1	50000	1	1
164	1	50000	1	5
165	1	50000	2	10
166	1	50000	10	500
167	1	50000	5	1
168	1	50000	10	2
169	1	50000	500	10
170	8	50000	1	1
171	8	50000	1	5
172	8	50000	2	10
173	8	50000	10	500
174	8	50000	5	1
175	8	50000	10	2
176	8	50000	500	10
177	0,015625	50000	10	500
178	0,000244	26000	10	500
179	0,015625	50000	1	5
180	0,001953	18000	2	10
181	0,015625	26000	5	1
182	0,015625	34000	5	1
183	0,000244	42000	10	2
184	0,001953	34000	5	1
185	0,000244	34000	5	1
186	0,001953	10000	2	10
187	0,001953	26000	1	5
188	0,015625	34000	10	500
189	0,015625	34000	500	10
190	0,125	34000	1	1
191	0,125	34000	1	5
192	0,125	34000	2	10
193	0,125	34000	10	500
194	0,125	34000	5	1
195	0,125	34000	10	2
196	0,125	34000	500	10
197	1	34000	1	1
198	1	34000	1	5
199	1	34000	2	10
200	1	34000	10	500
201	1	34000	5	1
202	1	34000	10	2
203	1	34000	500	10
204	8	34000	1	1
205	8	34000	1	5
206	8	34000	2	10
207	8	34000	10	500
208	8	34000	5	1
209	8	34000	10	2
210	8	34000	500	10
211	0,015625	26000	10	2
212	0,015625	10000	10	500
213	0,015625	10000	10	2
214	0,015625	10000	500	10
215	0,125	10000	1	1
216	0,125	10000	1	5
217	0,125	10000	2	10
218	0,125	10000	10	500
219	0,125	10000	5	1
220	0,125	10000	10	2
221	0,125	10000	500	10
222	1	10000	1	1

223	1	10000	1	5
224	1	10000	2	10
225	1	10000	10	500
226	1	10000	5	1
227	1	10000	10	2
228	1	10000	500	10
229	8	10000	1	1
230	8	10000	1	5
231	8	10000	2	10
232	8	10000	10	500
233	8	10000	5	1
234	8	10000	10	2
235	8	10000	500	10
236	0,015625	50000	5	1
237	0,001953	10000	500	10
238	0,001953	50000	10	500
239	0,015625	10000	2	10
240	0,125	50000	1	1
241	0,125	50000	1	5
242	0,000244	34000	10	500
243	0,000244	18000	10	2
244	0,001953	10000	1	5
245	0,000244	34000	10	2
246	0,001953	50000	2	10
247	0,015625	10000	5	1
248	0,001953	50000	5	1
249	0,000244	26000	10	2
250	0,000244	26000	5	1
251	0,000244	10000	5	1
252	0,000244	18000	5	1