UNIVERSIDADE DE SÃO PAULO – USP
ESCOLA DE ENGENHARIA DE SÃO CARLOS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Alex Antonio Affonso

**Reconhecimento facial em ambientes não controlados por
meio do "High-Boost Weber Descriptor" na região periocular**

São Carlos – SP

2018

Alex Antonio Affonso

Reconhecimento facial em ambientes não controlados por meio do "High-Boost Weber Descriptor" na região periocular

Tese de doutorado apresentada à Escola de Engenharia de São Carlos como parte dos requisitos para a obtenção do título de Doutor em Ciências, Programa de Engenharia Elétrica.

Área de Concentração: Processamento de Sinais e Instrumentação

Orientador: Prof. Dr. Evandro Luís Linhari Rodrigues

São Carlos

2018

Trata-se da versão corrigida da tese. A versão original se encontra disponível na EESC/USP que aloja o Programa de Pós-Graduação de Engenharia Elétrica.

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da EESC/USP com os dados inseridos pelo(a) autor(a).

A257r Affonso, Alex Antonio
Reconhecimento facial em ambientes não controlados por meio do "High-boost Weber Descriptor" na região periocular / Alex Antonio Affonso; orientador Evandro Luís Linhari Rodrigues. São Carlos, 2018.

Tese (Doutorado) - Programa de Pós-Graduação em Engenharia Elétrica e Área de Concentração em Processamento de Sinais e Instrumentação -- Escola de Engenharia de São Carlos da Universidade de São Paulo, 2018.

1. HBWLF. 2. MHBWLF. 3. HBWD. 4. Reconhecimento facial. 5. Localização dos olhos. 6. Visão Computacional. I. Título.

FOLHA DE JULGAMENTO

Candidato: Engenheiro **ALEX ANTONIO AFFONSO**.

Título da tese: "Reconhecimento facial em ambientes não controlados por meio do "High-boost Weber Descriptor" na região periocular".

Data da defesa: 27/04/2018.

Comissão Julgadora:

Resultado:

Prof. Associado **Evandro Luis Linhari Rodrigues**
(Orientador)
(Escola de Engenharia de São Carlos/EESC)

APROVADO

Profa. Titular **Agma Juci Machado Traina**
(Instituto de Ciências Matemáticas e de Computação/ICMC-USP)

APROVADO

Prof. Dr. **Aparecido Nilceu Marana**
(Universidade Estadual Paulista "Júlio de Mesquita Filho"/UNESP - Bauru)

Aprovado

Prof. Dr. **Maurilio Boaventura**
(Universidade Estadual Paulista "Júlio de Mesquita Filho"/UNESP – São José do Rio Preto)

APROVADO

Prof. Associado **Adilson Gonzaga**
(Escola de Engenharia de São Carlos/EESC)

APROVADO

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:
Prof. Associado **Luís Fernando Costa Alberto**

Presidente da Comissão de Pós-Graduação:
Prof. Associado **Luís Fernando Costa Alberto**

À minha família, com muito amor e carinho.

AGRADECIMENTOS

Expresso aqui nesta seção os meus mais sinceros agradecimentos a todos que me ajudaram ao longo deste caminho:

À Deus que, em sua infinita bondade, sempre foi meu protetor em todos os momentos de minha vida.

À Nossa Senhora Aparecida, pela saúde, proteção e conforto em toda a minha vida.

Ao Santo Expedito, pela proteção, força, coragem e serenidade ao longo de todo o caminho.

Aos meus pais, por serem minha referência e meu porto seguro. Meu muito obrigado por sempre acreditarem em mim e me apoiarem.

À minha filha, pelo seu amor, seu carinho, sua leveza e delicadeza, seus ensinamentos, sim, aprendo todo dia contigo. Descobri que ser pai não é só ensinar, mas também aprender, todos os dias. Meu muito obrigado por me fazer feliz, te amo!

À minha linda, querida e amada esposa, maravilhosa MÃE, pelo seu amor, sua paciência, seu carinho e inteligência em me ajudar em tudo.

Ao meu irmão, minha cunhada e meus sobrinhos queridos, que sempre me alegam com beijos e abraços que o tio adora.

Aos meus sogros, ao Allan e à Ana, por fazerem parte de minha família com muito amor e carinho.

Ao meu orientador, Prof. Dr. Evandro Luís Linhari Rodrigues, pelas sábias contribuições durante o desenvolvimento desta tese, pela confiança depositada, e pelas muitas horas dispendidas ao longo de todo o trabalho.

Aos professores da banca de qualificação, Prof. Dr. Adilson Gonzaga e Prof. Dr. Aparecido Nilceu Marana, pelas preciosas contribuições, as quais foram muito importantes para o desenvolvimento desta tese.

Aos professores da banca, pela gentileza e disponibilidade por aceitarem em contribuir para o desenvolvimento desta tese.

À todos os meus amigos e familiares, pelas palavras constantes de apoio e incentivo que sempre me fizeram acreditar.

Ao meu labrador Hulck, por sempre estar disponível para brincar e me fazer sorrir, incondicionalmente.

À todos os professores e funcionários do Departamento de Engenharia Elétrica e de Computação da Escola de Engenharia de São Carlos, da USP, pelo apoio e suporte ao longo de toda minha vida acadêmica.

À Universidade de São Paulo, por toda a sua estrutura e organização.

"First, think. Second, dream. Third, believe. And finally, dare." (Walt Disney)

RESUMO

AFFONSO, Alex Antonio. **Reconhecimento facial em ambientes não controlados por meio do "High-Boost Weber Descriptor" na região periocular.** 2018. 167 p. Tese (Doutorado em Ciências, Programa de Engenharia Elétrica) – Escola de Engenharia de São Carlos, Universidade de São Paulo, 2018.

O reconhecimento facial automático é uma tarefa muito importante para a sociedade moderna, pois possibilita o desenvolvimento de diversas aplicações, tais como o controle de imigração em aeroportos, a autenticação de documentos, etc. Muitas destas aplicações ocorrem em ambientes não controlados, onde as fotos são obtidas com grandes variações de poses e expressões faciais, de iluminação, no uso de maquiagem e acessórios, etc. A tarefa de reconhecimento facial automático em ambientes não controlados é ainda muito desafiadora e tem sido alvo de muitas pesquisas no mundo todo nos últimos anos. Dentro deste contexto, esta tese propõe e implementa um conjunto de novos métodos que visam contribuir para o avanço do estado da arte relacionado a este tema de pesquisa. Inicialmente foi proposto o HBWLF, um filtro para enfatizar as componentes de alta frequência da imagem, sem eliminar as de baixa, realçando assim os diversos detalhes das imagens faciais. Em seguida foi proposta uma versão mais geral deste filtro, o MHBWLF, que considera uma vizinhança circular, ao invés de uma grade regular de 3x3 pixels. O MHBWLF foi aplicado em conjunto com um filtro MOSSE no desenvolvimento de um método para a localização precisa dos centros dos olhos. Aproveitando as características do MHBWLF e inspirado em outros descritores foi proposto um novo descritor, o HBWD. Por fim, foi introduzido um novo método de reconhecimento facial, baseado no HBWD. O método proposto emprega o HBWD para descrever densamente a região periocular e, a fim de reduzir a dimensão dos dados, foi proposto um algoritmo de mapeamento baseado no método de agrupamento k-Means++. Os métodos propostos foram todos avaliados utilizando-se imagens das bases LFW, FGLFW e BioID e os resultados experimentais obtidos mostram que os métodos propostos tem desempenho superior a vários outros métodos estado da arte.

Palavras-chave: HBWLF. MHBWLF. HBWD. Reconhecimento facial. Localização dos olhos. Visão Computacional.

ABSTRACT

AFFONSO, Alex Antonio. **Face Recognition under Uncontrolled Scenarios using the new High-Boost Weber Descriptor in the Periocular Region**. 2018. 167 p. Tese (Doutorado em Ciências, Programa de Engenharia Elétrica) – Escola de Engenharia de São Carlos, Universidade de São Paulo, 2018.

The task of automatic face recognition is very important for modern society and very useful for many kind of applications, such as automatic recognition of credit card users, document authentication, security in big events and others. Further, this is a challenging task when performed in uncontrolled scenarios, which involve great variations in imaging conditions such as illumination, poses and facial expressions, partial occlusion due to hair or glasses, makeup, etc. This thesis first introduces the new High-Boost Weber Local Filter (HBWLF) that emphasizes high-frequency components, without eliminating the low-frequency ones, and thus enhances the details of a face. It is also introduced the new MHBWLF (Multiscale High-Boost Weber Local Filter), which is a multiscale version of HBWLF. A new method for precise eye localization is presented, where a MOSSE filter is used for learning the features enhanced by MHBWLF. This thesis also introduces a new local descriptor called HBWD (High-Boost Weber Descriptor) which combines some features of MHBWLF, SIFT and CS-LMP. Finally, a new method of face recognition is presented. The proposed method basically detects the faces, localizes their eyes and performs a face alignment. After that the region of interest (ROI) is more accurately cropped and described using the new HBWD in a dense sampling scheme (sampling each pixel). A new algorithm, based on the known clustering method k-Means++, reduces the dimensionality of the HBWD descriptors densely applied on each ROI, and produces a signature for the image pair being compared. Finally, a SVM is used to classify the images as a matched or mismatched pair. The proposed methods were evaluated using images from the well-known LFW, FGLFW and BioID databases and the experimental results show that the proposed methods outperform other state-of-the-art approaches.

Keywords: HBWLF. MHBWLF. HBWD. Face recognition. Eye Localization. Computer Vision.

LISTA DE ILUSTRAÇÕES

Figura 1 – Marcos históricos de trabalhos em reconhecimento facial.	35
Figura 2 – Comparação das técnicas PCA e LDA em um problema de classificação de dados para duas classes.	36
Figura 3 – Visão geral do método de reconhecimento facial apresentado em (SIMONYAN et al., 2013).	42
Figura 4 – Exemplos de região periocular: (a) e (b) sem sobrancelhas; (c) e (d) com sobrancelhas.	44
Figura 5 – Regiões utilizadas para reconhecimento facial: (a) face toda; (b) nove regiões locais definidas por Kumar et al. (2011).	45
Figura 6 – Codificação binária do LBP.	46
Figura 7 – Método LBP de vizinhança circular.	47
Figura 8 – Detecção de um ponto de máximo ou mínimo da DOG.	49
Figura 9 – Formação dos descritores SIFT.	51
Figura 10 – Ilustração do cálculo do WLD.	54
Figura 11 – Métodos LBP e CS-LBP para uma vizinhança de 8 pixels.	55
Figura 12 – Vizinhança 3x3 para cálculo do HBWLF.	60
Figura 13 – Exemplos de imagens filtradas: (a) imagens originais; (b) imagens filtradas pelo HBWLF, com $N \times N = 7 \times 7$ e $\beta = 0,2$	61
Figura 14 – Vizinhança circular com $NV = 8$	62
Figura 15 – Diagrama geral do método proposto para localização dos olhos.	63
Figura 16 – Fluxograma do bloco de pré-processamento.	64
Figura 17 – Face ajustada para 120x120 pixels e recorte da região de interesse. ...	64
Figura 18 – Grade regular com $N_C \times N_C$ células, ao redor do pixel central x_C . Cada célula contém $N_p \times N_p$ pixels.	68
Figura 19 – Exemplo de grade regular recortada: (a) imagem de entrada; (b) imagem filtrada pelo MHBWLF; (c) gaussiana de ponderação; (d) imagem filtrada ponderada.	68
Figura 20 – Exemplo de montagem do histograma final do HBWD.	69
Figura 21 – Diagrama geral do método proposto para reconhecimento facial em ambientes não controlados, visando o problema do <i>pair matching</i>	71
Figura 22 – Alinhamento espacial de uma imagem facial.	72

Figura 23 – Exemplos de regiões de interesse extraídas em faces da base LFW. ...	74
Figura 24 – Dense sampling na ROI usando o HBWD.	74
Figura 25 – Agrupamento dos kclusters para cada ROI-i.	75
Figura 26 – Mapeamento dos vetores da ROI-1 nos clusters da ROI-2.	76
Figura 27 – Histograma final para comparação das faces 1 e 2.	76
Figura 28 – Convenção dos eixos x e y adotados neste trabalho.	79
Figura 29 – Exemplos de imagens da atriz Jennifer Aniston contidas na base LFW.	81
Figura 30 – Detecção facial das imagens da Figura 29 usando o detector de Viola-Jones.	81
Figura 31 – Exemplos de outras imagens da base LFW.	82
Figura 32 – Exemplos de pares negativos (pessoas diferentes) da base LFW.	83
Figura 33 – Exemplos de pares negativos (pessoas diferentes) listados na FGLFW.	83
Figura 34 – Exemplos de imagens da mesma pessoa da base BioID.	85
Figura 35 – Imagens para caracterização dos filtros: (a) imagem sintética; (b) e (c) imagens da base LFW; (d) e (e) imagens da base BioID.	92
Figura 36 – Imagem da Figura 35 (a) filtrada pelo HBWLF, usando um filtro pela média da vizinhança de (I) 5x5; (II) 7x7 e (III) 9x9. O parâmetro β vale (a) 1,00; (b) 0,30 e (c) 0,10.	93
Figura 37 – Imagem da Figura 35 (b) filtrada pelo HBWLF, usando um filtro pela média da vizinhança de (I) 5x5; (II) 7x7 e (III) 9x9. O parâmetro β vale (a) 1,00; (b) 0,30 e (c) 0,10.	93
Figura 38 – Imagem da Figura 35 (c) filtrada pelo HBWLF, usando um filtro pela média da vizinhança de (I) 5x5; (II) 7x7 e (III) 9x9. O parâmetro β vale (a) 1,00; (b) 0,30 e (c) 0,10.	94
Figura 39 – Imagem da Figura 35 (d) filtrada pelo HBWLF, usando um filtro pela média da vizinhança de (I) 5x5; (II) 7x7 e (III) 9x9. O parâmetro β vale (a) 1,00; (b) 0,30 e (c) 0,10.	94
Figura 40 – Imagem da Figura 35 (e) filtrada pelo HBWLF, usando um filtro pela média da vizinhança de (I) 5x5; (II) 7x7 e (III) 9x9. O parâmetro β vale (a) 1,00; (b) 0,30 e (c) 0,10.	95

Figura 41 – Imagem da Figura 35 (a) filtrada pelo MHBWLF, usando um filtro pela média da vizinhança de 7×7 , β igual a 0,30 e R igual a (I) 1; (II) 2; (III) 3; (IV) 5 e (V) 7. O parâmetro NV vale (a) 1; (b) 4 e (c) 8.	96
Figura 42 – Comparação do filtro MHBWLF e WLD: (a) face original apresentada em Chen et al. (2008); (b) face filtrada pelo WLD; (c) face filtrada pelo MHBWLF.	96
Figura 43 – Resultados obtidos usando o método proposto (círculos cheios em vermelhos) e o <i>ground-truth</i> (marcas x em amarelo) para o conjunto de 2000 imagens da base LFW: (a) $D=0$; (b) $D \approx 0.05$; (c) $D \approx 0.10$; (d) $D \approx 0.15$ e (e) $D \approx 0.20$	101
Figura 44 – Grade de pontos utilizada no experimento.	102
Figura 45 – Imagens da base BioID selecionadas para o experimento: (I) Imagem 1, (II) Imagem 2, (III) Imagem 3, (IV) Imagem 4, (V) Imagem 5 e (VI) Imagem 6; (a) imagem original da base BioID, (b) imagem com grade de pontos.	105
Figura 46 – Adição de ruído Gaussiano: (a) imagem original, (b) imagem com ruído Gaussiano (Caso 1) e (c) imagem com ruído Gaussiano (Caso 2).	107
Figura 47 – Resultados gráficos do teste de adição de ruído gaussiano.	108
Figura 48 – Adição de ruído Speckle: (a) imagem original, (b) imagem com ruído Speckle (Caso 1) e (c) imagem com ruído Speckle (Caso 2).	110
Figura 49 – Resultados gráficos do teste de adição de ruído Speckle.	111
Figura 50 – Adição de ruído Salt & Pepper: (a) imagem original, (b) imagem com ruído Salt & Pepper (Caso 1) e (c) imagem com ruído Salt & Pepper (Caso 2).	113
Figura 51 – Resultados gráficos do teste de adição de ruído Salt & Pepper.	114
Figura 52 – Transformação de rotação: (a) imagem original. Imagem rotacionada: (b) 5° , (c) 10° , (d) 15° , (e) 20° , (f) 25° e (g) 30°	114
Figura 53 – Filtro unsharp: (a) imagens originais, (b) imagens filtradas.	118
Figura 54 – Filtro motion: (a) imagens originais, (b) imagens filtradas.	120

LISTA DE TABELAS

Tabela 1 – Ajuste do HBWLF usando a base LFW.....	97
Tabela 2 – Acurácia de localização (%) usando o conjunto de 2000 imagens da base LFW, com o mesmo protocolo para todos os filtros.....	98
Tabela 3 – Acurácia de localização (%) usando imagens da base LFW – uma comparação com outros resultados encontrados na literatura.	99
Tabela 4 – Acurácia de localização (%) usando a base BioID, com o mesmo protocolo para todos os filtros.	100
Tabela 5 – Acurácia de localização (%) usando a base BioID – uma comparação com outros resultados encontrados na literatura.....	100
Tabela 6 – Imagens selecionadas aleatoriamente da base BioID.....	102
Tabela 7 – Resultado do teste de correspondência de imagem com adição de ruído Gaussiano de média nula (50 amostras para cada teste).	108
Tabela 8 – Resultado do teste de correspondência de imagem com adição de ruído Speckle de média nula (50 amostras para cada teste).	111
Tabela 9 – Resultado do teste de correspondência de imagem com adição de ruído Salt & Pepper de média nula (50 amostras para cada teste).	112
Tabela 10 – Resultado do teste de correspondência de imagem com a transformação de rotação nas imagens.	115
Tabela 11 – Resultados do teste de correspondência para o filtro average de tamanho 5x5.....	117
Tabela 12 – Resultados do teste de correspondência para o filtro unsharp.....	119
Tabela 13 – Resultados do teste de correspondência para o filtro motion.....	121
Tabela 14 – Resultado do teste de correspondência de imagem com a transformação de escala nas imagens.....	122
Tabela 15 – Configuração inicial, baseada no teste de correspondência de imagens.	123
Tabela 16 – Ajuste de $N \times N$ do MHBWLF, na base LFW.....	124
Tabela 17 – Ajuste de β do MHBWLF, na base LFW.....	124
Tabela 18 – Ajuste de R e NV do MHBWLF, na base LFW.	124
Tabela 19 – Ajuste de THIL do HBWD, na base LFW.....	125
Tabela 20 – Ajuste de $N_c \times N_c$ do HBWD, na base LFW.....	125

Tabela 21 – Ajuste de $N_p \times N_p$ do HBWD, na base LFW.....	125
Tabela 22 – Ajuste de Nbins do HBWD, na base LFW.....	126
Tabela 23 – Ajuste de k -Clusters, na base LFW.....	126
Tabela 24 – Ajuste do mapeamento, na base LFW.....	127
Tabela 25 – Configuração final ajustada, na base LFW.	127
Tabela 26 – Resultados do método proposto e de outros métodos encontrados na literatura para a base LFW, no modo restrito e sem utilizar imagens externas.	128
Tabela 27 – Resultados do método proposto e de outros métodos encontrados na literatura para a base FGLFW.	130
Tabela 28 – Pares de imagens da base BioID para reconhecimento facial baseado no problema do <i>pair matching</i> . As colunas y representam: 1 – matched; 0 – mismatched.	153

LISTA DE SIGLAS

ASEF – *Average of Synthetic Exact Filters*
AUC – *Area Under Curve*
BoF – *Bag of Features*
CS-LBP – *Center Symmetric Local Binary Pattern*
CS-LMP – *Center Symmetric Local Mapped Pattern*
DOG – *Difference of Gaussians*
FERET – *Facial Recognition Technology*
FGLFW – *Fine-Grained Labeled Faces in the Wild*
FN – *False Negative*
FP – *False Positive*
FRVT – *Face Recognition Vendor Test*
FV – *Fisher Vector*
HBWD – *High-Boost Weber Descriptor*
HBWLF – *High-Boost Weber Local Filter*
HOG – *Histograms of Oriented Gradients*
IEEE – *Institute of Electrical and Electronics Engineers*
k-NN – *k-Nearest Neighbours*
LBP – *Local Binary Pattern*
LDA – *Linear Discriminant Analysis*
LFW – *Labeled Faces in the Wild*
MHBWLF – *Multiscale High-Boost Weber Local Filter*
MOSSE – *Minimum Output Sum of Squared Error*
MLBP – *Multi-Scale Local Binary Pattern*
MLP – *Multilayer Perceptron*
MRF – *Markov Random Field*
NIST – *National Institute of Standards and Technology*
ORL – *Olivetti Research Laboratory*
PCA – *Principal Component Analysis*
ROI – *Region of interest*
RSF – *Robust Statistical Frontalization*
SIFT – *Scale Invariant Feature Transform*

SVM – *Support Vector Machine*

TN – *True Negative*

TP – *True Positive*

WLD – *Weber Local Descriptor*

SUMÁRIO

Agradecimentos.....	9
Resumo	13
Abstract.....	15
Lista de Ilustrações.....	17
Lista de Tabelas	21
Lista de Siglas	23
Sumário	25
1 Introdução	29
1.1 Caracterização do Tema	29
1.2 Formulação do Problema de Pesquisa.....	30
1.3 Objetivos e Hipótese	31
1.4 Contribuições	32
1.5 Organização da Tese	33
2 Fundamentação Teórica	35
2.1 Reconhecimento Facial	35
2.2 Descritores Clássicos.....	45
2.2.1 LBP.....	46
2.2.2 SIFT.....	47
2.2.3 WLD	52
2.2.4 CS-LBP	55
2.2.5 CS-LMP.....	56
2.3 Considerações Finais.....	57
3 Métodos Propostos	59
3.1 High-Boost Weber Local Filter (HBWLF).....	59

3.2	Multiscale High-Boost Weber Local Filter (MHBWLF)	61
3.3	Método para Localização Precisa dos Olhos	62
3.3.1	Bloco de Pré-Processamento	63
3.3.2	Filtro Mosse	64
3.3.3	Avaliação do Pico de Correlação	66
3.3.4	Localização de Outros Pontos Faciais	66
3.4	High-Boost Weber Descriptor (HBWD)	67
3.5	Método para Reconhecimento Facial	70
3.5.1	Imagens de Entrada.....	70
3.5.2	Detecção de Pontos de Interesse	70
3.5.3	Aplicação do HBWD	72
3.5.4	Redução Dimensional por Agrupamento	74
3.5.5	Classificação usando Máquinas de Vetores de Suporte	76
3.6	Considerações Finais	77
4	Materiais e Métodos	79
4.1	Bases de Dados.....	79
4.1.1	<i>Labeled Faces in the Wild</i> (LFW).....	80
4.1.2	<i>Fine-Grained</i> LFW (FGLFW).....	82
4.1.3	BioID	84
4.1.4	Localização dos Centros dos Olhos: Ground-Truth	84
4.1.5	Reconhecimento Facial: <i>Pair Matching</i>	86
4.2	CrITÉRIOS de Avaliação	86
4.2.1	Ajuste Paramétrico dos Filtros HBWLF e MHBWLF	87
4.2.2	Localização Precisa dos Olhos.....	87
4.2.3	Correspondência de Imagens	88
4.2.4	Reconhecimento Facial	90
5	Resultados	91

5.1	Filtros: HBWLF e MHBWLF.....	91
5.2	Localização Precisa dos Olhos	97
5.3	Correspondência de Imagens com o HBWD.....	101
5.3.1	Adição de Ruído.....	106
5.3.2	Transformação de Rotação	114
5.3.3	Filtros.....	116
5.3.4	Transformação de Escala.....	121
5.4	Reconhecimento Facial	123
5.4.1	Resultados na Base LFW.....	123
5.4.2	Resultados na Base FGLFW	129
5.4.3	Resultados na Base BioID.....	131
5.5	Considerações Finais.....	132
6	Conclusões	137
6.1	Contribuições	140
6.2	Sugestões de Trabalhos Futuros.....	142
6.3	Lista de Artigos Publicados	143
	Referências.....	145
	Apêndice A	153

1 INTRODUÇÃO

Neste capítulo é apresentada uma introdução do tema de pesquisa abordado nesta tese de doutorado onde, em primeiro lugar, apresenta-se a caracterização do tema de pesquisa, demonstrando a importância de métodos de reconhecimento facial para a sociedade e como a academia tem tratado este problema. Em seguida apresentam-se a formulação do problema de pesquisa, os objetivos e a hipótese, bem como as contribuições alcançadas. Por fim descreve-se a organização da tese.

1.1 Caracterização do Tema

O ser humano é capaz de identificar as pessoas por meio de suas faces e faz isto muito bem, mesmo em ambientes não controlados, onde as pessoas são expostas a variações na iluminação, na pose, nas expressões faciais, no uso de acessórios e maquiagem, etc. Esta capacidade despertou um grande interesse de diversos pesquisadores e cientistas em estudar o sistema perceptual e cognitivo dos seres humanos a fim de melhor compreendê-lo, tratá-lo e mesmo reproduzi-lo artificialmente. A Universidade de Harvard, por exemplo, mantém um centro de estudo da *Prosopagnosia* (FACEBLIND.ORG, 2015), uma desordem que afeta a capacidade de reconhecimento facial de um indivíduo, produzindo resultados úteis para a utilização no tratamento de pacientes e também para pesquisadores de algoritmos de reconhecimento facial.

De fato, o reconhecimento facial automático é de grande importância para a sociedade moderna, possibilitando o desenvolvimento de diversas aplicações: reconhecimento de usuários de cartão de crédito, autenticação de documentos (licença de motorista, passaporte, etc.), segurança pública e privada (em bancos, lojas, eventos esportivos, etc.), controle de acesso (em aeroportos e rodoviárias, empresas, condomínios, universidades, etc.), entre diversas outras aplicações. Dado à grande importância de sistemas de reconhecimento facial, diversas empresas

surgiram oferecendo suas soluções ao mercado, o que pode ser comprovado no site FRVT (do inglês, *Face Recognition Vendor Test*), o qual é um programa de testes e avaliações de soluções comerciais em reconhecimento facial e contém uma série de relatórios comparando as diversas soluções comerciais existentes (FRVT, 2015). Atualmente este programa é suportado pelo NIST (do inglês, *National Institute of Standards and Technology*), vinculado ao Departamento de Comércio dos Estados Unidos da América.

Este tema, muito importante para a sociedade, despertou e continua despertando o interesse de muitos pesquisadores ao redor do mundo todo. Um dos primeiros trabalhos de grande relevância na área foi apresentado por Kanade (1973), onde as faces são descritas pelas suas características geométricas. Outros trabalhos surgiram, destacando-se: as *eigenfaces* (TURK; PENTLAND, 1991), que utilizam a técnica PCA; as *fisherfaces* (BELHUMEUR; HESPANHA; KRIEGMAN, 1997), que empregam a técnica LDA; vários métodos que utilizam descritores locais, tais como, por exemplo, LBP (HADID; PIETIKAINEN; AHONEN, 2004), SIFT (BICEGO et al., 2006) e WLD (GONG; LI; XIANG, 2011); uso de redes neurais profundas (TAIGMAN et al., 2014); etc. Uma pesquisa na base IEEE (do inglês, *Institute of Electrical and Electronics Engineers*) com as palavras chave “face” e “recognition” retornou 26.958 resultados, dos quais 9.891 estão compreendidos entre os anos de 2013 e 2018, demonstrando assim claramente o grande interesse atual pelo tema (IEEE Search: face recognition, 2018).

1.2 Formulação do Problema de Pesquisa

O desenvolvimento de novos métodos de reconhecimento facial em ambientes não controlados é ainda objeto de estudo de diversos autores ao redor do mundo todo, sendo um tema de grande importância para a sociedade, possibilitando o desenvolvimento de muitas aplicações práticas reais. Dentro deste contexto, este trabalho de doutorado aborda o reconhecimento facial em ambientes não controlados baseados no problema do *pair matching*.

O termo “ambientes não controlados” significa que as imagens de entrada do método de reconhecimento facial não foram obtidas somente em estúdios, onde

diversas variáveis (iluminação, pose, expressões faciais, parâmetros da câmera, entre outras) são controladas e mantidas constantes. Na verdade, tais imagens de entrada foram obtidas em ambientes diversos (estádios, praias, lojas, eventos esportivos, por exemplo), os quais apresentam grandes variações de iluminação, pose, expressões faciais, objetos de fundo, vestuário, utilização de acessórios e maquiagem, etc.

O outro termo apresentado, “*pair matching*”, se caracteriza pela comparação de duas imagens faciais, verificando se as mesmas pertencem à uma mesma pessoa ou não. Desta forma, verifica-se que neste trabalho as imagens são analisadas aos pares (*matching* de 1:1). A escolha do *pair matching* não significa que esta abordagem não possa ser utilizada também no problema de recuperação de imagens em um banco de dados (*matching* de 1:“várias imagens”), pois o núcleo do método de reconhecimento facial pode ser aplicado em ambos os casos.

1.3 Objetivos e Hipótese

O principal objetivo deste trabalho de doutorado consiste no desenvolvimento de um novo método de reconhecimento facial baseado no problema do *pair matching* e aplicado em ambientes não controlados. Para tanto, o método proposto deve executar algumas tarefas principais: detecção facial, detecção de pontos de interesse, alinhamento facial, descrição de uma região de interesse e classificação.

A detecção facial consiste na tarefa de detectar uma face, ou seja, de localizar uma face dentro de uma imagem de entrada. Em seguida, o método deve localizar pontos de interesse dentro desta face como, por exemplo, os centros dos olhos. Estes pontos servem para um algoritmo fazer o alinhamento facial, permitindo assim um recorte mais preciso da região de interesse, a qual deverá ser descrita para posterior classificação.

A hipótese considerada no desenvolvimento deste trabalho de doutorado foi: a utilização de descritores locais permite a extração de importantes características das faces em ambientes não controlados, possibilitando o seu reconhecimento.

1.4 Contribuições

A primeira contribuição deste trabalho foi o desenvolvimento do *High-Boost Weber Local Filter* (HBWLF) a fim de realçar os detalhes de uma imagem facial, realçando o seu conteúdo de alta frequência sem, no entanto, eliminar o seu conteúdo de baixa frequência. O HBWLF foi inspirado no WLD, porém ao invés de utilizar o pixel central utiliza a média da região vizinha ao mesmo, o que lhe confere uma maior robustez quanto ao ruído. O WLD utiliza a função arco-tangente como função de saída, enquanto no HBWLF utiliza-se a função sigmoide, o que lhe permite um melhor ajuste.

Posteriormente foi desenvolvido o *Multiscale High-Boost Weber Local Filter* (MHBWLF), que corresponde à segunda contribuição deste trabalho. Este filtro é uma evolução do HBWLF: enquanto o HBWLF considera uma grade regular de tamanho fixo igual a 3x3 pixels, no MHBWLF consideram-se N_V pixels distribuídos igualmente em uma circunferência de raio R .

Foi proposto ainda um método para a localização precisa dos centros dos olhos, o qual é a terceira contribuição deste trabalho. Neste método o MHBWLF é utilizado para realçar as características das faces e um filtro MOSSE foi utilizado para o aprendizado da localização dos centros dos olhos, utilizando as imagens filtradas pelo MHBWLF como entrada. O método foi avaliado nas bases BioID e LFW (do inglês, *Labeled Faces in the Wild*) e os resultados obtidos foram superiores aos métodos estado da arte encontrados na literatura.

A quarta contribuição deste trabalho constitui em um novo descritor local, chamado de *High-Boost Weber Descriptor* (HBWD), combinando características dos descritores SIFT, WLD e CS-LMP. O HBWD faz uso de uma imagem filtrada pelo MHBWLF e é montado em uma grade regular com $N_C \times N_C$ células, onde para cada célula é gerado um histograma com N_{bins} bins. O vetor final de saída do HBWD contém a dimensão de $N_C \times N_C \times N_{bins}$. Diversos testes de correspondência de imagens faciais foram realizados com o HBWD e o mesmo foi comparado com os descritores WLD e CS-LBP, obtendo melhores resultados, principalmente no aspecto da robustez quanto ao ruído.

A quinta contribuição deste trabalho corresponde ao desenvolvimento do método proposto para reconhecimento facial baseado no problema do *pair matching*.

O método proposto faz uso do algoritmo desenvolvido para a localização dos centros dos olhos a fim de realizar o alinhamento das faces de entrada, possibilitando desta forma um recorte mais preciso da região de interesse. A região de interesse utilizada no método foi a região periocular e a mesma foi descrita densamente pelo HBWD, ou seja, o HBWD foi aplicado em todos os pixels da região periocular. A fim de reduzir a dimensão dos dados, foi utilizado o método de k-Means++ em um mapeamento proposto, gerando assim uma assinatura de saída para o par de imagens em comparação. Por fim foi utilizado um classificador SVM para a decisão final (*matched* ou *mismatched*). O método proposto foi avaliado nas bases LFW, FGLFW e BioID e os resultados obtidos são superiores aos resultados de vários outros métodos estado da arte encontrados na literatura.

A sexta e última contribuição deste trabalho consiste na elaboração de uma lista de pares de imagens para o teste baseado no problema do *pair matching* com imagens oriundas da base BioID, o que dará a possibilidade de outros autores também testarem seus métodos na base BioID e compararem com os resultados obtidos neste trabalho.

1.5 Organização da Tese

Esta tese de doutorado apresenta a seguinte organização:

- ✓ **Capítulo 1 – Introdução:** apresenta uma introdução da tese propriamente dita, caracterizando o tema da pesquisa, apresentando seus objetivos e hipótese, bem como as principais contribuições desenvolvidas;
- ✓ **Capítulo 2 – Fundamentação Teórica:** neste capítulo é feita uma revisão de diversos métodos de reconhecimento facial desenvolvidos ao longo do tempo, bem como de vários descritores locais, permitindo assim melhor definir o desenvolvimento deste trabalho;
- ✓ **Capítulo 3 – Métodos Propostos:** são apresentados todos os métodos propostos que foram desenvolvidos ao longo deste trabalho;

- ✓ **Capítulo 4 – Materiais e Métodos:** descreve os materiais e os métodos utilizados ao longo dos experimentos realizados;
- ✓ **Capítulo 5 – Resultados:** concentra todos os resultados experimentais obtidos neste trabalho;
- ✓ **Capítulo 6 – Conclusões:** recapitula o trabalho apresentado nesta tese culminando nas conclusões finais e ainda, apresentando as limitações da abordagem proposta, bem como as sugestões de trabalhos futuros. Apresenta ainda uma lista de artigos publicados e submetidos.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta uma revisão de diversos métodos clássicos de reconhecimento facial, a qual foi muito importante para o desenvolvimento do trabalho proposto. Além da revisão de tais métodos, o capítulo descreve ainda diversos descritores locais, os quais foram utilizados neste trabalho como inspiração e também na comparação dos resultados obtidos em diversas etapas.

2.1 Reconhecimento Facial

A Figura 1 ilustra diversos marcos históricos de trabalhos na área de reconhecimento facial ao longo do tempo. Kanade (1973) apresentou um dos primeiros trabalhos de maior relevância na área, onde uma face é representada pelas distâncias e ângulos de alguns de seus pontos fiduciais como, por exemplo, os cantos dos olhos e os extremos da boca, sendo considerado um trabalho que faz o reconhecimento facial por meio das características geométricas faciais.

Figura 1 – Marcos históricos de trabalhos em reconhecimento facial.

Fonte: autoria própria, utilizando imagens de diversos artigos apresentados nesta seção.

Outro trabalho de grande importância foram as autofaces ou, do inglês, as *eigenfaces* (TURK; PENTLAND, 1991): este método consiste na aplicação da técnica PCA (do inglês, *Principal Component Analysis*) para a redução da dimensão dos dados no reconhecimento facial, onde uma face passa a ser representada por um vetor contendo suas componentes principais. Outra técnica muito conhecida são as *fisherfaces* (BELHUMEUR; HESPANHA; KRIEGMAN, 1997), onde é empregada a técnica de redução de dimensão dos dados conhecida como LDA (do inglês, *Linear Discriminant Analysis*).

A técnica LDA, mais do que reduzir a dimensão dos dados, é uma técnica que maximiza o raio de variância entre as classes (*between-class*) ao mesmo tempo em que minimiza a variância dentro das classes (*within-class*). Fazendo um comparativo, a técnica PCA reduz a dimensão dos dados pela representação dos mesmos pelas suas componentes principais sem, no entanto, reduzir a variância dos pontos de dentro de uma determinada classe, uma vez que não faz uso de informações que identifiquem uma determinada imagem dentro de um grupo ou classe. Por outro lado, a técnica LDA utiliza as informações de agrupamento, incrementando assim a separabilidade entre as classes.

Figura 2 – Comparação das técnicas PCA e LDA em um problema de classificação de dados para duas classes.

Fonte: adaptado de (BELHUMEUR; HESPANHA; KRIEGMAN, 1997).

A Figura 2 ilustra uma comparação das técnicas PCA e LDA aplicadas em um problema de classificação de duas classes distintas, com os dados originalmente distribuídos em um espaço 2D e projetados em um espaço 1D. Os dados projetados pela técnica PCA estão em uma reta, porém não são linearmente separáveis. Por outro lado, a técnica LDA faz com que os dados sejam mais facilmente classificados ou separados, uma vez que tal técnica aumenta a variância entre as classes 1 e 2, ao mesmo tempo que diminui a variância dentro de cada classe. Diversos experimentos realizados e apresentados em (BELHUMEUR; HESPANHA; KRIEGMAN, 1997) demonstram que a técnica *fisherfaces* apresenta melhores resultados que a técnica *eigenfaces*, mesmo em condições de variações de luminosidade e expressões faciais.

Conforme ilustrado na Figura 1, diversos métodos empregando descritores locais foram utilizados em métodos de reconhecimento facial. Dentro deste contexto, o método LBP (do inglês, *Local Binary Pattern*) proposto em (OJALA; PIETIKAINEN; HARWOOD, 1996) foi largamente explorado por diversos autores. Em (HADID; PIETIKAINEN; AHONEN, 2004) uma face é representada por suas características LBP, obtidas local e globalmente, e codificadas em um histograma. Por outro lado, em (AHONEN; HADID; PIETIKAINEN, 2006) a face é dividida em várias regiões e, para cada região, é extraído um vetor de características baseado no histograma LBP da mesma. Desta forma, cada região é descrita localmente e a combinação de todas as regiões leva a uma descrição global da face. Huang et al. (2011) apresentam uma revisão detalhada sobre a aplicação de descritores LBP e suas variantes na análise de imagens faciais (descrição, detecção, reconhecimento, análise de expressão e classificação demográfica facial).

Um dos descritores mais clássicos encontrados na literatura é o SIFT (do inglês, *Scale Invariant Feature Transform*) (LOWE, 2004). O primeiro trabalho aplicando SIFT em reconhecimento facial encontrado foi (BICEGO et al., 2006), onde foram utilizadas três estratégias diferentes de matching: mínima distância dos pontos chave, de características dos olhos e da boca e, por fim, em uma grade regular.

A estratégia de *matching* de mínima distância dos pontos chave é a mais simples, onde se calculam as distâncias dos pontos chave da imagem de teste com os pontos chave da imagem de referência, procurando-se pelos pontos chave correspondentes de menor distância. Para o cálculo da distância foi considerada a

distância Euclidiana. A segunda estratégia de *matching* utiliza apenas as características geradas pelos descritores SIFT ao redor dos olhos e da boca: os pontos chave ao redor dos olhos da imagem de teste são comparados com os pontos chave ao redor dos olhos da imagem de referência, pelo cálculo da distância Euclidiana dos descritores, sendo feito o mesmo raciocínio para os pontos chave ao redor da boca. A terceira estratégia de *matching* divide a imagem em grades regulares e calculam-se as distâncias Euclidianas dos descritores em suas respectivas grades na imagem de teste e de referência.

Os experimentos demonstraram que a melhor estratégia é a terceira, enquanto a primeira estratégia apresenta os piores resultados. A primeira estratégia, de fato, não considera a localização dos pontos chave na face e pode, portanto, fazer o *matching* de dois pontos chave de regiões distintas nas faces. A segunda estratégia resolve em parte este problema, porém não faz distinção de características dos olhos (direito e esquerdo) e, portanto, pode casar um ponto chave na região do olho direito de uma face com um ponto chave na região do olho esquerdo de outra face.

Os resultados apresentados em (BICEGO et al., 2006) demonstram uma viabilidade da aplicação dos descritores SIFT no reconhecimento facial, porém ressaltam uma necessidade de desenvolvimento de algoritmos de *matching* mais adequados, que levem em consideração algumas características peculiares das faces como, por exemplo, sua forma e variabilidade.

Em (LUO et al., 2007) os descritores SIFT são utilizados e agrupados pelo algoritmo de *k-means*. Este algoritmo foi testado na base FERET (do inglês, *Facial Recognition Technology*), avaliando-se as variações das expressões faciais, da iluminação e do envelhecimento. O algoritmo proposto por Luo et al. (2007) obteve um bom resultado quanto à variação das expressões faciais, porém os resultados para as variações de iluminação e de envelhecimento foram ruins. De fato, os descritores SIFT possuem características muito específicas e que são susceptíveis ao envelhecimento de uma pessoa (LUO et al., 2007).

Zhang et al. (2008) apresentaram um novo método de reconhecimento facial utilizando SIFT e SVM (do inglês, *Support Vector Machine*). O método proposto foi avaliado em duas bases de imagens: ORL (do inglês, *Olivetti Research Laboratory, Cambridge*) e Yale. Os resultados demonstram que o método proposto apresenta bons resultados, especialmente para a variação de expressões faciais e ainda, em

geral, apresentou melhores resultados que os métodos apresentados em (BICEGO et al., 2006) e (LUO et al., 2007). Contudo, o método apresentado por ZHANG et al. (2008) também apresenta resultados insatisfatórios para variações de iluminação e os autores inclusive destacam a necessidade de melhorias no método quanto à robustez neste aspecto.

Em (MAJUMDAR; WARD, 2009) é proposto um método para reconhecimento facial onde as características obtidas pelos descritores SIFT são ranqueadas a fim de selecionar apenas as mais representativas ao problema de reconhecimento: ao invés de usar todas as características SIFT para o reconhecimento facial, o método proposto utiliza apenas as características mais discriminantes, reduzindo assim o tempo necessário para o *matching*.

As características SIFT são ranqueadas em acordo com a relação de Fisher, que corresponde à razão da variação dos dados entre classes com a variação dos dados intra classe. Uma característica SIFT será considerada discriminante se ela for encontrada em mais do que uma imagem de uma mesma pessoa e não for encontrada em imagens de outras pessoas (MAJUMDAR; WARD, 2009). As características SIFT são ranqueadas durante a fase de treinamento, onde as características discriminantes são selecionadas e as demais descartadas.

Desta forma, a fase de treinamento do método proposto por Majumdar e Ward (2009) é mais lenta e, em contrapartida, a fase de teste se torna mais rápida, uma vez que são usadas menos características para o *matching*. Porém, para incluir uma nova pessoa no banco de dados torna-se necessária uma nova execução da fase de treinamento inteira, a fim de atualizar a base de características selecionadas.

Além da seleção de características SIFT, o método proposto por Majumdar e Ward (2009) também divide a imagem em grades regulares, dividindo uniformemente as faces em uma matriz de dimensão igual a 4x4. As características SIFT são sempre comparadas apenas com aquelas de suas respectivas grades. O método foi testado em três bases de imagens: ORL, JAFFE e Yale. Os resultados apontados em (MAJUMDAR; WARD, 2009) demonstram que o referido método é melhor que os métodos prévios utilizando a transformada SIFT para reconhecimento facial. No entanto, cabe ressaltar que o presente método não foi comparado com o método proposto por (ZHANG et al., 2008).

Geng e Jiang (2009) estudaram a versão original de SIFT e analisaram a mesma em aplicações de reconhecimento facial, apontando suas principais

deficiências. E ainda, são apresentados dois novos métodos para reconhecimento facial baseados em SIFT. A versão original de SIFT foi concebida para o reconhecimento e detecção de objetos em geral, os quais normalmente são rígidos, possuindo altos contrastes de um lado para o outro. As faces, por outro lado, não são rígidas, são suaves, possuindo menos estruturas com alto contraste (GENG; JIANG, 2009).

A versão original de SIFT proposta em (LOWE, 2004) remove os pontos chave com baixo contraste ou que são localizados ao longo de bordas dos objetos. No entanto, uma face não possui cantos ou linhas muito bem definidas, tão pouco possui muitas regiões com alto contraste. De fato, Geng e Jiang (2009) demonstram que a remoção de pontos chaves de baixo contraste e ao longo de bordas, em imagens faciais, remove importantes pontos chave, que carregam características importantes para o reconhecimento facial.

Desta forma, o primeiro método proposto em (GENG; JIANG, 2009) é o KPSIFT, o qual mantém todos os pontos chave inicialmente detectados pela aplicação de SIFT. O segundo método proposto é o PDSIFT, que descreve parcialmente os descritores cujos pontos chave são localizados em escalas grandes ou próximos aos limites da face.

Em (GENG; JIANG, 2009) os dois novos métodos são comparados com o método original SIFT, utilizando-se as bases ORL e AR. Os resultados apresentados demonstram que os métodos KPSIFT e PDSIFT obtiveram melhores resultados que o método original SIFT nestas duas bases. E ainda, os métodos KPSIFT e PDSIFT são comparados com métodos holísticos clássicos, apresentando também melhores resultados, principalmente o método PDSIFT.

Wang et al. (2010) apresentam um método para o reconhecimento do gênero a partir de imagens faciais. O método apresentado combina descritores baseados em SIFT com o contorno de imagens faciais a fim de reconhecer o gênero de uma pessoa. Em acordo com Geng e Jiang (2009), a detecção de pontos chave pelo algoritmo original de SIFT apresenta diversos problemas quando aplicado em reconhecimento facial devido à falta de textura, problemas de iluminação e baixa resolução das imagens faciais. Portanto, no método proposto os descritores não são obtidos somente ao redor dos pontos chave, mas sim em uma densa grade regular. O descritor baseado no contorno é invariante à escala e rotação e é baseado nas distâncias dos pontos pertencentes ao contorno das faces. O contorno das faces

possuem informações discriminantes para o reconhecimento do gênero de uma pessoa (WANG et al., 2010). Os descritores SIFT obtidos pela densa grade regular são concatenados em um vetor e então combinados com os descritores de contorno. Em seguida o método utiliza o algoritmo AdaBoost para selecionar as características mais importantes para a classificação (WANG et al., 2010).

Li; Park e Jain (2011) apresentam um modelo discriminativo para o reconhecimento facial com certa tolerância a variação de envelhecimento. O método proposto divide a imagem facial, já alinhada e recortada, em grades regulares sobrepostas e em seguida calcula as características SIFT e MLBP (do inglês, *Multi-Scale Local Binary Pattern*). Os descritores SIFT não são calculados ao redor de pontos chave, mas sim de maneira densa. Uma extensão do método LDA é utilizada para a redução da dimensão dos vetores de características. Os experimentos realizados demonstram que o modelo proposto apresenta bons resultados quanto à invariância ao envelhecimento, quando comparado com outros métodos existentes. No entanto, este método necessita ser aprimorado para se tornar mais robusto e preciso, cabendo ressaltar ainda que o método proposto é vulnerável para as variações de pose (LI; PARK; JAIN, 2011).

Simonyan et al. (2013) apresentam um método de reconhecimento facial baseado em SIFT e em FV (do inglês, *Fisher Vector*), conforme visão geral apresentada na Figura 3. O método é baseado na descrição dos pixels em grades regulares por meio da aplicação de SIFT, densamente, na escala e no espaço (do mesmo modo que em (WANG et al., 2010) utilizam-se apenas os descritores, sem calcular os pontos chave). Posteriormente, as características são codificadas em um vetor FV. Por fim, foi implementado um algoritmo para realizar a redução discriminante de dimensão dos FVs, melhorando assim o desempenho da tarefa de reconhecimento pela projeção dos dados em um sub-espço discriminativo.

O método apresentado em (SIMONYAN et al., 2013) foi testado na base LFW e obteve resultados melhores que aqueles apresentados por outros métodos estado da arte, tanto para os testes irrestritos quanto para os restritos (Labeled Faces in the Wild – Results, 2015). Segundo Simonyan et al. (2013), o uso de características densas permitiu obter um alto desempenho sem, no entanto, necessitar de sofisticados algoritmos para o cálculo dos marcos faciais.

Figura 3 – Visão geral do método de reconhecimento facial apresentado em (SIMONYAN et al., 2013).

Fonte: adaptado de (SIMONYAN et al., 2013).

O método ASEF (do inglês, *Average of Synthetic Exact Filters*), introduzido por Bolme, Draper e Beveridge (2009), utiliza filtros no domínio da frequência para localizar pontos de interesse em uma face como, por exemplo, os centros dos olhos. O método calcula um filtro exato de correlação para cada imagem de treinamento e produz um filtro final pela média dos filtros exatos. Bolme et al. (2010) introduziram ainda o filtro MOSSE (do inglês, *Minimum Output Sum of Squared Error*), que produz filtros similares ao ASEF, porém necessitam de menos imagens de treinamento para atingirem um melhor desempenho.

Outro descritor utilizado em reconhecimento facial é o WLD (do inglês, *Weber Local Descriptor*). Gong, Li e Xiang (2011) apresentam um método de reconhecimento facial baseado no WLD, onde uma imagem de uma face, já detectada, é dividida em um conjunto de regiões e as características de cada região são extraídas pelo WLD. A excitação diferencial é calculada do mesmo modo que em (CHEN et al., 2009), no entanto, a orientação do gradiente é calculada utilizando-se o operador de Sobel, tornando assim o método mais imune ao ruído da imagem de entrada (GONG; LI; XIANG, 2011).

Para o reconhecimento calcula-se a distância Euclidiana das características WLD de todas as regiões da imagem de teste com as imagens do banco de dados e aplica-se o método do vizinho mais próximo. O resultado de cada região é aplicado em uma lógica de decisão que, baseado na quantidade de regiões corretamente identificadas, faz a classificação da imagem como um todo. Tal método foi testado nas bases de dados ORL e Yale e os resultados demonstram que o mesmo apresenta uma acurácia melhor que o descritor LBP, sendo robusto para variações nas expressões faciais, na iluminação, nas poses e em oclusões parciais.

Ullah et al., (2012) utilizam o WLD para o reconhecimento do gênero a partir da imagem de uma face: a imagem facial é dividida em várias regiões, das quais se extraem as características WLD, características estas que são concatenadas e utilizadas para a classificação por meio de um classificador de mínima distância. Os resultados demonstram que, embora o método proposto seja relativamente simples, obtêm-se bons resultados, mesmo quando comparado com métodos mais tradicionais.

O descritor WLD também foi utilizado no reconhecimento de expressões faciais em (WANG et al., 2013) e (SUAISHI; YAUN; KEPING, 2014). Em (WANG et al., 2013) as características obtidas com o descritor são combinadas com aquelas obtidas pelo HOG (do inglês, *Histograms of Oriented Gradients*), enquanto Suaishi, Yaun e Keping (2014) utilizam apenas o WLD aplicado em sub-regiões da imagem, seguidos do classificador SVM.

Outra abordagem muito utilizada atualmente em reconhecimento facial é o emprego de redes neurais profundas. Taigman et al. (2014) apresentam um método de reconhecimento facial que emprega uma rede neural profunda de 9 camadas, com mais de 120 milhões de parâmetros, para a extração de características e classificação das imagens faciais. Este método foi testado na base LFW e obteve uma acurácia de 97,35% no modo irrestrito, com o uso de imagens externas: foram utilizadas 4 milhões de imagens faciais obtidas de mais de 4 mil pessoas. Cabe destacar que a base LFW contém apenas 13233 imagens no total, ou seja, um número de imagens muito menor que aquele utilizado pelos autores.

Os diversos métodos de reconhecimento facial estudados baseiam-se em determinadas regiões de uma face, ou mesmo na face total, incluindo a região do cabelo. Os métodos que se baseiam na face total são ditos métodos holísticos, os quais consideram características globais de uma face. Por outro lado, há outros métodos que se baseiam em características extraídas localmente como, por exemplo, os métodos baseados em SIFT. Há ainda métodos que combinam diversas regiões e características, ou mesmo extraem as características locais de uma maneira densa, ou seja, ao longo de toda a face, como ocorre em (SIMONYAN et al., 2013).

Uma região facial que tem despertado recentemente bastante interesse nos pesquisadores em aplicações de reconhecimento facial é a região periocular, pois é uma das regiões mais discriminantes de uma face (MILLER et al., 2010) (JUEFEI-

XU; LUU; SAVVIDES, 2015). A região periocular é aquela localizada na proximidade dos olhos e sua aquisição não requer muita cooperação do usuário e distâncias muito próximas de um sensor, como ocorre com a biometria da íris ou retina (PARK; ROSS; JAIN, 2009). No entanto, não há uma definição formal para a região periocular, de modo que a mesma pode incluir ou excluir as sobrancelhas ou outros componentes (PARK et al., 2010). A Figura 4 ilustra a região periocular de duas pessoas, sem e com sobrancelhas.

Figura 4 – Exemplos de região periocular: (a) e (b) sem sobrancelhas; (c) e (d) com sobrancelhas.

Fonte: adaptado (PARK; ROSS; JAIN, 2009).

A taxa de reconhecimento obtida pela utilização da região periocular é comparável com a taxa obtida pela utilização da região facial como um todo, usando-se características locais similares em ambos os casos (WOODARD et al., 2010). Em certos casos o uso da região periocular traz melhores resultados que o uso da face toda, especialmente em ambientes não controlados com usuários não cooperativos (XU; SAVVIDES, 2012).

Alguns trabalhos dividem uma face em diversas regiões para a extração de características como, por exemplo, em (KUMAR et al., 2011), conforme ilustra a Figura 5. A Figura 5 (a) mostra uma única região, a qual é utilizada para representar a face como um todo. A Figura 5 (b) mostra nove regiões de uma face, as quais são utilizadas para a extração de características locais – estas regiões são obtidas manualmente, uma única vez, no sistema de coordenadas do algoritmo de alinhamento, podendo ser utilizadas automaticamente para todas as faces alinhadas por este algoritmo (KUMAR et al., 2011). Note ainda que as regiões são relativamente grandes, o que contribui para minimizar os erros de alinhamento, diferenças entre indivíduos e variações nas poses.

Figura 5 – Regiões utilizadas para reconhecimento facial: (a) face toda; (b) nove regiões locais definidas por Kumar et al. (2011).

Fonte: adaptado de (KUMAR et al., 2011).

Verifica-se claramente a grande importância na escolha das regiões de uma face para um determinado sistema de reconhecimento facial e, ainda, que a região periocular é uma das regiões que apresenta as melhores características discriminantes.

2.2 Descritores Clássicos

Nesta seção são apresentados diversos descritores locais clássicos encontrados na literatura e que serviram como inspiração para o desenvolvimento dos métodos propostos nesta tese: LBP, SIFT, WLD, CS-LBP e CS-LMP. Alguns dos descritores aqui apresentados também foram utilizados nos experimentos realizados, permitindo assim uma comparação dos resultados obtidos pelos métodos propostos com descritores já encontrados na literatura dentro de um mesmo protocolo de teste.

2.2.1 LBP

A técnica LBP consiste de um descritor de textura baseado em uma codificação binária da vizinhança de tamanho 3x3 de um pixel, sendo invariante às variações na escala de cinza (OJALA; PIETIKAINEN; HARWOOD, 1996). A Figura 6 (a) ilustra os valores dos pixels de uma vizinhança de tamanho 3x3, na qual o descritor LBP toma o valor do pixel central, no caso igual a 6, e compara os oito vizinhos com este valor. Se o valor de um pixel vizinho for maior ou igual ao valor do pixel central, este terá um valor binário igual a '1', caso contrário, receberá o valor binário '0', gerando assim uma matriz binária de tamanho 3x3, conforme a Figura 6 (b). A Figura 6 (c) apresenta os pesos de cada vizinho e a Figura 6 (d) os resultados da multiplicação pixel a pixel da matriz binária com a matriz de pesos. Por fim, os valores destes últimos oito pixels são somados, resultando no valor LBP do pixel central que, no exemplo dado, é igual a 169.

Figura 6 – Codificação binária do LBP.

Fonte: adaptado de (OJALA; PIETIKAINEN; HARWOOD, 1996).

O procedimento descrito acima é realizado para a imagem inteira, gerando assim uma imagem com os valores LBP. Em seguida cria-se um histograma a partir desta imagem LBP, levantando-se as ocorrências dos diversos valores LBP da respectiva imagem – este histograma pode ser utilizado como vetor de características LBP para a caracterização de texturas como ocorre, por exemplo, em (VIEIRA et al., 2012).

Ojala, Pietikainen e Maenpaa (2002) adaptaram o método LBP para uma vizinhança circular, obtendo-se invariância quanto à rotação, sendo ainda possível a análise multiescalar. A Figura 7 ilustra uma vizinhança circular, de raio R, com P pontos distribuídos uniformemente no círculo de raio R, onde os valores de cinza

que não coincidem com o centro dos pixels são estimados por interpolação (OJALA; PIETIKAINEN; MAENPAA, 2002).

Figura 7 – Método LBP de vizinhança circular.

Fonte: adaptado de (OJALA; PIETIKAINEN; MAENPAA, 2002).

O operador LBP é definido então em função do raio R e do número de pontos P e sua nomenclatura é $LBP_{P,R}$. Sendo assim, o operador $LBP_{P,R}$ produz 2^P valores diferentes de saída. Por exemplo, o operador $LBP_{8,R}$ produz 256 valores diferentes de saída, da mesma forma que o descritor LBP original.

2.2.2 SIFT

SIFT é uma transformada invariante à escala e foi inicialmente introduzida por Lowe (1999) e apresentada detalhadamente em (LOWE, 2004) para reconhecimento de objetos ou cenas. Esta transformada é invariante à escala, translação e rotação e parcialmente invariante na adição de ruído, na variação da iluminação e da visada 3D de um objeto ou uma cena. SIFT pode ser sintetizada em uma transformada que busca por pontos chave estáveis, invariáveis quanto à escala, e que os descreve por meio de um descritor baseado na orientação deste ponto chave. Seu algoritmo é dividido em quatro partes principais:

- ✓ **Deteção de extremos no espaço-escala:** o primeiro estágio do algoritmo busca por candidatos a pontos chave, estáveis, em diversas escalas da imagem;

- ✓ **Localização dos pontos chave:** o segundo estágio calcula a localização e escala de cada ponto chave, bem como procura eliminar pontos chave que tendem a ser instáveis;
- ✓ **Orientação dos pontos chave:** o terceiro estágio computa as orientações de cada ponto chave estável a partir das direções da imagem de gradiente local;
- ✓ **Descritor dos pontos chave:** por fim, o quarto estágio cria os descritores dos pontos chave, descritores estes baseados em histogramas de orientações.

Em um primeiro passo o algoritmo faz a detecção de candidatos a pontos chave, encontrando sua localização e escala, as quais podem ser repetidamente encontradas para um mesmo objeto, mesmo para diferentes pontos de vista (LOWE, 2004). No entanto, diversos trabalhos demonstram que as faces não apresentam muitas estruturas com elevado contraste e, ainda, são muito susceptíveis ao envelhecimento, uso de maquiagem e acessórios, apresentam grandes variações de poses e expressões faciais, o que diminui consideravelmente a repetibilidade da detecção e localização dos pontos chave do algoritmo SIFT quando aplicado nas faces (GENG; JIANG, 2009) (WANG et al., 2010) (LI; PARK; JAIN, 2011) (SIMONYAN et al., 2013). Desta forma, em reconhecimento facial, diversos autores não trabalham com a localização de pontos chave, mas sim com a descrição de regiões de maneira densa.

Foi demonstrado em (LINDBERG, 1994) que, sob algumas suposições, o único núcleo possível do espaço escala é a função Gaussiana. Desta forma, o espaço escala de uma imagem $I(x, y)$ é definido como uma função $L(x, y, \sigma)$, a qual é produzida por meio da convolução de uma gaussiana de escala variável $G(x, y, \sigma)$ com a imagem $I(x, y)$, conforme dado pela Equação 1, com $G(x, y, \sigma)$ definido pela Equação 2.

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad \text{Equação 1}$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad \text{Equação 2}$$

A fim de detectar os pontos chave estáveis com eficiência foi proposto por Lowe (1999) o uso de extremos no espaço-escala na função DOG (do inglês, *Difference of Gaussians*). A DOG é calculada pela diferença de duas imagens filtradas por uma Gaussiana com escalas separadas por uma constante multiplicativa k , em acordo com a Equação 3. A função DOG nada mais é do que a diferença entre duas imagens suavizadas em escalas diferentes, $k\sigma$ e σ .

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad \text{Equação 3}$$

A Figura 8 ilustra o mecanismo de detecção de um ponto de máximo ou mínimo da DOG: um ponto x , candidato a ser um ponto chave, é comparado com seus 8 vizinhos da mesma escala e também com seus 18 vizinhos, sendo 9 de uma escala acima e outros 9 de uma escala abaixo do ponto x . O ponto x é seleccionado como candidato a ponto chave apenas se ele corresponder ao menor ou maior valor que todos os seus 26 vizinhos.

Figura 8 – Detecção de um ponto de máximo ou mínimo da DOG.

Fonte: adaptado de (LOWE, 2004).

O cálculo da orientação de cada ponto chave se faz muito importante para que o descritor deste ponto chave seja representado relativamente a esta orientação, conferindo assim uma invariância à rotação (LOWE, 2004). Para cada imagem suavizada $L(x, y, \sigma)$ são calculadas a magnitude $m(x, y)$, conforme a Equação 4, e a orientação $\Theta(x, y)$, conforme a Equação 5 (LOWE, 2004).

$$m(x, y) = \sqrt{(L(x + 1, y) - L(x - 1, y))^2 + (L(x, y + 1) - L(x, y - 1))^2} \quad \text{Equação 4}$$

$$\theta(x, y) = \tan^{-1} \left(\frac{(L(x, y + 1) - L(x, y - 1))}{(L(x + 1, y) - L(x - 1, y))} \right) \quad \text{Equação 5}$$

Para o cálculo da orientação de um ponto chave monta-se um histograma de orientações, dividindo-se os 360° em 36 faixas, com as orientações dos gradientes de uma região ao redor do ponto chave. As amostras na vizinhança do ponto chave são ponderadas pela magnitude e por uma janela Gaussiana, centrada no ponto chave, com σ uma vez e meia o valor da escala do ponto chave, reduzindo assim a importância dos gradientes mais distantes do ponto chave central.

O maior pico do histograma de orientações, bem como outros picos que sejam maiores que 80% do maior pico, são detectados e usados para calcular a orientação dominante do ponto chave. Caso os picos detectados tenham valores muito próximos em magnitude serão criados vários pontos chave com o mesmo local e escala, porém com orientações diferentes.

Foram descritas anteriormente as etapas responsáveis pelo cálculo da localização, da escala e da orientação dos pontos chave, restando o cálculo do descritor para cada ponto chave. O descritor apresentado por Lowe (2004) foi idealizado para tentar conferir uma determinada invariância quanto à variação de iluminação e ponto de vista 3D.

Para o cálculo do descritor de um determinado ponto chave tem-se primeiramente uma amostragem da imagem gradiente (magnitudes e orientações) em uma região ao redor do ponto chave, para a mesma escala e oitava do ponto chave. A fim de se obter a invariância quanto à rotação, as coordenadas do descritor e as orientações dos gradientes são rotacionadas em relação ao ponto chave. A Figura 9 ilustra o processo de formação de um descritor de um ponto chave.

Na imagem à esquerda da Figura 9 pode-se visualizar a imagem gradiente, onde as magnitudes estão representadas pelo tamanho das setas e as orientações pela orientação da própria seta. A janela circular representa os pesos de ponderação da magnitude dos gradientes e corresponde a uma função Gaussiana com σ igual a uma vez e meia a largura da janela do descritor: esta função Gaussiana dá menos ênfase para os gradientes que estão mais distantes do centro do descritor e evita

mudanças repentinas no descritor com pequenas mudanças na posição da janela (LOWE, 2004).

Figura 9 – Formação dos descritores SIFT.

Fonte: adaptado de (LOWE, 2004).

O descritor do ponto chave é representado na imagem à direita da Figura 9: observa-se que a região ao redor do ponto chave foi dividida em uma janela de 2x2 e que as orientações são acumuladas em um histograma com 8 direções, onde os tamanhos das setas correspondem às magnitudes acumuladas dos gradientes correspondentes. Por fim, o descritor é formado por um vetor contendo os valores das magnitudes para todas as orientações dos 4 histogramas da janela de 2x2 da imagem à direita.

Desta forma, a Figura 9 ilustra uma janela de 2x2, com 4 histogramas de 8 orientações cada e, portanto, o vetor do descritor terá, neste caso, 32 elementos para cada ponto chave. No entanto, os melhores resultados foram obtidos com uma janela de 4x4, contendo 16 histogramas de 8 orientações cada, resultando assim em um vetor de 128 elementos para cada ponto chave (LOWE, 2004).

Por fim, o vetor do descritor de cada ponto chave é modificado a fim de reduzir os efeitos de variações na iluminação. Primeiramente o vetor é normalizado, com comprimento unitário, o que lhe confere uma invariância para quando todos os pixels são multiplicados por uma constante (LOWE, 2004). Uma mudança no brilho, na qual se adiciona uma constante para cada pixel, também não afeta os valores dos gradientes, pois os mesmos são calculados pela diferença de pixels (LOWE, 2004). Note, no entanto, que isto não torna o algoritmo totalmente robusto à

invariância de brilho e contraste, pois na prática estas variações podem ser não homogêneas (não lineares) ao longo da imagem.

Os efeitos de não linearidade na iluminação causam grandes mudanças nas magnitudes de alguns gradientes, no entanto, afetam menos suas orientações (LOWE, 2004). Sendo assim, torna-se possível limitar a influência das grandes magnitudes do gradiente utilizando-se um valor de limiar, após a normalização do vetor – foi sugerido por Lowe (2004) a utilização de um limiar de 0,2 e em seguida normalizar, com comprimento unitário, o vetor do descritor novamente, reduzindo assim o efeito da não linearidade na iluminação pela limitação da influência das grandes magnitudes do gradiente. O limiar de 0,2 sugerido por Lowe (2004) foi obtido experimentalmente e este valor pode ser ajustado em função de uma determinada aplicação.

2.2.3 WLD

WLD foi introduzido por Chen et al. (2008) e consiste de um descritor local de texturas baseado na lei de Weber. A lei de Weber faz uma relação entre a mínima variação perceptível (ao ser humano) da intensidade de um dado fenômeno físico com relação à intensidade do mesmo: por exemplo, dado um pixel com intensidade I , haverá um incremento mínimo perceptível de sua intensidade, incremento este dado por ΔI . A Equação 6 representa a fração de Weber (CHEN et al., 2008), onde I é a intensidade de um estímulo, ΔI é a mínima intensidade perceptível e k uma constante.

$$k = \frac{\Delta I}{I}$$

Equação 6

O descritor proposto por Chen et al. (2008) é composto por duas componentes: excitação diferencial (ξ) e orientação (θ). A excitação diferencial (ξ) é uma fração de Weber da somatória das diferenças de intensidade de pixel em uma vizinhança 3x3 com relação à intensidade do pixel central, enquanto a orientação (θ)

é a orientação do próprio pixel central. A Figura 10, adaptada de (CHEN et al., 2009), ilustra o cálculo do WLD.

O filtro f_{00} calcula as diferenças de intensidade dos pixels vizinhos com relação ao pixel central, em uma vizinhança de 3x3, calculando assim v_S^{00} , conforme a Equação 7, onde x_i ($i=0, 1, \dots, p-1$) denotam os p vizinhos do pixel central x_c .

$$v_S^{00} = \sum_{i=0}^{p-1} (x_i - x_c) \quad \text{Equação 7}$$

O filtro f_{01} corresponde ao valor do pixel central, também representado por v_S^{01} . Segundo a lei de Weber, calcula-se a razão de v_S^{00} e v_S^{01} , dada por $G_r(x_c) = \frac{v_S^{00}}{v_S^{01}}$. O diferencial da excitação, $\xi(x_c)$, é calculado tomando-se a função arco-tangente de $G_r(x_c)$, conforme a Equação 8. O uso da função arco-tangente limita a saída $\xi(x_c)$, prevenindo que a mesma aumente ou diminua muito rapidamente quando a entrada se torna muito grande ou pequena.

$$\xi(x_c) = \arctan\left(\frac{\sum_{i=0}^{p-1} (x_i - x_c)}{x_c}\right) \quad \text{Equação 8}$$

A orientação do pixel central x_c é calculada através da orientação do gradiente dos pixels da vizinhança de 4: o filtro f_{10} calcula a diferença dos pixels na vertical (v_S^{10}), enquanto o filtro f_{11} calcula a diferença dos pixels na horizontal (v_S^{11}). A orientação $\theta(x_c)$ é dada pela Equação 9. Conforme pode ser observado na Figura 10 a orientação $\theta(x_c)$ é quantizada, obtendo-se Φ_t .

$$\theta(x_c) = \arctan\left(\frac{v_S^{11}}{v_S^{10}}\right) \quad \text{Equação 9}$$

Por fim é calculado o histograma 2D $\{WLD(\xi_j, \Phi_t)\}$, com $j = 0, 1, \dots, N-1$ e $t = 0, 1, \dots, T-1$, sendo N a dimensão da imagem de entrada e T a quantidade de orientações dominantes utilizada para a quantização de $\theta(x_c)$. A fim de se obter um

descritor com maior discriminabilidade e mais compacto, o referido histograma 2D é compactado em um histograma 1D (CHEN et al., 2009).

Figura 10 – Ilustração do cálculo do WLD.

Fonte: adaptado de (CHEN et al., 2008).

2.2.4 CS-LBP

O algoritmo CS-LBP (do inglês, *Center Symmetric Local Binary Pattern*) proposto por Heikkilä, Pietikäinen e Schmid (2006) combina características do descritor de texturas LBP com SIFT. O descritor de texturas LBP produz histogramas muito grandes, uma vez que o pixel central é comparado com toda a sua vizinhança e, a fim de resolver este problema, no CS-LBP apenas os pares simétricos são comparados (HEIKKILÄ, PIETIKÄINEN; SCHMID, 2009), conforme ilustrado na Figura 11.

Figura 11 – Métodos LBP e CS-LBP para uma vizinhança de 8 pixels.

Fonte: adaptado de (HEIKKILÄ, PIETIKÄINEN; SCHMID, 2009).

Da Figura 11 verifica-se que, para uma vizinhança com 8 pixels, enquanto o método LBP produz 256 padrões diferentes de saída, no CS-LBP são produzidos apenas 16 padrões. A característica, ou valor, gerado pelo método CS-LBP é dado conforme a Equação 10.

$$CS - LBP_{R,N,T}(x, y) = \sum_{i=0}^{(N/2)-1} s(n_i - n_{i+(N/2)})2^i \quad \text{Equação 10}$$

onde n_i e $n_{i+(N/2)}$ são os valores de níveis de cinza de pares simétricos ao centro, com N pixels na vizinhança, igualmente distribuídos em um círculo de raio R , sendo $s(x)$ dado pela Equação 11.

$$s(x) = \begin{cases} 1, & \text{se } x > T, \\ 0, & \text{caso contrário.} \end{cases}$$

Equação 11

O algoritmo CS-LBP utiliza quatro descritores de regiões: *Hessian-Affine*, *Harris-Affine*, *Hessian-Laplace* e *Harris-Laplace* (HEIKKILÄ, PIETIKÄINEN; SCHMID, 2009). Após serem detectadas, as regiões são mapeadas em uma região circular de raio constante a fim de se obter invariância quanto à escala e transformação afim e ainda, caso necessário, a região é rotacionada em relação ao gradiente dominante a fim de se obter invariância quanto à rotação (HEIKKILÄ, PIETIKÄINEN; SCHMID, 2009). Tais regiões possuem 41 x 41 pixels com intensidade normalizada entre o intervalo [0,1].

Para a construção do descritor CS-LBP, estas regiões são divididas em células, obtendo-se assim um histograma de valores CS-LBP para cada célula. A fim de distribuir o peso de cada característica entre *bins* adjacentes é usada uma interpolação bilinear. O histograma da região de interesse é dado então pela concatenação dos histogramas obtidos para cada célula e então este histograma é normalizado para comprimento unitário. Do mesmo modo que ocorre com os descritores SIFT, cada *bin* é limitado em 0,2 e o histograma final é novamente normalizado, gerando assim o descritor final. Considerando, por exemplo, que as regiões sejam divididas como no SIFT, em 4x4 células, o vetor de características gerado pelo CS-LBP terá 256 posições (4 x 4 x 16) contra 128 do SIFT (4 x 4 x 8).

2.2.5 CS-LMP

O descritor CS-LMP (do inglês, *Center Symmetric Local Mapped Pattern*) foi introduzido por Ferraz, Pereira e Gonzaga (2014) e possui características similares ao CS-LBP, porém corrigindo algumas de suas desvantagens. No descritor CS-LBP a diferença entre os pares simétricos em relação ao centro é binária, não capturando diversas nuances que podem ocorrer nas imagens. O valor CS-LMP é dado pela Equação 12.

$$CS - LMP_{R,N,f}(x, y) = \text{round}\left(\frac{\sum_{i=0}^{N/2-1} f(i)P(i)}{\sum_{i=0}^{N/2-1} P(i)}\right)b \quad \text{Equação 12}$$

onde b corresponde ao número de *bins* do histograma menos um, $P(i)$ ao i -ésimo elemento da matriz de pesos P dada pela Equação 13 e $f(i)$ a função de mapeamento sigmoide dada pela Equação 14.

$$P = [p_0 \ p_1 \ p_2 \ \dots \ p_{N/2-1}]^t \quad \text{Equação 13}$$

com $p_i = 2^i$.

$$f(i) = \frac{1}{1 + e^{\frac{-[n_i - n_{i+(N/2)}]}{\beta}}} \quad \text{Equação 14}$$

na qual n_i e $n_{i+(N/2)}$ corresponde à intensidade de nível de cinza dos pares simétricos ao centro, com N pixels igualmente distribuídos em um círculo de raio R e β definindo a inclinação da curva sigmoide. O descritor final é obtido de maneira análoga do CS-LBP, conforme explicado detalhadamente em (FERRAZ; PEREIRA; GONZAGA, 2014) e seu vetor de características possui 256 posições ($4 \times 4 \times 16$).

2.3 Considerações Finais

Neste capítulo foram apresentados os principais marcos históricos no desenvolvimento de algoritmos para o reconhecimento facial ao longo dos anos. O primeiro trabalho de grande relevância leva em consideração as características geométricas das faces (KANADE, 1973). Posteriormente dois trabalhos tiveram grande destaque: as *eigengaces*, que fazem uso da técnica PCA (TURK; PENTLAND, 1991) e as *fisherfaces*, que utilizam a técnica LDA (BELHUMEUR; HESPANHA; KRIEGMAN, 1997).

Em seguida foram demonstrados diversos trabalhos que utilizam descritores clássicos encontrados na literatura, tais como LBP, SIFT e WLD e suas respectivas

formas variantes, para o reconhecimento facial. Os diversos artigos estudados demonstram que o uso de descritores locais baseados em pontos chaves apresentam problemas quando empregados em problemas de reconhecimento facial, uma vez que uma face não é um objeto rígido e apresenta muitas variações pelo envelhecimento, uso de maquiagem e acessórios, expressões e poses faciais, etc (GENG; JIANG, 2009) (WANG et al., 2010) (LI; PARK; JAIN, 2011) (SIMONYAN et al., 2013). Desta forma, muitos trabalhos descrevem regiões, ao invés de pontos chave.

Alguns métodos se baseiam na face total e são ditos métodos holísticos, pois consideram características globais de uma face. Outros métodos, no entanto, se baseiam em determinadas regiões faciais: dentro deste contexto, a região periocular tem despertado muito interesse em diversos autores (MILLER et al., 2010) (PARK et al., 2010), pois apresenta características discriminantes que possibilitam resultados melhores que aqueles métodos que empregam outras regiões faciais (WOODARD et al., 2010) (XU; SAVVIDES, 2012).

Por fim, diversos descritores locais clássicos encontrados na literatura foram apresentados neste capítulo (LBP, SIFT e WLD), bem como algumas de suas variações (CS-LBP e CS-LMP). Estes descritores foram utilizados tanto como inspiração para o desenvolvimento dos métodos propostos nesta tese, como para a comparação dos resultados obtidos nos diversos experimentos realizados.

3 MÉTODOS PROPOSTOS

O objetivo principal deste trabalho é o desenvolvimento de um novo método de reconhecimento facial em ambientes não controlados. Para tanto foram desenvolvidos dois filtros de realce do conteúdo de alta frequência: o HBWLF (do inglês, *high-Boost Weber local filter*) e o MHBWLF (do inglês, *multiscale high-Boost Weber local filter*). Baseado no MHBWLF foi desenvolvido ainda um novo descritor, chamado de HBWD (do inglês, *high-Boost Weber descriptor*). Para o reconhecimento facial foi desenvolvido ainda um método de localização precisa dos olhos e da ponta do nariz de uma face, utilizando o MHBWLF. Por fim, é apresentado o método de reconhecimento facial em ambientes não controlados pela aplicação do HBWD densamente na região periocular.

3.1 High-Boost Weber Local Filter (HBWLF)

Neste trabalho foi desenvolvido o HBWLF com o objetivo de realçar os detalhes de uma imagem, realçando o conteúdo de alta frequência da mesma sem, no entanto, eliminar o seu conteúdo de baixa frequência (AFFONSO; RODRIGUES; PAIVA, 2018).

Para tanto, o desenvolvimento do HBWLF foi inspirado no filtro da excitação diferencial presente no WLD proposto por CHEN et al. (2008) descrito anteriormente na seção 2.2.3. A excitação diferencial do WLD realça as altas frequências pelo cálculo da somatória da diferença de um pixel central para sua vizinhança em uma grade retangular, o que o torna pouco robusto quanto ao ruído presente no pixel central. A fim de garantir uma maior robustez quanto ao ruído, realçando as altas frequências sem eliminar totalmente as componentes de baixa, o HBWLF calcula a somatória da diferença do valor médio de uma região ao redor do pixel central para a sua vizinhança em uma grade retangular.

O primeiro passo do HBWLF é o cálculo de uma imagem filtrada pela média da vizinhança, o que é feito por meio de um filtro de convolução de tamanho $N \times N$, como apresentado na Equação 15. Para um dado pixel x , o seu valor médio $\bar{x}$ em

uma vizinhança $N \times N$ também pode ser calculado conforme expresso na Equação 16, onde x_i denota a intensidade do i -ésimo vizinho de x .

$$h = \frac{1}{N^2} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}_{N \times N} \quad \text{Equação 15}$$

$$\bar{x} = \frac{1}{N^2} \sum_{i=0}^{N^2-1} x_i \quad \text{Equação 16}$$

A diferença de intensidade ΔI entre a vizinhança e o valor médio $\bar{x}$ é calculada conforme a Equação 17, onde x_c corresponde ao pixel central e x_i aos 8 vizinhos da região 3×3 , conforme ilustrado na Figura 12. A fração de Weber é dada conforme a Equação 18, onde α é uma constante real, positiva e de pequeno valor, utilizada apenas para evitar a divisão por zero.

Figura 12 – Vizinhança 3×3 para cálculo do HBWLF.

x_0	x_1	x_2
x_3	x_c	x_4
x_5	x_6	x_7

Fonte: adaptado de (AFFONSO; RODRIGUES; PAIVA, 2018).

$$\Delta I(x_c) = \sum_{i=0}^7 (x_i - \bar{x}) \quad \text{Equação 17}$$

$$k(x_c) = \frac{\sum_{i=0}^7 (x_i - \bar{x})}{\bar{x} + \alpha} \quad \text{Equação 18}$$

Por fim, a fração de Weber $k(x_c)$ é mapeada por meio da função sigmoide, conforme a Equação 19, onde β é uma constante real, positiva, utilizada para o controle da inclinação da curva sigmoide. Enquanto o WLD utiliza a função arco-tangente para o cálculo do valor final filtrado, a função sigmoide no HBWLF permite um melhor ajuste de sua saída, prevenindo que a mesma aumente ou diminua muito rapidamente quando a fração de Weber $k(x_c)$ se torna muito grande ou pequena.

$$f(x_c) = \frac{1}{1 + e^{-\frac{k(x_c)}{\beta}}}$$

Equação 19

Desta forma, para se filtrar uma região de interesse deve-se aplicar o procedimento explicado nesta seção para todos os pixels desta região e assim obter uma imagem filtrada pelo HBWLF. A Figura 13 ilustra alguns exemplos de imagens faciais filtradas pelo HBWLF.

Figura 13 – Exemplos de imagens filtradas: (a) imagens originais; (b) imagens filtradas pelo HBWLF, com $N \times N = 7 \times 7$ e $\beta = 0,2$.

Fonte: adaptado de (AFFONSO; RODRIGUES; PAIVA, 2018).

3.2 Multiscale High-Boost Weber Local Filter (MHBWLF)

O HBWLF é calculado usando-se uma grade regular de tamanho 3×3 , conforme ilustrado na Figura 12. A fim de produzir um filtro mais flexível para diversas aplicações, o MHBWLF foi desenvolvido considerando-se uma região circular, de raio R e com NV vizinhos igualmente distribuídos na vizinhança de x_c . A Figura 14 ilustra uma vizinhança circular com $NV = 8$ (8 pixels igualmente distribuídos ao redor de x_c) (AFFONSO; RODRIGUES, 2017).

Utilizando-se esta nova vizinhança, a diferença de intensidade ΔI no filtro MHBWLF é dada conforme a Equação 20. A fração de Weber $k(x_c)$, por sua vez, é dada pela Equação 21.

Figura 14 – Vizinhança circular com $NV = 8$.

Fonte: autoria própria.

$$\Delta I(x_c)_{NV,R} = \sum_{i=0}^{NV-1} (x_i - \bar{x}) \quad \text{Equação 20}$$

$$k(x_c)_{NV,R} = \frac{\sum_{i=0}^{NV-1} (x_i - \bar{x})}{\bar{x} + \alpha} \quad \text{Equação 21}$$

Do mesmo modo que no HBWLF, o MHBWLF também faz uso da função sigmoide (Equação 19) para filtrar a fração de Weber $k(x_c)$ dada pela Equação 21 e assim produzir o resultado final de saída. O filtro MHBWLF possui assim quatro parâmetros de ajustes: tamanho do filtro ($N \times N$), inclinação da curva sigmoide (β), número de pixels (NV) e raio (R) na vizinhança circular. No caso particular em que $NV = 8$ e $R = 1$, tem-se que o filtro MHBWLF torna-se o filtro HBWLF e, desta forma, pode-se dizer que o filtro HBWLF é um caso particular do filtro MHBWLF.

3.3 Método para Localização Precisa dos Olhos

Nesta seção descreve-se o método desenvolvido para a localização precisa dos centros dos olhos em ambientes não controlados. O método proposto é do tipo passivo, ou seja, não utiliza nenhum aparato específico de hardware, mas apenas imagens capturadas por câmeras em condições normais do dia a dia; extrai características locais da região periocular (região ao redor dos olhos da face); realça os detalhes da região periocular pelo uso do MHBWLF; utiliza o filtro MOSSE (BOLME et al., 2010) para o aprendizado das características realçadas pelo MHBWLF. A Figura 15 apresenta o diagrama geral do método proposto, o qual é dividido em duas fases distintas: fase de treinamento e fase de testes. A fase de

treinamento recebe um conjunto de imagens, com suas respectivas posições dos centros dos olhos, a fim de treinar um filtro MOSSE para localizar os centros dos olhos em uma imagem de entrada durante a fase de teste.

Na Figura 15 verifica-se que o método proposto possui os seguintes principais blocos, os quais serão descritos em detalhes nas seções a seguir: pré-processamento, MHBWLF (já descrito na seção 3.2), filtro MOSSE e avaliação do pico de correlação.

Figura 15 – Diagrama geral do método proposto para localização dos olhos.

Fonte: adaptado de (AFFONSO; RODRIGUES; PAIVA, 2018).

3.3.1 Bloco de Pré-Processamento

A Figura 16 ilustra o fluxograma do bloco de pré-processamento. A primeira etapa do bloco de pré-processamento converte a imagem de entrada colorida em uma imagem em níveis de cinza, se necessário. Na segunda etapa o bloco faz a detecção facial por meio do algoritmo de Viola-Jones (VIOLA; JONES, 2001): caso o detector facial encontre mais de uma face, o bloco de pré-processamento irá

considerar apenas a face central e ignorar as demais faces. Posteriormente, na terceira etapa, a face detectada tem o seu tamanho ajustado para 120×120 pixels por meio de interpolação bicúbica e a região periocular é recortada empiricamente, conforme recorte ilustrado na Figura 17.

Figura 16 – Fluxograma do bloco de pré-processamento.

Fonte: adaptado de (AFFONSO; RODRIGUES; PAIVA, 2018).

Figura 17 – Face ajustada para 120×120 pixels e recorte da região de interesse.

Fonte: adaptado de (AFFONSO; RODRIGUES; PAIVA, 2018).

3.3.2 Filtro Mosse

Após o pré-processamento a ROI (do inglês, *region of interest*) selecionada é filtrada pelo MHBWLF, descrito na seção 3.2, e esta imagem filtrada é então utilizada como imagem de entrada para o filtro MOSSE. O filtro MOSSE produz um filtro ASEF-like, contudo, requer menos imagens de treinamento que o filtro ASEF – todo o detalhamento do filtro MOSSE pode ser encontrado em (BOLME et al., 2010), no entanto, o desenvolvimento principal será aqui resumido.

As imagens de entrada do filtro MOSSE, nas fases de treinamento e execução, são obtidas pela saída do MHBWLF e são chamadas de f_i . Na fase de treinamento, as posições dos olhos de cada imagem são utilizadas para gerar as imagens g_i : estas imagens são consideradas como saídas desejadas e são

calculadas pelo bloco “Saídas Sintetizadas”, apresentado anteriormente na Figura 15, e correspondem ao cálculo de uma gaussiana bidimensional, conforme a Equação 22.

$$g_i(x,y) = e^{-\frac{(x-x_i)^2+(y-y_i)^2}{\sigma^2}} \quad \text{Equação 22}$$

onde i denota a i -ésima imagem do conjunto de imagens de treinamento e σ define o raio do pico. O problema de aprendizado do filtro MOSSE pode ser resumido no aprendizado do filtro h , definido na Equação 23.

$$g_i = h \star f_i \quad \text{Equação 23}$$

No entanto, esta tarefa é muito difícil de ser realizada no domínio do espaço e, portanto, é realizada no domínio da frequência. Usando letras maiúsculas para indicar as imagens no domínio da frequência, a Equação 23 pode ser escrita conforme a Equação 24.

$$G_i = F_i \odot H^* \quad \text{Equação 24}$$

no qual $\odot$ representa uma multiplicação de elemento por elemento e $*$ corresponde ao complexo conjugado. A fim de encontrar um filtro que melhor mapeia o conjunto de imagens de treinamento e suas respectivas saídas, Bolme et al. (2010) apresentaram um filtro H^* que minimiza a soma do erro quadrático entre a saída atual ($F_i \odot H^*$) e a saída desejada (G_i), problema este que pode ser resumido conforme $\min_{H^*} \sum_i |F_i \odot H^* - G_i|^2$. E, resolvendo o problema apresentado, o filtro MOSSE é descrito conforme a Equação 25.

$$H^* = \frac{\sum_i G_i \odot F_i^*}{\sum_i F_i \odot F_i^* + \epsilon} \quad \text{Equação 25}$$

em que ϵ é uma constante real, positiva, usada como um parâmetro de regularização. O filtro treinado H^* é utilizado, durante a fase de teste, para gerar uma imagem de saída G_o , a partir de uma imagem de teste F_q , conforme a Equação 26

$$Go = Fq \odot H^*$$

Equação 26

onde Fq representa a imagem de teste já filtrada pelo MHBWLF, no domínio da frequência. A imagem Go está no domínio da frequência e é convertida para o domínio do espaço por meio da transformada inversa de Fourier, gerando assim a imagem g_o no domínio do espaço. Ressalta-se aqui que o filtro MOSSE foi aplicado individualmente para cada olho e, portanto, a solução apresentada na Figura 15 possui dois filtros H^* , sendo um para o olho direito e outro para o olho esquerdo.

3.3.3 Avaliação do Pico de Correlação

Este bloco realiza uma avaliação do pico de correlação de saída presente nas imagens g_o (uma imagem para o olho direito e outra para o olho esquerdo), gerando as posições (x, y) de cada olho. Para cada imagem g_o , o bloco faz uma busca pelos picos de correlação maiores que o limiar T_H e a posição final do centro do olho será dada pela média dos pontos que tiverem um pico de correlação maior que T_H . Se nenhum pixel possuir um pico de correlação maior que T_H , então a posição (x, y) do referido olho será dada pela posição média de todas as imagens usadas durante a fase de treinamento.

3.3.4 Localização de Outros Pontos Faciais

O método de localização dos centros dos olhos, descrito na seção 3.3, também pode ser aplicado para a localização de outros pontos na face como, por exemplo, os cantos dos olhos, a ponta do nariz, etc. De fato, neste trabalho este método também foi utilizado para a localização da ponta do nariz durante o método de reconhecimento facial que será explicado na seção 3.5.

Para localizar outro ponto, diferente dos centros dos olhos, basta apenas trocar as posições dos centros dos olhos (Figura 15) pelas posições deste novo

ponto, gerando assim as imagens desejadas para o filtro MOSSE pelo bloco “Saídas Sintetizadas”. Um novo treinamento do filtro MOSSE deve ser realizado, a fim de gerar um novo filtro H^* capaz de localizar este novo ponto desejado. Cabe ainda ressaltar que dependendo da localização do novo ponto desejado na face pode-se fazer necessário um ajuste na ROI indicada na Figura 17.

3.4 High-Boost Weber Descriptor (HBWD)

Neste trabalho é introduzido o HBWD, um descritor local baseado no MHBWLF descrito na seção 3.2. O desenvolvimento do HBWD foi inspirado nos conhecidos descritores SIFT (LOWE, 2004) e CS-LMP (FERRAZ; PEREIRA; GONZAGA, 2014), reunindo algumas de suas características, conforme descrição apresentada a seguir.

Primeiramente deve-se extrair um ponto de interesse, o qual deseja-se descrevê-lo usando o HBWD. Esta tarefa pode ser realizada pela utilização de diversos detectores de regiões de interesse, tais como: *Hessian-Affine*, *Harris-Affine*, *Hessian-Laplace* ou *dense sampling*. Conforme será descrito na seção 3.5, para o reconhecimento facial foi utilizada neste trabalho a técnica *dense sampling* de detecção de pontos de interesse.

O primeiro passo do HBWD consiste em filtrar a imagem de entrada utilizando o já apresentado MHBWLF. Em seguida, monta-se uma grade regular com $N_c \times N_c$ células, centrada no pixel de interesse x_c , onde cada célula contém $N_p \times N_p$ pixels, conforme ilustrado na Figura 18. Deve-se verificar se todos os pixels da grade estão localizados dentro da imagem filtrada inicialmente, caso contrário não é possível calcular o descritor para o ponto de interesse x_c .

Do mesmo modo que ocorre no descritor SIFT, a grade regular montada é ponderada, pixel a pixel, por uma gaussiana centrada em x_c : o objetivo desta gaussiana é de ponderar a imagem filtrada, fazendo com que os pixels próximos ao da região central tenham maior influência no descritor final que os pixels das bordas da grade. A Figura 19 ilustra uma grade regular recortada, com 4×4 células, onde cada célula contém 12×12 pixels, gerando assim uma grade de 48×48 pixels. O filtro MHBWLF foi configurado com $N \times N = 13 \times 13$, $\beta = 0,3$, $R = 1$ e $NV = 8$.

Figura 18 – Grade regular com $N_C \times N_C$ células, ao redor do pixel central x_C . Cada célula contém $N_p \times N_p$ pixels.

Fonte: autoria própria.

Figura 19 – Exemplo de grade regular recortada: (a) imagem de entrada; (b) imagem filtrada pelo MHBWLF; (c) gaussiana de ponderação; (d) imagem filtrada ponderada.

Fonte: autoria própria.

Posteriormente monta-se um histograma com o número de *bins* igual a N_{bins} para cada uma das $N_C \times N_C$ células. O histograma final é obtido pela concatenação dos histogramas individuais de cada célula e, sendo assim, sua dimensão será dada por $N_C \times N_C \times N_{bins}$. A Figura 20 ilustra a montagem do histograma final do HBWD para o caso de $N_C \times N_C = 3 \times 3$ células, com cada célula contendo $N_p \times N_p = 9 \times 9$ pixels, e $N_{bins} = 5$, resultando em um histograma final com 45 elementos. Os *bins* adjacentes do histograma final são interpolados, distribuindo assim o peso de suas características na composição final do histograma.

Do mesmo modo que no SIFT (LOWE, 2004), o vetor do descritor gerado no passo anterior é normalizado para o comprimento unitário, pela divisão de cada um de seus elementos pelo seu módulo. E ainda, cada elemento do vetor final é

comparado com um limiar (T_{H_IL}) a fim de se limitar os elementos com grande magnitude, reduzindo os efeitos de não linearidades na iluminação. Por fim, o vetor é novamente normalizado para o comprimento unitário e será então utilizado como descritor do ponto de interesse x_C .

Figura 20 – Exemplo de montagem do histograma final do HBWD.

Fonte: autoria própria.

Cabe ressaltar que o HBWD não possui nenhum artifício que o torne invariante quanto à rotação como ocorre, por exemplo, com o SIFT (LOWE, 2004), que calcula a orientação dominante de um ponto chave, tornando-o invariante quanto à rotação. Esta propriedade poderia ser adicionada ao HBWD, no entanto, uma vez que o mesmo foi desenvolvido especialmente para o reconhecimento facial, em faces já alinhadas quando de sua aplicação, isto se torna desnecessário, conforme será demonstrado no capítulo 5 durante a análise dos resultados experimentais. E ainda, incluir o cálculo da orientação do pixel poderia até mesmo diminuir a robustez do HBWD quando ao ruído ou variações na iluminação, prejudicando uma propriedade muito importante do descritor proposto na presente aplicação.

3.5 Método para Reconhecimento Facial

A Figura 21 apresenta o diagrama geral do método proposto para reconhecimento facial em ambientes não controlados baseados no problema do pair matching (quando se compara um par de imagens para verificar se correspondem à uma mesma pessoa ou não). O método proposto recebe duas imagens de entrada (imagens 1 e 2) e gera um histograma final de saída, o qual é utilizado para a classificação das imagens como sendo *matched* (quando as imagens de entrada 1 e 2 são da mesma pessoa) ou *mismatched* (quando as imagens de entrada 1 e 2 são de pessoas diferentes). As seções a seguir apresentam uma descrição das principais etapas do método proposto.

3.5.1 Imagens de Entrada

O método proposto recebe duas imagens de entrada, imagens 1 e 2, que podem ser coloridas ou em tons de cinza. O método irá detectar as faces de cada imagem e considerará como faces de interesse apenas as faces centrais de cada imagem, ignorando as demais faces, caso existam. Ressalta-se que a detecção facial ocorre na etapa de localização dos pontos de interesse, conforme descrito na seção 3.7.

3.5.2 Detecção de Pontos de Interesse

O método proposto realiza a localização de três pontos de interesse em cada face: centro do olho direito, centro do olho esquerdo e ponta do nariz. Para tanto, foi empregado o método descrito anteriormente na seção 3.3.

Figura 21 – Diagrama geral do método proposto para reconhecimento facial em ambientes não controlados, visando o problema do *pair matching*.

Fonte: autoria própria.

3.5.3 Aplicação do HBWD

Antes da aplicação do HBWD a imagem facial é alinhada espacialmente, usando o método de interpolação do vizinho mais próximo (em inglês, *nearest-neighbor interpolation*): para tanto utiliza-se apenas as coordenadas dos olhos como referência e a imagem alinhada deve conter os dois olhos alinhados horizontalmente, conforme ilustrado na Figura 22.

Figura 22 – Alinhamento espacial de uma imagem facial.

Fonte: autoria própria.

Em seguida é realizado o recorte da ROI: neste trabalho utiliza-se como região de interesse para reconhecimento facial apenas a região periocular, umas das regiões mais discriminantes de uma face (MILLER et al., 2010). Cabe ressaltar que não há uma definição formal para a região periocular de uma face, de modo que a mesma pode incluir ou excluir as sobrancelhas ou outros componentes da face (PARK et. al., 2010). Neste trabalho a região da boca foi desconsiderada a fim de reduzir a influência de suas variações nas diversas expressões faciais (exemplo: pessoa sorrindo, falando, gritando, chorando, etc.). Sendo as coordenadas obtidas pela detecção dos pontos de interesse dadas por:

- ✓ (x_R, y_R) : centro do olho direito;
- ✓ (x_L, y_L) : centro do olho esquerdo;
- ✓ (x_N, y_N) : ponta do nariz.

Após o alinhamento espacial, têm-se as novas coordenadas:

- ✓ (x_R^A, y_R^A) : centro do olho direito na imagem alinhada;
- ✓ (x_L^A, y_L^A) : centro do olho esquerdo na imagem alinhada;
- ✓ (x_N^A, y_N^A) : ponta do nariz na imagem alinhada.

O canto superior esquerdo da ROI possui as suas coordenadas dadas por $(X_0 - d_X, Y_0)$, enquanto as coordenadas do canto inferior direito são $(X_0 + d_X, Y_0 + d_Y)$, com X_0 , d_X , Y_0 e d_Y dados, respectivamente, pela Equação 27, Equação 28, Equação 29 e Equação 30. A Figura 23 ilustra alguns exemplos de regiões de interesse extraídas em faces da base LFW, onde pode-se notar uma grande diversidade de ângulos e poses faciais, no entanto, todas elas com a extração da região periorcular como região de interesse.

$$X_0 = \frac{(x_R^A + x_L^A)}{2} \quad \text{Equação 27}$$

$$d_X = x_N^A - X_0 \quad \text{Equação 28}$$

$$Y_0 = y_R^A - \frac{(y_N^A - y_R^A)}{2} \quad \text{Equação 29}$$

$$d_Y = y_L^A + \frac{(y_L^A - y_N^A)}{2} - Y_0 \quad \text{Equação 30}$$

Neste trabalho foi utilizada a técnica de *dense sampling* para a detecção dos pontos de interesse a serem descritos pelo HBWD, ou seja, todos os pixels pertencentes à ROI de uma determinada face são descritos pelo HBWD. Esta abordagem traz o inconveniente de produzir muitos vetores de características da ROI a ser descrita, porém apresenta a vantagem de descrever a ROI por completo.

Desta forma, cada ROI- i detectada será densamente descrita por meio do HBWD. Se uma ROI- i possui $NROI_i$ pixels, logo a ROI- i será representada por $NROI_i$ vetores Hi , onde a dimensão de cada vetor Hi será dada pela configuração do HBWD escolhida ($NHBWD = N_C \times N_C \times N_{bins}$).

Figura 23 – Exemplos de regiões de interesse extraídas em faces da base LFW.

Fonte: autoria própria.

3.5.4 Redução Dimensional por Agrupamento

Conforme explicado anteriormente, cada ROI- i será representada por $NROI_i$ vetores H_i , de dimensão $NHBWD$, conforme ilustrado na Figura 24.

Figura 24 – Dense sampling na ROI usando o HBWD.

Fonte: autoria própria.

Inspirado no método de BoF (do inglês, *Bag-of-Features*) empregado por Ferraz et al. (2014), os $NROI_i$ vetores H_i de cada ROI- i são agrupados pelo método de k-Means++ (ARTHUR; VASSILVITSKII, 2007), gerando $k_{clusters}$ vetores base

para cada ROI-i, conforme ilustrado na Figura 25, onde o índice i corresponde ao mesmo índice da referida ROI-i (ROI-1 ou ROI-2).

Figura 25 – Agrupamento dos kclusters para cada ROI-i.

Fonte: autoria própria.

Antes de se proceder ao mapeamento das regiões de interesse faz-se necessário ordenar os clusters. Para tanto foi utilizada a função definida na Equação 31: calcula-se o valor de C_k para $k = 1$ até $k_{clusters}$ e os clusters são ordenados pela ordem crescente de C_k . Este procedimento é realizado tanto para a ROI-1 quanto para a ROI-2. Todas as operações a seguir são realizadas nos clusters já ordenados.

$$C_k = \sum_{j=1}^{N_C \times N_C \times N_{bins}} Cluster_k(j) \times j \quad \text{Equação 31}$$

Os $NROI_1$ vetores $H1$ da ROI-1 são mapeados nos $k_{clusters}$ vetores bases da ROI-2: a seguir este mapeamento será explicado em detalhes e o mesmo raciocínio pode ser aplicado para o caso inverso, ou seja, quando os $NROI_2$ vetores $H2$ da ROI-2 são mapeados nos $k_{clusters}$ vetores bases da ROI-1.

O mapeamento dos $NROI_1$ vetores $H1$ da ROI-1 nos clusters (vetores base) da ROI-2 consiste na montagem de um histograma com $k_{clusters}$ bins. Para um j-ésimo vetor $H1$, com j variando de 1 até $NROI_1$, calcula-se a distância euclidiana deste vetor j para cada um dos $k_{clusters}$ vetores base da ROI-2 e soma-se uma unidade ao bin do histograma que corresponder à menor distância euclidiana, montando-se assim o histograma de mapeamento ilustrado na Figura 26.

Desta forma, a ROI-1 que possui $NROI_1$ vetores $H1$, onde cada vetor tem dimensão $NHBWD$, é agora representada por um histograma com $k_{clusters}$ bins. O mesmo procedimento é aplicado na ROI-2, e assim é gerado o histograma final, com

$2 \times k_{clusters}$ bins, para a comparação das faces de entrada (faces 1 e 2), conforme ilustrado na Figura 27.

Figura 26 – Mapeamento dos vetores da ROI-1 nos clusters da ROI-2.

Fonte: autoria própria.

Figura 27 – Histograma final para comparação das faces 1 e 2.

Fonte: autoria própria.

3.5.5 Classificação usando Máquinas de Vetores de Suporte

Para o classificador foi selecionada uma máquina de vetores de suporte, SVM, a qual constitui uma técnica de aprendizado que vem sendo muito utilizada pela comunidade científica em diversos trabalhos de visão computacional, por exemplo. As SVMs são embasadas pela teoria de aprendizado estatístico, gerando classificadores com boa generalização, ou seja, que possuem uma boa capacidade de prever corretamente a classe de novos dados de entrada, desde que estes estejam no mesmo domínio em que foi realizado o aprendizado.

A SVM é utilizada em problemas onde deseja-se classificar uma amostra dentre duas classes possíveis, o que corresponde exatamente ao caso em estudo: classificar um histograma de entrada indicando se as duas faces de entrada são da

mesma pessoa (saída = *matched*) ou de pessoas distintas (saída = *mismatched*). O classificador SVM irá encontrar o melhor hiperplano que separa, dentro de uma certa margem, os dados das duas classes. Neste trabalho foi escolhido o *kernel* linear para o SVM.

3.6 Considerações Finais

Neste capítulo foram introduzidos dois novos filtros, o HBWLF e o MHBWLF, ambos inspirados no filtro da excitação diferencial do WLD. Estes dois filtros enfatizam o conteúdo de alta frequência das imagens sem, no entanto, eliminar o de baixa, realçando assim os detalhes da região periocular de uma face. O HBWLF calcula a imagem filtrada localmente, utilizando uma grade regular de 3×3 pixels, enquanto o MHBWLF faz uso de uma região circular de raio R com NV vizinhos distribuídos uniformemente na vizinhança circular.

Enquanto no WLD considera-se o pixel central da região para o cálculo da imagem filtrada, nos filtros desenvolvidos utiliza-se o valor médio da região ao redor do pixel central, o que lhe confere uma maior robustez quanto ao ruído presente no pixel central. E ainda, o WLD utiliza a função arco-tangente para o cálculo do valor final filtrado, enquanto o HBWLF e o MHBWLF fazem uso da função sigmoide, que permite um melhor ajuste de sua saída.

O desenvolvimento destes filtros possibilitou a criação de um novo método para a localização, com grande precisão, de diversos pontos de interesse em uma face, especialmente dos centros dos olhos e da ponta do nariz, em imagens faciais obtidas em ambientes não controlados.

Foi introduzido também neste capítulo um novo descritor local, o HBWD, o qual utiliza a imagem filtrada pelo MHBWLF para descrever um determinado pixel localmente. Este descritor foi utilizado no método proposto para o reconhecimento facial, baseado no problema do *pair matching*: o método recebe duas imagens faciais de entrada e tem como saída a informação se tais imagens são de uma mesma pessoa ou não. No método proposto o HBWD foi aplicado densamente apenas na região periocular das faces e uma redução da dimensão dos dados foi

obtida pelo uso da técnica de agrupamento por meio de k-Means++ e um mapeamento inspirado no uso de BoF.

4 MATERIAIS E MÉTODOS

Neste capítulo são apresentados todos os materiais e métodos utilizados nesta tese de doutorado para a realização dos experimentos descritos no capítulo 5. Há uma descrição das bases utilizadas, tanto da aplicação em problemas de localização dos centros dos olhos, quanto no reconhecimento facial propriamente dito. E ainda são descritos os métodos utilizados nos experimentos realizados nesta tese.

4.1 Bases de Dados

Este trabalho de doutorado foi desenvolvido com ênfase em ambientes não controlados, onde as imagens faciais são obtidas nas mais diversas condições do mundo real (por exemplo, imagens em estúdios, praias, lojas, eventos esportivos, entre outras), apresentando grande variação na iluminação, nas poses e expressões faciais, nos objetos de fundo (*background*), na raça e etnia, na idade, no gênero, no vestuário, no estilo de cabelo, nos parâmetros da câmera, entre outras variáveis.

Esta seção descreve as bases de dados utilizadas a fim de cumprir tais requisitos e melhor avaliar os métodos propostos. A Figura 28, retirada da base LFW, ilustra a convenção dos eixos x e y utilizada ao longo desta tese, onde se verifica que a origem $(1,1)$ localiza-se no canto superior esquerdo da imagem.

Figura 28 – Convenção dos eixos x e y adotados neste trabalho.

Fonte: adaptado de imagem da base LFW.

4.1.1 *Labeled Faces in the Wild (LFW)*

A base de dados LFW (Labeled Faces in the Wild, 2015) foi desenvolvida para ser uma base específica para o teste de algoritmos de reconhecimento facial (HUANG et al., 2007). Suas imagens faciais apresentam grandes variações na pose, na iluminação, nas expressões faciais, na idade, no gênero, na raça, no uso de acessórios e maquiagem, no *background*, nas oclusões e em diversos parâmetros da câmera. Esta base foi concebida para o estudo do problema do *pair matching*, que consiste em comparar duas fotos de faces e verificar se estas são da mesma pessoa ou não. Além disto, cada par de imagens no conjunto de teste nunca foi utilizado anteriormente no conjunto de treinamento, caracterizando assim o paradigma do *matching* de imagens nunca vistas anteriormente.

A base LFW contém 13233 imagens, as quais possuem uma ou mais faces, onde apenas a face central deve ser considerada, e as demais faces devem ser ignoradas ou tratadas como *background* da imagem da face central. A base contém imagens faciais de 5749 pessoas diferentes, das quais 1680 pessoas têm duas ou mais imagens faciais na base e as demais 4069 pessoas possuem apenas uma única imagem facial na base. Todas as imagens estão no formato JPEG, com tamanho de 250x250 pixels. Apenas algumas imagens estão em escala de cinza, enquanto as demais são coloridas.

Todas as imagens faciais da base LFW foram previamente detectadas pelo detector facial Viola-Jones (VIOLA; JONES, 2001) aplicado em imagens extraídas da internet. Os casos de falsos positivos extraídos do detector facial foram removidos manualmente, assim como também foram removidas as imagens em duplicidade. As imagens resultantes são todas salvas com informações do nome da pessoa identificada pela face central. A região da face detectada é expandida por um fator igual a 2,2 em cada dimensão e a imagem final é recortada para o tamanho final de 250x250 pixels. Além das imagens originais é possível também fazer o download de imagens previamente alinhadas (HUANG; JAIN; LEARNED-MILLER, 2007), (HUANG et al., 2012).

A Figura 29 ilustra nove imagens diferentes da atriz Jennifer Aniston, contidas na base LFW, notando-se claramente uma grande variação de poses e expressões faciais, na iluminação, nos objetos de fundo, no estilo do cabelo, etc. A Figura 30

apresenta a detecção facial de todas as imagens da Figura 29 por meio do detector facial de Viola-Jones (VIOLA; JONES, 2001).

Figura 29 – Exemplos de imagens da atriz Jennifer Aniston contidas na base LFW.

Fonte: autoria própria.

Figura 30 – Detecção facial das imagens da Figura 29 usando o detector de Viola-Jones.

Fonte: autoria própria.

A Figura 31 ilustra outros exemplos de imagens extraídas da base LFW, onde pode-se verificar facilmente a grande variedade de situações em que as fotos foram

retiradas, demonstrando assim que a base LFW é adequada para a avaliação de algoritmos de reconhecimento facial em ambientes não controlados.

Figura 31 – Exemplos de outras imagens da base LFW.

Fonte: autoria própria.

4.1.2 *Fine-Grained* LFW (FGLFW)

A base FGLFW (do inglês, *Fine-Grained Labeled Faces in the Wild*) (Fine-Grained Labeled Faces in the Wild, 2017) utiliza as mesmas imagens da base LFW, no entanto, esta base muda os pares de comparação para os testes de algoritmos de reconhecimento facial baseados no problema do *pair matching*. Após Deng et al. (2017) investigarem a base LFW eles identificaram que os pares negativos (pares de imagens de pessoas diferentes) são relativamente fáceis de serem classificados como negativos, tornando assim mais fácil o aprendizado: isto ocorre porque os pares de testes da base LFW foram gerados randomicamente, sem levar nenhum critério de semelhança em consideração. Desta forma, Deng et al. (2017) propuseram novos pares de teste, onde os pares negativos são compostos de imagens de pessoas que apresentam uma maior semelhança, trazendo assim um maior desafio para o teste dos algoritmos de reconhecimento facial baseados no problema do *pair matching*.

Figura 32 – Exemplos de pares negativos (pessoas diferentes) da base LFW.

Fonte: extraído de (Fine-Grained Labeled Faces in the Wild, 2017).

Figura 33 – Exemplos de pares negativos (pessoas diferentes) listados na FGLFW.

Fonte: extraído de (Fine-Grained Labeled Faces in the Wild, 2017).

A Figura 32 ilustra vários pares negativos de teste, ou seja, pares que não são da mesma pessoa, extraídos da base LFW: de fato, verifica-se que vários destes pares são de pessoas com pouca semelhança, o que acaba por facilitar a tarefa de classificação. Por outro lado, a Figura 33 apresenta alguns pares negativos listados na base FGLFW, onde nota-se claramente uma maior semelhança dentro dos pares, o que dificulta a tarefa de classificar corretamente tais pares como negativos. Cabe ainda ressaltar que todas as imagens são da base LFW, no entanto, a base FGLFW apenas muda os pares de teste a fim de apresentar um conjunto de testes com maior dificuldade.

4.1.3 BioID

A base BioID (BioID, 2017) foi desenvolvida para a avaliação de diversos métodos de detecção facial, incluindo o problema de localização dos olhos, pois nesta base estão disponíveis as coordenadas dos centros dos olhos para todas as suas imagens faciais.

As imagens contidas na base BioID foram gravadas dando uma grande ênfase nas condições do mundo real, possuindo assim imagens com variações na iluminação, nos objetos de fundo, nas poses e expressões faciais, etc. A Figura 34 ilustra diversas imagens da base BioID, todas elas para uma mesma pessoa.

A base BioID contém imagens em tons de cinza com resolução de 384x286 pixels. As imagens foram tiradas de vistas frontais de 23 pessoas diferentes, com variações em poses e expressões faciais, iluminação e tamanho das faces.

4.1.4 Localização dos Centros dos Olhos: Ground-Truth

O método de localização dos centros dos olhos proposto e descrito no capítulo 3 necessita das coordenadas dos centros dos olhos das faces (*ground-truth*), tanto para o treinamento, quanto para a avaliação do seu desempenho para um dado conjunto de imagens.

Figura 34 – Exemplos de imagens da mesma pessoa da base BioID.

Fonte: autoria própria.

Como explicado anteriormente, a base LFW não apresenta as coordenadas dos centros dos olhos de suas imagens e, por isto, foi utilizado um conjunto de 2000 imagens proposto por Tan et al. (2009), com as coordenadas dos centros dos olhos devidamente anotadas. Tan et al. (2009) selecionaram randomicamente 2000 imagens da base LFW e anotaram manualmente as suas respectivas coordenadas dos centros dos olhos, criando assim um conjunto de 2000 imagens da base LFW com seu respectivo *ground-truth* para serem usados em avaliação e treinamento de algoritmos de localização dos centros dos olhos em ambientes não controlados. E ainda, Tan et al. (2009) dividiu este conjunto em dois grupos de 1000 imagens cada, onde um grupo deve ser utilizado para treinamento e o outro grupo para teste, ressaltando que não há sobreposição de imagens entre os grupos de treinamento e teste.

A base BioID, por sua vez, já possui as coordenadas dos centros dos olhos para todas as suas imagens e, portanto, pode ser utilizada diretamente para treinamento e avaliação do método proposto para a localização dos centros dos olhos.

4.1.5 Reconhecimento Facial: *Pair Matching*

O problema do *pair matching* corresponde à classificação de duas imagens, verificando se estas imagens são de uma mesma pessoa (*matched*) ou de pessoas distintas (*mismatched*). Este problema também é conhecido na literatura como *matching* de 1:1, indicando que as faces são comparadas aos pares (uma face contra a outra).

As bases LFW e FGLFW já apresentam uma lista de pares de imagens para serem utilizadas durante o treinamento e teste baseado no problema do *pair matching*. No entanto, a base BioID não possui tais listas e, a fim de também utilizá-la neste mesmo propósito, uma lista de pares de imagens da base BioID foi criada e apresentada no Apêndice A desta tese.

Esta lista criada contendo os pares de imagens foi elaborada manualmente a partir das imagens da base BioID, gerando 1774 pares de imagens, os quais foram divididos em dois grupos de 887 imagens cada, sendo um grupo utilizado para treinamento e outro para teste, sem sobreposição das imagens entre os grupos.

4.2 Critérios de Avaliação

Nesta seção são descritos os métodos de avaliação empregados ao longo desta tese. Os métodos foram divididos em acordo com os experimentos realizados e descritos no capítulo 5:

- ✓ Ajuste paramétrico dos filtros HBWLF e MHBWLF;
- ✓ Avaliação do método proposto para a localização das coordenadas dos centros dos olhos;
- ✓ Teste de correspondência de imagem para caracterização do desempenho do descritor proposto HBWD;
- ✓ Análise do desempenho do método proposto para o reconhecimento facial baseado no problema do *pair matching*.

4.2.1 Ajuste Paramétrico dos Filtros HBWLF e MHBWLF

O experimento relacionado ao ajuste paramétrico dos filtros propostos HBWLF e MHBWLF visa uma caracterização dos filtros propostos em função do ajuste de seus diversos parâmetros. O experimento utiliza imagens das bases LFW e BioID, bem como uma imagem sintética a fim de melhor entender o comportamento do filtro quanto às bordas (verticais, horizontais, inclinadas e curvilíneas).

Em um primeiro momento deve-se avaliar o desempenho do caso particular do MHBWLF, quando o mesmo converge para o HBWLF, ou seja, fazendo-se $R = 1$ e $NV = 8$. Neste caso a caracterização deve acontecer quanto à variação do tamanho do filtro pela média da vizinhança ($N \times N$) e da inclinação da curva sigmoide (β). Posteriormente deve-se escolher um valor para $N \times N$ e β e analisar o efeito da variação do raio (R) e do número de vizinhos (NV) no desempenho do MHBWLF. Em todos os casos deve-se caracterizar o comportamento dos filtros propostos quanto ao realce das bordas e do conteúdo de ruído presente na imagem.

Por fim, os filtros propostos HBWLF e MHBWLF devem ser comparados com o filtro gerado pela excitação diferencial do WLD, uma vez que os filtros propostos também fazem uso da fração de Weber sem, no entanto, considerar o valor do pixel central, mas sim o valor médio de sua vizinhança, o que deve torná-lo mais robusto quanto ao ruído.

4.2.2 Localização Precisa dos Olhos

A fim de medir o desempenho do método proposto para a localização precisa dos centros dos olhos será utilizada a medida normalizada proposta por Jesorsky, Kirchberg e Frischholz (2001), apresentada na Equação 32.

$$D = \frac{\max(\text{dist}(\tilde{c}_R, C_R), \text{dist}(\tilde{c}_L, C_L))}{\|C_R - C_L\|} \quad \text{Equação 32}$$

Na Equação 32 $dist(\tilde{c}_R, C_R)$ e $dist(\tilde{c}_L, C_L)$ correspondem às distâncias Euclidianas dos centros dos olhos calculados ($\tilde{c}_R$ e $\tilde{c}_L$) aos respectivos centros reais dados pelo *ground-truth* das bases (C_R e C_L). O operador $\max(\dots)$ computa a máxima distância Euclidiana calculada e o resultado é normalizado pela distância Euclidiana dos olhos reais, tornando o erro D independente quanto ao tamanho da face.

Neste trabalho os resultados serão reportados calculando-se a acurácia para três pontos de erro relativo ($D < 0,05$, $D < 0,10$ e $D < 0,25$), método largamente empregado em diversos trabalhos correlatos encontrados na literatura (TAN et al., 2009) (YI; LEI; LI, 2011) (VAZQUEZ-FERNANDEZ; GONZALEZ-JIMENEZ; YU, 2013) (ZHANG et al., 2014) (FLOREA; FLOREA; VERTAN, 2015). Sendo assim, a acurácia reportada é calculada, para um dado erro D ($D < 0,05$, $D < 0,10$ ou $D < 0,25$), como sendo a quantidade de imagens no experimento que apresenta um erro menor do que D .

Por exemplo, um experimento com 100 imagens, onde 80 imagens apresentaram um erro $D < 0,05$, significa que sua acurácia foi de 80% no nível de erro $D < 0,05$. A adoção desta métrica traz importantes informações sobre o método em teste, pois permite avalia-lo rapidamente olhando apenas para estes três pontos de erro relativo e ainda, é possível comparar facilmente o método em teste com diversos outros trabalhos já reportados na literatura.

4.2.3 Correspondência de Imagens

O teste de correspondência de imagens deve ser empregado a fim de avaliar o descritor proposto HBWD quanto a diversas transformações (adição de ruído, rotação, uso de filtros e escala) e compará-lo com outros descritores encontrados na literatura (CS-LBP e WLD). Uma vez que o descritor proposto foi desenvolvido inicialmente para uma aplicação de reconhecimento facial por meio de pontos da região periocular, torna-se interessante realizar o presente teste fazendo-se o uso de pontos faciais dentro da própria região periocular.

Dada uma imagem facial A , o descritor proposto HBWD deve ser aplicado localmente em uma grade com N_{PONTOS} e, então, a imagem A é transformada em uma imagem B (pela adição de ruído, ou rotação, ou aplicação de filtros, ou

mudança de escala) e seus pontos são também descritos. Em seguida calcula-se a distância Euclidiana entre todos os pontos, gerando assim N_{PONTOS}^2 de distâncias. Pelo *ground-truth* sabe-se que destas N_{PONTOS}^2 distâncias, N_{PONTOS} correspondem às correspondências verdadeiras. Em seguida utiliza-se um classificador do tipo k-NN (do inglês, *k-Nearest Neighbour*): define-se um limiar de distância Euclidiana de tal modo que os k elementos mais próximos são classificados como positivos. A fim de avaliar o classificador e o desempenho do descritor proposto HBWD é utilizada a curva *recall x precision*. Inicialmente definem-se as seguintes nomenclaturas:

- ✓ TP (do inglês, *true positive*): elementos classificados como positivos e que são realmente positivos;
- ✓ FP (do inglês, *false positive*): elementos classificados como positivos, mas que são negativos;
- ✓ TN (do inglês, *true negative*): elementos classificados como negativos e que são realmente negativos;
- ✓ FN (do inglês, *false negative*): elementos classificados como negativos, mas que são positivos.

A variável *recall* mede a probabilidade de uma variável relevante ser recuperada pelo classificador e é dada pela Equação 33. Por outro lado, a variável *precision* mede a probabilidade de uma amostra recuperada pelo classificador ser relevante, e é dada pela Equação 34.

$$\text{recall} = \frac{TP}{TP + FN} \quad \text{Equação 33}$$

$$\text{precision} = \frac{TP}{TP + FP} \quad \text{Equação 34}$$

A curva *recall x precision* é obtida variando-se o valor do limiar de distância Euclidiana utilizado no classificador k-NN, plotando-se os valores de *recall* e *precision* em um mesmo gráfico. Quanto mais próximo do topo do gráfico, melhor é o desempenho do classificador: um classificador ideal tocaria o topo do gráfico em seu canto superior direito, com coordenadas iguais a (1,1). Por fim, deve-se calcular a

AUC (do inglês, *area under curve*) da curva *recall x precision*, onde quanto mais próximo do valor unitário, melhor será o classificador.

4.2.4 Reconhecimento Facial

O experimento relacionado ao método de reconhecimento facial visa calcular a acurácia do método proposto, a qual se dá conforme a Equação 35.

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{Equação 35}$$

onde:

- ✓ TP: correspondem aos pares de imagens do tipo *matched* e classificados como *matched*;
- ✓ TN: correspondem aos pares de imagens do tipo *mismatched* e classificados como *mismatched*;
- ✓ FP: correspondem aos pares de imagens do tipo *mismatched*, porém classificados como *matched*;
- ✓ FN: correspondem aos pares de imagens do tipo *matched*, porém classificados como *mismatched*.

Todos os parâmetros do método proposto devem ser ajustados utilizando-se o conjunto *view 1* da base LFW e a melhor configuração obtida deve ser considerada como configuração final para os demais testes. Para reportar o resultado final na base LFW deverá ser utilizado o conjunto *view 2* da mesma, evitando assim um *overfitting* na solução proposta.

Todos os testes e treinamentos devem ser conduzidos no modo restrito, quando não é permitido utilizar os nomes das faces para a composição de novos pares de testes. E ainda, os testes e treinamento devem ser executados sem a utilização de imagens externas adicionais aos conjuntos fornecidos pela base. Esta mesma metodologia de teste deve ser aplicada também nas bases FGLFW e BioID.

5 RESULTADOS

Neste capítulo são apresentados os resultados experimentais dos métodos propostos, utilizando-se os materiais e métodos descritos no capítulo 4. Em um primeiro momento os filtros propostos (HBWLF e MHBWLF) são caracterizados para, em seguida, apresentar os resultados de suas aplicações no método proposto para a localização dos olhos de uma face em ambientes não controlados, comparando os resultados obtidos com o estado da arte.

Posteriormente apresenta-se um experimento de correspondência de imagens para avaliação do descritor proposto (HBWD), onde o mesmo é comparado com outros descritores clássicos na literatura (WLD e CS-LBP). Neste experimento são avaliados também a robustez do HBWD quanto ao ruído, escala, rotação e movimentação da câmera.

Por fim, é apresentado um experimento para avaliação do método proposto para o reconhecimento facial em ambientes não controlados, baseado no problema do *pair matching*. O experimento foi realizado usando-se as bases LFW, FGLFW e BioID e os resultados são comparados com aqueles encontrados na literatura.

5.1 Filtros: HBWLF e MHBWLF

Nesta seção os filtros propostos são avaliados a fim de se verificar o seu desempenho e, principalmente, seus diversos ajustes paramétricos. Os testes foram realizados com as imagens apresentadas na Figura 35:

- ✓ Figura 35 (a): imagem sintética, desenvolvida neste trabalho;
- ✓ Figura 35 (b): imagem Abel_Pacheco_0001 da base LFW;
- ✓ Figura 35 (c): imagem Martha_Lucia_Ramirez_0002 da base LFW;
- ✓ Figura 35 (d): imagem BioID_0680 da base BioID;
- ✓ Figura 35 (e): imagem BioID_0010 da base BioID.

O primeiro experimento foi realizado para o caso particular em que $NV = 8$ e $R = 1$, ou seja, quando o MHBWLF torna-se o HBWLF. As imagens da Figura 35

foram assim filtradas pelo HBWLF, variando-se o filtro pela média da vizinhança e a inclinação da curva sigmoide:

- ✓ Filtros pela média da vizinhança utilizados: $N \times N = 5 \times 5$; 7×7 e 9×9 ;
- ✓ Inclinação da curva sigmoide: $\beta = 1,00$; $0,30$ e $0,10$.

Os resultados estão apresentados conforme a seguir:

- ✓ Figura 36: resultados da aplicação do HBWLF na Figura 35 (a);
- ✓ Figura 37: resultados da aplicação do HBWLF na Figura 35 (b);
- ✓ Figura 38: resultados da aplicação do HBWLF na Figura 35 (c);
- ✓ Figura 39: resultados da aplicação do HBWLF na Figura 35 (d);
- ✓ Figura 40: resultados da aplicação do HBWLF na Figura 35 (e).

Figura 35 – Imagens para caracterização dos filtros: (a) imagem sintética; (b) e (c) imagens da base LFW; (d) e (e) imagens da base BioID.

Fonte: autoria própria.

As imagens filtradas demonstram que, quanto maior o tamanho do filtro pela média da vizinhança, maior será o tamanho das bordas realçadas, e vice-versa. O parâmetro β tem um efeito direto na suavização da imagem filtrada: quanto maior o valor de β , mais suave será a imagem filtrada de saída. Por outro lado, a medida que se diminui o valor de β há um maior realce das componentes de alta frequência, o que também acarreta no aumento do conteúdo de ruído que estava presente na imagem original. Desta forma, o ajuste de β deve ser feito com o compromisso de ajustar os detalhes desejados sem, no entanto, aumentar demasiadamente o ruído da imagem.

Figura 36 – Imagem da Figura 35 (a) filtrada pelo HBWLF, usando um filtro pela média da vizinhança de (I) 5x5; (II) 7x7 e (III) 9x9. O parâmetro β vale (a) 1,00; (b) 0,30 e (c) 0,10.

Fonte: autoria própria.

Figura 37 – Imagem da Figura 35 (b) filtrada pelo HBWLF, usando um filtro pela média da vizinhança de (I) 5x5; (II) 7x7 e (III) 9x9. O parâmetro β vale (a) 1,00; (b) 0,30 e (c) 0,10.

Fonte: autoria própria.

Figura 38 – Imagem da Figura 35 (c) filtrada pelo HBWLF, usando um filtro pela média da vizinhança de (I) 5x5; (II) 7x7 e (III) 9x9. O parâmetro β vale (a) 1,00; (b) 0,30 e (c) 0,10.

Fonte: autoria própria.

Figura 39 – Imagem da Figura 35 (d) filtrada pelo HBWLF, usando um filtro pela média da vizinhança de (I) 5x5; (II) 7x7 e (III) 9x9. O parâmetro β vale (a) 1,00; (b) 0,30 e (c) 0,10.

Fonte: autoria própria.

Figura 40 – Imagem da Figura 35 (e) filtrada pelo HBWLF, usando um filtro pela média da vizinhança de (I) 5x5; (II) 7x7 e (III) 9x9. O parâmetro β vale (a) 1,00; (b) 0,30 e (c) 0,10.

Fonte: autoria própria.

A Figura 41 mostra a imagem da Figura 35 (a) filtrada pelo MHBWLF usando um filtro pela média da vizinhança de tamanho $N \times N = 7 \times 7$ e $\beta = 0,30$. O raio foi variado ($R = 1, 2, 3, 5$ e 7), bem como a quantidade de pixels na vizinhança ($NV = 1, 4$ e 8).

O aumento do raio R provoca um deslocamento da borda realçada, bem como o seu alargamento: nota-se na Figura 41 que as bordas claras se invertem de posição com as bordas escuras e ficam mais largas com o aumento do raio. A quantidade de vizinhos NV influencia diretamente na continuidade das bordas detectadas: um aumento em NV proporciona a detecção de bordas mais lisas e contínuas.

A Figura 42 apresenta uma comparação de uma imagem facial filtrada pelo WLD proposto por Chen et al. (2008) e pelo MHBWLF ($N \times N = 7 \times 7$, $\beta = 0,30$, $R = 2$ e $NV = 8$), na qual verifica-se claramente que a imagem filtrada pelo MHBWLF possui muito mais detalhes realçados do que aquela filtrada pelo WLD.

Figura 41 – Imagem da Figura 35 (a) filtrada pelo MHBWLF, usando um filtro pela média da vizinhança de 7×7 , β igual a 0,30 e R igual a (I) 1; (II) 2; (III) 3; (IV) 5 e (V) 7. O parâmetro NV vale (a) 1; (b) 4 e (c) 8.

Fonte: autoria própria.

Figura 42 – Comparação do filtro MHBWLF e WLD: (a) face original apresentada em Chen et al. (2008); (b) face filtrada pelo WLD; (c) face filtrada pelo MHBWLF.

Fonte: autoria própria.

5.2 Localização Precisa dos Olhos

O método proposto para a localização precisa dos olhos foi avaliado usando-se as imagens das bases LFW e BioID: os resultados aqui apresentados foram publicados no artigo (AFFONSO; RODRIGUES; PAIVA, 2018). O primeiro passo desta avaliação foi a realização do ajuste dos parâmetros do filtro MHBWLF e, para este propósito, foi utilizado o conjunto de 2000 imagens da base LFW com as coordenadas dos olhos já conhecidas, o que compõe o chamado *ground-truth*, proposto por Tan et al. (2009).

O MHBWLF foi ajustado para o caso particular (com $R = 1$ e $NV = 8$), quando o MHBWLF converge para o HBWLF, enquanto $\alpha = 0,001$ (Equação 21) e $\epsilon = 0,1$ (Equação 25), uma vez que ambas constantes devem ser ajustadas para valores pequenos, reais e positivos. O limiar para avaliação dos picos de correlação foi ajustado em $T_H = 0,98$, a fim de considerar todos os pixels com intensidade maior que 98% da intensidade do maior pico de correlação. Foram testados dois filtros pela média da vizinhança (3×3 e 5×5), variando o valor de β (Equação 19). A Tabela 1 resume os resultados obtidos, demonstrando que a melhor configuração foi obtida usando-se um filtro de tamanho 5×5 e $\beta = 0,3$.

Tabela 1 – Ajuste do HBWLF usando a base LFW.

N x N	β	Acurácia (%) na Base LFW		
		D < 0,05	D < 0,10	D < 0,25
3x3	0,01	79,8%	96,7%	99,9%
3x3	0,1	82,2%	96,9%	99,9%
3x3	0,2	79,8%	96,3%	99,8%
3x3	0,3	77,3%	95,3%	99,7%
5x5	0,01	82,4%	97,4%	99,9%
5x5	0,1	84,4%	97,8%	99,9%
5x5	0,2	85,4%	98,3%	99,9%
5x5	0,25	85,4%	98,2%	99,9%
5x5	0,3	85,5%	98,1%	99,9%
5x5	0,35	85,2%	98,1%	99,9%
5x5	0,4	85,3%	98,2%	99,9%
5x5	0,5	85,3%	98,0%	99,9%

Fonte: adaptado de (AFFONSO; RODRIGUES; PAIVA, 2018).

A fim de produzir uma comparação mais justa com outros métodos clássicos encontrados na literatura, foi feito um primeiro experimento, onde todos os métodos foram avaliados usando-se o mesmo protocolo: os métodos foram avaliados usando-se o conjunto de 2000 imagens da base LFW (TAN et al., 2009) e um esquema de validação cruzada de *2-fold*. Nesta avaliação foi testado o método proposto, variando-se o filtro da Figura 15: HBWLF, SIFT (LOWE, 2004), CS-LBP (HEIKKILÄ; PIETIKÄINEN; SCHMID, 2009), CLAHE (ZUIDERVELD, 1994) e WLD (CHEN et al., 2008). Os resultados estão apresentados na Tabela 2.

Tabela 2 – Acurácia de localização (%) usando o conjunto de 2000 imagens da base LFW, com o mesmo protocolo para todos os filtros.

Filtro	Acurácia (%) na Base LFW		
	D < 0,05	D < 0,10	D < 0,25
SIFT (Gradient)	80,2%	94,9%	97,9%
CS-LBP	78,1%	96,0%	99,9%
CLAHE	81,6%	96,8%	99,9%
WLD	81,8%	96,2%	99,9%
HBWLF Proposto	85,5%	98,1%	99,9%

Fonte: adaptado de (AFFONSO; RODRIGUES; PAIVA, 2018).

Os resultados apresentados na Tabela 2 demonstram que o método proposto para a localização dos olhos usando o filtro HBWLF atinge os melhores resultados, especialmente para os menores erros ($D < 0,05$ e $D < 0,10$). A Tabela 3 apresenta uma comparação dos resultados obtidos com aqueles encontrados na literatura, para trabalhos que também fizeram uso de imagens da base LFW.

Os resultados dos métodos Original ASEF, ASEF+Msk20x20 e iASEF (VAZQUEZ-FERNANDEZ; GONZALEZ-JIMENEZ; YU, 2013); Fast LL-SVM (ZHANG et al., 2014); RDHV e D-RDHV (JIN; TAN, 2016) foram todos obtidos usando-se o mesmo conjunto de 2000 imagens da base LFW proposto por Tan et al. (2009). No método Enhanced PS, Tan et al. (2009) fizeram o uso do mesmo conjunto de 2000 imagens da base LFW, porém com várias transformações a fim de criar um novo conjunto com mais de 17.000 imagens. No método Intensity Filtering + K-means, Qian e Xu (2010) usaram outro conjunto de 1200 imagens da base LFW. E, por fim, Wang, Xiong e Chu (2015), no método Gaussian Mixture Models, avaliaram seu método em um conjunto com 10 pontos faciais marcados em imagens da base LFW.

Tabela 3 – Acurácia de localização (%) usando imagens da base LFW – uma comparação com outros resultados encontrados na literatura.

Método	Acurácia (%) na Base LFW		
	D < 0,05	D < 0,10	D < 0,25
Original ASEF	74,0%	90,7%	97,2%
ASEF + Mask20x20	74,4%	91,9%	99,0%
Intensity Filtering + k-means	75,1%	90,6%	96,1%
Gaussian Mixture Models	78,0%	97,0%	-
RDHV	78,4%	97,5%	-
iASEF	78,5%	96,0%	99,6%
Enhanced OS	80,2%	98,4%	-
Fast LL-SVM	81,0%	93,3%	98,9%
D-RDHV	84,4%	98,1%	-
Método Proposto	85,5%	98,1%	99,9%

Fonte: adaptado de (AFFONSO; RODRIGUES; PAIVA, 2018).

O método proposto obteve o melhor desempenho para os erros $D < 0,05$ e $D < 0,25$: obteve uma acurácia de 85,5% para um erro de $D < 0,05$ e uma acurácia de 99,9% para $D < 0,25$. Para o erro $D < 0,10$ o método Enhanced PS apresentou o melhor resultado, uma acurácia de 98,4%, muito próxima da acurácia apresentada pelo método proposto, que foi de 98,1%. Conforme apresentado anteriormente, os resultados resumidos na Tabela 3 não foram obtidos exatamente sob os mesmos protocolos, mas todos foram obtidos usando-se imagens da base LFW, e a maioria deles empregando o mesmo conjunto de imagens. Estes resultados demonstram claramente que o método proposto é muito competitivo para aplicações de localização dos centros dos olhos em cenários não controlados, mesmo quando comparado com vários outros métodos estado da arte encontrados na literatura.

Os mesmos experimentos descritos anteriormente e realizados com imagens da base LFW foram também realizados com as imagens da base BioID, também utilizando a validação cruzada de *2-fold*. A Tabela 4 resume os resultados para o teste do método proposto para localização dos olhos na base BioID utilizando-se os mesmos filtros que no caso anterior. Os resultados obtidos demonstram que o método proposto utilizando o HBWLF obteve a melhor acurácia, para todos os erros considerados ($D < 0,05$, $D < 0,10$ e $D < 0,25$). A Tabela 5 mostra uma comparação dos resultados obtidos com outros métodos encontrados na literatura considerando apenas a utilização de imagens da base BioID.

Tabela 4 – Acurácia de localização (%) usando a base BioID, com o mesmo protocolo para todos os filtros.

	Acurácia (%) na Base BioID		
Filtro	D < 0,05	D < 0,10	D < 0,25
SIFT (Gradient)	94,4%	95,9%	99,5%
CS-LBP	93,8%	96,9%	99,5%
CLAHE	95,6%	97,8%	99,7%
WLD	91,9%	95,8%	99,5%
HBWLF Proposto	96,6%	98,8%	99,7%

Fonte: adaptado de (AFFONSO; RODRIGUES; PAIVA, 2018).

Tabela 5 – Acurácia de localização (%) usando a base BioID – uma comparação com outros resultados encontrados na literatura.

	Acurácia (%) na Base BioID		
Método	D < 0,05	D < 0,10	D < 0,25
ZEP-MLP	70,5%	91,9%	98,9%
MICs + SIFT	86,1%	91,7%	97,9%
ASEF + Mask20x20	87,2%	94,8%	99,6%
MSD	89,6%	95,5%	99,1%
iASEF	93,0%	98,7%	100,0%
Método Proposto	96,6%	98,8%	99,7%

Fonte: adaptado de (AFFONSO; RODRIGUES; PAIVA, 2018).

Cabe ressaltar que os métodos apresentados na Tabela 5 foram todos testados com as imagens da base BioID: ZEP-MLP (FLOREA; FLOREA; VERTAN, 2015), MICS+SIFT (VALENTI; GVERS, 2012), ASEF+Mask20x20 (VAZQUEZ-FERNANDEZ; GONZALEZ-JIMENEZ; YU, 2013), MSD (YANG et al., 2011) e iASEF (VAZQUEZ-FERNANDEZ; GONZALEZ-JIMENEZ; YU, 2013). O método proposto atingiu a melhor acurácia para os erros $D < 0,05$ (96,6%) e $D < 0,10$ (98,8%). Para o erro $D < 0,25$ o método iASEF obteve a melhor acurácia, de 100,0%, porém muito próxima da acurácia obtida pelo método proposto, que foi de 99,7%.

Desta forma, os resultados obtidos com os testes usando as imagens da base BioID corroboram os testes anteriores, realizados com as imagens da base LFW, demonstrando que o método proposto é muito competitivo, mesmo quando comparado com outros métodos que representam o estado da arte em algoritmos de localização dos centros dos olhos de uma face. A Figura 43 ilustra alguns exemplos de localização dos centros dos olhos em imagens da base LFW.

Figura 43 – Resultados obtidos usando o método proposto (círculos cheios em vermelhos) e o *ground-truth* (marcas x em amarelo) para o conjunto de 2000 imagens da base LFW: (a) $D=0$; (b) $D \approx 0.05$; (c) $D \approx 0.10$; (d) $D \approx 0.15$ e (e) $D \approx 0.20$.

Fonte: Extraído de (AFFONSO; RODRIGUES; PAIVA, 2018).

5.3 Correspondência de Imagens com o HBWD

O descritor proposto HBWD foi avaliado em um experimento de correspondência de imagens e para isto foram selecionadas aleatoriamente seis imagens da base BioID, conforme resumido na Tabela 6. Para o teste de correspondência de imagens foi montada uma grade artificial, baseada nas posições dos centros dos olhos e ponta do nariz de cada imagem facial, conforme ilustrado na Figura 44 para a Imagem 1 da Tabela 6. A Figura 45 apresenta todas as seis imagens da Tabela 6, bem como suas respectivas grades de pontos.

Tabela 6 – Imagens selecionadas aleatoriamente da base BioID.

Imagem	Nome na Base BioID
1	BioID_0078
2	BioID_0127
3	BioID_0680
4	BioID_0999
5	BioID_1379
6	BioID_1519

Fonte: autoria própria.

Figura 44 – Grade de pontos utilizada no experimento.

Fonte: autoria própria.

Considerando as seguintes coordenadas:

- ✓ (x_R, y_R) : centro do olho direito;
- ✓ (x_L, y_L) : centro do olho esquerdo;
- ✓ (x_N, y_N) : ponta do nariz.

E, desta forma, as coordenadas (x_i, y_i) dos onze pontos da grade representada na Figura 44, onde i corresponde ao i -ésimo ponto, são dadas por (Equação 36 até Equação 57):

$$x_1 = x_R - (x_N - x_R)$$

Equação 36

$$y_1 = y_R \quad \text{Equação 37}$$

$$x_2 = \frac{(x_R + x_L)}{2} - \left(x_N - \frac{(x_R + x_L)}{2} \right) \quad \text{Equação 38}$$

$$y_2 = y_N \quad \text{Equação 39}$$

$$x_3 = x_L - (x_N - x_L) \quad \text{Equação 40}$$

$$y_3 = y_L \quad \text{Equação 41}$$

$$x_4 = x_R \quad \text{Equação 42}$$

$$y_4 = y_R \quad \text{Equação 43}$$

$$x_5 = \frac{(x_R + x_L)}{2} \quad \text{Equação 44}$$

$$y_5 = y_N \quad \text{Equação 45}$$

$$x_6 = x_L \quad \text{Equação 46}$$

$$y_6 = y_L \quad \text{Equação 47}$$

$$x_7 = \frac{(x_R + x_N)}{2} \quad \text{Equação 48}$$

$$y_7 = \frac{(y_R + y_N)}{2} \quad \text{Equação 49}$$

$$x_8 = \frac{(x_L + x_N)}{2} \quad \text{Equação 50}$$

$$y_8 = \frac{(y_L + y_N)}{2} \quad \text{Equação 51}$$

$$x_9 = x_N \quad \text{Equação 52}$$

$$y_9 = y_R \quad \text{Equação 53}$$

$$x_{10} = x_N \quad \text{Equação 54}$$

$$y_{10} = y_N \quad \text{Equação 55}$$

$$x_{11} = x_N \quad \text{Equação 56}$$

$$y_{11} = y_L \quad \text{Equação 57}$$

Para cada imagem da Tabela 6 calculam-se os descritores HBWD para todos os pontos de sua respectiva grade. Em seguida a imagem passa por uma transformação e calculam-se novamente os descritores HBWD na nova imagem. Posteriormente calculam-se as distâncias Euclidianas de cada descritor da imagem original para cada um da imagem transformada, gerando assim um vetor de 121 distâncias. Estas 121 distâncias correspondem a 11 pontos de correspondências verdadeiras e a 110 pontos de correspondências falsas. Os resultados são avaliados por meio da AUC das curvas *recall x precision*, conforme detalhado no capítulo 4.

O experimento foi realizado aplicando-se as seguintes transformações nas imagens de entrada:

- ✓ Adição de ruído: foram adicionados três tipos diferentes de ruído na imagem original (ruído Gaussiano, *Speckle* e *Salt & Pepper*);
- ✓ Rotação: foram testadas seis rotações diferentes (5°, 10°, 15°, 20°, 25° e 30°);
- ✓ Filtros: três tipos diferentes foram empregados (*average*, *unsharp* e *motion*);
- ✓ Escala: seis fatores de escala foram avaliados (0,7, 0,8, 0,9, 1,1, 1,2 e 1,3).

Figura 45 – Imagens da base BioID selecionadas para o experimento: (I) Imagem 1, (II) Imagem 2, (III) Imagem 3, (IV) Imagem 4, (V) Imagem 5 e (VI) Imagem 6; (a) imagem original da base BioID, (b) imagem com grade de pontos.

Fonte: autoria própria.

5.3.1 Adição de Ruído

Conforme explicado anteriormente, foram utilizados três tipos diferentes de ruído: Gaussiano, Speckle e Salt & Pepper. As seções a seguir detalham cada um destes experimentos.

5.3.1.1 Adição de Ruído Gaussiano

Neste experimento foi adicionado um ruído do tipo Gaussiano na imagem original a fim de avaliar o desempenho do descritor proposto nesta situação. A função de densidade de probabilidade do ruído Gaussiano, $f(x)$, é igual à distribuição normal, com uma média μ e desvio padrão σ , conforme a Equação 58.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{Equação 58}$$

Foram testados dois casos, ambos com média nula ($\mu = 0$):

- ✓ Caso 1: $\sigma^2 = 0,005$;
- ✓ Caso 2: $\sigma^2 = 0,010$.

A Figura 46 ilustra as imagens originais, sem ruído, bem como as imagens com ruído Gaussiano adicionado (casos 1 e 2). O descritor proposto HBWD foi ajustado e testado manualmente e, dentre as configurações testadas, a configuração abaixo apresentou o melhor desempenho:

- ✓ Filtro pela média da vizinhança: $N \times N = 13 \times 13$;
- ✓ Vizinhança circular: $R = 1$ e $NV = 8$;
- ✓ Curva sigmoide: $\beta = 0,30$;
- ✓ Histograma: $N_{bins} = 8$;
- ✓ Células: $N_C \times N_C = 4 \times 4$ e $N_p \times N_p = 12 \times 12$.
- ✓ Limiar de iluminação: $T_{H_{IL}} = 0,20$.

Figura 46 – Adição de ruído Gaussiano: (a) imagem original, (b) imagem com ruído Gaussiano (Caso 1) e (c) imagem com ruído Gaussiano (Caso 2).

Fonte: autoria própria.

Para cada imagem foram rodadas 50 amostras diferentes de ruído Gaussiano, uma vez que foi utilizado um gerador randômico de ruído. A Tabela 7 e a Figura 47 apresentam os resultados obtidos pelos descritores HBWD, CS-LBP (HEIKKILÄ; PIETIKÄINEN; SCHMID, 2009) e WLD (CHEN et al., 2008). A AUC Média corresponde à média das AUCs para as 50 amostras de cada imagem. A coluna Desv Pad AUC corresponde ao desvio padrão das AUCs calculadas. Verifica-se que o descritor HBWD foi aquele que apresentou o melhor desempenho global quanto ao teste de adição de ruído Gaussiano, com uma AUC Média bem superior aos demais descritores, para os dois níveis de ruídos estudados (casos 1 e 2).

Tabela 7 – Resultado do teste de correspondência de imagem com adição de ruído Gaussiano de média nula (50 amostras para cada teste).

		HBWD		CS-LBP		WLD	
Imagem	Sigma	AUC Média	Desv Pad AUC	AUC Média	Desv Pad AUC	AUC Média	Desv Pad AUC
1	sigma = 0,005	0,9720	0,0119	0,9743	0,0211	0,8947	0,0293
	sigma = 0,010	0,8941	0,0205	0,8808	0,0507	0,8164	0,0522
2	sigma = 0,005	0,7203	0,0258	0,4084	0,1068	0,4609	0,1215
	sigma = 0,010	0,6125	0,0374	0,2561	0,0739	0,2976	0,0883
3	sigma = 0,005	0,7992	0,0258	0,6255	0,0801	0,5510	0,1031
	sigma = 0,010	0,6583	0,0471	0,4031	0,0805	0,3863	0,0915
4	sigma = 0,005	1,0000	0,0000	0,9282	0,0391	0,9040	0,0547
	sigma = 0,010	1,0000	0,0000	0,7599	0,0766	0,6898	0,1021
5	sigma = 0,005	0,9796	0,0121	0,9160	0,0420	0,8640	0,0697
	sigma = 0,010	0,9160	0,0194	0,6943	0,0819	0,6325	0,1054
6	sigma = 0,005	0,8584	0,0179	0,7281	0,0743	0,6407	0,0755
	sigma = 0,010	0,7360	0,0311	0,4540	0,0759	0,3940	0,0749
AUC Média Total		0,8455		0,6691		0,6277	
AUC (média e desvio padrão) com ruído Gaussiano de média nula (50 amostras)							

Fonte: autoria própria.

Figura 47 – Resultados gráficos do teste de adição de ruído gaussiano.

Fonte: autoria própria.

5.3.1.2 Adição de Ruído Speckle

Neste experimento o descritor proposto HBWD foi avaliado quanto ao ruído multiplicativo, também conhecido como ruído Speckle. Este ruído, por ser multiplicativo, é proporcional à intensidade do pixel ao qual ele afeta e o seu cálculo pode ser expresso conforme a Equação 59, onde I é a imagem de entrada, I_{NOISE} a imagem de saída, com ruído Speckle, e n um ruído randômico com distribuição uniforme, média zero ($\mu = 0$) e variância igual a σ^2 .

$$I_{NOISE} = I + n \cdot I \quad \text{Equação 59}$$

A Figura 48 apresenta, para todas as seis imagens selecionadas da base BioID, a imagem original e as imagens com ruído Speckle, para os dois casos avaliados:

- ✓ Caso 1: $\sigma^2 = 0,005$;
- ✓ Caso 2: $\sigma^2 = 0,010$.

Da mesma forma que no experimento com adição de ruído Gaussiano, este experimento também utilizou um gerador randômico de ruído e o mesmo foi executado considerando-se 50 amostras para cada uma das imagens avaliadas. O descritor proposto HBWD foi avaliado com a mesma configuração do caso anterior, ou seja, do caso de adição de ruído Gaussiano. A Tabela 8 e a Figura 49 apresentam os resultados obtidos, onde verifica-se que o descritor proposto HBWD apresenta resultados ligeiramente superiores àqueles obtidos com os descritores CS-LBP e WLD.

Cabe ressaltar que o descritor proposto HBWD não foi ajustado para este teste, ou seja, foi utilizada exatamente a mesma configuração empregada no teste de imagens com adição de ruído Gaussiano.

Figura 48 – Adição de ruído Speckle: (a) imagem original, (b) imagem com ruído Speckle (Caso 1) e (c) imagem com ruído Speckle (Caso 2).

Fonte: autoria própria.

Tabela 8 – Resultado do teste de correspondência de imagem com adição de ruído Speckle de média nula (50 amostras para cada teste).

		HBWD		CS-LBP		WLD	
Imagem	Sigma	AUC Média	Desv Pad AUC	AUC Média	Desv Pad AUC	AUC Média	Desv Pad AUC
1	sigma = 0,005	0,9917	0,0087	0,9944	0,0068	0,9264	0,0212
	sigma = 0,010	0,9196	0,0178	0,9647	0,0178	0,8751	0,0379
2	sigma = 0,005	0,9584	0,0208	0,9666	0,0222	0,9671	0,0215
	sigma = 0,010	0,8278	0,0166	0,8568	0,0520	0,8927	0,0499
3	sigma = 0,005	0,9987	0,0018	0,9901	0,0101	0,9814	0,0119
	sigma = 0,010	0,9768	0,0114	0,9404	0,0248	0,9022	0,0367
4	sigma = 0,005	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000
	sigma = 0,010	1,0000	0,0000	0,9999	0,0005	0,9987	0,0028
5	sigma = 0,005	1,0000	0,0000	0,9981	0,0033	0,9940	0,0084
	sigma = 0,010	0,9930	0,0103	0,9812	0,0160	0,9423	0,0307
6	sigma = 0,005	0,9723	0,0164	0,9682	0,0207	0,9528	0,0221
	sigma = 0,010	0,8847	0,0198	0,8499	0,0526	0,8225	0,0649
AUC Média Total		0,9603		0,9592		0,9379	
AUC (média e desvio padrão) com ruído Speckle de média nula (50 amostras)							

Fonte: autoria própria.

Figura 49 – Resultados gráficos do teste de adição de ruído Speckle.

Fonte: autoria própria.

5.3.1.3 Adição de Ruído Salt & Pepper

Neste experimento o descritor proposto HBWD foi avaliado quanto ao ruído impulsivo, conhecido como ruído Salt & Pepper. Este ruído é caracterizado por possuir randomicamente pixels brancos ou escuros ao longo da imagem, em uma representação direta de *dead pixels* (pixels danificados permanentemente), erros de conversão analógica/digital ou mesmo erros durante a transmissão. Os pixels com

ruído são distribuídos randomicamente ao longo da imagem com uma densidade d . Neste experimento foram avaliados dois casos, conforme ilustrado na Figura 50:

- ✓ Caso 1: $d = 0,010$;
- ✓ Caso 2: $d = 0,020$.

Do mesmo modo que nos testes anteriores, o experimento foi realizado com a utilização de 50 amostras de imagens com ruído Salt & Pepper e o descritor proposto HBWD foi ajustado para a mesma configuração do teste com adição de ruído Gaussiano. A Tabela 9 e a Figura 51 resumem os resultados obtidos para os descritores HBWD, CS-LBP e WLD, onde se verifica que os resultados obtidos com o descritor proposto HBWD são ligeiramente superiores que aqueles obtidos pelos demais descritores.

Tabela 9 – Resultado do teste de correspondência de imagem com adição de ruído Salt & Pepper de média nula (50 amostras para cada teste).

		HBWD		CS-LBP		WLD	
Imagem	Densidade	AUC Média	Desv Pad AUC	AUC Média	Desv Pad AUC	AUC Média	Desv Pad AUC
1	d= 0,010	1,0000	0,0000	1,0000	0,0000	1,0000	0,0000
	d = 0,020	0,9846	0,0159	0,9868	0,0114	0,9789	0,0319
2	d= 0,010	0,9929	0,0126	0,9850	0,0175	0,9909	0,0100
	d = 0,020	0,8505	0,0552	0,7370	0,0504	0,8410	0,0643
3	d= 0,010	0,9993	0,0022	0,9993	0,0028	0,9980	0,0039
	d = 0,020	0,9705	0,0203	0,9508	0,0280	0,9594	0,0286
4	d= 0,010	0,9998	0,0014	1,0000	0,0000	0,9998	0,0014
	d = 0,020	0,9678	0,0264	0,9729	0,0209	0,9598	0,0315
5	d= 0,010	0,9993	0,0024	0,9969	0,0050	0,9972	0,0041
	d = 0,020	0,9317	0,0371	0,8610	0,0418	0,9170	0,0471
6	d= 0,010	0,9985	0,0036	0,9982	0,0032	0,9951	0,0079
	d = 0,020	0,9387	0,0313	0,8765	0,0335	0,9184	0,0342
AUC Média Total		0,9695		0,9470		0,9630	
AUC (média e desvio padrão) com ruído Salt & Pepper (50 amostras)							

Fonte: autoria própria.

Figura 50 – Adição de ruído Salt & Pepper: (a) imagem original, (b) imagem com ruído Salt & Pepper (Caso 1) e (c) imagem com ruído Salt & Pepper (Caso 2).

Fonte: autoria própria.

Figura 51 – Resultados gráficos do teste de adição de ruído Salt & Pepper.

Fonte: autoria própria.

5.3.2 Transformação de Rotação

Neste experimento o descritor proposto HBWD foi avaliado quando à rotação da imagem, considerando-se seis rotações: 5°, 10°, 15°, 20°, 25° e 30°. A Figura 52 ilustra a imagem original, bem como as imagens rotacionadas: esta figura ilustra a rotação da imagem 1 da Tabela 6, no entanto, todas as imagens da Tabela 6 foram utilizadas no experimento.

Figura 52 – Transformação de rotação: (a) imagem original. Imagem rotacionada: (b) 5°, (c) 10°, (d) 15°, (e) 20°, (f) 25° e (g) 30°.

Fonte: autoria própria.

A grade de pontos também foi rotacionada, utilizando-se o mesmo ângulo de rotação empregado nas imagens. A rotação das imagens foi feita usando-se o método de interpolação do vizinho mais próximo e os pontos que estavam fora da imagem foram todos ajustados para zero. O descritor proposto HBWD foi ajustado para a mesma configuração utilizada no experimento de adição de ruído Gaussiano.

A Tabela 10 resume os resultados obtidos. Considerando-se os valores médios de AUC verifica-se que para baixas rotações (5° e 10°) o descritor proposto HBWD apresentou os melhores resultados. De 15° a 25° os melhores resultados foram obtidos pelo descritor CS-LBP e, por fim, para 30° o melhor resultado foi obtido pelo HBWD.

Tabela 10 – Resultado do teste de correspondência de imagem com a transformação de rotação nas imagens.

		AUC para diversas Rotações					
Imagem	Método	5°	10°	15°	20°	25°	30°
1	HBWD	1,0000	1,0000	1,0000	0,9805	0,9146	0,8655
	CS-LBP	1,0000	1,0000	1,0000	1,0000	0,9226	0,8042
	WLD	1,0000	1,0000	0,9962	0,9062	0,8516	0,6486
2	HBWD	1,0000	1,0000	0,9542	0,8274	0,5291	0,3537
	CS-LBP	1,0000	0,9930	0,9646	0,8392	0,7023	0,5014
	WLD	1,0000	0,9962	0,9062	0,8036	0,6939	0,4586
3	HBWD	1,0000	1,0000	0,9506	0,9116	0,7689	0,6427
	CS-LBP	1,0000	1,0000	0,9924	0,9120	0,6887	0,4871
	WLD	1,0000	1,0000	1,0000	1,0000	0,8959	0,6017
4	HBWD	1,0000	1,0000	0,9372	0,8300	0,7676	0,4977
	CS-LBP	1,0000	1,0000	0,9924	0,9633	0,8491	0,4186
	WLD	1,0000	1,0000	0,9962	0,9574	0,7079	0,4906
5	HBWD	1,0000	1,0000	1,0000	0,9879	0,8770	0,7207
	CS-LBP	1,0000	1,0000	0,9924	0,8951	0,6880	0,4401
	WLD	1,0000	1,0000	0,9962	0,8878	0,7293	0,4687
6	HBWD	1,0000	1,0000	0,9903	0,9010	0,6614	0,4992
	CS-LBP	1,0000	1,0000	0,9924	0,9146	0,8875	0,4961
	WLD	1,0000	1,0000	1,0000	0,9596	0,8340	0,5367
		AUC Média					
		5°	10°	15°	20°	25°	30°
	HBWD	1,0000	1,0000	0,9721	0,9064	0,7531	0,5966
	CS-LBP	1,0000	0,9988	0,9890	0,9207	0,7897	0,5246
	WLD	1,0000	0,9994	0,9825	0,9191	0,7854	0,5342

Fonte: autoria própria.

Cabe ainda destacar que os resultados obtidos foram muito próximos para os três descritores testados e que a AUC média se manteve próxima do valor unitário para uma rotação de até 15°. Com relação ao descritor proposto HBWD, conforme explicado anteriormente, o mesmo não possui nenhum artifício que o torne invariante quanto à rotação, como ocorre com o SIFT (LOWE, 2004). No entanto, conforme ficou demonstrado neste experimento, não se espera grande perda de desempenho do descritor proposto HBWD para rotações pequenas nas imagens faciais (até 15°). E ainda, o descritor proposto será utilizado no reconhecimento facial de imagens já alinhadas, conforme o método de reconhecimento facial proposto e descrito na seção 3.5. Desta forma, este teste demonstra que o descritor proposto HBWD é robusto para pequenas rotações nas imagens faciais.

5.3.3 Filtros

O descritor proposto HBWD foi avaliado também em imagens que foram filtradas por três tipos de filtros diferentes: *average*, *unsharp* e *motion*. Estes três tipos de filtros modelam muito bem diversas situações reais: o filtro *average* é muito utilizado para a redução de ruído em imagens e pode degradar as bordas das mesmas; o filtro *unsharp*, por outro lado, é um filtro de aguçamento e que realça as bordas, muito utilizado em diversas câmeras digitais; por fim, o filtro *motion*, produz um borramento na imagem, similar ao borramento devido ao movimento da câmera, ou objeto, durante a captura da imagem.

Desta forma, o teste de correspondência de imagens foi executado comparando-se as imagens originais com aquelas filtradas, o que corresponde a uma interessante situação real, onde a imagem original estaria salva em um banco de dados e a imagem filtrada seria uma nova imagem adquirida e que se deseja verificar sua correspondência.

5.3.3.1 Filtro Average

Para este teste foi utilizado um filtro de correlação de tamanho 5x5, conforme descrito na Equação 60. As seis imagens da Tabela 6 foram filtradas por este filtro e o teste de correspondência de imagens foi executado, cujos resultados podem ser encontrados na Tabela 11.

$$h = \frac{1}{5^2} \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}_{5 \times 5} \quad \text{Equação 60}$$

Foi utilizada a mesma configuração do descritor proposto HBWD utilizada no teste de adição de ruído Gaussiano e verifica-se que o resultado obtido pelo descritor proposto é o melhor, quando comparado com os descritores CS-LBP e WLD. E ainda, a AUC Média obtida foi igual a 1,00, indicando que o descritor proposto não teve seu desempenho degradado pelo filtro *average* utilizado.

Tabela 11 – Resultados do teste de correspondência para o filtro average de tamanho 5x5.

	AUC		
Imagem	HBWD	CS-LBP	WLD
1	1,0000	0,9757	0,9245
2	1,0000	0,9417	0,9924
3	1,0000	0,9667	1,0000
4	1,0000	1,0000	1,0000
5	1,0000	0,9137	0,9174
6	1,0000	1,0000	1,0000
AUC Média	1,0000	0,9663	0,9724

Fonte: autoria própria.

5.3.3.2 Filtro Unsharp

Neste experimento foi utilizado um filtro de correlação de tamanho 3x3 para o aguçamento, ou realce das altas frequências, conforme a Equação 61. A Figura 53 apresenta as imagens originais e as imagens filtradas.

Figura 53 – Filtro unsharp: (a) imagens originais, (b) imagens filtradas.

Fonte: autoria própria.

Na Figura 53 foi empregada a mesma configuração do descritor proposto HBWD utilizada no teste de adição de ruído Gaussiano e a Tabela 12 resume os resultados obtidos no teste de correspondência.

$$h = \frac{1}{11} \begin{bmatrix} -1 & -9 & -1 \\ -9 & 51 & -9 \\ -1 & -9 & -1 \end{bmatrix} \quad \text{Equação 61}$$

Tabela 12 – Resultados do teste de correspondência para o filtro unsharp.

Imagem	AUC		
	HBWD	CS-LBP	WLD
1	1,0000	0,9924	0,9439
2	0,9930	0,8729	0,9291
3	0,9962	0,8943	0,8671
4	1,0000	0,9962	0,9886
5	1,0000	0,9636	0,9074
6	1,0000	0,8491	0,6650
AUC Média	0,9982	0,9281	0,8835

Fonte: autoria própria.

Os melhores resultados foram obtidos utilizando-se o descritor proposto HBWD, quando comparado com os descritores CS-LBP e WLD. E ainda, ressalta-se que a AUC Média obtida para o descritor proposto foi alta, igual a 0,9982, o que demonstra que o seu desempenho foi muito pouco afetado para o uso do filtro *unsharp*.

5.3.3.3 Filtro Motion

Neste experimento foi utilizado um filtro *motion* a fim de simular o movimento linear da câmera durante a captura da imagem: o movimento simulado corresponde ao deslocamento de $N_{DPixels}$ em uma determinada direção $\theta_{DPixels}$. O algoritmo empregado para o filtro *motion* corresponde ao mesmo algoritmo da função *fspecial* do software Matlab, para o filtro *motion* (FSPECIAL MOTION, 2018). O experimento foi realizado considerando um deslocamento de $N_{DPixels} = 21$ pixels na direção $\theta_{DPixels} = 45^\circ$: a Figura 54 apresenta as imagens originais e as imagens filtradas.

Figura 54 – Filtro motion: (a) imagens originais, (b) imagens filtradas.

Fonte: autoria própria.

O teste de correspondência foi realizado utilizando-se a mesma configuração do descritor proposto HBWD empregada no teste com adição de ruído Gaussiano e a Tabela 13 resume os resultados experimentais, os quais demonstram que o melhor resultado global foi obtido empregando-se o descritor proposto HBWD. A AUC Média usando o descritor proposto foi ligeiramente superior àquela usando o descritor WLD e obteve um resultado ainda melhor quando comparado com o descritor CS-LBP.

Tabela 13 – Resultados do teste de correspondência para o filtro motion.

Imagem	AUC		
	HBWD	CS-LBP	WLD
1	1,0000	0,8975	0,8422
2	0,7641	0,6882	0,8139
3	0,8847	0,8234	1,0000
4	0,8582	0,7087	0,8348
5	1,0000	0,9487	0,9500
6	0,8922	0,9790	0,9482
AUC Média	0,8999	0,8409	0,8982

Fonte: autoria própria.

5.3.4 Transformação de Escala

Neste teste foi avaliado o desempenho do descritor proposto HBWD quanto à escala das imagens, variando-se os tamanhos das imagens em seis níveis: 0,7, 0,8, 0,9, 1,1, 1,2 e 1,3. Os tamanhos foram variados utilizando-se o método de interpolação bicúbica. O descritor proposto HBWD foi empregado na mesma configuração utilizada no teste de adição de ruído Gaussiano e a Tabela 14 resume os resultados obtidos no experimento.

Os resultados obtidos para todos os descritores foram muito parecidos e pode-se verificar que uma variação na escala em torno de 10% praticamente não afeta o desempenho da AUC Média, para todos os descritores. E ainda, uma variação de 20% na escala afeta muito pouco a AUC Média para o HBWD (menos que 5%).

Tabela 14 – Resultado do teste de correspondência de imagem com a transformação de escala nas imagens.

Imagem	Método	AUC para diversas Escalas					
		70%	80%	90%	110%	120%	130%
1	HBWD	0,9848	1,0000	1,0000	1,0000	1,0000	0,8070
	CS-LBP	0,8115	0,9520	0,9930	1,0000	1,0000	0,8343
	WLD	0,8387	0,9848	1,0000	1,0000	0,9742	0,7599
2	HBWD	0,6633	0,9102	0,9930	1,0000	0,9333	0,7467
	CS-LBP	0,6654	0,9177	0,9924	1,0000	0,9479	0,7059
	WLD	0,4913	0,9028	1,0000	1,0000	0,9145	0,7381
3	HBWD	0,7420	0,9886	1,0000	1,0000	0,9962	0,9924
	CS-LBP	0,4767	0,8973	1,0000	1,0000	0,9287	0,8499
	WLD	0,6345	0,9537	1,0000	1,0000	0,9886	0,9374
4	HBWD	0,4661	0,9121	1,0000	1,0000	0,9886	0,8389
	CS-LBP	0,6744	0,9184	1,0000	1,0000	0,9708	0,8283
	WLD	0,8509	0,9795	1,0000	1,0000	0,9924	0,9470
5	HBWD	0,7823	0,9666	1,0000	0,9197	0,9741	0,8233
	CS-LBP	0,7328	0,8706	1,0000	0,9188	0,9930	0,9586
	WLD	0,7721	0,9811	1,0000	0,9171	0,9714	0,8955
6	HBWD	0,9080	0,9769	1,0000	1,0000	0,9276	0,7321
	CS-LBP	0,7269	0,9858	1,0000	1,0000	0,9962	0,8885
	WLD	0,9344	0,9930	1,0000	1,0000	0,9612	0,9066
		AUC Média					
		70%	80%	90%	110%	120%	130%
	HBWD	0,7578	0,9591	0,9988	0,9866	0,9700	0,8234
	CS-LBP	0,6813	0,9236	0,9976	0,9865	0,9728	0,8443
	WLD	0,7537	0,9658	1,0000	0,9862	0,9671	0,8641

Fonte: autoria própria.

Cabe ressaltar que o descritor proposto HBWD não é invariante à escala como ocorre com os descritores SIFT: para tanto, o descritor proposto poderia ser modificado a fim de se utilizar uma abordagem invariante à escala, fazendo uso de pontos chave analogamente ao que é feito com o SIFT, por exemplo. Neste trabalho o descritor proposto HBWD será utilizado densamente, ou seja, não será utilizado para descrever um ponto chave, mas sim uma região por meio da técnica de *dense sampling*.

5.4 Reconhecimento Facial

Os parâmetros do método de reconhecimento facial proposto e descrito na seção 3.5 foram primeiramente ajustados para o conjunto de imagens que constituem o *view 1* da base LFW. Em seguida, o método proposto foi avaliado no conjunto *view 2* da base LFW e os resultados obtidos foram comparados com aqueles encontrados na literatura. A mesma configuração ajustada com as imagens da base LFW foi também avaliada empregando-se as imagens das bases FGLFW e BioID e seus resultados são discutidos nesta seção.

5.4.1 Resultados na Base LFW

Em primeiro lugar foi realizado o ajuste dos parâmetros do método proposto para reconhecimento facial: este procedimento foi feito utilizando-se o conjunto de imagens conhecido como *view 1* da base LFW. A configuração inicial utilizada nos testes foi oriunda dos testes de correspondência de imagens apresentado na seção 5.3 e sintetizado na Tabela 15.

Tabela 15 – Configuração inicial, baseada no teste de correspondência de imagens.

Config.	MHBWLF				HBWD				k-Means++	
	N x N	Beta	R	NV	Np x Np	Nc x Nc	Nbins	THIL	k-Clusters	Map.
0	13 x 13	0,30	1	8	12 x 12	4 x 4	8	0,20	32	Duplo

Fonte: autoria própria.

O primeiro parâmetro ajustado foi o tamanho do filtro pela média da vizinhança, do MHBWLF: este parâmetro foi variado de 5×5 até 15×15 , conforme ilustrado na Tabela 16, com todos os demais parâmetros mantidos constantes. A configuração 3, com $N \times N = 11 \times 11$, foi a que apresentou a melhor acurácia. Desta forma, manteve-se $N \times N = 11 \times 11$ e foi feito o ajuste da inclinação da curva sigmoide, β , conforme ilustrado na Tabela 17.

Tabela 16 – Ajuste de N x N do MHBWLF, na base LFW.

Config.	MHBWLF				HBWD				k-Means++		Acurácia (%)
	N x N	Beta	R	NV	Np x Np	Nc x Nc	Nbins	THIL	k-Clusters	Map.	
1	15 x 15	0,30	1	8	12 x 12	4 x 4	8	0,20	32	Duplo	85,4%
2	13 x 13	0,30	1	8	12 x 12	4 x 4	8	0,20	32	Duplo	88,5%
3	11 x 11	0,30	1	8	12 x 12	4 x 4	8	0,20	32	Duplo	88,6%
4	9 x 9	0,30	1	8	12 x 12	4 x 4	8	0,20	32	Duplo	87,9%
5	5 x 5	0,30	1	8	12 x 12	4 x 4	8	0,20	32	Duplo	79,5%

Fonte: autoria própria.

Tabela 17 – Ajuste de β do MHBWLF, na base LFW.

Config.	MHBWLF				HBWD				k-Means++		Acurácia (%)
	N x N	Beta	R	NV	Np x Np	Nc x Nc	Nbins	THIL	k-Clusters	Map.	
6	11 x 11	0,10	1	8	12 x 12	4 x 4	8	0,20	32	Duplo	86,1%
3	11 x 11	0,30	1	8	12 x 12	4 x 4	8	0,20	32	Duplo	88,6%
7	11 x 11	1,00	1	8	12 x 12	4 x 4	8	0,20	32	Duplo	84,8%

Fonte: autoria própria.

A configuração 3, com $\beta = 0,30$, foi a que apresentou a melhor acurácia. Em seguida realizou-se um teste para avaliar os parâmetros R e NV do MHBWLF, conforme ilustrado na Tabela 18. Os resultados demonstram que a configuração 3, com $R = 1$ e $NV = 8$, obtiveram a melhor acurácia. Desta forma, todos os parâmetros da configuração 3 para o MHBWLF foram fixados nos experimentos seguintes.

Tabela 18 – Ajuste de R e NV do MHBWLF, na base LFW.

Config.	MHBWLF				HBWD				k-Means++		Acurácia (%)
	N x N	Beta	R	NV	Np x Np	Nc x Nc	Nbins	THIL	k-Clusters	Map.	
3	11 x 11	0,30	1	8	12 x 12	4 x 4	8	0,20	32	Duplo	88,6%
8	11 x 11	0,30	2	8	12 x 12	4 x 4	8	0,20	32	Duplo	87,7%
9	11 x 11	0,30	2	4	12 x 12	4 x 4	8	0,20	32	Duplo	85,8%

Fonte: autoria própria.

O próximo experimento realizado procurou ajustar o limiar de iluminação do HBWD, T_{H_IL} , conforme ilustrado na Tabela 19. E novamente, a configuração 3 foi a que apresentou melhor resultado, com $T_{H_IL} = 0,20$. Posteriormente, foi feito um ajuste da quantidade de células do HBWD, $N_C \times N_C$, conforme os resultados apresentados na Tabela 20.

Tabela 19 – Ajuste de THIL do HBWD, na base LFW.

Config.	MHBWLF				HBWD				k-Means++		Acurácia (%)
	N x N	Beta	R	NV	Np x Np	Nc x Nc	Nbins	THIL	k-Clusters	Mapeamento	
10	11 x 11	0,30	1	8	12 x 12	4 x 4	8	0,10	32	Duplo	87,9%
3	11 x 11	0,30	1	8	12 x 12	4 x 4	8	0,20	32	Duplo	88,6%
11	11 x 11	0,30	1	8	12 x 12	4 x 4	8	0,40	32	Duplo	88,3%

Fonte: autoria própria.

Tabela 20 – Ajuste de Nc x Nc do HBWD, na base LFW.

Config.	MHBWLF				HBWD				k-Means++		Acurácia (%)
	N x N	Beta	R	NV	Np x Np	Nc x Nc	Nbins	THIL	k-Clusters	Mapeamento	
3	11 x 11	0,30	1	8	12 x 12	4 x 4	8	0,20	32	Duplo	88,6%
12	11 x 11	0,30	1	8	12 x 12	3 x 3	8	0,20	32	Duplo	88,8%
13	11 x 11	0,30	1	8	12 x 12	2 x 2	8	0,20	32	Duplo	86,3%

Fonte: autoria própria.

Em acordo com os resultados apresentados na Tabela 20, a configuração 12, com $N_c \times N_c = 3 \times 3$ células, foi a que produziu uma melhor acurácia. Quando comparado com a configuração 3, que possui $N_c \times N_c = 4 \times 4$ células, verifica-se ainda uma redução de 7 células, o que impacta diretamente em uma redução de tempo de processamento. Com o número de células fixado, conforme a configuração 12, realizou-se um teste a fim de melhor ajustar a quantidade de pixels em cada célula, $N_p \times N_p$, com os resultados obtidos apresentados na Tabela 21.

Tabela 21 – Ajuste de Np x Np do HBWD, na base LFW.

Config.	MHBWLF				HBWD				k-Means++		Acurácia (%)
	N x N	Beta	R	NV	Np x Np	Nc x Nc	Nbins	THIL	k-Clusters	Mapeamento	
14	11 x 11	0,30	1	8	4 x 4	3 x 3	8	0,20	32	Duplo	83,0%
15	11 x 11	0,30	1	8	6 x 6	3 x 3	8	0,20	32	Duplo	88,6%
16	11 x 11	0,30	1	8	8 x 8	3 x 3	8	0,20	32	Duplo	90,9%
17	11 x 11	0,30	1	8	10 x 10	3 x 3	8	0,20	32	Duplo	89,5%
12	11 x 11	0,30	1	8	12 x 12	3 x 3	8	0,20	32	Duplo	88,8%

Fonte: autoria própria.

Os resultados obtidos e apresentados na Tabela 21 indicam que a configuração 16, que utiliza $N_p \times N_p = 8 \times 8$ pixels por célula, obteve a melhor acurácia, sendo igual a 90,9%. Esta configuração, quando comparada com a anterior (configuração 12), também possui uma redução no tempo de processamento, pois apresenta uma redução do número de pixels por célula. Por

fim, foi ajustado o número de *bins* do histograma do HBWD, N_{bins} , conforme os resultados apresentado na Tabela 22.

Tabela 22 – Ajuste de Nbins do HBWD, na base LFW.

Config.	MHBWLF				HBWD				k-Means++		Acurácia (%)
	N x N	Beta	R	NV	Np x Np	Nc x Nc	Nbins	THIL	k-Clusters	Map.	
18	11 x 11	0,30	1	8	8 x 8	3 x 3	4	0,20	32	Duplo	90,9%
16	11 x 11	0,30	1	8	8 x 8	3 x 3	8	0,20	32	Duplo	90,9%

Fonte: autoria própria.

A configuração 18, com $N_{bins} = 4$ apresentou o mesmo resultado que a configuração 16, com $N_{bins} = 8$: ambas configurações apresentaram uma acurácia igual a 90,9%. No entanto, na configuração 18 o vetor de características final do HBWD possui 36 elementos, contra 72 elementos da configuração 16. Desta forma, a configuração 18 foi escolhida para ser mantida nos testes seguintes, pois produz um vetor de características menor.

O próximo teste realizado avaliou a variação na quantidade de *clusters* agrupados pelo método de k-Means++, $k_{clusters}$, conforme os resultados apresentados na Tabela 23. Dos resultados obtidos verifica-se que a melhor acurácia foi obtida com $k_{clusters} = 32$ *clusters*, conforme a configuração 18.

Tabela 23 – Ajuste de *k-Clusters*, na base LFW.

Config.	MHBWLF				HBWD				k-Means++		Acurácia (%)
	N x N	Beta	R	NV	Np x Np	Nc x Nc	Nbins	THIL	k-Clusters	Map.	
19	11 x 11	0,30	1	8	8 x 8	3 x 3	4	0,20	8	Duplo	75,8%
20	11 x 11	0,30	1	8	8 x 8	3 x 3	4	0,20	16	Duplo	85,6%
21	11 x 11	0,30	1	8	8 x 8	3 x 3	4	0,20	24	Duplo	90,3%
18	11 x 11	0,30	1	8	8 x 8	3 x 3	4	0,20	32	Duplo	90,9%
22	11 x 11	0,30	1	8	8 x 8	3 x 3	4	0,20	40	Duplo	89,6%
23	11 x 11	0,30	1	8	8 x 8	3 x 3	4	0,20	64	Duplo	90,1%

Fonte: autoria própria.

Posteriormente foi feita uma análise no mapeamento final. O mapeamento duplo consiste naquele onde os descritores HBWD extraídos da ROI-1 são mapeados nos *clusters* da ROI-2 e vice-versa. Por outro lado, no mapeamento simples apenas os descritores HBWD extraídos da ROI-1 são mapeados nos

clusters da ROI-2. Os resultados estão ilustrados na Tabela 24, onde verifica-se que o mapeamento duplo apresentou uma melhor acurácia.

Tabela 24 – Ajuste do mapeamento, na base LFW.

Config.	MHBWLF				HBWD				k-Means++		Acurácia (%)
	N x N	Beta	R	NV	Np x Np	Nc x Nc	Nbins	THIL	k-Clusters	Map.	
18	11 x 11	0,30	1	8	8 x 8	3 x 3	4	0,20	32	Duplo	90,9%
24	11 x 11	0,30	1	8	8 x 8	3 x 3	4	0,20	32	Simples	90,1%

Fonte: autoria própria.

Desta forma, todos os parâmetros do método de reconhecimento facial proposto foram ajustados, conforme sintetizado na Tabela 25, que apresenta a acurácia obtida no conjunto de imagens *view 1* da base LFW.

Tabela 25 – Configuração final ajustada, na base LFW.

Config.	MHBWLF				HBWD				k-Means++		Acurácia (%)
	N x N	Beta	R	NV	Np x Np	Nc x Nc	Nbins	THIL	k-Clusters	Map.	
18	11 x 11	0,30	1	8	8 x 8	3 x 3	4	0,20	32	Duplo	90,9%

Fonte: autoria própria.

A configuração ajustada foi então testada no conjunto *view 2* da base LFW e os resultados obtidos são apresentados e comparados com outros resultados encontrados na literatura, conforme ilustrado na Tabela 26. Cabe ressaltar aqui que os resultados apresentados na Tabela 26 foram obtidos no modo restrito aos pares de treinamento e teste fornecidos pelo *view 2* da base LFW, sem a utilização dos *labels* das faces (nomes dos arquivos que representam as faces na base LFW) para formar novos pares de treinamento ou teste e, ainda, sem utilizar nenhuma outra imagem externa.

As *eigenfaces* (TURK; PENTLAND, 1991) obtiveram uma acurácia de 60,0% neste modo de teste da base LFW. Uma abordagem utilizando múltiplos *kernels* de aprendizado obteve uma acurácia 68,2% quando codificando uma representação simplificada de pixels (Pixels/MKL) e 79,4% usando um modelo V1-like (PINTO; DICARLO; COX, 2009).

Tabela 26 – Resultados do método proposto e de outros métodos encontrados na literatura para a base LFW, no modo restrito e sem utilizar imagens externas.

Método	Acurácia (%)
(TURK; PENTLAND, 1991)	60,0%
(PINTO; DICARLO; COX, 2009) - Pixels/MKL	68,2%
(NOWAK; JURIE, 2007) - original	72,5%
(SANDERSON; LOVELL, 2009)	73,0%
(NOWAK; JURIE, 2007) - funneled	73,9%
(AFFONSO; RODRIGUES, 2015)	75,9%
(WOLF; HASSNER; TAIGMAN, 2008)	78,5%
(ARASHLOO; KITTLER, 2013)	79,1%
(PINTO; DICARLO; COX, 2009) - V1-like/MKL	79,4%
(LI et al., 2013)	84,1%
(SIMONYAN et al., 2013)	87,5%
(JUEFEI-XU; LUU; SAVVIDES, 2015)	87,6%
(SAGONAS et al., 2016)	88,8%
(LI et al., 2014)	89,0%
Método Proposto	90,4%
(LI; HUA, 2015)	91,1%

Fonte: adaptado de (Labeled Faces in the Wild – Results, 2018).

Nowak e Jurie (2007) introduziram uma nova medida de similaridade, a qual foi avaliada no conjunto de imagens original da base LFW e alcançou uma acurácia de 72,5%. Este método também foi testado no conjunto de imagens *funneled* da base LFW, um conjunto já previamente alinhado, e obteve uma acurácia de 73,9%.

Um histograma de probabilidade de múltiplas regiões foi empregado em (SANDERSON; LOVELL, 2009), atingindo uma acurácia de 73,0%. Affonso e Rodrigues (2015) utilizaram o WLD em uma grade regular, sem nenhum alinhamento prévio, em conjunto com SVM, e obtiveram uma acurácia de 75,9%. Em (WOLF; HASSNER; TAIGMAN, 2008) é apresentado um método baseado no aprendizado de similaridades utilizando vários descritores, alcançando uma acurácia de 78,5%. Já Arashloo e Kittler (2013) utilizaram a modelagem MRF (do inglês, *Markov Random Field*) em conjunto com o descritor MLBP para a extração das características, e a acurácia obtida foi de 79,1%.

Li et al. (2013) usaram os descritores LBP e SIFT extraídos densamente e empregados em um modelo de probabilidade elástica, atingindo uma acurácia de 84,1% no conjunto de imagens alinhadas *funneled* da base LFW. Simonyan et al.

(2013) empregaram somente descritores baseados em SIFT, também extraídos densamente, obtendo uma acurácia de 87,5%.

Em (JUEFEI-XU; LUU; SAVVIDES, 2015) a região periocular é utilizada em um novo framework proposto baseado em um modelo elástico 3D, alcançando uma acurácia de 87,6%. Sagonas et al. (2016) empregaram um método estatístico chamado RSF (do inglês, *Robust Statistical Frontalization*) e obtiveram uma acurácia de 88,8%. Li et al. (2014) usaram um modelo elástico probabilístico em conjunto com a técnica PCA, atingindo uma acurácia de 89,0%.

O método proposto nesta tese de doutorado, empregando o novo HBWD densamente na região periocular e fazendo a redução da dimensão dos dados pelo uso da técnica k-Means++ de agrupamento em um algoritmo de mapeamento, obteve uma acurácia de 90,4% no conjunto *view 2* da base LFW. Foram utilizadas as imagens originais da base LFW, sem nenhum alinhamento prévio, pois o próprio método proposto realiza o alinhamento das faces a partir da localização precisa dos centros dos olhos, o que é feito por meio do uso do MHBWLF em conjunto com um filtro MOSSE.

Cabe destacar ainda que Li e Hua (2015) propuseram um novo método chamado de *Hierarchical-PEP*, onde uma face é decomposta em várias camadas hierárquicas: considerando apenas uma camada os autores obtiveram uma acurácia de 87,2%; com duas camadas a acurácia foi de 90,0%; com três camadas de 90,4%; e, por fim, empregando um algoritmo de fusão a acurácia final foi de 91,4%. Os resultados apresentados por Li e Hua (2015) foram ligeiramente superiores aos obtidos pelo método proposto, no entanto, o método proposto é relativamente mais simples, o que gera a possibilidade de trabalhos futuros a fim de incrementar o método proposto de modo a melhorar o seu desempenho.

5.4.2 Resultados na Base FGLFW

O método proposto também foi avaliado na base FGLFW, utilizando a mesma configuração ajustada na avaliação anterior na base LFW, e os resultados estão resumidos na Tabela 27, na qual observa-se que, exceto para o método proposto, todos os outros métodos foram avaliados no modo irrestrito, com a utilização de

imagens externas. Não foram encontrados na literatura resultados de outros métodos no modo restrito, sem a utilização de imagens externas.

Tabela 27 – Resultados do método proposto e de outros métodos encontrados na literatura para a base FGLFW.

Método	Modo de Teste	Quantidade de Imagens	Acurácia (%)		
			LFW	FGLFW	Redução
DeepFace (TAIGMAN, 2014)	Irrestrito, com imagens externas	500.000	92,9%	78,8%	14,1%
DeepID2 (SUN; WANG; TANG, 2014)	Irrestrito, com imagens externas	200.000	95,0%	78,3%	16,7%
VGG-Face (PARKHI; VEDALDI; ZISSERMAN, 2015)	Irrestrito, com imagens externas	2.600.000	96,7%	85,8%	10,9%
DCMN (DENG ET AL., 2017)	Irrestrito, com imagens externas	500.000	98,0%	91,0%	7,0%
Noisy Softmax (CHEN; DENG; DU, 2017)	Irrestrito, com imagens externas	500.000	99,2%	94,5%	4,7%
Método Proposto	Restrito, sem imagens externas	12.000	90,4%	87,9%	2,5%

Fonte: adaptado de (DENG et al., 2017).

Analisando-se os resultados dos diversos métodos apresentados na Tabela 27, verifica-se que o método proposto foi aquele que apresentou uma menor queda na acurácia do teste realizado na base LFW para o teste da base FGLFW, demonstrando claramente que o método proposto foi capaz de generalizar muito bem a solução encontrada na base LFW quando aplicado na base FGLFW. A redução obtida foi de apenas 2,5% no método proposto, sendo quase a metade da redução sofrida para o método de melhor acurácia (Noisy, com redução de 4,7%).

O método proposto foi avaliado utilizando-se apenas 12.000 imagens, enquanto os outros métodos foram avaliados com muito mais imagens. Por exemplo, o método VGG-Face (PARKHI; VEDALDI; ZISSERMAN, 2015) foi avaliado utilizando-se 2.600.000 imagens. De fato, os métodos apresentados na Tabela 27 fazem uso de redes neurais profundas, que normalmente necessitam de muitas imagens para o aprendizado.

Pode-se notar nos resultados apresentados na Tabela 27 que a base FGLFW realmente representa um maior desafio que a base LFW, onde geralmente o desempenho obtido na base LFW cai consideravelmente na base FGLFW. Por exemplo, o método DeepID2 (SUN et al., 2014) atingiu uma acurácia de 95,0% na base LFW, a qual caiu para apenas 78,3% na base FGLFW.

O método proposto obteve uma acurácia de 87,9% na base FGLFW, no entanto, destaca-se que este método foi o único avaliado no modo restrito, quando não é permitido o uso dos *labels* das faces para formar novos pares de teste e treinamento. E ainda, o método proposto foi avaliado sem a inclusão de imagens

externas, utilizando apenas 12.000 imagens, enquanto os demais métodos empregaram uma quantidade muito superior de imagens, o que torna difícil uma comparação direta dos resultados.

No entanto, cabe destacar que, mesmo sendo avaliado no modo restrito e utilizando bem menos imagens que os outros métodos comparados na Tabela 27, o método proposto apresentou um bom desempenho, sendo superior a diversos outros métodos tidos como estado da arte na literatura.

5.4.3 Resultados na Base BioID

O método proposto para reconhecimento facial também foi avaliado na base BioID, empregando os pares de imagens apresentados no apêndice A. Durante este experimento foi utilizada a mesma configuração ajustada durante os testes realizados com a base LFW. O método proposto atingiu uma acurácia de 96,8%.

Não foi possível comparar o resultado obtido para a base BioID neste modo de teste, baseado no problema do *pair matching*, com o resultado de outros autores, pois a base BioID não fornece os pares de teste para este tipo de problema. De fato, estes pares de imagens, apresentados no apêndice A, foram elaborados neste trabalho de doutorado, gerando a possibilidade de que outros autores possam utilizá-lo futuramente.

Cabe destacar aqui que o método proposto apresentou uma acurácia maior para os pares de imagens da base BioID, quando comparados com os pares das bases LFW e FGLFW. De fato, as imagens da base LFW, que são as mesmas imagens utilizadas pela base FGLFW, possuem muito mais variações do que as imagens da base BioID. Isto pode ser facilmente observado após uma análise de exemplos de imagens extraídas de tais bases como, por exemplo, ilustram a Figura 31, Figura 33 e Figura 34. As imagens da base LFW apresentam grandes variações em poses e expressões faciais, no ângulo de captura das fotos, no uso de acessórios e maquiagens, etc. No entanto, é importante ainda destacar que os pares de imagens de teste aqui utilizados, e apresentados no apêndice A, constitui de um novo conjunto para que outros autores possam também avaliar seus trabalhos na base BioID dentro do problema baseado no *pair matching*.

5.5 Considerações Finais

Neste capítulo foram apresentados os experimentos realizados e seus resultados foram discutidos detalhadamente. Os experimentos foram divididos em quatro partes principais: filtros HBWLF e MHBWLF; localização precisa dos olhos; correspondência de imagens com o HBWD; e reconhecimento facial.

O primeiro experimento realizado caracterizou os filtros HBWLF e MHBWLF. Quanto maior o tamanho do filtro pela média da vizinhança, maior é o tamanho das bordas realçadas, e vice-versa. A inclinação da curva sigmoide, β , produz um efeito direto na suavização da imagem filtrada: quanto maior o valor de β , mais suave será a imagem filtrada e, por outro lado, um valor menor acarreta em um maior realce das componentes de alta frequência da imagem, o que acaba também por provocar um aumento do conteúdo de ruído presente na imagem original. O valor de β deve ser ajustado de modo a realçar os detalhes desejados sem, no entanto, aumentar demasiadamente o ruído na imagem. O raio R provoca um deslocamento da borda realçada, bem como o seu alargamento e a quantidade de vizinhos amostrados, NV , influencia diretamente na continuidade das bordas detectadas.

Foi destacado também que o HBWLF é um caso particular do MHBWLF, o qual é obtido fazendo-se $R = 1$ e $NV = 8$. Por fim, o MHBWLF foi comparado com a excitação diferencial do WLD, verificando-se que o MHBWLF é capaz de realçar mais detalhes na imagem filtrada. De fato, os ajustes possíveis no MHBWLF possibilitam também que o mesmo seja melhor configurado para uma determinada aplicação.

O segundo experimento avaliou o método proposto para a localização precisa dos centros dos olhos. O método proposto inova ao empregar o MHBWLF para realçar os detalhes das imagens e, em conjunto com um filtro MOSSE, localizar com precisão os centros dos olhos. O método proposto foi avaliado nas bases LFW e BioID e os resultados obtidos foram melhores que vários métodos que representam o estado da arte em localização dos centros dos olhos. Os resultados apresentados nesta tese também foram publicados no artigo (AFFONSO; RODRIGUES; PAIVA, 2018).

O terceiro experimento realizado foi um teste de correspondência de imagens baseado no descritor proposto HBWD. Para este teste foram selecionadas

aleatoriamente seis imagens faciais da base BioID e empregada uma grade com onze pontos distribuídos ao redor da região periocular de cada face. O teste foi realizado aplicando-se as seguintes transformações: adição de ruído (Gaussiano, *Speckle* e *Salt & Pepper*); rotação (5°, 10°, 15°, 20°, 25° e 30°); uso de filtros (*average*, *unsharp* e *motion*); e escala (0,7, 0,8, 0,9, 1,1, 1,2 e 1,3). O descritor proposto HBWD foi comparado com os descritores WLD e CS-LBP, todos eles amostrados nos pontos da grade regular pré-definida no experimento, sem nenhum artifício de invariância quanto à escala e rotação.

Para os testes de adição de ruído foram geradas 50 amostras, para cada tipo de ruído diferente. O HBWD foi ajustado inicialmente para o teste de adição de ruído gaussiano e a sua configuração (parâmetros ajustados) foi mantida constante para todos os outros testes de correspondência de imagens. Para o teste de adição de ruído do tipo gaussiano o desempenho do HBWD foi muito superior, com uma $AUC = 0,8455$, contra $AUC = 0,6691$ para o descritor CS-LBP e $AUC = 0,6277$ para o WLD. Para a adição de ruído do tipo *Speckle* e *Salt & Pepper*, o desempenho do HBWD foi ligeiramente superior aos demais com $AUC = 0,9603$ e $AUC = 0,9695$, respectivamente.

Para o teste de rotação os resultados obtidos foram muito próximos para os três descritores testados e a AUC média ficou praticamente constante e próxima do valor unitário para uma rotação de até 15°. Cabe ressaltar aqui que em nenhum método testado foi utilizada alguma técnica de invariância quanto à rotação, pois neste trabalho o descritor desenvolvido deverá ser aplicado em faces já anteriormente alinhadas dentro do próprio método proposto para o reconhecimento facial.

Os testes quanto aos filtros *average*, *unsharp* e *motion* tentam simular situações reais de captura de imagens faciais nas mais diversas condições do dia a dia e por diversos equipamentos diferentes. O HBWD obteve o melhor desempenho, quando comparado com os outros dois descritores nos testes apresentados. No filtro *average*, o HBWD apresentou uma $AUC = 1,0000$, contra $AUC = 0,9663$ do CS-LBP e $AUC = 0,9724$ do WLD. Já no teste com o filtro *unsharp*, o HBWD apresentou uma $AUC = 0,9982$, contra $AUC = 0,9281$ do CS-LBP e $AUC = 0,8835$ do WLD. E, por último, no teste com o filtro *motion*, o HBWD apresentou uma $AUC = 0,8999$, contra $AUC = 0,8409$ do CS-LBP e $AUC = 0,8982$ do WLD.

Por fim, o teste de correspondência quanto à variação de escala apresentou resultados muito parecidos para os três descritores, verificando-se que uma variação na escala em torno de 10% praticamente não afeta o desempenho da AUC média, para todos os três descritores. E ainda, uma variação de 20% na escala afeta muito pouco a AUC Média para o HBWD (menos que 5%). Este resultado sugere, no entanto, que o emprego de uma técnica de invariância quanto à escala pode contribuir para melhorar o desempenho da aplicação do HBWD em problemas de reconhecimento facial.

O quarto e último experimento realizado foi a avaliação do método proposto para o reconhecimento facial. Para esta avaliação foram empregadas três bases: LFW, FGLFW e BioID. A base LFW é considerada uma das bases com imagens mais desafiadoras para o estudo do problema de reconhecimento facial, com imagens com grandes variações encontradas no mundo real (variações de poses e expressões, iluminação, uso de acessórios, etc.). A base FGLFW emprega as mesmas imagens da base LFW, porém modifica os pares negativos (*mismatched*) da base LFW, apresentando assim um conjunto com pares de imagens muito mais difíceis de serem classificadas. Por fim, a base BioID também foi empregada e para isto foi gerada neste trabalho uma lista com os pares de imagens a serem considerados (apêndice A).

Primeiramente foi feito um ajuste de todos os parâmetros do método proposto (MHBWLF, HBWD e Mapeamento com k-Means++) utilizando as imagens do conjunto *view 1* da base LFW. A configuração ajustada foi então testada no conjunto *view 2* da base LFW, evitando assim um *overfitting*, o que ocorreria caso o método proposto fosse especializado manualmente apenas para ter uma melhor resposta no conjunto de imagens *view 1*. Desta forma, a acurácia obtida foi de 90,4%, sendo maior que a acurácia obtida por diversos outros métodos estado da arte que empregam outros descritores como, por exemplo, SIFT (SIMONYAN et al., 2013) e LBP (LI et al., 2013). Cabe ressaltar aqui que os valores reportados para a base LFW foram todos comparados no modo restrito de teste, quando não é permitido usar o nome dos arquivos para compor novos pares de testes, e sem a adição de imagens externas.

O método proposto foi também avaliado na base FGLFW e a acurácia obtida foi de 87,9%. A queda na acurácia de 90,4% na base LFW para 87,9% na base FGLFW foi relativamente pequena e demonstra que a base FGLFW realmente

possui pares de imagens com maior desafio que a base LFW. O resultado obtido pelo método proposto foi melhor que o resultado de alguns métodos estado da arte, no entanto, esta comparação não é totalmente justa, pois o método testado foi avaliado com apenas 12.000 imagens, no modo restrito e sem adição de imagens externas, enquanto todos os outros métodos foram avaliados no modo irrestrito e com adição de imagens externas. Ou seja, os outros métodos utilizaram muito mais imagens durante seu processo de treinamento e teste, o que pode lhes conferir alguma vantagem adicional. Ressalta-se ainda que não foram encontrados na literatura outros métodos que foram avaliados na base FGLFW no modo restrito e sem usar imagens externas, como fora feito neste trabalho, impossibilitando assim uma melhor comparação dos resultados obtidos.

O método proposto foi ainda avaliado nos pares de imagens da base BioID introduzidos no apêndice A desta tese e a acurácia obtida foi de 96,8%. Este resultado demonstra que os pares de imagens da base BioID apresentados no apêndice A foram classificados mais facilmente que aqueles pares dados nas bases LFW e FGLFW. Ressalta-se que não foram encontrados na literatura outros resultados na base BioID que permitissem uma comparação com os resultados aqui apresentados e que a disponibilização do apêndice A nesta tese torna possível que novos trabalhos surjam e possam ser comparados.

Cabe ressaltar que uma limitação dos resultados apresentados quanto ao método de reconhecimento facial proposto foi na abordagem de ajuste dos seus parâmetros, de tal modo que não é possível garantir que foi atingido o melhor desempenho do método proposto para estas mesmas bases de teste. O método foi todo ajustado para a base LFW, ajustando um parâmetro por vez: primeiramente foram ajustados os parâmetros do MHBWLF, depois do HBWD e, por fim, do algoritmo de mapeamento com k-Means++. Contudo, esta abordagem não garante que foi encontrada uma solução ótima para o método proposto. Ressalta-se ainda que esta abordagem partiu de uma configuração pré-ajustada nos experimentos realizados para o teste de correspondência de imagens do HBWD, o que permitiu que a abordagem empregada fizesse um ajuste um pouco mais fino nos diversos parâmetros sem, no entanto, garantir que foi encontrada a solução ótima de ajuste dos mesmos.

6 CONCLUSÕES

O principal objetivo deste trabalho de doutorado foi o desenvolvimento de um novo método de reconhecimento facial, baseado no problema do *pair matching*, para ser aplicado em ambientes não controlados. A hipótese considerada no desenvolvimento deste trabalho foi que a utilização de descritores locais permite a extração de importantes características das faces em ambientes não controlados, possibilitando o seu reconhecimento.

Em seguida foi realizado um estudo dos principais marcos históricos no desenvolvimento de algoritmos para o reconhecimento facial ao longo os anos. Neste estudo foram observados que muitos autores empregaram diversos descritores locais em seus métodos de reconhecimento facial como, por exemplo, LBP, SIFT e WLD. Os artigos estudados demonstraram que o uso de descritores locais baseados em pontos chave apresentam problemas quando empregados em reconhecimento facial, pois uma face não é um objeto rígido, apresentando muitas variações pelo envelhecimento, uso de maquiagem e acessórios, poses e expressões faciais, etc. Desta forma, os descritores locais, quando utilizados em reconhecimento facial, geralmente não são empregados para descreverem pontos chave, mas sim uma determinada região de interesse.

Este estudo também revelou que a região periocular de uma face é aquela que apresenta as características mais discriminantes para serem usadas no reconhecimento facial, sendo menos susceptível às expressões faciais. Foi observado, por exemplo, que a região da boca é muito afetada pelas expressões faciais de uma pessoa. A fundamentação teórica realizada foi finalizada com um estudo detalhado de cinco descritores locais: LBP, SIFT, WLD, CS-LBP e CS-LMP. Este estudo foi importante para caracterizar alguns descritores utilizados como comparação nos experimentos realizados e também para servir como uma inspiração no desenvolvimento dos métodos propostos.

Desta forma, o trabalho todo foi conduzido para o desenvolvimento de um novo método de reconhecimento facial. Em um primeiro momento foram introduzidos os filtros HBWLF e MHBWLF para realçar os detalhes de uma imagem facial. O filtro MHBWLF foi então empregado em um algoritmo proposto para a localização, com grande precisão, dos centros dos olhos de uma face: a imagem facial é filtrada pelo

MHBWLF e tem seus detalhes realçados para, em seguida, um filtro MOSSE realizar a localização dos centros dos olhos com maior precisão.

Posteriormente foi proposto um novo descritor, o HBWD, o qual faz uso do MHBWLF para descrever localmente um pixel. Este descritor foi desenvolvido para ser utilizado densamente no método proposto para o reconhecimento facial. O HBWD utiliza a imagem filtrada pelo MHBWLF e reúne algumas características de outros descritores como, por exemplo, os descritores SIFT para a montagem do vetor de características.

E, por fim, foi introduzido um novo método de reconhecimento facial para comparação de duas faces (*pair matching*). Em um primeiro momento as faces comparadas são recortadas e os centros de seus olhos são detectados pelo algoritmo proposto neste trabalho. De posse dos centros dos olhos, as faces são alinhadas e uma região de interesse, ao redor da região periocular, é extraída para cada face. Cada região de interesse é descrita densamente pelo HBWD e, a fim de reduzir a dimensão dos dados, um algoritmo de mapeamento baseado na técnica de agrupamento de k-Means++ é proposto, produzindo assim uma assinatura para o par de faces de entrada. Uma máquina de vetores de suporte, SVM, é utilizada para realizar a classificação final.

Este trabalho visa o reconhecimento facial em ambientes não controlados e, para tanto, foram selecionadas três bases para serem utilizadas nos experimentos realizados: LFW, FGLFW e BioID. Estas bases foram e ainda são muito utilizadas por diversos autores e representam um grande desafio, pois possuem imagens com grandes variações de poses e expressões faciais, de iluminação, de objetos de fundo, uso de maquiagem e acessórios, etc.

Dos diversos experimentos realizados, o primeiro deles caracterizou os filtros propostos HBWLF e MHBWLF, demonstrando a grande capacidade dos mesmos em realçar detalhes das imagens, com a possibilidade de ajuste de diversos parâmetros a fim de melhor sintonizá-los para uma determinada aplicação. O MHBWLF foi comparado com o WLD, verificando que o filtro proposto é capaz de realçar mais detalhes na imagem filtrada.

O método proposto para a localização precisa dos centros dos olhos foi avaliado e os resultados obtidos demonstram que o mesmo possui um desempenho superior ao desempenho de vários métodos que representam o estado da arte neste

tipo de problema. O presente método e seus resultados foi publicado em um periódico internacional importante (AFFONSO; RODRIGUES; PAIVA, 2018).

O descritor proposto, HBWD, foi avaliado individualmente em um teste de correspondência de imagens. Para tanto foram selecionadas seis imagens da base BioID e empregada uma grade com onze pontos para cada imagem. O teste foi realizado aplicando-se as seguintes transformações: adição de ruído (Gaussiano, *Speckle* e *Salt & Pepper*), rotação, uso de filtros (*average*, *unsharp* e *motion*) e escala. O descritor proposto foi comparado com os descritores WLD e CS-LBP. Todos os parâmetros do HBWD foram ajustados uma única vez, durante o teste de adição de ruído Gaussiano e, posteriormente, mantidos constantes. Os resultados obtidos para o HBWD foram, em geral, melhores que os resultados dos descritores WLD e CS-LBP, dentro das condições do teste apresentado na seção 5.3.

O último experimento realizado avaliou o método proposto para o reconhecimento facial nas bases LFW, FGLFW e BioID. O método proposto apresentou uma acurácia de 90,4% na base LFW, para o modo irrestrito de teste, sem a adição de imagens externas. Comparando-se o resultado obtido com os resultados encontrados na literatura verifica-se que o método proposto apresenta um desempenho superior à vários outros métodos estado da arte.

A acurácia do método proposto na base FGLFW foi de 87,9%, apresentando também um melhor desempenho que outros métodos. Cabe ressaltar que não foi encontrado na literatura nenhum outro método avaliado na base FGLFW no modo irrestrito e sem a adição de imagens externas. O método proposto foi avaliado na base FGLFW utilizando apenas as 12.000 imagens propostas na base, enquanto nos outros trabalhos comparados a quantidade de imagens utilizada foi muito superior, pois os mesmos usaram muitas outras imagens externas à base: por exemplo, foram utilizadas 2.600.000 imagens no método VGG-Face (PARKHI; VEDALDI; ZISSERMAN, 2015).

O método proposto foi ainda avaliado na base BioID, empregando os pares de imagens definidos neste trabalho e apresentado no apêndice A: a acurácia obtida foi de 96,8%. Neste caso, cabe destacar que a base BioID não disponibiliza os pares de imagens para o problema do *pair matching* e, portanto, não foram encontrados outros resultados para serem comparados com o presente trabalho.

Analisando-se os resultados obtidos verifica-se que a hipótese assumida neste trabalho de doutorado é aceita, pois o emprego do descritor local proposto,

HBWD, extrai importantes características das faces em ambientes não controlados, possibilitando assim obter resultados competitivos quanto à acurácia do reconhecimento facial, mesmo quando comparados com os resultados de outros métodos que representam o estado da arte.

Desta forma, além de apresentar um novo método para o reconhecimento facial em ambientes não controlados, esta tese também introduz dois novos filtros (HBWLF e MHBWLF), um novo descritor (HBWD) e um novo método para a localização precisa dos centros dos olhos de uma face. Os experimentos foram realizados utilizando-se as imagens das bases LFW, FGFW e BioID, as quais possuem imagens que representam muito bem ambientes não controlados. Por fim, os resultados obtidos demonstram que o desempenho dos métodos propostos são superiores a muitos métodos estado da arte encontrados na literatura.

6.1 Contribuições

Nesta tese de doutorado foram propostos dois filtros, um descritor, um método de localização precisa dos centros dos olhos e um método de reconhecimento facial baseado no problema do *pair matching*:

- ✓ *High-Boost Weber Local Filter* (HBWLF): este filtro salienta os detalhes de uma imagem facial, realçando o seu conteúdo de alta frequência sem, no entanto, eliminar o seu conteúdo de baixa frequência. O HBWLF foi inspirado no WLD, porém ao invés de utilizar o pixel central utiliza a média da região vizinha ao mesmo, o que lhe confere uma maior robustez quanto ao ruído. O WLD utiliza a função arco-tangente como função de saída, enquanto no HBWLF utiliza-se a função sigmoide, o que lhe permite um melhor ajuste;
- ✓ *Multiscale High-Boost Weber Local Filter* (MHBWLF): este filtro é uma evolução do HBWLF, pois enquanto o HBWLF considera uma grade regular de tamanho fixo igual a 3x3 pixels, no MHBWLF consideram-se NV pixels distribuídos igualmente em uma circunferência de raio R ;
- ✓ Método para a localização precisa dos centros dos olhos: neste método o MHBWLF é utilizado para realçar as características das faces e um

filtro MOSSE foi utilizado para o aprendizado da localização dos centros dos olhos, utilizando as imagens filtradas pelo MHBWLF como entrada. O método foi avaliado nas bases BioID e LFW e os resultados obtidos foram superiores aos métodos estado da arte encontrados na literatura;

- ✓ *High-Boost Weber Descriptor* (HBWD): este descritor combina características dos descritores SIFT, WLD e CS-LMP. O HBWD faz uso de uma imagem filtrada pelo MHBWLF e é montado em uma grade regular com $N_C \times N_C$ células, onde para cada célula é gerado um histograma com N_{bins} bins. O vetor final de saída do HBWD contém a dimensão de $N_C \times N_C \times N_{bins}$. Diversos testes de correspondência de imagens faciais foram realizados com o HBWD e o mesmo foi comparado com os descritores WLD e CS-LBP, obtendo melhores resultados, principalmente no aspecto da robustez quanto ao ruído;
- ✓ Método proposto para reconhecimento facial baseado no problema do *pair matching*: o método proposto faz uso do algoritmo desenvolvido para a localização dos centros dos olhos a fim de realizar o alinhamento das faces de entrada, possibilitando desta forma um recorte mais preciso da região de interesse. A região de interesse utilizada no método foi a região periocular e a mesma foi descrita densamente pelo HBWD, ou seja, o HBWD foi aplicado em todos os pixels da região periocular. A fim de reduzir a dimensão dos dados, foi utilizado o método de k-Means++ em um mapeamento proposto, gerando assim uma assinatura de saída para o par de imagens de entrada em comparação. Por fim foi utilizado um classificador SVM para a decisão final (*matched* ou *mismatched*). O método proposto foi avaliado nas bases LFW, FGLFW e BioID e os resultados obtidos são superiores aos resultados de vários outros métodos estado da arte encontrados na literatura;
- ✓ A última contribuição deste trabalho consiste na elaboração de uma lista de pares de imagens para o teste baseado no problema do *pair matching* com imagens oriundas da base BioID, o que dará a possibilidade de outros autores também testarem seus métodos na base BioID e compararem com os resultados obtidos neste trabalho.

6.2 Sugestões de Trabalhos Futuros

A execução deste trabalho de doutorado apresentado nesta tese gera algumas possibilidades de novas propostas de trabalho:

- ✓ Uma limitação apresentada no método proposto para o reconhecimento facial consiste na solução utilizada para o ajuste paramétrico do método proposto, o qual não garantiu que foi encontrada a sua solução ótima. Desta forma, uma proposta de trabalho futuro seria a otimização dos diversos parâmetros do método proposto a fim de encontrar a sua solução ótima;
- ✓ Uma outra proposta de trabalho futuro seria a investigação do uso do MHBWLF em conjunto com redes neurais profundas para o desenvolvimento de um novo método de reconhecimento facial. Muitos métodos que empregam redes neurais profundas utilizam as imagens faciais puras como entrada da rede – a proposta aqui sugerida seria a investigação de se utilizar imagens faciais filtradas pelo MHBWLF como entrada para as redes neurais profundas a fim de verificar se as características ressaltadas pelo MHBWLF auxiliam na tarefa de reconhecimento facial;
- ✓ O descritor proposto HBWD foi empregado nesta tese em problemas de reconhecimento facial, com faces já alinhadas dentro do método de reconhecimento facial proposto. Além disto, o HBWD foi empregado densamente, ou seja, em todos os pixels da região de interesse. Uma proposta de trabalho futuro seria o emprego do HBWD em outros tipos de problemas de reconhecimento de padrões e utilizando outras abordagem como, por exemplo, a utilização de pontos chave;
- ✓ O método de reconhecimento facial proposto analisa as imagens faciais nas escalas originais em que são apresentadas. Sugere-se aqui a adaptação do presente método para uma versão multiescalar empregando, por exemplo, uma pirâmide de imagens filtradas pelo MHBWLF.

6.3 Lista de Artigos Publicados

Até o presente momento, este trabalho de doutorado gerou as seguintes publicações:

- **WVC (2015):**

AFFONSO, ALEX ANTONIO; RODRIGUES, EVANDRO LUÍS LINHARI . Unconstrained Face Recognition using Weber Local Descriptor and SVM. In: XI Workshop de Visão Computacional, 2015, São Carlos. XI Workshop de Visão Computacional, 2015. p. 146-151.

- **WVC (2017):**

AFFONSO, ALEX ANTONIO; RODRIGUES, EVANDRO LUÍS LINHARI . Precise eye localization using the new Multiscale High-Boost Weber Local Filter. In: XIII Workshop de Visão Computacional: anais do 13o Workshop de Visão Computacional realizado na UFRN. Natal: EDUFRN, 2018. 167 p.

- **Pattern Recognition Letters – Elsevier (2018):**

AFFONSO, ALEX ANTONIO; RODRIGUES, EVANDRO LUÍS LINHARI ; PAIVA, MARIA STELA VELUDO DE . High-Boost Weber local filter for precise eye localization under uncontrolled scenarios. PATTERN RECOGNITION LETTERS, v. 102, p. 50-57, 2018, <https://doi.org/10.1016/j.patrec.2017.12.015>

REFERÊNCIAS

AFFONSO, A. A.; RODRIGUES, E. L. L. Unconstrained Face Recognition using Weber Local Descriptor and SVM. In: XI Workshop de Visão Computacional, 2015, São Carlos. XI Workshop de Visão Computacional, p. 146-151, 2015.

AFFONSO, A. A.; RODRIGUES, E. L. L. Precise eye localization using the new Multiscale High-Boost Weber Local Filter. In: XIII Workshop de Visão Computacional: anais do 13o Workshop de Visão Computacional realizado na UFRN. Natal: EDUFRN, 2018. 167 p.

AFFONSO, A. A., RODRIGUES, E. L. L., PAIVA, M. S. High-Boost Weber Local Filter for Precise Eye Localization under Uncontrolled Scenarios. Pattern Recognition Letters, v. 102, p. 50-57, January 2018. DOI: <https://doi.org/10.1016/j.patrec.2017.12.015>.

AHONEN, T.; HADID, A.; PIETIKAINEN, M. Face Recognition with Local Binary Patterns. In: Proceedings of 8th European Conference on Computer Vision (ECCV), Prague, Czech Republic, p. 469-481, May 2004. DOI: 10.1007/978-3-540-24670-1_36.

AHONEN, T.; HADID, A.; PIETIKAINEN, M. Face Description with Local Binary Patterns: Application to Face Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 28, n. 12, p. 2037-2041, December 2006. DOI: 10.1109/TPAMI.2006.244

ARASHLOO, S. R.; KITTLER, J. Efficient Processing of MRFs for Unconstrained-Pose Face Recognition. In: Biometrics: Theory, Applications and Systems (BTAS), 2013 IEEE Sixth International Conference on, 2013. DOI: 10.1109/BTAS.2013.6712721.

ARTHUR, D., VASSILVITSKII, S. K-means++: The Advantages of Careful Seeding. In: SODA 2007: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on DISCRETE ALGORITHMS. P. 1027–1035, 2007. DOI: 10.1145/1283383.1283494

BELHUMEUR, P.; HESPAÑA, J.; KRIEGMAN, D. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 19, n. 7, p. 711-720, July 1997. DOI: 10.1109/34.598228.

BICEGO, M. et al. On the Use of SIFT Features for Face Authentication. In: Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW), p. 35, 2006. DOI: 10.1109/CVPRW.2006.149

BioID, 2017. Disponível em: < <https://www.bioid.com> >. Acesso em: 10 Janeiro 2017.
BOLME, D. S., DRAPER, B. A., BEVERIDGE, J. R. Average of Synthetic Exact Filters. IEEE Conference on Computer Vision and Pattern Recognition, p. 2105-2112, 2009. DOI 10.1109/CVPR.2009.5206701.

Bolme, D. S. et al. Visual object tracking using adaptive correlation filters. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, p. 2544-2550, 2010. DOI 10.1109/CVPR.2010.5539960.

CHEN, J. et al. A Robust Descriptor Based on Weber's Law. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p. 1-7, 2008. DOI: 10.1109/CVPR.2008.4587644.

CHEN, J. et al. WLD: A Robust Local Image Descriptor. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 32, n. 9, p. 1705-1720, August 2009. DOI: 10.1109/TPAMI.2009.155.

CHEN, B.; DENG, W.; DU, J. Noisy softmax: Improving the Generalization Ability of DCNN via Postponing the Early Softmax Saturation. In: Computer Vision and Pattern Recognition (CVPR), 2017 IEEE Conference on, 2017. DOI: 10.1109/CVPR.2017.428

DENG, W. et al. Fine-grained face verification: FGLFW database, baselines, and human-DCNN partnership. Pattern Recognition, v. 66, p. 63-73, June 2017. DOI: <https://doi.org/10.1016/j.patcog.2016.11.023>

FACEBLIND.ORG – Prosopagnosia Research Centers at Dartmouth College, Harvard University and University College London. Disponível em: <<http://www.faceblind.org/>>. Acesso em: 28/02/2015.

Ferraz, C. T., Pereira, O., Gonzaga, A. Feature Description based on Center-Symmetric Local Mapped Patterns. In: 29th Symposium on Applied Computing, 2014, Gyeongju, South Korea. Proceedings of the 29th Symposium on Applied Computing. New York: ACM, 2014. p. 39-44

Ferraz, C. T. et al. Object Recognition Based on Bag of Features and a New Local Pattern Descriptor. International Journal of Pattern Recognition and Artificial Intelligence, v. 28, n. 8 December 2014. DOI: <http://www.worldscientific.com/doi/abs/10.1142/S0218001414550106>.

Fine-Grained Labeled Faces in the Wild, 2017. Disponível em: <<http://www.whdeng.cn/FGLFW/FGLFW.html>>. Acesso em: 01 Novembro 2017.

Florea, L. Florea, C., Vertan, C. Robust eye centers localization with zero-crossing encoded image projections. Pattern Analysis and Applications, v. 20, p. 127-143, 2015. DOI 10.1007/s10044-015-0479-x.

FRVT – Face Recognition Vendor Test (NIST – National Institute of Standards and Technology) Disponível em <<http://www.nist.gov/itl/iad/ig/frvt-home.cfm>>. Acesso em: 28/02/2015.

FSPECIAL MOTION, MATLAB. Disponível em: <<https://www.mathworks.com/help/images/ref/fspecial.html>>. Acesso em 17/01/2018.

GENG, C.; JIANG, X. SIFT features for face recognition. In: IEEE International Conference on Computer Science and Information Technology (ICCSIT), 2nd, p. 598-602, 2009. DOI: 10.1109/ICCSIT.2009.5234877.

GONG, D.; LI, S.; XIANG, Y. Face recognition using the Weber Local Descriptor. In: Asian Conference on Pattern Recognition (ACPR), p. 589-592, 2011. DOI: 10.1109/ACPR.2011.6166675.

HADID, A.; PIETIKAINEN, M.; AHONEN, T. A discriminative feature space for detecting and recognizing faces. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), p. 797-804, 2004. DOI: 10.1109/CVPR.2004.1315246.

HEIKKILÄ, M.; PIETIKÄINEN, M.; SCHMID, C. Description of interest regions with center-symmetric local binary patterns. Springer-Verlag Berlin Heidelberg, 2006.

HEIKKILÄ, M.; PIETIKÄINEN, M.; SCHMID, C. Description of interest regions with local binary patterns. Pattern Recognition, v. 42, p. 425-436, 2009.

HUANG, G. B. et al. Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments. Technical Report 07-49, University of Massachusetts, Amherst, October, 2007.

HUANG, G.; JAIN, V.; LEARNED-MILLER, E. Unsupervised joint alignment of complex images. In: IEEE International Conference on Computer Vision (ICCV), p. 1-8, 2007. DOI: 10.1109/ICCV.2007.4408858.

HUANG, D. et al. Local Binary Patterns and Its Application to Facial Image Analysis: A Survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, v. 41, n. 6, p. 765-781, March 2011. DOI: 10.1109/TSMCC.2011.2118750.

HUANG, G. B. et al. Learning to Align from Scratch. Advances in Neural Information Processing Systems (NIPS), 2012.

IEEE Search: face recognition. Disponível em:
<<http://ieeexplore.ieee.org/search/searchresult.jsp?queryText=face%20recognition&newsearch=true>>. Acesso em: 02/02/2018.

JESORSKY, O.; KIRCHBERG, K. J.; FRISCHHOLZ, R. W. Robust face detection using the hausdorff distance. In: J. Bigun, F. Smeraldi (Eds.), Audio-and Video-Based Bio-metric Person Authentication. AVBPA 2001. Lecture Notes in Computer Science, 2091, Springer, Berlin, Heidelberg, 2001. doi: 10.1007/3-540-45344-X_14.

Jin, X., Tan, X. Face alignment by robust discriminative Hough voting. Pattern Recognition, v. 60, p. 318-333, 2016. DOI <https://doi.org/10.1016/j.patcog.2016.05.017>.

JUEFEI-XU, F.; LUU, K.; SAVVIDES, M. Spartans: Single-Sample Periocular-Based Alignment-Robust Recognition Technique Applied to Non-Frontal Scenarios. IEEE Transactions on Image Processing, v. 24, Issue 12, August 2015. DOI: 10.1109/TIP.2015.2468173

KANADE, T. Picture Processing by Computer Complex and Recognition of Human Faces. Ph.D. dissertation, Kyoto University, 1973.

KUMAR, N. et al. Describable Visual Attributes for Face Verification and Image Search. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 33, n. 10, p. 1962-1977, March 2011. DOI: 10.1109/TPAMI.2011.48.

Labeled Faces in the Wild, 2015. Disponível em: <<http://vis-www.cs.umass.edu/lfw/index.html>>. Acesso em: 25 Fevereiro 2015.

Labeled Faces in the Wild – Results, 2015. Disponível em: <<http://vis-www.cs.umass.edu/lfw/results.html>>. Acesso em: 25 Fevereiro 2015.

LI, Z.; PARK, U.; JAIN, A. A Discriminative Model for Age Invariant Face Recognition. IEEE Transactions on Information Forensics and Security, v. 6, n. 3, p. 1028-1037, May 2011. DOI: 10.1109/TIFS.2011.2156787.

LI, H. et al. Probabilistic Elastic Matching for Pose Variant Face Verification. In: Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, 2013. DOI: 10.1109/CVPR.2013.449

LI, H. et al. Eigen-PEP for Video Face Recognition. In: Asian Conference on Computer Vision (ACCV), 2014. DOI: DOI 10.1007/978-3-319-16811-1_2.

Lindeberg, T. Scale-space theory: A basic tool for analysing structures at different scales. Journal of Applied Statistics, v. 21, n. 2, p. 224-270, 1994.

LOWE, D. G. Object recognition from local scale-invariant features. In: The Proceedings of the Seventh IEEE International Conference on Computer Vision, p. 1150-1157, 1999. DOI: 10.1109/ICCV.1999.790410.

LOWE, D. G. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, v. 60, n.2, p. 91-110, November 2004. DOI: 10.1023/B:VISI.0000029664.99615.94.

LUO, J. et al. Person-Specific SIFT Features for Face Recognition. In: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) vol. 2, p. 593-596, 2007. DOI: 10.1109/ICASSP.2007.366305.

MAJUMDAR, A.; WARD, R. Discriminative SIFT features for face recognition. In: Canadian Conference on Electrical and Computer Engineering (CCECE), 9th, p. 27-30, 2009. DOI: 10.1109/CCECE.2009.5090085.

MILLER, P. et al. Performance Evaluation of Local Appearance Based Periocular Recognition. In: IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS), p. 1-6, 2010. DOI: 10.1109/BTAS.2010.5634536

NOWAK, E.; JURIE, F. Learning visual similarity measures for comparing never seen objects. In: Computer Vision and Pattern Recognition (CVPR), IEEE Conference on. 2007. DOI: 10.1109/CVPR.2007.382969

OJALA, T.; PIETIKAINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. Pattern Recognition, v. 29, n.1, p. 51-59, January 1996. DOI: 10.1016/0031-3203(95)00067-4.

OJALA, T.; PIETIKAINEN, M.; MAENPAA, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 24, n. 7, p. 971-987, July 2002. DOI: 10.1109/TPAMI.2002.1017623.

PARK, U.; ROSS, A.; JAIN, A. Periocular Biometrics in the Visible Spectrum: A Feasibility Study. In: IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems (BTAS), p. 1-6, 2009. DOI: 10.1109/BTAS.2009.5339068.

PARK, U. et al. Periocular Biometrics in the Visible Spectrum. IEEE Transactions on Information Forensics and Security, v. 6, n. 1, p. 96-106, December 2010. DOI: 10.1109/TIFS.2010.2096810.

PARKHI, O. M.; VEDALDI, A.; ZISSERMAN, A. Deep face recognition. In British Machine Vision Conference, 2015. DOI: 10.5244/C.29.41.

PINTO, N.; DICARLO, J.; COX, D. How far can you get with a modern face recognition test set using only simple features? In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p. 2591-2598, 2009. DOI: 10.1109/CVPR.2009.5206605.

Qian, Z., Xu, D. Automatic eye detection using intensity filtering and k-means clustering. Pattern Recognition Letters, v. 31, Issue 12, p. 1633-1640, 2010. DOI <http://dx.doi.org/10.1016/j.patrec.2010.05.012>.

SAGONAS, C. et al. Robust Statistical Frontalization of Human and Animal Faces. International Journal of Computer Vision (IJCV), July 2016. DOI: <https://doi.org/10.1007/s11263-016-0920-7>.

SHUAISHI, L.; YAUN, Z.; KEPING, L. Facial expression recognition under partial occlusion based on Weber Local Descriptor histogram and decision fusion. In: Chinese Control Conference (CCC), p. 4664-4668, 2014. DOI: 10.1109/ChiCC.2014.6895725.

SIMONYAN, K. et al. Fisher Vector Faces in the Wild. In: British Machine Vision Conference (BMVC), 2013.

SUN, Y. et al. Deep learning face representation by joint identification-verification. In: Advances in Neural Information Processing Systems, p. 1988–1996, 2014.

TAIGMAN, Y. et al. DeepFace: Closing the Gap to Human-Level Performance in Face Verification. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), p. 1701-1708, 2014. DOI: 10.1109/CVPR.2014.220.

Tan., X. et al., Enhanced Pictorial Structures for precise eye localization under uncontrolled conditions. IEEE Conference on Computer Vision and Pattern Recognition, p. 1621-1628, 2009. DOI 10.1109/CVPR.2009.5206818.

TURK, M. A.; PENTLAND, A P.. Eigenfaces for Recognition. Journal of Cognitive Neuroscience, v. 3, n. 1, p. 71-86, January 1991. DOI: 10.1162/jocn.1991.3.1.71.

ULLAH, I. et al. Gender recognition from face images with local WLD descriptor. In: International Conference on Systems, Signals and Image Processing (IWSSIP), 2012. p. 417-420.

Valenti, R., Gvers, T. Accurate eye center localization through invariant isocentric patterns. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 34, Issue 9, p. 1785-1798, 2012. DOI 10.1109/TPAMI.2011.251.

Vazquez-Fernandez, E., Gonzalez-Jimenez, D., Yu, L. L. Improved average of synthetic exact filters for precise eye localization under realistic conditions. IET Biometrics, v. 2, Issue 1, p. 10-20, 2013. DOI: 10.1049/iet-bmt.2011.0006.

VIEIRA, R. T. et al. Uma nova proposta para análise de textura local por conjuntos fuzzy. In: XI Workshop de Visão Computacional, 2012, Goiânia. XI Workshop de Visão Computacional, p. 146-151, 2012.

VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), p. 511-518, 2001. DOI: 10.1109/CVPR.2001.990517.

WANG, J. et al. Boosting dense SIFT descriptors and shape contexts of face images for gender recognition. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), p. 96-102, 2010. DOI: 10.1109/CVPRW.2010.5543238

WANG, X. et al. Feature fusion of HOG and WLD for facial expression recognition. In: IEEE/SICE International Symposium on System Integration (SII), p. 227-232, 2013. DOI: 10.1109/SII.2013.6776664.

Wang, J., Xiong, R., Chu, J. Facial feature points detecting based on Gaussian Mixture Models. Pattern Recognition Letters, v. 53, p. 62-68, 2015. DOI <https://doi.org/10.1016/j.patrec.2014.11.004>.

WOLF, L.; HASSNER, T.; TAIGMAN, Y. Descriptor Based Methods in the Wild. Faces in Real-Life Images Workshop in European Conference on Computer Vision (ECCV), 2008.

WOODARD, D. et al. Periocular Region Appearance Cues for Biometric Identification. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), p. 162-169, 2010. DOI: 10.1109/CVPRW.2010.5544621.

XU, J.; SAVVIDES, M. Unconstrained periocular biometric acquisition and recognition using COTS PTZ camera for uncooperative and non-cooperative subjects. In: IEEE Workshop on Applications of Computer Vision (WACV), p. 201-208, 2012. DOI: 10.1109/WACV.2012.6163051.

Yang, F. et al. Eye localization through multiscale sparse dictionaries. IEEE International Conference on Automatic Face & Gesture Recognition and Workshops (FG 2011), p. 514-518, 2011. DOI 10.1109/FG.2011.5771450.

YI, D.; LEI, Z.; LI, S. Z. A robust eye localization method for low quality face images, In: International Joint Conference on Biometrics (IJCB), 2011, p. 1–6, 2011. Doi: 10.1109/IJCB.2011.6117499.

ZHANG, L. et al. Face Recognition Using Scale Invariant Feature Transform and Support Vector Machine. In: International Conference for Young Computer Scientists (ICYCS), 9th, p. 1766-1770, 2008. DOI: 10.1109/ICYCS.2008.481.

Zhang, C. et al. Precise eye localization by fast local linear SVM. IEEE International Conference on Multimedia and Expo (ICME), 2014. DOI 10.1109/ICME.2014.6890174.

ZUIDERVELD, K. Contrast limited adaptive histogram equalization. Graphics Gems IV, p. 474-485, 1994.

APÊNDICE A

A Tabela 28 apresenta os pares de imagens da base BioID gerados neste trabalho de doutorado para métodos de reconhecimento facial baseados no problema do *pair matching*, pois a base BioID não apresenta nenhuma lista com pares de imagens para este tipo de aplicação.

Tabela 28 – Pares de imagens da base BioID para reconhecimento facial baseado no problema do *pair matching*. As colunas *y* representam: 1 – matched; 0 – mismatched.

Imagem 1	Imagem 2	y	Imagem 1	Imagem 2	y	Imagem 1	Imagem 2	y
BioID_0005.pgm	BioID_0004.pgm	1	BioID_0480.pgm	BioID_0166.pgm	0	BioID_0980.pgm	BioID_0585.pgm	0
BioID_0005.pgm	BioID_0006.pgm	1	BioID_0485.pgm	BioID_1397.pgm	0	BioID_0980.pgm	BioID_0752.pgm	0
BioID_0005.pgm	BioID_0622.pgm	0	BioID_0485.pgm	BioID_1499.pgm	0	BioID_0980.pgm	BioID_1294.pgm	0
BioID_0005.pgm	BioID_0977.pgm	0	BioID_0490.pgm	BioID_0488.pgm	1	BioID_0985.pgm	BioID_0983.pgm	1
BioID_0005.pgm	BioID_0696.pgm	0	BioID_0490.pgm	BioID_0489.pgm	1	BioID_0985.pgm	BioID_0984.pgm	1
BioID_0005.pgm	BioID_0004.pgm	1	BioID_0490.pgm	BioID_0491.pgm	1	BioID_0985.pgm	BioID_0996.pgm	1
BioID_0010.pgm	BioID_1081.pgm	0	BioID_0490.pgm	BioID_1034.pgm	0	BioID_0985.pgm	BioID_0002.pgm	0
BioID_0010.pgm	BioID_1015.pgm	0	BioID_0490.pgm	BioID_1508.pgm	0	BioID_0985.pgm	BioID_0755.pgm	0
BioID_0015.pgm	BioID_0016.pgm	0	BioID_0495.pgm	BioID_0493.pgm	1	BioID_0985.pgm	BioID_0745.pgm	0
BioID_0015.pgm	BioID_0949.pgm	0	BioID_0495.pgm	BioID_0494.pgm	1	BioID_0990.pgm	BioID_0988.pgm	1
BioID_0015.pgm	BioID_1169.pgm	0	BioID_0495.pgm	BioID_0982.pgm	0	BioID_0990.pgm	BioID_0044.pgm	0
BioID_0020.pgm	BioID_0019.pgm	1	BioID_0495.pgm	BioID_0790.pgm	1	BioID_0990.pgm	BioID_1104.pgm	0
BioID_0020.pgm	BioID_0021.pgm	1	BioID_0495.pgm	BioID_0559.pgm	0	BioID_0990.pgm	BioID_0767.pgm	0
BioID_0020.pgm	BioID_0973.pgm	0	BioID_0495.pgm	BioID_0517.pgm	1	BioID_0990.pgm	BioID_0490.pgm	0
BioID_0020.pgm	BioID_0947.pgm	1	BioID_0500.pgm	BioID_0498.pgm	1	BioID_0995.pgm	BioID_0994.pgm	1
BioID_0020.pgm	BioID_0856.pgm	0	BioID_0500.pgm	BioID_0511.pgm	1	BioID_0995.pgm	BioID_0996.pgm	1
BioID_0025.pgm	BioID_1406.pgm	0	BioID_0505.pgm	BioID_0503.pgm	1	BioID_0995.pgm	BioID_0743.pgm	0
BioID_0025.pgm	BioID_0324.pgm	0	BioID_0505.pgm	BioID_0988.pgm	0	BioID_0995.pgm	BioID_0572.pgm	0
BioID_0030.pgm	BioID_0029.pgm	1	BioID_0505.pgm	BioID_0889.pgm	0	BioID_1000.pgm	BioID_0998.pgm	1
BioID_0030.pgm	BioID_0031.pgm	1	BioID_0510.pgm	BioID_0509.pgm	1	BioID_1000.pgm	BioID_0760.pgm	0
BioID_0030.pgm	BioID_0377.pgm	0	BioID_0510.pgm	BioID_0755.pgm	0	BioID_1000.pgm	BioID_1006.pgm	1
BioID_0030.pgm	BioID_0326.pgm	0	BioID_0515.pgm	BioID_0513.pgm	1	BioID_1000.pgm	BioID_0302.pgm	0
BioID_0030.pgm	BioID_0931.pgm	0	BioID_0515.pgm	BioID_0516.pgm	1	BioID_1000.pgm	BioID_0951.pgm	0
BioID_0030.pgm	BioID_1181.pgm	0	BioID_0515.pgm	BioID_0573.pgm	0	BioID_1005.pgm	BioID_1003.pgm	1
BioID_0030.pgm	BioID_0461.pgm	0	BioID_0515.pgm	BioID_0506.pgm	1	BioID_1005.pgm	BioID_1004.pgm	1
BioID_0030.pgm	BioID_1402.pgm	0	BioID_0520.pgm	BioID_0519.pgm	1	BioID_1005.pgm	BioID_1006.pgm	1
BioID_0035.pgm	BioID_0036.pgm	1	BioID_0520.pgm	BioID_0521.pgm	1	BioID_1005.pgm	BioID_0505.pgm	0
BioID_0035.pgm	BioID_1504.pgm	0	BioID_0520.pgm	BioID_0593.pgm	0	BioID_1005.pgm	BioID_1106.pgm	0
BioID_0035.pgm	BioID_1498.pgm	0	BioID_0520.pgm	BioID_0001.pgm	0	BioID_1010.pgm	BioID_1009.pgm	1
BioID_0040.pgm	BioID_1219.pgm	0	BioID_0520.pgm	BioID_0686.pgm	0	BioID_1010.pgm	BioID_1011.pgm	1

Continua na próxima página

Continua na próxima página

Continua na próxima página

Continua na próxima página

Continua na próxima página

Continua na próxima página

Continua na próxima página

Imagem 1	Imagem 2	y	Imagem 1	Imagem 2	y	Imagem 1	Imagem 2	y
BioID_0255.pgm	BioID_0869.pgm	0	BioID_0730.pgm	BioID_0180.pgm	0	BioID_1260.pgm	BioID_0720.pgm	0
BioID_0255.pgm	BioID_0933.pgm	0	BioID_0735.pgm	BioID_0733.pgm	1	BioID_1260.pgm	BioID_1261.pgm	1
BioID_0260.pgm	BioID_0261.pgm	1	BioID_0735.pgm	BioID_0734.pgm	1	BioID_1260.pgm	BioID_1438.pgm	0
BioID_0260.pgm	BioID_0327.pgm	0	BioID_0735.pgm	BioID_0311.pgm	0	BioID_1260.pgm	BioID_0605.pgm	0
BioID_0260.pgm	BioID_0290.pgm	1	BioID_0735.pgm	BioID_0734.pgm	1	BioID_1265.pgm	BioID_1263.pgm	1
BioID_0260.pgm	BioID_1365.pgm	0	BioID_0735.pgm	BioID_0654.pgm	0	BioID_1265.pgm	BioID_1264.pgm	1
BioID_0265.pgm	BioID_0264.pgm	1	BioID_0735.pgm	BioID_0087.pgm	0	BioID_1265.pgm	BioID_1266.pgm	1
BioID_0265.pgm	BioID_0266.pgm	1	BioID_0740.pgm	BioID_0738.pgm	1	BioID_1265.pgm	BioID_1246.pgm	1
BioID_0265.pgm	BioID_0158.pgm	0	BioID_0740.pgm	BioID_0739.pgm	1	BioID_1265.pgm	BioID_1265.pgm	1
BioID_0265.pgm	BioID_0343.pgm	0	BioID_0740.pgm	BioID_1278.pgm	0	BioID_1265.pgm	BioID_1376.pgm	0
BioID_0265.pgm	BioID_0264.pgm	1	BioID_0740.pgm	BioID_0602.pgm	0	BioID_1265.pgm	BioID_0857.pgm	0
BioID_0265.pgm	BioID_0229.pgm	0	BioID_0740.pgm	BioID_0722.pgm	1	BioID_1265.pgm	BioID_0718.pgm	0
BioID_0265.pgm	BioID_1055.pgm	0	BioID_0740.pgm	BioID_0936.pgm	0	BioID_1265.pgm	BioID_0428.pgm	0
BioID_0265.pgm	BioID_0459.pgm	0	BioID_0740.pgm	BioID_1136.pgm	1	BioID_1265.pgm	BioID_1287.pgm	1
BioID_0265.pgm	BioID_0469.pgm	0	BioID_0740.pgm	BioID_0927.pgm	0	BioID_1270.pgm	BioID_1268.pgm	1
BioID_0270.pgm	BioID_0268.pgm	1	BioID_0740.pgm	BioID_1130.pgm	1	BioID_1270.pgm	BioID_0652.pgm	0
BioID_0270.pgm	BioID_0269.pgm	1	BioID_0740.pgm	BioID_1138.pgm	1	BioID_1270.pgm	BioID_0241.pgm	0
BioID_0270.pgm	BioID_0271.pgm	1	BioID_0740.pgm	BioID_0902.pgm	0	BioID_1270.pgm	BioID_0407.pgm	0
BioID_0270.pgm	BioID_1368.pgm	0	BioID_0745.pgm	BioID_0743.pgm	1	BioID_1270.pgm	BioID_0195.pgm	0
BioID_0275.pgm	BioID_0273.pgm	1	BioID_0745.pgm	BioID_0744.pgm	1	BioID_1270.pgm	BioID_0460.pgm	0
BioID_0275.pgm	BioID_0274.pgm	1	BioID_0745.pgm	BioID_1038.pgm	0	BioID_1270.pgm	BioID_1291.pgm	1
BioID_0275.pgm	BioID_0276.pgm	1	BioID_0745.pgm	BioID_0766.pgm	1	BioID_1270.pgm	BioID_0214.pgm	0
BioID_0275.pgm	BioID_0389.pgm	0	BioID_0745.pgm	BioID_0775.pgm	1	BioID_1275.pgm	BioID_1274.pgm	1
BioID_0275.pgm	BioID_1365.pgm	0	BioID_0750.pgm	BioID_0748.pgm	1	BioID_1275.pgm	BioID_1276.pgm	1
BioID_0275.pgm	BioID_1390.pgm	0	BioID_0750.pgm	BioID_0749.pgm	1	BioID_1275.pgm	BioID_0726.pgm	0
BioID_0275.pgm	BioID_0277.pgm	1	BioID_0750.pgm	BioID_0751.pgm	1	BioID_1275.pgm	BioID_0851.pgm	0
BioID_0275.pgm	BioID_0222.pgm	0	BioID_0750.pgm	BioID_0300.pgm	0	BioID_1275.pgm	BioID_0637.pgm	0
BioID_0275.pgm	BioID_1360.pgm	0	BioID_0750.pgm	BioID_0478.pgm	0	BioID_1275.pgm	BioID_0722.pgm	0
BioID_0275.pgm	BioID_0386.pgm	0	BioID_0750.pgm	BioID_0297.pgm	0	BioID_1275.pgm	BioID_1212.pgm	0
BioID_0275.pgm	BioID_0160.pgm	0	BioID_0750.pgm	BioID_0504.pgm	0	BioID_1275.pgm	BioID_0711.pgm	0
BioID_0275.pgm	BioID_0382.pgm	0	BioID_0750.pgm	BioID_0764.pgm	1	BioID_1275.pgm	BioID_0080.pgm	0
BioID_0280.pgm	BioID_0278.pgm	1	BioID_0755.pgm	BioID_0753.pgm	1	BioID_1280.pgm	BioID_1278.pgm	1
BioID_0280.pgm	BioID_0281.pgm	1	BioID_0755.pgm	BioID_0756.pgm	1	BioID_1280.pgm	BioID_1279.pgm	1
BioID_0280.pgm	BioID_1312.pgm	0	BioID_0755.pgm	BioID_1036.pgm	0	BioID_1280.pgm	BioID_1281.pgm	1
BioID_0280.pgm	BioID_0276.pgm	1	BioID_0755.pgm	BioID_0484.pgm	0	BioID_1280.pgm	BioID_1308.pgm	0
BioID_0280.pgm	BioID_0284.pgm	1	BioID_0755.pgm	BioID_0304.pgm	0	BioID_1280.pgm	BioID_0398.pgm	0
BioID_0280.pgm	BioID_0946.pgm	0	BioID_0760.pgm	BioID_0758.pgm	1	BioID_1280.pgm	BioID_0422.pgm	0
BioID_0280.pgm	BioID_1406.pgm	0	BioID_0760.pgm	BioID_0759.pgm	1	BioID_1280.pgm	BioID_0204.pgm	0
BioID_0285.pgm	BioID_0283.pgm	1	BioID_0760.pgm	BioID_0761.pgm	1	BioID_1280.pgm	BioID_0911.pgm	0
BioID_0285.pgm	BioID_0284.pgm	1	BioID_0760.pgm	BioID_0298.pgm	0	BioID_1280.pgm	BioID_1184.pgm	0
BioID_0285.pgm	BioID_0577.pgm	0	BioID_0760.pgm	BioID_1341.pgm	0	BioID_1280.pgm	BioID_0212.pgm	0
BioID_0285.pgm	BioID_0944.pgm	0	BioID_0760.pgm	BioID_0776.pgm	1	BioID_1280.pgm	BioID_0961.pgm	0
BioID_0290.pgm	BioID_0288.pgm	1	BioID_0760.pgm	BioID_0497.pgm	0	BioID_1285.pgm	BioID_1283.pgm	1

Continua na próxima página

Continua na próxima página

Continua na próxima página

Continua na próxima página

Imagem 1	Imagem 2	y	Imagem 1	Imagem 2	y	Imagem 1	Imagem 2	y
BiolD_0400.pgm	BiolD_0398.pgm	1	BiolD_0875.pgm	BiolD_0474.pgm	0	BiolD_1380.pgm	BiolD_1381.pgm	1
BiolD_0400.pgm	BiolD_0399.pgm	1	BiolD_0875.pgm	BiolD_0391.pgm	0	BiolD_1380.pgm	BiolD_0673.pgm	0
BiolD_0400.pgm	BiolD_0601.pgm	1	BiolD_0875.pgm	BiolD_1418.pgm	0	BiolD_1380.pgm	BiolD_0902.pgm	0
BiolD_0400.pgm	BiolD_0003.pgm	0	BiolD_0875.pgm	BiolD_0877.pgm	1	BiolD_1380.pgm	BiolD_1149.pgm	0
BiolD_0400.pgm	BiolD_0342.pgm	1	BiolD_0880.pgm	BiolD_0878.pgm	1	BiolD_1380.pgm	BiolD_0257.pgm	0
BiolD_0400.pgm	BiolD_1497.pgm	0	BiolD_0880.pgm	BiolD_0881.pgm	1	BiolD_1380.pgm	BiolD_1362.pgm	1
BiolD_0400.pgm	BiolD_1122.pgm	0	BiolD_0880.pgm	BiolD_1396.pgm	0	BiolD_1380.pgm	BiolD_0363.pgm	0
BiolD_0400.pgm	BiolD_0241.pgm	0	BiolD_0880.pgm	BiolD_0772.pgm	0	BiolD_1385.pgm	BiolD_1383.pgm	1
BiolD_0400.pgm	BiolD_0830.pgm	0	BiolD_0880.pgm	BiolD_0262.pgm	1	BiolD_1385.pgm	BiolD_0532.pgm	0
BiolD_0400.pgm	BiolD_0377.pgm	1	BiolD_0880.pgm	BiolD_0563.pgm	0	BiolD_1385.pgm	BiolD_0005.pgm	0
BiolD_0400.pgm	BiolD_1175.pgm	1	BiolD_0880.pgm	BiolD_0944.pgm	0	BiolD_1385.pgm	BiolD_0965.pgm	0
BiolD_0405.pgm	BiolD_0403.pgm	1	BiolD_0880.pgm	BiolD_0882.pgm	1	BiolD_1385.pgm	BiolD_0140.pgm	0
BiolD_0405.pgm	BiolD_0404.pgm	1	BiolD_0880.pgm	BiolD_0748.pgm	0	BiolD_1390.pgm	BiolD_1388.pgm	1
BiolD_0405.pgm	BiolD_0406.pgm	1	BiolD_0885.pgm	BiolD_0883.pgm	1	BiolD_1390.pgm	BiolD_1389.pgm	1
BiolD_0405.pgm	BiolD_1501.pgm	0	BiolD_0885.pgm	BiolD_0998.pgm	0	BiolD_1390.pgm	BiolD_1391.pgm	1
BiolD_0405.pgm	BiolD_1418.pgm	0	BiolD_0885.pgm	BiolD_0951.pgm	0	BiolD_1390.pgm	BiolD_0647.pgm	0
BiolD_0405.pgm	BiolD_0159.pgm	0	BiolD_0885.pgm	BiolD_0753.pgm	0	BiolD_1390.pgm	BiolD_0270.pgm	0
BiolD_0405.pgm	BiolD_0041.pgm	0	BiolD_0890.pgm	BiolD_0889.pgm	1	BiolD_1390.pgm	BiolD_0479.pgm	1
BiolD_0405.pgm	BiolD_1053.pgm	0	BiolD_0890.pgm	BiolD_0891.pgm	1	BiolD_1390.pgm	BiolD_0225.pgm	0
BiolD_0405.pgm	BiolD_1388.pgm	0	BiolD_0890.pgm	BiolD_1115.pgm	0	BiolD_1390.pgm	BiolD_0266.pgm	0
BiolD_0405.pgm	BiolD_0404.pgm	1	BiolD_0890.pgm	BiolD_0494.pgm	0	BiolD_1390.pgm	BiolD_0474.pgm	1
BiolD_0405.pgm	BiolD_0094.pgm	0	BiolD_0890.pgm	BiolD_1221.pgm	1	BiolD_1390.pgm	BiolD_0413.pgm	0
BiolD_0405.pgm	BiolD_1513.pgm	0	BiolD_0895.pgm	BiolD_0894.pgm	1	BiolD_1395.pgm	BiolD_1393.pgm	1
BiolD_0405.pgm	BiolD_0932.pgm	0	BiolD_0895.pgm	BiolD_0896.pgm	1	BiolD_1395.pgm	BiolD_1394.pgm	1
BiolD_0405.pgm	BiolD_1410.pgm	0	BiolD_0895.pgm	BiolD_1364.pgm	0	BiolD_1395.pgm	BiolD_0474.pgm	1
BiolD_0410.pgm	BiolD_1053.pgm	0	BiolD_0895.pgm	BiolD_0538.pgm	0	BiolD_1395.pgm	BiolD_0449.pgm	1
BiolD_0410.pgm	BiolD_0964.pgm	0	BiolD_0895.pgm	BiolD_0848.pgm	0	BiolD_1395.pgm	BiolD_0581.pgm	0
BiolD_0410.pgm	BiolD_0160.pgm	0	BiolD_0895.pgm	BiolD_0894.pgm	1	BiolD_1395.pgm	BiolD_1312.pgm	0
BiolD_0410.pgm	BiolD_0485.pgm	0	BiolD_0895.pgm	BiolD_0852.pgm	0	BiolD_1395.pgm	BiolD_0462.pgm	1
BiolD_0410.pgm	BiolD_0756.pgm	0	BiolD_0895.pgm	BiolD_0894.pgm	1	BiolD_1400.pgm	BiolD_1398.pgm	1
BiolD_0415.pgm	BiolD_0413.pgm	1	BiolD_0900.pgm	BiolD_0898.pgm	1	BiolD_1400.pgm	BiolD_1399.pgm	1
BiolD_0415.pgm	BiolD_0414.pgm	1	BiolD_0900.pgm	BiolD_0899.pgm	1	BiolD_1400.pgm	BiolD_1401.pgm	1
BiolD_0415.pgm	BiolD_0416.pgm	1	BiolD_0900.pgm	BiolD_0901.pgm	1	BiolD_1400.pgm	BiolD_1410.pgm	1
BiolD_0415.pgm	BiolD_1048.pgm	0	BiolD_0900.pgm	BiolD_0326.pgm	0	BiolD_1400.pgm	BiolD_0464.pgm	1
BiolD_0415.pgm	BiolD_0468.pgm	0	BiolD_0900.pgm	BiolD_0395.pgm	0	BiolD_1400.pgm	BiolD_0403.pgm	0
BiolD_0415.pgm	BiolD_1368.pgm	0	BiolD_0900.pgm	BiolD_0625.pgm	0	BiolD_1400.pgm	BiolD_0945.pgm	0
BiolD_0415.pgm	BiolD_0649.pgm	0	BiolD_0900.pgm	BiolD_0605.pgm	0	BiolD_1400.pgm	BiolD_1504.pgm	0
BiolD_0415.pgm	BiolD_0451.pgm	0	BiolD_0900.pgm	BiolD_1509.pgm	0	BiolD_1400.pgm	BiolD_1349.pgm	0
BiolD_0415.pgm	BiolD_0479.pgm	0	BiolD_0905.pgm	BiolD_0903.pgm	1	BiolD_1400.pgm	BiolD_1053.pgm	0
BiolD_0415.pgm	BiolD_0391.pgm	1	BiolD_0905.pgm	BiolD_0904.pgm	1	BiolD_1400.pgm	BiolD_0650.pgm	0
BiolD_0420.pgm	BiolD_0419.pgm	1	BiolD_0905.pgm	BiolD_0906.pgm	1	BiolD_1405.pgm	BiolD_1403.pgm	1
BiolD_0420.pgm	BiolD_0421.pgm	1	BiolD_0905.pgm	BiolD_0194.pgm	0	BiolD_1405.pgm	BiolD_1404.pgm	1
BiolD_0420.pgm	BiolD_0395.pgm	0	BiolD_0905.pgm	BiolD_1133.pgm	0	BiolD_1405.pgm	BiolD_1406.pgm	1

Continua na próxima página

Continua na próxima página

Continua na próxima página

Imagem 1	Imagem 2	y	Imagem 1	Imagem 2	y	Imagem 1	Imagem 2	y
BiolD_0975.pgm	BiolD_1000.pgm	1	BiolD_0980.pgm	BiolD_0979.pgm	1	BiolD_0980.pgm	BiolD_0168.pgm	0
BiolD_0980.pgm	BiolD_0283.pgm	0	BiolD_0980.pgm	BiolD_0166.pgm	0	BiolD_0980.pgm	BiolD_1323.pgm	0
Conclusão								

Fonte: autoria própria.