

**Universidade da Beira Interior**  
Departamento de Informática



## **Análise de Imagens: Reconhecimento Biométrico através da Forma da Orelha**

**Adolfo Barbosa do Amaral Peixinho, nº 12458**  
Licenciatura em Engenharia Informática

**Orientador do Projecto: Prof. Doutor Hugo Pedro Martins  
Cariço Proença**

Covilhã Julho de 2010



# Agradecimentos

Dedico a todos os que tornarão possível a realização deste projecto, um sincero agradecimento.

Em primeiro lugar, gostaria de agradecer ao meu orientador de projecto, o Professor Doutor Hugo Proença, por toda a informação que me forneceu, pela forma clara e precisa com que me transmitiu a sabedoria e conhecimento científico, pela sua disponibilidade de tempo e acima de tudo pelos seus ensinamentos e conselhos, sendo que foram cruciais para a realização deste projecto.

Quero também agradecer à minha família, em especial à minha mãe, pois sempre me apoiou e ajudou ao longo desta etapa da minha vida.

Quero agradecer a todos os membros e colaboradores do grupo *Soft Computing and Image Analysis Laboratory* (SOCIA Lab). pelo óptimo espírito de equipa e pela inter-ajuda existente no laboratório, especialmente ao Silvío e ao Gil.

Por fim, mas não menos importante à minha querida namorada Francisca Neves (FrancisJo), por toda e imensa paciência e apoio que me deu ao longo deste tempo, e mais não digo.

Obrigado!



# Conteúdo

<b>Agradecimentos</b>	<b>i</b>
<b>Conteúdo</b>	<b>iii</b>
<b>Lista de Tabelas</b>	<b>v</b>
<b>Lista de Figuras</b>	<b>vii</b>
<b>Lista de Algoritmos</b>	<b>xxi</b>
<b>Acrónimos</b>	<b>xxiii</b>
<b>Glossário</b>	<b>xxv</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Objectivo . . . . .	1
1.2 Motivação . . . . .	4
1.3 Abordagem . . . . .	6
1.4 Organização do relatório . . . . .	8
<b>2 Método de Extracção/Reconhecimento de Características de Imagens Digitais</b>	<b>9</b>
2.1 Banco de Imagens Digitais . . . . .	9

---

2.2	Pesquisa Relacionada . . . . .	11
2.3	Método de Extração de Características . . . . .	13
2.3.1	Detecção dos Extremos no Espaço-Escala . . . . .	13
2.3.2	Localização de Keypoints . . . . .	16
2.3.3	Atribuição de Orientações . . . . .	21
2.3.4	Descritores de Keypoints . . . . .	22
2.4	Método de Reconhecimento de Keypoints . . . . .	26
<b>3</b>	<b>Implementação</b>	<b>31</b>
3.1	Comparação da <i>Application Programming Interface</i> (API) utilizada e o Método de Lowe . . . . .	32
3.2	Parâmetros Utilizados na Extração de Características . . . . .	34
3.3	Algoritmos Implementados . . . . .	38
<b>4</b>	<b>Resultados</b>	<b>43</b>
4.1	Conjunto de Características . . . . .	43
4.2	Discussão de Resultados . . . . .	52
<b>5</b>	<b>Conclusão e Trabalho Futuro</b>	<b>61</b>
5.1	Conclusão . . . . .	61
5.1.1	Trabalho Futuro . . . . .	63
<b>A</b>	<b>Anexos</b>	<b>65</b>
A.1	Banco de Imagens A . . . . .	65
A.2	Banco de Imagens B . . . . .	86
A.3	Banco de Imagens C . . . . .	107
	<b>Bibliografia</b>	<b>127</b>

# Lista de Tabelas

2.1	distribuição das imagens da Base de Dados de imagens da <i>Universidade de Notre Dame (UND)</i> . . . . .	10
3.1	Configurações de extracção de características testadas . . . .	34



# Lista de Figuras

1.1	(a) Anatomia (b) Localizações das medidas antropomórficas usadas no sistema de Iannarelli. . . . .	2
1.2	Avaliação de Dados Biométricos [10] . . . . .	3
2.1	Processo de segmentação da orelha . . . . .	10
2.2	Ilustração da função <i>Difference of Gaussian</i> (DoG) . . . . .	15
2.3	Máxima e mínima da função DoG . . . . .	16
2.4	Gráfico de avaliação/repetibilidade de escalas no espaço-escala	17
2.5	Parâmetros de Extracção de Keypoints . . . . .	18
2.6	Histograma de orientações dominantes dos gradientes [7] . . . . .	22
2.7	Múltiplas direcções dominantes do histograma.[7] . . . . .	23
2.8	Descritor de imagem local . . . . .	24
2.9	Gráfico histograma de orientações . . . . .	26
2.10	<i>Probability Density Function</i> (PDF) para uma correspondência correcta . . . . .	28
3.1	Comparação entre Keypoints . . . . .	32
3.2	Percentagem de Keypoints Coincidentes . . . . .	33
3.3	Percentagem de keypoints em Relação à Distância Média do Descritor . . . . .	33
3.4	Exemplos de keypoints extraídos nas experiências, do Banco de Imagem A. . . . .	37

4.1	Comparação Intra-Classe . . . . .	44
4.2	Comparação Intra-Classe . . . . .	45
4.3	Comparação Intra-Classe . . . . .	46
4.4	Comparação Intra-Classe . . . . .	46
4.5	Comparação Intra-Classe . . . . .	47
4.6	Keypoints extraídos com um $peakthreshold = 0.02$ . . . . .	47
4.7	Comparação Intra-Classe orelha oclusa . . . . .	48
4.8	Reconhecimento Falhado/Impossível . . . . .	48
4.9	Comparação Intra-Classe orelhas rodadas . . . . .	49
4.10	Comparação Intra-Classe orelhas rodadas . . . . .	49
4.11	Comparação Inter-Classe . . . . .	50
4.12	Comparação Inter-Classe . . . . .	50
4.13	Tabela de Resultados do Banco de Imagens A, configurações impossíveis marcadas a vermelho, melhores resultados marcados a verde . . . . .	53
4.14	Avaliação do Classificador na Configuração 1 para 10 keypoints (cardinalidade = 10). Sendo esta a configuração que apresenta melhores resultados no Banco de Imagens A. . . . .	54
4.15	Avaliação do Classificador na Configuração 1 para 20 keypoints (cardinalidade = 20). Podemos verificar que com a mesma configuração e com um aumento da cardinalidade de correspondências o resultado piora . . . . .	54
4.16	Avaliação do Classificador na Configuração 2, com máscara, para 5 keypoints (cardinalidade = 5). Esta configuração e a configuração 13 foram as que apresentaram melhores resultados, com máscara, para o Banco de Imagens A. . . . .	55
4.17	Avaliação do Classificador na Configuração 13, com máscara, rotação entre $-3\pi/2 < \theta < \pi/2$ para 5 keypoints (cardinalidade = 5). Esta e configuração 2 foram as que apresentaram melhores resultados, com máscara, para o Banco de Imagens A. . . . .	55
4.18	Tabela de Resultados do Banco de Imagens B . . . . .	56

---

4.19	Avaliação do Classificador na Configuração 1 para 10 key-points (cardinalidade = 10). Sendo que é a melhor configuração. . . . .	57
4.20	Avaliação do Classificador na Configuração 2 para 5 key-points (cardinalidade = 5). Uma boa classificação com máscara no Banco de Imagens B . . . . .	57
4.21	Tabela de Resultados do Banco de Imagens C . . . . .	58
4.22	Avaliação do Classificador na Configuração 1 para 20 key-points (cardinalidade = 20). . . . .	58
4.23	Avaliação do Classificador na Configuração 6 para 20 key-points (cardinalidade = 20). Sendo que é o segundo pior resultado, a seguir à configuração 9. . . . .	59
4.24	Avaliação do Classificador na Configuração 6 para 10 key-points (cardinalidade = 10). Sendo um dos melhores resultados apresentado por configurações com máscara. . . . .	59
A.1	Avaliação do Classificador na Configuração 1 para 20 key-points (cardinalidade = 20). . . . .	65
A.2	Avaliação do Classificador na Configuração 1 para 10 key-points (cardinalidade = 10). . . . .	66
A.3	Avaliação do Classificador na Configuração 1 para 5 key-points (cardinalidade = 5). . . . .	66
A.4	Avaliação do Classificador na Configuração 2 para 20 key-points (cardinalidade = 20), com máscara . . . . .	67
A.5	Avaliação do Classificador na Configuração 2 para 10 key-points (cardinalidade = 10), com máscara . . . . .	67
A.6	Avaliação do Classificador na Configuração 2 para 5 key-points (cardinalidade = 5), com máscara . . . . .	68
A.7	Avaliação do Classificador na Configuração 3 para 20 key-points (cardinalidade = 20), com máscara, peakthreshold = 0.01 . . . . .	68
A.8	Avaliação do Classificador na Configuração 3 para 10 key-points (cardinalidade = 10), com máscara, peakthreshold = 0.01 . . . . .	69

---

A.9	Avaliação do Classificador na Configuração 3 para 5 key-points (cardinalidade = 5), com máscara, peakthreshold = 0.01 . . . . .	69
A.10	Avaliação do Classificador na Configuração 4 para 20 key-points (cardinalidade = 20), peakthreshold = 0.01 e escala > 2 . . . . .	70
A.11	Avaliação do Classificador na Configuração 4 para 10 key-points (cardinalidade = 10), peakthreshold = 0.01 e escala > 2 . . . . .	70
A.12	Avaliação do Classificador na Configuração 4 para 5 key-points (cardinalidade = 5), peakthreshold = 0.01 e escala > 2 . . . . .	71
A.13	Avaliação do Classificador na Configuração 5 para 20 key-points (cardinalidade = 20), peakthreshold = 0.01 . . . . .	71
A.14	Avaliação do Classificador na Configuração 5 para 10 key-points (cardinalidade = 10), peakthreshold = 0.01 . . . . .	72
A.15	Avaliação do Classificador na Configuração 5 para 5 key-points (cardinalidade = 5), peakthreshold = 0.01 . . . . .	72
A.16	Avaliação do Classificador na Configuração 6 para 20 key-points (cardinalidade = 20), com máscara, peakthreshold = 0.01 e escala $\geq 2$ . . . . .	73
A.17	Avaliação do Classificador na Configuração 6 para 10 key-points (cardinalidade = 10), com máscara, peakthreshold = 0.01 e escala $\geq 2$ . . . . .	73
A.18	Avaliação do Classificador na Configuração 6 para 5 key-points (cardinalidade = 5), com máscara, peakthreshold = 0.01 e escala $\geq 2$ . . . . .	74
A.19	Avaliação do Classificador na Configuração 7 para 10 key-points (cardinalidade = 10), com máscara, peakthreshold = 0.02 . . . . .	74
A.20	Avaliação do Classificador na Configuração 7 para 5 key-points (cardinalidade = 5), com máscara, peakthreshold = 0.02 . . . . .	75

---

A.21 Avaliação do Classificador na Configuração 8 para 10 keypoints (cardinalidade = 10), com máscara, peakthreshold = 0.02 e escala $\geq 2$ . . . . .	75
A.22 Avaliação do Classificador na Configuração 8 para 5 keypoints (cardinalidade = 5), com máscara, peakthreshold = 0.02 e escala $\geq 2$ . . . . .	76
A.23 Avaliação do Classificador na Configuração 9 para 20 keypoints (cardinalidade = 20), peakthreshold = 0.02 . . . . .	76
A.24 Avaliação do Classificador na Configuração 9 para 10 keypoints (cardinalidade = 10), peakthreshold = 0.02 . . . . .	77
A.25 Avaliação do Classificador na Configuração 9 para 5 keypoints (cardinalidade = 5), peakthreshold = 0.02 . . . . .	77
A.26 Avaliação do Classificador na Configuração 10 para 20 keypoints (cardinalidade = 20), orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	78
A.27 Avaliação do Classificador na Configuração 10 para 10 keypoints (cardinalidade = 10), orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	78
A.28 Avaliação do Classificador na Configuração 10 para 5 keypoints (cardinalidade = 5), orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	79
A.29 Avaliação do Classificador na Configuração 11 para 20 keypoints (cardinalidade = 20), peakthreshold = 0.02 e escala $\geq 2$ . . . . .	79
A.30 Avaliação do Classificador na Configuração 11 para 10 keypoints (cardinalidade = 10), peakthreshold = 0.02 e escala $\geq 2$ . . . . .	80
A.31 Avaliação do Classificador na Configuração 11 para 5 keypoints (cardinalidade = 5), peakthreshold = 0.02 e escala $\geq 2$ . . . . .	80
A.32 Avaliação do Classificador na Configuração 12 para 20 keypoints (cardinalidade = 20), com máscara, orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	81

A.33 Avaliação do Classificador na Configuração 12 para 10 key-points (cardinalidade = 10), com máscara, orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	81
A.34 Avaliação do Classificador na Configuração 12 para 5 key-points (cardinalidade = 5), com máscara, orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	82
A.35 Avaliação do Classificador na Configuração 13 para 20 key-points (cardinalidade = 20), com máscara, orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	82
A.36 Avaliação do Classificador na Configuração 13 para 10 key-points (cardinalidade = 10), com máscara, orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	83
A.37 Avaliação do Classificador na Configuração 13 para 5 key-points (cardinalidade = 5), com máscara, orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	83
A.38 Avaliação do Classificador na Configuração 14 para 20 key-points (cardinalidade = 20), orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	84
A.39 Avaliação do Classificador na Configuração 14 para 10 key-points (cardinalidade = 10), orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	84
A.40 Avaliação do Classificador na Configuração 14 para 5 key-points (cardinalidade = 5), orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	85
A.41 Avaliação do Classificador na Configuração 1 para 20 key-points (cardinalidade = 20). . . . .	86
A.42 Avaliação do Classificador na Configuração 1 para 10 key-points (cardinalidade = 10). . . . .	87
A.43 Avaliação do Classificador na Configuração 1 para 5 key-points (cardinalidade = 5). . . . .	87
A.44 Avaliação do Classificador na Configuração 2 para 20 key-points (cardinalidade = 20), com máscara . . . . .	88
A.45 Avaliação do Classificador na Configuração 2 para 10 key-points (cardinalidade = 10), com máscara . . . . .	88

---

A.46 Avaliação do Classificador na Configuração 2 para 5 key-points (cardinalidade = 5), com máscara . . . . .	89
A.47 Avaliação do Classificador na Configuração 3 para 20 key-points (cardinalidade = 20), com máscara, peakthreshold = 0.01 . . . . .	89
A.48 Avaliação do Classificador na Configuração 3 para 10 key-points (cardinalidade = 10), com máscara, peakthreshold = 0.01 . . . . .	90
A.49 Avaliação do Classificador na Configuração 3 para 5 key-points (cardinalidade = 5), com máscara, peakthreshold = 0.01 . . . . .	90
A.50 Avaliação do Classificador na Configuração 4 para 20 key-points (cardinalidade = 20), peakthreshold = 0.01 e escala > 2 . . . . .	91
A.51 Avaliação do Classificador na Configuração 4 para 10 key-points (cardinalidade = 10), peakthreshold = 0.01 e escala > 2 . . . . .	91
A.52 Avaliação do Classificador na Configuração 4 para 5 key-points (cardinalidade = 5), peakthreshold = 0.01 e escala > 2 . . . . .	92
A.53 Avaliação do Classificador na Configuração 5 para 20 key-points (cardinalidade = 20), peakthreshold = 0.01 . . . . .	92
A.54 Avaliação do Classificador na Configuração 5 para 10 key-points (cardinalidade = 10), peakthreshold = 0.01 . . . . .	93
A.55 Avaliação do Classificador na Configuração 5 para 5 key-points (cardinalidade = 5), peakthreshold = 0.01 . . . . .	93
A.56 Avaliação do Classificador na Configuração 6 para 20 key-points (cardinalidade = 20), com máscara, peakthreshold = 0.01 e escala >= 2 . . . . .	94
A.57 Avaliação do Classificador na Configuração 6 para 10 key-points (cardinalidade = 10), com máscara, peakthreshold = 0.01 e escala >= 2 . . . . .	94

A.58 Avaliação do Classificador na Configuração 6 para 5 key-points (cardinalidade = 5), com máscara, peakthreshold = 0.01 e escala $\geq 2$ . . . . .	95
A.59 Avaliação do Classificador na Configuração 7 para 10 key-points (cardinalidade = 10), com máscara, peakthreshold = 0.02 . . . . .	95
A.60 Avaliação do Classificador na Configuração 7 para 5 key-points (cardinalidade = 5), com máscara, peakthreshold = 0.02 . . . . .	96
A.61 Avaliação do Classificador na Configuração 8 para 10 key-points (cardinalidade = 10), com máscara, peakthreshold = 0.02 e escala $\geq 2$ . . . . .	96
A.62 Avaliação do Classificador na Configuração 8 para 5 key-points (cardinalidade = 5), com máscara, peakthreshold = 0.02 e escala $\geq 2$ . . . . .	97
A.63 Avaliação do Classificador na Configuração 9 para 20 key-points (cardinalidade = 20), peakthreshold = 0.02 . . . . .	97
A.64 Avaliação do Classificador na Configuração 9 para 10 key-points (cardinalidade = 10), peakthreshold = 0.02 . . . . .	98
A.65 Avaliação do Classificador na Configuração 9 para 5 key-points (cardinalidade = 5), peakthreshold = 0.02 . . . . .	98
A.66 Avaliação do Classificador na Configuração 10 para 20 key-points (cardinalidade = 20), orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	99
A.67 Avaliação do Classificador na Configuração 10 para 10 key-points (cardinalidade = 10), orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	99
A.68 Avaliação do Classificador na Configuração 10 para 5 key-points (cardinalidade = 5), orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	100
A.69 Avaliação do Classificador na Configuração 11 para 20 key-points (cardinalidade = 20), peakthreshold = 0.02 e escala $\geq 2$ . . . . .	100

---

A.70 Avaliação do Classificador na Configuração 11 para 10 key-points (cardinalidade = 10), peakthreshold = 0.02 e escala $\geq 2$ . . . . .	101
A.71 Avaliação do Classificador na Configuração 11 para 5 key-points (cardinalidade = 5), peakthreshold = 0.02 e escala $\geq 2$ . . . . .	101
A.72 Avaliação do Classificador na Configuração 12 para 20 key-points (cardinalidade = 20), com máscara, orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	102
A.73 Avaliação do Classificador na Configuração 12 para 10 key-points (cardinalidade = 10), com máscara, orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	102
A.74 Avaliação do Classificador na Configuração 12 para 5 key-points (cardinalidade = 5), com máscara, orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	103
A.75 Avaliação do Classificador na Configuração 13 para 20 key-points (cardinalidade = 20), com máscara, orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	103
A.76 Avaliação do Classificador na Configuração 13 para 10 key-points (cardinalidade = 10), com máscara, orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	104
A.77 Avaliação do Classificador na Configuração 13 para 5 key-points (cardinalidade = 5), com máscara, orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	104
A.78 Avaliação do Classificador na Configuração 14 para 20 key-points (cardinalidade = 20), orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	105
A.79 Avaliação do Classificador na Configuração 14 para 10 key-points (cardinalidade = 10), orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	105
A.80 Avaliação do Classificador na Configuração 14 para 5 key-points (cardinalidade = 5), orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	106

---

A.81 Avaliação do Classificador na Configuração 1 para 20 key-points (cardinalidade = 20). . . . .	107
A.82 Avaliação do Classificador na Configuração 1 para 10 key-points (cardinalidade = 10). . . . .	108
A.83 Avaliação do Classificador na Configuração 1 para 5 key-points (cardinalidade = 5). . . . .	108
A.84 Avaliação do Classificador na Configuração 2 para 20 key-points (cardinalidade = 20), com máscara . . . . .	109
A.85 Avaliação do Classificador na Configuração 2 para 10 key-points (cardinalidade = 10), com máscara . . . . .	109
A.86 Avaliação do Classificador na Configuração 2 para 5 key-points (cardinalidade = 5), com máscara . . . . .	110
A.87 Avaliação do Classificador na Configuração 3 para 20 key-points (cardinalidade = 20), com máscara, peakthreshold = 0.01 . . . . .	110
A.88 Avaliação do Classificador na Configuração 3 para 10 key-points (cardinalidade = 10), com máscara, peakthreshold = 0.01 . . . . .	111
A.89 Avaliação do Classificador na Configuração 3 para 5 key-points (cardinalidade = 5), com máscara, peakthreshold = 0.01 . . . . .	111
A.90 Avaliação do Classificador na Configuração 4 para 20 key-points (cardinalidade = 20), peakthreshold = 0.01 e escala > 2 . . . . .	112
A.91 Avaliação do Classificador na Configuração 4 para 10 key-points (cardinalidade = 10), peakthreshold = 0.01 e escala > 2 . . . . .	112
A.92 Avaliação do Classificador na Configuração 4 para 5 key-points (cardinalidade = 5), peakthreshold = 0.01 e escala > 2 . . . . .	113
A.93 Avaliação do Classificador na Configuração 5 para 20 key-points (cardinalidade = 20), peakthreshold = 0.01 . . . . .	113
A.94 Avaliação do Classificador na Configuração 5 para 10 key-points (cardinalidade = 10), peakthreshold = 0.01 . . . . .	114

---

A.95 Avaliação do Classificador na Configuração 5 para 5 key-points (cardinalidade = 5), peakthreshold = 0.01 . . . . .	114
A.96 Avaliação do Classificador na Configuração 6 para 20 key-points (cardinalidade = 20), com máscara, peakthreshold = 0.01 e escala $\geq 2$ . . . . .	115
A.97 Avaliação do Classificador na Configuração 6 para 10 key-points (cardinalidade = 10), com máscara, peakthreshold = 0.01 e escala $\geq 2$ . . . . .	115
A.98 Avaliação do Classificador na Configuração 6 para 5 key-points (cardinalidade = 5), com máscara, peakthreshold = 0.01 e escala $\geq 2$ . . . . .	116
A.99 Avaliação do Classificador na Configuração 7 para 10 key-points (cardinalidade = 10), com máscara, peakthreshold = 0.02 . . . . .	116
A.100 Avaliação do Classificador na Configuração 7 para 5 key-points (cardinalidade = 5), com máscara, peakthreshold = 0.02 . . . . .	117
A.101 Avaliação do Classificador na Configuração 8 para 10 key-points (cardinalidade = 10), com máscara, peakthreshold = 0.02 e escala $\geq 2$ . . . . .	117
A.102 Avaliação do Classificador na Configuração 8 para 5 key-points (cardinalidade = 5), com máscara, peakthreshold = 0.02 e escala $\geq 2$ . . . . .	118
A.103 Avaliação do Classificador na Configuração 9 para 20 key-points (cardinalidade = 20), peakthreshold = 0.02 . . . . .	118
A.104 Avaliação do Classificador na Configuração 9 para 10 key-points (cardinalidade = 10), peakthreshold = 0.02 . . . . .	119
A.105 Avaliação do Classificador na Configuração 9 para 5 key-points (cardinalidade = 5), peakthreshold = 0.02 . . . . .	119
A.106 Avaliação do Classificador na Configuração 10 para 20 key-points (cardinalidade = 20), orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	120

A.107	Avaliação do Classificador na Configuração 10 para 10 keypoints (cardinalidade = 10), orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	120
A.108	Avaliação do Classificador na Configuração 10 para 5 keypoints (cardinalidade = 5), orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	121
A.109	Avaliação do Classificador na Configuração 11 para 10 keypoints (cardinalidade = 10), peakthreshold = 0.02 e escala $\geq 2$ . . . . .	121
A.110	Avaliação do Classificador na Configuração 11 para 5 keypoints (cardinalidade = 5), peakthreshold = 0.02 e escala $\geq 2$ . . . . .	122
A.111	Avaliação do Classificador na Configuração 12 para 20 keypoints (cardinalidade = 20), com máscara, orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	122
A.112	Avaliação do Classificador na Configuração 12 para 10 keypoints (cardinalidade = 10), com máscara, orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	123
A.113	Avaliação do Classificador na Configuração 12 para 5 keypoints (cardinalidade = 5), com máscara, orientação de keypoints $-\pi/2 < \theta < 3\pi/2$ . . . . .	123
A.114	Avaliação do Classificador na Configuração 13 para 20 keypoints (cardinalidade = 20), com máscara, orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	124
A.115	Avaliação do Classificador na Configuração 13 para 10 keypoints (cardinalidade = 10), com máscara, orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	124
A.116	Avaliação do Classificador na Configuração 13 para 5 keypoints (cardinalidade = 5), com máscara, orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	125
A.117	Avaliação do Classificador na Configuração 14 para 20 keypoints (cardinalidade = 20), orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	125

---

A.118	Avaliação do Classificador na Configuração 14 para 10 keypoints (cardinalidade = 10), orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	126
A.119	Avaliação do Classificador na Configuração 14 para 5 keypoints (cardinalidade = 5), orientação de keypoints $-3\pi/2 < \theta < \pi/2$ . . . . .	126



# Lista de Algoritmos

1	Algoritmo responsável pela extracção de características com os parâmetros definidos. . . . .	38
2	Algoritmo responsável pelo cálculo das distâncias entre descritores de duas imagens. . . . .	39
3	Algoritmo responsável pelo avaliação do critério aplicado ao reconhecimento. . . . .	40



# Acrónimos

<b>PCA</b>	<i>Principal Component Analysis</i>
<b>SIFT</b>	<i>Scale Invariant Feature Transform</i>
<b>DoG</b>	<i>Difference of Gaussian</i>
<b>LLS</b>	<i>Linear-Least-Squares</i>
<b>EER</b>	<i>Equal Error Rate</i>
<b>FNR</b>	<i>False Negative Rate</i>
<b>FPR</b>	<i>False Positive Rate</i>
<b>API</b>	<i>Application Programming Interface</i>
<b>UND</b>	<i>Universidade de Notre Dame</i>
<b>LoG</b>	<i>Laplacian of Gaussian</i>
<b>MSER</b>	<i>Maximally-Stable-Extremal-Regions</i>
<b>ROC</b>	<i>Receiver Operating Characteristic</i>
<b>AUC</b>	<i>Area Under Curve</i>
<b>SOCIA Lab</b>	<i>Soft Computing and Image Analysis Laboratory</i>
<b>PDF</b>	<i>Probability Density Function</i>
<b>BBF</b>	<i>Best-Bin-First</i>
<b>RANSAC</b>	<i>RANdom SAmple Consensus</i>

**LMS** *Least Mean Squares*

**UCLA** *University of California Los Angeles*

**VGG** *Visual Geometry Group*

**MatLab** *Matrix Laboratory*

# Glossário

***Equal Error Rate (EER)*** - É a taxa de erro para a verificação do funcionamento de um sistema para se ajustar o limiar para a aceitação / rejeição da decisão, é adaptada de tal forma a que as taxas de falsos positivos e falsos negativos se tornam iguais.

***True Negative Rate (TNR)*** - ou taxa de verdadeiros negativos é a taxa de orelhas detectadas como sendo de indivíduos diferentes, sendo que são de indivíduos diferentes.

***False Positive Rate (FPR)*** - ou taxa de falsos positivos é a taxa de orelhas detectadas como pertencentes ao mesmo indivíduo, não sendo pertencentes ao mesmo indivíduo.

***False Negative Rate (FNR)*** - ou taxa de falsos negativos é a taxa de orelhas detectadas como orelhas de indivíduos diferentes, sendo do mesmo indivíduo.

***True Positive Rate (TPR)*** - ou taxa de verdadeiros positivos é a taxa de orelhas detectadas como pertencentes ao mesmo indivíduo, sendo mesmo indivíduo.

***Receiver Operating Characteristic (ROC)*** - ou simplesmente curva ROC é a representação gráfica da taxa de verdadeiros positivos com a taxa de falsos positivos.



# Capítulo 1

## Introdução

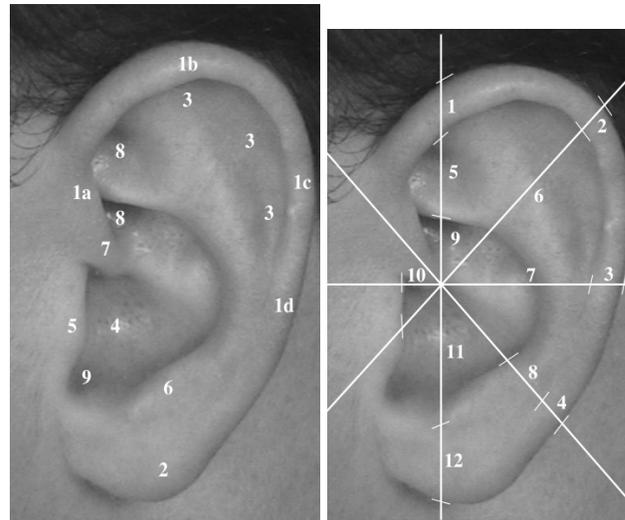
### 1.1 Objectivo

O objectivo deste projecto é o desenvolvimento de um sistema de detecção de orelhas humanas.

No domínio dos sistemas biométricos, a detecção de orelhas humanas manifesta um potencial inicialmente reconhecido por um inspector criminal Alphonse Bertillon. [9]

No campo da Visão Computacional, comparando com técnicas mais populares de reconhecimento, como a face, o olho, ou as impressões digitais, o reconhecimento biométrico da orelha recebeu uma menor atenção. No entanto as orelhas desempenham um papel significativo na ciência forense há muito tempo, particularmente nos Estados Unidos, onde um sistema de classificação baseado em medidas, extraídas manualmente, foi desenvolvido por Iannarelli, à mais de 40 anos. [23]

Não é considerado um método cientificamente comprovado e foi recentemente contestado em tribunal. [14]



(a) 1- arco helix; 2- lóbulos; 3- anti-helix; 4- concha; 5- tragus; 6- antitragus; 7- cruz de helix; 8- fossa triangular; 9- incisão intertrágica.

(b) medidas

**Figura 1.1:** (a) Anatomia (b) Localizações das medidas antropomórficas usadas no sistema de Iannarelli.

O "Sistema de Iannarelli" baseia-se em 12 medidas ilustradas nas seguintes figuras 1.1(a) 1.1(b) [23].

Dentro deste contexto, o processo de reconhecimento biométrico de orelhas humanas tem crescido significativamente em importância nos últimos anos.

O propósito dos investigadores não está voltado apenas para a detecção da orelha, mas também, das suas características intrínsecas (helix, anti-helix, cruz-helix, concha, etc).

<i>Biometric Trait</i>	<i>Permanence over time</i>
Fingerprint (Minutia)	000000
Signature (dynamic)	0000
Facial structure	00000
Iris pattern	000000000
Retina	00000000
Hand geometry	0000000
Finger geometry	0000000
Vein structure of the back of the hand	000000
Ear form	000000
Voice (Tone)	000
DNA	000000000
Odor	000000?
Keyboard strokes	0000
Comparison: Password	00000

**Figura 1.2:** *Variação no tempo, da permanência de diferentes dados biométricos. A melhor permanência possui mais 0-símbolos.[10]*

Uma empresa especializada na área da biometria humana *Bromba GmbH* (2008)[10] fez um estudo comparativo de diferentes dados biométricos, incluindo a orelha.

A orelha manifesta o seu valor biométrico pela sua consistente estrutura ao longo do tempo, desde o nascimento até idade avançada. A orelha não sofre alterações de diferentes expressões faciais, e a sua localização facilita a previsibilidade do fundo, em relação à face que normalmente precisa de uma captura controlada de fundo.

A extracção de dados não está associada a problemas de higiene como pode acontecer com a leitura de impressões digitais, e não é propensa a criar ansiedade ou stress, como pode acontecer com extracção de medidas de retina ou íris.

A orelha é amplamente comparada à íris, retina e impressões digitais, sendo mais fácil de capturar.

## 1.2 Motivação

Desde tempos imemoriais que o ser humano tenta desenvolver a sua tecnologia de modo a torná-la extensão, ou mesmo substituta, do seu próprio corpo. A ideia de máquinas e seres inanimados automatizados habita o imaginário humano, desde longa data. Tais progressos assentam em motivações diversas, desde a segurança e comodidade, a económico-sociais e económico-comerciais, à pura curiosidade do Homem.

Inserido num processo de verosimilhança e aperfeiçoamento da máquina humana, o estudo da visão exerce um grande fascínio. Neste ponto, o interesse está não só na percepção do estímulo luminoso, mas também, no reconhecimento lógico dos objectos circundantes. Muitos são os animais dotados de visão, sendo sistemas mais simples ou complexos mais específicos ou gerais. A visão é usada para inúmeras tarefas, muitas delas relacionadas com a sobrevivência e perpetuação da espécie. Os sistemas de visão de mamíferos superiores, como os primatas, apresentam grande complexidade e mecanismos sofisticados de reconhecimento de padrões.

Além do interesse pela visão artificial, os processos automatizados de reconhecimento de padrões ganham importância frente à crescente oposição entre a velocidade da informação e a limitação temporal. Esses processos requerem eficiência, eficácia e agilidade crescentes. Nesse sentido, as tarefas visuais automatizadas tornam-se excelentes soluções, especialmente pelo suporte conferido e pelo avanço da computação e dos sistemas automáticos.

Os frutos da pesquisa nessa área do conhecimento são compartilhados no campo científico e comercial. Sob o aspecto financeiro, a aplicação crescente ocorre em sistemas de segurança e vigilância, motivada pela violência social, tráfico, terrorismo, etc.

Na simulação do comportamento visual humano, denominado genericamente por Visão Computacional, o reconhecimento do padrão de uma orelha numa pessoa constitui uma tarefa com crescente potencial.

Dos primeiros reconhecer o potencial da orelha humana, e a apresentar um "modelo biométrico" foram Burge e Burger [13][12]. O seu modelo considerava comparações entre grafos, num diagrama de Voroni de curvas extraídas de um mapa de arestas, do detector de arestas Canny.

Hurley et al., [22] usaram um sistema denominado "Force Field Feature Extraction", usando "campos de força", focando nos seus "potenciais poços" e "potenciais canais" como características [21].

Yan et al.,[34] explorou a estrutura 3D da orelha humana. Yuizono et all[37] tratou o problema como uma tarefa de optimização e desenvolveu uma procura local genética. Moreno et al., [4] usou diferentes combinações de vários classificadores neuronais.

Principal Component Analysis (PCA) também foi utilizado em diferentes estudos [35] [24], no entanto PCA não apresenta características invariantes, logo depende da aquisição e pré-processamento para enquadrar e alinhar dados.

A comunidade científica tem publicado um grande número de trabalhos nesta área. A biometria de orelhas humanas, ou seja, a técnica de medidas de características intrínsecas à orelha é, actualmente, uma das frentes de pesquisa na área da Visão Computacional.

*"Biometria são medidas de quantidades ou padrões biológicos, mas também a medição de características de um indivíduo, tais como impressões digitais, que podem identificar ou autenticar univocamente uma pessoa. Usada desta forma, uma biometria é uma senha que não pode ser esquecida, perdida ou roubada."*[16]

A indexação e a recuperação de imagens contendo actividades humanas, as actividades computacionais em tempo real, como a visão de robôs, poderão incluir técnicas de detecção de orelhas humanas para obter respostas rápidas e tomar decisões num curto espaço de tempo. Da mesma forma, sistemas de segurança e vigilância, de identificação pessoal também requerem técnicas biométricas.

## 1.3 Abordagem

Para o tema proposto, foi utilizada a abordagem descrita por David G. Lowe [28] [29], devido à sua aproximação baseada em modelos, capitalizando numa estrutura específica e com a vantagem de ser robusta com ruído e oclusão.

A robustez referida anteriormente é alcançada principalmente graças ao método extracção de características invariantes, *Scale Invariant Feature Transform* (SIFT).

Esta aproximação transforma uma imagem numa larga colecção de vectores de características, as quais apresentam invariância em relação à translação, escala, rotação de uma imagem, e parcialmente invariantes a mudanças de luminância, projecção 3D e enviesamento (transformada afim). As características são bem localizadas no domínio espacial e frequência, reduzindo a probabilidade de disrupção por oclusão, bloqueio ou ruído, promovendo a sustentabilidade de reconhecimento de objectos ou cenas.

O custo de extracção desta características é minimizado tendo em conta uma aproximação de filtragem em cascata, em que operações computacionalmente mais custosas são apenas aplicadas a localizações que passam num teste inicial.

O reconhecimento processa-se correspondendo características individuais com uma base dados de características de objectos conhecidos, implementando um algoritmo de procura vizinho-mais-próximo, seguido de por uma filtragem através da função *transformada de Hough*, identificando "clusters" pertencentes a um determinado objecto, e verificando poses consistentes nos parâmetros através do *Linear-Least-Squares* (LLS). Permitindo um reconhecimento robusto em situações de oclusão ou ruído.

De seguida são enumerados os passos computados, na extracção/detecção destas características.

1. **Detecção dos Extremos no Espaço-Escala:** a primeira etapa computacional procura em todas as escala do Espaço-Escala e em todas as molduras de imagem, é implementada usando a função DoG , na identificação de potenciais pontos de interesse, que são invariantes à escala e orientação.
2. **Localização de Keypoints (pontos-chave):** a cada moldura candidata, é enquadrado um modelo detalhado para determinar a localização e escala. Keypoints são seleccionados baseados em medidas de estabilidade.
3. **Atribuição de Orientações:** uma ou mais orientações são atribuídas a cada localização de keypoint, baseadas em gradientes de direcções da imagem local. Todas as operações futuras nos dados da imagem serão transformadas relativamente à atribuição de escala, orientação e localização de cada característica, logo manifestando invariância relativamente a estas transformações;
4. **Descriptor de Keypoints (pontos-chave):** gradientes de imagens locais são medidos na escala seleccionada na região à volta de cada keypoint. Estes são transformados numa representação que permite algum nível de distorção da forma e mudanças de iluminação.

Esta aproximação foi denominada SIFT, já que transforma dados de imagem em coordenadas invariantes à escala relativas a características locais.

## 1.4 Organização do relatório

Este relatório encontra-se dividido por capítulos. O primeiro capítulo apresenta uma introdução ao tema, frisando vários aspectos com ele relacionados.

**Capítulo 2 - Método de Extração/Reconhecimento de Características de Orelhas em Imagens Digitais** Este capítulo apresenta em maior detalhe o método utilizado na extração das características de imagens digitais de orelhas humanas, sendo que foi dividido em etapas, dependendo dos procedimentos a utilizar.

**Capítulo 3 - Implementação** Apresentação do método e a API que foi usada na elaboração deste projecto e alterações efectuadas na procura da obtenção de melhores resultados.

**Capítulo 4 - Experiências e Resultados** Neste capítulo descreve-se as experiências realizadas, bem como os resultados obtidos pelos diversos métodos.

**Capítulo 5 - Conclusão e trabalho futuro** Como referido, pretende-se retirar algumas conclusões / argumentos sobre a realização deste projecto bem como descrever o que pode ser acrescentado a este trabalho na tentativa da obtenção de melhores resultados.

## Capítulo 2

# Método de Extracção/Reconhecimento de Características de Imagens Digitais

### 2.1 Banco de Imagens Digitais

O banco de imagens digitais usado, cedido pela UND, é composto por 464 imagens de 114 indivíduos capturadas no ano de 2002, cujas imagens apresentam algum grau de rotação.

Algumas imagens foram retiradas do conjunto inicial, 9 imagens de um indivíduo e 2 imagens doutro indivíduo. A escolha de retirar estas imagens, 11 no total, foi considerada por uma melhor distribuição de imagens por indivíduo.

Esta base de dados foi cedida pela UND para efectuar as experiências relacionadas com o projecto.

## Método de Extracção/Reconhecimento de Características de Imagens Digitais

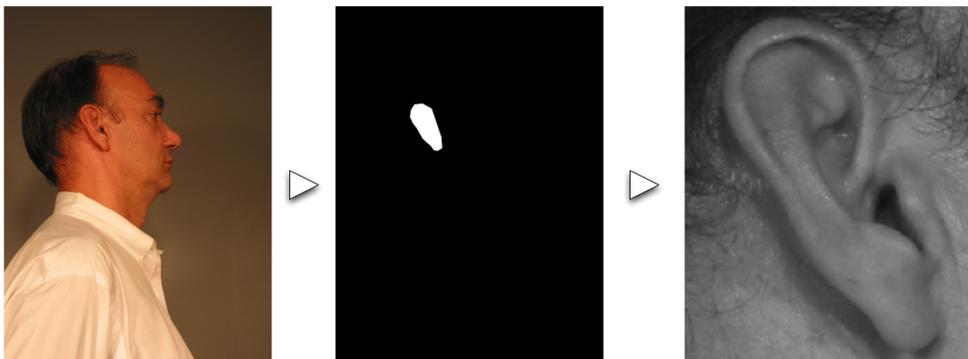
As imagens foram divididas em Bases de Dados diferentes com o propósito de melhor avaliar o modelo de reconhecimento, sendo que assim é mais fácil diferenciar e avaliar os resultados relativos a um determinado banco de dados.

A seguinte tabela 2.1 ilustra a denominação atribuída ao conjunto de dados, bem como o total de imagens por conjunto, se o conjunto apresenta rotação de cena e o nº de imagens por indivíduo.

Denominação	total de imagens	rotação	imagens por indivíduo
Base de Dados A	453	sim	3,4,6
Base de Dados B	311	sim	3,4,6
Base de Dados C	142	não	3,4,6

**Tabela 2.1:** distribuição das imagens da Base de Dados de imagens da UND.

Na figura seguinte encontra-se ilustrado o processo de selecção (*segmentação*) da área da orelha 2.1.



**Figura 2.1:** Exemplo do processo efectuado na captura da área da orelha. Após segmentação manual obtemos uma "máscara" dessa orelha, em que podemos, normalizando, seleccionar uma secção da imagem, o SIFT não apresenta sensibilidade a esta normalização já que apresenta invariância à escala

## 2.2 Pesquisa Relacionada

O desenvolvimento do reconhecimento de imagens, usando um conjunto de "*pontos locais de interesse*", foi inicialmente estudado por *Moravec(1981)* [33].

O detector de *Moravec* foi melhorado por *Harris e Stevens(1988)*[19], tornando-o mais robusto sob pequenas variações na imagem e na proximidade de arestas.

Apesar destes "*detectores de características*" serem denominados "*detectores de arestas*", não estão apenas a detectar arestas, mas também a seleccionar localizações da imagem que contenham gradientes grandes em todas as direcções numa escala predeterminada.

O trabalho inovador de *Schmid e Mohr(1997)*[36] mostrou que o reconhecimento de características locais invariantes poderia ser estendido ao reconhecimento geral de imagens em que uma característica era reconhecida numa base de dados de imagens grande.

O "*detector de arestas Harris*" também foi usado na selecção de "*pontos de interesse*", mas em vez de fazerem o reconhecimento com uma janela de correlação, usaram descritores invariantes à rotação numa região local da imagem. Esta inovação permitiu o reconhecimento de características sob uma orientação arbitrária entre duas imagens.

Ainda mais, demonstrou que múltiplas características poderiam permitir um reconhecimento sob oclusão e ruído, identificando "*clusters*" consistentes das características reconhecidas.

O problema de identificação de uma escala apropriada e consistente à extracção de características foi estudado a fundo por *Lindeberg (1993,1994)* [26] [27].

Mais recentemente, vários artigos têm sido publicados focando a extensão de características locais de forma a manifestarem completa invariância à transformação afim Baumberg(2000) [6]; Mikolajczyk e Schmid(2002) [31]; Brown e Lowe(2002) [11].

Muitos outros tipos de características têm sido propostos para uso no reconhecimento, alguns dos quais poderiam ser usados em junção às características descritas fornecendo uma maior taxa de reconhecimento sob diferentes circunstâncias.

Outra classe de características, é a que faz uso dos contornos da imagem ou limites de região de imagem, o que faz com que sejam menos susceptíveis a disrupções de fundo perto do limite do objecto. Matas et al.,(2002) [30] mostrou com seu *Maximally-Stable-Extremal-Regions* (MSER) que consegue produzir um grande número de características de reconhecimento com boa estabilidade.

Mikolajczyk et al.,(2003) [32] desenvolveram um novo descritor que usa arestas locais, ignorando arestas adjacentes não relacionadas, fornecendo a habilidade de encontrar características estáveis mesmo perto de fronteiras de imagem de forma estreita sobrepostas num fundo desordenado.

A aproximação a características locais pode facilmente incorporar novos tipos de características, sendo que características extra contribuem para a robustez quando fornecem um melhor reconhecimento, senão pouco mal advém da sua adição a não ser o seu custo computacional.

Assim sendo sistemas futuros provavelmente combinarão vários tipos de características.

## 2.3 Método de Extração de Características

Como previamente enunciado o método de extração de características escolhido é o SIFT.

O SIFT encontra-se dividido em 4 etapas principais:

1. **Detecção dos Extremos no Espaço-Escala.**
2. **Localização de Keypoints.**
3. **Atribuição de Orientações.**
4. **Descritores de Keypoints.**

Cada etapa encontra-se explanada nas seguintes secções.

### 2.3.1 Detecção dos Extremos no Espaço-Escala

Esta primeira etapa de detecção de keypoints serve para identificar localizações e escalas que poderão ser repetidamente atribuídas sob diferentes vistas do mesmo objecto.

Detectar localizações que são invariantes à escala da imagem é realizado através da procura por características estáveis através de todas as escalas possíveis, usando um função de escala contínua denominada por *Espaço-Escala*. [39]

Foi demonstrado por *Koenderink(1984)* [25] e *Lindeberg(1994)* [27] sob uma variedade de suposições que o único *kernel* aplicável ao *Espaço-Escala* é a *função Gaussiana*. Logo, o *Espaço-Escala* de uma imagem é definido como a função,  $L(x, y, \sigma)$ , que é produzida da convolução de um *Gaussiano* de escala variável,  $G(x, y, \sigma)$ , com a imagem,  $I(x, y)$  :

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y),$$

onde  $*$  é a operação de convolução em  $x$  e  $y$ , e

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

Afim de detectar eficientemente localizações estáveis de keypoints no Espaço-Escala, foi proposto por *Lowe(1999)* [28], usar os extremos do Espaço-Escala na função DoG e a convolução com a imagem  $D(x, y, \sigma)$ , que pode ser computada pela diferença de duas escalas adjacentes separadas por uma constante multiplicativa de factor  $k$ :

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \quad (2.1)$$

$$= L(x, y, k\sigma) - L(x, y, \sigma) \quad (2.2)$$

Existem inúmeras razões para escolher esta função. Primeiro, é particularmente eficiente de computar, já que as imagens suavizadas,  $L$ , precisam de ser computados, em qualquer caso, para descrição da característica no Espaço-Escala, e  $D$  pode então ser computado por simples subtração de imagem.

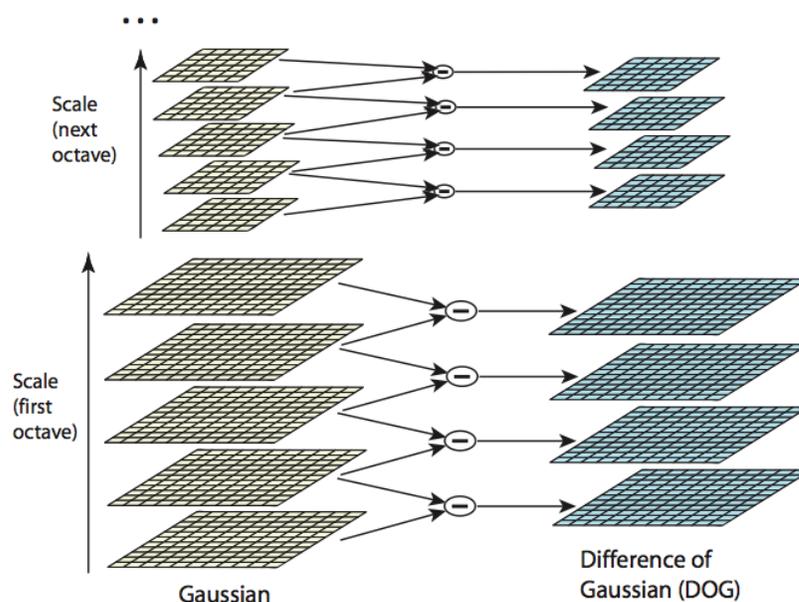
Em suma, a função DoG fornece uma boa aproximação ao *Laplacian of Gaussian* (LoG) normalizado à escala,  $\sigma^2 \nabla^2 G$ , como estudado por *Lindeberg (1994)* [27]. Lindeberg mostrou que a normalização do Laplaciano com um factor de  $\sigma^2$  é requerido a uma verdadeira invariância à escala. Em comparações experimentais detalhadas, *Mikolajczyk (2002)* [31] conclui que a máxima e mínima de  $\sigma^2 \nabla^2 G$  produz as características de imagem mais estáveis, comparando com um leque de possíveis funções de imagem, como o gradiente, função de arestas, Hessian ou Harris.

A relação entre  $D$  e  $\sigma^2 \nabla^2 G$  pode ser compreendida da equação de difusão térmica (parametrizada em termos de  $\sigma$  em vez de o mais usual  $t = \sigma^2$ ):

$$\frac{\partial G}{\partial \sigma} = \sigma \nabla^2 G$$

Daqui, podemos ver que  $\nabla^2 G$  pode ser computado da aproximação de subtrações finitas a  $\partial G / \partial \sigma$ , usando a subtração de escalas subjacentes a  $k\sigma$  a  $\sigma$ :

$$\sigma \nabla^2 G = \frac{\partial G}{\partial \sigma} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma}$$



**Figura 2.2:** Para cada oitava do espaço-escala, à imagem inicial é aplicada a convolução com Gaussianos para produzir o conjunto de imagens de espaço-escala mostradas à esquerda. Imagens Gaussianas adjacentes são subtraídas afim de produzir a DoG apresentada à direita. Após cada oitava, a imagem Gaussiana é reduzida em amostragem por um factor de 2, e o processo repetido. [29]

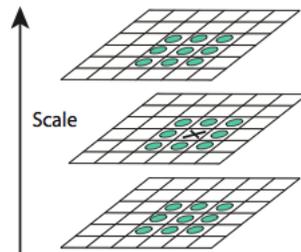
e então,

$$\sigma \nabla^2 G = \frac{\partial G}{\partial \sigma} \approx (k-1) \sigma^2 \nabla^2 G$$

Tal demonstra que quando a função DoG possui escalas diferenciadas por um factor constante, incorpora a normalização de escala  $\sigma^2$ , requerida pelo Laplaciano invariante à escala.

### Detecção dos extremo local

De modo a detectar o máximo e mínimo local de  $D(x, y, \sigma)$ , cada ponto amostral é comparado com os seus oito vizinhos na imagem actual e com os nove vizinhos da escala acima e abaixo.



**Figura 2.3:** A máxima e mínima da DoG é detectada comparando um pixel (marcado com X) aos seus 26 vizinhos em regiões de 3x3 na escala actual em escalas adjacentes (marcados com círculos). [29]

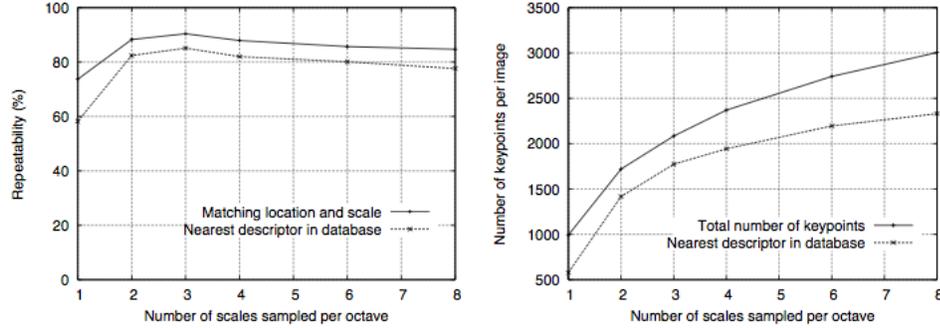
É seleccionada apenas se for maior ou menor que todos estes vizinhos. O custo desta verificação é relativamente baixo tendo em conta que maior parte dos pontos amostrais são eliminados aquando das primeiras verificações.

Uma questão importante é a determinação da frequência de amostragem na imagem e domínios da escala. Infelizmente acontece que não existe um mínimo de espaçamento de amostras que detecte todos os extremos, tanto que estes extremos podem encontrar-se arbitrariamente próximos.

Pela análise da figura 2.4 podemos concluir que a melhor repetibilidade é alcançada com 3 escalas por oitava. Portanto a solução estabelecida corresponde a um compromisso entre a eficiência e a completude.

### 2.3.2 Localização de Keypoints

Quando um candidato a keypoint é encontrado, comparando um pixel ao seus vizinhos, o próximo passo é realizar um ajuste detalhado aos dados para localização, escala e razão da curvatura principal. Esta informação permite que sejam rejeitados pontos que possuem um baixo contraste (sendo assim sensíveis ao ruído) ou que estão deficientemente localizados ao longo de uma aresta.



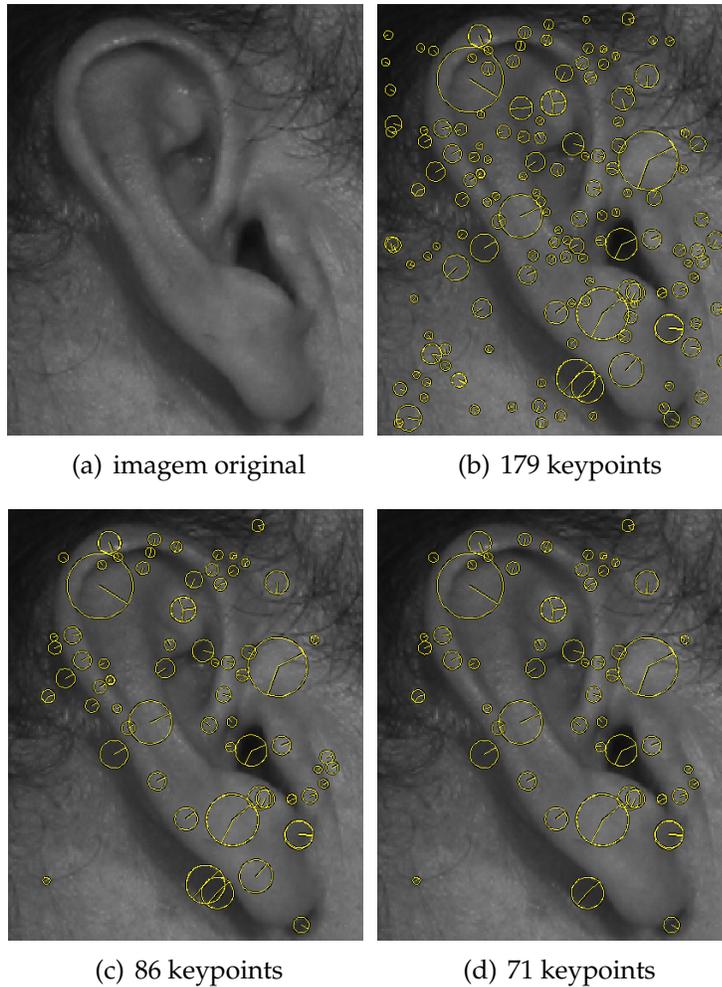
**Figura 2.4:** A linha superior do gráfico da esquerda mostra a percentagem de keypoints que são repetidamente detectados na mesma localização e escala na imagem transformada como função do número de escalas processadas por oitava. A linha inferior mostra a percentagem de keypoints que têm os seus descritores correctamente reconhecidos numa base de dados grande. O gráfico da direita mostra o número total de keypoints detectados numa imagem como função do número escalas por oitava.

A implementação inicial desta aproximação de Lowe (1999) [28] simplesmente localizava keypoints na localização e escala no ponto central da amostra.

No entanto, recentemente Brown desenvolveu um método (Brown e Lowe (2002)) [11] para ajustar uma função 3D quadrática aos pontos de amostra locais afim de determinar a posição interpolada do máximo, e as suas experiências mostraram que fornecem um substancial melhoramento do reconhecimento e estabilidade. Esta aproximação usa a expansão de Taylor da função de Espaço-Escala,  $D(x, y, \sigma)$ , deslocada para que a origem esteja na amostra do ponto:

$$D(X) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x. \quad (2.3)$$

onde  $D$  e as suas derivadas são avaliadas na amostra do ponto e  $X = (x, y, \sigma)^T$  é o desvio a partir deste ponto.



**Figura 2.5:** Esta figura mostra vários parâmetros de extração de keypoints. (a) imagem de orelha original. (b) 179 keypoints iniciais localizados no máximo e mínimo da função DoG. Keypoints são mostrados como vectores indicando escala, orientação e localização. (c) Depois de aplicar um threshold ao contraste mínimo,restam 86 keypoints. (d) Os 71 keypoints finais adicionando um threshold à razão da curvatura principal

A localização do extremo,  $\hat{X}$  é determinado usando a derivada desta função em relação a  $X$  e pondo-a a zero, dando:

$$\hat{X} = -\frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x} \quad (2.4)$$

Como sugerido por *Brown*, a Hessiana e a derivada de  $D$  são aproximadas usando a diferença de pixels vizinhos no ponto amostral.

O sistema linear 3x3 resultante pode ser resolvido com custo mínimo.

Se o deslocamento de  $\hat{X}$  é maior que 0.5 em qualquer dimensão, querendo isto dizer que o extremo encontra-se mais perto de um ponto amostral diferente. O deslocamento final  $\hat{X}$  é adicionado à localização do seu ponto amostral, sendo a estimativa interpolada para a localização do extremo. O valor da função no extremo,  $D(\hat{X})$ , é útil para rejeitar extremos instáveis com baixo contraste. Isto é obtido substituindo a equação (2.4) na equação (2.3), ficando:

$$D(\hat{X}) = D + \frac{1}{2} \frac{\partial D^T}{\partial x} \hat{x} \quad (2.5)$$

Para a implementação de *Lowe*, todos os extremos com um valor de  $|D(\hat{X})|$  menor que 0.03 são descartados (assumindo que valores de pixels num limite de  $[0,1]$ ).

### Eliminação de respostas nas arestas

Para efeitos de estabilidade, não é suficiente eliminar keypoints com baixo contraste. A função DoG apresenta fortes respostas ao longo de arestas, mesmo que a localização ao longo da aresta seja deficientemente determinada e portanto susceptíveis a pequenas quantidades de ruído.

Um pico deficientemente definido na função DoG terá uma grande curvatura principal ao longo da aresta, mas uma pequena na direcção perpendicular.

A curvatura principal pode ser computada da matriz Hessiana de 2x2,  $\mathbf{H}$ , computada na localização e escala do keypoint.

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

As derivadas são estimadas tirando as diferenças de pontos vizinhos. Os *eigenvalues* de  $\mathbf{H}$  são proporcionais à principal curvatura de  $D$ . Inspirado na aproximação de *Harris e Stevens (1988)* [19], podemos evitar computar explicitamente os *eigenvalues* já que apenas estamos interessados com a sua razão. Seja  $\alpha$  o *eigenvalue* com a maior magnitude e  $\beta$  o de menor magnitude. Então, podemos computar a soma dos *eigenvalues* do índice de  $\mathbf{H}$  e o seu produto do determinante:

$$\begin{aligned} Tr(\mathbf{H}) &= D_{xx} + D_{yy} = \alpha + \beta, \\ Det(\mathbf{H}) &= D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \end{aligned}$$

No caso improvável de o determinante ser negativo, as curvaturas apresentam sinais diferentes logo os pontos serão descartados já que não pertencem ao extremo. Seja  $r$  a razão entre o maior *eigenvalue* e o menor, tal que  $\alpha = r\beta$ . Então:

$$\frac{Tr(\mathbf{H})^2}{Det(\mathbf{H})} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r\beta + \beta)^2}{r\beta^2} = \frac{(r + 1)^2}{r},$$

Dependendo apenas da razão entre os *eigenvalues* em vez dos seus valores individuais. A quantidade  $(r + 1)^2/2$  está no seu mínimo quando os dois *eigenvalues* são iguais e aumenta com o  $r$ . Portanto, para verificar que a razão de curvaturas principais está abaixo de dado limite,  $r$ , basta verificar:

$$\frac{Tr(\mathbf{H})^2}{Det(\mathbf{H})} < \frac{(r + 1)^2}{r}$$

A implementação de *Lowe* usa um valor de  $r = 10$ , que elimina keypoints que possuem uma razão entre as principais curvaturas maiores que 10.

### 2.3.3 Atribuição de Orientações

Atribuindo uma orientação constante a cada keypoint baseado em propriedades locais da imagem, o descritor do keypoint pode ser representado relativamente à sua orientação e portanto apresentar invariância à rotação de imagem.

A escala do keypoint é usada na selecção da imagem suavizada Gaussiana,  $L$ , com a escala mais próxima, para que todas as computações sejam realizadas de maneira que sejam invariantes à escala. Para cada imagem,  $L(x, y)$ , nesta escala, o gradiente de magnitude,  $m(x, y)$ , e a orientação,  $\theta(x, y)$ , são precomputados usando a diferença de pixels:

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

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y)))$$

Um histograma de orientações é formado a partir do gradiente de orientações de pontos amostrais dentro de uma região à volta do keypoint. O histograma de orientações possui 36 porções, cobrindo 360 graus de orientações. Cada amostra adicionada ao histograma é pesada pela seu gradiente de magnitude e por janela Gaussiana pesada circular com  $\sigma$  de 1.5 vezes a escala do keypoint.

Picos no histograma de orientações correspondem a direcções dominantes dos gradientes locais.

O pico mais alto no histograma é detectado, e todos os outros picos locais que atinjam 80% do pico mais alto também são usados para criar um keypoint com essa orientação. Portanto, para localizações com múltiplos picos de magnitudes semelhantes, irão haver múltiplos keypoints criadas com a mesma localização e escala, mas diferentes orientações. Apenas cerca de 15% dos pontos atribuídos contém múltiplas orientações, mas estes contribuem significativamente para a estabilidade do reconhecimento. Finalmente uma parábola é ajustada aos 3 valores do histograma mais próximos de cada pico, para interpolar a posição do pico para melhor precisão.

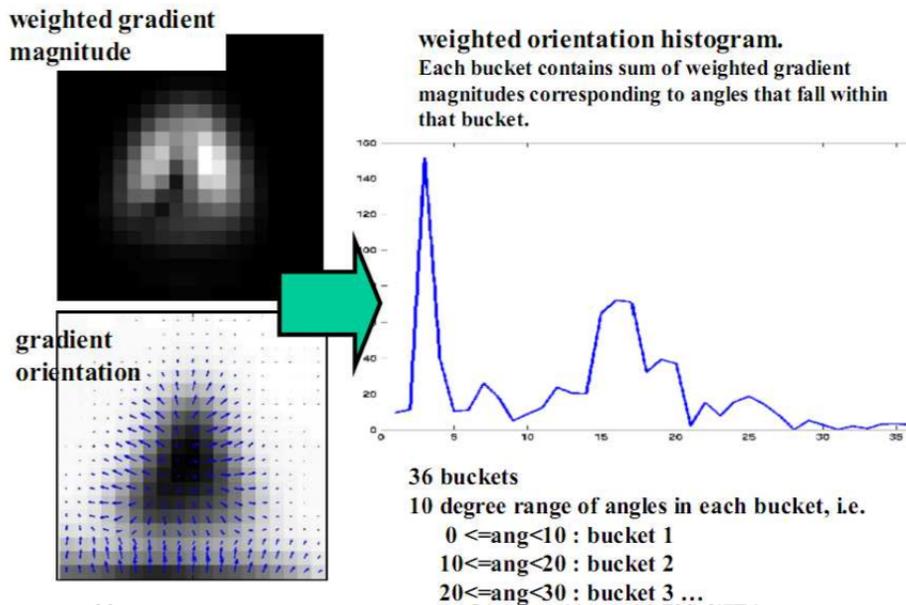


Figura 2.6: Histograma de orientações dominantes dos gradientes [7]

### 2.3.4 Descritores de Keypoints

#### Descritor da imagem local

As operações anteriores atribuíram uma localização, uma escala, uma orientação a cada keypoint. Estes parâmetros impõem um sistema de coordenadas 2D local repetível em que é descrita uma região local da imagem, e portanto fornecendo invariância a estes parâmetros.

O passo seguinte é computar um descritor para a região local da imagem que é altamente distinto, no entanto é tão invariante quanto possível às restantes variações, como mudança de iluminação ou perspectiva 3D. Uma aproximação óbvia seria processar as intensidades da imagem local à volta dos keypoints na escala apropriada e correspondendo estes a uma medida de correlação normalizada.

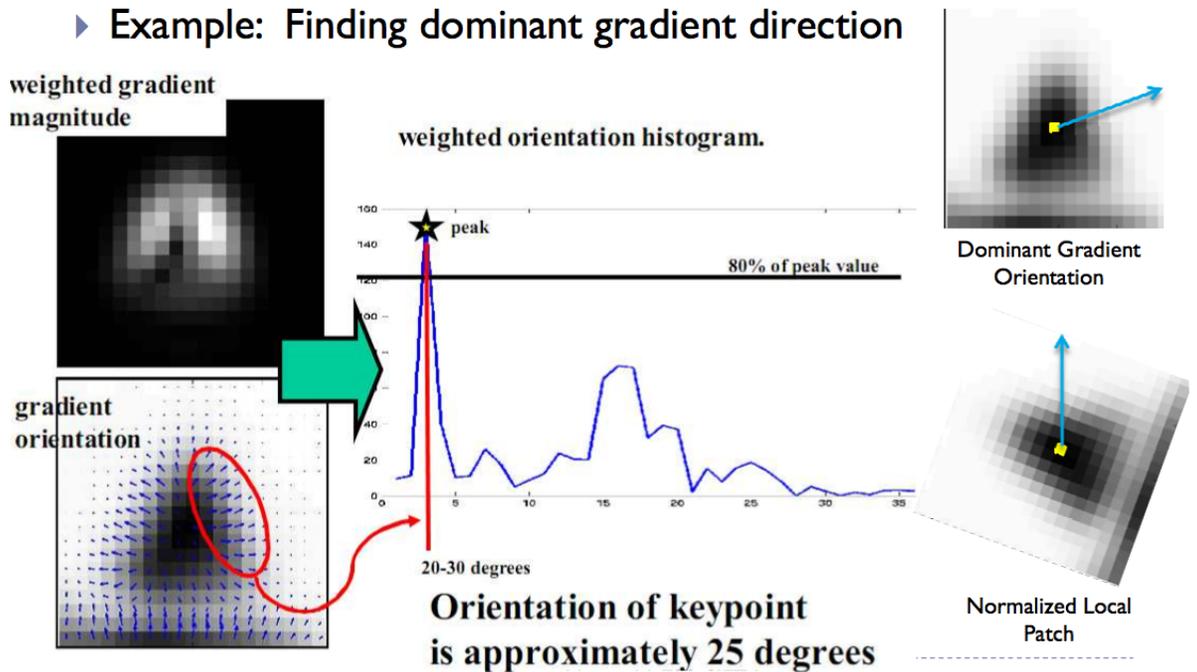
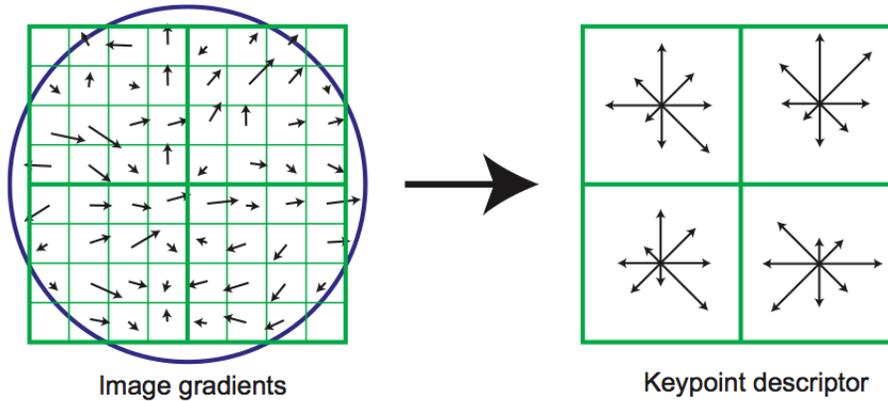


Figura 2.7: Múltiplas direcções dominantes do histograma.[7]

No entanto, a simples correlação de fragmentos de imagem é altamente sensível a mudanças que causam um incorrecto registo de amostra, tais como transformações afim, ou perspectiva 3D.

Uma melhor aproximação foi demonstrada por *Edelman, Intrator, Poggio (1997)* [17]. A representação proposta é baseada num modelo biológico da visão, particularmente em neurónios complexos no córtex visual primário. Estes neurónios complexos reagem a um gradiente numa orientação e frequência espacial específicas, mas a localização deste na retina pode sofrer alterações sob um pequeno campo receptivo, em vez de ser precisamente localizado.

*Edelman et al.* [17] hipotetizaram que a função destes neurónios complexos era permitir a correspondência e o reconhecimento de objectos 3D, dum variedade de pontos de vista.



**Figura 2.8:** O descritor de keypoint é criado computando primeiro o gradiente de magnitude e orientação de cada ponto amostral numa região à volta da localização do keypoint, como mostrada à esquerda. Estes são pesados por uma janela Gaussiana circular. Estas amostras são acumuladas em histogramas de orientação resumindo o conteúdo em sub-regiões de 4x4, mostrado à direita, com o comprimento de cada seta correspondendo à soma das magnitudes do gradiente próximo dessa orientação na região. Esta figura mostra um descritor de 2x2 computado a partir de um conjunto de amostras de 8x8, enquanto a implementação de Lowe usa descritores de 4x4 de um conjunto de amostras de 16x16.

### Representação do Descritor

A figura 2.8 ilustra a computação do descritor de um keypoint. Primeiro o gradiente de magnitudes e orientações é processado à volta da localização do keypoint, usando a escala do keypoint para seleccionar o nível de suavização do Gaussiano da imagem. A fim de obter invariância à orientação as coordenadas do descritor e as orientações do gradiente são rodadas relativamente à rotação do keypoint. Para efeitos de eficiência os gradientes são pre-computados para todos os níveis da pirâmide como descrito na secção 2.3.1. Estes são ilustrados com pequenas setas em cada localização da amostra como ilustra a figura 2.8.

Uma função Gaussiana pesada com  $\sigma$  igual a metade da largura da janela do descritor é usado para atribuir o peso da magnitude de cada ponto amostral. Isto é ilustrado com a janela circular do lado esquerdo da figura 2.8.

O propósito desta janela Gaussiana é evitar mudanças subitas no descritor com pequenas mudanças na posição da janela, dando menos ênfase a gradientes que se encontram afastados do centro do descritor, já que estes são mais afectados por um registo incorrecto.

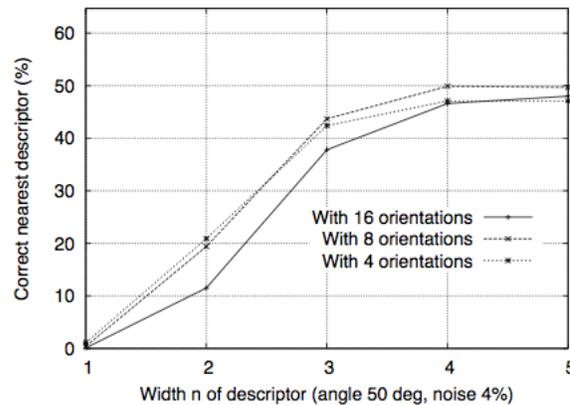
O descritor é formado por um vector contendo os valores de todas as entradas do histograma de orientações, correspondendo ao comprimento das setas, no lado direito da figura 2.8. A figura mostra um array de 2x2 de histogramas de orientação, enquanto a implementação de Lowe mostra que os melhores resultados são alcançados com um array de 4x4 de histogramas com 8 fatias de orientações em cada.

Finalmente o vector de características é modificado de modo a reduzir os efeitos de variações de iluminação. De inicio o vector é normalizado a um comprimento unitário. Uma mudança no contraste da imagem, em cada valor de um pixel é multiplicado por uma constante, não irá afectar os valores de gradiente já que são computados através da diferença de pixels. Assim, o descritor é invariante à transformação afim na iluminação.

### Testes ao Descritor

Existem dois parâmetros que podem ser usadas para variar a complexidade: o número de orientações  $r$ , em histogramas e a largura  $n$ , do array de  $n \times n$  do histograma de orientações. O tamanho do vector descritor resultante é  $rn^2$ . À medida que a complexidade do descritor aumenta, será mais capaz de discriminar numa base de dados grande, mas será também mais sensíveis a distorções na forma e oclusão.

A figura 2.9 mostra resultados experimentais em que o número de orientações e o tamanho do descritor são variáveis. O gráfico foi gerado para uma transformação do ponto em que a superfície plana está inclinada a 50 graus do ponto de observação e adicionada de 4% de ruído. Este caso encontra-se próximo do limite fidedigno de reconhecimento, já que é nestes casos mais difíceis que a performance do descritor é mais importante.



**Figura 2.9:** Este gráfico mostra a percentagem de keypoints que fornecem um reconhecimento correcto com uma base de dados de 40.000 keypoints, como função de largura de  $n \times n$  descritores de keypoints e um número de orientações em cada histograma. O gráfico é computado para imagem com ponto de vista de transformada afim de 50 graus com adição de 4% de ruído

Os resultados mostram a percentagem de keypoints que encontram uma correspondência correcta ao vizinho mais próximo numa base de dado de 40.000 keypoints.

O gráfico mostra que um único histograma de orientação ( $n = 1$ ) é bastante fraco a discriminar, mas os resultados continuam a melhorar até alcançar um array de  $4 \times 4$  de histogramas com 8 orientações. Adicionando mais orientações ou um maior descritor pode efectivamente prejudicar o reconhecimento, tornando o descritor mais sensível à distorção.

## 2.4 Método de Reconhecimento de Keypoints

O melhor candidato correspondente a cada keypoint é encontrado identificando o seu vizinho mais próximo na base de dados de keypoints no conjunto de imagem de treino. O vizinho mais próximo é definido como o keypoint com a mínima distância Euclideana para o vector descritor invariante, como descrito anteriormente.

No entanto algumas características de uma imagem não possuem nenhuma correspondência correcta no conjunto de treino porque resultam de "*barulho de fundo*" ou não foram detectadas no conjunto de treino. Portanto, seria útil uma maneira de descartar características que não possuem nenhuma boa correspondência à base de dados. Um limite global à distância da característica mais próxima, não resulta numa boa performance, já que alguns descritores são mais discriminantes que outros. Uma medida mais eficiente é obtida pela comparação da distância do vizinho mais próximo ao segundo vizinho mais próximo.

Se há múltiplas imagens do conjunto de treino do mesmo objecto, então define-se o segundo vizinho mais próximo como sendo o vizinho mais próximo de um objecto diferente do primeiro. Esta medida apresenta uma boa performance já que correspondências precisam de ter o vizinho mais próximo, significativamente mais perto que o correspondência incorrecta mais próxima, para alcançar um reconhecimento fidedigno.

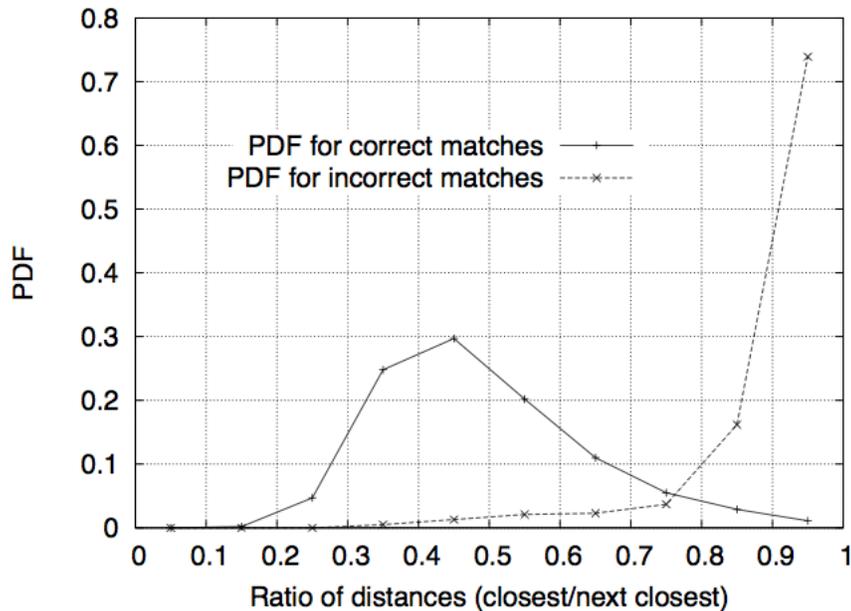
Para correspondências falsas, irão provavelmente existir outras correspondências com distâncias semelhantes devido à alta dimensionalidade das características do espaço. Pode-se considerar a segunda melhor correspondência como uma estimativa da densidade das falsas correspondências dentro desta porção de características espaciais e ao mesmo tempo identificando instancias específicas da ambiguidade das características.

A figura 2.10 mostra o valor desta medida para dados de imagens reais. A PDF para correspondências correctas e incorrectas são demonstradas em termos da razão entre o vizinho mais próximo e segundo mais próximo de cada keypoint.

Na implementação de *Lowe*, são rejeitadas todas as correspondências em que a razão das distancias é superior a 0.8, o que elimina 90% das correspondências falsas, enquanto rejeita apenas 5% das correctas.

### **Indexação Eficiente do Vizinho mais próximo**

Não são conhecidos algoritmos que consigam identificar um vizinho mais próximo num espaço dimensional elevado que sejam mais eficientes que a procura exhaustiva.



**Figura 2.10:** A probabilidade que uma correspondência seja correcta pode ser determinada através da razão entre a distancia do vizinho mais próximo, com a distancia do segundo vizinho mais próximo. Usando uma Base de Dados de 40.000 keypoints, a linha a cheio mostra que a PDF desta razão para correspondências correctas, enquanto que a linha tracejado mostra as correspondências incorrectas. [29]

O descritor de keypoint tem um vector de características de dimensão 128, e os melhores algoritmos como k-d tree (*Friedman (1997)*) [18] não fornecem uma maior rapidez em relação à procura exaustiva para mais que 10 espaços dimensionais. Portanto, foi usada uma aproximação do algoritmo, denominada *Best-Bin-First (BBF)* *Beis e Lowe (1997)* [8]. É uma aproximação no sentido que retorna o vizinho mais próximo com alta probabilidade. O BBF algoritmo usa uma procura modificada ordenada do algoritmo k-d tree, em que "porções" (*bins*) do espaço de características são procuradas por ordem da distância mais próxima do local de procura. Esta procura por ordem de prioridade foi primeiro examinada por *Arya e Mount (1993)* [2] e *Arya et al. (1998)* [3].

Uma razão para qual o algoritmo BBF funciona particularmente bem neste problema é, que apenas são consideradas correspondências em que o vizinho mais próximo é menor que 0.8 vezes a distância ao segundo vizinho mais próximo (como descrito na secção anterior), e logo não existe a necessidade de resolver os casos mais difíceis, em que muitos dos vizinhos estão a distâncias similares.

### *Clustering* com a Transformada de Hough

Uma imagem usual, típica, contém 2,000 ou mais características que podem surgir de objectos diferentes, como de ruído de fundo. Enquanto o teste da razão aplicado às distâncias, descrito na secção anterior, permitirá descartar muitas das falsas correspondências que surgem do ruído de fundo, isto não remove correspondências de outros objectos válidos, e ainda existe a necessidade de identificar subconjuntos de correspondências correctas de menos de 1% de "*inliers*" entre 99% de "*outliers*" ( "*inliers*", ou seja, os dados cuja distribuição pode ser explicada por um conjunto de parâmetros do modelo, e "*outliers*", que são dados que não se encaixam no modelo).

Muitos dos mais conhecidos métodos de ajuste robusto, como *RANdom SAmple Consensus* (RANSAC) ou *Least Mean Squares* (LMS), apresentam um desempenho insatisfatório quando a percentagem de "*inliers*" cai muito abaixo de 50%. Felizmente, um desempenho muito melhor pode ser obtido por "*clustering*" de características no "*espaço-pose*" utilizando a transformada Hough (*Hough (1962) [20], Ballard (1981),[5]*).

A transformada de Hough identifica "*clusters*" de características com uma interpretação coerente, usando cada característica no voto de todas "*poses*" de objectos que são consistentes com a característica.

Quando "*clusters*" de características são encontradas para a mesma pose de um objeto, a probabilidade de que seja uma interpretação correcta é muito maior do que para qualquer característica única.

Cada um dos keypoints especifica quatro parâmetros: localização (2D), escala e orientação, e cada keypoint correspondente no banco de dados tem um registro dos parâmetros do keypoint em relação à imagem em que foi encontrado.

Na maioria das implementações da transformada de Hough, uma matriz multi-dimensional é usada para representar *bins*. No entanto, muitas das potenciais *bins* ficarão vazias, sendo difícil de calcular o intervalo de valores de possíveis *bins* devido à sua dependência mútua (por exemplo, a dependência da discretização da localização na escala seleccionada).

Estes problemas podem ser evitados usando uma função hash pseudo-aleatória dos valores de *bins* para inserir numa tabela hash uni-dimensional, em que colisões são facilmente detectadas.

## Capítulo 3

# Implementação

A API usada esta disponível on-line, e é suportada pela *University of California Los Angeles (UCLA) Vision Lab* e pela *Oxford Visual Geometry Group (VGG)*.

Foi implementada por *A. Vedaldi e B. Fulkerson* [38].

Consiste numa biblioteca *OpenSource* de algoritmos de *Visão Computacional* em *Matrix Laboratory (MatLab)*, entre os quais se encontra o SIFT.

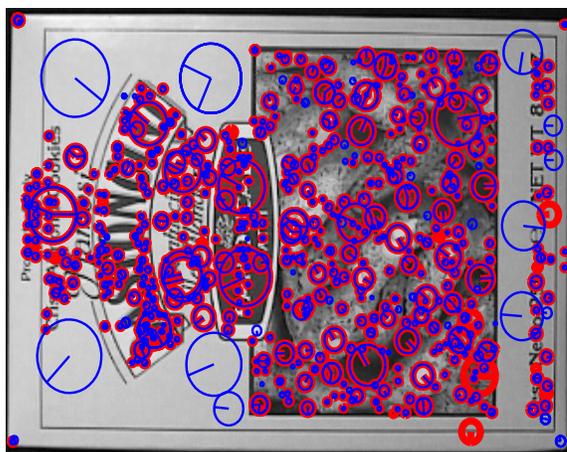
Esta API foi considerada em detrimento da implementação de *Lowe* por apresentar uma boa documentação, e resultados similares à implementação de *Lowe*, sendo que é mais facilmente parametrizável.

### 3.1 Comparação da API utilizada e o Método de Lowe

Os exemplos aqui apresentadas estão presentes na documentação da API [38].

A biblioteca de implementação do SIFT é amplamente compatível com a implementação de Lowe (porém, os keypoints são armazenados num formato ligeiramente diferente, consultar [38] ).

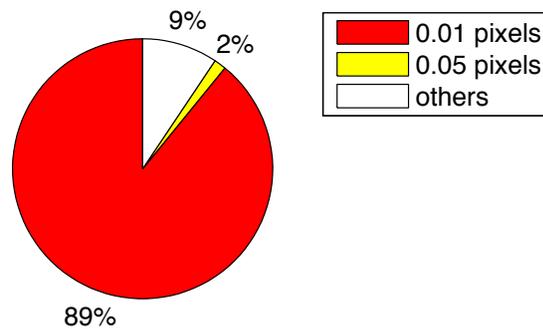
A figura a seguir compara keypoints calculados pela função SIFT da API (azul) e a *Implementação de Lowe* (vermelho) 3.1.



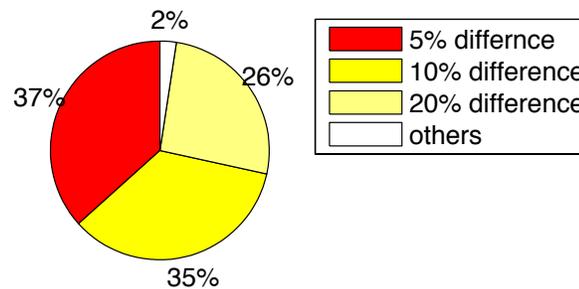
**Figura 3.1:** keypoints API (azul) sobrepostos a keypoints Lowe (vermelho). A maioria dos keypoints coincide exactamente [38]

A grande maioria dos keypoints coincide quase exactamente. A figura a seguir mostra o percentagem de keypoints calculados pelas duas implementações, cujo centro coincide com uma precisão de pelo menos 0,01 pixels e 0,05 pixels respectivamente 3.2.

Os descritores também apresentam bastantes semelhanças. A figura a seguir mostra o percentagem de descritores calculados pelas duas implementações cuja distância seja inferior a 5 %, 10% e 20% da distância média do descritor 3.3.



**Figura 3.2:** Percentagem de keypoints obtidos a partir das duas implementações que coincidem com uma precisão de 0,01 e 0,05 pixels [38]



**Figura 3.3:** Percentagem de descritores obtidos a partir das duas implementações cuja distância é inferior a 10% e 20% da distância média do descritor. [38]

## 3.2 Parâmetros Utilizados na Extração de Características

Nesta secção estão definidos os parâmetros de extração utilizados pelo SIFT.

Foram executadas as "alterações" ("override") necessárias à API afim duma melhor avaliação das características extraídas, de modo a concluir quais dessas características teriam uma maior influência sobre o reconhecimento.

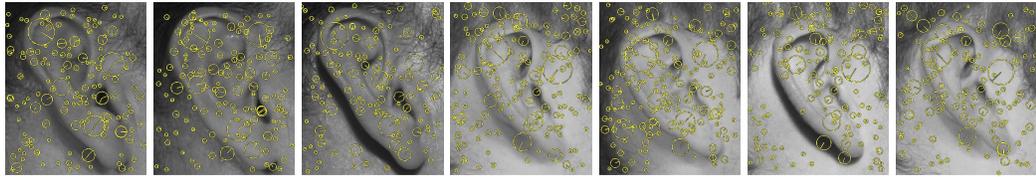
As configurações aqui apresentadas são apenas algumas das configurações possíveis, mesmo algumas destas configurações não fornecem as melhores características como iremos concluir.

Parâmetros de Extração					
	<i>Peakthreshold</i>	<i>Edgethreshold</i>	<i>Orientação</i>	<i>Escala</i>	<i>Máscara</i>
Configuração 1	0	10	•	•	•
Configuração 2	0	10	•	•	✓
Configuração 3	0.01	10	•	•	✓
Configuração 4	0.01	10	•	$\geq 2$	•
Configuração 5	0.01	10	•	•	•
Configuração 6	0.01	10	•	$\geq 2$	✓
Configuração 7	0.02	10	•	•	✓
Configuração 8	0.02	10	•	$\geq 2$	✓
Configuração 9	0.02	10	•	•	•
Configuração 10	0	10	$-\frac{\pi}{2} < \theta < \frac{3\pi}{2}$	•	•
Configuração 11	0.02	10	•	$\geq 2$	•
Configuração 12	0	10	$-\frac{\pi}{2} < \theta < \frac{3\pi}{2}$	•	✓
Configuração 13	0	10	$-\frac{3\pi}{2} < \theta < \frac{\pi}{2}$	•	✓
Configuração 14	0	10	$-\frac{3\pi}{2} < \theta < \frac{\pi}{2}$	•	•

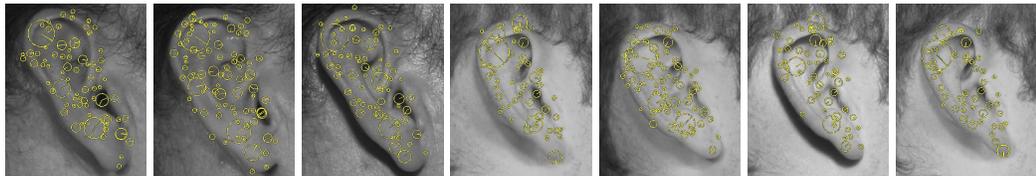
Tabela 3.1: Configurações de extração de características testadas

Algumas destas configurações como poderemos constatar adiante na avaliação de resultados serão *impossíveis*, como por exemplo um n° de keypoints elevado (*cardinalidade grande*) com um *peak threshold* igualmente elevado. <sup>1</sup>

<sup>1</sup> O valor do *edgethreshold* não sofreu alterações nas várias configurações, já que a sua alteração em valores significativos não demonstrou uma relevância a considerar.



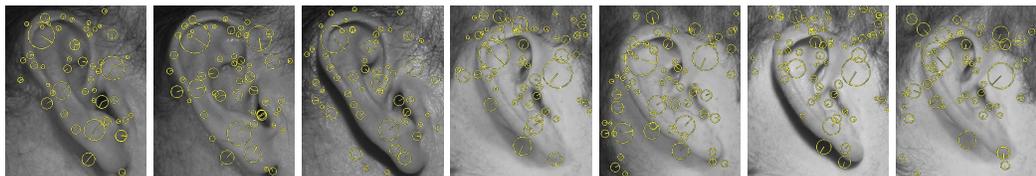
(a) Configuração 1 - keypoints com peakthreshold = 0 e edgethreshold = 10



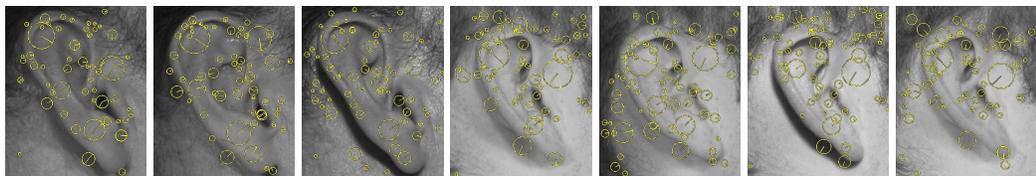
(b) Configuração 2 - keypoints com máscara, peakthreshold = 0 e edgethreshold = 10



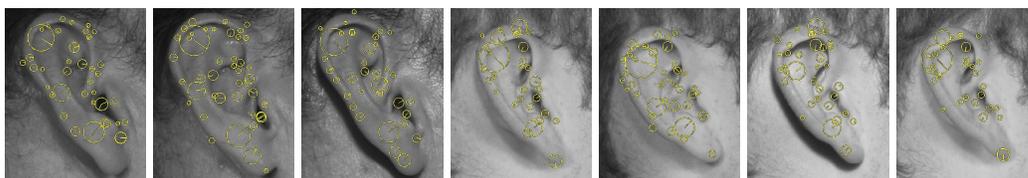
(c) Configuração 3 - keypoints com máscara, peakthreshold = 0.001



(d) Configuração 4 - keypoints peakthreshold = 0.001 e escala > 2



(e) Configuração 5 - keypoints peakthreshold = 0.001



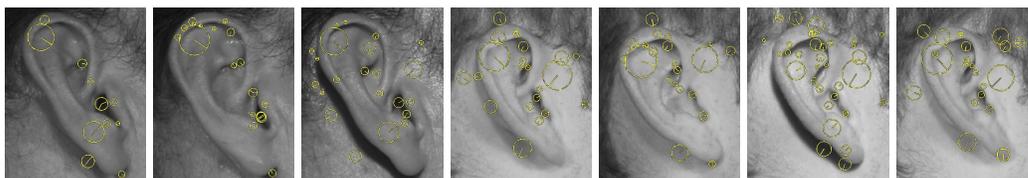
(f) Configuração 6 - keypoints com máscara, peakthreshold = 0.001 e escala > 2



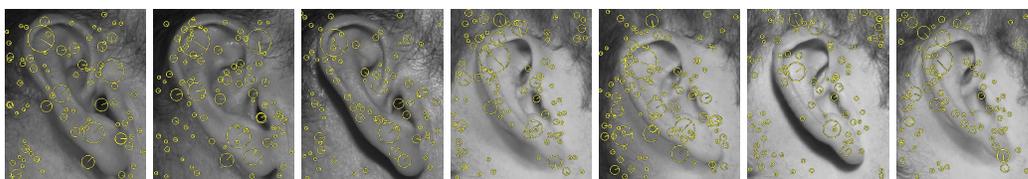
(g) Configuração 7 - keypoints com máscara, peakthreshold = 0.002



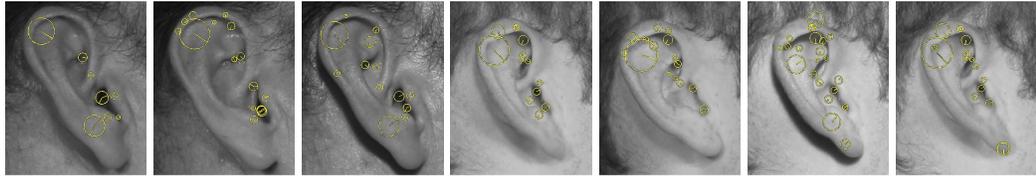
(h) Configuração 8 - keypoints com máscara, peakthreshold = 0.002 e escala > 2



(i) Configuração 9 - keypoints peakthreshold = 0.002



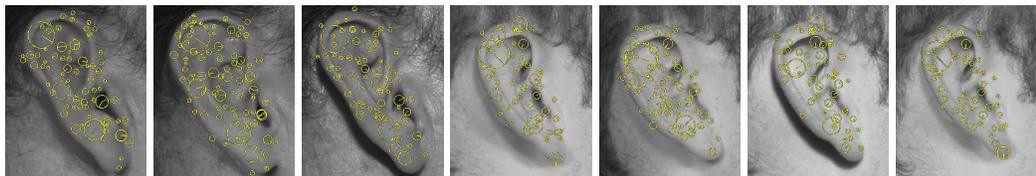
(j) Configuração 10 - keypoints com rotação entre  $-\pi/2 < \theta < 3\pi/2$



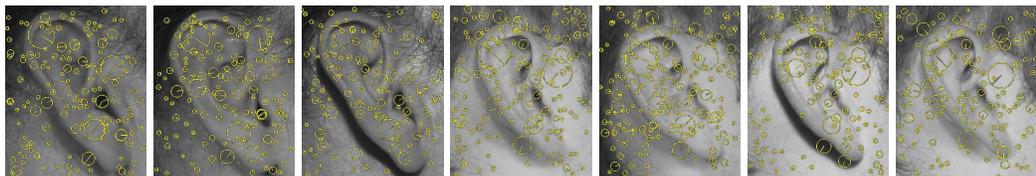
(k) Configuração 11 - keypoints com máscara, peakthreshold = 0.002 e escala > 2



(l) Configuração 12 - keypoints com máscara, keypoints com rotação entre  $-\pi/2 < \theta < 3\pi/2$



(m) Configuração 13 - keypoints com máscara, keypoints com rotação entre  $-3\pi/2 < \theta < \pi/2$



(n) Configuração 14 - keypoints com rotação entre  $-3\pi/2 < \theta < \pi/2$

**Figura 3.4:** Exemplos de keypoints extraídos nas experiências, do Banco de Imagem A.

### 3.3 Algoritmos Implementados

Em seguida encontram-se ilustrados os algoritmos implementados, para extracção de localizações  $(x, y, escala, \theta)$  e descritores  $[keypoints \times 128]$ , bem como o algoritmo de reconhecimento implementado, este algoritmo foi *inspirado* num modelo apresentado por Mark S. Nixon, David J. Hurley e Banafshe Arbab-Zavar (2007) [1].

O método aplicado por Nixon S.M., et al [1], não utiliza o classificador de Lowe, utilizando um critério de discriminação algo diferente, resumindo: também se baseia numa medida de distância, em que se o resultado dessa medida, normalizada, para um determinado keypoint for maior que 3 vezes o desvio padrão no seu conjunto de treino é descartado.

O modelo proposto por Nixon S.M., et al apresenta uma cardinalidade de 20, já que utiliza um número constante de keypoints no modelo para classificar.

Em relação ao critério de reconhecimento, é utilizada a média das distâncias entre descritores correspondentes.

---

**Algoritmo 1** Algoritmo responsável pela extracção de características com os parâmetros definidos.

---

**Require:** index

$\forall Imagens \leftarrow$  Imagem A ... Imagem Z

**for** index = 1 to total( *Imagens* ) **do**

    Extrair localizações e descritores da imagem, com os parâmetros definidos.

    [frames,descritores]  $\leftarrow$  SIFT(Imagem, edgethreshold, peakthreshold, mask, nkeypoints, escala,  $\theta_a, \theta_b$  )

**end for**

---

---

**Algoritmo 2** Algoritmo responsável pelo cálculo das distâncias entre descritores de duas imagens.

---

**Require:**  $index_i = 1, index_j$

$\forall$  [descritores] de um dado conjunto (Imagem A ... Imagem Z)

```
for Imagem( $index_i$ ) to total(Imagens) do
   $index_j = index_i + 1$ 
  for Imagem( $index_j$ ) to total(Imagens) do
    Computar distância Euclidiana entre os descritores, rejeitando key-
    points através do método de Lowe

    [matches, scores]  $\leftarrow$  MATCH(Imagem( $index_i$ ).descritores,
    imagem( $index_j$ ).descritores)
    if  $ImagemA_{assinatura} \in ImagemB_{assinatura}$  then
      Imagem do mesmo indivíduo, atribuir resultado a intra-classe
       $intra_{score} \leftarrow scores$ 
       $intra_{match} \leftarrow matches$ 

    else
      Imagem de indivíduos diferentes, atribuir resultado a inter-classe
       $inter_{score} \leftarrow scores$ 
       $inter_{score} \leftarrow matches$ 

    end if
  end for
end for
```

---

---

**Algoritmo 3** Algoritmo responsável pela avaliação do critério aplicado ao reconhecimento.

---

Para todo o conjunto da intra-classe e da inter-classe.

Definir média mínima , média máxima

Definir o número de keypoints (matches), cardinalidade = número de keypoints da função de reconhecimento

**Require:** cardinalidade, index, média-max, média-min

**for** index = 1 to total(intra-classe) **do**

**if** scores =  $\emptyset$  **then**

$intra_{vazios} \leftarrow index$

**else if** tamanho(scores-intra)  $\geq$  cardinalidade **then**

        Ordenar scores (distâncias) crescente.

        [scores]  $\leftarrow$  ordenar(scores)

        matcher(index)  $\leftarrow$  media(scores(1 ... cardinalidade))

**if** média(scores(1 ... cardinalidade))  $>$  média-max **then**

            média-max  $\leftarrow$  média

**end if**

**if** média(scores(1 ... cardinalidade))  $<$  média-min **then**

            média-min  $\leftarrow$  média

**end if**

**else**

        guardar indexes dos que não cumprem o critério da cardinalidade

$intra_{cardinalidade} \leftarrow index$

**if** média(scores)  $>$  media máxima **then**

            média máxima  $\leftarrow$  média

**end if**

**if** média(scores)  $<$  média mínima **then**

            média mínima  $\leftarrow$  média

**end if**

**end if**

**end for**

---

---

```
for index = 1 to total(inter-classe) do
  if scores =  $\emptyset$  then
     $inter_{vazios} \leftarrow index$ 
  else if tamanho(scores inter)  $\geq$  cardinalidade then
    Ordenar scores (distâncias) crescente.

    [scores]  $\leftarrow$  ordenar(scores)
    matcher(index)  $\leftarrow$  media(scores(1 ... cardinalidade))
    if média(scores(1 ... cardinalidade))  $>$  média-max then
      média-max  $\leftarrow$  média
    end if
    if média(scores(1 ... cardinalidade))  $<$  média-min then
      média-min  $\leftarrow$  média
    end if
  else
    guardar indexes que não cumprem o critério da cardinalidade
     $inter_{cardinalidade} \leftarrow index$ 
    if média(scores)  $>$  média-max then
      média-max  $\leftarrow$  média
    end if
    if média(scores)  $<$  média-min then
      média-max  $\leftarrow$  média
    end if
  end if
end for
if  $intra_{vazios} \neq \emptyset$  then
   $matcher(intra_{vazios}) \leftarrow$  média máxima
end if
if  $inter_{vazios} \neq \emptyset$  then
   $matcher(inter_{vazios}) \leftarrow$  média máxima
end if
if  $intra_{cardinalidade} \neq \emptyset$  then
   $matcher(intra_{cardinalidade}) \leftarrow$  média máxima
end if
if  $inter_{cardinalidade} \neq \emptyset$  then
   $matcher(inter_{cardinalidade}) \leftarrow$  média máxima
end if
```

---



# Capítulo 4

## Resultados

### 4.1 Conjunto de Características

Para o testar o método proposto por *Lowe e Brown* [29], foram extraídos os conjuntos de características, com os vários parâmetros (*14 configurações*), referidos anteriormente 3.1.

As *configurações* que revelam interesse no campo da *identificação biométrica*, são as contendo *máscara*, já que não é do interesse uma correspondência de características provenientes do *fundo da cena*. A *configuração 1* sendo que é a configuração padrão, sem qualquer alteração dos parâmetros SIFT, apresenta um maior número de características extraídas, e como iremos concluir adiante este facto influencia, em muito, os resultados.

O valor como *Extractor / Detector* de características, que apresenta propriedades de invariância a algumas dessas características, poderá então ser avaliado, sendo que é justificável considerar uma configuração padrão sem qualquer alteração dos parâmetros de extracção.<sup>1</sup>

---

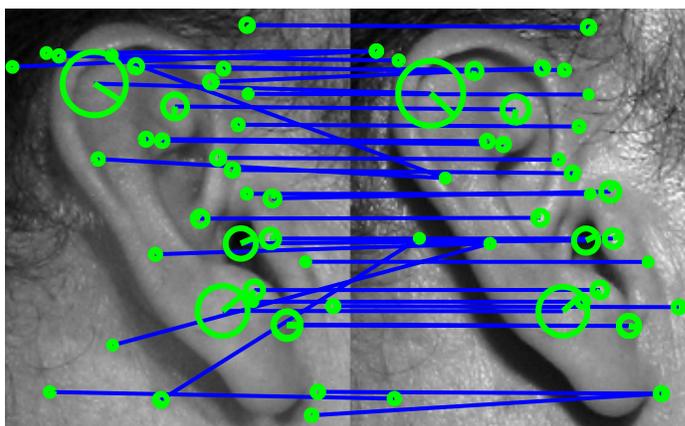
<sup>1</sup> As *configurações* contendo máscara são: 2,3,6,7,8,12,13. A *configuração padrão* é a *configuração 1*. 3.1.

As características SIFT possuem um largo espectro de interesse desde: reconhecimento de objectos em cena com alta percentagem de oclusão, a *image stitching* alinhando imagens de diferentes panoramas, etc. [15]

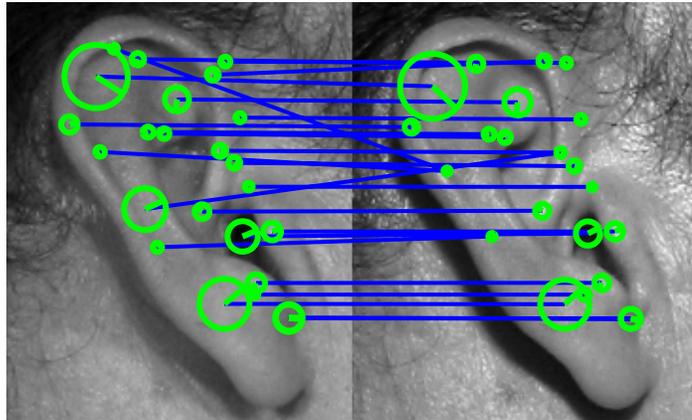
Sendo que um sistema biométrico baseado em características SIFT, é uma das vertentes da pesquisa no campo da biometria, este projecto representa mais uma avaliação destas características bem como a metodologia aplicada na sua obtenção.

Estes parâmetros foram escolhidos, entre outras dezenas, ou centenas, que também poderiam ser parametrizadas para que tivessem a mais variada complexidade e que apresentassem diferentes focos de características escolhidas, só com as mais variadas condições é possível verificar a eficácia de um método.

Nas seguintes figuras é ilustrado o resultado de comparações encontradas entre imagens diferentes do mesmo indivíduo (*Intra-Classe*). 4.1 4.2 4.3 4.4 4.5 4.7 4.8 4.9 4.10



**Figura 4.1:** Configuração 1 (padrão) - comparação *Intra-Classe*, analisando a imagem podemos concluir que existe um elevado nº de correspondências, sendo que uma grande parte dessas provém de zonas da imagem não contendo orelha. Facilitando a detecção de um objecto numa dada cena, dado que a cena não sofra alterações existe alta probabilidade de reconhecimento do objecto. [29]

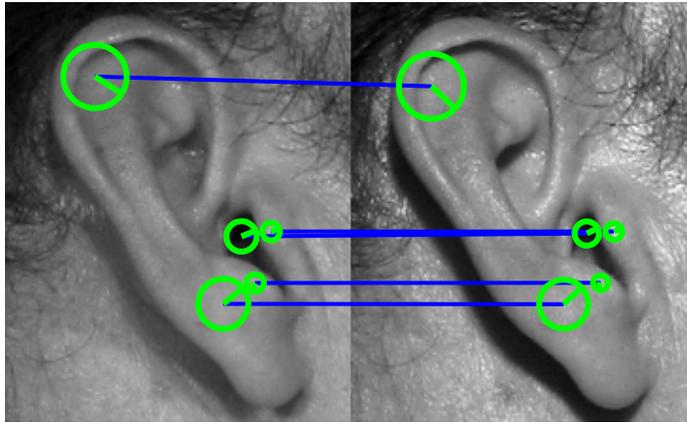


**Figura 4.2:** Comparação Intra-Classe, com máscara, apresentando sombra na imagem da direita, apesar de alguma oclusão não é o pior caso, continuando a existir ainda um elevado nº de correspondências.

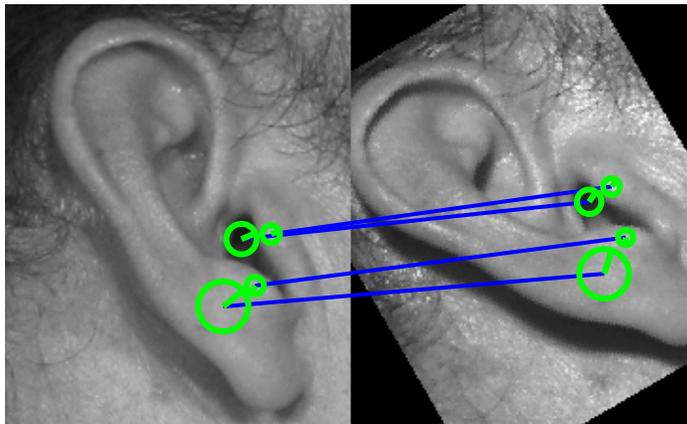
As ilustrações servem para uma melhor visualização de alguns dos parâmetros configurados, e para sobre estes obter uma melhor perspectiva, as experiências aqui demonstradas revelam algumas das dificuldades na obtenção e interpretação dos resultados.

Os casos que se seguem tentam ilustrar a declarada invariância de algumas das características SIFT, sendo que a invariância à escala é demonstrada pelo processo de *segmentação* da imagem da orelha (*crop*), como enunciado anteriormente 2.1, já que todas as imagens possuem o mesmo tamanho, sendo normalizadas a partir do centro da sua máscara, podendo por isso apresentar pequenas variações de escala.

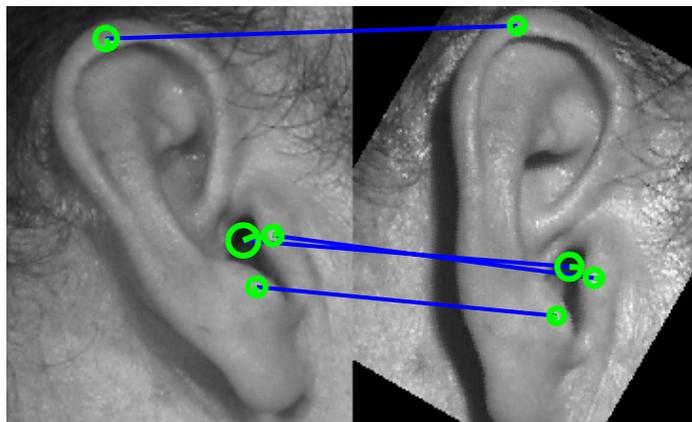
Relativamente a alguma invariância à rotação esta também foi testada ilustrando uma rotação artificialmente adicionada à imagem. Como iremos concluir rotação em alguns casos poderá ser sinónimo de oclusão.



**Figura 4.3:** A mesma comparação, com  $peakthreshold = 0.02$ , este depende da luminosidade, iluminação da imagem tem um maior grau de sensibilidade, nesta imagem estão ilustradas 5 correspondências.



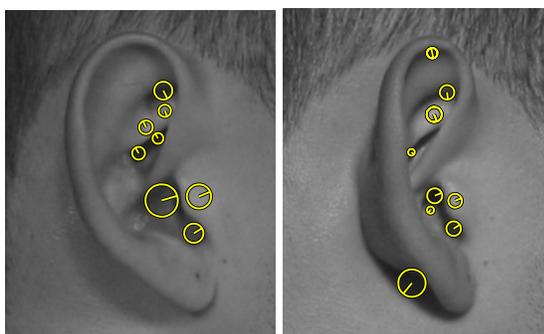
**Figura 4.4:** A mesma comparação, com  $peakthreshold = 0.02$  e rotação de  $30^\circ$  artificialmente adicionada, com um alto  $peakthreshold$ , 4 correspondências ainda são visíveis.



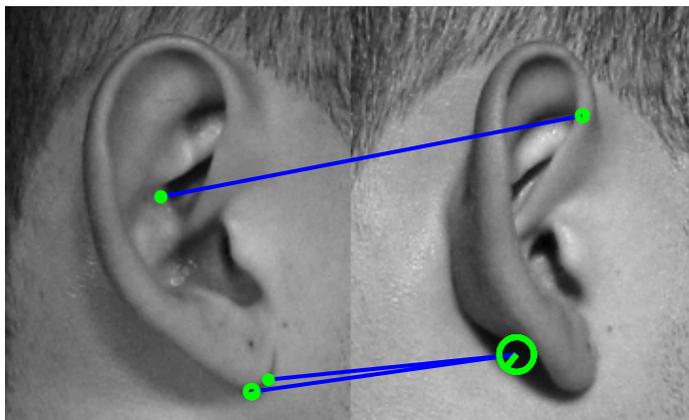
**Figura 4.5:** A mesma comparação, com  $peakthreshold = 0.02$  e rotação de  $330^\circ$  artificialmente adicionada, com um alto  $peakthreshold$ , 4 correspondências, algumas diferentes da anterior rotação, ainda são visíveis.

As figuras que se seguem referem um exemplo de comparação de imagens *Intra-Classe* de difícil reconhecimento já que a elevada rotação demonstrada pela orelha, ocluí partes da mesma.

Também é demonstrado na seguinte figura 4.6 que as orelhas descritas possuem características detectadas/extraídas por SIFT.



**Figura 4.6:** Keypoints extraídos com um  $peakthreshold = 0.02$ .

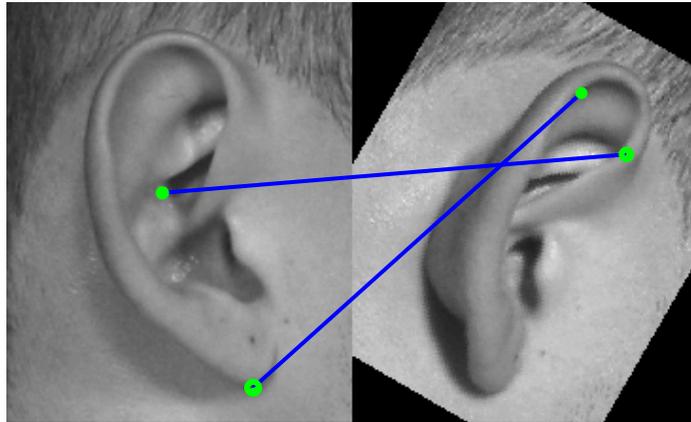


**Figura 4.7:** Comparação intra-classe, com  $peakthreshold = 0.01$ , muito poucas correspondências visíveis, tendo em conta que é uma comparação intra-classe, podemos verificar que o reconhecimento desta orelha seja considerado muito difícil.

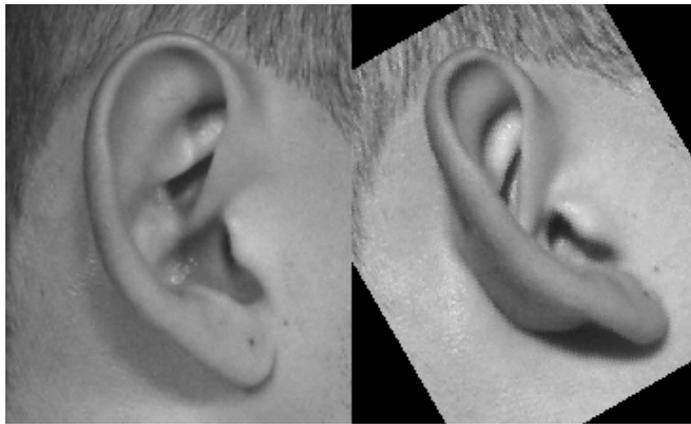
Na mesma comparação elevando o  $peakthreshold$ , verificamos a perda de capacidade de reconhecimento de características SIFT.



**Figura 4.8:** Comparação intra-classe, com  $peakthreshold = 0.02$ , como verificado anteriormente apesar das imagens possuírem keypoints, todos eles são descartados no critério de reconhecimento, fazendo desta comparação, que apesar de ser do mesmo indivíduo, impossível.

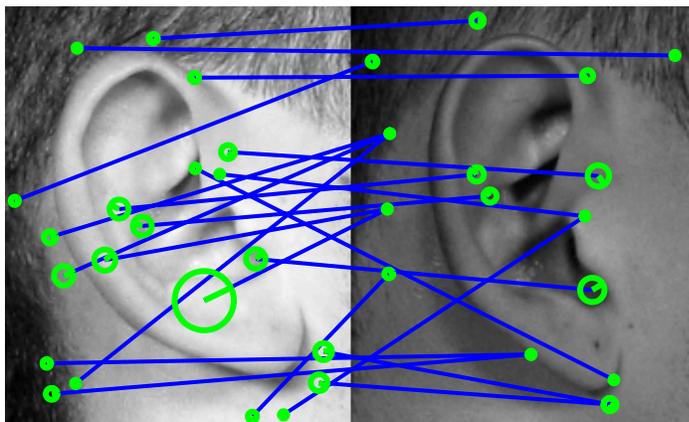


**Figura 4.9:** Comparação intra-classe, rotação de  $30^\circ$ , duas correspondências visíveis, sendo que deverão apresentar um resultado elevado para a distância entre eles, ou seja um fraco resultado, poucas correspondências resultado elevado.

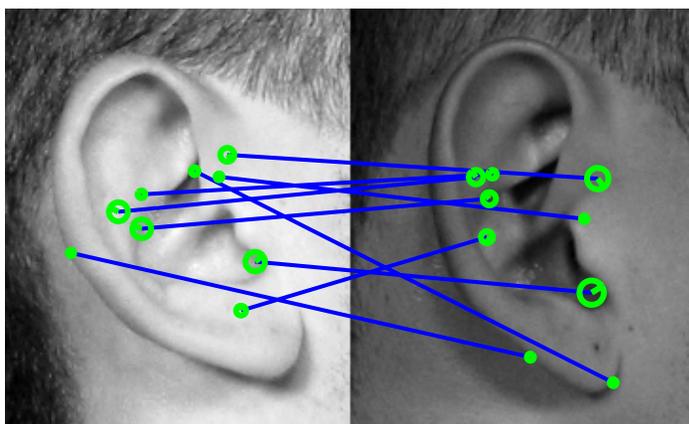


**Figura 4.10:** Comparação intra-classe, rotação de  $330^\circ$ , nenhuma correspondência visível, reconhecimento impossível.

Nas seguintes figuras é ilustrado o resultado de comparações encontradas entre imagens diferentes de indivíduos distintos (*Inter-Classe*) 4.11 4.12.



**Figura 4.11:** Configuração 1 (padrão) - comparação *Inter-Classe* , correspondências com o fundo visíveis.



**Figura 4.12:** Configuração 2 (com máscara) - comparação *Inter-Classe*, correspondências apenas na área da máscara.

Das seguintes ilustrações podemos referenciar que apesar de características SIFT poderem ser extraídas de imagens estas poderão não ser suficientemente consistentes para fornecer um reconhecimento satisfatório, mesmo sendo de imagens do mesmo indivíduo.

Algumas das propriedades referidas anteriormente, como invariância à rotação também foram demonstradas, referenciando o número de correspondências possíveis mesmo com um elevado grau de rotação e modificando o parâmetro de *peakthreshold*. 4.4 4.5

## 4.2 Discussão de Resultados

Segundo *Lowe* [29], a rotação num plano é suportada até  $30^\circ$ , sem perda significativa da repetibilidade dos resultados. Basicamente *Lowe* refere-se ao *enviesamento* ou *transformação afim*, sendo de enfatizar que estas transformações, serão efectuadas segundo um plano, não existindo oclusão de informação.

Sendo a orelha uma superfície não-plana, e uma imagem caracterizada por um referencial 2D, alguma informação é perdida nesta transição para um universo 2D, e apesar do SIFT demonstrar alguma invariância a rotação, e até à oclusão, o facto da rotação ocultar uma parte da orelha diminuí a capacidade de extracção de características. 4.8 4.10

A superfície não-plana da orelha é uma limitação presente no SIFT, na implementação de um sistema biométrico.

O classificador proposto por *Lowe*, baseia-se numa medida de distância entre os descritores, as experiências efectuadas, além da parametrização do SIFT na extracção de características, também focaram a cardinalidade das correspondências entre descritores, efectuando o reconhecimento com um número variado de correspondências (5, 10, 20).

Um maior número de correspondências com o resultado baixo equivale ao melhor reconhecimento, esta avaliação depende dos parâmetros de extracção, logo uma variação dos parâmetros, é equiparada a uma variação das correspondências.

Como referenciado anteriormente os parâmetros de extracção estão correlacionadas com a capacidade de reconhecimento, mas concretamente com a cardinalidade das correspondências e o resultado atribuído às suas distâncias.

Nas seguintes tabelas de resultados podemos ver o valores obtidos nas várias configurações para a referida cardinalidade do conjunto de *keypoints* coincidente nos vários Bancos de Imagem A,B,C.

Todos os dados de todas as configurações podem ser consultados em anexo.

São utilizadas medidas de avaliação de sistemas como *Equal Error Rate* (EER), onde a taxa de *False Negative Rate* (FNR) e *False Positive Rate* (FPR) é igual, curvas *Receiver Operating Characteristic* (ROC) que ilustram a relação entre a FPR e a FNR. Sendo a medida de decidibilidade definida tal que:  $Decidibilidade = \frac{|\mu_{inter} - \mu_{intra}|}{\sqrt{1/2(\sigma_{inter}^2 + \sigma_{intra}^2)}}$

Pela avaliação dos resultados podemos ver que o conjunto de características inicial influencia, em muito o resultado esperado, sendo a cardinalidade do conjunto de características um dos factores mais relevante, senão o mais relevante. 4.13 4.18 4.21

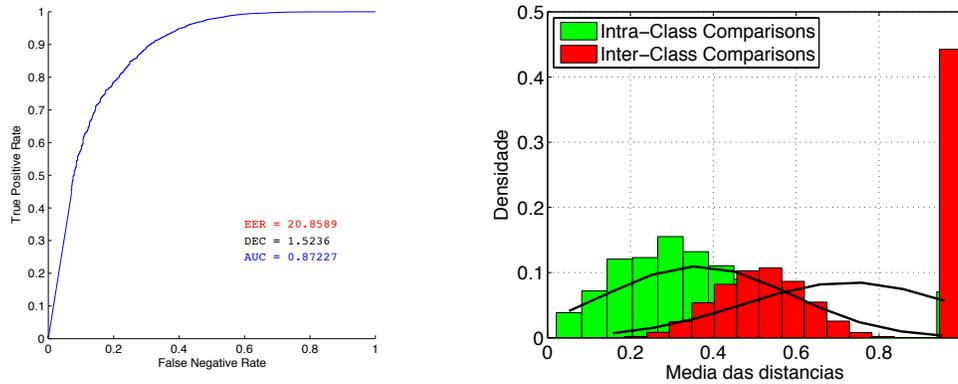
Este Banco de dados é constituído por todo o conjunto de imagens disponível, com rotação e sem rotação, com 3, 4, e 6 imagens por indivíduo, sendo que é o banco mais heterogéneo.

Podemos ainda verificar configurações impossíveis (marcadas a vermelho na tabela), como resultado de parâmetros de extracção inviáveis para a cardinalidade do conjunto requerida.

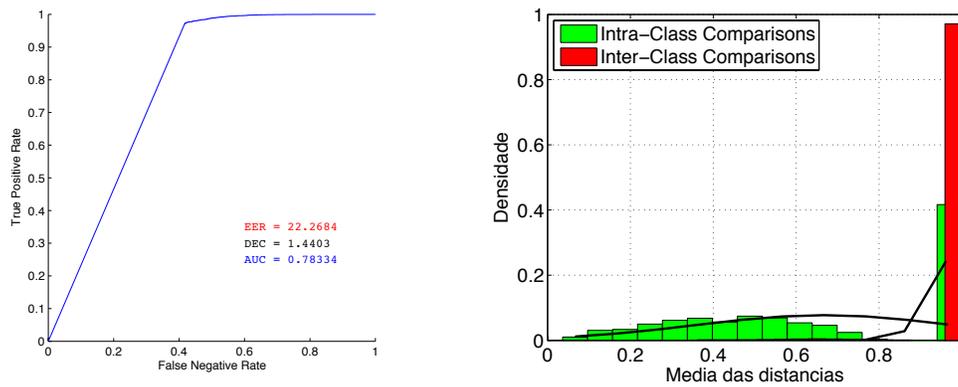
Banco de Imagens A									
	20 keypoints			10 keypoints			5 keypoints		
	EER	DEC	AUC	EER	DEC	AUC	EER	DEC	AUC
Configuração 1	22,26840	1,44030	0,78334	20,86530	1,52360	0,87227	23,06480	1,27190	0,85706
Configuração 2	38,40810	0,76591	0,61607	27,21580	1,18550	0,76711	26,95150	1,02400	0,81122
Configuração 3	45,90090	0,41970	0,54099	32,04280	0,96818	0,69098	29,11010	0,95685	0,77866
Configuração 4	39,32130	0,72516	0,60696	25,00330	1,34870	0,79764	23,45690	1,16080	0,83921
Configuração 5	37,82570	0,78893	0,62204	25,80110	1,32290	0,80161	23,31750	1,14240	0,84058
Configuração 6	46,66960	0,37551	0,53331	32,65130	0,95089	0,68206	29,10320	0,99445	0,77945
Configuração 7	45,31560	0,43859	0,54706	32,07620	0,89551	0,71520	32,07620	0,89551	0,71520
Configuração 8	----	----	----	45,58820	0,42513	0,54419	32,23120	0,89580	0,71193
Configuração 9	49,61780	0,12333	0,50382	37,00020	0,78776	0,63375	28,08500	1,13720	0,78507
Configuração 10	37,70600	0,78375	0,62310	23,89310	1,46620	0,81554	24,19440	1,05640	0,84916
Configuração 11	49,61740	0,12360	0,50383	37,44150	0,76505	0,62856	27,56380	1,14430	0,78386
Configuração 12	46,99110	0,35644	0,53009	32,60970	0,95646	0,68205	29,48350	0,93345	0,76891
Configuração 13	38,40810	0,76591	0,61607	27,21580	1,18550	0,76110	26,95150	1,02400	0,81122
Configuração 14	22,26840	1,50710	0,78334	20,86380	1,41270	0,87227	23,05980	1,19050	0,85706

Figura 4.13: Tabela de Resultados do Banco de Imagens A, configurações impossíveis marcadas a vermelho, melhores resultados marcados a verde

As configurações que não possuem máscara como seria de esperar apresentam resultados mais satisfatórios, sendo justificável por um maior número de características iniciais extraídas. 4.14 4.15

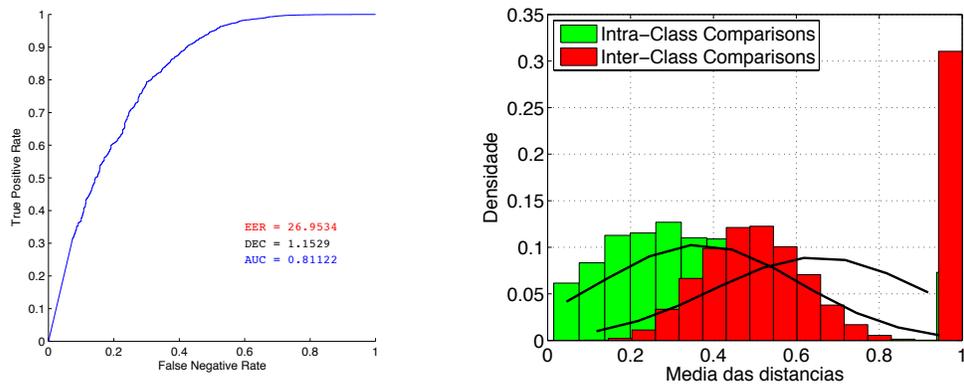


**Figura 4.14:** Avaliação do Classificador na Configuração 1 para 10 keypoints (cardinalidade = 10). Sendo esta a configuração que apresenta melhores resultados no Banco de Imagens A.

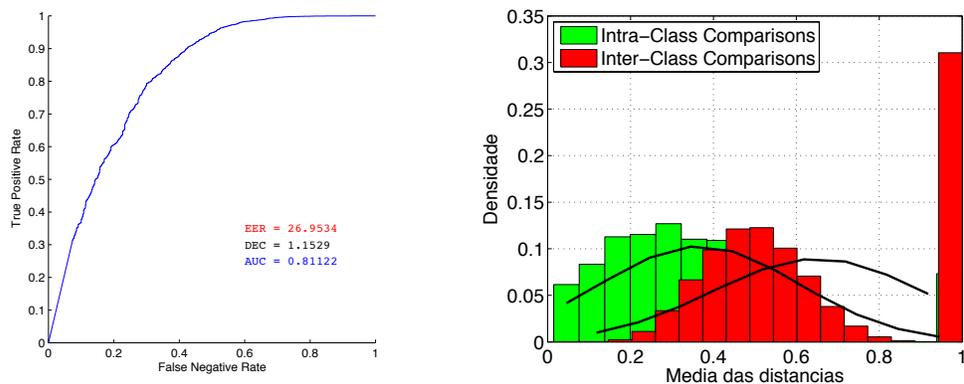


**Figura 4.15:** Avaliação do Classificador na Configuração 1 para 20 keypoints (cardinalidade = 20). Podemos verificar que com a mesma configuração e com um aumento da cardinalidade de correspondências o resultado piora

Diminuindo a cardinalidade para 5, continuamos a ter a configuração 1 como a melhor, mas os resultados entre diferentes configurações apresentam um menor desvio entre si. Havendo assim configurações com máscara que apresentam *Area Under Curve* (AUC) acima dos 80 %. 4.16 4.17



**Figura 4.16:** Avaliação do Classificador na Configuração 2, com máscara, para 5 keypoints (cardinalidade = 5). Esta configuração e a configuração 13 foram as que apresentaram melhores resultados, com máscara, para o Banco de Imagens A.



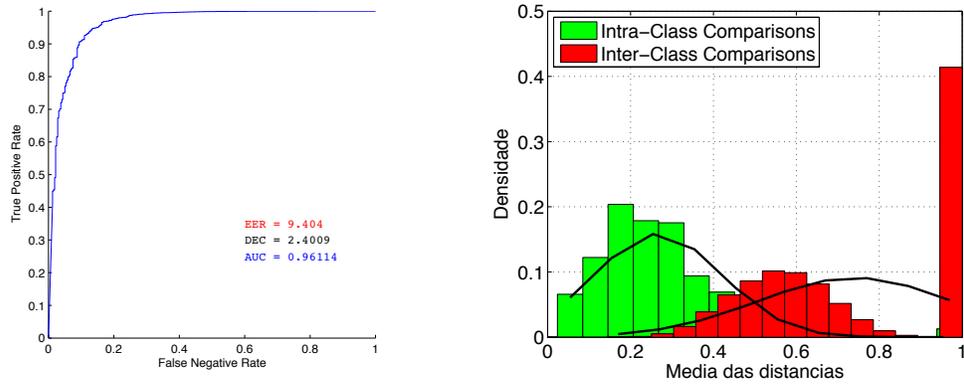
**Figura 4.17:** Avaliação do Classificador na Configuração 13, com máscara, rotação entre  $-3\pi/2 < \theta < \pi/2$  para 5 keypoints (cardinalidade = 5). Esta e configuração 2 foram as que apresentaram melhores resultados, com máscara, para o Banco de Imagens A.

Este banco de imagens caracteriza-se por 311 imagens com algum grau de rotação. Também podemos verificar um grande vazio, assinalado a vermelho na tabela, na impossibilidade de obtenção de características sob certos parâmetros.

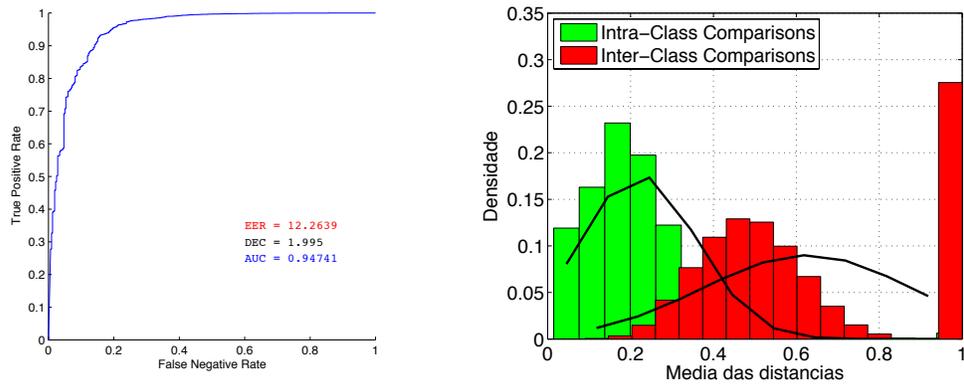
São visíveis alguns resultados bastante satisfatórios para configurações com máscara. 4.20

Banco de Imagens B									
	20 keypoints			10 keypoints			5 keypoints		
	EER	DEC	AUC	EER	DEC	AUC	EER	DEC	AUC
Configuração 1	8,14290	3,09890	0,93138	9,40400	2,02190	0,96114	12,22320	1,81980	0,94689
Configuração 2	26,89330	1,28540	0,73144	12,53,97	2,39900	0,92793	12,26180	1,76200	0,94741
Configuração 3	42,01150	0,61158	0,57989	19,41470	1,71470	0,83171	15,12280	1,57660	0,91253
Configuração 4	30,20050	1,12700	0,69839	13,47770	2,20810	0,91666	13,13820	1,64580	0,93822
Configuração 5	27,91420	1,23720	0,72157	12,85100	2,11760	0,91840	12,85510	1,60600	0,93972
Configuração 6	43,57890	0,53915	0,56422	20,29310	1,64030	0,81627	15,43510	1,65090	0,91326
Configuração 7	---	---	---	41,40410	0,62679	0,58650	25,55000	1,30420	0,80762
Configuração 8	---	---	---	41,82320	0,60929	0,58220	25,32350	1,31400	0,80725
Configuração 9	49,06160	0,19475	0,50938	30,83330	1,05370	0,69924	19,70110	1,46830	0,86981
Configuração 10	25,93720	1,34170	0,74096	9,37790	2,60920	0,95045	11,24760	1,72930	0,95601
Configuração 11	49,06160	0,19472	0,50938	31,74790	1,01770	0,68866	19,44200	1,46720	0,86503
Configuração 12	44,36150	0,50140	0,55639	17,96310	1,83710	0,84000	13,79110	1,73920	0,92679
Configuração 13	26,89330	1,28540	0,73144	12,53970	2,39900	0,92793	12,26180	1,76200	0,94741
Configuração 14	8,14290	3,09890	0,93138	9,40400	2,02190	0,96114	12,22320	1,81980	0,94689

Figura 4.18: Tabela de Resultados do Banco de Imagens B



**Figura 4.19:** Avaliação do Classificador na Configuração 1 para 10 keypoints (cardinalidade = 10). Sendo que é a melhor configuração.



**Figura 4.20:** Avaliação do Classificador na Configuração 2 para 5 keypoints (cardinalidade = 5). Uma boa classificação com máscara no Banco de Imagens B

Este banco caracteriza-se por apenas possuir imagens com pouca, ou nenhuma rotação, iremos constatar que esta propriedade, irá melhorar em muito os resultados, existindo inúmeras configurações com bons resultados.

Seja o caso de melhoria mais flagrante a configuração 6, de cardinalidade = 20 para cardinalidade = 10, onde vemos uma melhoria considerável dos resultados.

Banco de Imagens C									
	20 keypoints			10 keypoints			5 keypoints		
	EER	DEC	AUC	EER	DEC	AUC	EER	DEC	AUC
Configuração 1	3,13450	6,07210	0,99783	3,86580	2,56310	0,99546	6,66470	2,54910	0,98897
Configuração 2	8,59380	2,94770	0,91620	2,84890	3,43800	0,99892	3,04930	2,26880	0,99810
Configuração 3	33,34340	0,97900	0,66660	3,33480	4,82000	0,99696	3,18460	2,48050	0,99685
Configuração 4	18,55880	1,78140	0,81578	3,43500	3,38850	0,99576	3,92090	2,10970	0,99032
Configuração 5	16,94220	1,90990	0,83234	3,78570	3,20400	0,99484	5,92330	2,06500	0,98885
Configuração 6	35,00500	0,90595	0,64996	3,00920	5,17450	0,99770	2,79880	2,52660	0,99764
Configuração 7	---	---	---	33,77920	0,95281	0,66366	3,67550	2,68790	0,97443
Configuração 8	---	---	---	35,41580	0,88463	0,64709	3,64540	2,68190	0,97384
Configuração 9	50,00500	0,01416	0,49995	17,35450	1,88920	0,84215	6,66970	2,40120	0,98888
Configuração 10	18,53870	1,76800	0,81574	5,93330	3,14340	0,99188	6,66470	1,98820	0,98689
Configuração 11	---	---	---	18,83580	1,77250	0,82574	9,25310	2,40460	0,98512
Configuração 12	33,34840	0,90330	0,66657	6,66970	3,57470	0,96213	3,33480	2,25720	0,99535
Configuração 13	8,59380	2,94770	0,91620	2,84890	3,43800	0,99892	3,04930	2,26880	0,99810
Configuração 14	3,13450	6,07210	0,99783	3,86580	2,56310	0,99546	6,66470	2,54910	0,98897

Figura 4.21: Tabela de Resultados do Banco de Imagens C

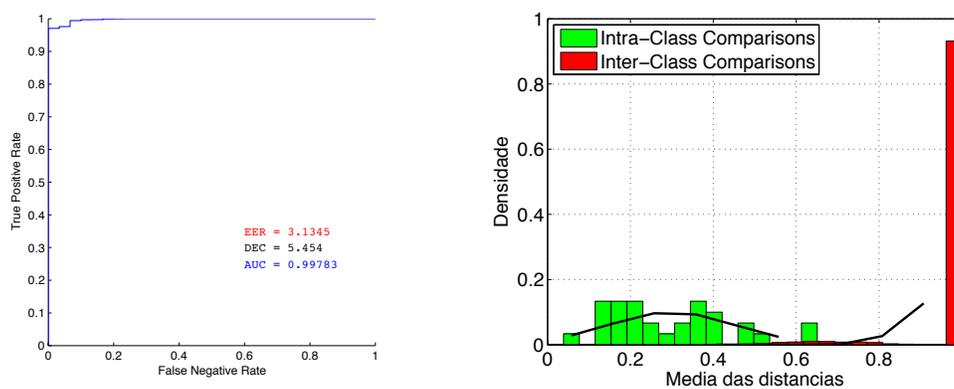
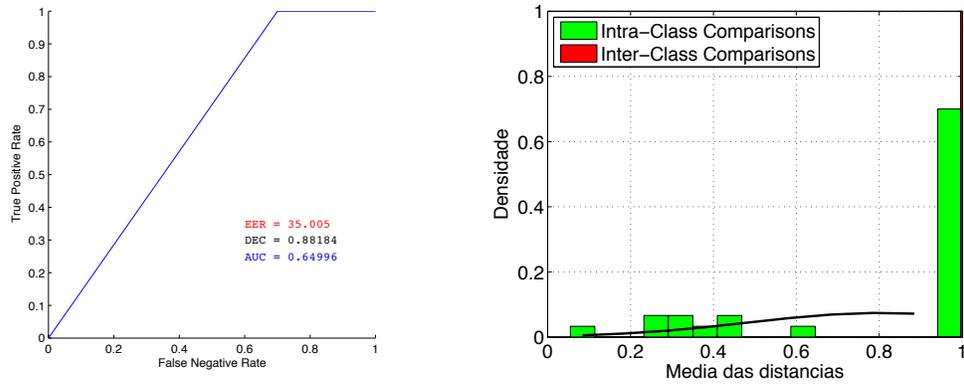
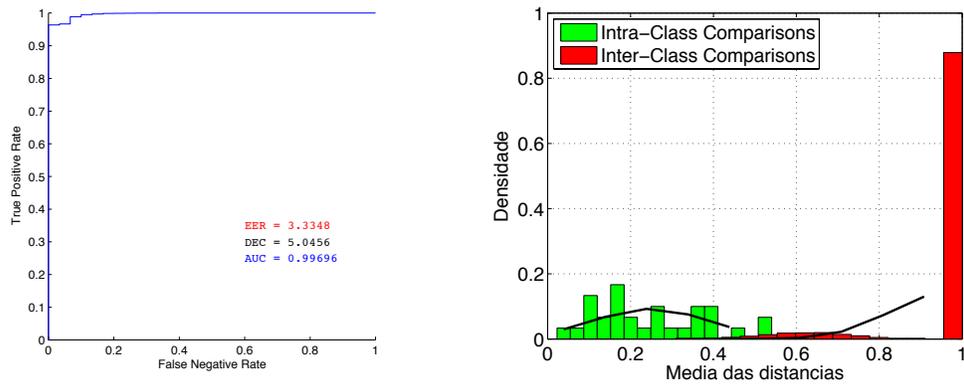


Figura 4.22: Avaliação do Classificador na Configuração 1 para 20 keypoints (cardinalidade = 20).



**Figura 4.23:** Avaliação do Classificador na Configuração 6 para 20 keypoints (cardinalidade = 20). Sendo que é o segundo pior resultado, a seguir à configuração 9.



**Figura 4.24:** Avaliação do Classificador na Configuração 6 para 10 keypoints (cardinalidade = 10). Sendo um dos melhores resultados apresentado por configurações com máscara.



# Capítulo 5

## Conclusão e Trabalho Futuro

### 5.1 Conclusão

O objectivo deste projecto, sendo um desenvolvimento dum sistema de detecção de orelhas, foi amplamente cumprido.

O valor biométrico da orelha humana voltou a relevar a sua importância, à semelhança do que foi concluído por diversos autores,[1] comparativamente por exemplo à face ou à íris.

A avaliação do SIFT como método de extracção de características foi bem sucedida, bem como a sua avaliação em inúmeros parâmetros, e o método de reconhecimento de *Lowe* [29].

Os resultados obtidos permitem concluir que é possível a elaboração de um "*modelo de reconhecimento*" através de características extraídas do SIFT. Alguns valores de medidas de avaliação de sistemas apresentam resultados satisfatórios em AUC elevadas, com baixas EER e uma boa decidibilidade.

Uma das maiores dificuldades na elaboração de um *sistema biométrico*, é na obtenção de características "fiáveis", que permitam uma abstracção da realidade para um modelo computacional.

Sendo que não existe um conjunto de características "mágicas", as características SIFT revelam valor biométrico.

- são computacionalmente rápidas a extrair/detectar;
- podem ser extraídas em maior ou menor cardinalidade, dependendo dos parâmetros;
- são baseadas em apenas 4 propriedades localização  $(x, y)$ , escala, rotação  $\theta$ , algumas destas propriedades apresentam invariância de valores, sendo esta uma propriedade de maior valência.

Como dito anteriormente, a abstracção de um sistema computacional é uma das propriedades desejadas inerentes a qualquer sistema.

Duas vertentes revelam-se à primeira vista, na procura de um "modelo de fusão" com maior robustez:

- um "modelo de fusão" com outro conjunto de características, em paralelismo com as características SIFT;
- um "modelo de fusão" com outros sistemas biométricos de reconhecimento: íris, face, etc.

Também se poderá concluir que deve existir um compromisso entre o número de correspondências a procurar (espaço de procura = cardinalidade), e os parâmetros de extracção do SIFT.

Logicamente parâmetros de extracção pouco limitativos, correspondem a um maior número de keypoints, podendo então o espaço de procura ser mais alargado, executando correspondências com uma maior cardinalidade ou vice-versa.

Em suma, um maior número de keypoints extraídos duma imagem implica uma melhor probabilidade de reconhecimento.

### 5.1.1 Trabalho Futuro

Sendo que as experiências aqui efectuadas demonstram um potencial biométrico das características SIFT, seria de considerar a continuação do desenvolvimento de "modelo biométrico" proposto.

Numa perspectiva de trabalho futuro podem ser referidas as seguintes tarefas:

- implementar um modelo de classificação diferente do já implementado;
- utilizar um nº de imagens maior, uma grande Base de Dados de Imagens;
- imagens captadas de diferentes escalas , com iluminações variadas, apresentando vários graus de rotação.

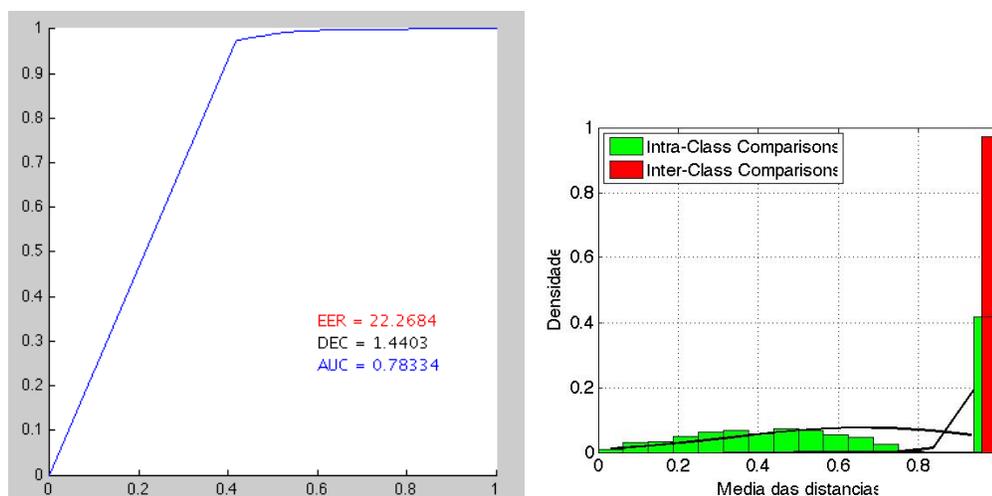


# Apêndice A

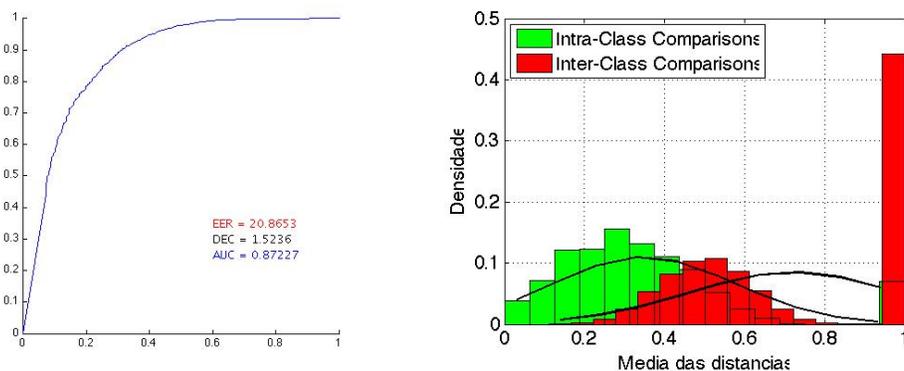
## Anexos

### A.1 Banco de Imagens A

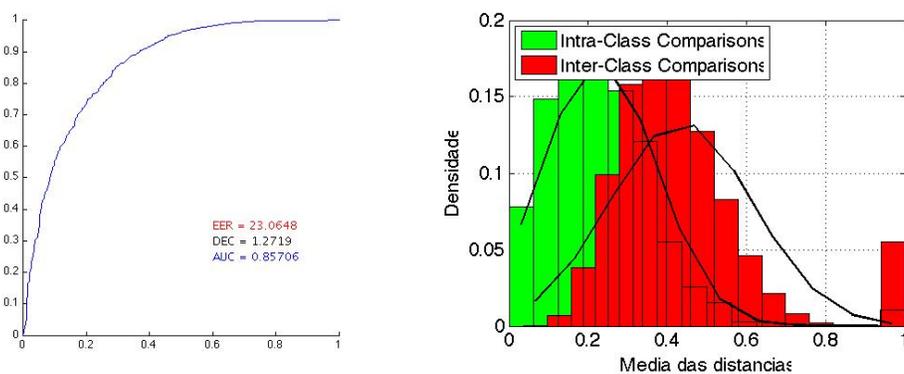
Resultados das avaliações feitas a todo o conjunto de imagens, com rotação e sem rotação.



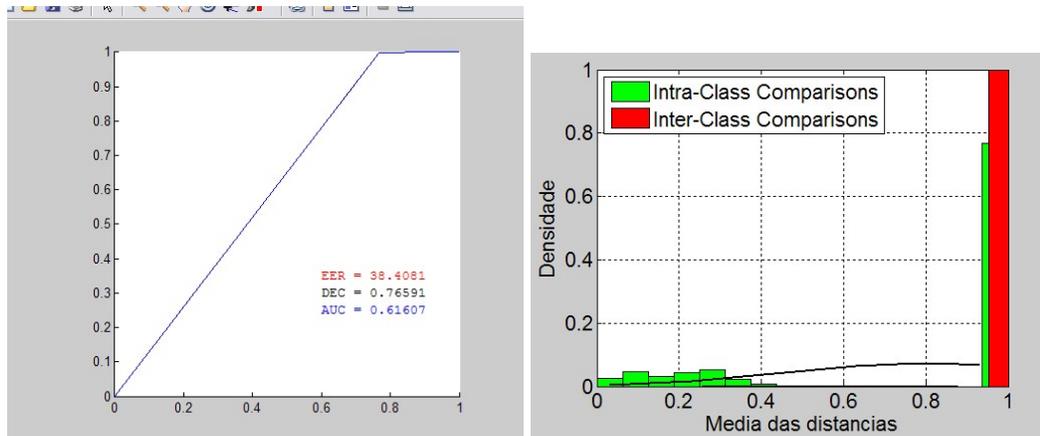
**Figura A.1:** Avaliação do Classificador na Configuração 1 para 20 keypoints (cardinalidade = 20).



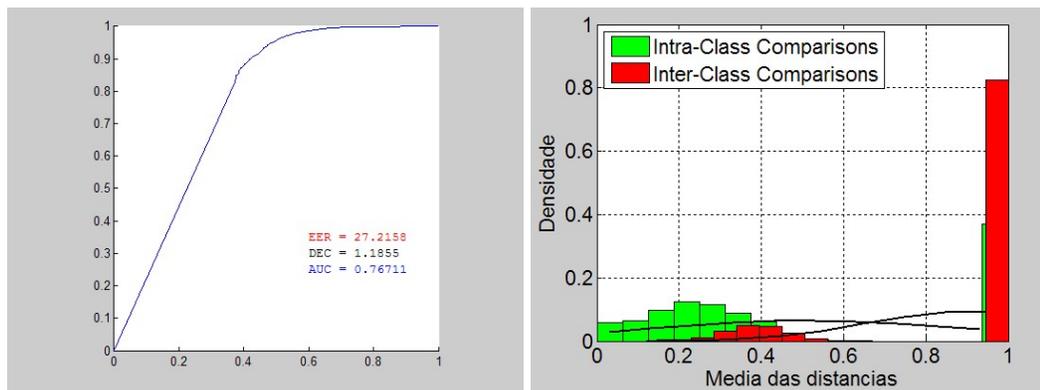
**Figura A.2:** Avaliação do Classificador na Configuração 1 para 10 keypoints (cardinalidade = 10).



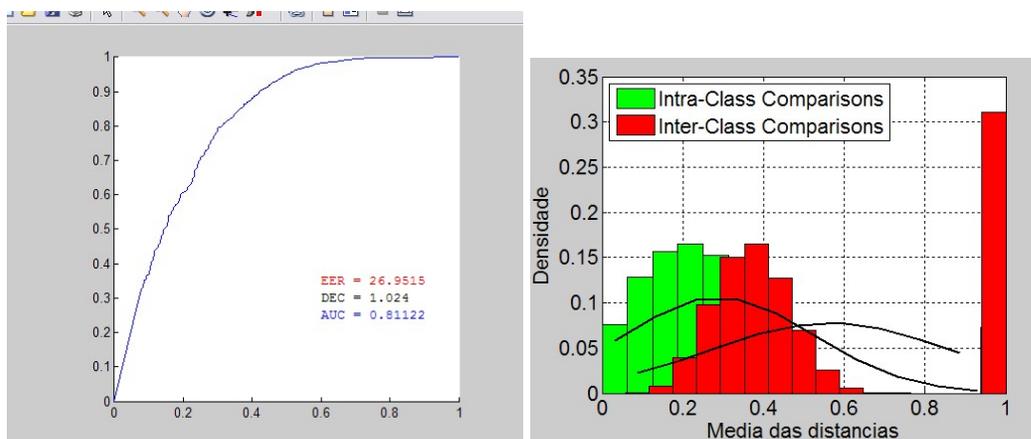
**Figura A.3:** Avaliação do Classificador na Configuração 1 para 5 keypoints (cardinalidade = 5).



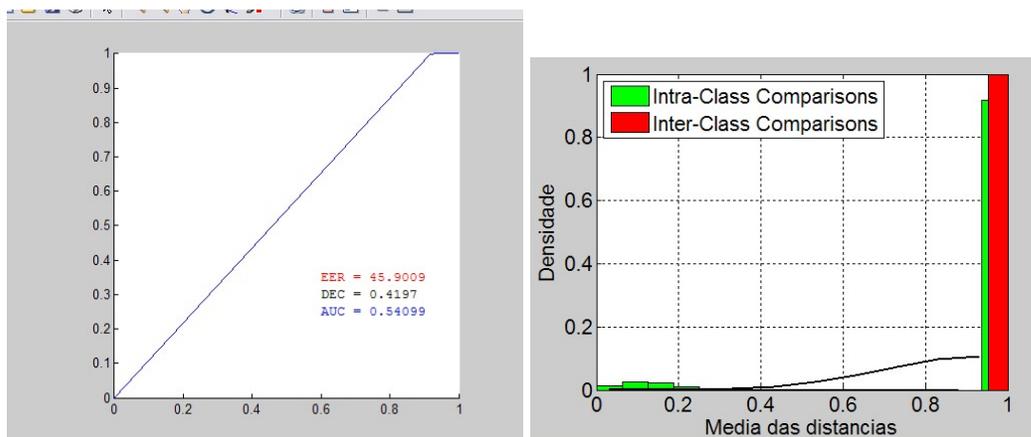
**Figura A.4:** Avaliação do Classificador na Configuração 2 para 20 keypoints (cardinalidade = 20), com máscara



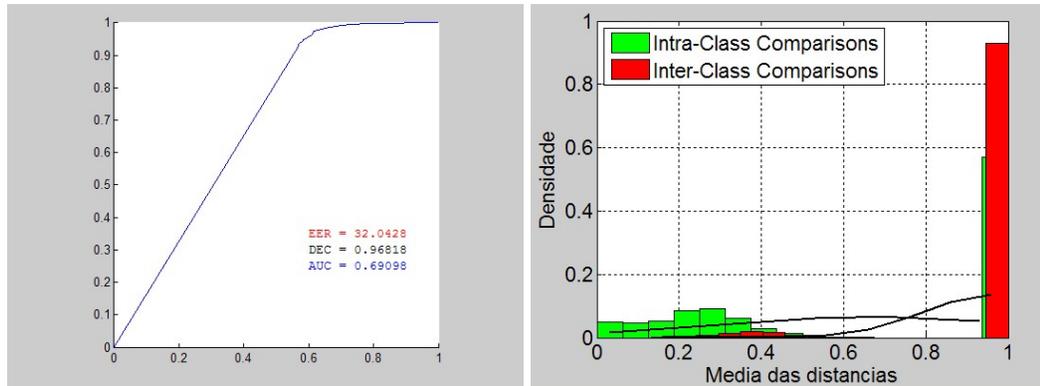
**Figura A.5:** Avaliação do Classificador na Configuração 2 para 10 keypoints (cardinalidade = 10), com máscara



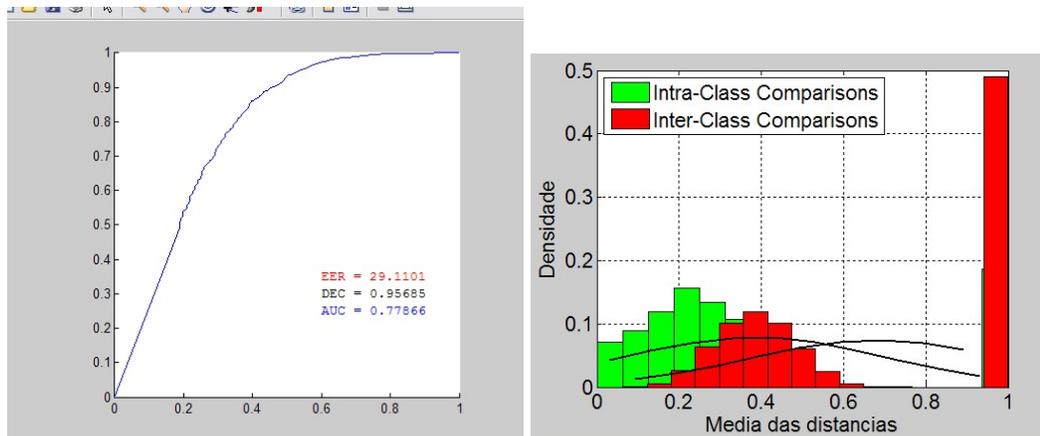
**Figura A.6:** Avaliação do Classificador na Configuração 2 para 5 keypoints (cardinalidade = 5), com máscara



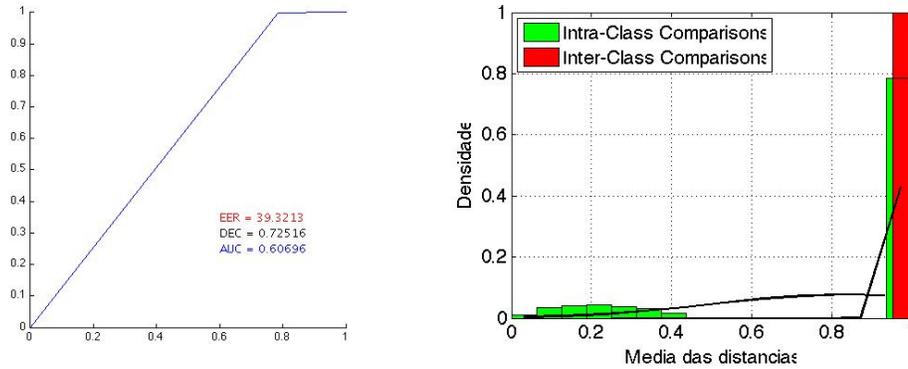
**Figura A.7:** Avaliação do Classificador na Configuração 3 para 20 keypoints (cardinalidade = 20), com máscara, peakthreshold = 0.01



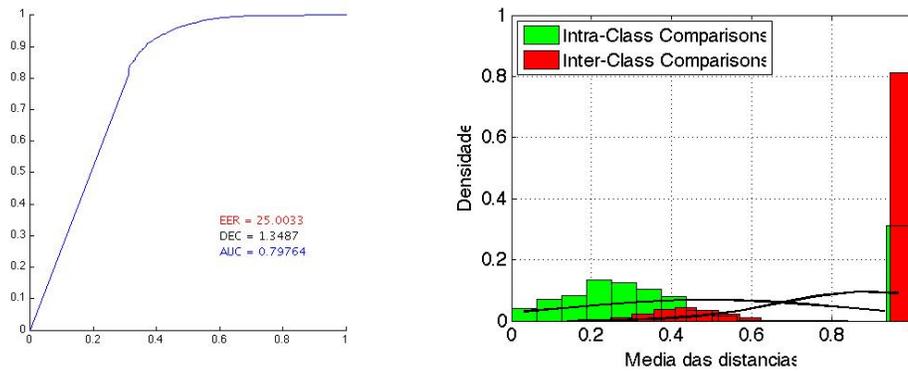
**Figura A.8:** Avaliação do Classificador na Configuração 3 para 10 keypoints (cardinalidade = 10), com máscara, peakthreshold = 0.01



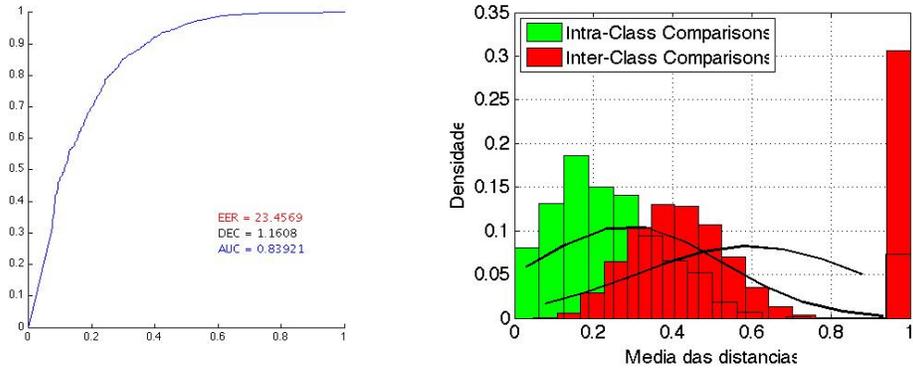
**Figura A.9:** Avaliação do Classificador na Configuração 3 para 5 keypoints (cardinalidade = 5), com máscara, peakthreshold = 0.01



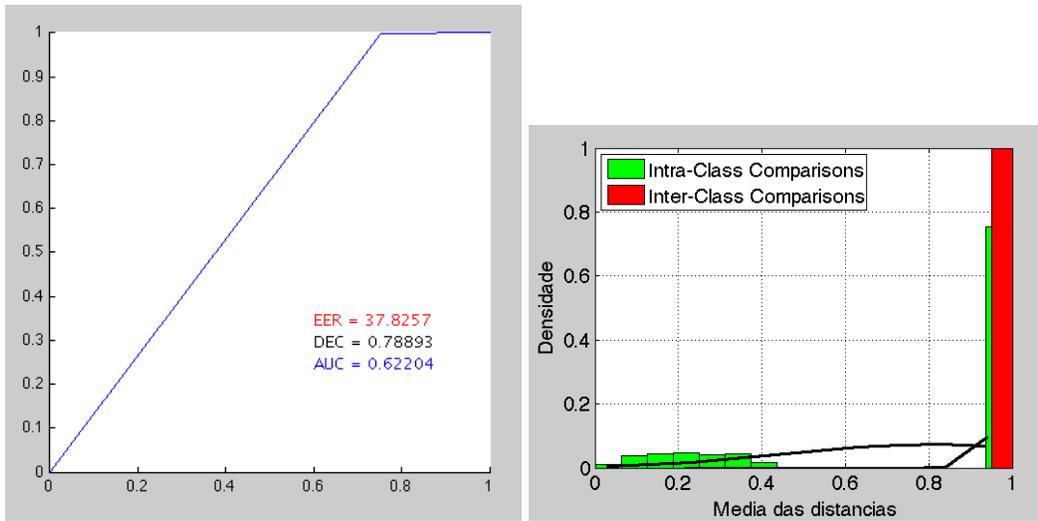
**Figura A.10:** Avaliação do Classificador na Configuração 4 para 20 keypoints (*cardinalidade* = 20), *peakthreshold* = 0.01 e *escala* > 2



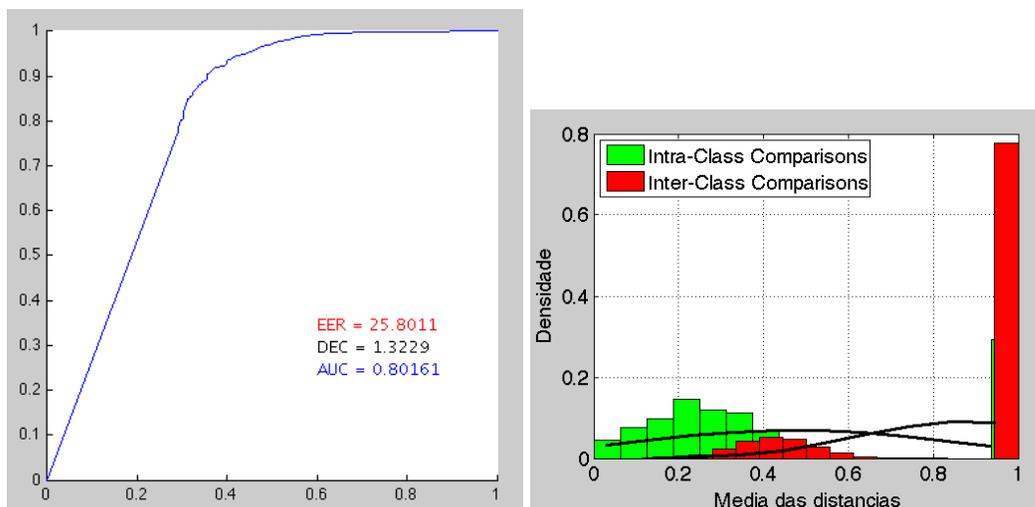
**Figura A.11:** Avaliação do Classificador na Configuração 4 para 10 keypoints (*cardinalidade* = 10), *peakthreshold* = 0.01 e *escala* > 2



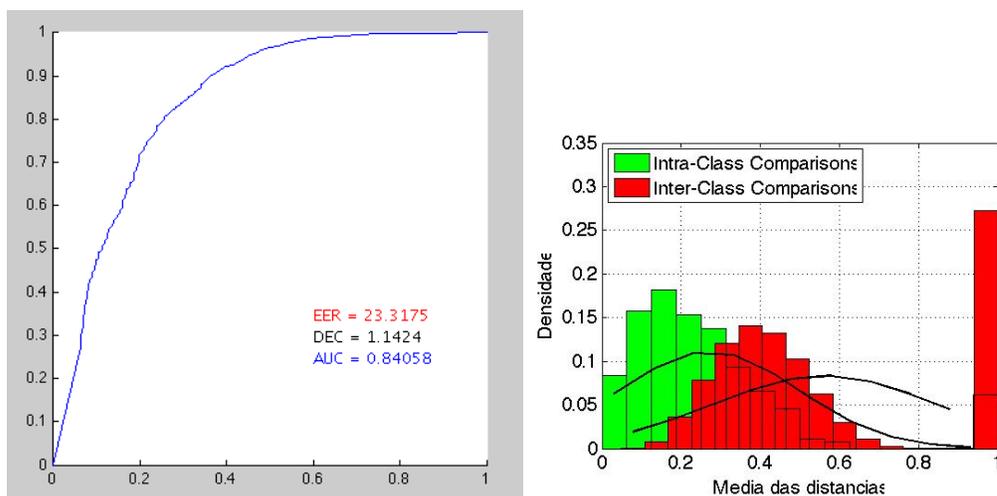
**Figura A.12:** Avaliação do Classificador na Configuração 4 para 5 keypoints (cardinalidade = 5), peakthreshold = 0.01 e escala > 2



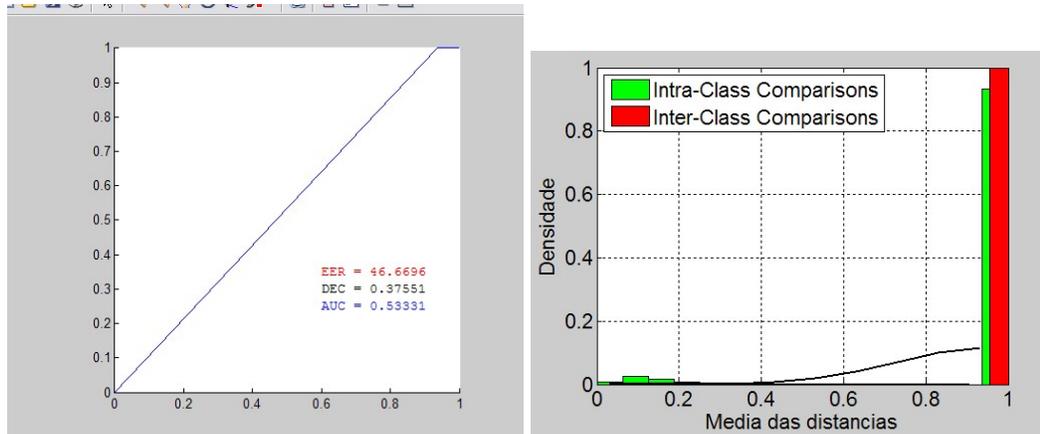
**Figura A.13:** Avaliação do Classificador na Configuração 5 para 20 keypoints (cardinalidade = 20), peakthreshold = 0.01



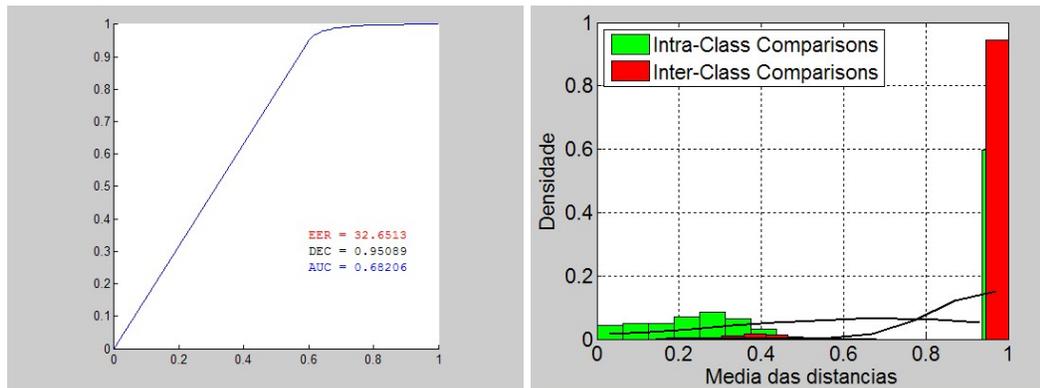
**Figura A.14:** Avaliação do Classificador na Configuração 5 para 10 keypoints (cardinalidade = 10),  $peakthreshold = 0.01$



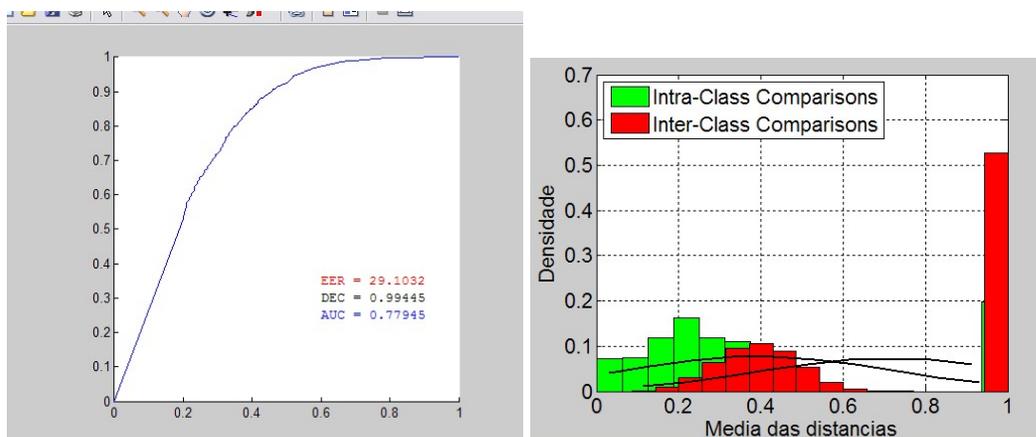
**Figura A.15:** Avaliação do Classificador na Configuração 5 para 5 keypoints (cardinalidade = 5),  $peakthreshold = 0.01$



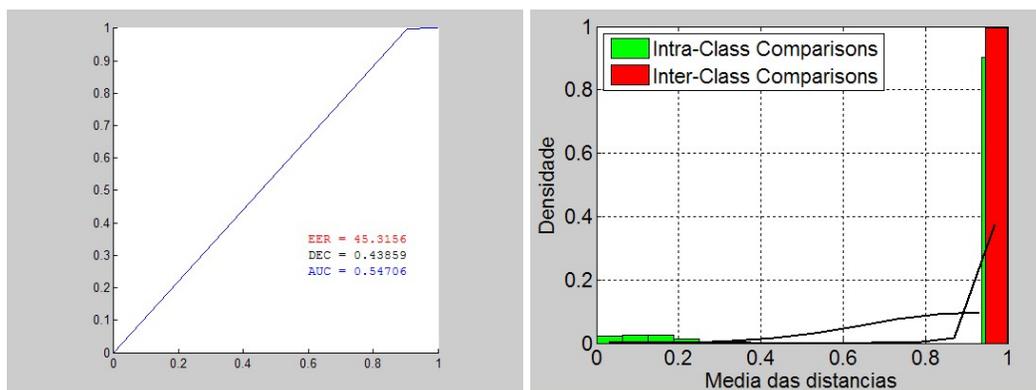
**Figura A.16:** Avaliação do Classificador na Configuração 6 para 20 keypoints (cardinalidade = 20), com máscara,  $peakthreshold = 0.01$  e escala  $\geq 2$



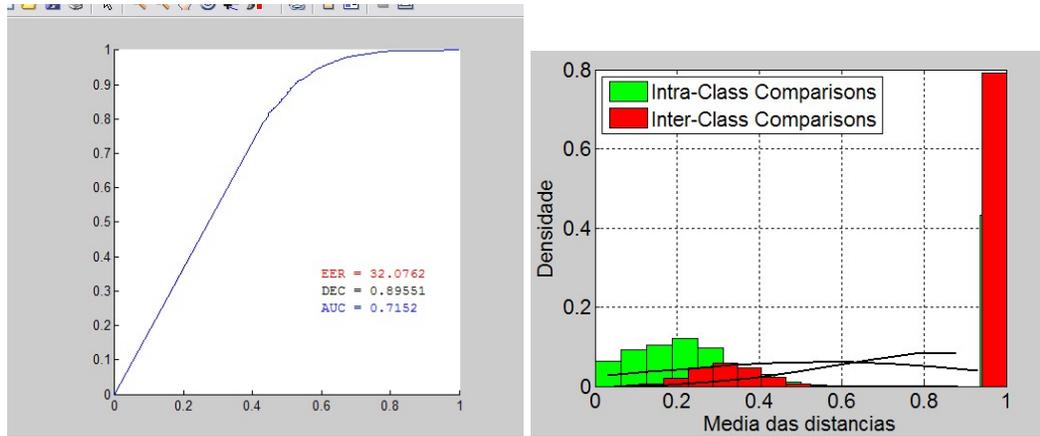
**Figura A.17:** Avaliação do Classificador na Configuração 6 para 10 keypoints (cardinalidade = 10), com máscara,  $peakthreshold = 0.01$  e escala  $\geq 2$



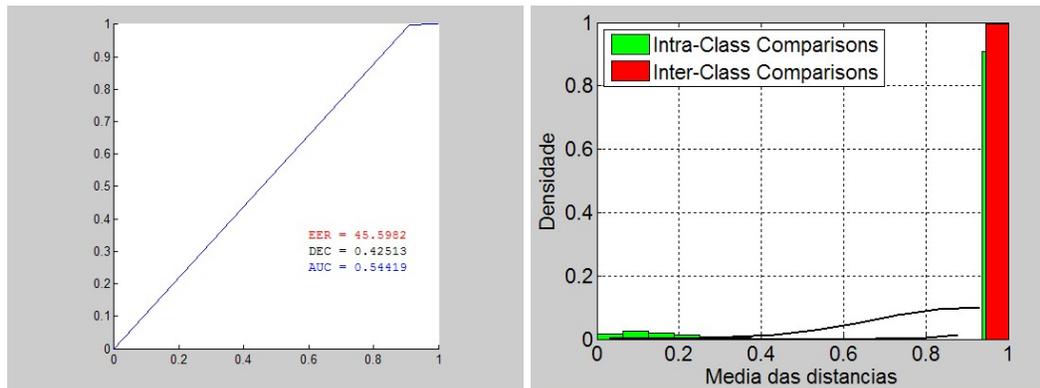
**Figura A.18:** Avaliação do Classificador na Configuração 6 para 5 keypoints (cardinalidade = 5), com máscara,  $peakthreshold = 0.01$  e escala  $\geq 2$



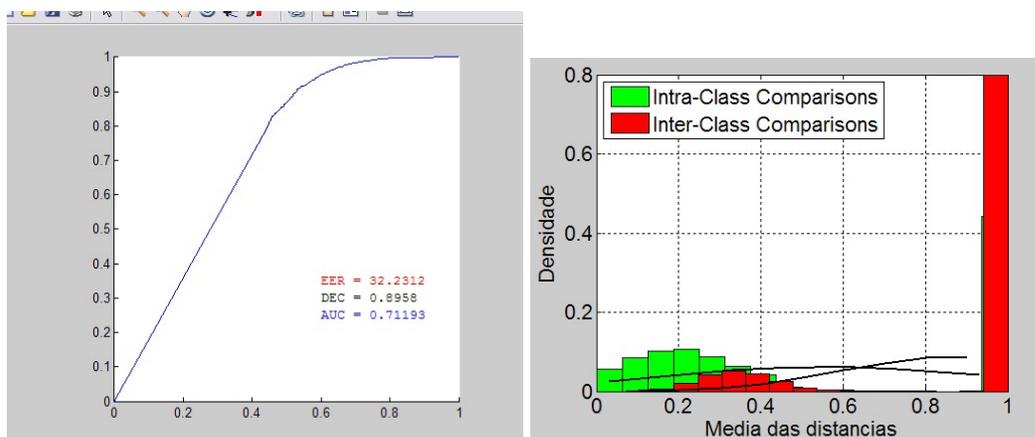
**Figura A.19:** Avaliação do Classificador na Configuração 7 para 10 keypoints (cardinalidade = 10), com máscara,  $peakthreshold = 0.02$



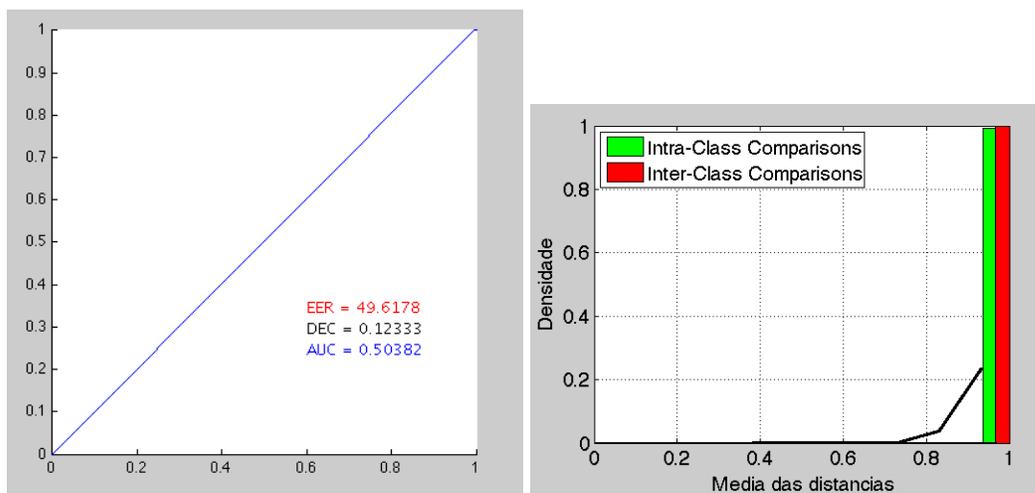
**Figura A.20:** Avaliação do Classificador na Configuração 7 para 5 keypoints (cardinalidade = 5), com máscara,  $peakthreshold = 0.02$



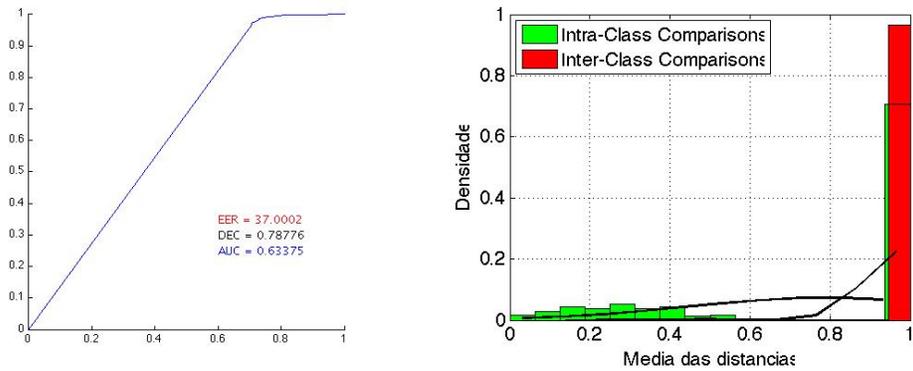
**Figura A.21:** Avaliação do Classificador na Configuração 8 para 10 keypoints (cardinalidade = 10), com máscara,  $peakthreshold = 0.02$  e escala  $\geq 2$



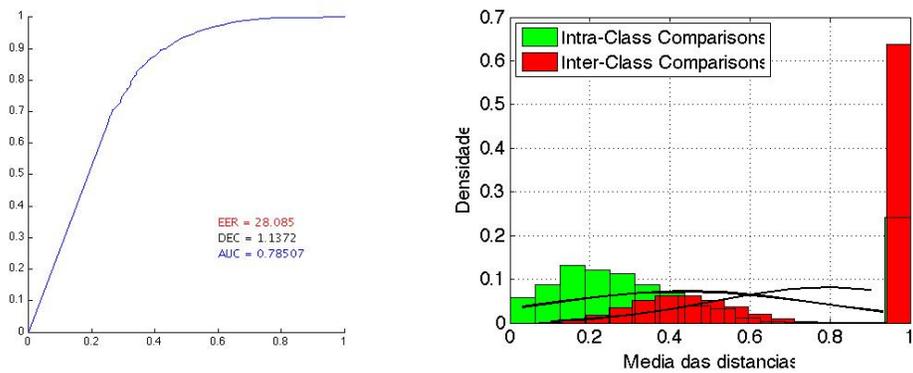
**Figura A.22:** Avaliação do Classificador na Configuração 8 para 5 keypoints (cardinalidade = 5), com máscara,  $peakthreshold = 0.02$  e escala  $\geq 2$



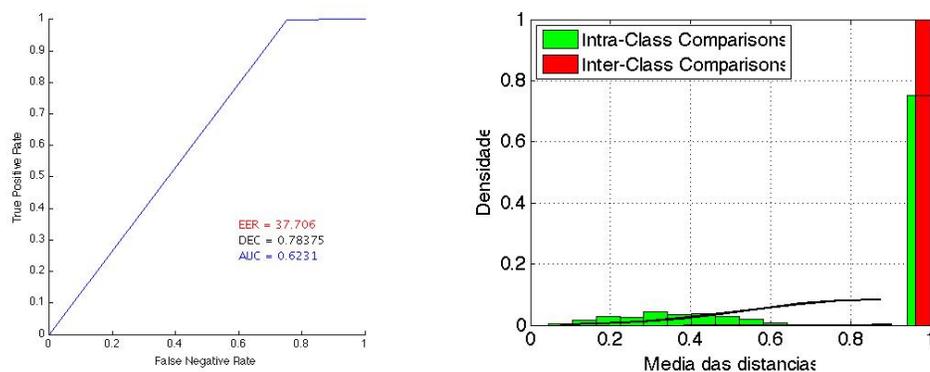
**Figura A.23:** Avaliação do Classificador na Configuração 9 para 20 keypoints (cardinalidade = 20),  $peakthreshold = 0.02$



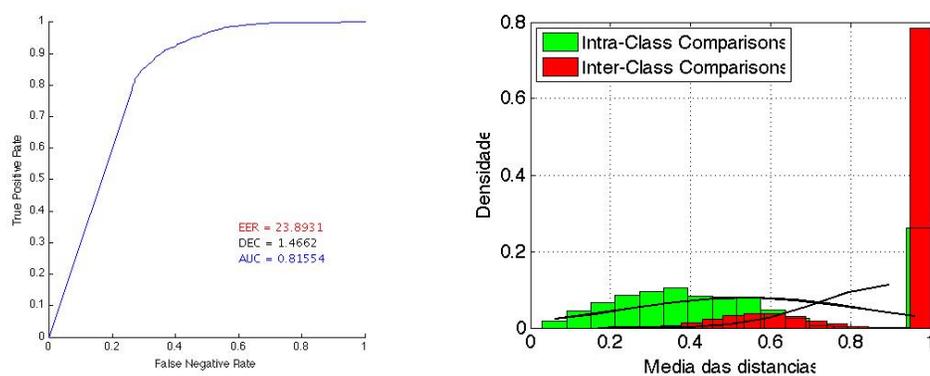
**Figura A.24:** Avaliação do Classificador na Configuração 9 para 10 keypoints (cardinalidade = 10), peakthreshold = 0.02



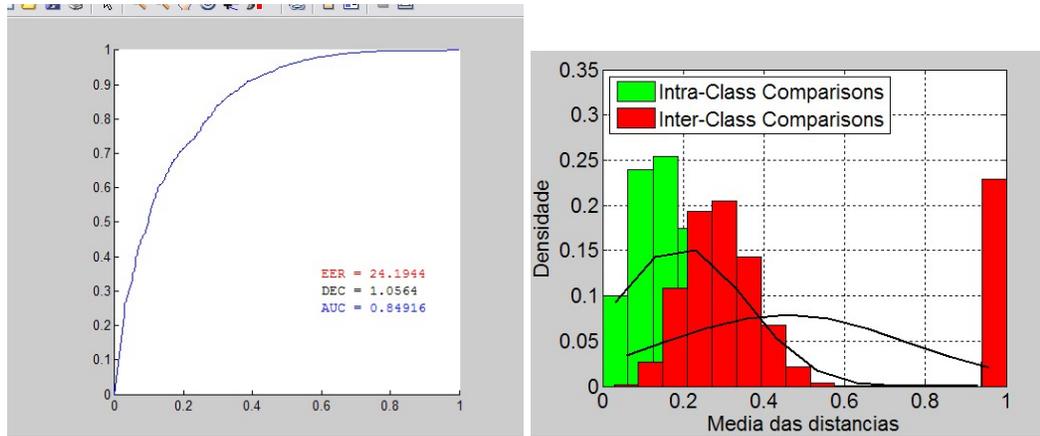
**Figura A.25:** Avaliação do Classificador na Configuração 9 para 5 keypoints (cardinalidade = 5), peakthreshold = 0.02



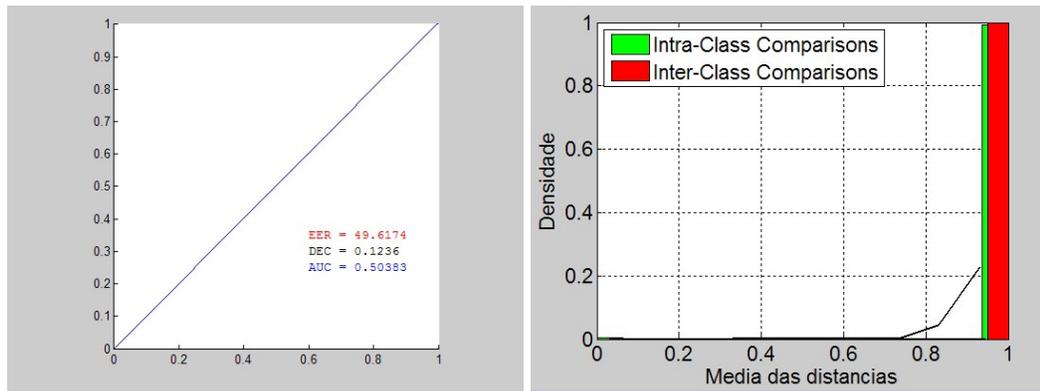
**Figura A.26:** Avaliação do Classificador na Configuração 10 para 20 keypoints (cardinalidade = 20), orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



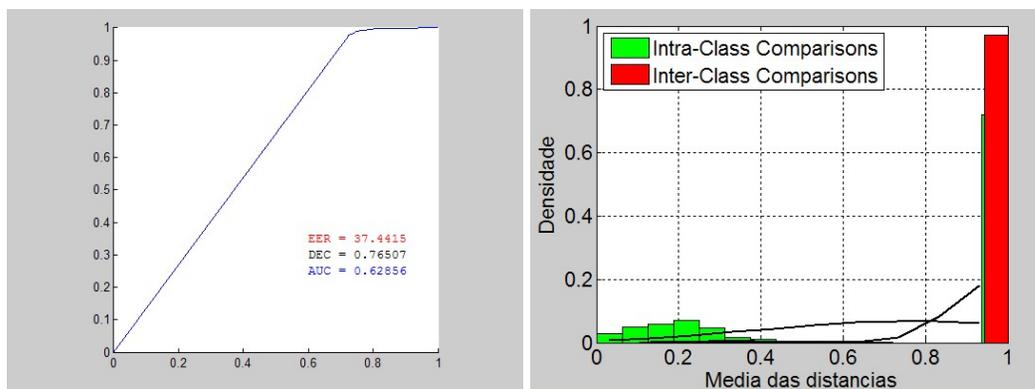
**Figura A.27:** Avaliação do Classificador na Configuração 10 para 10 keypoints (cardinalidade = 10), orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



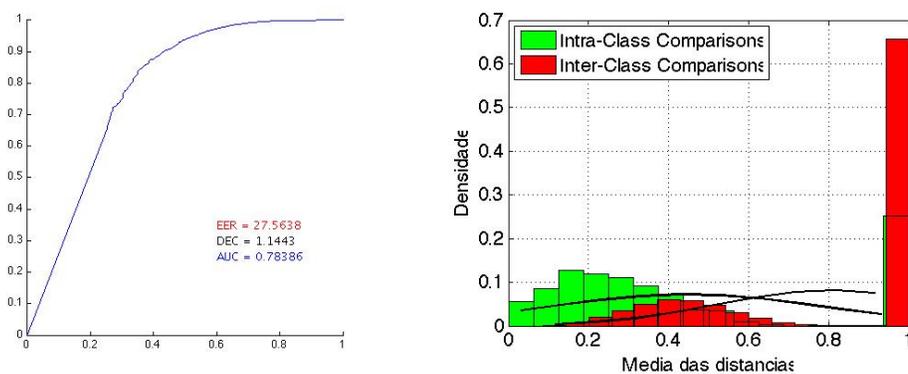
**Figura A.28:** Avaliação do Classificador na Configuração 10 para 5 keypoints (cardinalidade = 5), orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



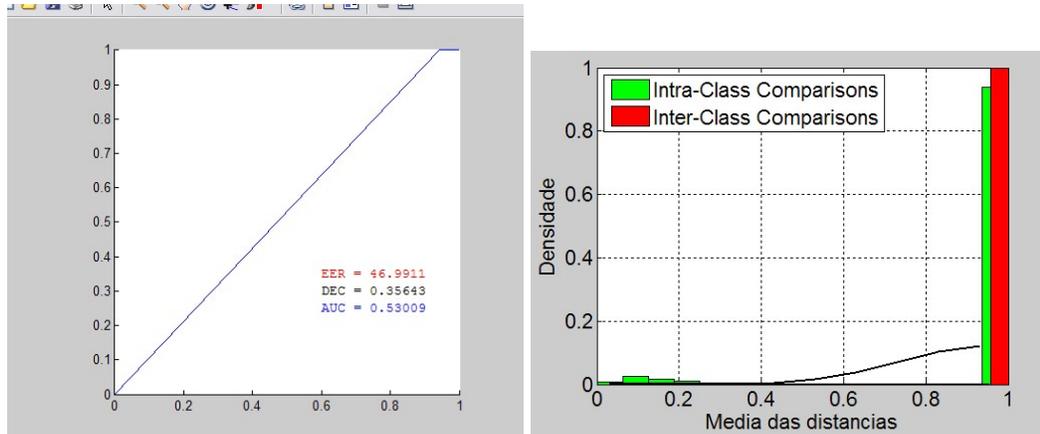
**Figura A.29:** Avaliação do Classificador na Configuração 11 para 20 keypoints (cardinalidade = 20),  $peakthreshold = 0.02$  e  $escala \geq 2$



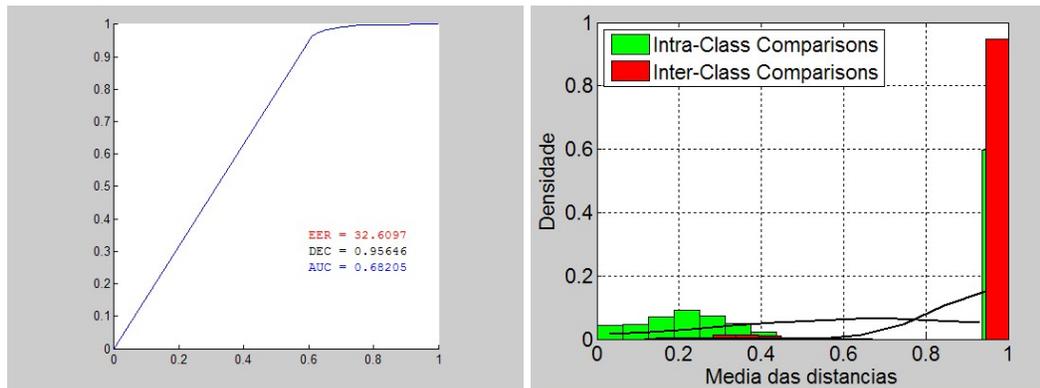
**Figura A.30:** Avaliação do Classificador na Configuração 11 para 10 keypoints (cardinalidade = 10),  $peakthreshold = 0.02$  e escala  $\geq 2$



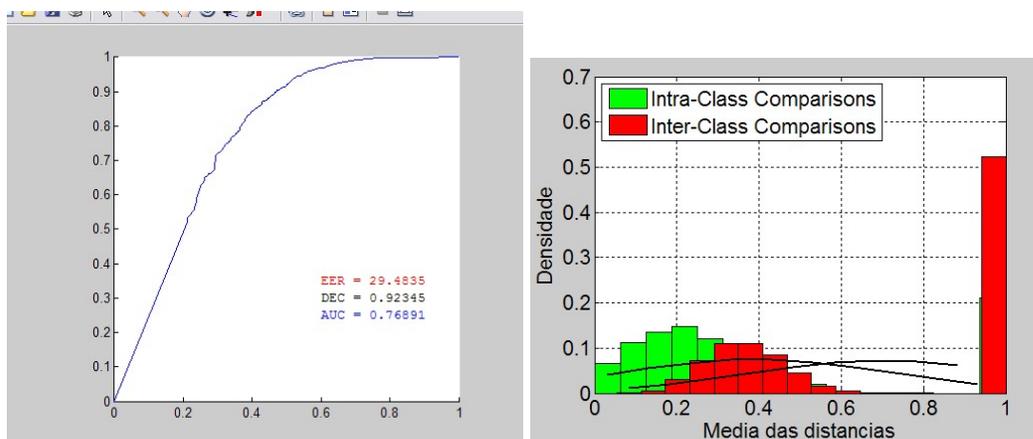
**Figura A.31:** Avaliação do Classificador na Configuração 11 para 5 keypoints (cardinalidade = 5),  $peakthreshold = 0.02$  e escala  $\geq 2$



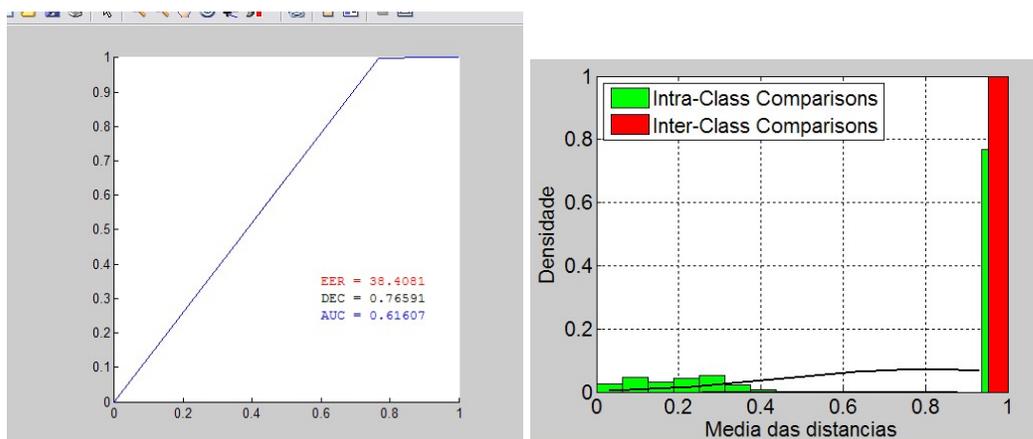
**Figura A.32:** Avaliação do Classificador na Configuração 12 para 20 keypoints (cardinalidade = 20), com máscara, orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



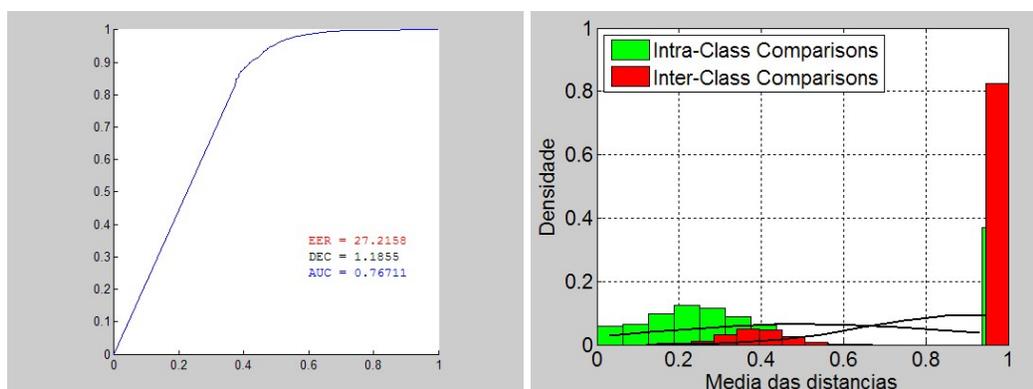
**Figura A.33:** Avaliação do Classificador na Configuração 12 para 10 keypoints (cardinalidade = 10), com máscara, orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



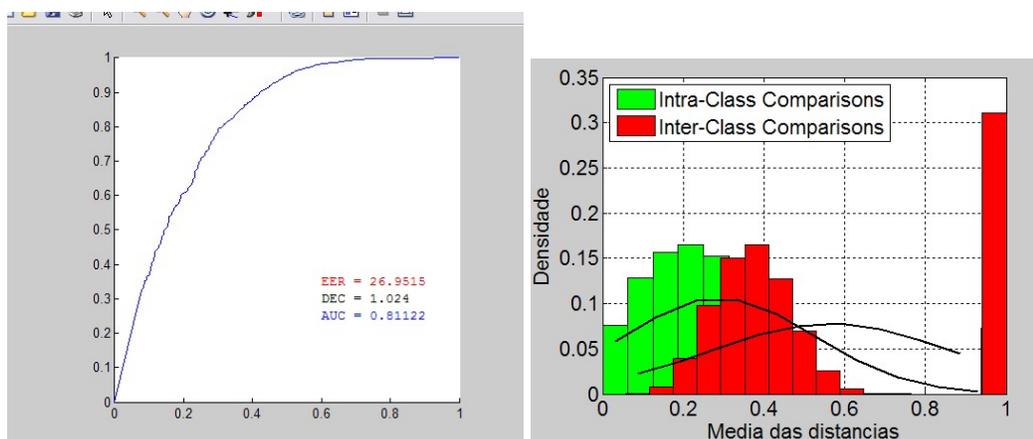
**Figura A.34:** Avaliação do Classificador na Configuração 12 para 5 keypoints (cardinalidade = 5), com máscara, orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



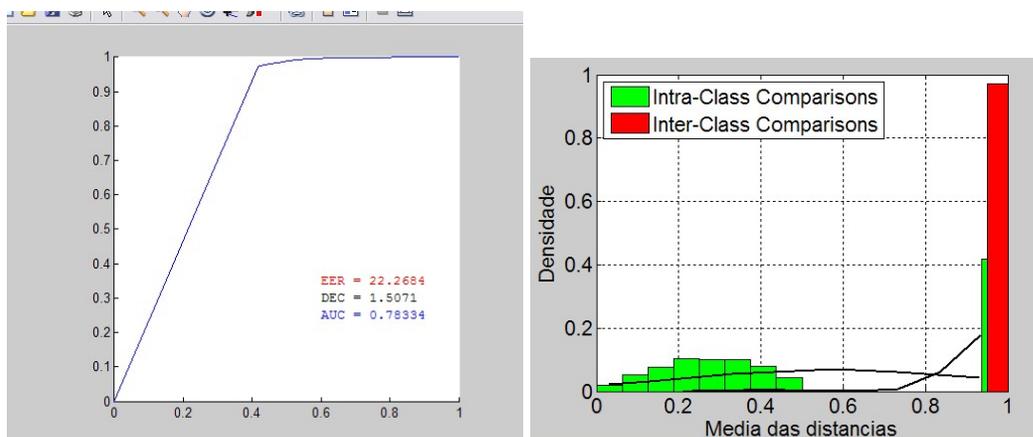
**Figura A.35:** Avaliação do Classificador na Configuração 13 para 20 keypoints (cardinalidade = 20), com máscara, orientação de keypoints  $-3\pi/2 < \theta < \pi/2$



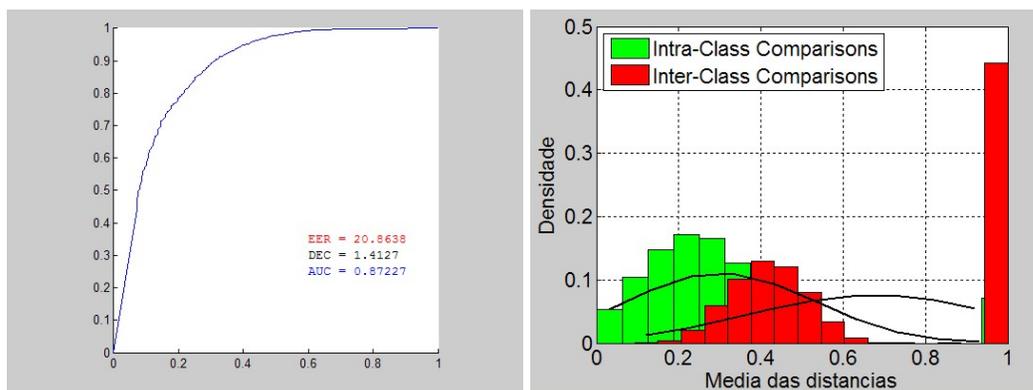
**Figura A.36:** Avaliação do Classificador na Configuração 13 para 10 keypoints (cardinalidade = 10), com máscara, orientação de keypoints  $-3\pi/2 < \theta < \pi/2$



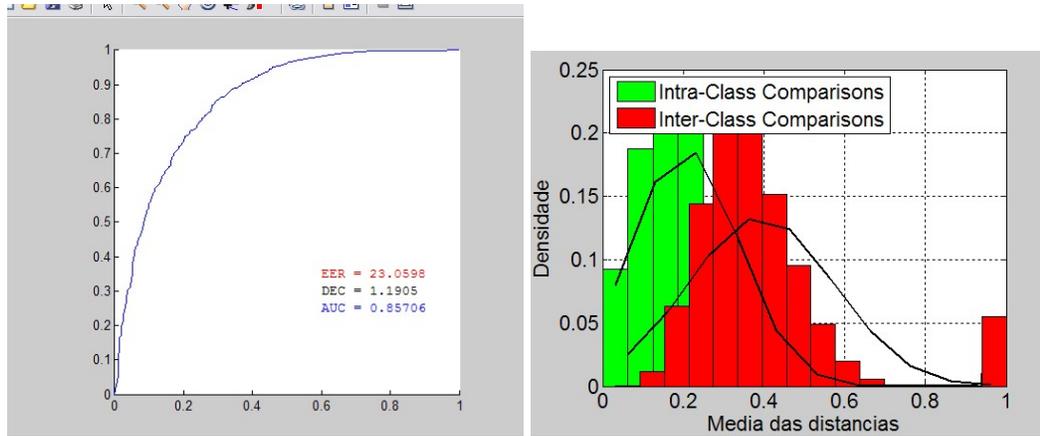
**Figura A.37:** Avaliação do Classificador na Configuração 13 para 5 keypoints (cardinalidade = 5), com máscara, orientação de keypoints  $-3\pi/2 < \theta < \pi/2$



**Figura A.38:** Avaliação do Classificador na Configuração 14 para 20 keypoints (cardinalidade = 20), orientação de keypoints  $-3\pi/2 < \theta < \pi/2$



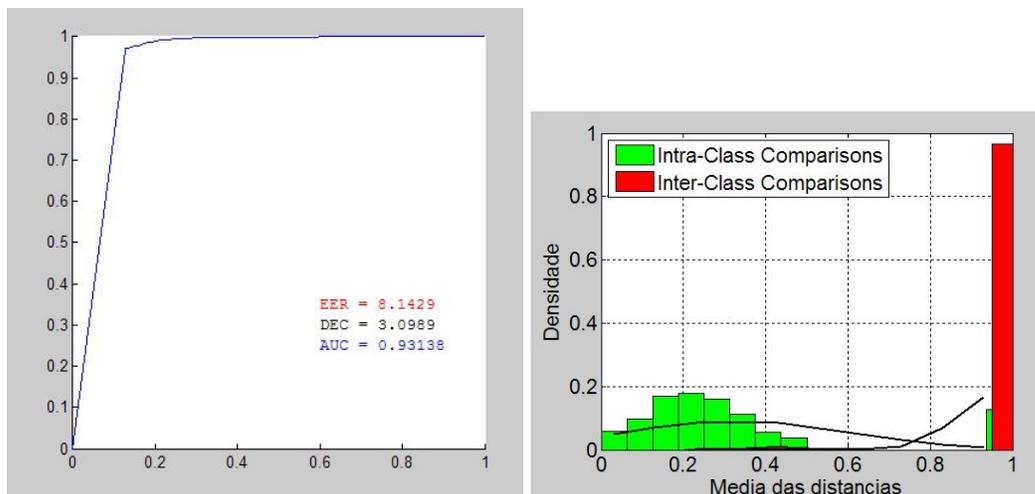
**Figura A.39:** Avaliação do Classificador na Configuração 14 para 10 keypoints (cardinalidade = 10), orientação de keypoints  $-3\pi/2 < \theta < \pi/2$



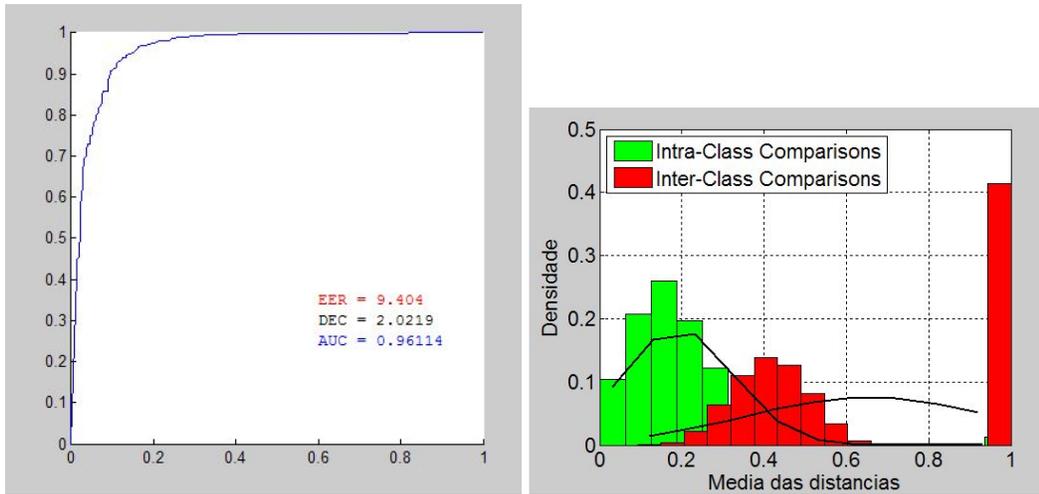
**Figura A.40:** Avaliação do Classificador na Configuração 14 para 5 keypoints (cardinalidade = 5), orientação de keypoints  $-3\pi/2 < \theta < \pi/2$

## A.2 Banco de Imagens B

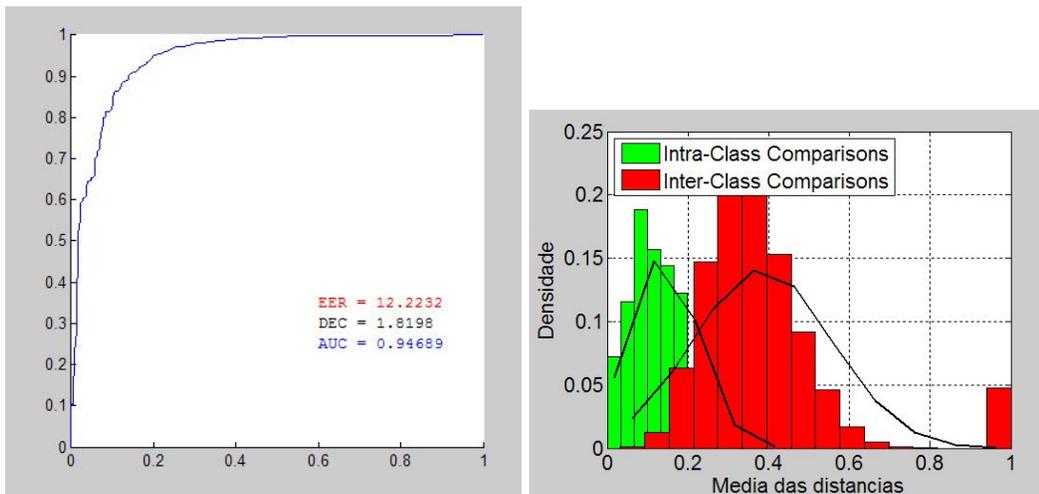
Resultado das avaliações feitas ao conjunto de imagens, apenas com rotação.



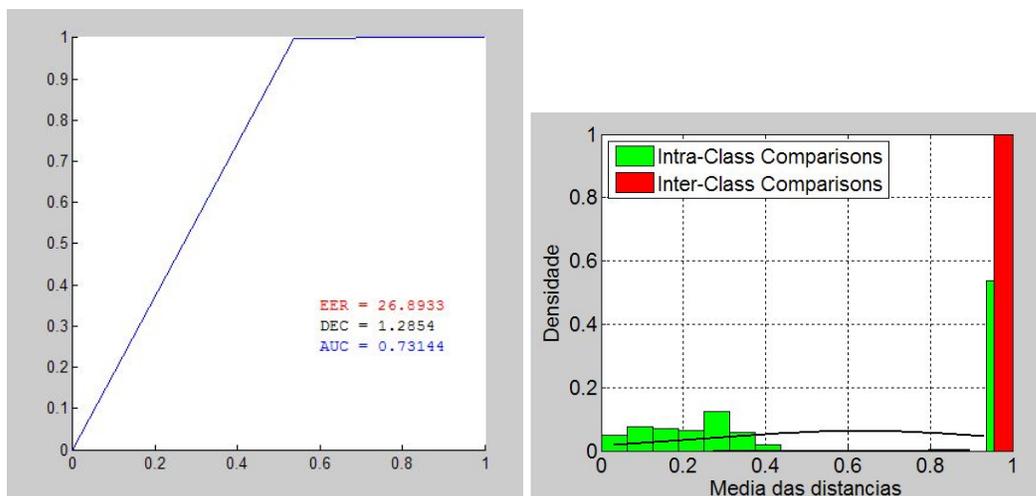
**Figura A.41:** Avaliação do Classificador na Configuração 1 para 20 keypoints (cardinalidade = 20).



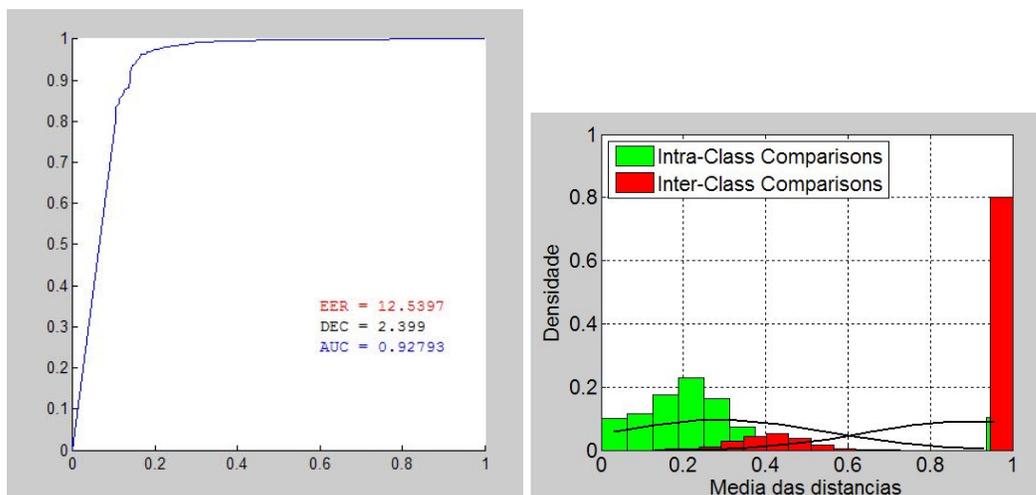
**Figura A.42:** Avaliação do Classificador na Configuração 1 para 10 keypoints (cardinalidade = 10).



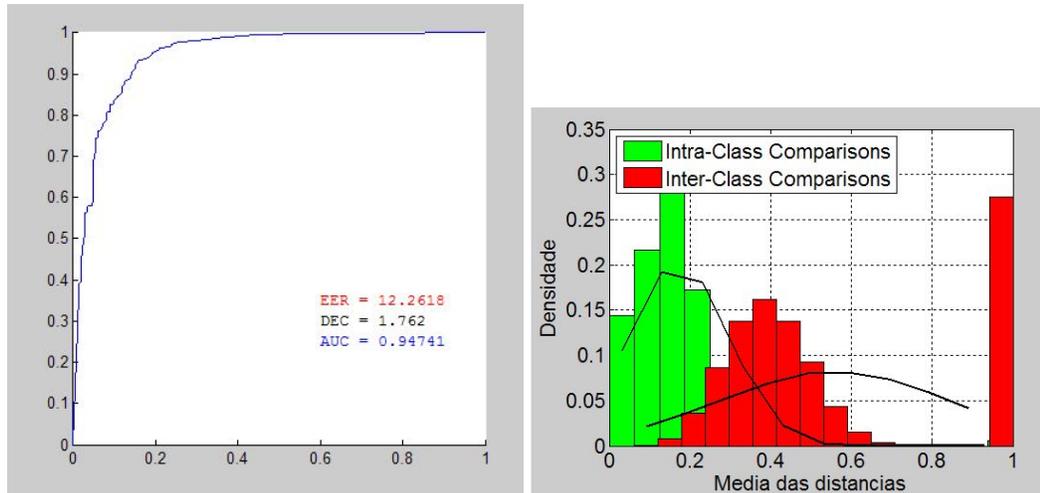
**Figura A.43:** Avaliação do Classificador na Configuração 1 para 5 keypoints (cardinalidade = 5).



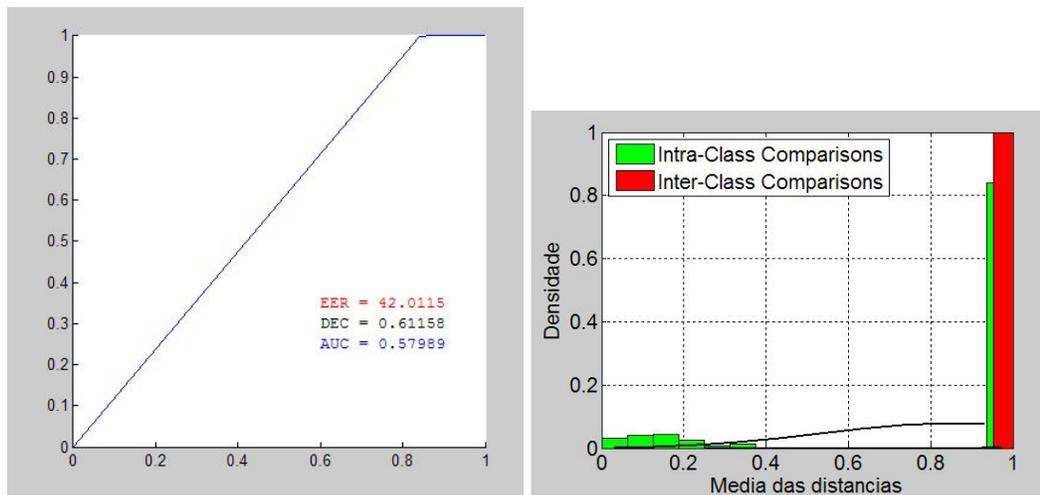
**Figura A.44:** Avaliação do Classificador na Configuração 2 para 20 keypoints (cardinalidade = 20), com máscara



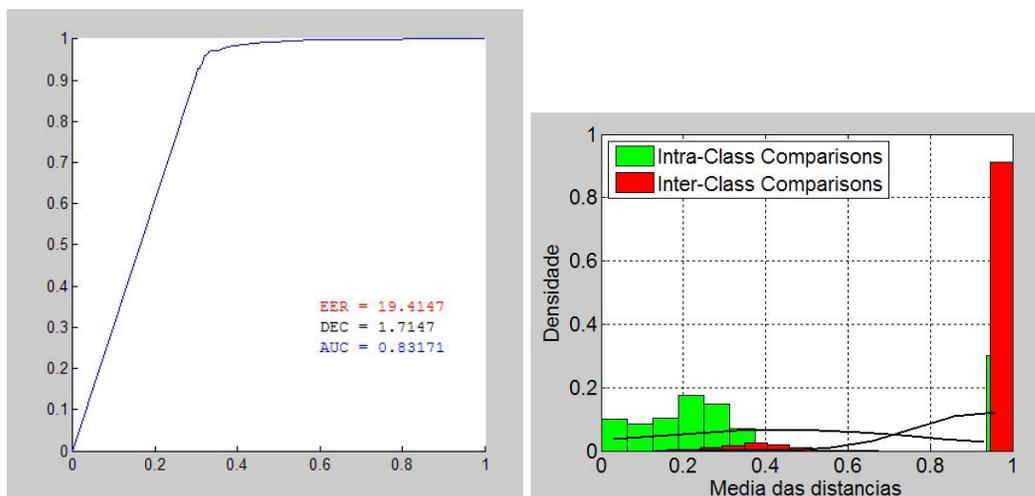
**Figura A.45:** Avaliação do Classificador na Configuração 2 para 10 keypoints (cardinalidade = 10), com máscara



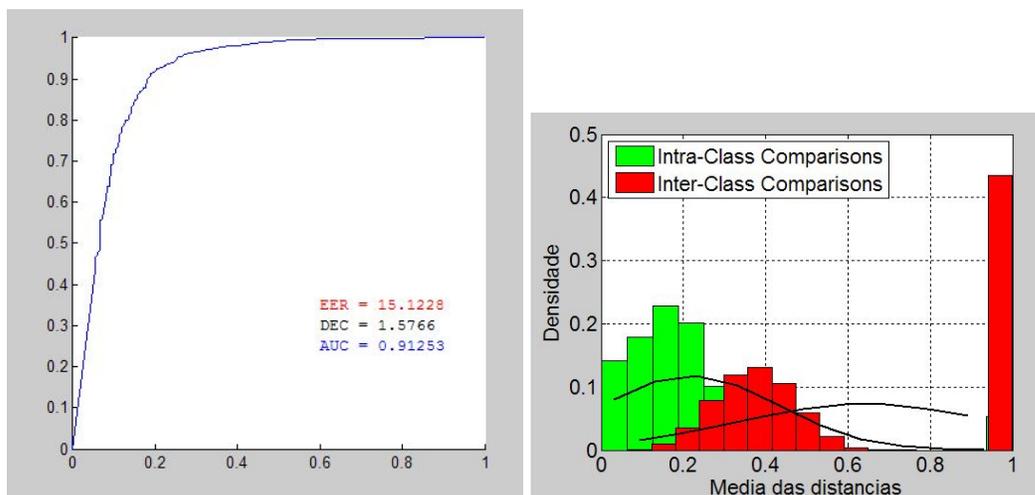
**Figura A.46:** Avaliação do Classificador na Configuração 2 para 5 keypoints (cardinalidade = 5), com máscara



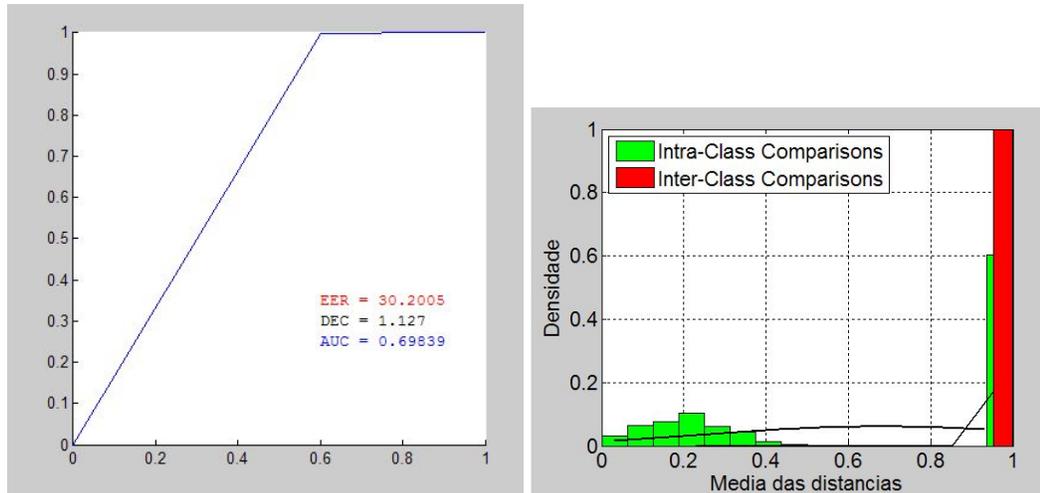
**Figura A.47:** Avaliação do Classificador na Configuração 3 para 20 keypoints (cardinalidade = 20), com máscara, peakthreshold = 0.01



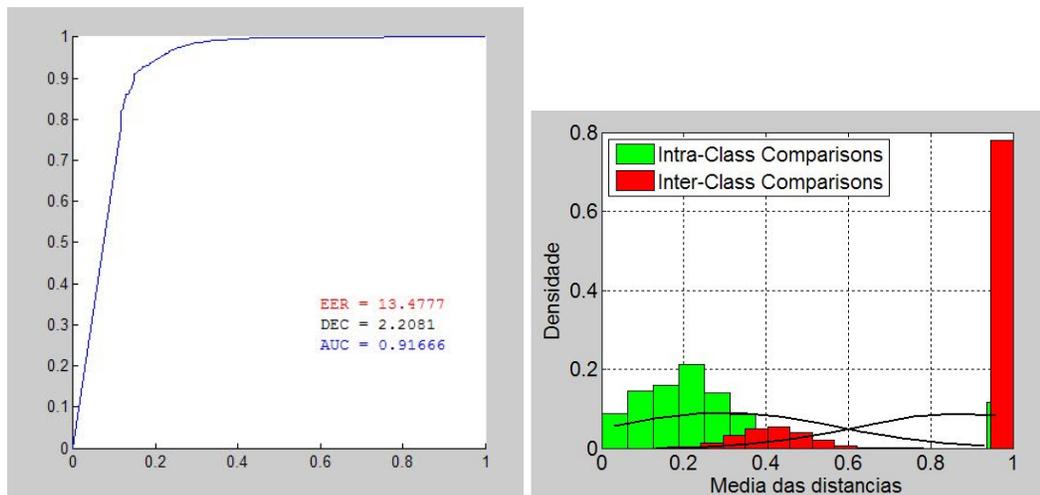
**Figura A.48:** Avaliação do Classificador na Configuração 3 para 10 keypoints (cardinalidade = 10), com máscara,  $peakthreshold = 0.01$



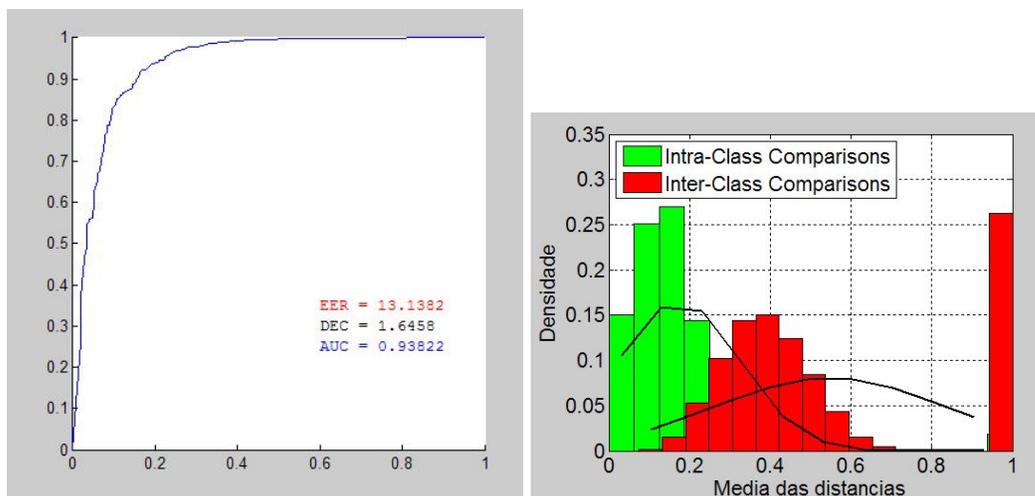
**Figura A.49:** Avaliação do Classificador na Configuração 3 para 5 keypoints (cardinalidade = 5), com máscara,  $peakthreshold = 0.01$



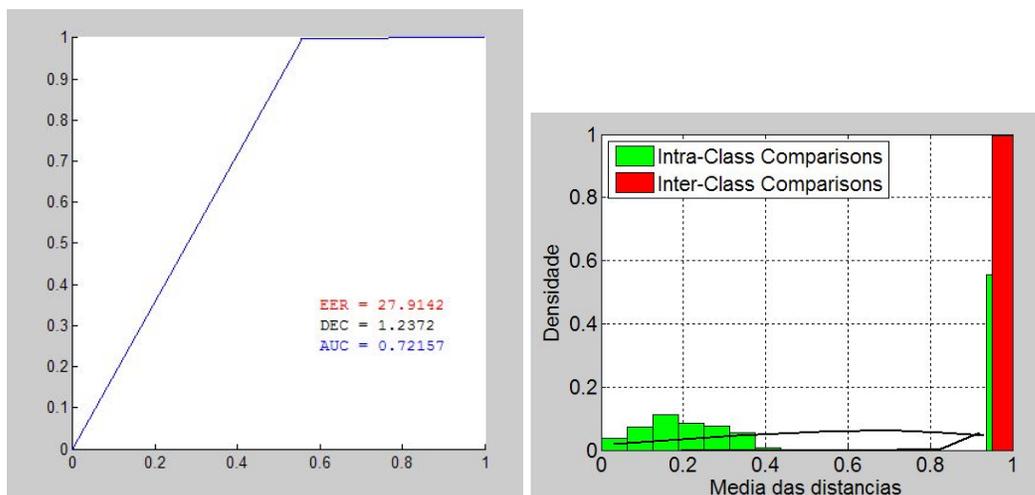
**Figura A.50:** Avaliação do Classificador na Configuração 4 para 20 keypoints (cardinalidade = 20),  $peakthreshold = 0.01$  e  $escala > 2$



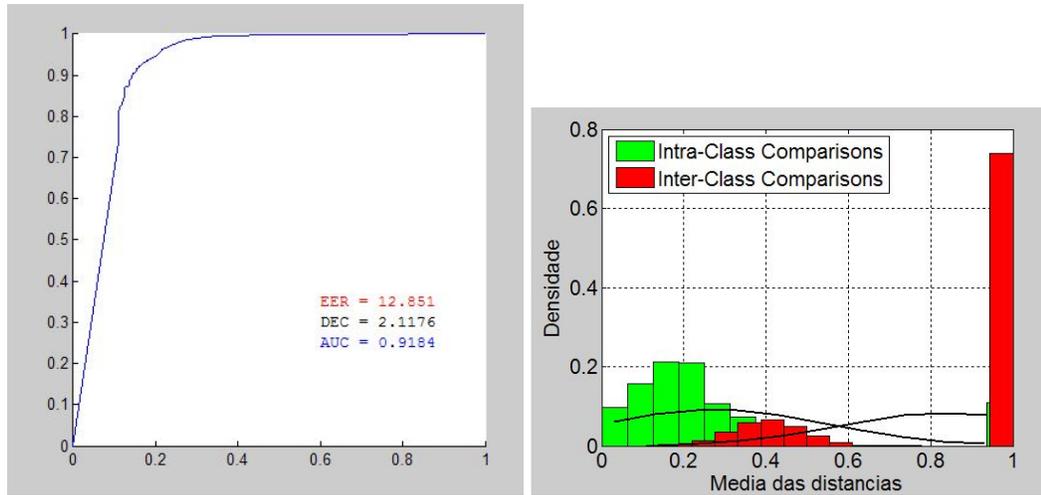
**Figura A.51:** Avaliação do Classificador na Configuração 4 para 10 keypoints (cardinalidade = 10),  $peakthreshold = 0.01$  e  $escala > 2$



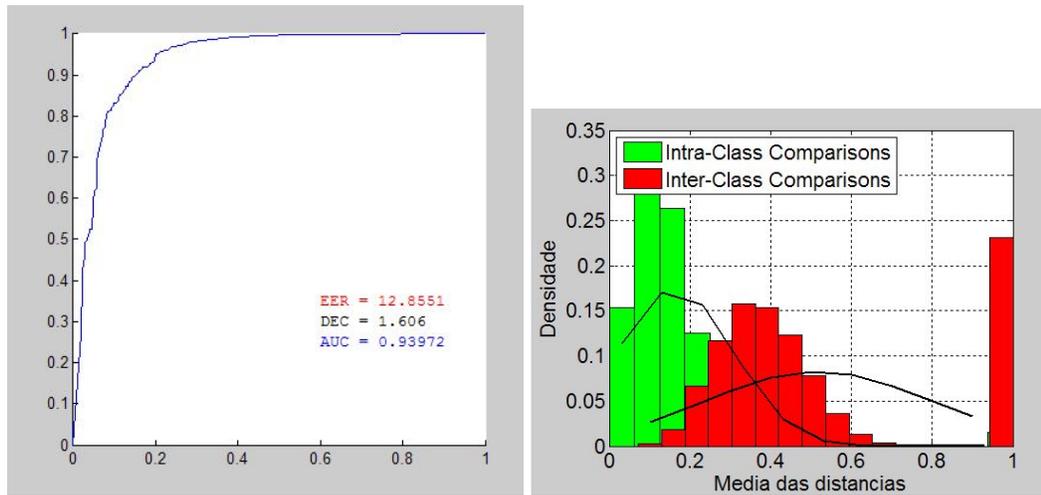
**Figura A.52:** Avaliação do Classificador na Configuração 4 para 5 keypoints (cardinalidade = 5),  $peakthreshold = 0.01$  e escala  $> 2$



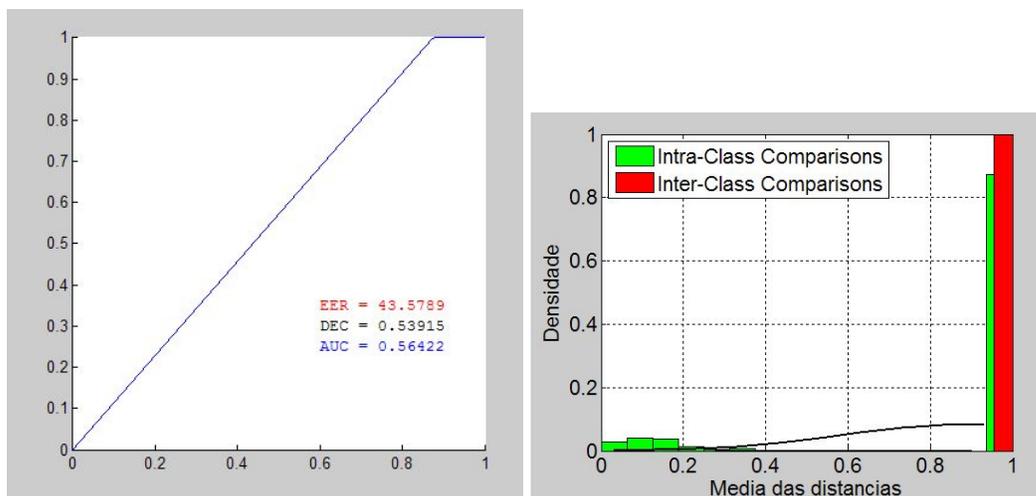
**Figura A.53:** Avaliação do Classificador na Configuração 5 para 20 keypoints (cardinalidade = 20),  $peakthreshold = 0.01$



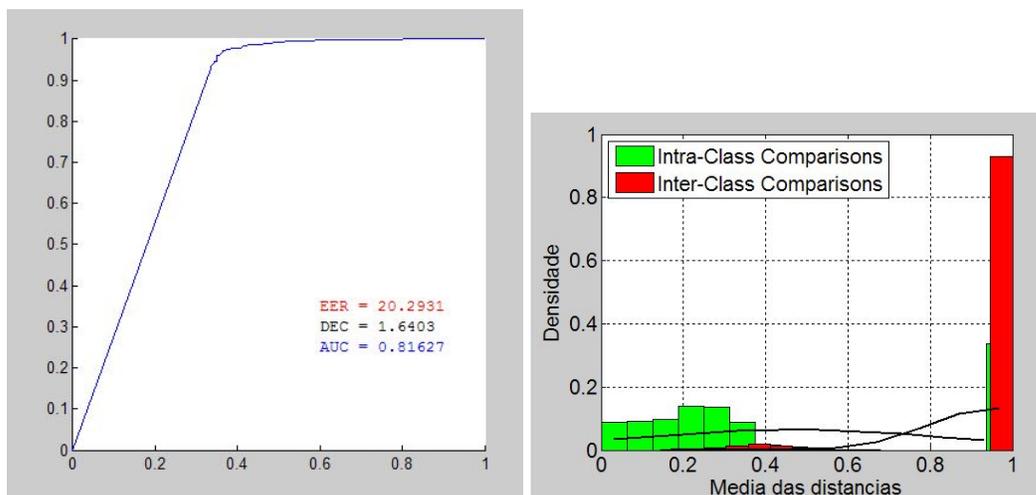
**Figura A.54:** Avaliação do Classificador na Configuração 5 para 10 keypoints (cardinalidade = 10),  $peakthreshold = 0.01$



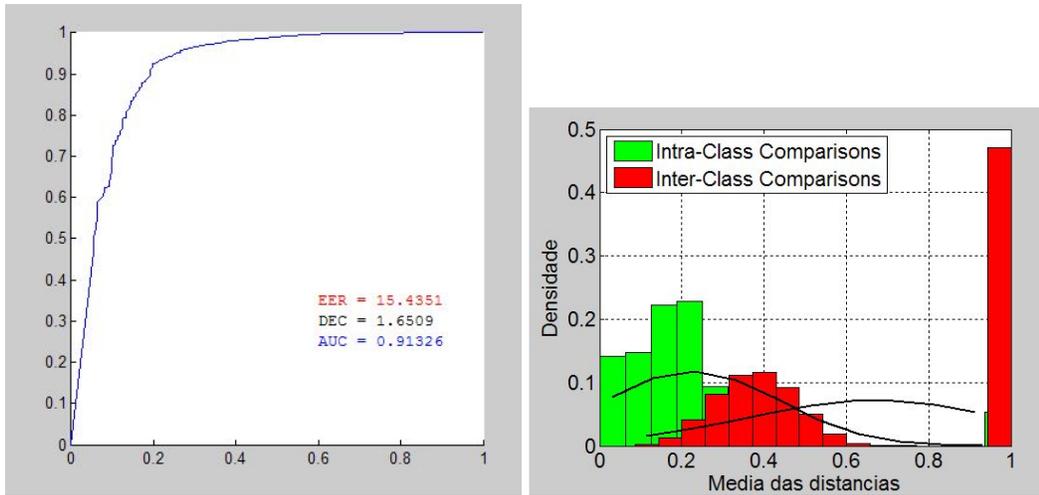
**Figura A.55:** Avaliação do Classificador na Configuração 5 para 5 keypoints (cardinalidade = 5),  $peakthreshold = 0.01$



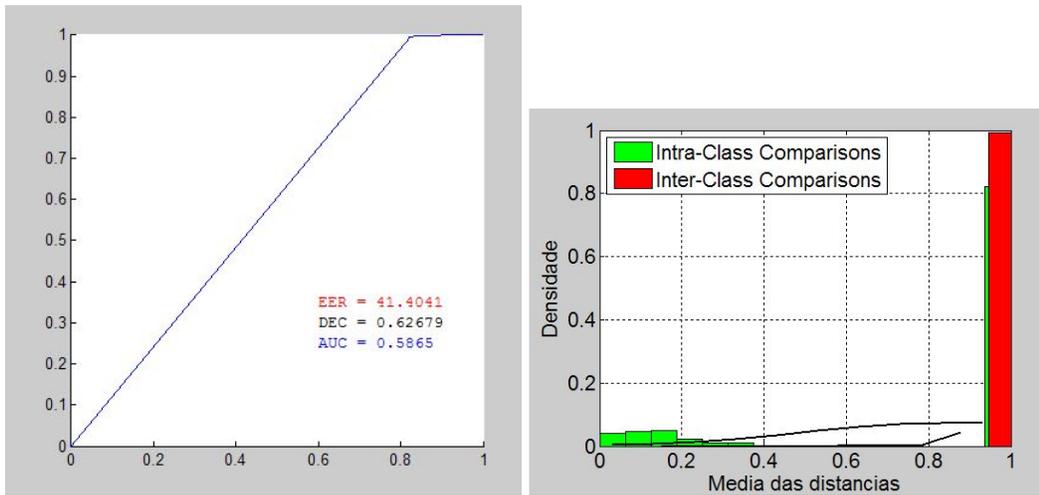
**Figura A.56:** Avaliação do Classificador na Configuração 6 para 20 keypoints (cardinalidade = 20), com máscara, peakthreshold = 0.01 e escala  $\geq 2$



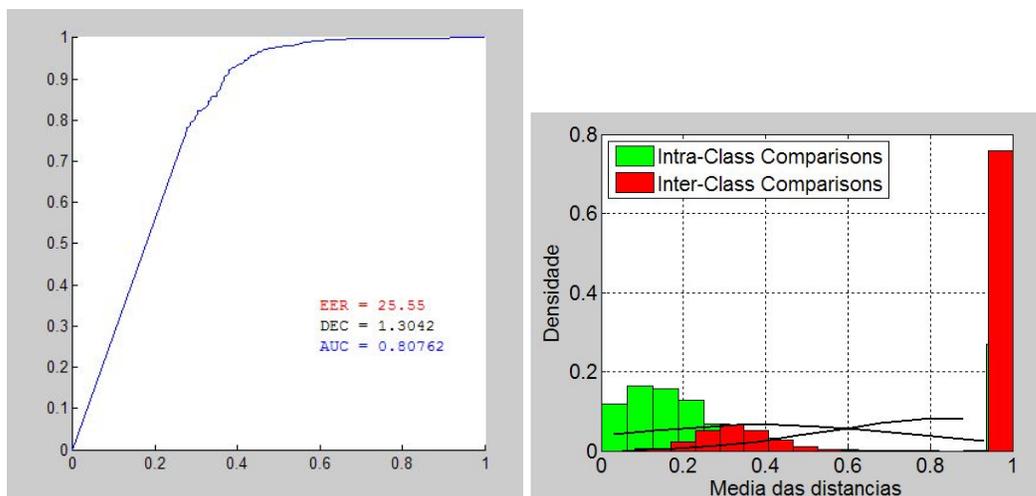
**Figura A.57:** Avaliação do Classificador na Configuração 6 para 10 keypoints (cardinalidade = 10), com máscara, peakthreshold = 0.01 e escala  $\geq 2$



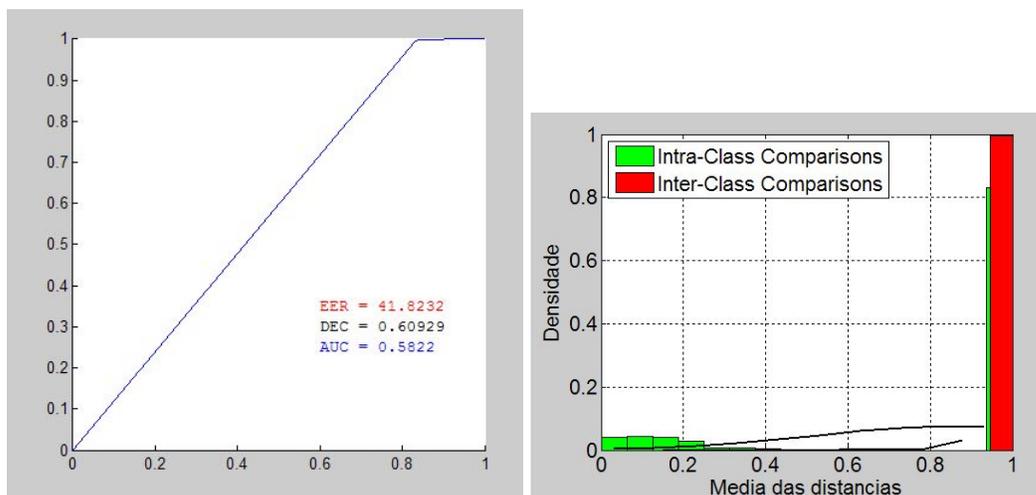
**Figura A.58:** Avaliação do Classificador na Configuração 6 para 5 keypoints (cardinalidade = 5), com máscara, peakthreshold = 0.01 e escala  $\geq 2$



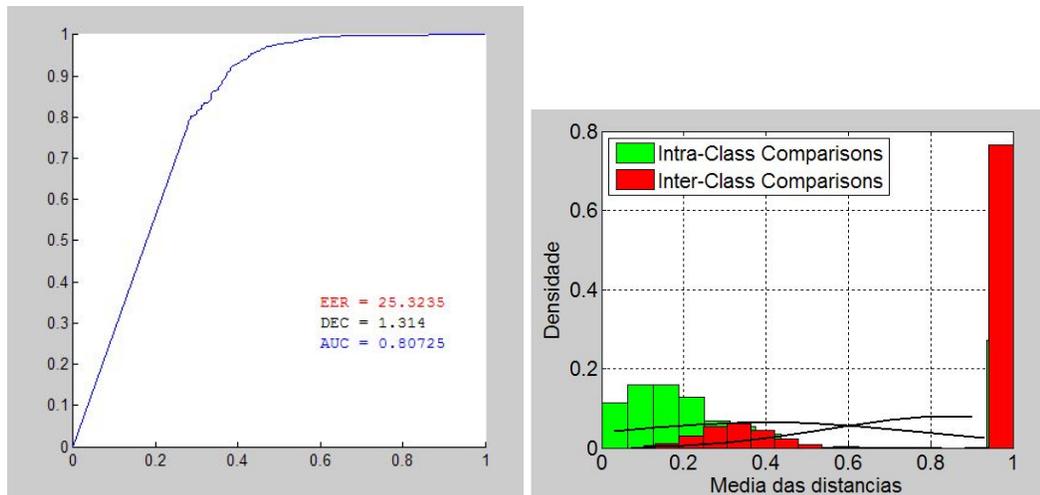
**Figura A.59:** Avaliação do Classificador na Configuração 7 para 10 keypoints (cardinalidade = 10), com máscara, peakthreshold = 0.02



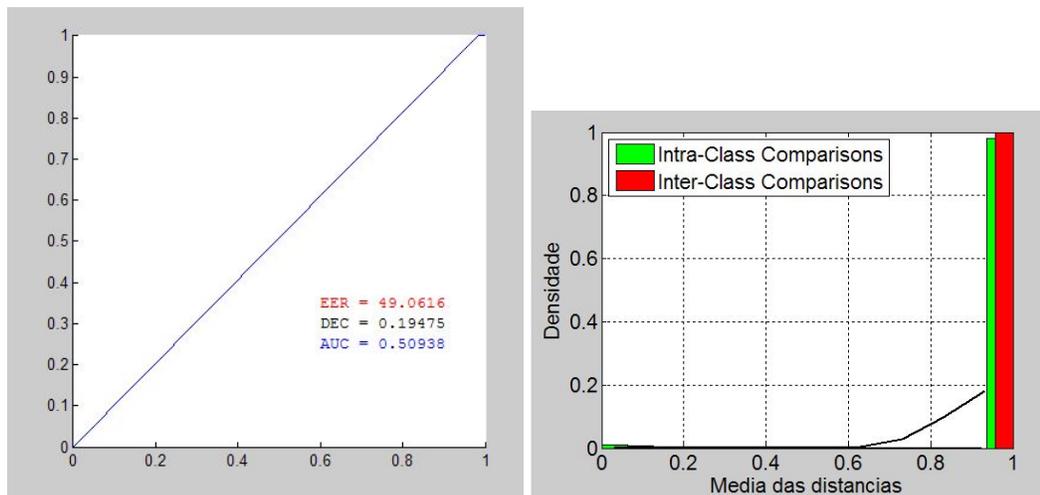
**Figura A.60:** Avaliação do Classificador na Configuração 7 para 5 keypoints (cardinalidade = 5), com máscara,  $peakthreshold = 0.02$



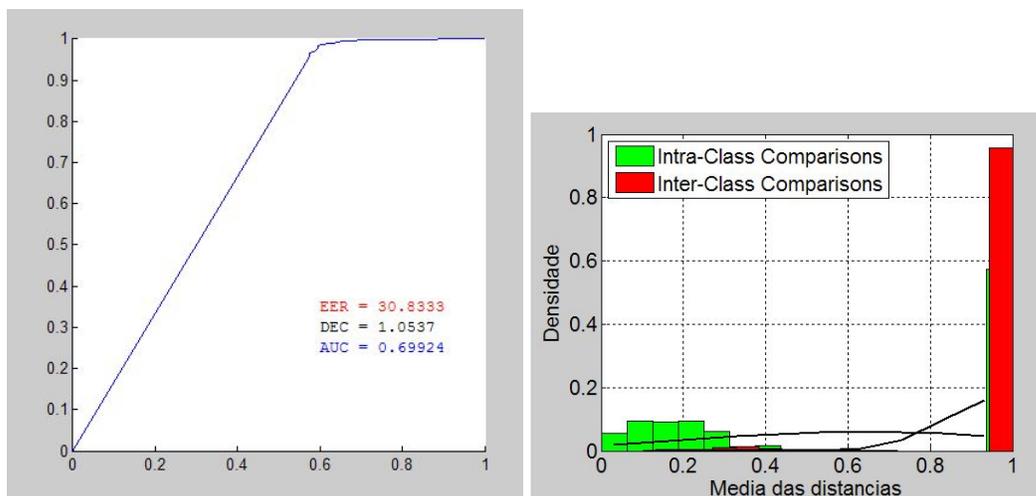
**Figura A.61:** Avaliação do Classificador na Configuração 8 para 10 keypoints (cardinalidade = 10), com máscara,  $peakthreshold = 0.02$  e escala  $\geq 2$



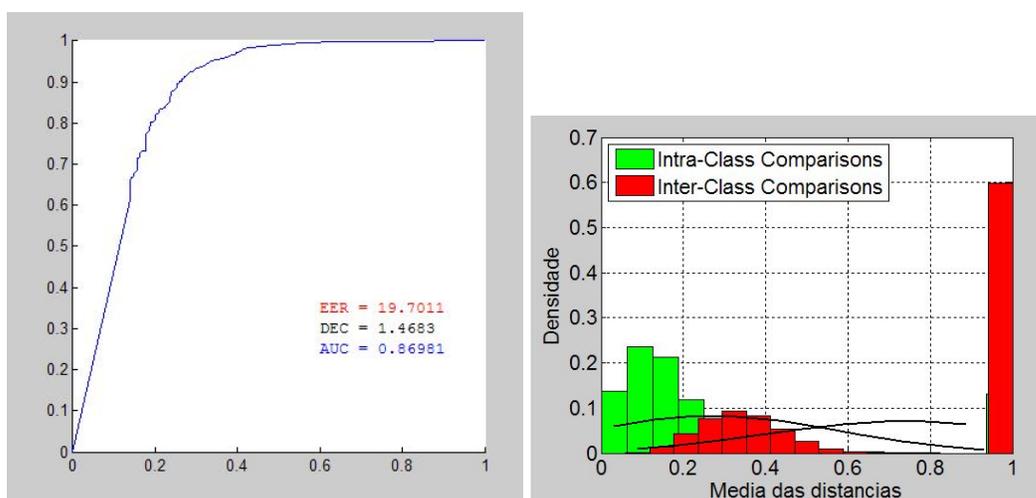
**Figura A.62:** Avaliação do Classificador na Configuração 8 para 5 keypoints (cardinalidade = 5), com máscara,  $peakthreshold = 0.02$  e escala  $\geq 2$



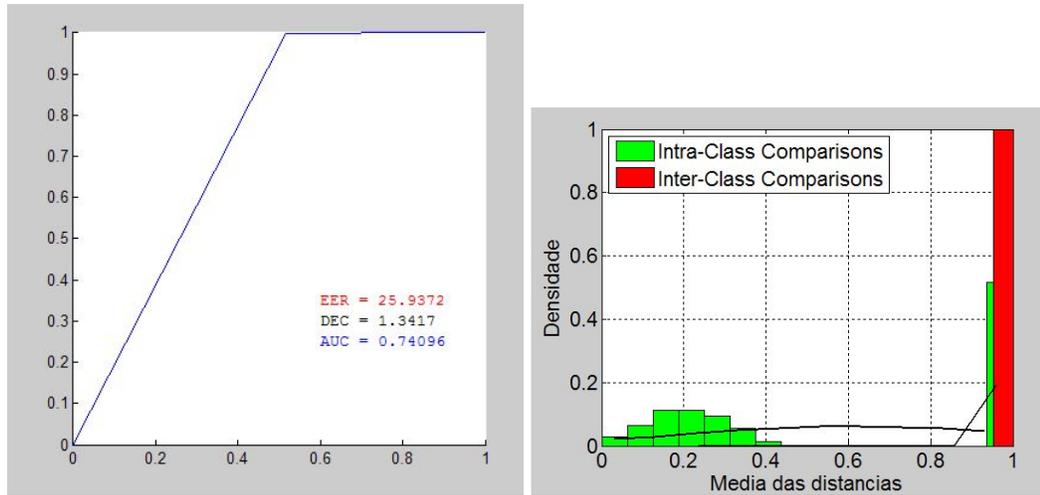
**Figura A.63:** Avaliação do Classificador na Configuração 9 para 20 keypoints (cardinalidade = 20),  $peakthreshold = 0.02$



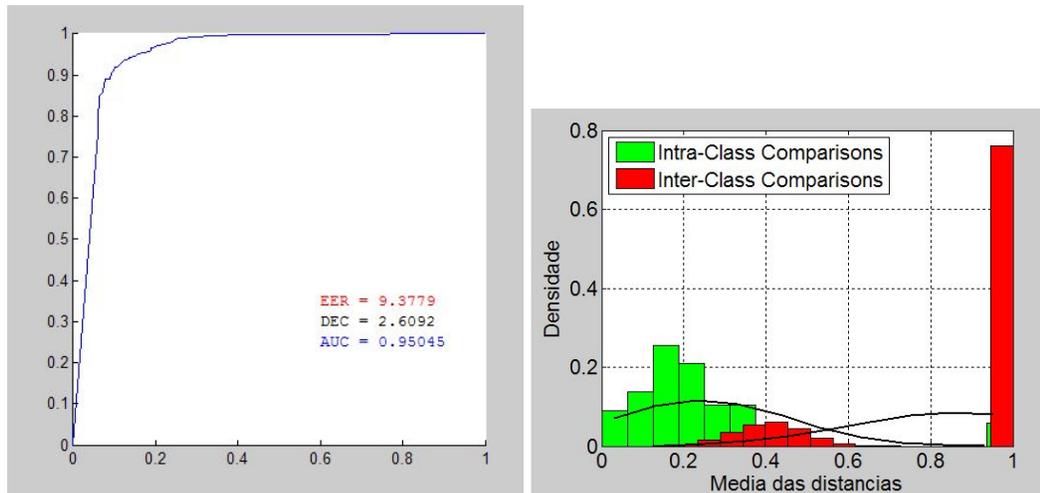
**Figura A.64:** Avaliação do Classificador na Configuração 9 para 10 keypoints (cardinalidade = 10),  $peakthreshold = 0.02$



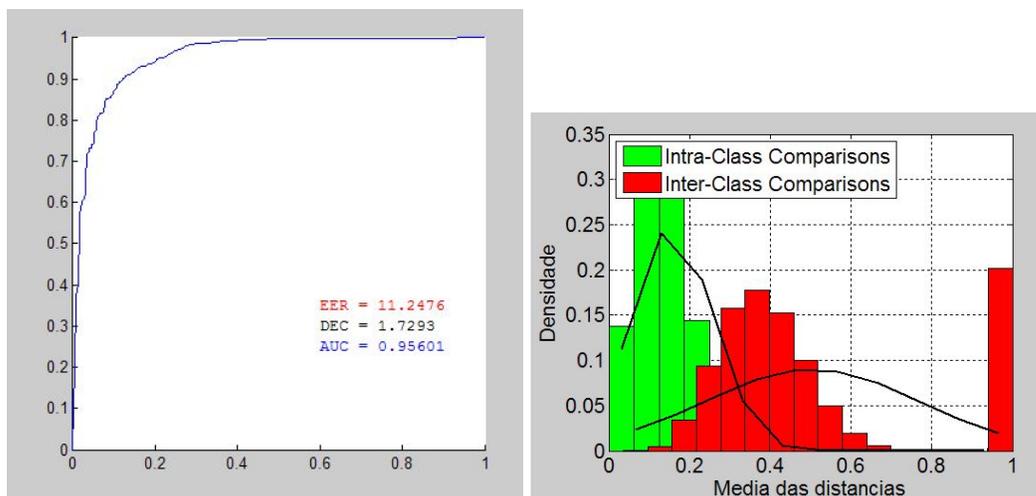
**Figura A.65:** Avaliação do Classificador na Configuração 9 para 5 keypoints (cardinalidade = 5),  $peakthreshold = 0.02$



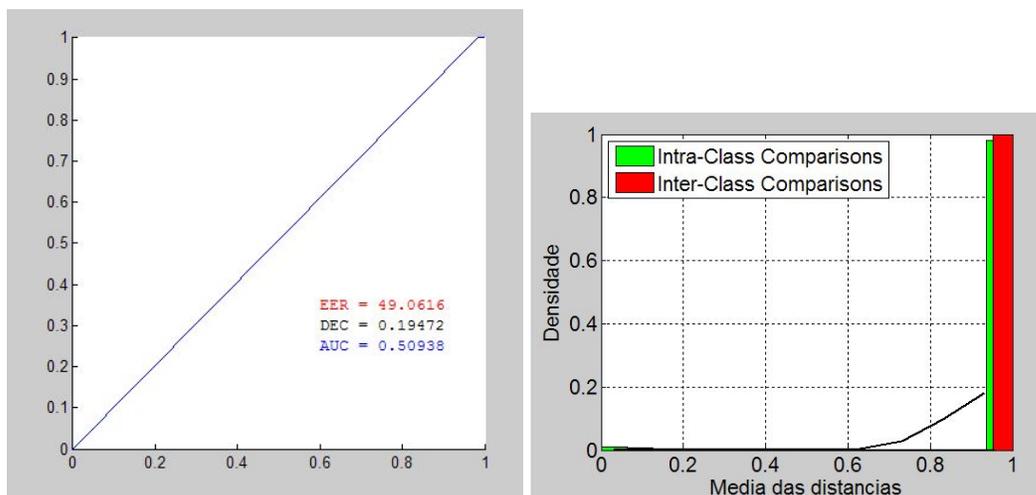
**Figura A.66:** Avaliação do Classificador na Configuração 10 para 20 keypoints (cardinalidade = 20), orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



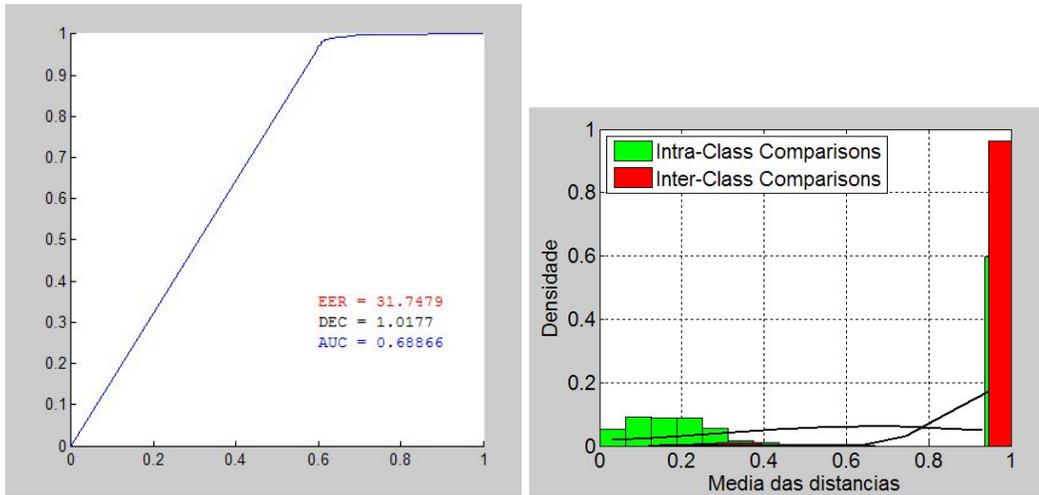
**Figura A.67:** Avaliação do Classificador na Configuração 10 para 10 keypoints (cardinalidade = 10), orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



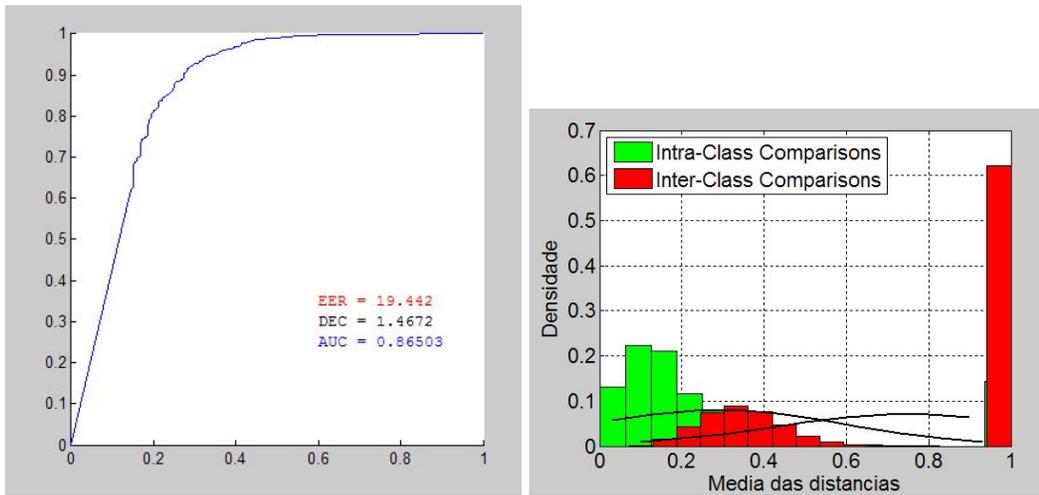
**Figura A.68:** Avaliação do Classificador na Configuração 10 para 5 keypoints (cardinalidade = 5), orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



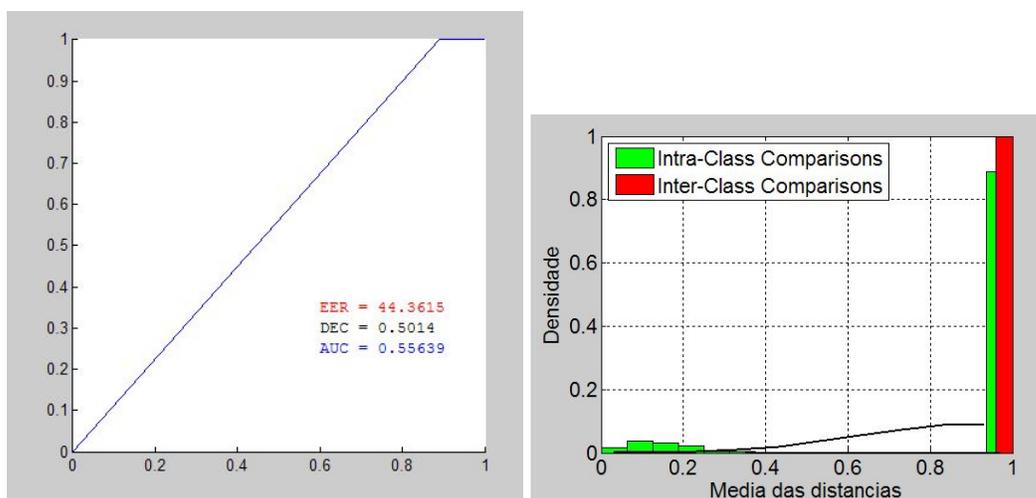
**Figura A.69:** Avaliação do Classificador na Configuração 11 para 20 keypoints (cardinalidade = 20),  $peakthreshold = 0.02$  e  $escala \geq 2$



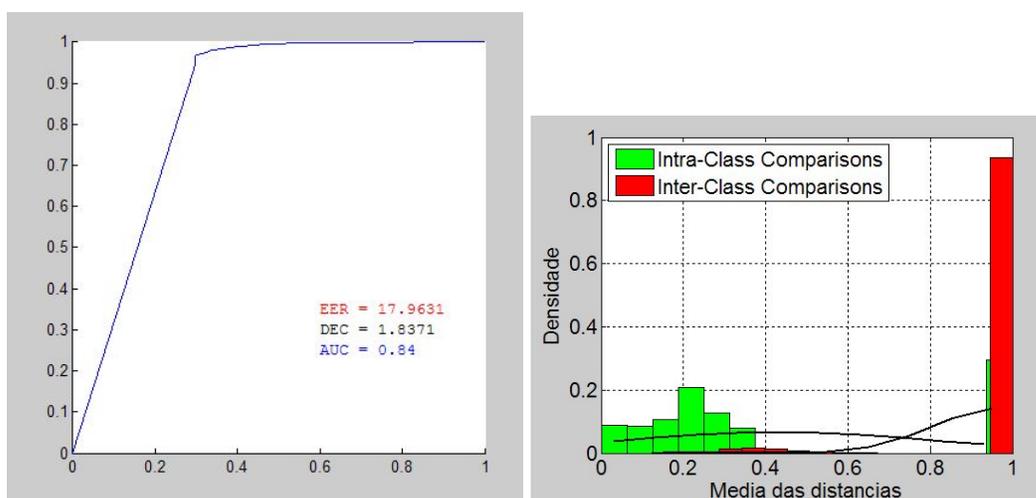
**Figura A.70:** Avaliação do Classificador na Configuração 11 para 10 keypoints (cardinalidade = 10), peakthreshold = 0.02 e escala  $\geq 2$



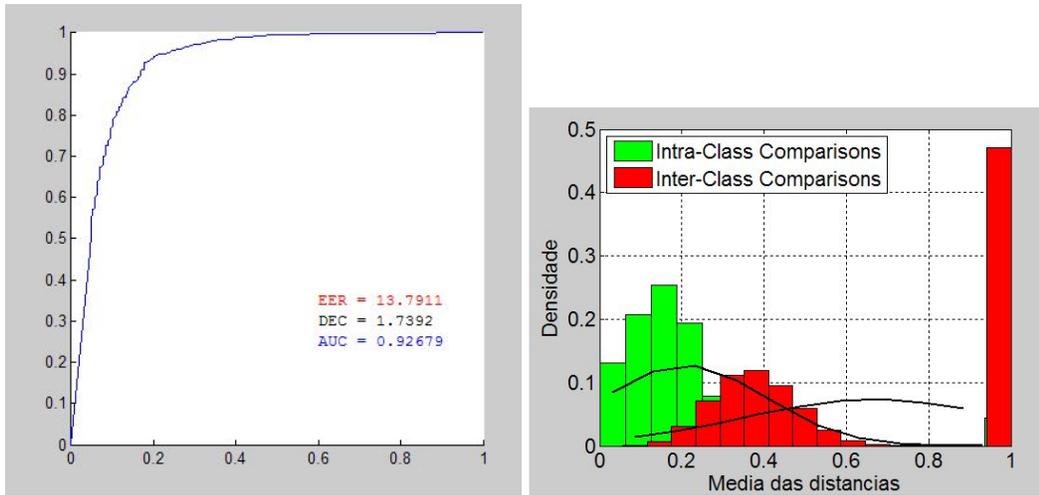
**Figura A.71:** Avaliação do Classificador na Configuração 11 para 5 keypoints (cardinalidade = 5), peakthreshold = 0.02 e escala  $\geq 2$



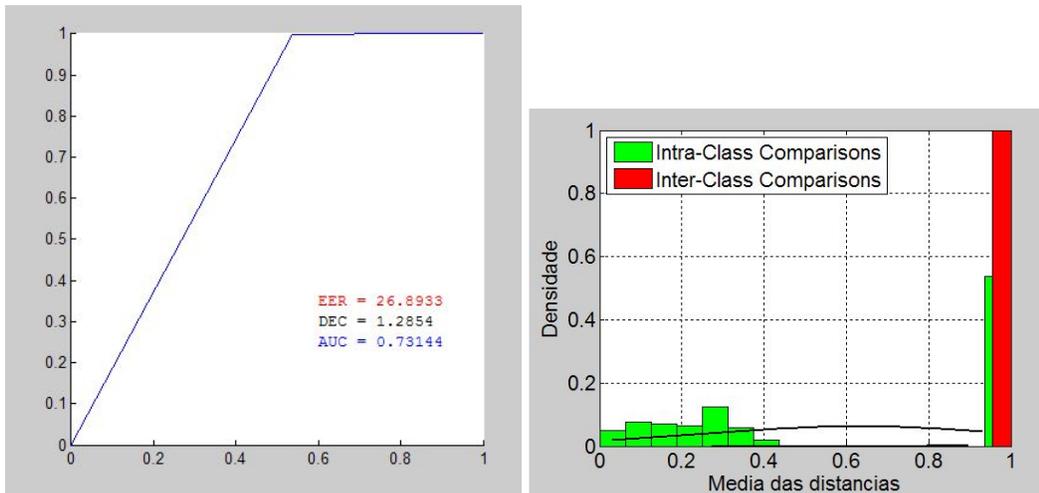
**Figura A.72:** Avaliação do Classificador na Configuração 12 para 20 keypoints (cardinalidade = 20), com máscara, orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



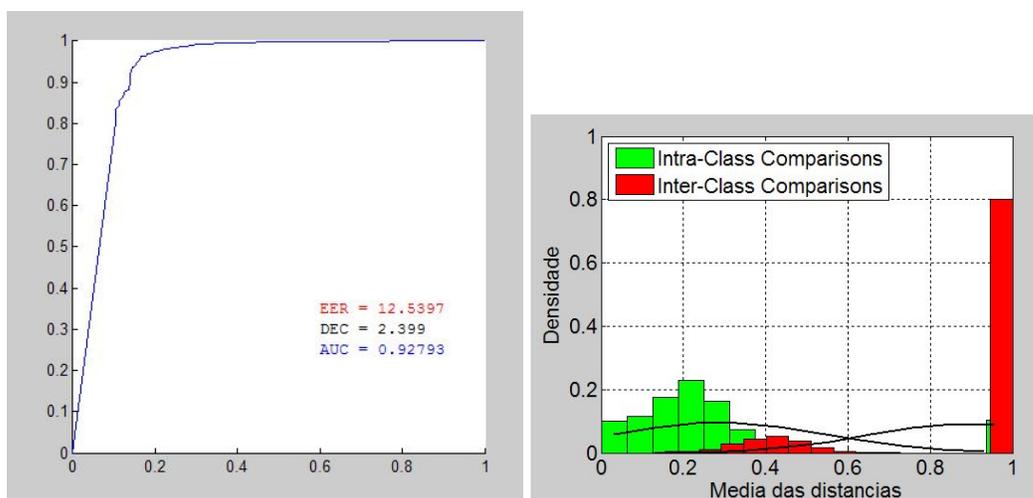
**Figura A.73:** Avaliação do Classificador na Configuração 12 para 10 keypoints (cardinalidade = 10), com máscara, orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



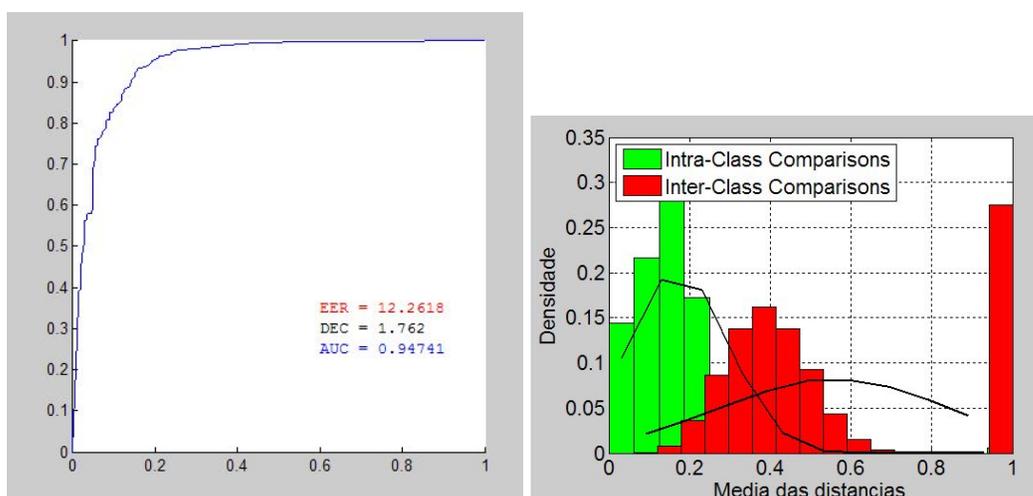
**Figura A.74:** Avaliação do Classificador na Configuração 12 para 5 keypoints (cardinalidade = 5), com máscara, orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



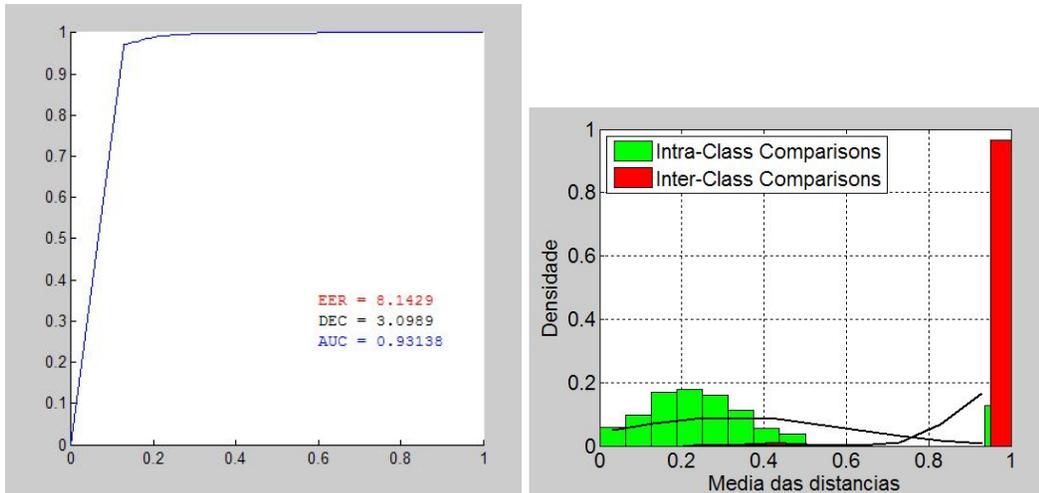
**Figura A.75:** Avaliação do Classificador na Configuração 13 para 20 keypoints (cardinalidade = 20), com máscara, orientação de keypoints  $-3\pi/2 < \theta < \pi/2$



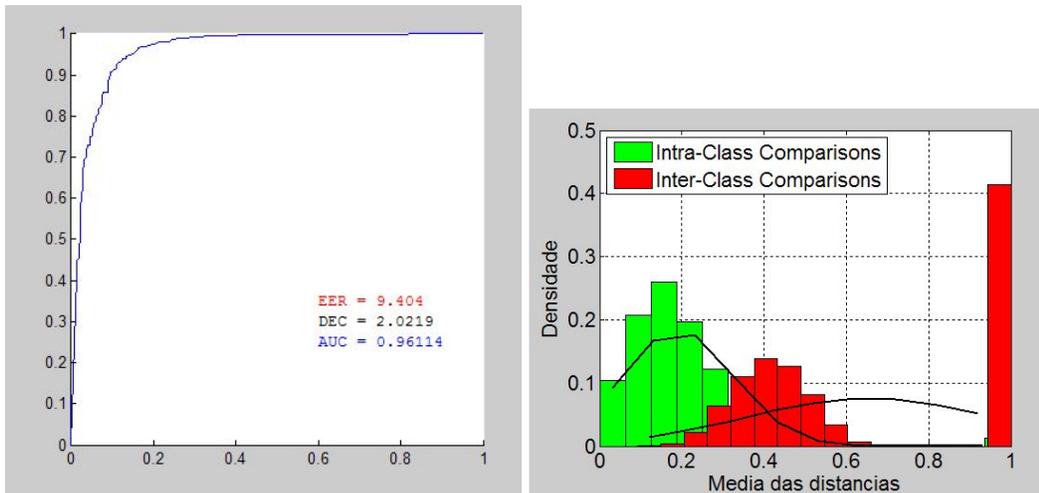
**Figura A.76:** Avaliação do Classificador na Configuração 13 para 10 keypoints (cardinalidade = 10), com máscara, orientação de keypoints  $-3\pi/2 < \theta < \pi/2$



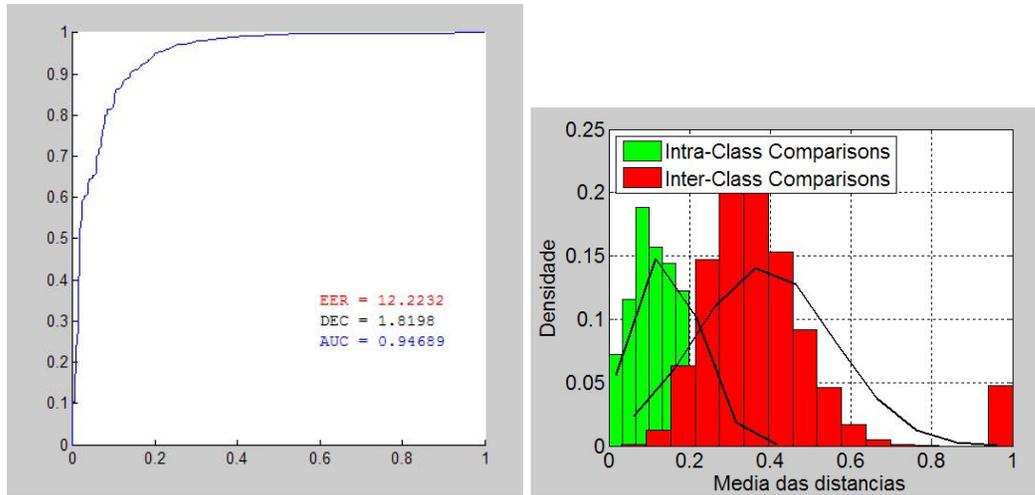
**Figura A.77:** Avaliação do Classificador na Configuração 13 para 5 keypoints (cardinalidade = 5), com máscara, orientação de keypoints  $-3\pi/2 < \theta < \pi/2$



**Figura A.78:** Avaliação do Classificador na Configuração 14 para 20 keypoints (cardinalidade = 20), orientação de keypoints  $-3\pi/2 < \theta < \pi/2$



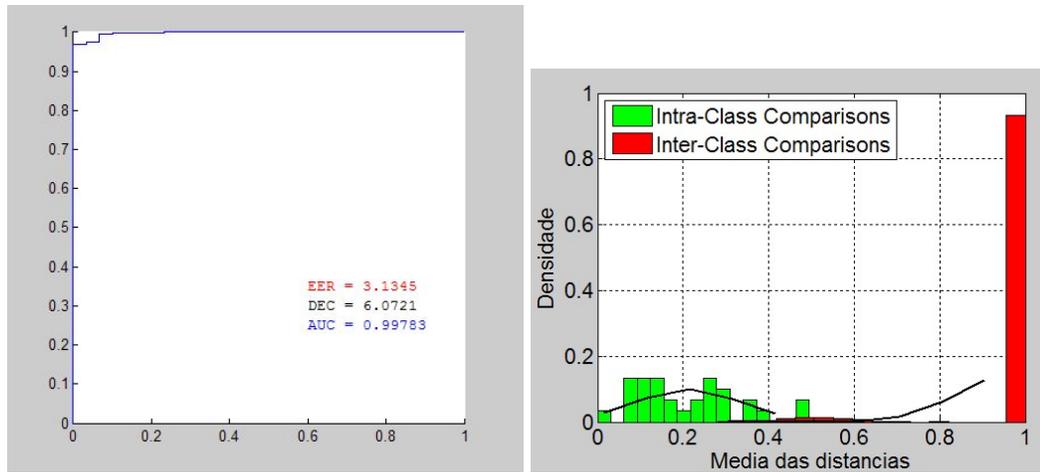
**Figura A.79:** Avaliação do Classificador na Configuração 14 para 10 keypoints (cardinalidade = 10), orientação de keypoints  $-3\pi/2 < \theta < \pi/2$



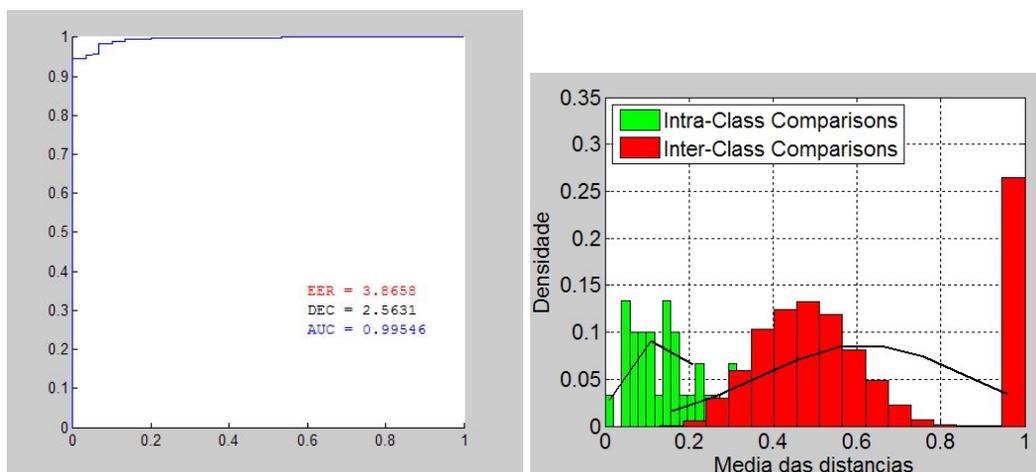
**Figura A.80:** Avaliação do Classificador na Configuração 14 para 5 keypoints (cardinalidade = 5), orientação de keypoints  $-3\pi/2 < \theta < \pi/2$

## A.3 Banco de Imagens C

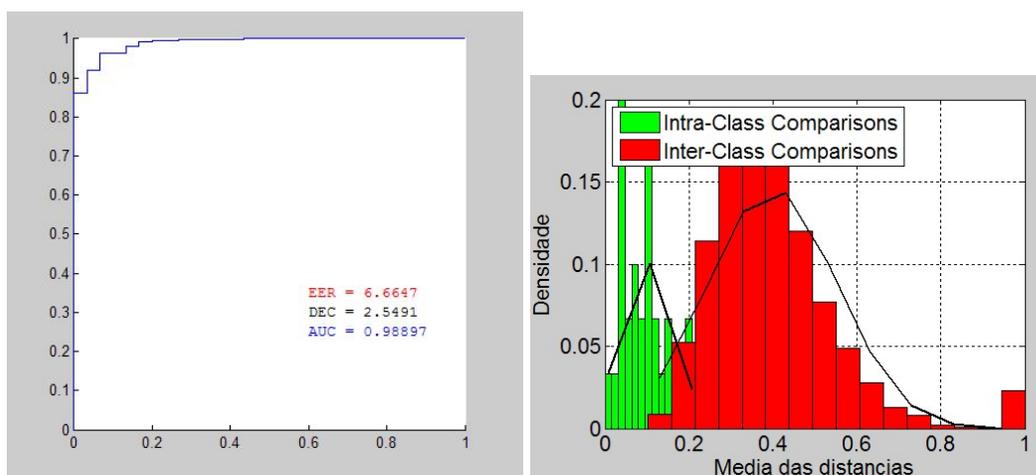
Resultado das avaliações feitas ao conjunto de imagens, com pouco ou quase nenhuma rotação.



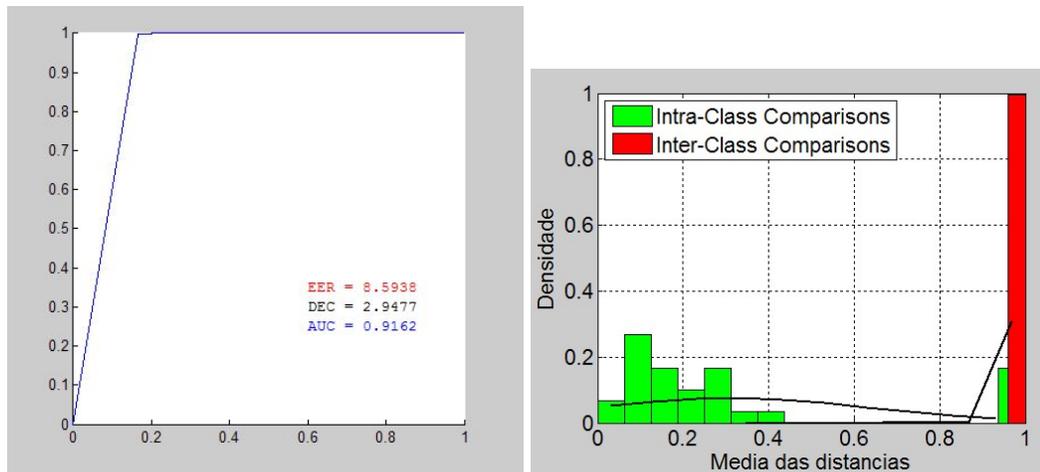
**Figura A.81:** Avaliação do Classificador na Configuração 1 para 20 keypoints (cardinalidade = 20).



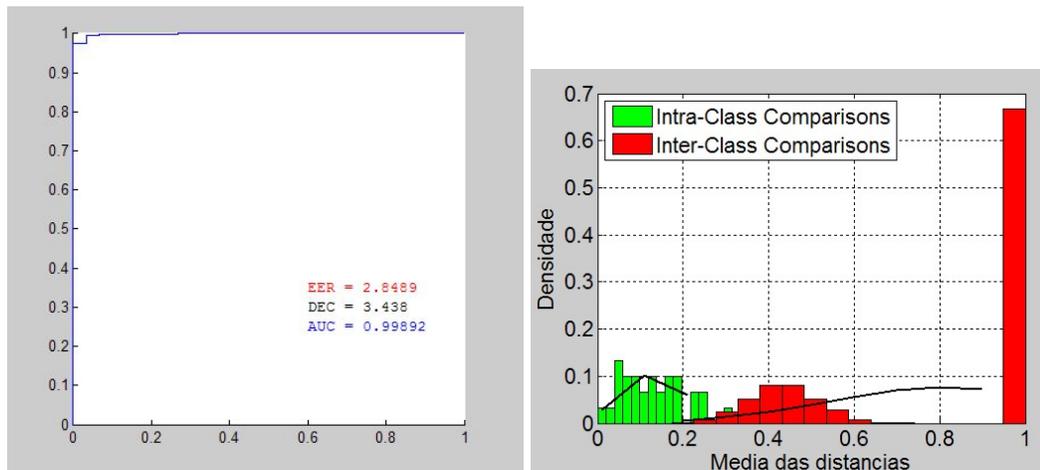
**Figura A.82:** Avaliação do Classificador na Configuração 1 para 10 keypoints (cardinalidade = 10).



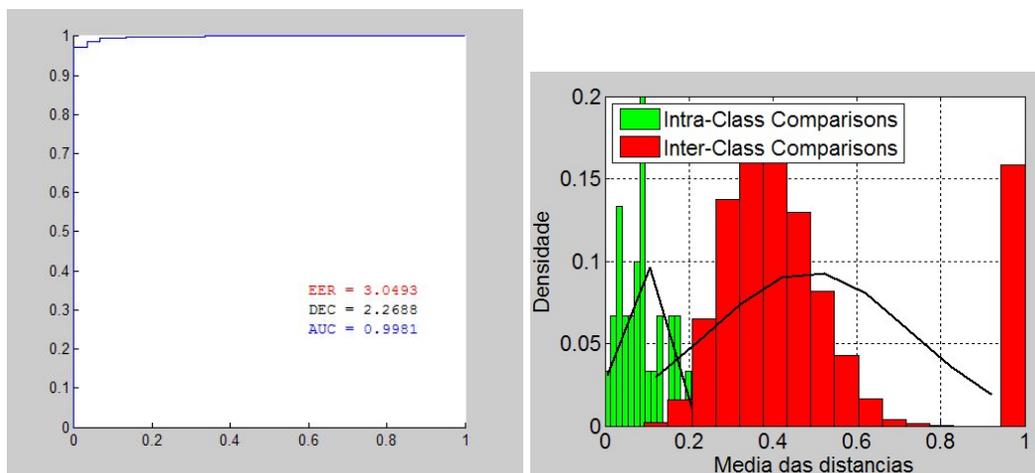
**Figura A.83:** Avaliação do Classificador na Configuração 1 para 5 keypoints (cardinalidade = 5).



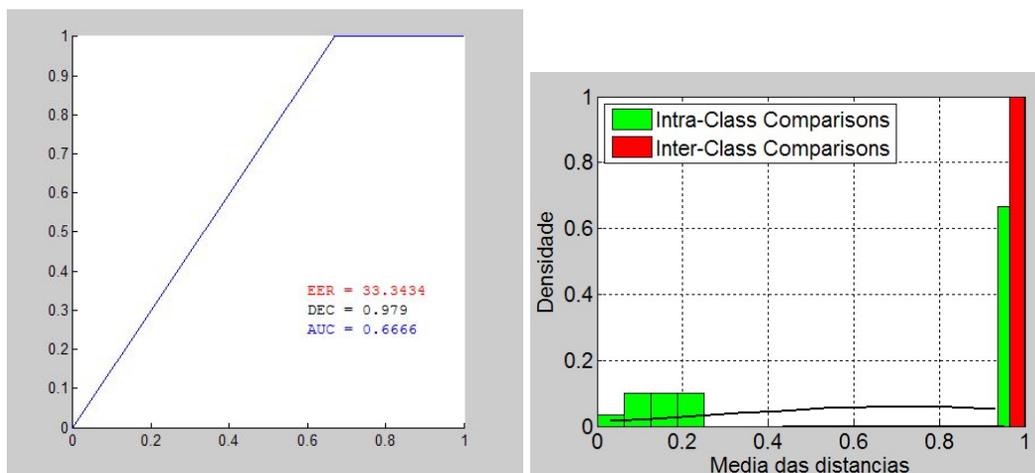
**Figura A.84:** Avaliação do Classificador na Configuração 2 para 20 keypoints (cardinalidade = 20), com máscara



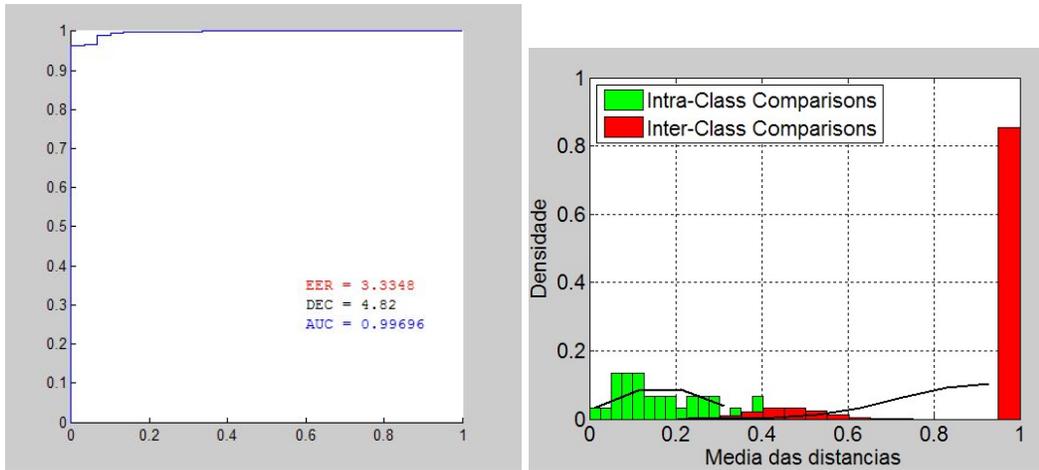
**Figura A.85:** Avaliação do Classificador na Configuração 2 para 10 keypoints (cardinalidade = 10), com máscara



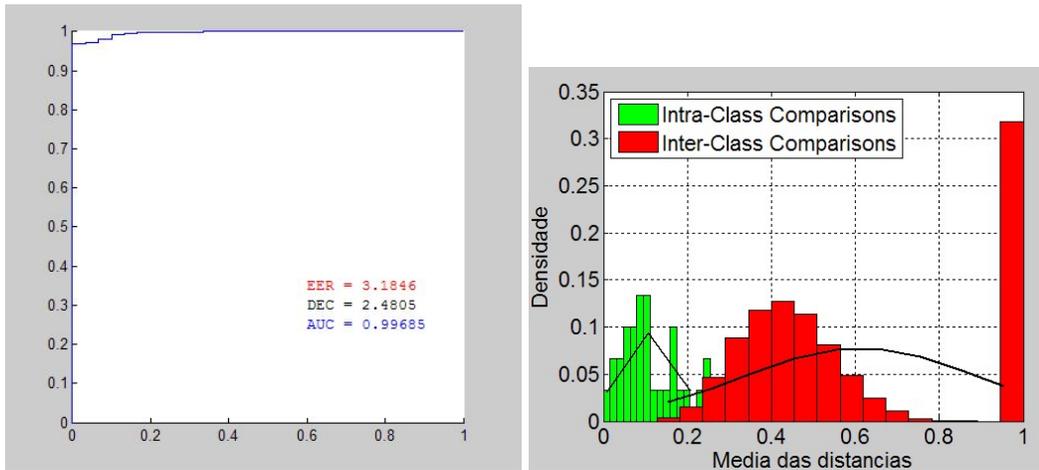
**Figura A.86:** Avaliação do Classificador na Configuração 2 para 5 keypoints (cardinalidade = 5), com máscara



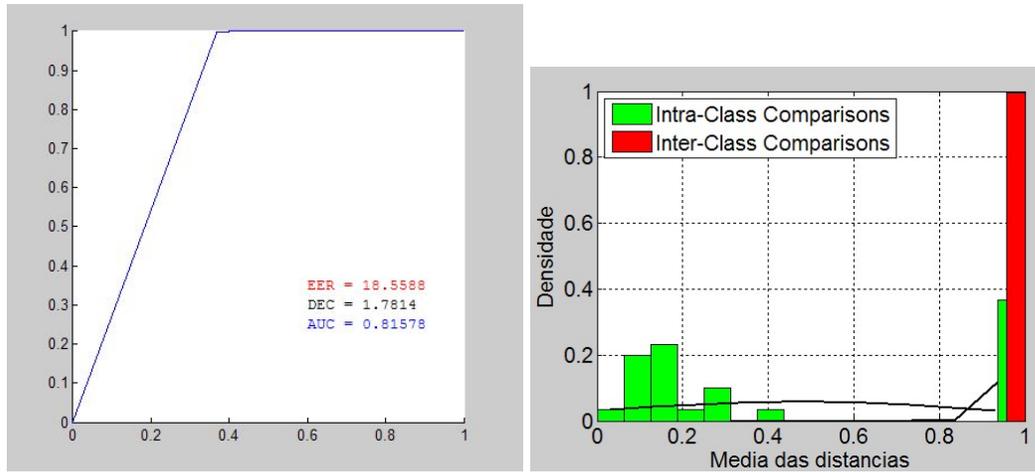
**Figura A.87:** Avaliação do Classificador na Configuração 3 para 20 keypoints (cardinalidade = 20), com máscara, peakthreshold = 0.01



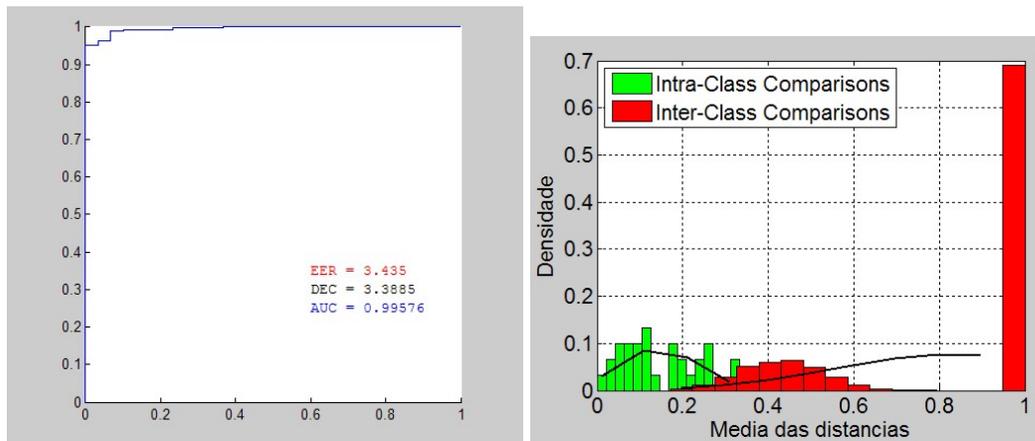
**Figura A.88:** Avaliação do Classificador na Configuração 3 para 10 keypoints (cardinalidade = 10), com máscara, peakthreshold = 0.01



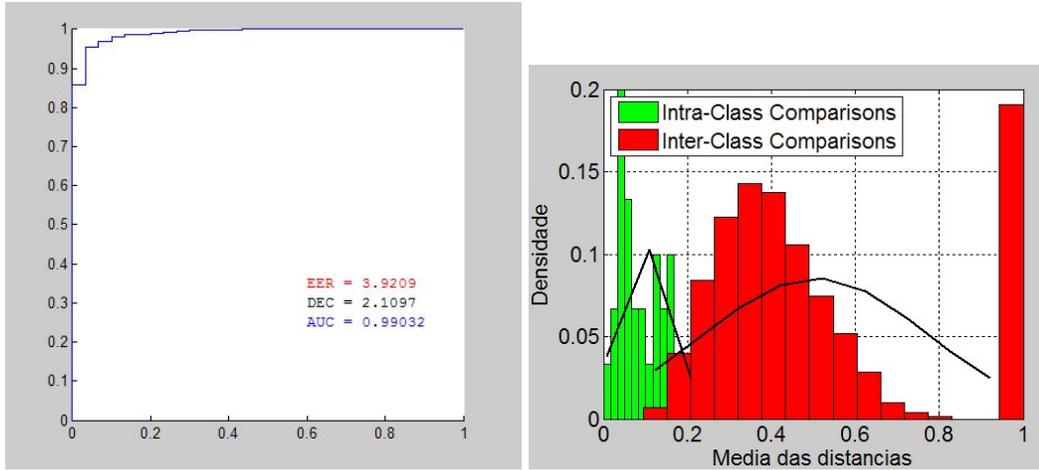
**Figura A.89:** Avaliação do Classificador na Configuração 3 para 5 keypoints (cardinalidade = 5), com máscara, peakthreshold = 0.01



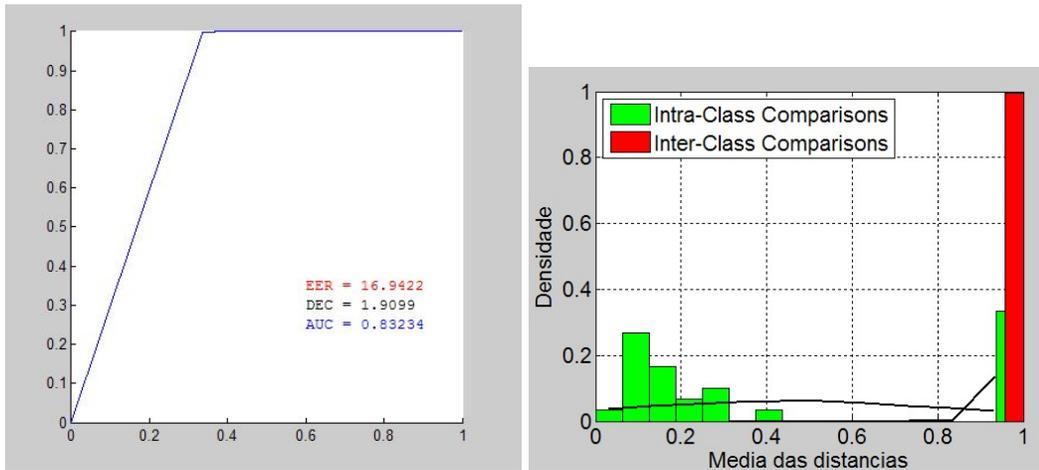
**Figura A.90:** Avaliação do Classificador na Configuração 4 para 20 keypoints (cardinalidade = 20),  $peakthreshold = 0.01$  e escala > 2



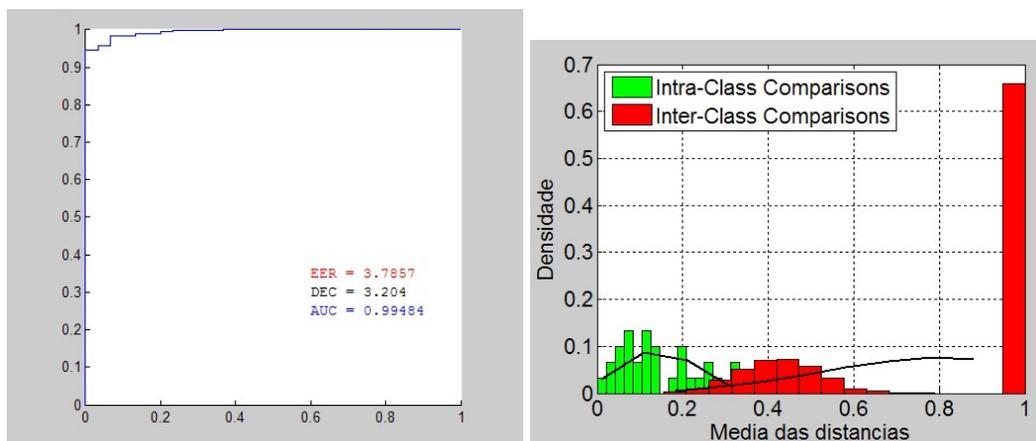
**Figura A.91:** Avaliação do Classificador na Configuração 4 para 10 keypoints (cardinalidade = 10),  $peakthreshold = 0.01$  e escala > 2



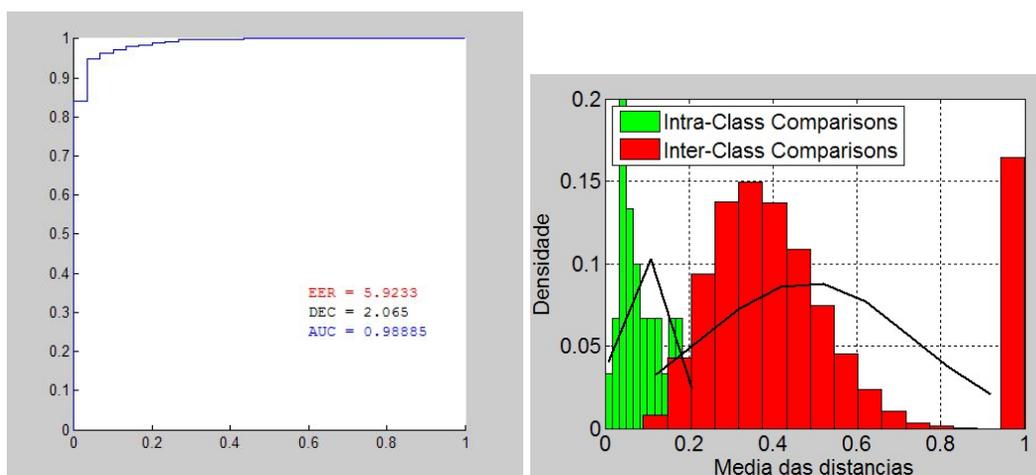
**Figura A.92:** Avaliação do Classificador na Configuração 4 para 5 keypoints (cardinalidade = 5),  $peakthreshold = 0.01$  e  $escala > 2$



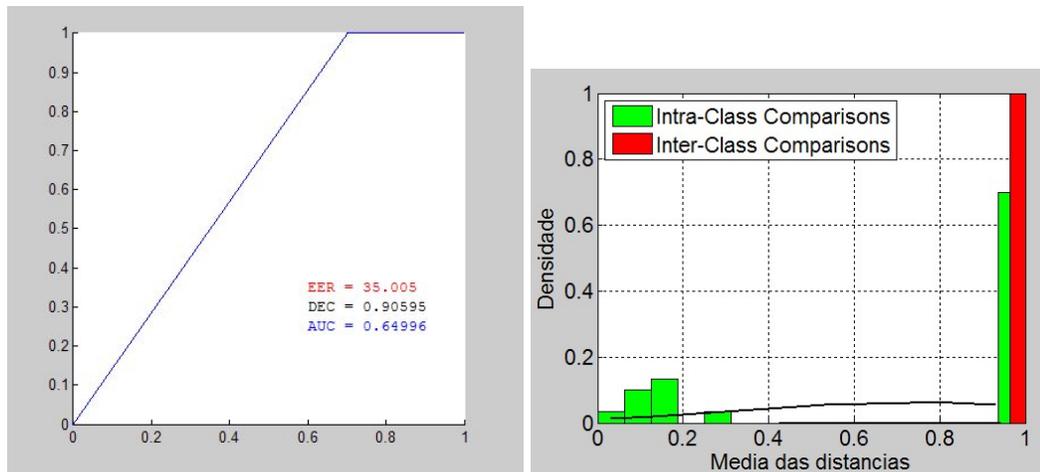
**Figura A.93:** Avaliação do Classificador na Configuração 5 para 20 keypoints (cardinalidade = 20),  $peakthreshold = 0.01$



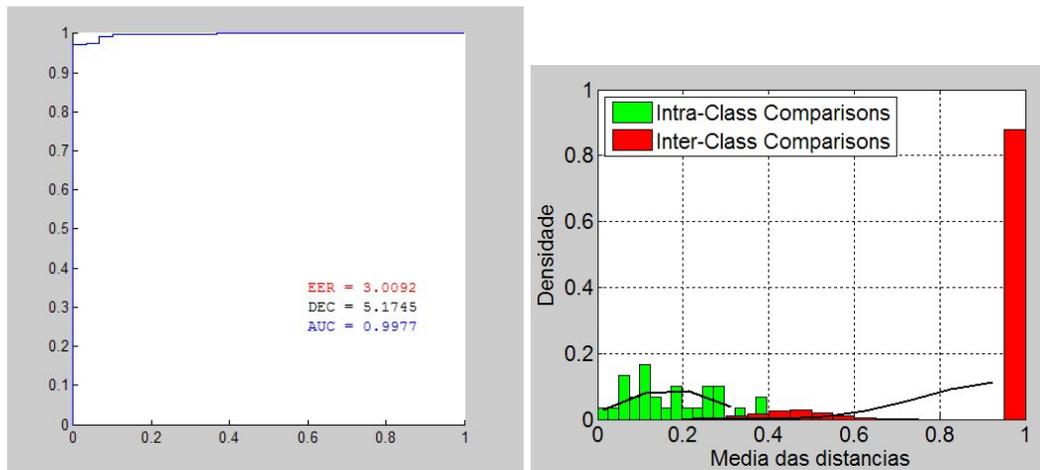
**Figura A.94:** Avaliação do Classificador na Configuração 5 para 10 keypoints (cardinalidade = 10),  $peakthreshold = 0.01$



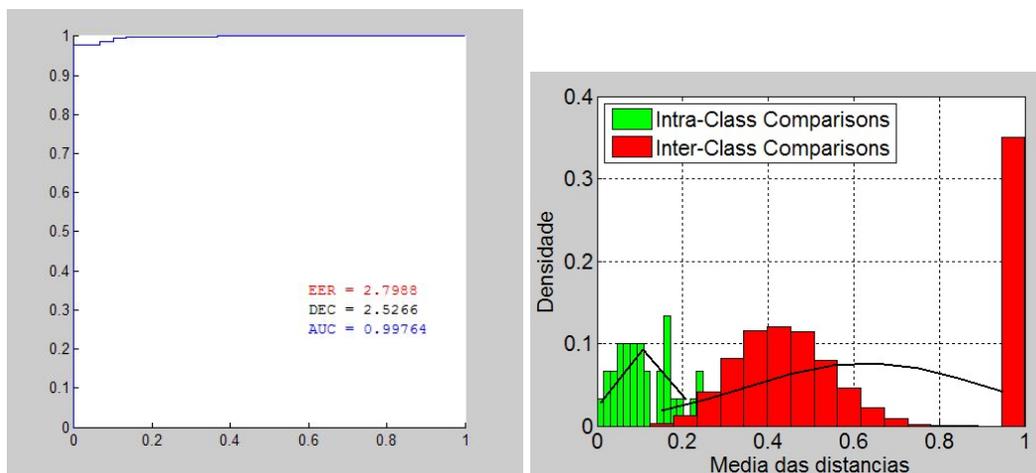
**Figura A.95:** Avaliação do Classificador na Configuração 5 para 5 keypoints (cardinalidade = 5),  $peakthreshold = 0.01$



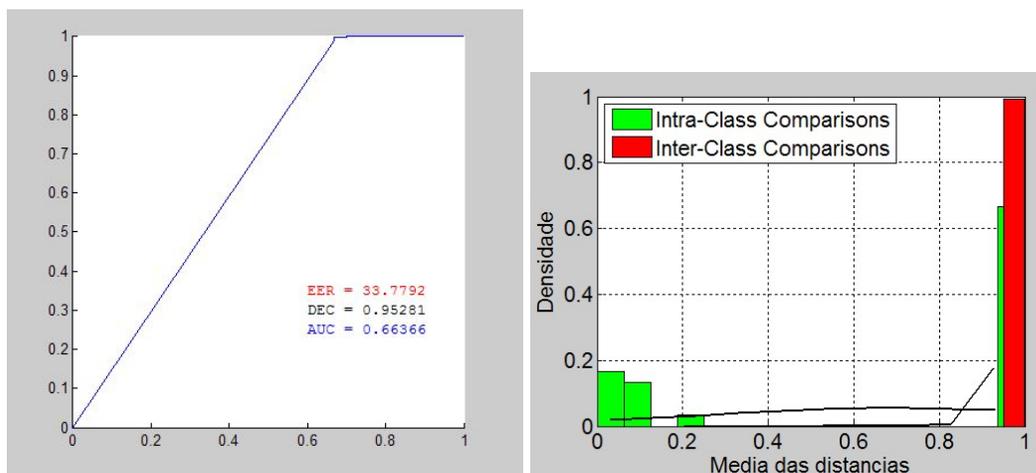
**Figura A.96:** Avaliação do Classificador na Configuração 6 para 20 keypoints (cardinalidade = 20), com máscara,  $peakthreshold = 0.01$  e escala  $\geq 2$



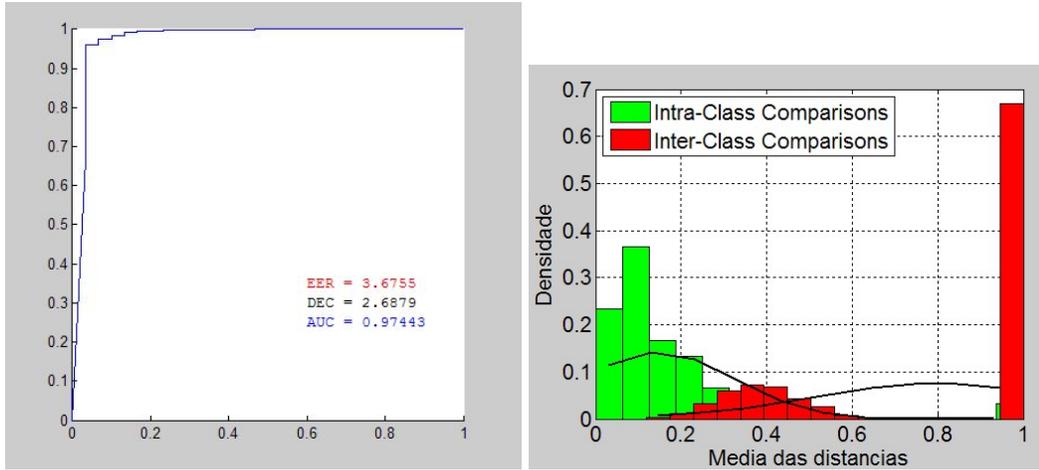
**Figura A.97:** Avaliação do Classificador na Configuração 6 para 10 keypoints (cardinalidade = 10), com máscara,  $peakthreshold = 0.01$  e escala  $\geq 2$



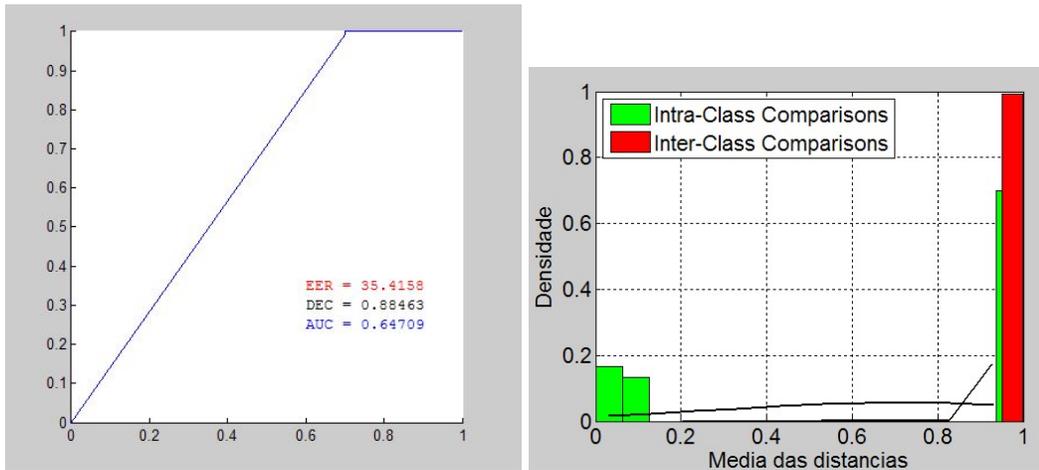
**Figura A.98:** Avaliação do Classificador na Configuração 6 para 5 keypoints (cardinalidade = 5), com máscara,  $peakthreshold = 0.01$  e escala  $\geq 2$



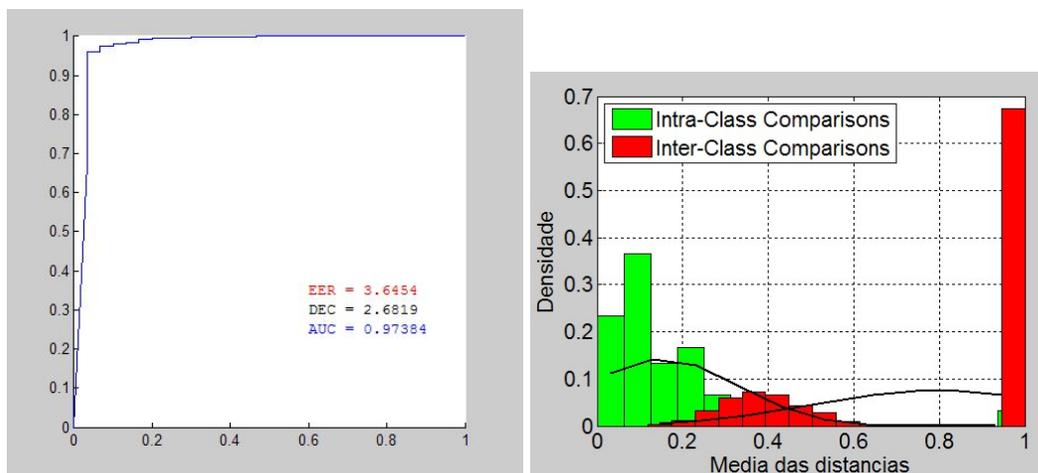
**Figura A.99:** Avaliação do Classificador na Configuração 7 para 10 keypoints (cardinalidade = 10), com máscara,  $peakthreshold = 0.02$



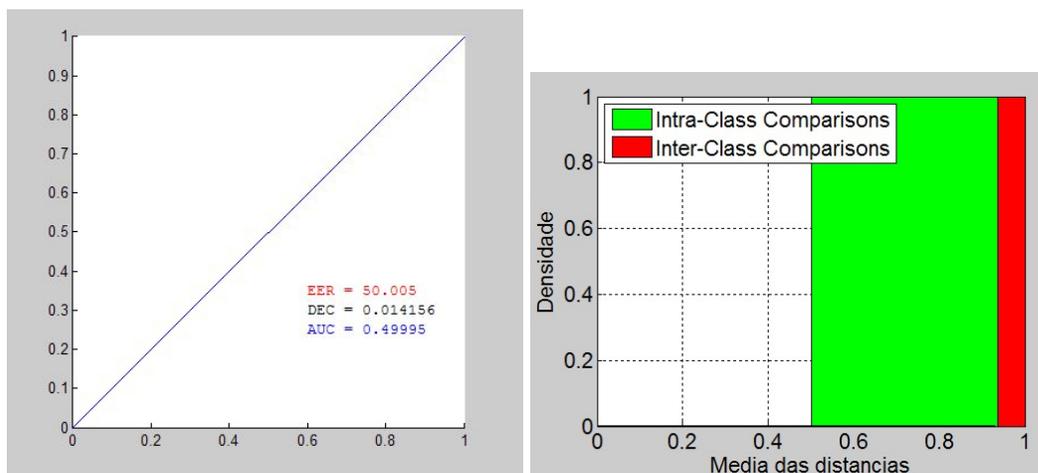
**Figura A.100:** Avaliação do Classificador na Configuração 7 para 5 keypoints (cardinalidade = 5), com máscara, peakthreshold = 0.02



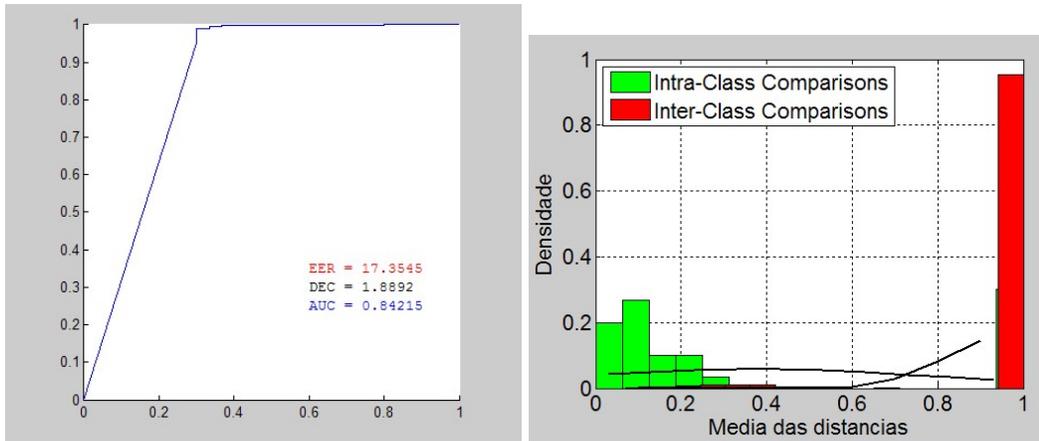
**Figura A.101:** Avaliação do Classificador na Configuração 8 para 10 keypoints (cardinalidade = 10), com máscara, peakthreshold = 0.02 e escala  $\geq 2$



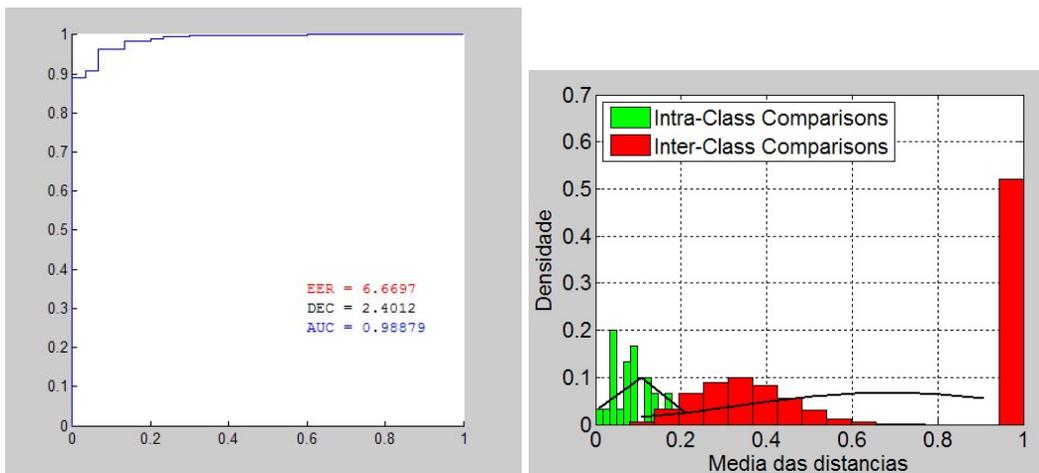
**Figura A.102:** Avaliação do Classificador na Configuração 8 para 5 keypoints (cardinalidade = 5), com máscara,  $peakthreshold = 0.02$  e escala  $\geq 2$



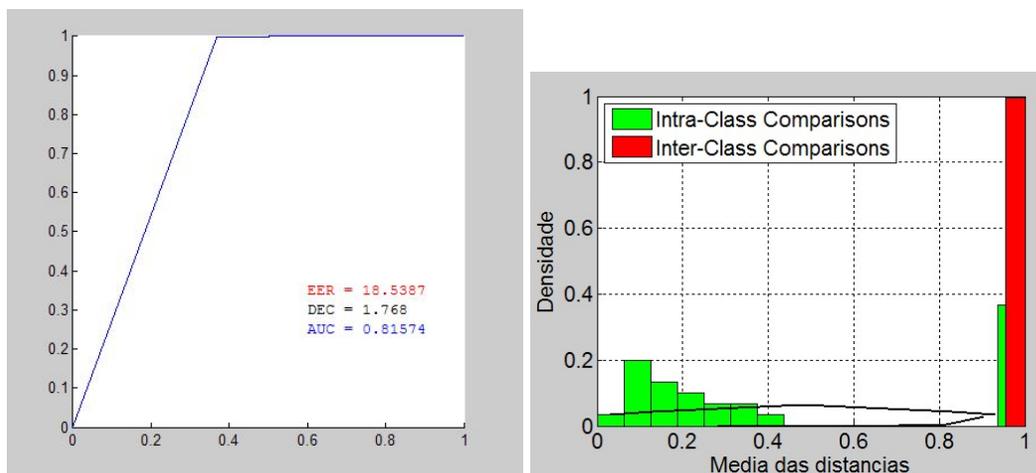
**Figura A.103:** Avaliação do Classificador na Configuração 9 para 20 keypoints (cardinalidade = 20),  $peakthreshold = 0.02$



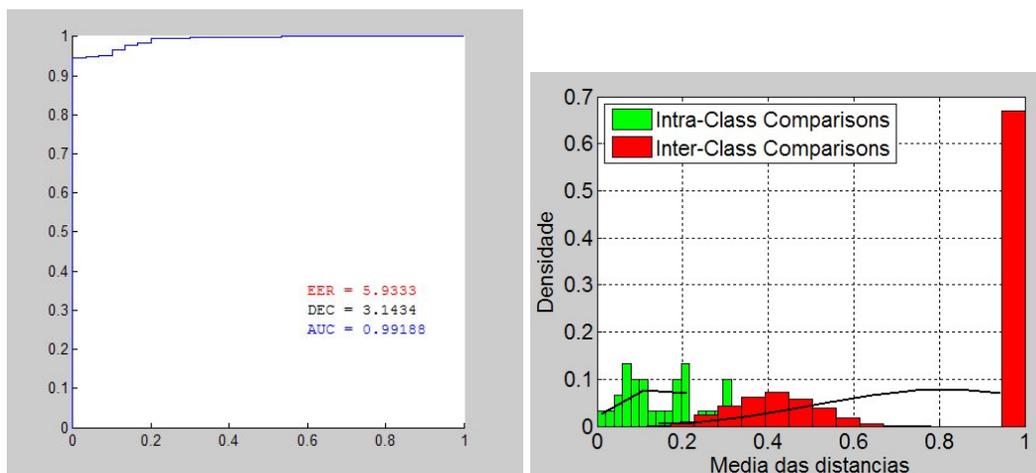
**Figura A.104:** Avaliação do Classificador na Configuração 9 para 10 keypoints (cardinalidade = 10),  $peakthreshold = 0.02$



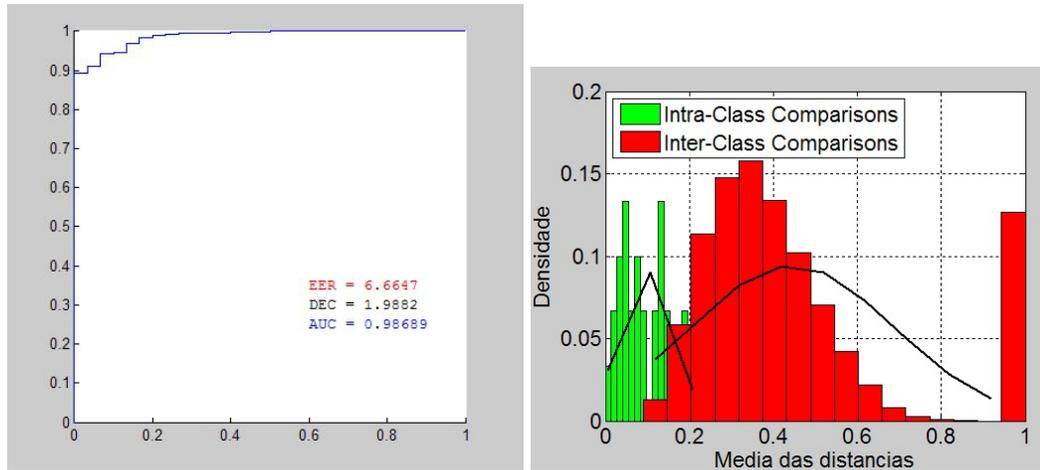
**Figura A.105:** Avaliação do Classificador na Configuração 9 para 5 keypoints (cardinalidade = 5),  $peakthreshold = 0.02$



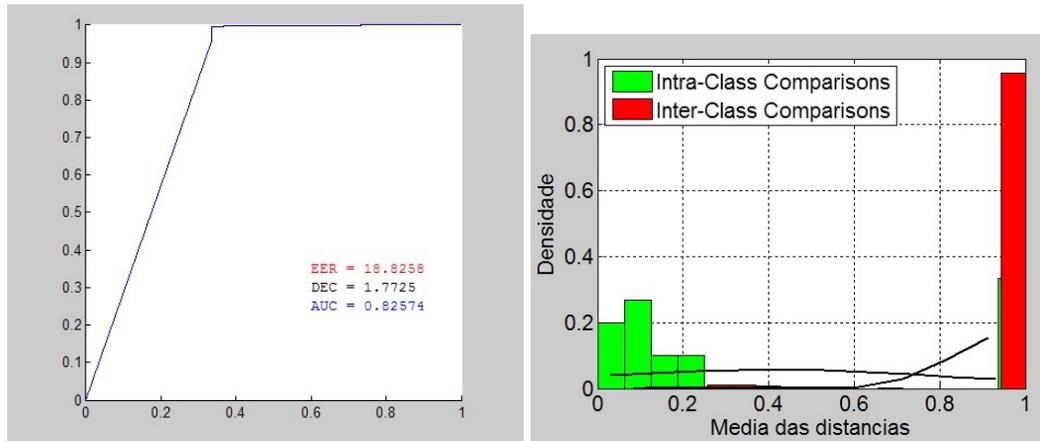
**Figura A.106:** Avaliação do Classificador na Configuração 10 para 20 keypoints (cardinalidade = 20), orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



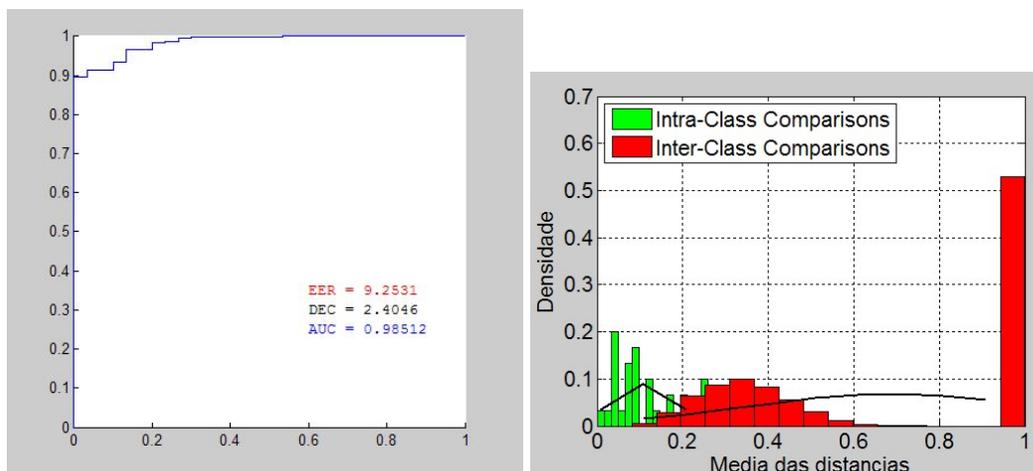
**Figura A.107:** Avaliação do Classificador na Configuração 10 para 10 keypoints (cardinalidade = 10), orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



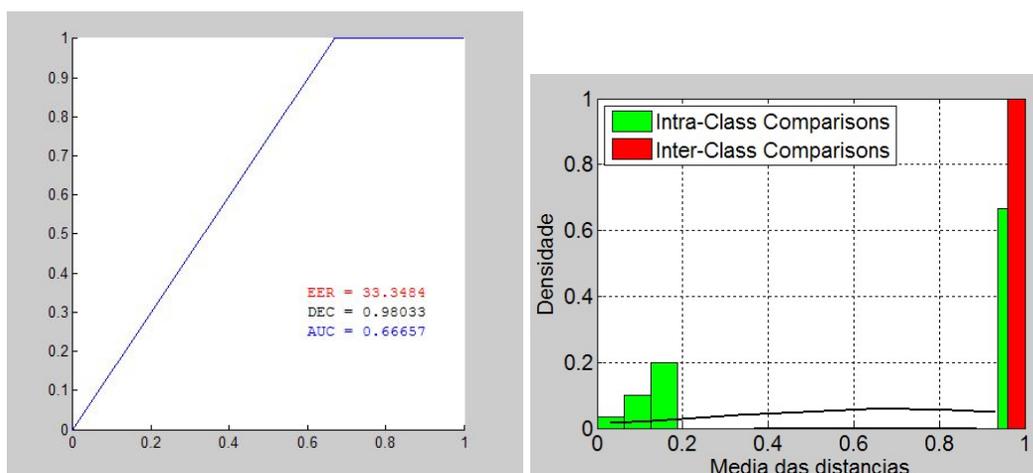
**Figura A.108:** Avaliação do Classificador na Configuração 10 para 5 keypoints (cardinalidade = 5), orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



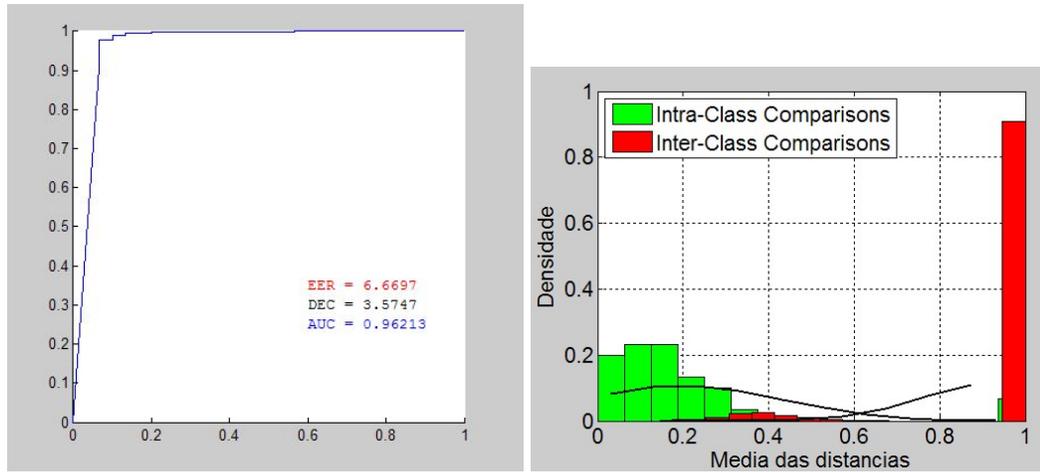
**Figura A.109:** Avaliação do Classificador na Configuração 11 para 10 keypoints (cardinalidade = 10),  $peakthreshold = 0.02$  e escala  $\geq 2$



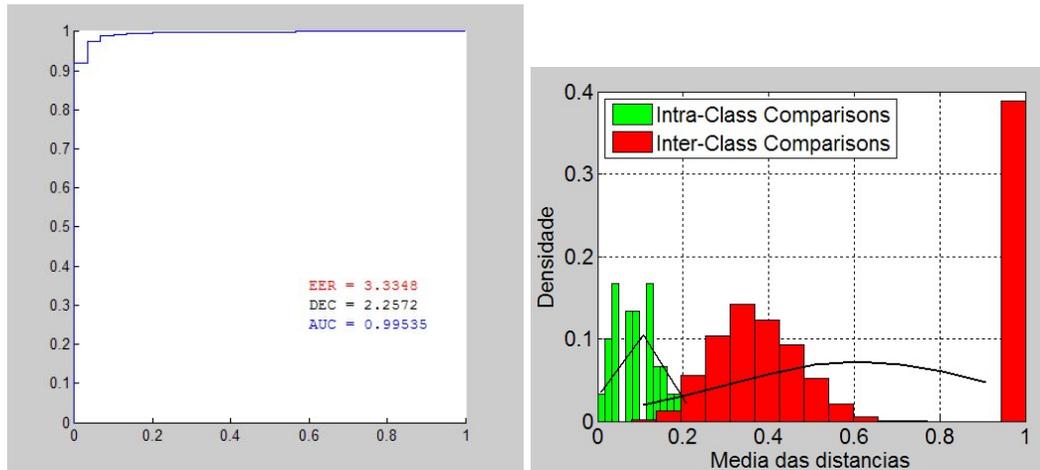
**Figura A.110:** Avaliação do Classificador na Configuração 11 para 5 keypoints (cardinalidade = 5),  $peakthreshold = 0.02$  e  $escala \geq 2$



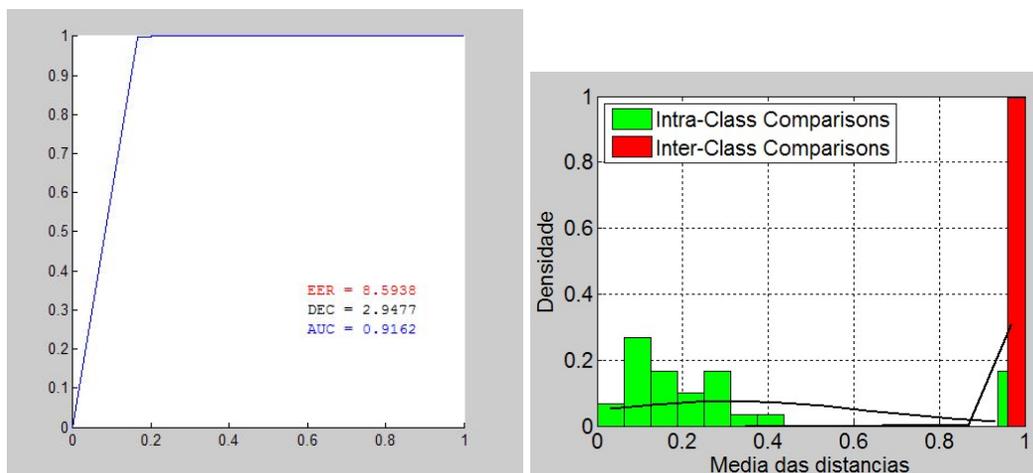
**Figura A.111:** Avaliação do Classificador na Configuração 12 para 20 keypoints (cardinalidade = 20), com máscara, orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



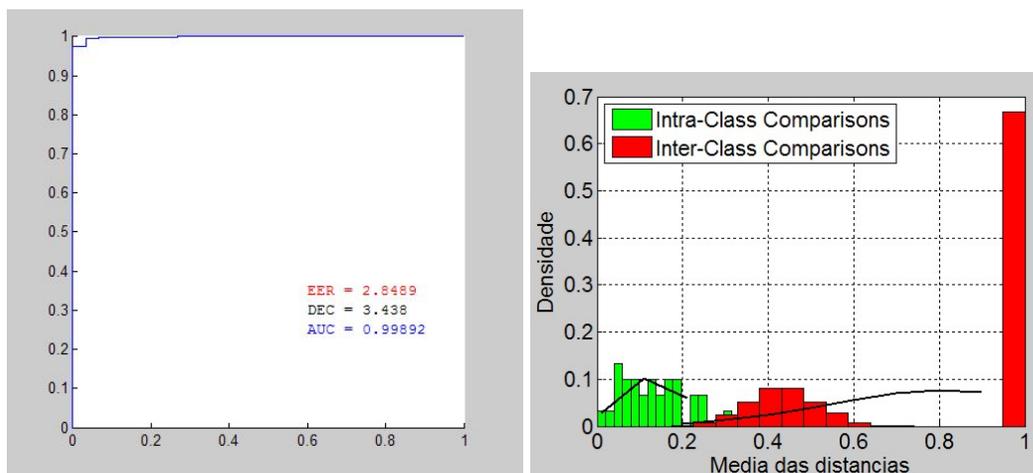
**Figura A.112:** Avaliação do Classificador na Configuração 12 para 10 keypoints (cardinalidade = 10), com máscara, orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



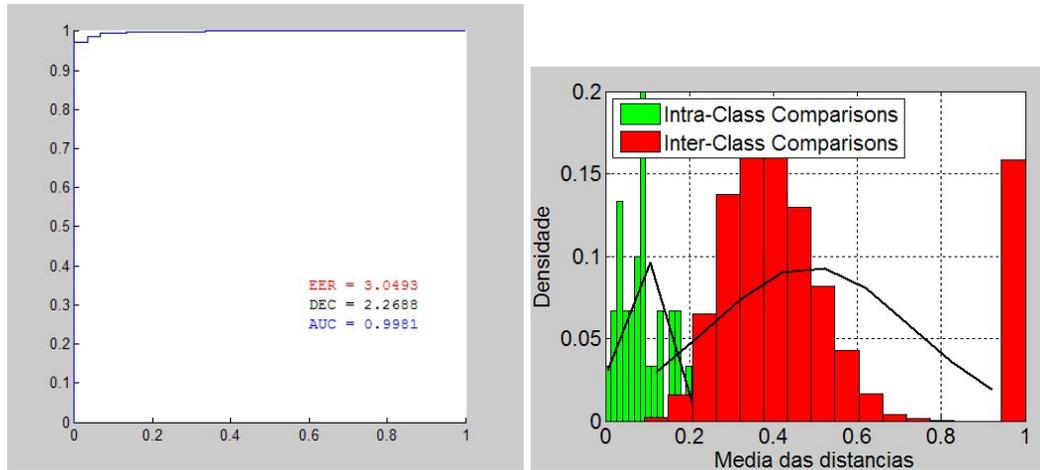
**Figura A.113:** Avaliação do Classificador na Configuração 12 para 5 keypoints (cardinalidade = 5), com máscara, orientação de keypoints  $-\pi/2 < \theta < 3\pi/2$



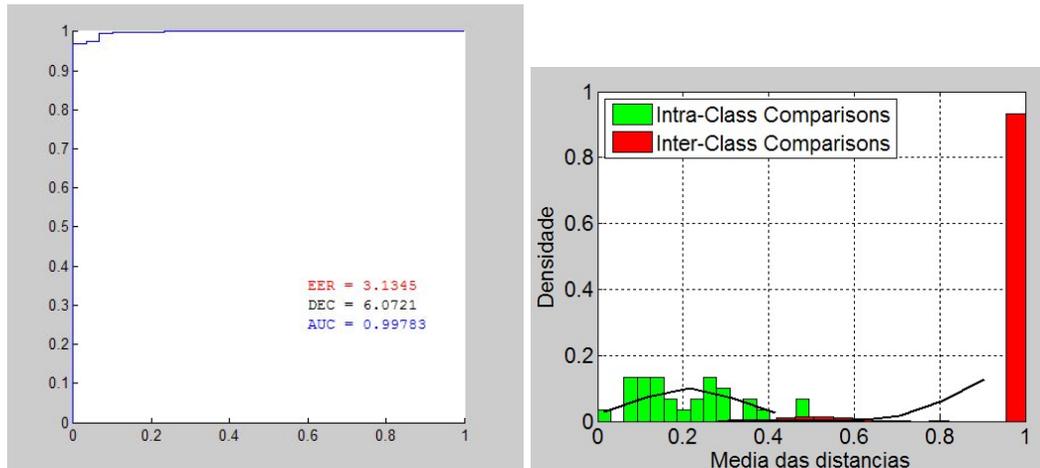
**Figura A.114:** Avaliação do Classificador na Configuração 13 para 20 keypoints (cardinalidade = 20), com máscara, orientação de keypoints  $-3\pi/2 < \theta < \pi/2$



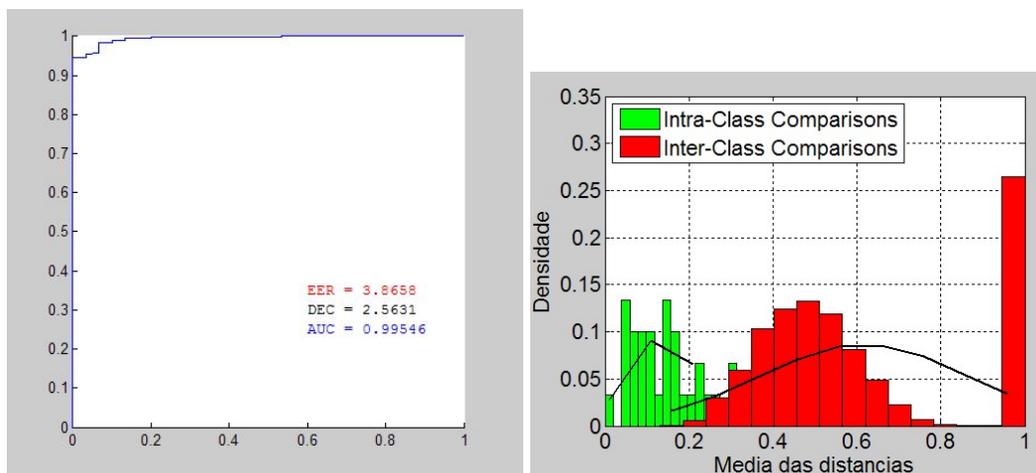
**Figura A.115:** Avaliação do Classificador na Configuração 13 para 10 keypoints (cardinalidade = 10), com máscara, orientação de keypoints  $-3\pi/2 < \theta < \pi/2$



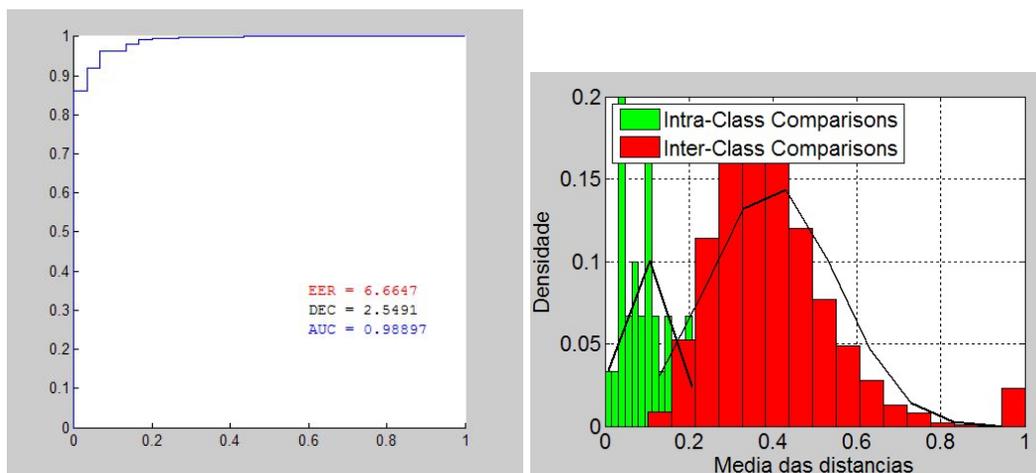
**Figura A.116:** Avaliação do Classificador na Configuração 13 para 5 keypoints (cardinalidade = 5), com máscara, orientação de keypoints  $-3\pi/2 < \theta < \pi/2$



**Figura A.117:** Avaliação do Classificador na Configuração 14 para 20 keypoints (cardinalidade = 20), orientação de keypoints  $-3\pi/2 < \theta < \pi/2$



**Figura A.118:** Avaliação do Classificador na Configuração 14 para 10 keypoints (cardinalidade = 10), orientação de keypoints  $-3\pi/2 < \theta < \pi/2$



**Figura A.119:** Avaliação do Classificador na Configuração 14 para 5 keypoints (cardinalidade = 5), orientação de keypoints  $-3\pi/2 < \theta < \pi/2$

# Bibliografia

- [1] Banafshe Arbab-Zavar, Mark S. Nixon, and David J. Hurley. On model-based analysis of ear biometrics. 2007.
- [2] S. Arya and D.M Mount. Approximate nearest neighbor queries in fixed dimensions. A.C.M, 1993.
- [3] S. Arya, Netanyahu N.S. Silverman R. Mount, D.M., and A.Y. Wu. An optimal algorithm for approximate nearest neighbor searching. A.C.M, 1998.
- [4] A. Sanchez B. Moreno. On the use of outer ear images for personal identification in security applications. In *Security Technology*. IEEE, 1999.
- [5] D.H. Ballard. Generalizing the hough transform to detect arbitrary patterns. 1981.
- [6] A. Baumberg. Reliable feature matching across widely separated views. Conference on Computer Vision and Pattern Recognition, 2000.
- [7] Brian C. Becker. Computer vision sift lecture, 2009.
- [8] J. Beis and D.G. Lowe. Shape indexing using approximate nearest-neighbor search in high-dimensional spaces. Conference on Computer Vision and Pattern Recognition, 1997.
- [9] Alphonse Bertillon. *La photographie judiciaire, avec un appendice sur la classification et l'identification anthropométriques*. Gauthier-Villars, Paris, 1890.
- [10] GmbH Bromba. Bioidentification frequently asked questions, 2008.

- 
- [11] M. Brown and D.G. Lowe. Invariant features from interest point groups. British Machine Vision Conference, 2002.
- [12] M. Burge and Burger. *Biometrics: Personal Identification in a Networked Society*. Kluwer Academic, 1998.
- [13] M. Burge and Burger. In *Ear Biometrics in Computer Vision in 15th International Conference of Pattern Recognition*. ICPR, 2000.
- [14] Court of Appeals of Washington, Division 2. *State vs. David Wayne Kunze*, 1999.
- [15] Matthew David G. Lowe, Brown. Automatic panoramic image stitching using invariant features. 2007.
- [16] Ben Dawson. Biometrics measures physical traits, 2001.
- [17] Intrator N. Edelman, S. and Poggio. Complex cells and object recognition. 1997.
- [18] Bentley J.L. Friedman, J.H. and R.A. Finkel. An algorithm for finding best matches in logarithmic expected time. A.C.M, 1977.
- [19] C. Harris and M. Stephens. A combined corner and edge detector. Fourth Alvey Vision, 1988.
- [20] P.V.C Hough. Method and means for recognizing complex patterns. 1962.
- [21] Nixon M.S. Carter J.N. Hurley, D.J. Automated ear recognition by force field transformations. Colloquium on Visual Biometrics. IEEE, 2000.
- [22] Nixon M.S. Carter J.N. Hurley, D.J. A new force field transform for ear and face recognition. In *International Conference on Image Processing*. IEEE, 2000.
- [23] A. Iannarelli. *Ear Identification*. Paramount Publishing Company, 1989.
- [24] S. Sankar B. Victor K. Chang, K.W. Bowyer. Comparison and combination of ear and face images. In *Pattern Analysis and Machine Intelligence*. IEEE, 2003.

- 
- [25] J.J Koenderink. The structure of images. In *Biological Cybernetics*, 1984.
- [26] T. Lindeberg. Detecting salient blob-like image structures and their scales with a scale-space. *International Journal of Computer Vision*, 1993.
- [27] T. Lindeberg. Scale space theory. In *A basic tool for analysing structures at different scales*. *Journal of Applied Statistics*, 1994.
- [28] David G. Lowe. Object recognition from local scale-invariant features. *International Journal of Computer Vision*, 1999.
- [29] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004.
- [30] Chum O. Urban M. Matas, J. and Pajdla. Robust wide baseline stereo from maximally stable extremal regions. *British Machine Vision Conference*, 2002.
- [31] K. Mikolajczyk and C. Schmid. An affine invariant interest point detector. *European Conference on Computer Vision*, 2002.
- [32] Zisserman A. Mikolajczyk, K. and C. Schmid. Shape recognition with edge-based features. *British Machine Vision Conference*, 2003.
- [33] H. Moravec. Rover visual obstacle avoidance. 1981.
- [34] K.W. Bowyer P. Yan. 2nd and 3rd ear recognition. *Biometric Consortium*, 2004.
- [35] K.W. Bowyer P. Yan. Biometric recognition using 3d ears shape. *IEEE*, 2007.
- [36] C. Schmid and R. Mohr. Local gray value invariants for image retrieval. *IEEE*, 1997.
- [37] K. Satoh S. Nakayama T. Yuizono, Y. Wang. Individual recognition of ear images using genetic local search. *Congress on Evolutionary Computation*, 2002.
- [38] A. Vedaldi and B. Fulkerson. Vlfeat: An open and portable library of computer vision algorithms. <http://www.vlfeat.org>, 2008.
- [39] A.P. Witkin. Scale-space filtering. *International Joint Conference on Artificial Intelligence*, 1983.