



UNIVERSIDADE FEDERAL DA PARAÍBA

CENTRO DE INFORMÁTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

**Classificação e Verificação Multibiométrica por Geometria da Mão e Impressão
Palmar com Otimização por Algoritmos Genéticos**

Arnaldo Gualberto de Andrade e Silva

JOÃO PESSOA – PB

OUTUBRO DE 2015



UNIVERSIDADE FEDERAL DA PARAÍBA

CENTRO DE INFORMÁTICA

PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

**Classificação e Verificação Multibiométrica por Geometria da Mão e Impressão
Palmar com Otimização por Algoritmos Genéticos**

Defesa de Mestrado apresentada ao Centro de Informática da Universidade Federal de Paraíba por Arnaldo Gualberto de Andrade e Silva, sob a orientação do Prof. Dr. Leonardo Vidal Batista, como parte dos requisitos para obtenção do título de Mestre em Informática.

Orientador: Prof. Dr. Leonardo Vidal Batista
Aluno: Arnaldo Gualberto de Andrade e Silva

JOÃO PESSOA – PB

OUTUBRO DE 2015

S586c Silva, Arnaldo Gualberto de Andrade e.

Classificação e verificação multibiométrica por geometria da mão e impressão palmar com otimização por algoritmos genéticos / Arnaldo Gualberto de Andrade e Silva.- João Pessoa, 2015.

128f. : il.

Orientador: Leonardo Vidal Batista

Dissertação (Mestrado) - UFPB/CI

1. Informática. 2. Biometria. 3. Geometria de mão.
4. Impressão palmar. 5. Algoritmos genéticos. 6. Verificação e classificação.

UFPB/BC

CDU: 004(043)

ARNALDO GUALBERTO DE ANDRADE E SILVA

**CLASSIFICAÇÃO E VERIFICAÇÃO MULTIBIOMÉTRICA POR GEOMETRIA DA MÃO E IMPRESSÃO
PALMAR COM OTIMIZAÇÃO POR ALGORITMOS GENÉTICOS**

Aprovado em 23 de Outubro de 2015.

Banca Examinadora:

Prof. Leonardo Vidal Batista, Dr. – UFPB

(Orientador)

Prof.^a Thaís Gaudencio do Rêgo, Dra. – UFPB

(Membro)

Prof. Tiago Pereira do Nascimento, Dr. – UFPB

(Membro)

Prof. Tsang Ing Ren, Dr. – UFPE

(Membro Externo)

Dedico este trabalho aos meus pais por me ensinar os principais princípios da minha vida

Agradecimentos

À minha família. Ao meu pai, Arnaldo Gualberto, por me dar muito mais que apoio financeiro, demonstrando seu amor e preocupação com a minha vida e meu futuro. A minha mãe, Luzia de Andrade, por todo o carinho e cuidado comigo e com as minhas coisas desde o dia em que nasci. E ao meu irmão, Juan Gualberto, por todo orgulho, preocupação e confiança que ele tem por mim e por ser muito mais que um irmão para mim todos os dias.

À minha noiva, Sabrina Figueiredo, por todas as surpresas, por todas as conversas e por todos os momentos em que ficamos juntos; Por ser mais que minha esposa -, sendo minha melhor amiga em todas as horas. E aos nossos amigos, Alberto de Albuquerque, Bárbara Carla e Raísa Vitcel, por estarem presentes em todos esses momentos.

Aos meus amigos, Adriano Marinho, Braulio Siebra, Glauco de Sousa, Hugo Neves, Igor Malheiros, Iron Araújo, João Janduy, José Ivan, Ramon Celestino, Rajiv Albino, Raul Felipe, Rodrigo Parente, Teobaldo Leite, e Yuri Gil por me darem a oportunidade de realizar sonhos, por me ajudar em todas as dificuldades e, simplesmente, por estarem presentes nos dias mais felizes da minha vida.

Aos meus amigos Wesley, Eddie, Henrique, Paiva e Nogueira por fazerem parte da construção do meu caráter e personalidade, contribuindo para a concretização desse trabalho.

Ao meu tutor, o professor Dr. Leonardo Vidal Batista, por acreditar e confiar em mim para realização dos seus projetos; Por todo aprendizado pessoal e profissional como meu tutor e orientador do PET e PIBIC; E por me fazer entender a essência da alma de um pesquisador.

Ao PET.Com por me tornar muito mais que um profissional competente e um pesquisador dedicado, amadurecendo meu lado pessoal abrindo minha cabeça para novos pensamentos e pontos de vista.

A VSoft Tecnologia por me dar a oportunidade de trabalhar em projetos inovadores que estão entre os melhores do Brasil e do mundo, ampliando meu conhecimento e melhorando minhas qualidades como pesquisador.

Ao CAPES e ao CNPq pelo financiamento estudantil durante boa parte da minha vida acadêmica.

A todos que, de alguma forma, contribuíram para a realização desse objetivo.

“A mente que se abre a uma nova ideia jamais
voltará ao seu tamanho original.”

Albert Einstein

RESUMO

A Biometria oferece um mecanismo de autenticação confiável utilizando traços (físicos ou comportamentais) que permitem identificar usuários baseados em suas características naturais. Serviços biométricos de classificação e verificação de usuários são considerados, a princípio, mais seguros que sistemas baseados em políticas de senha, por exigirem a apresentação de uma característica física única e, portanto, a presença do usuário ao menos no momento da autenticação. No entanto, tais métodos apresentam vulnerabilidades e podem resultar em alta taxa de verificação falsa, mesmo em sistemas mais modernos. Algoritmos genéticos (GA), por sua vez, são uma abordagem de otimização baseada no princípio da seleção natural de Charles Darwin e que vêm provando, ao longo dos anos, ser uma ferramenta útil na busca de soluções em problemas complexos. Além disso, seu uso em sistemas biométricos também vem crescendo por se mostrar uma alternativa para seleção de características. Este trabalho utiliza os algoritmos genéticos como ferramenta na otimização de atributos para classificação e verificação biométrica por geometria da mão e impressão palmar. A Base Multibiométrica BioPass-UFPB é apresentada e empregada para teste e validação do método proposto. Ao todo, 99 atributos – sendo 85 geométricos e 14 de textura - extraídos de cada imagem são utilizados e análises sobre a importância desses atributos são realizadas. Os resultados mostraram que, nas duas abordagens de verificação empregadas, os algoritmos genéticos conseguiram melhoras superiores a 30% e 90% da EER em relação ao caso em que o GA não era aplicado. Na classificação, o uso de algoritmos genéticos conseguiu reduzir na média o número de *templates* a serem recuperados pelo sistema para garantir que ao menos um desses seja da mesma classe da amostra de referência. Por fim, espera-se que os resultados deste trabalho, bem como a base BioPass-UFPB, sirvam de referência na implementação de novos sistemas de reconhecimento biométrico baseados na geometria da mão.

Palavras-chave: Biometria, Geometria da Mão, Impressão Palmar, Algoritmos Genéticos, Verificação, Classificação

ABSTRACT

Biometrics provides a trusted authentication mechanism by using traits (physical or behavioral) which identify users based on their natural characteristics. Biometric services of classification and verification of users are considered, in principle, more secure than password-based systems, requiring the presentation of a unique physical characteristic and, therefore, the presence of the user at least in the moment of authentication. However, such methods have vulnerabilities that result in high rates of false verification, even in the most modern systems. On the other hand, Genetic algorithms (GA) are an optimization approach based on the principle of natural selection proposed by Charles Darwin which has been proving to be a useful tool in finding solutions to complex problems. Moreover, the use of genetic algorithms in biometrics systems has also been growing as they are an interesting alternative for selecting features. This work applies a genetic algorithm-based approach to optimizing parameters of classification and verification of a hand dataset. The BioPass-UFPB multi-biometric dataset is presented and used to test and validate the proposed method. In total, 99 features – 85 geometric features and 14 texture features - extracted from each hand image were used. Additionally, the importance of each feature is also analyzed. The results showed relative improvements of EER greater than 30% and 90% in the best cases of the two verification approaches performed, respectively. As for classification, the use of genetic algorithms were able to reduce, on average, the number of templates to be recovered by the system to ensure that at least one of these is of the same class of the reference sample. In conclusion, both the results showed and the BioPass-UFPB dataset might help the development of new hand geometry-based biometric recognition systems.

Keywords: Biometrics, Hand Geometry, Palmprint, Genetic Algorithms, Verification, Classification

Lista de Figuras

Figura 1: Processamento digital e imagens e áreas correlatas. Fonte: (Batista, 2006).	23
Figura 2: Direções de adjacência mais comuns no cálculo da GLCM.....	24
Figura 3: GLCM de uma imagem 4x4 com 4 níveis de cinza para distância $d = 1$ e direção $\vartheta = 0^\circ$	25
Figura 4: Exemplo de cálculo do LBP. Adaptado de (OpenCV, 2008).....	25
Figura 5: LBP de imagens artificialmente modificadas. Fonte: (OpenCV, 2008).....	26
Figura 6: As seis características biométricas mais comuns e outras, que são utilizadas com menor frequência ou que estão em estágios iniciais de pesquisa. Fonte: (Costa et al., 2006).....	28
Figura 7: Posição de alguns sistemas biométricos em um gráfico de segurança versus conveniência. Fonte: (Kulp & Braskamp, 2012).	29
Figura 8: Um modelo simples de sistemas biométricos. Adaptado de Costa et al. (2006).....	30
Figura 9: Exemplos de Impressões Palmares	32
Figura 10: Medidas típicas da geometria da mão. O modelo esquemático (esquerda) pode ser apreciado na imagem real (centro) obtida de um dispositivo (direita)	33
Figura 11: Dispositivo de captura guiado por pinos. Fonte: (Subcommittee on Biometrics of National Science and Technology Council (NSTC), 2006).	34
Figura 12: As curvas típicas das taxas de erro FAR e FRR, plotadas uma ao lado da outra, em relação ao limiar T configurado para o sistema. A EER é representada pelo ponto de interseção das curvas. Fonte: (Costa et al., 2006).	37
Figura 13: Curva ROC. As taxas de erro FAR e (1-FRR) podem ser plotadas uma contra outra numa curva bidimensional.	38
Figura 14: Exemplificação dos tipos de cruzamento implementados nos algoritmos genéticos.	41
Figura 15: Exemplo da operação de mutação.....	41
Figura 16: (a) Dispositivo de captura composto por negatoscópio, caixa em MDF e câmera fotográfica. (b) Detalhe da câmera posicionada na caixa.	50
Figura 17: Exemplo de imagens capturadas pelo dispositivo: (a) geometria da mão. (b) impressão palmar	51
Figura 18: Ângulo entre os segmentos que representam a inclinação natural dos dedos. Fonte: do Nascimento et al. (2014).	53
Figura 19: Ilustração dos 14 pontos do contorno da palma definidos por (do Nascimento et al., 2014).	54
Figura 20: Ilustração da rotina de testes da primeira abordagem da verificação. Cada caixa representa um conjunto de 10 classes, isto é, a caixa “1-10” representa as 10 primeiras classes; a caixa “11-20” representa as classes de 11 a 20, e assim por diante. Ademais, as caixas destacadas em verde representam os conjuntos de treinamento, enquanto as caixas azuis representam os conjuntos de validação.	56
Figura 21: Ilustração da rotina de testes da segunda abordagem da verificação. Cada caixa representa agora todas as classes da n-ésima amostra, isto é, a caixa “1ª” é formada pela 1ª amostra de todas as classes, a caixa “2ª” é formada pela 2ª amostra de todas as classes, e assim por diante. Novamente, as caixas destacadas em verde representam os conjuntos de treinamento, enquanto as caixas azuis representam os conjuntos de validação.	57
Figura 22: Ilustração dos testes de identificação 1:N. Cada amostra SCi do banco Bi (banco de <i>template</i>) é comparada com todas as amostras do banco Bj (banco de referência). O resultado para cada amostra é ordenado em ordem crescente e encontra-se o número de templates a serem	

recuperados pelo sistema (k) cujas amostras S_i e S_j pertencem a mesma classe. O objetivo do GA nesse contexto é minimizar o maior número de <i>templates</i> recuperados.	58
Figura 23: (a) Ilustração da demarcação da ROI na palma da mão. (b) ROI extraída	59
Figura 24: Imagens das classes 96 (a)-(e) e 97 (f)-(j).	67
Figura 25: Gráficos do tipo radar que analisam o valor dos pesos atribuídos aos 85 atributos pelo GA. Atributos (a) da palma, (b) do dedo mínimo, (c) do dedo anelar, (d) do dedo médio, (e) do dedo indicador e (f) do polegar. A numeração dos atributos segue a numeração definida na Seção 4.3. Os valores dos atributos foram normalizados na escala 30,67342:1.	71
Figura 26: Gráficos do tipo radar que analisam o valor dos pesos atribuídos aos 85 atributos pelo GA na 1ª execução da 6ª combinação mostrada na Tabela 38. Os valores dos atributos foram normalizados na escala 70,61395:1.	76
Figura 27: Gráficos do tipo radar que analisam o valor dos pesos atribuídos aos 85 atributos pelo GA na 3ª execução da 6ª combinação mostrada na Tabela 38. Os valores dos atributos foram normalizados na escala 52,06996:1.	78
Figura 28: Gráficos do tipo radar que analisam os valores dos pesos dados pelo GA aos atributos de textura nos melhores casos das duas primeiras abordagens da verificação, (a) e (b), respectivamente.	82
Figura 29: Gráficos do tipo radar que analisam o valor dos pesos atribuídos aos 85 atributos pelo GA no conjunto de treinamento "2ax1a". Os valores dos atributos foram normalizados na escala 22,77208:1.	89
Figura 30: Gráfico do tipo radar que analisa os valores dos pesos dados pelo GA aos atributos de textura no melhor caso da segunda abordagem da classificação. Os valores dos atributos foram normalizados na escala 13,88468:1.	91

Lista de Equações

Equação 1: Cálculo da FAR	36
Equação 2: Cálculo da FRR	36
Equação 3: Taxa de acerto	38
Equação 4: Medida de similaridade entre dois <i>templates</i> A e B.....	48
Equação 5: Correlação de Histograma	49

Lista de Gráficos

Gráfico 1: Histograma da combinação de parâmetros para a menor EER nos bancos de treinamento e teste na primeira abordagem da verificação	66
Gráfico 2: Resultados das menores EER nos conjunto de treinamento na primeira abordagem da verificação. O conjunto de treinamento 1 é formado pelas 10 primeiras classes; o segundo é formado pelas 20 primeiras classes, e assim por diante.	68
Gráfico 3: Resultados das menores EER nos conjunto de testes na primeira abordagem da verificação. O conjunto de teste 1 é formado pelas 90 últimas classes; o segundo é formado pelas 80 últimas classes, e assim por diante.	68
Gráfico 4: Resultado dos ganhos de performance nos conjuntos de treinamento e teste para cada rodada da primeira abordagem da verificação. Cada rodada representa os conjuntos de treinamento e teste conforme o Gráfico 2 e Gráfico 3.	69
Gráfico 5: Curvas FAR e FRR do conjunto de teste formado pelas 50 últimas classes.	72
Gráfico 6: Histograma da combinação de parâmetros para a menor EER nos bancos de treinamento e teste na segunda abordagem da verificação	73
Gráfico 7: Resultados das menores EER nos conjuntos de treinamento na 2ª abordagem da verificação. Cada conjunto de treinamento representa cada permutação possível das instâncias.	74
Gráfico 8: Resultados das menores EER nos conjuntos de teste na 2ª abordagem da verificação. Cada conjunto de teste é formado pelas instâncias não utilizadas em cada permutação do conjunto de treinamento.	74
Gráfico 9: Resultado dos ganhos de performance nos conjuntos de treinamento e teste para cada rodada da validação cruzada da segunda abordagem da verificação.	75
Gráfico 10: Curvas FAR e FRR do conjunto de teste formado pelas 3ª e 5ª instâncias	80
Gráfico 11: Histograma da combinação de parâmetros do GA para o vizinho mais próximo ($k = 1$) nos bancos de treinamento da segunda abordagem da classificação.	87

Lista de Tabelas

Tabela 1: Distribuição por finalidade das principais aplicações biométricas. Fonte: (BITE, 2005)	30
Tabela 2: Estatísticas das idades dos voluntários da base BioPass-UFPB	51
Tabela 3: Parâmetros de teste para o algoritmo genético	54
Tabela 4: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (10 primeiras classes) e no banco de testes (90 últimas classes)	60
Tabela 5: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (20 primeiras classes) e no banco de testes (80 últimas classes)	61
Tabela 6: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (30 primeiras classes) e no banco de testes (70 últimas classes)	61
Tabela 7: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (40 primeiras classes) e no banco de testes (60 últimas classes)	62
Tabela 8: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (50 primeiras classes) e no banco de testes (50 últimas classes)	62
Tabela 9: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (60 primeiras classes) e no banco de testes (40 últimas classes)	63
Tabela 10: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (70 primeiras classes) e no banco de testes (30 últimas classes)	64
Tabela 11: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (80 primeiras classes) e no banco de testes (20 últimas classes)	64
Tabela 12: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (90 primeiras classes) e no banco de testes (10 últimas classes)	65
Tabela 13: Média dos resultados obtidos na primeira abordagem da verificação.....	69
Tabela 14: Análise estatística dos pesos atribuídos a cada parte integrante da mão sem normalização.	71
Tabela 15: Média dos resultados obtidos na segunda abordagem da verificação.....	73
Tabela 16: Análise estatística dos pesos atribuídos a cada parte integrante da mão sem normalização. Os valores mostrados são correspondentes a 1ª execução da 6ª combinação mostrada na Tabela 38.	77
Tabela 17: Análise estatística dos pesos atribuídos a cada parte integrante da mão sem normalização. Os valores mostrados são correspondentes a 3ª execução da 6ª combinação mostrada na Tabela 38.	79
Tabela 18: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (50 primeiras classes) e no banco de testes (50 últimas classes) com a inclusão dos atributos de textura.....	81
Tabela 19: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (1ª, 2ª e 4ª instâncias) e no banco de testes (3ª e 5ª instâncias) com a inclusão dos atributos de textura	82
Tabela 20: Análise estatística dos pesos atribuídos pelo GA aos 99 atributos no melhor caso da primeira abordagem da verificação, sem normalização.....	83
Tabela 21: Análise estatística dos pesos atribuídos pelo GA aos 99 atributos no melhor caso da segunda abordagem da verificação, sem normalização.	83
Tabela 22: Visão geral dos resultados obtidos nas três abordagens da verificação. Resultados mostrados para o melhor caso nos bancos de testes.....	84
Tabela 23: Resultados da classificação 1:N utilizando pesos unitários.....	86

Tabela 24: Dados estatísticos dos conjuntos de teste para cada conjunto de treinamento da segunda abordagem da classificação.	88
Tabela 25: Análise estatística dos pesos atribuídos a cada parte integrante da mão sem normalização. Os valores mostrados são correspondentes ao conjunto de treinamento “2ax1a”	90
Tabela 26: Análise estatística dos pesos atribuídos pelo GA aos 99 atributos no melhor caso da segunda abordagem da classificação, sem normalização.....	91
Tabela 27: Visão geral dos resultados obtidos nas duas últimas abordagens da classificação. O valor de k representa o número de <i>templates</i> que o sistema deve retornar para garantir que a classe da amostra de referência está presente nos k templates retornados. Resultados mostrados para o melhor caso nos bancos de testes.	92
Tabela 28: Resultado da Aplicação do GA utilizando as 10 primeiras classes para treinamento e as 90 últimas classes para teste.	102
Tabela 29: Resultado da Aplicação do GA utilizando as 20 primeiras classes para treinamento e as 80 últimas classes para teste.	103
Tabela 30: Resultado da Aplicação do GA utilizando as 30 primeiras classes para treinamento e as 70 últimas classes para teste.	104
Tabela 31: Resultado da Aplicação do GA utilizando as 40 primeiras classes para treinamento e as 60 últimas classes para teste.	105
Tabela 32: Resultado da Aplicação do GA utilizando as 50 primeiras classes para treinamento e as 50 últimas classes para teste.	106
Tabela 33: Resultado da Aplicação do GA utilizando as 60 primeiras classes para treinamento e as 40 últimas classes para teste.	107
Tabela 34: Resultado da Aplicação do GA utilizando as 70 primeiras classes para treinamento e as 30 últimas classes para teste.	108
Tabela 35: Resultado da Aplicação do GA utilizando as 80 primeiras classes para treinamento e as 20 últimas classes para teste.	109
Tabela 36: Resultado da Aplicação do GA utilizando as 90 primeiras classes para treinamento e as 10 últimas classes para teste.	110
Tabela 37: Resultado da Aplicação do GA utilizando as 1ª, 2ª e 3ª instâncias de cada classe para treinamento e as 4ª e 5ª instâncias para teste.	111
Tabela 38: Resultado da Aplicação do GA utilizando as 1ª, 2ª e 4ª instâncias de cada classe para treinamento e as 3ª e 5ª instâncias para teste.	112
Tabela 39: Resultado da Aplicação do GA utilizando as 1ª, 2ª e 5ª instâncias de cada classe para treinamento e as 3ª e 4ª instâncias para teste.	113
Tabela 40: Resultado da Aplicação do GA utilizando as 1ª, 3ª e 4ª instâncias de cada classe para treinamento e as 2ª e 5ª instâncias para teste.	114
Tabela 41: Resultado da Aplicação do GA utilizando as 1ª, 3ª e 5ª instâncias de cada classe para treinamento e as 2ª e 4ª instâncias para teste.	115
Tabela 42: Resultado da Aplicação do GA utilizando as 1ª, 4ª e 5ª instâncias de cada classe para treinamento e as 2ª e 3ª instâncias para teste.	116
Tabela 43: Resultado da Aplicação do GA utilizando as 2ª, 3ª e 4ª instâncias de cada classe para treinamento e as 1ª e 5ª instâncias para teste.	117
Tabela 44: Resultado da Aplicação do GA utilizando as 2ª, 3ª e 5ª instâncias de cada classe para treinamento e as 1ª e 4ª instâncias para teste.	118

Tabela 45: Resultado da Aplicação do GA utilizando as 2ª, 4ª e 5ª instâncias de cada classe para treinamento e as 1ª e 3ª instâncias para teste.	119
Tabela 46: Resultado da Aplicação do GA utilizando as 3ª, 4ª e 5ª instâncias de cada classe para treinamento e as 1ª e 2ª instâncias para teste.	120
Tabela 47: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (1ª, 2ª e 3ª instâncias) e no banco de testes (4ª e 5ª instâncias)	121
Tabela 48: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (1ª, 2ª e 4ª instâncias) e no banco de testes (3ª e 5ª instâncias)	121
Tabela 49: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (1ª, 2ª e 5ª instâncias) e no banco de testes (3ª e 4ª instâncias)	121
Tabela 50: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (1ª, 3ª e 4ª instâncias) e no banco de testes (2ª e 5ª instâncias)	122
Tabela 51: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (1ª, 3ª e 5ª instâncias) e no banco de testes (2ª e 4ª instâncias)	122
Tabela 52: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (1ª, 4ª e 5ª instâncias) e no banco de testes (2ª e 3ª instâncias)	122
Tabela 53: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (2ª, 3ª e 4ª instâncias) e no banco de testes (1ª e 5ª instâncias)	123
Tabela 54: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (2ª, 3ª e 5ª instâncias) e no banco de testes (1ª e 4ª instâncias)	123
Tabela 55: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (2ª, 4ª e 5ª instâncias) e no banco de testes (1ª e 3ª instâncias)	123
Tabela 56: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (3ª, 4ª e 5ª instâncias) e no banco de testes (1ª e 2ª instâncias)	124

Lista de Siglas

AUC	<i>Area Under the Curve (Área Sob a Curva)</i>
CMR	<i>Correct Match Rate (Taxa de Casamento Correto)</i>
CRR	<i>Correct Recognition Rate (Taxa de Reconhecimento Correto)</i>
DDR3	<i>Double Data Rate 3</i>
EER	<i>Equal Error Rate (Taxa de Erro Igual)</i>
FA	<i>Falsa Aceitação</i>
FR	<i>Falsa Rejeição</i>
FAR	<i>False Acceptance Rate (Taxa de Falsa Aceitação)</i>
FRR	<i>False Rejection Rate (Taxa de Falsa Rejeição)</i>
GA	<i>Genetic Algorithm (Algoritmo Genético)</i>
GB	<i>Giga Byte</i>
GHz	<i>Giga Hertz</i>
GLCM	<i>Gray Level Co-occurrence Matrix (Matrix de Coocorrência de Níveis de Cinza)</i>
IDE	<i>Integrated Development Environment (Ambiente de Desenvolvimento Integrado)</i>
K-NN	<i>k-Nearest Neighbour (k-Vizinhos mais Próximos)</i>
LBP	<i>Local Binary Pattern (Padrão Binário Local)</i>
LDA	<i>Linear Discriminant Analysis (Análise de Discriminância Linear)</i>
RAM	<i>Random Access Memory (Memória de Acesso aleatório)</i>
ROC	<i>Receiver Operating Characteristic (Característica de Operação do Receptor)</i>
ROI	<i>Region Of Interest (Região de Interesse)</i>
SVM	<i>Support Vector Machine (Máquina de Vetor de Suporte)</i>
TA	<i>Taxa de Acerto</i>

SUMÁRIO

Introdução	19
1.1 Objetivos	20
1.2 Estrutura da Dissertação	21
Fundamentação Teórica.....	22
2.1 Imagens, Processamento Digital de Imagens e Visão Computacional.....	22
2.1.1 Atributos de Haralick	24
2.1.2 Padrão Binário Local (Local Binary Pattern)	25
2.2 Biometria.....	26
2.2.1 Sistemas Biométricos	29
2.2.2 Impressão Palmar	31
2.2.3 Geometria da Mão.....	33
2.2.4 Autenticação versus Identificação	35
2.2.5 Métricas de Erro.....	36
2.3 Algoritmos Genéticos	38
2.3.1 Conceitos.....	39
2.3.2 Algoritmo	42
2.3.3 Aplicações	42
2.3.4 Influência dos Parâmetros dos Algoritmos Genéticos.....	43
Trabalhos Relacionados	45
Materiais e Métodos	48
4.1 Ambiente de Desenvolvimento.....	48
4.2 Medida de Similaridade	48
4.3 Base de dados BioPass-UFPB	49
4.4 Extração da ROI	53
4.5 Rotina de Testes	54
4.5.1 Verificação	55

4.5.2	<i>Classificação</i>	57
Resultados e Discussões.....		59
5.1	Extração da ROI	59
5.2	Verificação.....	60
5.2.1	<i>Primeira Abordagem</i>	60
5.2.2	<i>Segunda Abordagem</i>	72
5.2.3	<i>Terceira Abordagem</i>	80
5.2.4	<i>Comparação Entre as Abordagens de Verificação</i>	83
5.3	Classificação	85
5.3.1	<i>Primeira Abordagem</i>	85
5.3.2	<i>Segunda Abordagem</i>	85
5.3.3	<i>Terceira Abordagem</i>	90
5.3.4	<i>Comparação Entre as Abordagens de Classificação</i>	91
5.4	Análise dos Atributos	93
Considerações Finais		94
Referências.....		97
APÊNDICE A: Resultados Do GA Na Verificação		102
ANEXO A: Parâmetros Da Câmera De Captura		125
ANEXO B: Termo De Consentimento Livre E Esclarecido		126
ANEXO C: Instrumento		127
ANEXO D: Certidão De Aprovação De Projeto De Pesquisa		128

Capítulo 1

Introdução

O mundo passa por um processo intenso de virtualização da economia, em que cada vez menos dinheiro circula em sua apresentação clássica de papel-moeda. Os computadores exercem um papel fundamental nesse processo, pois automatizam diversas funções organizacionais, dentre elas, funções bancárias de transferência de ativos.

Atualmente, serviços Web de autenticação baseados em políticas de senha, mecanismos de criptografia de dados e assinaturas digitais são massivamente utilizados. No entanto, a metodologia tradicional de autenticação baseada em senha permite que qualquer pessoa munida de um *login* e sua respectiva senha, tenha acesso a um serviço, independentemente de ser o usuário a quem o *login* e senha se destinam. Há também serviços biométricos de classificação e verificação de usuários, considerados, a princípio, mais seguros, por exigirem a apresentação de uma característica física única e, portanto, a presença do usuário ao menos no momento da autenticação.

No entanto, tais métodos também apresentam vulnerabilidades. Os serviços biométricos que utilizam a impressão digital como característica biométrica, por exemplo, podem permitir que uma reprodução da digital de um usuário em material de silicone seja usada para acesso indevido. A adoção de métodos para identificar o uso de dedos falsos (ausência de fluxo sanguíneo, perspiração, odor etc.) (Antonelli et al., 2006) torna o processo complicado e custoso. Um trabalho recente apontou 90% de taxa de verificação falsa em sistemas modernos de reconhecimento de digitais, utilizando-se dedos de cadáveres, de plástico, de gelatina e de outros materiais de modelagem (Roberts, 2007).

Assim como impressões digitais, impressões palmares são únicas e invariantes em relação à idade do indivíduo, e técnicas de reconhecimento utilizando impressões palmares vem aumentando ao longo dos anos. A classificação pela palma da mão é considerada uma alternativa em potencial para métodos de classificação biométrica em relação aos métodos mais populares, como impressões digitais (Yunyong & Yodchanan, 2009). Por exemplo, se a

mão de uma pessoa para classificação está suja, a precisão do método baseado em impressões digitais é prejudicada, enquanto que o método baseado na impressão palmar ainda mantém alta a precisão de classificação (Zhang et al., 2003).

Por serem consideradas relativamente novas, as técnicas de reconhecimento pela palma e geometria da mão ainda deixam lacunas nessa área que podem ser melhoradas tanto na classificação, quanto na verificação de impressão palmares. Como exemplo, pode-se citar a otimização dos pesos que são aplicados para ponderar a importância dos fatores na mensuração do grau de similaridade entre dois *templates*. Em geral, esses pesos são computados de maneira empírica ou estatisticamente. No entanto, a otimização desses pesos pode tornar-se uma difícil tarefa em amplos espaços de atributos N-dimensional, fazendo-se necessária a utilização de estratégias de otimização como, por exemplo, os algoritmos genéticos.

Algoritmos genéticos (em inglês, *Genetic Algorithms*, GA) são uma abordagem de otimização baseada no princípio da seleção natural de Charles Darwin (1951). Ao longo dos anos, os algoritmos genéticos têm provado ser uma significativa ferramenta na busca de soluções em problemas complexos, como: caixeiro viajante, problema da mochila, escalonamento de eventos, etc. (Dawid, 1996). Na literatura, encontram-se trabalhos que aplicam o GA desde a detecção de objetos (Sun et al., 2004) até detecção de genes em dados de *microarray* (McLachlan et al., 2002). Na seleção de características, por sua vez, espera-se que a aplicação do GA seja capaz de atingir 3 objetivos: melhorar a taxa de precisão dos classificadores; melhorar a velocidade e a eficiência dos preditores através da redução de características; e fornecer uma melhor compreensão do processo subjacente que gerou os dados (Luque et al., 2011).

Este trabalho analisa a aplicação de algoritmos genéticos como ferramenta na otimização de parâmetros da classificação e verificação. A base de dados multibiométricos BioPass-UFPB foi utilizada para validação do método proposto.

1.1 Objetivos

O objetivo geral do presente trabalho é aplicar, em diferentes abordagens, os GAs no banco de imagens de mão da base multibiométrica BioPass-UFPB na tentativa de otimizar os parâmetros da classificação e verificação.

Os objetivos específicos são:

1. Desenvolver um método de cálculo de escore entre dois *templates* da base BioPass-UFPB;
2. Aplicar o GA em diferentes abordagens de verificação e classificação para validar o método proposto;
3. Analisar os resultados do GA para identificar os principais atributos geométricos e de textura extraídos da base BioPass-UFPB em cada abordagem.

1.2 Estrutura da Dissertação

Neste trabalho, além deste Capítulo introdutório, estão dispostos mais 5 Capítulos. O Capítulo 2 apresenta, dentre outros, os principais conceitos acerca de processamento digital de imagens, visão computacional, biometria e algoritmos genéticos, necessários para o entendimento deste trabalho. No Capítulo 3, os principais trabalhos relacionados são apresentados. No Capítulo 4 estão descritos os materiais utilizados e os métodos desenvolvidos no decorrer da pesquisa e o Capítulo 5 mostra e discute os principais resultados do método proposto. Por fim, no Capítulo 6 são feitas considerações a respeito de todo o trabalho.

Capítulo

2

Fundamentação Teórica

Neste Capítulo são apresentados os principais conceitos teóricos empregados na realização desse trabalho. Primeiramente, são apresentados conceitos acerca do processamento digital de imagens e algumas de suas vertentes. Logo após, definições a respeito de biometria e sistemas biométricos são apresentadas. Por fim, noções acerca de algoritmos genéticos são dissertadas.

2.1 Imagens, Processamento Digital de Imagens e Visão Computacional

Imagens são sinais, ou seja, funções que conduzem alguma informação a respeito de algo com um interesse. Uma imagem monocromática é uma função $f(x,y)$ onde x e y representam as suas coordenadas espaciais e o valor de $f(x,y)$ representa um valor de intensidade luminosa, geralmente chamada nível de cinza. Para serem representadas no computador, é necessário digitalizar as imagens, o que poder ser feito facilmente através dos processos de amostragem e quantização (Batista, 2006). Amostragem refere-se a capturar, em intervalos de tempo, o sinal analógico que representa a imagem, enquanto que quantização refere-se a definir os valores que o sinal pode assumir no meio digital (Lordão, 2009). O resultado de tais processos é uma imagem digital monocromática, representada da seguinte forma:

$$\begin{pmatrix} f(0,0) & \cdots & f(0,C-1) \\ \vdots & \ddots & \vdots \\ f(L-1,0) & \cdots & f(L-1,C-1) \end{pmatrix}$$

Os valores L e C representam, respectivamente, o número máximo de linhas e colunas da imagem digitalizada, e são determinados no processo de digitalização. Cada elemento da matriz $f(x,y)$ é conhecido como *pixel*, e, como dito, representa o valor do nível de cinza naquele ponto. O número total de níveis de cinza N também é determinado no processo de

digitalização. Em geral, o menor valor é 0 (zero), e representa o preto, enquanto o maior valor, $N - 1$, representa o branco.

O *Processamento Digital de Imagens* (PDI) é o campo da Ciência da Computação que se dedica ao processamento de imagens digitais em um computador digital (Gonzales et al., 2004). Em síntese, pode-se dizer que o PDI consiste em qualquer forma de processamento de dados no qual a entrada e saída são imagens tais como fotografias ou quadros de vídeo. Ao contrário do tratamento de imagens, que se preocupa somente com a manipulação de figuras para sua representação final, o PDI é um estágio para novos processamentos de dados tais como aprendizagem de máquina ou reconhecimento de padrões.

O Processamento Digital de Imagens é altamente multidisciplinar (Batista, 2006) e tem relações muito próximas com as áreas de Computação Gráfica (a qual consiste em obter imagens a partir de dados) e Visão Computacional (a qual consiste em obter dados a partir de imagens), como mostra a Figura 1. Muitas vezes um sistema de PDI exerce tarefas dessas duas áreas ao mesmo tempo. Outros exemplos de áreas ligados ao PDI são inteligência artificial, animação, reconhecimento de padrões e a indústria do entretenimento.

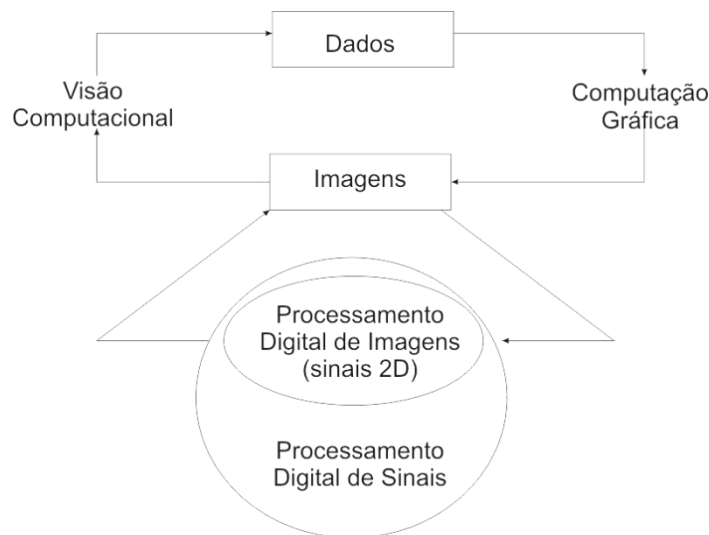


Figura 1: Processamento digital e imagens e áreas correlatas. Fonte: (Batista, 2006).

No âmbito das aplicações da Visão Computacional, uma das áreas mais importantes é a análise de textura. Tradicionalmente, quando é realizada a classificação de uma imagem, apenas os tons de cinza são analisados e não se tira vantagem do conteúdo contextual das imagens como faria o foto intérprete (Sette & Maillard, 2014). A análise de textura visual é

uma das maneiras de levar em conta o contexto no qual cada pixel é inserido. Vários estudos já comprovaram que a textura melhora a acurácia da classificação (Franklin et al., 2001; Palubinskas et al., 1995; Haralick et al., 1973). Além da textura, atributos de cor, forma e movimento também podem ser utilizados na classificação de imagens.

Neste trabalho, técnicas de análise de textura de imagens foram aplicadas na extração de atributos de uma imagem. As próximas subseções descrevem as técnicas utilizadas para tal fim.

2.1.1 Atributos de Haralick

Haralick et al. (1973) descrevem a textura de uma imagem como uma das características mais importantes usada na identificação de objetos ou regiões de interesse em uma imagem. Além disso, uma nova técnica de extração de atributos de textura é apresentada, onde até 14 características podem ser calculadas por essa técnica.

O primeiro passo para extração das características de textura é o cálculo da Matrix de Coocorrência de Níveis de Cinza (em inglês, *Gray Level Co-occurrence Matrix*, GLCM) que considera o relacionamento entre dois pixels vizinhos, cujo primeiro pixel é conhecido como referência e o segundo como pixel vizinho. Dado uma imagem com N níveis de cinza, o elemento (i, j) de uma GLCM representa o número de ocorrências do par de níveis i e j dado uma distância d e um ângulo θ entre i e j . O resultado final é uma GLCM de tamanho $N \times N$ e normalizada pela soma dos elementos não nulos. A Figura 2 apresenta os oito ângulos de adjacência mais utilizados e a Figura 3 exemplifica o cálculo de uma GLCM.

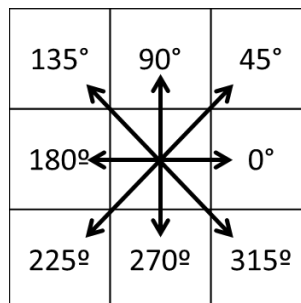


Figura 2: Direções de adjacência mais comuns no cálculo da GLCM

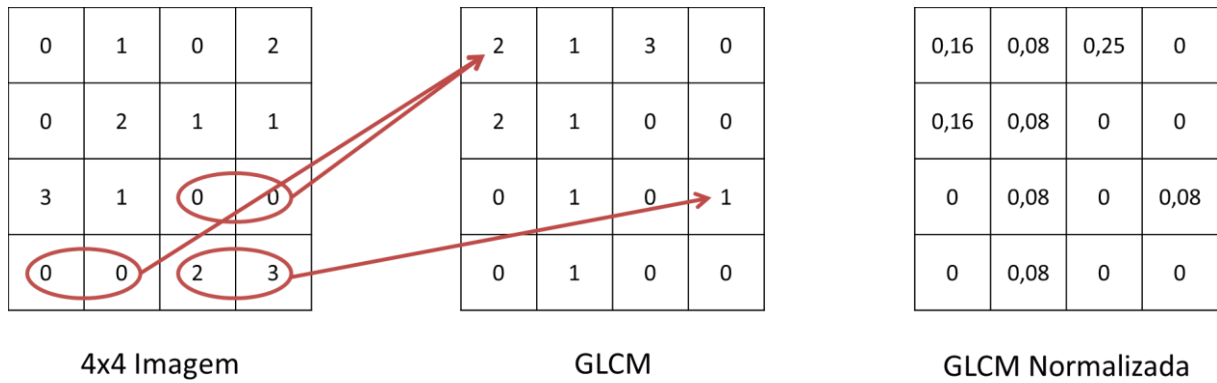


Figura 3: GLCM de uma imagem 4x4 com 4 níveis de cinza para distância $d = 1$ e direção $\theta = 0^\circ$.

A GLCM normalizada é utilizada para o cálculo dos 14 atributos de Haralick descritos em Haralick et al. (1973).

2.1.2 Padrão Binário Local (*Local Binary Pattern*)

O Padrão Binário Local (do inglês, *Local Binary Pattern*, LBP), proposto em Ojala et al. (1996), é um método utilizado, principalmente, na análise de textura em imagens. Basicamente, para cada pixel da imagem, o LBP aplica um processo de limiarização binária dos oito vizinhos mais próximos pelo valor do pixel atual. Os valores binarizados são, então, compostos num único valor de 8 dígitos que, quando convertidos para decimal, tornam-se o valor de textura presente na imagem resultante do LBP. A Figura 4 exemplifica o processo de cálculo do LBP para um determinado pixel.

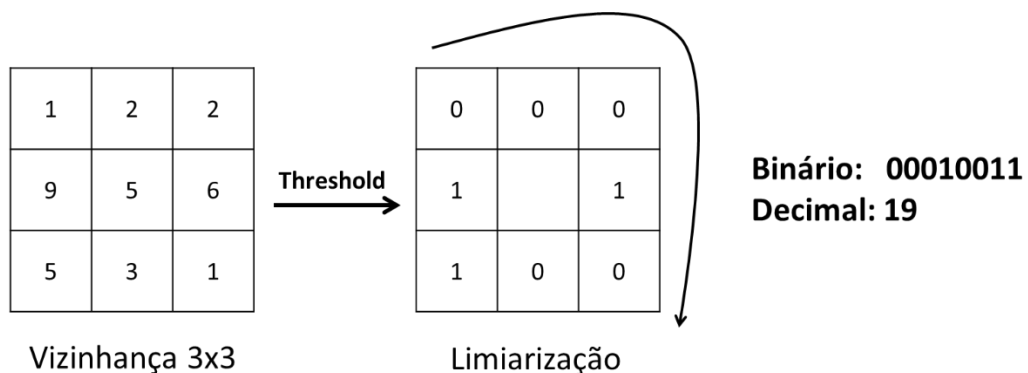


Figura 4: Exemplo de cálculo do LBP. Adaptado de (OpenCV, 2008).

Por definição, o LBP é robusto a variação uniforme de luminosidade, o que pode ser visualmente verificado na Figura 5

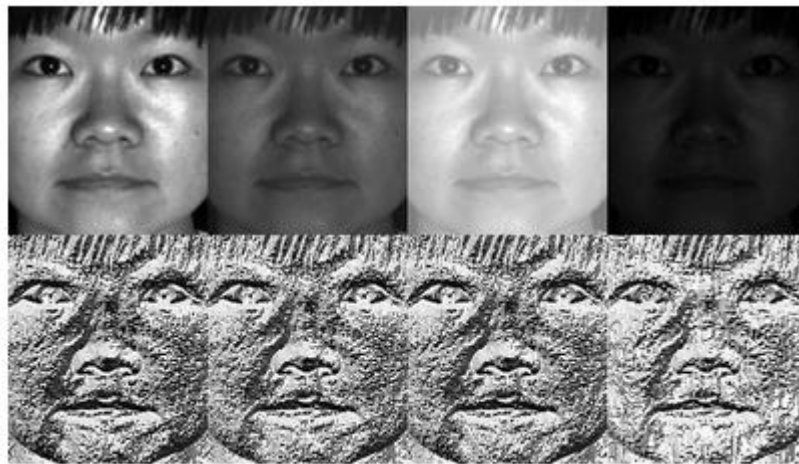


Figura 5: LBP de imagens artificialmente modificadas. Fonte: (OpenCV, 2008).

Uma nova versão do LBP foi introduzida no trabalho de Ojala et al. (2002). Esta versão possibilita a extensão do raio de análise bem como a escolha do número de vizinhos a serem analisados. No entanto, neste trabalho optou-se por utilizar a versão original do LBP.

A próxima Seção apresenta os conceitos acerca da biometria e sistemas biométricos, enfatizando as impressões palmares e geometria da mão.

2.2 Biometria

O termo Biometria se refere ao uso de características físicas ou comportamentais, tais como face, íris, impressão digital, voz e forma de digitar (*keystroke*), para identificar pessoas automaticamente. Conforme Clarke (1994), qualquer dessas características humanas pode ser usada como característica biométrica desde que ela satisfaça alguns requisitos básicos:

- **Universalidade:** toda a população (a ser autenticada) deve possuir a característica. Na prática, temos pessoas que não possuem impressões digitais, por exemplo.
- **Unicidade:** uma característica biométrica deve ser única para cada indivíduo, ou seja, a possibilidade de pessoas distintas possuírem características idênticas, deve ser nula ou desprezível. Assim, a altura de uma pessoa não é uma boa característica para autenticação, já que várias pessoas podem possuir a mesma altura. Na prática, as características biométricas podem apresentar maior ou menor

grau de unicidade, mas nenhuma delas pode ser considerada absolutamente única para cada indivíduo.¹

- **Permanência:** a característica deve ser imutável. Na prática, existem alterações ocasionadas pelo envelhecimento, pela mudança das condições de saúde ou mesmo emocionais das pessoas e por mudanças nas condições do ambiente de coleta.
- **Coleta:** a característica tem que ser passível de mensuração por meio de um dispositivo. Na prática, todas as características biométricas utilizadas comercialmente atendem a este requisito.
- **Aceitação :** a coleta da característica deve ser tolerada pelo indivíduo em questão. Na prática, existem preocupações com higiene, com privacidade e questões culturais que diminuem a aceitação da coleta.

Além destas, Jain et al. (2004) ainda definem mais duas características ideais:

- **Desempenho:** referem-se à precisão do reconhecimento e velocidade, os recursos requeridos para conseguir uma precisão e velocidade aceitáveis, e ao trabalho ou fatores ambientes que afetam a precisão e a velocidade.
- **Proteção:** relacionada à facilidade/dificuldade de enganar o sistema com técnicas fraudulentas.

É importante salientar que, na prática, nenhuma característica biométrica atende com perfeição a todos os requisitos de uma característica biométrica ideal.

As características biométricas ainda são divididas em dois grupos, ilustrados pela Figura 6. As características fisiológicas (ou estáticas) são traços fisiológicos, originários da carga genética do indivíduo e, essencialmente, variam pouco (ou nada) ao longo do tempo. De tal maneira, as principais características estáticas são representadas pela aparência facial, a geometria da mão, o padrão da íris e as impressões digitais.

¹ A quantidade de variação devida à genética e ao ambiente muda de biometria para biometria. Cada pessoa é única, se analisada com suficiente detalhe. É próximo do impossível que duas pessoas diferentes tenham a mesma, idêntica, representação biométrica em qualquer sistema razoável. Contudo, ao lidar com tecnologias práticas de autenticação, encontramos limites na resolução das imagens extraídas, na capacidade de armazenamento e na habilidade de comparação entre dados extraídos. Na prática, isto extermina a noção de unicidade absoluta para todas as características biométricas (Costa et al., 2006).

Outras características estáticas também são utilizadas em menor grau ou estão em estágios iniciais de pesquisa, como o DNA (Bolle et al., 2004), o formato das orelhas (Burge & Burger, 2000), o padrão vascular da retina (Hill, 2002), o odor do corpo (Korotkaya, 2003), o padrão da arcada dentária (Chen & Jain, 2005) e o padrão de calor do corpo ou de partes dele (Prokoski & Riedel, 1999).

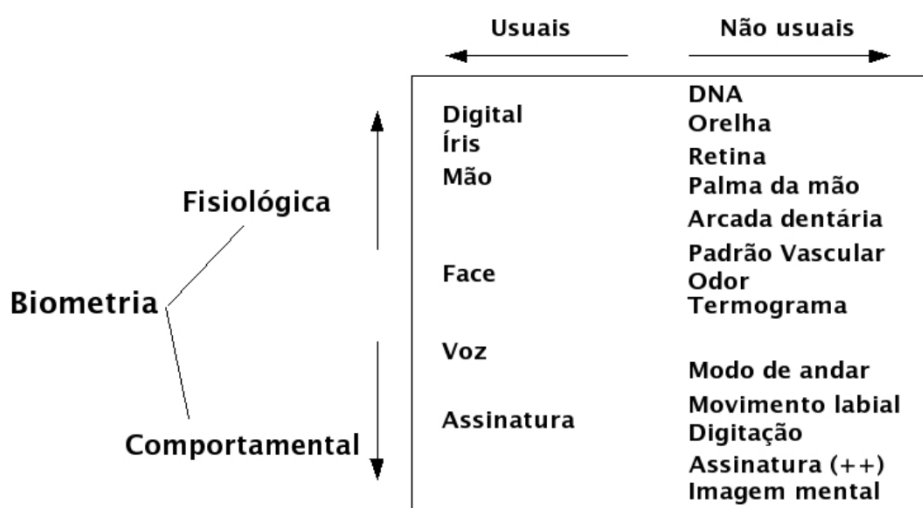


Figura 6: As seis características biométricas mais comuns e outras, que são utilizadas com menor frequência ou que estão em estágios iniciais de pesquisa. Fonte: (Costa et al., 2006)

O segundo grupo de características biométricas são chamadas de comportamentais (ou dinâmicas). São características aprendidas ou desenvolvidas através da utilização constante, e que podem variar fortemente ao longo do tempo. Além disso, podem ser facilmente alteradas pela vontade ou estado do usuário. Assim, até mesmo duas amostras consecutivas podem mudar bastante. As principais características dinâmicas utilizadas são o padrão de voz e a dinâmica da assinatura (Costa et al., 2006).

Outras características dinâmicas também são utilizadas em menor grau ou estão em estágios iniciais de pesquisa, como dinâmica de digitação (*keystroke dynamics*) (Bergadano et al., 2002), modo de andar (Phillips et al., 2002), movimento labial (BioID, 2005; Valid, 2006), som da assinatura (Ling, 2003), vídeo da assinatura (Fink et al., 2001) e imagens mentais (*pass-thoughts*) (Thorpe et al., 2005).

Para uso dessas características para fim de reconhecimento biométrico, entretanto, mais alguns critérios devem ser atendidos: o processo de identificação através dessa característica deve possuir um desempenho aceitável, no que diz respeito à precisão e velocidade; é preciso saber até que ponto as pessoas aceitam determinado identificador

biométrico em suas vidas, disponibilizando determinada característica sua para identificação; é preciso observar também o quão fácil é enganar o sistema com técnicas fraudulentas (Jain et al., 2004).

Além de introduzir os principais conceitos sobre biometria, este Capítulo visa também definir os sistemas biométricos, seus tipos e aplicações; e definir o traço biométrico utilizado nessa tese: a geometria da mão.

2.2.1 Sistemas Biométricos

Segundo Jain et al. (2004), um sistema biométrico é essencialmente um sistema de reconhecimento de padrões que opera através da aquisição de dados biométricos de um indivíduo, extraindo um conjunto de características a partir dos dados adquiridos, e comparando esse conjunto de características com o conjunto modelo no banco de dados. Uma vez que os identificadores biométricos não podem ser facilmente extraviados, forjados, ou compartilhados, métodos de identificação biométricos são considerados mais confiáveis do que métodos baseados em *tokens* (como *smartcards*) ou senhas (Maltoni, 2009). Assim, os sistemas de reconhecimento biométrico estão sendo cada vez mais implantados em um grande número de aplicações governamentais, civis e forenses.

São diversos os tipos de sistemas biométricos existentes. A Figura 7 apresenta um gráfico de comparação de alguns desses tipos, considerando segurança *versus* conveniência.



Figura 7: Posição de alguns sistemas biométricos em um gráfico de segurança versus conveniência. Fonte: (Kulp & Braskamp, 2012).

Os sistemas biométricos têm diversas utilidades, entre elas, o aumento na segurança, prevenção do roubo de identidade, identificação e verificação de indivíduos e proteção de

ativos nacionais e organizacionais, além de autenticação para utilização de aplicativos (Rahman et al., 2008).

De uma maneira geral, os sistemas biométricos podem ser divididos em aplicações de nicho Governamental, Comercial e Forense que podem ser extendidas para a classificação por finalidade. O relatório *BITE Market Report* (BITE, 2005) dividiu essas finalidades em sete grupos e mensurou a utilização de cada uma delas. Tais dados podem ser vistos na Tabela 1.

Tabela 1: Distribuição por finalidade das principais aplicações biométricas. Fonte: (BITE, 2005)

Finalidade	Utilização
Identificação Criminal	28%
Controle de acesso e atendimento	22%
Identificação Civil	21%
Segurança de redes e de computadores	19%
Autenticação em ponto de vendas, ATM's e varejo	4%
Autenticação telefônica e comércio eletrônico	3%
Vigilância e filtragem	3%

Um sistema biométrico pode ser encarado como um sistema de reconhecimento de padrões de propósito específico (Bolle et al., 2002). Independente da finalidade, seu modelo conceitual leva em consideração os dados e processos básicos comuns a qualquer sistema biométrico. Em uma primeira utilização, o usuário é previamente registrado e seu perfil biométrico fica armazenado. Na utilização posterior do sistema, o processo de aquisição obtém os dados biométricos apresentados e características particulares dos dados são extraídas para comparação com o perfil armazenado. O processo de comparação decide se os dados apresentados são suficientemente similares ao perfil registrado (Costa et al., 2006).

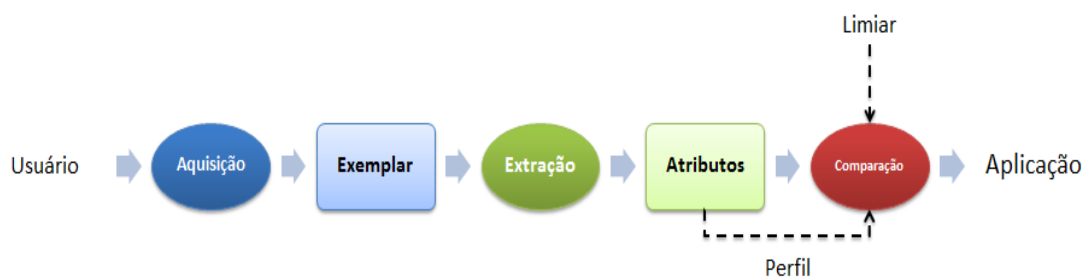


Figura 8: Um modelo simples de sistemas biométricos. Adaptado de Costa et al. (2006)

(Costa et al., 2006) define cada etapa do modelo como segue:

- **Aquisição e exemplar:** O processo de aquisição ou apresentação é o processo de obtenção dos dados da característica biométrica oferecida. Normalmente, a dificuldade deste processo é balancear adequadamente a qualidade da amostra sem causar excesso de inconveniência para o usuário. Neste módulo é geralmente embutido um controle da qualidade da amostra adquirida (viabilidade de processamento). O exemplar ou amostra (*sample*) é o resultado do processo de aquisição.
- **Extração e atributos:** O processo de extração produz uma representação computacional do exemplar obtido, que chamaremos de atributos, ou características extraídas (*features*). A extração de características é a redução de um conjunto de medidas formado por uma grande quantidade de dados que contém uma pequena quantidade de informação útil para um conjunto que contém menos dados, mas praticamente a mesma quantidade de informação (Patrick, 1972).
- **Registro e perfil:** O processo de registro, ou *enrollment*, obtém previamente os dados biométricos do usuário para cadastramento no sistema. O perfil biométrico obtido, ou *template*, é armazenado para uma comparação posterior. O processo de registro é necessário para o estabelecimento do perfil para posterior comparação.
- **Comparação, limiar e decisão:** O processo de comparação, ou *matching*, verifica qual é o grau de similaridade entre as características extraídas da amostra do usuário e o perfil armazenado previamente. Este processo fornece um escore representativo da similaridade entre os dois conjuntos de dados e segue a mesma lógica da binarização de imagens. Caso a similaridade seja superior a certo limite previamente determinado, conhecido como limiar, ou *threshold*, a decisão é aceitar o usuário, ou seja, uma autenticação válida. Caso a similaridade seja inferior ao limiar, a decisão é não aceitar o usuário, e então temos um usuário não autenticado.

Neste trabalho, por enquanto, somente a geometria da mão foi utilizada para o reconhecimento biométrico, mas espera-se utilizar também a impressão palmar. Detalhes e diferenças acerca desses traços são particularizados nas seções a seguir.

2.2.2 Impressão Palmar

Impressão palmar (em inglês, *palmprint*), é caracterizada como a superfície interna da palma da mão – localizada entre os dedos e os pulsos – que normalmente possui três linhas

principais (denominadas linhas de flexão), vincos e cumes (doravante denominados *ridges*) (Kong et al., 2009). A Figura 9 mostra exemplos de impressões palmares onde pode-se observar a presença dessas características. As impressões palmares são únicas em cada indivíduo, sendo diferentes, inclusive, entre gêmeos univitelinos. A informação extraída da palma da mão é estável desde o nascimento, e permite distinguir indivíduos de uma população grande. Elas também fornecem atributos ricos e variados, incluindo textura, papilas dérmicas e vincos. Tais características as tornam interessantes como forma de identificação de pessoas.

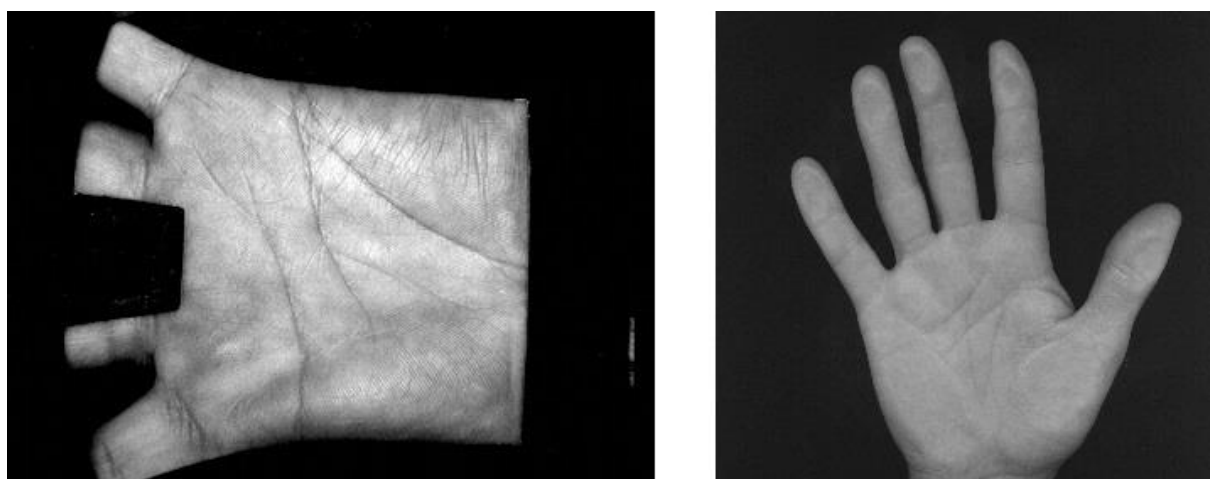


Figura 9: Exemplos de Impressões Palmares

Na área de impressão palmar, um típico sistema de reconhecimento consiste de cinco partes: um *scanner* de impressão palmar, pré-processamento, extração de características, classificador e o banco de dados. O *scanner* de impressão palmar coleta as imagens palmares. O pré-processamento define o sistema de coordenadas para ajustes das imagens palmares e realiza a segmentação para a extração de características que, por sua vez, obtém as características efetivas das imagens pré-processadas. O classificador relaciona duas impressões palmares e o banco de dados armazena os *templates* (Kong et al., 2009).

Como nova característica biométrica, impressões palmares têm diversas vantagens em comparação com outras características disponíveis: imagens de baixa resolução podem ser capturadas; dispositivos de captura de baixo custo podem ser usados; é muito difícil ou impossível falsificar uma impressão palmar; e as linhas da impressão palmar são estáveis (Wu et al., 2004).

Deve-se destacar que o reconhecimento por impressão palmar constitui uma abordagem biométrica completamente diversa do reconhecimento por geometria da mão. De

fato, os problemas relativos ao processo guardam semelhanças muito mais marcantes com os de reconhecimento por impressão digital. De qualquer modo, as informações fornecidas pela impressão palmar são diferentes daquelas fornecidas pela impressão digital e pela geometria da mão.

2.2.3 Geometria da Mão

Segundo Jain et al. (2004), sistemas de reconhecimento pela geometria da mão são baseados nas medições feitas a partir da mão humana, incluindo sua forma, tamanho da palma da mão, comprimentos e larguras dos dedos. Ademais, a espessura e curvatura dos dedos também podem ser utilizadas para reconhecimento. Tais características podem ser observadas na Figura 10.



Figura 10: Medidas típicas da geometria da mão. O modelo esquemático (esquerda) pode ser apreciado na imagem real (centro) obtida de um dispositivo (direita)

No processo de aquisição da geometria da mão, em geral, o dispositivo leitor de geometria da mão utiliza uma câmera para capturar imagens em preto e branco da silhueta da mão. Não são registrados detalhes a respeito de textura, impressões digitais, linhas e cores. Quando combinadas com um refletor e espelhos laterais, duas imagens distintas são produzidas, sendo uma vista de cima e outra de lado. Este método é conhecido como orto-leitura (Figura 11).



Figura 11: Dispositivo de captura guiado por pinos. Fonte: (Subcommittee on Biometrics of National Science and Technology Council (NSTC), 2006).

O processo de extração trabalha sobre a imagem adquirida. A imagem obtida é convertida para preto e branco, caso seja colorida, e pequenos desvios eventuais são corrigidos. Para estes ajustes, em alguns casos, são úteis as imagens dos pinos existentes na plataforma. Um algoritmo de detecção de bordas, em geral, é aplicado para extrair o contorno da mão (Costa et al., 2006).

No processo de comparação, a representação obtida é comparada com o perfil armazenado. A comparação pode envolver, por exemplo, acumulação de diferenças absolutas nas características individuais, entre a representação de entrada e o perfil armazenado. Para o cálculo da similaridade entre os dois vetores, geralmente são utilizados algoritmos baseados em distância euclidiana, distância de Hamming, ou redes neurais (Costa et al., 2006).

Dentre as vantagens do uso da geometria da mão para reconhecimento biométrico, (Ross et al., 2006) citam que a técnica utilizada é simples, relativamente fácil de usar e barata, o que fez com que sistemas comerciais de verificação baseados na geometria da mão fossem instalados em centenas de locais em todo o mundo. Fatores ambientais, como tempo seco ou anomalias individuais, como pele seca, não afetam adversamente tais sistemas. Além disso, a coleta das características é considerada fácil e não intrusiva.

Entretanto, de acordo com Ross et al. (2006) a geometria da mão não é considerada como uma característica distintiva, o que dificulta a utilização dos sistemas de reconhecimento baseado na geometria da mão para uma ampla população. Ademais, a informação da geometria da mão não é invariante durante o período de crescimento de uma

criança e a presença de joias (por exemplo, anéis) ou limitações em destreza (por exemplo, artrites) podem representar desafios na correta extração de informações da geometria da mão.

2.2.4 Autenticação versus Identificação

Os sistemas biométricos são usados para a autenticação de pessoas. Nestes sistemas, existem dois modos de autenticação: a verificação e a classificação (Bolle et al., 2004).

Na verificação, a característica biométrica é apresentada pelo usuário juntamente com uma identidade alegada, usualmente por meio da digitação de um código de identificação. Esta abordagem de autenticação é dita uma busca um-para-um (1:1), ou busca fechada, em um banco de dados de perfis biométricos e é tipicamente utilizada em aplicações de reconhecimento positivo, cujo objetivo é prevenir que múltiplas pessoas usem a mesma identidade (Jain et al., 2004). O princípio da verificação está fundamentado na resposta à questão: “O usuário é quem alega ser?”.

Na classificação, por sua vez, dada uma característica biométrica pelo usuário, o sistema busca todos os registros do banco de dados e retorna uma lista de registros com características suficientemente similares à característica biométrica apresentada. A lista retornada pode ser refinada posteriormente por comparação adicional, biometria adicional ou intervenção humana (Costa et al., 2006). Agora, comparações do tipo um-para-muitos (1:N) são realizadas. Por conta disso, geralmente a precisão do sistema diminui com o aumento da base de dados. De acordo com Jain et al. (2004), a classificação é um componente crítico em aplicações de reconhecimento de negativos, onde o sistema verifica se uma pessoa é quem ela nega ser. O propósito de reconhecimento negativo é prevenir que uma pessoa tenha múltiplas identidades. Basicamente, a classificação corresponde a responder à questão: “Quem é o usuário?”.

Por fim, ainda há outro conceito utilizado em sistemas biométricos: a identificação. Nesta, o objetivo do sistema é determinar se um dado *template* está presente na base de dados utilizada e, em caso afirmativo, definir a quem pertence esse *template* através da classificação. A identificação, além de não ser comumente implementada em sistemas biométricos, é também considerada um problema mais difícil, pois o sistema agora deve verificar se a amostra em questão é diferente de todas as demais, já que ela não foi cadastrada previamente. Na identificação, o sistema deve responder à questão: “O usuário está cadastrado?”.

2.2.5 Métricas de Erro

O desempenho da verificação de sistemas biométricos geralmente é medido utilizando métricas como taxas de *Falsa Aceitação* (do inglês, *False Acceptance Rate*, FAR), *Falsa Rejeição* (do inglês, *False Rejection Rate*, FRR) e *Erro Igual* (EER). Falsa aceitação ocorre quando um usuário não registrado obtém o acesso ao sistema protegido, enquanto falsa rejeição refere-se quando um usuário registrado não consegue obter o acesso legítimo. A EER é geralmente utilizada para comparar o desempenho de diferentes sistemas biométricos (Zang, 2004) e geralmente é calculada pelo ponto de interseção entre as curvas FAR e FRR (Figura 12). Além disso, o sistema pode operar nas faixas de “conveniência” ou de “segurança”, conforme calibração do limiar. Todas essas medidas são inversamente proporcionais em relação à qualidade do desempenho do sistema biométrico.

De acordo com Ross et al. (2006), as taxas FAR e FRR podem ser calculadas pelas Equações 1 e 2 a seguir:

$$FAR(\eta) = \int_{\eta}^{\infty} p(s|impostor) ds \quad (1)$$

$$FRR(\eta) = \int_{-\infty}^{\eta} p(s|genuíno) ds \quad (2)$$

Onde $p(s|impostor)$ e $p(s|genuíno)$ representam distribuições de probabilidade de um escore s sobre condições de genuíno e impostor, respectivamente, e η é um limiar que define se um indivíduo é genuíno (se $s > \eta$) ou impostor².

² Em geral, o cálculo do escore em sistemas biométricos é diretamente proporcional às amostras, isto é, quanto maior o valor de escore entre duas amostras, mais parecidas elas são consideradas. Neste trabalho, no entanto, o cálculo do escore é dado por uma métrica de similaridade onde quanto menor o valor de escore entre duas amostras, maior o grau de similaridade entre elas. Isso inverte a condição de genuíno e impostor a respeito do limiar como também as curvas de FAR e FRR, porém optou-se por utilizar as definições das referências em questão.

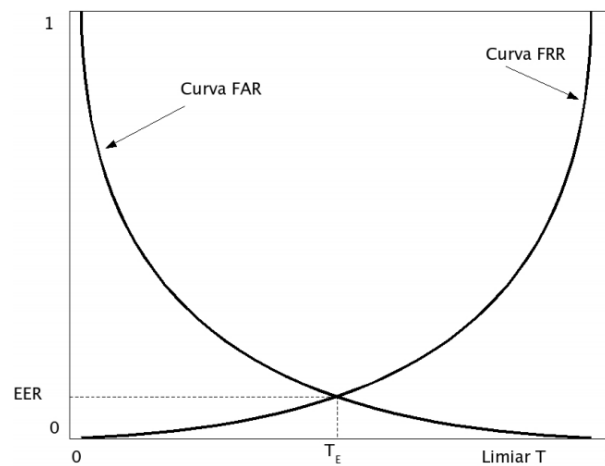


Figura 12: As curvas típicas das taxas de erro FAR e FRR, plotadas uma ao lado da outra, em relação ao limiar T configurado para o sistema. A EER é representada pelo ponto de interseção das curvas. Fonte: (Costa et al., 2006).

É importante destacar que o ajuste do limiar permite a calibragem do sistema, mas também implica consequências opostas. Sistemas onde brechas na segurança devem ser evitadas ao máximo (e.g., sistemas bancários), ajustam o limiar para que a FAR seja mínima. No entanto, isso pode resultar em problemas de conveniência, pois usuários genuínos poderão ter seu acesso negado até que o sistema tenha a total certeza sobre o perfil do usuário. Por outro lado, sistemas que buscam minimizar a FRR no sentido de melhorar a conveniência de utilização por parte do usuário, poderão ter problemas de segurança, autorizando o acesso de usuários não legítimos.

Outra maneira de se comparar as taxas FAR e FRR é produzindo uma curva bidimensional conhecida como *Característica de Operação do Receptor* (do inglês, *Receiver Operating Characteristic, ROC*). Diferente da forma apresentada na Figura 12, essa abordagem compara as taxas FAR e $1 - \text{FRR}$ uma contra a outra, permitindo visualizar o desempenho de um sistema biométrico independentemente do valor de *threshold* (Figura 13). Entretanto, ao comparar dois ou mais sistemas biométricos através dessa curva, a área sob a curva (do inglês, *Area Under the Curve, AUC*) ROC é usualmente utilizada. Nessa métrica, quanto maior a área sob a curva ROC produzida, melhor o desempenho do sistema (Maltoni, 2009).

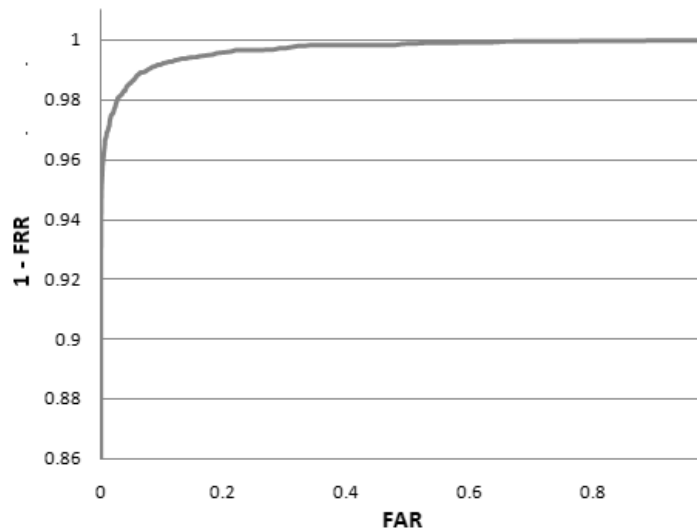


Figura 13: Curva ROC. As taxas de erro FAR e (1-FRR) podem ser plotadas uma contra outra numa curva bidimensional.

Por outro lado, o desempenho da classificação de sistemas biométricos é usualmente dado pela taxa de acerto (TA) dos classificadores utilizados. Em geral, essa taxa de acerto é dada pela seguinte equação:

$$TA = \frac{TP}{n} * 100\% \quad (3)$$

Onde TP é o número de verdadeiros positivos e n é o número de amostras.

Na próxima subseção, são apresentados os principais conceitos sobre algoritmos genéticos, suas aplicações e os fatores que envolvem o projeto de algoritmos genéticos.

2.3 Algoritmos Genéticos

Os algoritmos genéticos (GA) foram inicialmente propostos por Holland (1962) como uma ferramenta de busca de soluções em problemas de otimização. Eles são baseados no princípio da seleção natural descrito por Charles Darwin (1951). Em seu livro, Darwin descreve a aptidão basicamente como a capacidade de sobreviver. E, embora este princípio pareça ser ligeiramente tautológico, ele faz todo sentido no mundo de problemas de otimização, onde a aptidão de um “indivíduo” é dada pelo valor da função a ser otimizada.

No contexto da computação, os algoritmos genéticos são, em geral, simples de se descrever e programar, embora seu comportamento em tempo de execução possa se tornar complicado de ser observado. Por conta disso, muitos trabalhos têm sido feitos para explicar como e por que algoritmos genéticos funcionam e para quais aplicações eles são adequados (Holland, 1975; Golberg, 1989; Banzhaf, 1999; Whitley, 1993). Em termos gerais, acredita-se que os GAs funcionem pela descoberta, ênfase e recombinação de boas soluções que tendem a ser feitas de boas soluções anteriores.

Neste Capítulo, os principais conceitos acerca dos algoritmos genéticos são abordados. Além disso, um pseudocódigo de algoritmo genético é apresentado e exemplos da aplicação dos GA também são mostrados.

2.3.1 Conceitos

Tipicamente, um algoritmo genético trabalha numa população de *indivíduos*. Cada indivíduo é representado por um ou mais *cromossomos* formados por um conjunto de genes representando os parâmetros a serem otimizados. Certas operações são realizadas com o objetivo de produzir novas gerações de indivíduos baseado nas suas capacidades de gerar melhores resultados, são elas:

- **Cruzamento:** é a operação chave para gerar novos indivíduos na população. Inspirado pelo exemplo da natureza, o cruzamento é projetado para juntar material genético de cromossomos com um alto valor de *fitness* a fim de produzir indivíduos ainda melhores. Usualmente, os algoritmos genéticos implementam 3 tipos de cruzamento (Figura 14), tais quais:
 - **Um ponto:** no cruzamento de um ponto, os cromossomos pais são divididos aleatoriamente em duas partes pelo mesmo gene e, então, os cromossomos filho são gerados pela junção das partes direita de um pai com a parte esquerda do outro pai.
 - **Dois pontos:** semelhante ao de cruzamento de um ponto, sendo que, nesse caso, cada cromossomo pai é dividido em 3 partes pela escolha aleatória de 2 genes. O processo de geração do cromossomo filho também é semelhante ao cruzamento de um ponto.
 - **Uniforme:** o cruzamento uniforme realiza randomicamente o cruzamento dos cromossomos pais na criação dos filhos. Primeiramente, os

cromossomos pais são divididos em duas ou mais partes aleatórias na primeira geração, cuja divisões são repetidas nas próximas gerações. Os filhos, então, são criados pela união da parte genética dos pais.

- **Seleção:** o operador de seleção destina-se a implementar a ideia de “sobrevivência do mais apto”. Esse operador basicamente determina qual dos cromossomos na população atual tem permissão para herdar o seu material genético na próxima geração. Há 4 tipos usuais de métodos de seleção:
 - **Roleta:** nesse tipo de seleção, os cromossomos são selecionados de acordo com seu *fitness*. Quanto mais apto o cromossomo, maior é a chance de ele ser selecionado para o cruzamento.
 - **Classificação:** na seleção por classificação, os cromossomos são, primeiramente, classificados por seu valor de *fitness*. Assim, o pior cromossomo recebe o valor de *fitness* igual a 1, o segundo pior recebe 2, e assim por diante até que o melhor cromossomo tenha valor de *fitness* igual ao número de cromossomos na população. Esse método é mais indicado quando há grandes diferenças entre os valores de *fitness* dos cromossomos, porém pode resultar em menor convergência, pois os melhores cromossomos não se distinguem muito uns dos outros.
 - **Estado estacionário:** diferente dos anteriores, esse método não é voltado particularmente à seleção de pais. Nesse caso, a ideia é garantir que uma nova população tenha uma grande parte de cromossomos que sobreviverão a próxima geração. Em cada nova geração, alguns poucos bons cromossomos (alto valor de *fitness*) são selecionados para criação da descendência, enquanto cromossomos de baixa aptidão são removidos e novos são colocados em seus lugares. Todo o resto da população sobrevive para a próxima geração.
- **Mutação:** o operador de mutação altera aleatoriamente um ou mais genes de um cromossomo para mudar aleatoriamente a descendência criada pelo cruzamento. Isso permite ao GA encontrar soluções que contenham genes com valores não existentes na população inicial. O parâmetro que regula esse operador é chamado *probabilidade de mutação*. Caso o operador de seleção reduza a diversidade na população, o operador de mutação eleva a diversidade novamente. Além disso, quanto maior a probabilidade de mutação, menor é a chance de convergência prematura. Entretanto, a alta probabilidade de mutação faz com que o GA torne-se

um algoritmo de busca aleatória (Dawid, 1996). Um exemplo de operação de mutação pode ser visto na Figura 15.

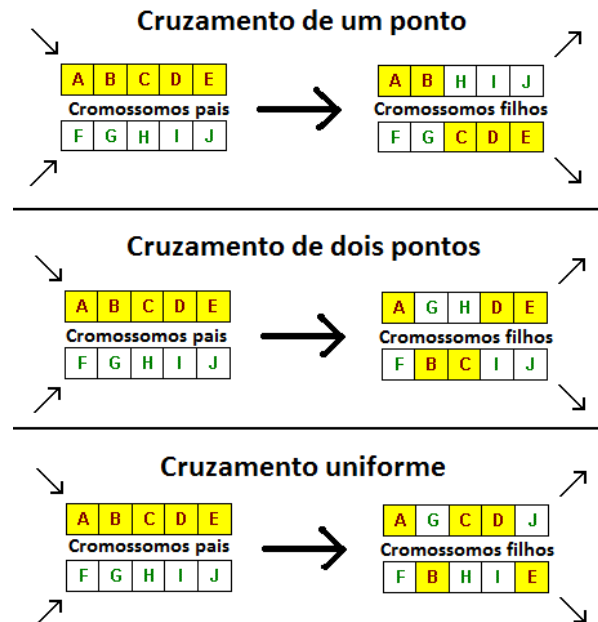


Figura 14: Exemplificação dos tipos de cruzamento implementados nos algoritmos genéticos.

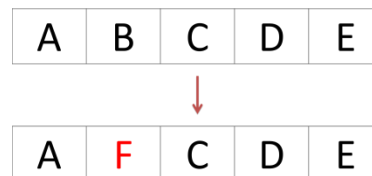


Figura 15: Exemplo da operação de mutação.

Nos algoritmos genéticos, a seleção dos indivíduos é dada por uma função denominada *função de fitness*, responsável por atribuir um escore (*fitness*) para cada cromossomo na população atual (Mitchell, 1995). O *fitness* de um cromossomo depende de quão bem tal cromossomo resolve o problema. Em geral, os algoritmos genéticos são implementados visando à maximização dessa função. No entanto, quando o objetivo é minimizar tal valor, geralmente retorna-se o valor negativo que iria ser retornado pela função de *fitness*. Por convenção, o melhor cromossomo de todas as gerações, isto é, que gera o melhor valor de *fitness*, é chamado de cromossomo ótimo.

2.3.2 Algoritmo

Seja P uma população aleatória de N cromossomos $(c_1, c_2, c_3, \dots, c_N)$ e $f(c)$ a função de *fitness* para um cromossomo c . O seguinte pseudocódigo descreve os passos gerais dos algoritmos genéticos:

- (1) Crie uma população aleatória P de N cromossomos (soluções candidatas para o problema).
- (2) Avalie $f(c)$ para cada cromossomo c na população.
- (3) Gere uma nova população pela repetição dos seguintes passos até que a nova população atinja o tamanho N :
 - a. Selecione dois cromossomos pais da população dando preferência aos de maior valor de *fitness*. Automaticamente copie o melhor cromossomo para a próxima geração.
 - b. Dada a probabilidade de cruzamento, faça o cruzamento de dois cromossomos pais para gerar dois novos filhos. Se o cruzamento não foi executado, os cromossomos filhos são exatamente a cópia dos pais.
 - c. Dada a probabilidade de mutação, randomicamente altere um ou mais genes no cromossomo filho.
 - d. Copie o valor do cromossomo filho na nova geração.
- (4) Copie a nova geração criada sobre a população anterior
- (5) Se a condição de término do *loop* é satisfeita, então pare e retorne a melhor solução na população atual. Caso contrário, vá para o passo (2).

O simples algoritmo descrito acima é a base para a maior parte de aplicações dos algoritmos genéticos. No entanto, existem versões mais complexas de algoritmos genéticos, como aqueles que possuem diferentes tipos de operadores de cruzamento e mutação (Mitchell, 1995).

2.3.3 Aplicações

O trabalho de Mitchell (1995) cita que as variações do algoritmo básico dos GAs podem ser aplicadas em diversos campos científicos, dentre eles:

- **Otimização:** algoritmos genéticos têm sido utilizados em uma ampla variedade de tarefas de otimização, incluindo otimizações numéricas e problemas

combinacionais como *layout* de circuitos e agendamento de trabalho (Mazumder & Rudnick, 1999).

- **Programação automática:** GAs têm sido usados para desenvolver programas de computador com propósito específico e projetar outras estruturas computacionais como triagem de redes e autômatos celulares (Mitchell et al., 1996).
- **Aprendizagem de máquina:** algoritmos genéticos também são aplicados em diversas aplicações de aprendizagem de máquina como, por exemplo, previsão do tempo ou estrutura de proteínas (Pedersen & Moult, 1996). Além disso, também se aplica GAs na computação de pesos de redes neurais (Leung et al., 2003); sistemas de classificadores de aprendizagem (Booker et al., 1989); e sensores robóticos (Manikas et al., 2007).
- **Modelos econômicos:** nesse campo, os GAs são utilizados na inovação de modelos de processos, desenvolvimento de estratégias de licitação e no surgimento de mercados econômicos (Richter Jr & Sheblé, 1998).
- **Interações entre evolução e aprendizagem:** algoritmos genéticos são envidados para estudar como aprendizagem de indivíduos e evoluções de espécies afetam uns aos outros (Ackley & Littman, 1991).
- **Modelos de sistemas sociais:** GAs também são aplicados nos estudos de aspectos evolucionários de sistemas sociais como o comportamento de formigas no seguimento de trilhas (Dorigo & Gambardella, 1997), por exemplo.

A lista acima não mostra todas as aplicações onde os algoritmos genéticos podem ser utilizados, mas apresenta alguns exemplos onde os GAs são aplicados tanto na solução de problemas quanto em contextos científicos.

2.3.4 *Influência dos Parâmetros dos Algoritmos Genéticos*

A utilização de algoritmos genéticos envolve a escolha de diversos fatores, entre eles, o número de cromossomos, o número de gerações, a probabilidade de mutação, a probabilidade de cruzamento. A escolha correta desses fatores é crucial para que o algoritmo genético possa ser capaz de encontrar o resultado esperado.

A escolha do número de indivíduos é importante para a variabilidade da população. Quanto maior o número de indivíduos, espera-se que maior seja a diferença entre eles, já que os N indivíduos iniciais são, em geral, criados de forma aleatória. No entanto, um elevado

número de indivíduos pode piorar o desempenho dos algoritmos genéticos, visto que mais cruzamentos e mutações podem ser efetuados. Valores entre 100 e 1000 são frequentemente usados em trabalhos da literatura (ver Capítulo sobre Trabalhos Relacionados).

O número de gerações, por sua vez, é crucial para garantir que os algoritmos genéticos não converjam para algum mínimo local no espaço de resultados. Quando o número de gerações é consideravelmente baixo, a variabilidade da população pode não ser suficiente para que o algoritmo genético explore de forma mais satisfatória tal espaço. Por outro lado, quanto maior o número de gerações, maior o tempo despendido pelo algoritmo genético. Embora não seja possível garantir que os algoritmos genéticos encontrem o melhor resultado (máximo/mínimo global, a depender do caso) dentro do espaço de resultados, uma vez que ele nem sempre é totalmente conhecido, em geral, cem gerações são suficientes para que o algoritmo genético convirja para o resultado considerado ótimo. Por esse motivo, esse resultado encontrado é normalmente denominado de resultado ótimo.

As probabilidades de cruzamento e mutação também são importantes na diversidade da população, visto que elas são responsáveis pela geração de novos indivíduos. Uma baixa probabilidade desses dois fatores pode impedir o algoritmo genético de explorar o espaço de resultados, convergindo prematuramente. Por outro lado, a alta probabilidade de mutação, como dito anteriormente, pode tornar o algoritmo genético um simples algoritmo de busca aleatória. Neste trabalho, diferentes probabilidades foram testadas para verificação desses fatores.

Há ainda outros fatores que influenciam na execução dos algoritmos genéticos como, por exemplo, a escolha do tipo de cruzamento. Porém, os fatores citados acima são considerados os mais importantes no projeto de algoritmos genéticos. Este trabalho buscou combinar diferentes variações de parâmetros nas execuções dos GA nos testes executados. Tais parâmetros, assim como os testes executados, são apresentados no Capítulo 4. O próximo Capítulo apresenta os principais trabalhos relacionados ao método proposto.

Capítulo

3

Trabalhos Relacionados

Nesse Capítulo são descritos alguns trabalhos relacionados com a presente proposta. Primeiramente, trabalhos relacionados aos algoritmos genéticos são citados, seguidos por trabalhos que os aplicam em sistemas biométricos.

John Holland (1962) foi o primeiro a descrever a ideia de algoritmos evolucionários (ou evolutivos). Tal trabalho provê a base do objetivo de Holland sobre o entendimento da adaptação da vida, como isso ocorre na natureza e as maneiras de se desenvolver sistemas computacionais baseados nesses princípios.

Uma significativa fundamentação teórica a respeito de algoritmos evolucionários para otimização pode ser encontrada no trabalho de Back & Schwefel (1993). Esse trabalho fornece uma visão geral dos três principais ramos dos algoritmos evolutivos: estratégias evolutivas, programação evolutiva e algoritmos genéticos. Além disso, conceitos sobre o esquema de representação de variáveis, mutação, recombinação e operadores de seleção também são abordados.

Em Behroozmand & Almasganj (2007) é descrito um método de seleção ótima de características para sinais de voz de pessoas com paralisia vocal unilateral. O GA é utilizado para encontrar um conjunto ótimo de atributos que maximizem a taxa de reconhecimento através de um classificador baseado na máquina de vetor de suporte (em inglês, *Support Vector Machine*, SVM). Os resultados mostram que o atributo de entropia, em comparação com o de energia, demonstra uma descrição mais eficiente da patologia da voz e fornece uma ferramenta de diagnóstico de paralisia da laringe.

Considerando a importância da otimização de parâmetros nos sistemas biométricos, o trabalho de Goranin & Cenys (2010) analisa a aplicação dos algoritmos genéticos nesse contexto. De acordo com os autores, o uso de algoritmos genéticos pode garantir um aumento

qualitativo e quantitativo das medidas dos sistemas biométricos, como velocidade, taxa de erro e flexibilidade.

O banco empregado na validação deste trabalho, BioPass-UFPB, já foi utilizado no trabalho de do Nascimento et al. (2014). Tal trabalho aplica e compara diferentes métodos de classificação no reconhecimento biométrico baseado na geometria da mão. Utilizando os atributos extraídos pelos autores, os resultados mostram-se competitivos quando comparados a outros métodos considerados estado da arte.

Entre os trabalhos mais parecidos com o método proposto, pode-se citar Luque et al. (2011). Neste trabalho, entre características geométricas e outros tipos de descritores, 403 características são extraídas das imagens palmares de acordo com o método descrito pelos autores. Utilizando a base de dados IIT Delhi Palmprint Image Database (137 indivíduos, com seis amostras/indivíduo), sendo 30% das amostras para treinamento e 70% para validação através dos classificadores SVM, *k*-vizinhos mais próximos (do inglês, *k-Nearest Neighbour*, *k-NN*), e análise discriminante linear (do inglês. *Linear Discriminant Analysis*, *LDA*), os autores conseguiram reduzir as 403 características extraídas para apenas 25 mantendo a taxa de acerto próxima a 97,5% com a aplicação do GA. Os resultados do GA mostraram-se superior aos outros seletores de características utilizados pelos autores, cujas taxas de acerto não foram superior a 75%.

Há também o trabalho de Jin et al. (2008). Neste trabalho, a Transformada Discreta de Onda (do inglês, *Discrete Wavelet Transform*, *DWT*) é aplicada sobre uma região de interesse (do inglês, *Region of Interest*, *ROI*) da palma da mão que é, então, dividida em 16 blocos de tamanho 16x16 e, por fim, 64 valores são extraídos de todos os blocos e utilizados para representar cada indivíduo. Trezentas imagens palmares de 30 pessoas diferentes são empregadas na validação do método dos autores, onde 5 amostras/usuário são reservadas para treinamento. Utilizando o algoritmo genético como seletor de características e aplicando o SVM como classificador, os autores alcançaram até 100% de Taxa de Casamento Correto (do inglês, *correct match rate*, *CMR*).

O trabalho de Du et al. (2009) apresenta um sistema de verificação de impressões palmares específico para cada usuário. Os algoritmos genéticos são aplicados para encontrar parâmetros ótimos do filtro de Gabor (orientação e frequência) na região de interesse da palma. Assim, parâmetros ótimos são calculados no treinamento para cada usuário e, no

momento da verificação, os parâmetros são informados para extração das características e classificação pelo SVM. De acordo com os autores, houve redução da taxa de erro e melhoria significativa da performance do SVM.

O método apresentado no trabalho de Luque-Baena et al. (2013) também foca na seleção de características de imagens palmares utilizando os algoritmos genéticos. Utilizando um conjunto de dados com 150 indivíduos com 10 imagens/indivíduo e aplicando o LDA para classificação e verificação, os autores atingiram 100% de sucesso na taxa de acerto e taxa de erro igual (do inglês, *Equal Error Rate*, EER) superiores a 4,5% nos bancos utilizados.

Como se pode ver nos trabalhos acima, a utilização dos GAs vem mostrando bons resultados na classificação de imagens palmares. No entanto, a maior parte dos trabalhos citados foca somente na classificação. De fato, apenas o trabalho de Luque-Baena et al. (2013) aplica os algoritmos genéticos tanto na classificação quanto na verificação de sistemas biométricos da geometria da mão. O presente trabalho, por sua vez, busca analisar a aplicação dos GAs em ambos os casos.

Capítulo

4

Materiais e Métodos

Este Capítulo descreve os recursos utilizados e os métodos desenvolvidos. Inicialmente, o ambiente de desenvolvimento é apresentado, seguido pela descrição das amostras utilizadas. Afinal, os métodos desenvolvidos são anunciados e aprofundados.

4.1 Ambiente de Desenvolvimento

Foi utilizada a IDE *Netbeans* com a linguagem de programação Java®, versão 1.8.0, instalada em um computador com processador Intel® Core™ i5-2310 2.9 GHz, 4GB de memória RAM DDR3, sistema operacional Microsoft® Windows 7 *Ultimate* de 64 bits *Service Pack 1*. O Matlab® foi utilizado para cálculo da AUC através da função *trapz*.

Para aplicação dos algoritmos genéticos, a biblioteca *GALib* (Smith, 2013) foi utilizada. Essa biblioteca aplica o algoritmo genético para uma dada população de cromossomos, um número de gerações, a probabilidade de mutação, a probabilidade de cruzamento, um número mínimo e máximo de execuções preliminares, a chance de seleção randômica e o número de casas decimais de precisão. Além disso, também é possível escolher se a função de *fitness* deve usar somente valores positivos ou não.

4.2 Medida de Similaridade

Para se calcular a similaridade entre os atributos extraídos de dois *templates* da base de dados empregada neste trabalho, optou-se pela ponderação da distância de Manhattan (também conhecida como distância L_1) que pode ser vista na seguinte equação:

$$d_{A,B} = \begin{cases} \sum_{i=1}^N w_i * |a_i^A - a_i^B| & , \quad N = 85 \\ \sum_{i=1}^{N-1} w_i * |a_i^A - a_i^B| + w_N * (1 - Corr(H_A, H_B)) & , \quad N = 99 \end{cases} \quad (4)$$

Onde, w_i é i -ésimo coeficiente multiplicador (peso); a_i^A e a_i^B representam o i -ésimo atributo dos *templates* A e B , respectivamente; $N \in \{85,99\}$ é o número de atributos (ver Seção 4.3); $Corr(H_A, H_B)$ representa a Correlação dos Histogramas do LBP de A e B ; e $d_{A,B}$ representa o grau de similaridade entre dois *templates* A e B ; Quanto menor o escore entre dois *templates*, maior é o grau de semelhança entre eles.

A correlação dos histogramas é utilizada para calcular o grau de semelhança entre os histogramas gerados pelo LBP aplicado em cada *template*. A correlação retorna um valor $\in [-1,1]$ e é definida pela seguinte equação:

$$Corr(H_A, H_B) = \frac{\sum_I (H_A(I) - \overline{H_A})(H_B(I) - \overline{H_B})}{\sqrt{\sum_I (H_A(I) - \overline{H_A})^2 \sum_I (H_B(I) - \overline{H_B})^2}} \quad (5)$$

Onde $\overline{H_A}$ e $\overline{H_B}$ representam as médias dos histogramas supracitados. Vale ressaltar que a correlação de histograma só é aplicada quando os atributos de textura (Haralick e LBP) são incluídos. Mais detalhes serão explicados na Seção 4.5.

4.3 Base de dados BioPass-UFPB

Para o presente trabalho, a base de dados BioPass-UFPB foi utilizada para teste e validação do método proposto. A BioPass-UFPB é uma base de dados multibiométrica formada por imagens de impressão digital, impressão palmar e geometria da mão. A base foi construída pelo autor deste trabalho em parceria com a empresa Vsoft Tecnologia LTDA® e com ajuda de diversos colaboradores. Detalhes a respeito da construção da base são apresentados a seguir.

Para captura das imagens de impressão palmar e geometria da mão, foi construído um dispositivo formado por três partes principais: uma câmera, um negatoscópio e uma caixa de *Medium Density Fiberboard* (MDF). O negatoscópio, aparelho dotado de iluminação especial para perfeita observação dos negativos ou chapas radiográficas, é responsável por distribuir a luz uniformemente dentro do dispositivo. A caixa de MDF, sobre a superfície do negatoscópio, cria um ambiente controlado, aumentando o contraste entre a superfície do negatoscópio e a mão sobre ele, e isolando o negatoscópio de iluminação externa. A caixa também possui duas aberturas: uma para colocação da mão e outra, no topo da caixa, para

fixação da câmera que captura as imagens. Na Figura 16 é apresentado o dispositivo de captura, composto pelo negatoscópio, a caixa em MDF e a câmera escolhida.

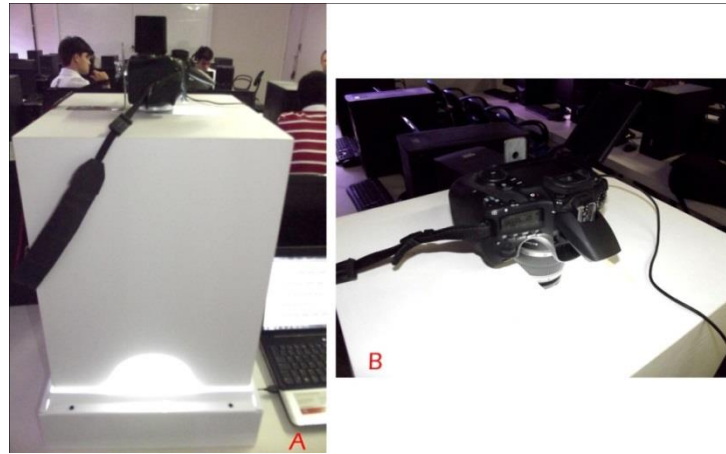


Figura 16: (a) Dispositivo de captura composto por negatoscópio, caixa em MDF e câmera fotográfica. (b) Detalhe da câmera posicionada na caixa.

O interior do dispositivo apresenta ainda régua de vinte centímetros e um cartão fotográfico de cores reais *colorchecker classic* da *X-rite* fixados à superfície do negatoscópio. As régua fornecem informações acerca da escala das imagens capturadas bem como a área resultante de cada *pixel* através de simples cálculos matemáticos. O cartão de cores, por sua vez, pode ser utilizado para correção de brilho e balanço de branco.

A câmera fotográfica utilizada para captura das imagens palmares e de geometria da mão foi uma *DSLR Canon 60D*. O CD que acompanha a câmera contém um *software* em que podem ser definidos diferentes parâmetros para a captura de imagens. Os parâmetros utilizados encontram-se no Anexo A do presente trabalho. Eles foram escolhidos baseados em testes empíricos preliminares.

Como a construção da base envolveu pesquisa com seres humanos, o projeto teve de ser aprovado pela Plataforma Brasil e, em seguida, liberado pelo Comitê de Ética em Pesquisa com Seres Humanos do Hospital Universitário Lauro Wanderley da UFPB. Cada participante do banco ainda assinou um termo de consentimento e respondeu um breve questionário, também aprovados pelos órgãos citados. O termo de consentimento, o questionário, e o instrumento completo do protocolo de captura aprovado, encontram-se nos Anexos B, C e D, respectivamente, do presente trabalho.

Ao todo, a base BioPass-UFPB possui 1200 imagens da mão esquerda de 100 pessoas diferentes – sendo 50 homens e 50 mulheres –, com 12 amostras/pessoas. Para cada pessoa, foram capturadas 7 imagens do dorso da mão e 5 imagens da palma. A Figura 17 apresenta um exemplo de imagens capturadas pelo dispositivo e a Tabela 2 apresenta dados estatísticos a respeito das idades dos voluntários da base BioPass-UFPB. Neste trabalho, apenas as imagens palmares foram utilizadas por apresentar um desafio maior na extração de atributos, uma vez que, no caso da base BioPass-UFPB, o posicionamento das mãos sobre a superfície do negatoscópio é livre.

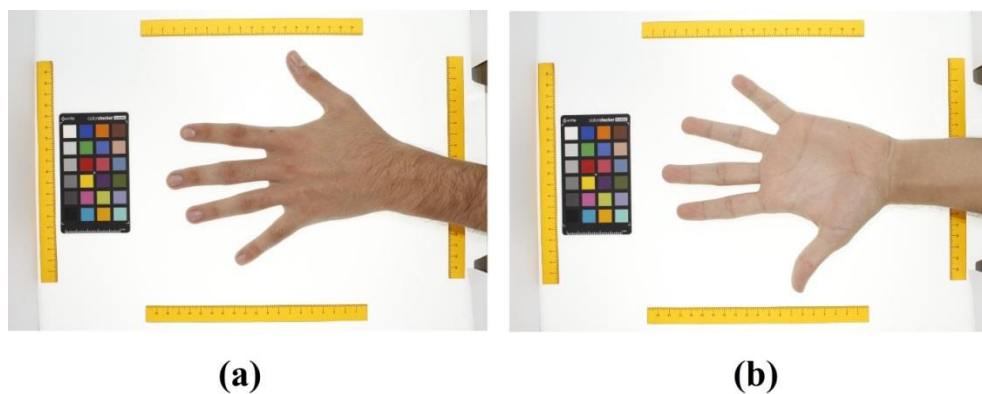


Figura 17: Exemplo de imagens capturadas pelo dispositivo: (a) geometria da mão. (b) impressão palmar

Tabela 2: Estatísticas das idades dos voluntários da base BioPass-UFPB

	Todas	Homens	Mulheres
Menor Idade	17	17	19
Maior Idade	59	59	49
Média das Idades	26	26	26
Mediana das Idades	24	24	23
Desvio Padrão das Idades	7,15	7,43	6,94

Para representar os *templates* de cada imagem, atributos geométricos foram extraídos para servir como características representativas da palma e dos dedos. O trabalho de Nascimento et al. (2014) extrai 54 atributos do banco BioPass-UFPB, porém, neste trabalho, mais 45 atributos foram utilizados, totalizando 99 atributos. A descrição dos atributos, bem como a numeração dada a eles, foi definida como segue:

- Da Palma

1. Área da mão;
2. Perímetro da mão;
3. Área convexa (área do menor polígono convexo que contém a mão);
4. Excentricidade (da elipse que tem o mesmo segundo momento central normalizado da mão);
5. Comprimento do maior eixo da elipse (descrita acima);
6. Comprimento do menor eixo da elipse; e
7. Largura da mão;

- De cada dedo:

1. Área do dedo;
2. Perímetro do dedo;
3. Área convexa do dedo (similar a da palma);
4. Excentricidade
5. Comprimento do maior eixo da elipse
6. Comprimento do menor eixo da elipse
7. Comprimento do dedo
8. Largura Inferior (base do dedo)
- 9-14. 6 medidas de larguras mensuradas da base ao topo do dedo; e
- 15-16. 2 ângulos que representam a inclinação do dedo, exceto polegar (ver Figura 18);

- De textura:

- 1-13. Atributos de Haralick; e
14. Comparação dos Histogramas dado pelo LBP;

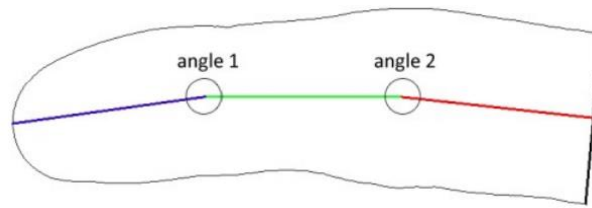


Figura 18: Ângulo entre os segmentos que representam a inclinação natural dos dedos. Fonte: do Nascimento et al. (2014).

Vale salientar que somente um atributo de textura de Haralick não foi utilizado, o Coeficiente de Correlação Máxima, pelo motivo que as outras 13 medidas serem as mais adotadas em outros trabalhos que utilizam os atributos de Haralick (Doyle et al., 2007; Miyamoto & Merryman, 2005). Além disso, o cálculo da GLCM foi parametrizado para $d = 1$ e $\theta = 0^\circ$ como recomendado pela literatura (ver Seção 2.1.1). Por fim, para extração dos atributos de textura, tanto Haralick quanto LBP, uma ROI é extraída de cada imagem palmar. Detalhes a respeito do método de extração são definidos na subseção que segue.

4.4 Extração da ROI

O objetivo de extração da ROI na imagem palmar é definir uma região que esteja preferencialmente contida somente dentro da palma da mão. Isso evita que a extração dos atributos de Haralick e LBP considerem texturas diferentes da palma da mão, como a do negastoscópio ou do cartão de cores no caso do banco BioPass-UFPB.

A Figura 19 apresenta os 14 pontos do contorno da mão extraídos de acordo com o trabalho de do Nascimento et al. (2014), representados aqui por $P_0, P_1, \dots, P_{13}$. O método de extração da ROI desenvolvido neste trabalho leva em consideração os pontos P_0, P_3, P_7 , e P_{13} . Primeiramente, os pontos médios entre os pontos P_3 e P_7 , denominado $PM_{3,7}$, e P_0 e P_{13} , denominado $PM_{0,13}$, são calculados para definir os limites aproximados da palma da mão. Logo após, o ponto médio entre esses dois novos pontos também é calculado para aproximar o centro da palma. Por fim, uma área retangular de tamanho pré-definido (550x750) e rotacionado pelo ângulo entre P_3 e P_7 é definida em torno do centro da palma.

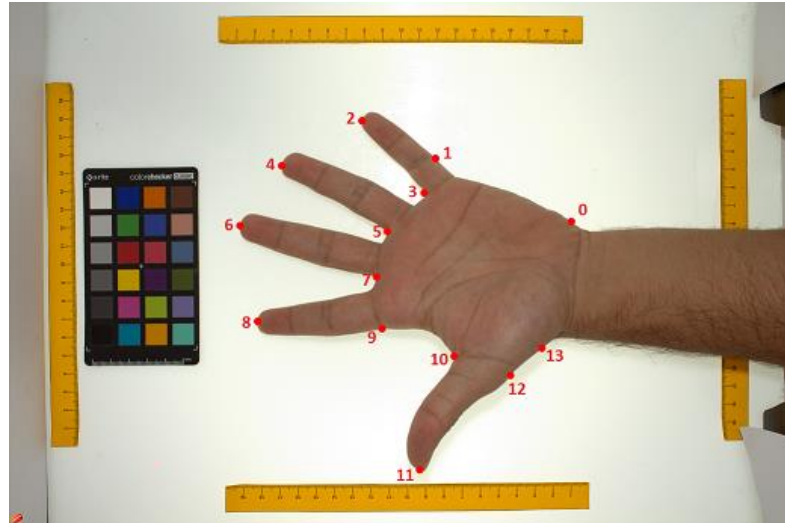


Figura 19: Ilustração dos 14 pontos do contorno da palma definidos por (do Nascimento et al., 2014).

4.5 Rotina de Testes

Para validar o método proposto, rotinas de testes foram pensadas para analisar seu comportamento perante diferentes situações de verificação e classificação. Em comum, todos os testes de verificação ou classificação foram realizados com as configurações do GA presentes na Tabela 3.

Tabela 3: Parâmetros de teste para o algoritmo genético

Comb.	Nº de Cromossomos	Probabilidade de Cruzamento (%)	Nº de Gerações	Probabilidade de Mutação (%)	Tipo de Cruzamento
1	100	0,5	10	0,5	Um ponto
2	100	0,5	10	0,5	Dois pontos
3	100	0,5	10	0,5	Uniforme
4	100	0,5	100	0,5	Uniforme
5	100	0,5	1.000	0,5	Uniforme
6	1.000	0,5	100	0,5	Uniforme
7	1.000	0,5	1.000	0,5	Uniforme
8	100	0,75	100	0,5	Uniforme
9	100	1,0	100	0,5	Uniforme
10	100	0,5	100	0,75	Uniforme
11	100	0,5	100	1,0	Uniforme
12	100	1,0	100	1,0	Uniforme

O restante dos parâmetros possíveis da biblioteca GALib são mantidos como padrão, isto é, “*random selection chance*” = 5 e “*num. Prelim. Runs*” = “*max prelim runs*” = 0. Cinco

casas decimais de precisão são utilizadas e somente valores positivos gerados pela função de *fitness* são permitidos.

As três primeiras combinações de testes variam o tipo de cruzamento a ser efetuado pelo algoritmo genético para verificar qual dos tipos converge melhor em poucas gerações. Testes empíricos preliminares mostraram que o tipo de cruzamento “uniforme” obteve melhores resultados na média em relação aos outros dois tipos. Por causa disso, o cruzamento “uniforme” é mantido no restante das combinações. Da quarta a sétima combinação, varia-se o número de cromossomos e gerações entre 100 ou 1.000 para verificar se esse aumento contribui para a melhora do desempenho do método proposto. Por fim, as últimas cinco combinações permutam a probabilidade de cruzamento e/ou de mutação entre os valores 0,5%, 0,75% e 1,0% para o número de cromossomos e gerações iguais a 100 com o mesmo propósito. Todas as 12 combinações são executadas três vezes cada e o melhor resultado é utilizado para avaliação. O tempo decorrido de cada execução do algoritmo genético também é computado.

O objetivo do GA no presente trabalho é definir os valores dos pesos $w_i \in \{w_1, w_2, \dots, w_N\}$ (ver Equação 4) que minimizam a EER no caso da verificação e maximizam a taxa de acerto na classificação. Pesos diferentes foram calculados para cada caso, respectivamente. Para medir o ganho de performance obtido pelo GA, as métricas de desempenho da verificação e classificação foram calculadas utilizando pesos unitários, isto é, $w_1 = w_2 = \dots = w_n = 1$. Os pesos unitários representam o mesmo grau de importância para todos os atributos.

Nas subseções a seguir são diferenciados os testes executados para validar a verificação e a classificação.

4.5.1 Verificação

No sentido de analisar o comportamento do algoritmo genético em diferentes cenários de verificação, inicialmente duas abordagens baseadas nas técnicas de avaliação de modelos *holdout* e validação cruzada foram empregadas (Kohavi, 1995). Por fim, uma terceira abordagem analisa a inclusão dos atributos de textura nos melhores casos das duas primeiras abordagens. Em todas, o GA foi executado visando a minimização da EER. Dessa forma, a função de *fitness* foi projetada para retornar a EER dada pela aplicação dos pesos da geração

atual no cálculo do escore para todos os *templates* em questão. Detalhes de cada abordagem são descritos como segue.

A primeira abordagem, baseada na técnica *holdout*, separa as P primeiras classes, em conjuntos de 10 em 10, para treinamento e o restante para teste. Ou seja, na primeira rodada dessa abordagem, as 10 primeiras classes são usadas para treinamento enquanto as 90 últimas são usadas para teste; na segunda rodada, as 20 primeiras classes são usadas para treinamento e as 80 classes restantes para teste; e assim por diante (ver Figura 20). Este tipo de validação simula, a princípio, como o treinamento do sistema seria dado na prática, onde certa quantidade de classes seria inicialmente cadastrada no banco para definição dos pesos utilizados no cálculo do escore e, a partir dessa definição, cada nova classe cadastrada já poderia ser aplicada na prática sem necessidade de novo treinamento.

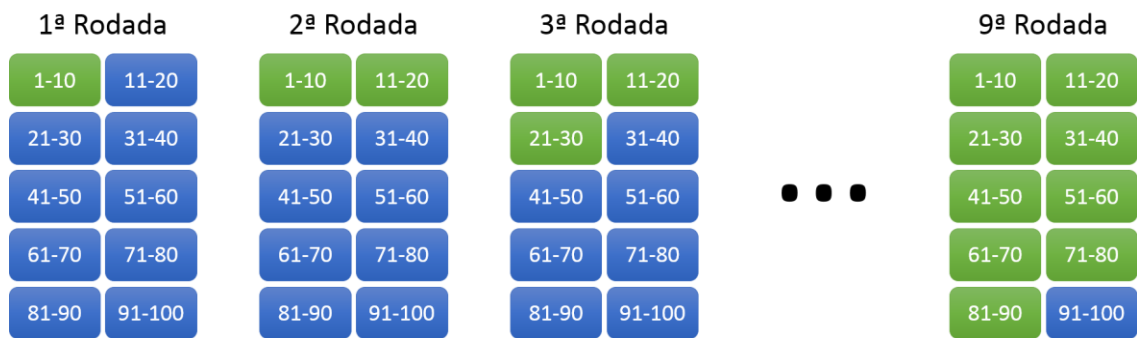


Figura 20: Ilustração da rotina de testes da primeira abordagem da verificação. Cada caixa representa um conjunto de 10 classes, isto é, a caixa “1-10” representa as 10 primeiras classes; a caixa “11-20” representa as classes de 11 a 20, e assim por diante. Ademais, as caixas destacadas em verde representam os conjuntos de treinamento, enquanto as caixas azuis representam os conjuntos de validação.

A segunda abordagem, por sua vez, baseada na validação cruzada, combina as instâncias das classes, isto é, para cada combinação, três amostras de todas as classe são escolhidas para treinamento e o restante é utilizado para teste. Neste caso, todas as 10 combinações possíveis foram efetuadas (ver Figura 21). Os resultados desses testes serviram para averiguar a heterogeneidade das instâncias no treinamento e sua importância para o resultado final. Quanto mais heterogêneas são as amostras para o treinamento, espera-se que melhor seja o modelo de treinamento gerado como também a predição das classes no momento do teste.



Figura 21: Ilustração da rotina de testes da segunda abordagem da verificação. Cada caixa representa agora todas as classes da n-ésima amostra, isto é, a caixa “1ª” é formada pela 1ª amostra de todas as classes, a caixa “2ª” é formada pela 2ª amostra de todas as classes, e assim por diante. Novamente, as caixas destacadas em verde representam os conjuntos de treinamento, enquanto as caixas azuis representam os conjuntos de validação.

Por fim, a terceira abordagem da verificação analisa a inclusão dos atributos de textura – Haralick e LBP – nos melhores casos das duas primeiras abordagens separadamente. Dessa forma, pode-se averiguar a influência dos atributos de textura na verificação e saber o nível de importância frente aos atributos geométricos.

Acredita-se que as três abordagens utilizadas foram suficientes para medir de forma quantitativa o comportamento do modelo na prática. Com eles, foi possível aferir separadamente a importância tanto do número de classes quanto do número de instâncias na verificação. Além disso, também foi possível observar quais atributos dos *templates* são mais importantes para a verificação. Tais resultados serão mostrados posteriormente.

4.5.2 Classificação

Na classificação, o algoritmo genético também foi aplicado em três abordagens diferentes. A primeira visa aumentar a taxa de acerto, enquanto o objetivo da segunda é minimizar o número de *templates* que devem ser recuperados pelo sistema de forma que haja ao menos uma amostra da mesma classe do *template* em questão. No primeiro caso, o classificador *k-NN* (Fix & Jr., 1951) foi utilizado. Assim como na verificação, uma terceira abordagem foi efetuada para aferir a influência dos atributos de textura nos melhores casos. Somente na classificação, a execução do algoritmo genético era interrompida quando, em alguma geração, o GA alcançava 100% na taxa de acerto para o *template* mais próximo.

A primeira abordagem segue a mesma ideia da primeira abordagem da verificação. De fato, os mesmos conjuntos foram utilizados, porém, o objetivo agora é maximizar a taxa de acerto. Dessa forma, tanto no conjunto de treinamento quanto no teste, as similaridades entre

todos os *templates* são calculadas e o classificador k -NN, para $k = 1$, era aplicado para mensurar a taxa de acerto.

A segunda abordagem, por outro lado, foca na classificação do tipo 1: N (ver Seção 2.2.4). De tal maneira, todas as classes foram divididas em cinco bancos diferentes, isto é, o primeiro banco possui a primeira instância de todas as classes; o segundo banco é formado pela segunda instância, e assim por diante. Todas as permutações 1: N foram testadas de maneira mutuamente exclusiva. Assim, uma permutação é utilizada para treinamento e as permutações restantes são adotadas no teste. Como dito, o objetivo nessa abordagem era minimizar o número de *templates* recuperados pelo sistema no conjunto de treinamento de forma que essa otimização refletisse nas permutações de testes. A Figura 22 ilustra o processo dessa abordagem.

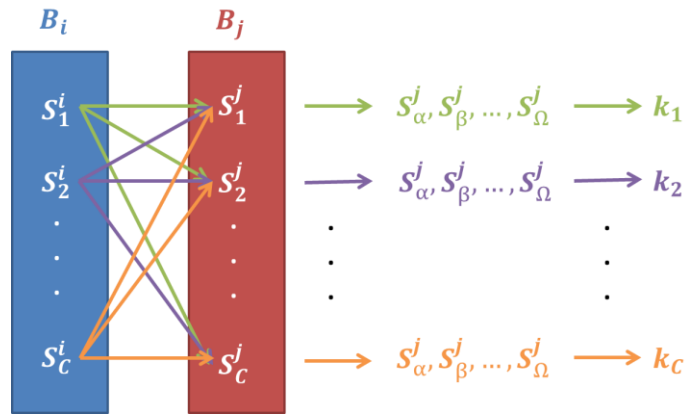


Figura 22: Ilustração dos testes de identificação 1:N. Cada amostra S_C^i do banco B_i (banco de *template*) é comparada com todas as amostras do banco B_j (banco de referência). O resultado para cada amostra é ordenado em ordem crescente e encontra-se o número de *templates* a serem recuperados pelo sistema (k) cujas amostras S^i e S^j pertencem a mesma classe. O objetivo do GA nesse contexto é minimizar o maior número de *templates* recuperados.

Do mesmo modo na verificação, a terceira abordagem da classificação averigua a influência dos atributos de textura em relação ao melhores resultados das duas primeiras abordagens. Vale lembrar que os atributos de textura são dados pelos atributos de Haralick e LBP.

Assim como na rotina de testes da verificação, acredita-se que a rotina de testes da classificação também foi suficiente para analisar o comportamento do sistema na prática. Da mesma forma, será possível medir a importância do número de classes, do número de instância e de cada atributo para classificação. Tais análises serão mostradas e discutidas no Capítulo de resultados.

Capítulo 5

Resultados e Discussões

Este trabalho analisou a utilização de algoritmos genéticos como ferramenta na melhoria de desempenho da verificação e classificação sobre a base BioPass-UFPB. As técnicas de avaliação de modelos *holdout* e validação cruzada, juntamente com o classificador *k-NN* e a recuperação de amostras similares no caso da classificação, foram utilizadas para validar o método proposto. Esse Capítulo apresenta e discute os resultados obtidos na extração da ROI, verificação e classificação, nesta ordem.

5.1 Extração da ROI

O método de extração da ROI desenvolvido neste trabalho funcionou da maneira esperada, pois em nenhuma imagem a ROI demarcada ficou fora da região palmar. A suficiente precisão dos pontos demarcados no contorno da palma da mão foi importante para que tal sucesso ocorresse, pois o método desenvolvido só leva em consideração tais pontos. Ademais, a forma como foi implementado o método, além de simples, permite uma fácil adaptação para outros tipos de banco de dados com resolução e escala diferentes. A Figura 23 ilustra o processo de extração da ROI.

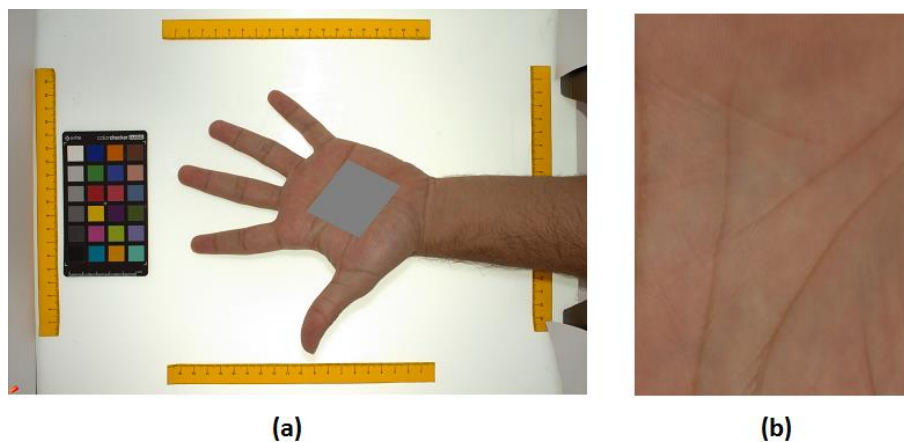


Figura 23: (a) Ilustração da demarcação da ROI na palma da mão. (b) ROI extraída

5.2 Verificação

Esta Seção apresenta resumidamente os resultados das três abordagens de rotinas de testes aplicadas na verificação do presente trabalho (ver Seção 4.5.1). É importante ressaltar que os resultados completos da aplicação do GA nos conjuntos de treinamento em todas as abordagens da verificação encontram-se no Apêndice A do presente trabalho.

5.2.1 Primeira Abordagem

Apresentam-se aqui os resultados da aplicação do GA na primeira abordagem da verificação aplicada sobre o banco, isto é, a técnica *holdout* aplicada em relação às classes. A título de recordação, cada classe têm cinco *templates* com 85 atributos/*template*.

A Tabela 4 mostra que a EER no banco de teste, na primeira rodada de testes, foi de aproximadamente 1,11735%. Além disso, é possível ver que o GA não conseguiu melhorar o resultado da EER quando pesos unitários foram aplicados no conjunto de treinamento. Já para as amostras de teste, a aplicação do GA melhorou a EER de 1,15481% pra 1,11735%, o que representa um ganho relativo de 3,243%.

Tabela 4: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (10 primeiras classes) e no banco de testes (90 últimas classes)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	10 primeiras	1,11111%	0,000%
Sim	Treinamento	10 primeiras	1,11111%	
Não	Teste	90 últimas	1,15481%	3,243%
Sim	Teste	90 últimas	1,11735%	

A Tabela 5 apresenta o resultado da segunda rodada de testes da primeira abordagem, onde as 20 primeiras classes foram utilizadas para treinamento e o restante para teste. No banco de treinamento, o GA permitiu alcançar uma EER de aproximadamente 0,26315% em quase todas as variações de parâmetros, melhorando em 79,592% o desempenho do sistema. No banco de teste, por sua vez, a EER atingiu a marca de 1,17959% quando utilizado os pesos calculados pelo GA no treinamento. Isso representa uma piora relativa de aproximadamente

2,686%, já que, quando aplicado pesos unitários no cálculo da similaridade, a EER é de 1,14873%.

Tabela 5: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (20 primeiras classes) e no banco de testes (80 últimas classes)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	20 primeiras	1,28947%	79,592%
Sim	Treinamento	20 primeiras	0,26316%	
Não	Teste	80 últimas	1,14873%	-2,686%
Sim	Teste	80 últimas	1,17959%	

Na Tabela 6 podem-se observar as EERs obtidas no banco de treinamento e teste com e sem aplicação dos algoritmos genéticos para as 30 primeiras (treinamento) e 70 últimas classes (teste). Quando utilizados pesos unitários para o cálculo do escore, a EER totaliza 1,07471% e 1,21235% nos bancos de treinamento e teste, respectivamente. No entanto, quando o GA foi aplicado, essas taxas reduziram para 0,11494% e 1,05901%, na devida ordem. Isso constitui um ganho relativo de performance de 89,305% no banco de treinamento e 12,713% no banco de teste.

Tabela 6: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (30 primeiras classes) e no banco de testes (70 últimas classes)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	30 primeiras	1,07471%	89,305%
Sim	Treinamento	30 primeiras	0,11494%	
Não	Teste	70 últimas	1,21325%	12,713%
Sim	Teste	70 últimas	1,05901%	

Quando utilizadas as 40 primeiras classes para treinamento e as 60 últimas para teste, o GA atingiu a EER de aproximadamente 0,06410% ao menos uma vez para quase todas as

combinações, exceto as combinações 1, 2, 3 e 8 (ver Tabela 3). No banco de teste, que continha as 60 últimas classes, a EER auferiu 1,20480%. Ao todo, o GA conseguiu melhorar a performance em 94,382% no banco de treinamento e 10,399% no banco de teste. Tais resultados podem ser conferidos na Tabela 7.

Tabela 7: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (40 primeiras classes) e no banco de testes (60 últimas classes)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	40 primeiras	1,14103%	94,382%
Sim	Treinamento	40 primeiras	0,06410%	
Não	Teste	60 últimas	1,34463%	10,399%
Sim	Teste	60 últimas	1,20480%	

É possível ver na Tabela 8 que o ganho de desempenho obtido pelo GA foi de consideráveis 95,192% no banco de treinamento composto pelas 50 primeiras classes, reduzindo a EER de 0,84898% para 0,04082%. No banco de teste, a EER decresceu de 1,73878% para 1,21224%, ou seja, o ganho relativo chegou a 30,282%. A menor EER foi atingida ao menos uma vez para quase todas as combinações de teste do GA, exceto para as 3 primeiras. No banco de teste, a menor EER (1,21224%) foi obtida para os pesos ótimos calculados pelo GA na primeira execução da combinação oito no banco de treinamento.

Tabela 8: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (50 primeiras classes) e no banco de testes (50 últimas classes)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	50 primeiras	0,84898%	95,192%
Sim	Treinamento	50 primeiras	0,04082%	
Não	Teste	50 últimas	1,73878%	30,282%
Sim	Teste	50 últimas	1,21224%	

A Tabela 9 mostra que, quando utilizados pesos unitários no cálculo do escore para as 60 primeiras classes, a EER é de 0,89689%, enquanto que no banco de teste, a EER é mais que o dobro do banco de treinamento (1,90064%). A aplicação dos algoritmos genéticos reduziram essas taxas para 0,02825% e 1,51923%, respectivamente. Vale ressaltar que esses valores só foram atingidos quando utilizados 1.000 cromossomos. A variação do número de gerações não contribuiu para a redução ainda maior da EER. Também é importante destacar que dois conjuntos de pesos diferentes resultaram na mesma EER no banco de treinamento, 1,51923%.

Tabela 9: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (60 primeiras classes) e no banco de testes (40 últimas classes)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	60 primeiras	0,89689%	96,850%
Sim	Treinamento	60 primeiras	0,02825%	
Não	Teste	40 últimas	1,90064%	20,067%
Sim	Teste	40 últimas	1,51923%	

Quando o banco de treinamento foi composto pelas 70 primeiras classes e o banco de teste pelas 30 últimas, a EER obtida foi de 0,77122% e 2,66092% sem a aplicação do GA. A Tabela 10 mostra que os pesos calculados pelo GA conseguiram reduzir essas taxas em 97,315%, no banco de treinamento, e 8,639% no banco de teste. Os menores valores de EER tanto no banco de treinamento, quanto no teste, foram obtidos em mais de uma combinação de parâmetros. Entretanto, no mínimo 1000 cromossomos ou gerações foram necessários para obtenção do menor valor no banco de treinamento, ou probabilidades de 1,0% de mutação e cruzamento. No banco de teste, a menor EER foi alcançada em pelo menos uma das execuções das combinações 5, 6 e 7.

Tabela 10: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (70 primeiras classes) e no banco de testes (30 últimas classes)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	70 primeiras	0,77122%	97,315%
Sim	Treinamento	70 primeiras	0,02070%	
Não	Teste	30 últimas	2,66092%	8,639%
Sim	Teste	30 últimas	2,43103%	

A menor EER (0,01582%) no banco de treinamento da Tabela 11 foi obtida somente nas combinações 6 e 7, onde o número de cromossomos utilizados pelo algoritmo genético foi igual a 1.000. Por coincidência, dois conjuntos de pesos calculados pelo GA também geraram a menor EER no banco de teste, 3,59211%. Sem a aplicação do GA, a EER é aproximadamente 48 vezes maior no banco de treinamento. Contudo, o ganho relativo de performance no banco de teste foi de apenas 6,826%, diminuindo a EER de 3,85526% para 3,59211% (ver Tabela 11).

Tabela 11: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (80 primeiras classes) e no banco de testes (20 últimas classes)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	80 primeiras	0,77057%	97,947%
Sim	Treinamento	80 primeiras	0,01582%	
Não	Teste	20 últimas	3,85526%	6,826%
Sim	Teste	20 últimas	3,59211%	

O último teste da primeira abordagem da verificação – utilizando as 90 primeiras classes para treinamento e as 10 últimas para teste – mostra que o GA não conseguiu melhorar a performance no banco de teste (Tabela 12), assim como no caso em que as 10 primeiras classes foram usadas para treinamento (ver Tabela 4). Todavia, o ganho de performance foi maior que 98% no banco de treinamento. A menor EER no conjunto de treinamento

(0,01248%) foi alcançada somente na terceira execução da combinação 7 do GA. Afinal, usando os pesos calculados pelo GA nessa execução, a EER é igual a 7,94444% no banco de teste.

Tabela 12: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (90 primeiras classes) e no banco de testes (10 últimas classes)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	90 primeiras	0,70787%	98,236%
Sim	Treinamento	90 primeiras	0,01248%	
Não	Teste	10 últimas	7,94444%	0,000%
Sim	Teste	10 últimas	7,94444%	

O Gráfico 1 apresenta o número de ocorrências (ao menos uma vez) de cada combinação dos parâmetros do GA para o melhor resultado da EER na primeira abordagem da verificação mostrada até aqui. As combinações 6 (100 gerações) e 7 (1.000 gerações), ambas com 1.000 cromossomos, tipo de cruzamento uniforme e probabilidade de mutação e cruzamento igual a 0,5%, foram as que mais influenciaram nos resultados da EER. Nos bancos de teste, a combinação 10 também se mostrou tão importante quanto às combinações 6 e 7, resultando quatro vezes na menor EER. Também é importante destacar que as combinações que parametrizavam um baixo número de gerações - presentes nas combinações 1, 2 e 3, assim como a probabilidade de 1% para mutação (combinações 11 e 12) não conseguiram encontrar pesos no banco de treinamento que resultassem na menor EER no conjunto de validação. A combinação 7 foi a única que resultou na menor EER em todos os bancos de treinamento utilizados.

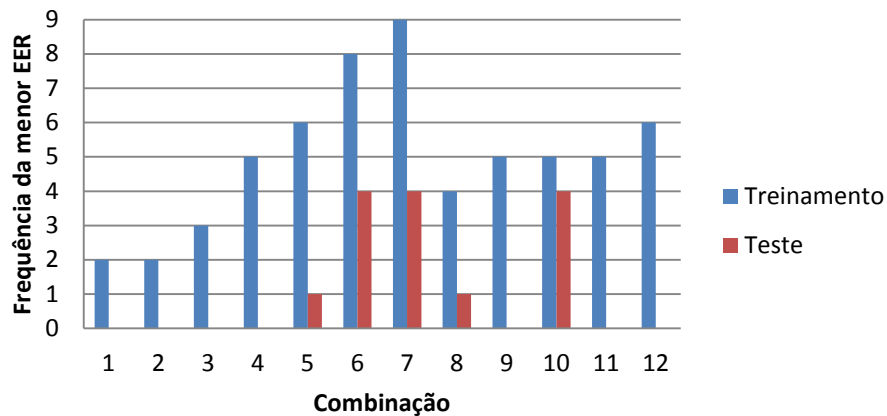


Gráfico 1: Histograma da combinação de parâmetros para a menor EER nos bancos de treinamento e teste na primeira abordagem da verificação

O Gráfico 2 e Gráfico 3 apresentam as EER obtidas em todos os conjuntos de treinamento e teste da primeira abordagem da verificação. Nos bancos de treinamento (Gráfico 2) a diminuição da EER pelo uso do GA fica mais evidente a partir do segundo banco (20 primeiras classes para treinamento). Apenas no primeiro conjunto a EER se manteve igual para a aplicação ou não dos algoritmos genéticos (ver Tabela 4). Nos conjuntos de teste, o mesmo efeito ocorreu no nono banco. Nos outros bancos de testes ocorreu a redução da EER, exceto no segundo (80 últimas classes). Apesar de não importantes do ponto vista prático, os resultados no banco de treinamento demonstram o poder de generalização do modelo pelo GA, que melhorou consideravelmente a EER em praticamente todos os casos.

Um ponto importante a se destacar do Gráfico 3 é o aumento da EER no conjunto de teste. A princípio, espera-se que a EER diminua – ou ao menos mantenha-se na mesma taxa – à medida que o número de amostras no banco de treinamento aumentava enquanto o inverso ocorria com o banco de teste. De tal maneira, quanto maior o número de amostras de treinamento, melhor deveria ser a generalização do modelo pelo GA, tornando o modelo apto o suficiente na verificação das demais classes. No sentido de analisar esse caso, um experimento utilizando a validação cruzada foi executada para analisar a influência desse último conjunto sobre os demais. Exclusivamente nesse caso, o GA foi executado com 100 gerações, 100 cromossomos, probabilidades de mutação e cruzamento iguais a 0,5% e tipo de cruzamento uniforme. Os resultados dessa validação cruzada mostraram que as 10 últimas classes são responsáveis proporcionalmente pelo maior número de FA e FR em relação aos outros bancos (7 FR e 4 FA). Ademais, as classes 96 e 97 estão presentes em nove desses 11 casos. As imagens da Figura 24 mostram as classes em questão. Nessas imagens é possível

observar a presença de dedos parcialmente dobrados, unhas maiores que os dedos e sombras atrás da mão (devido à mão não estar totalmente apoiada na superfície do negatoscópio). Tais erros na captura também ocasionaram erros no processamento das imagens e, consequentemente, na extração dos atributos. A presença dessas características indesejáveis alteram consideravelmente atributos como, por exemplo, altura e/ou largura dos dedos, área da mão/dedo, perímetro da mão/dedo e os ângulos dos dedos. Logo, isso explica a não-uniformidade do banco e o aumento da EER para os conjuntos de treinamento.

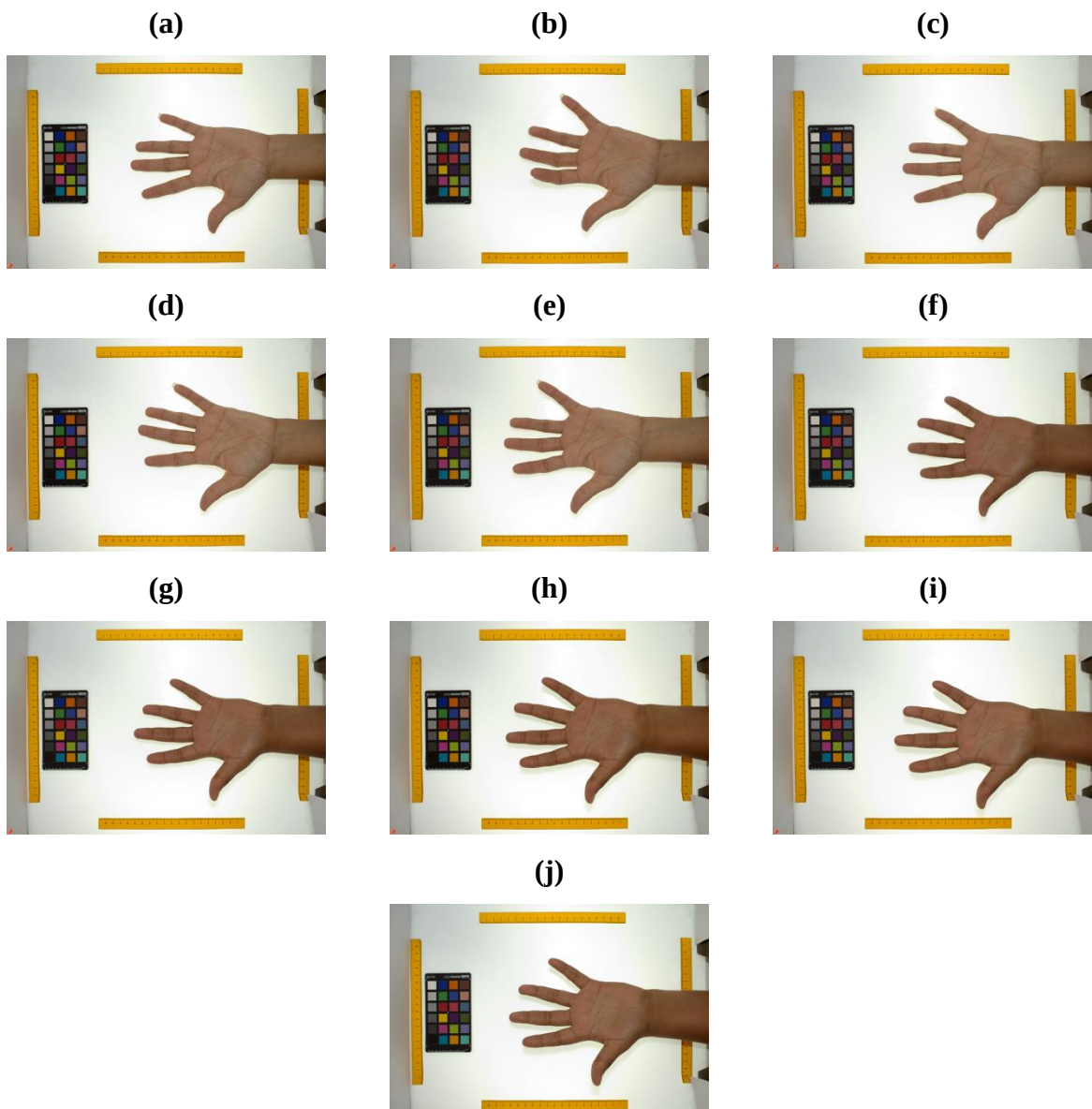


Figura 24: Imagens das classes 96 (a)-(e) e 97 (f)-(j).

O Gráfico 4, por sua vez, apresenta o ganho de desempenho nos conjuntos de treinamento e teste para cada rodada da primeira abordagem da verificação. Como é possível observar, o ganho de performance nos conjuntos de treinamento cresce à uma tendência

logarítmica á medida que se aumentaram as amostras de treinamento. Por outro lado, no conjunto de validação houve um aumento gradativo do ganho de performance nas cinco primeiras rodadas, seguido por um decrescimento nas últimas quatro rodadas.

Por fim, a Tabela 13 apresenta as medidas estatísticas da EER e dos ganhos de performance na primeira abordagem da verificação.

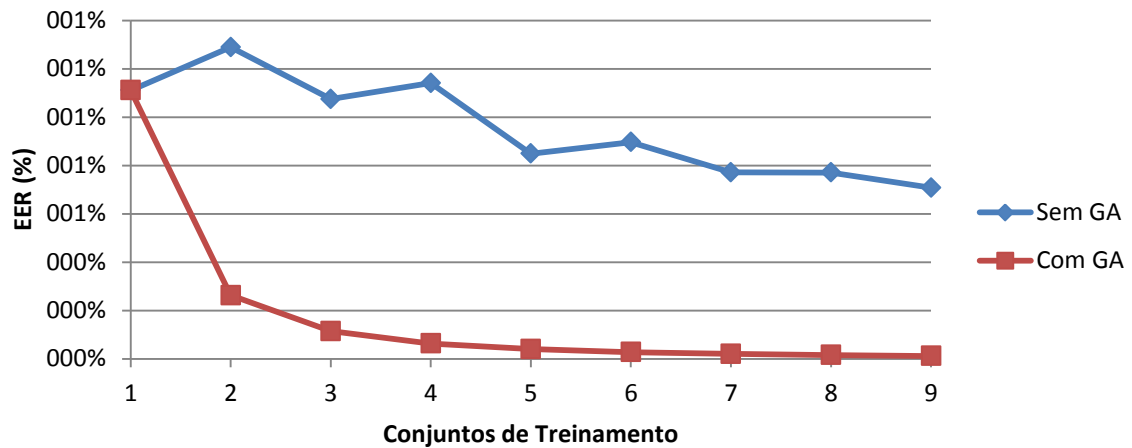


Gráfico 2: Resultados das menores EER nos conjunto de treinamento na primeira abordagem da verificação. O conjunto de treinamento 1 é formado pelas 10 primeiras classes; o segundo é formado pelas 20 primeiras classes, e assim por diante.

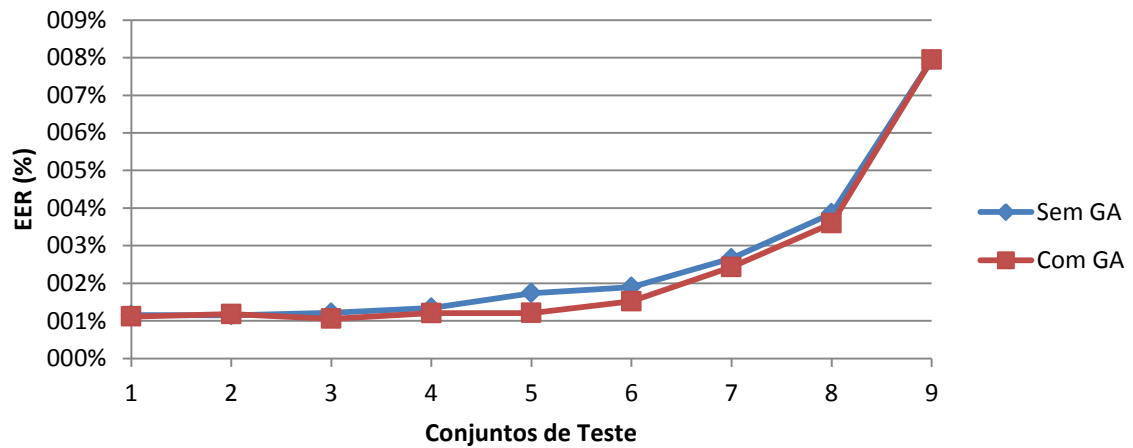


Gráfico 3: Resultados das menores EER nos conjunto de testes na primeira abordagem da verificação. O conjunto de teste 1 é formado pelas 90 últimas classes; o segundo é formado pelas 80 últimas classes, e assim por diante.

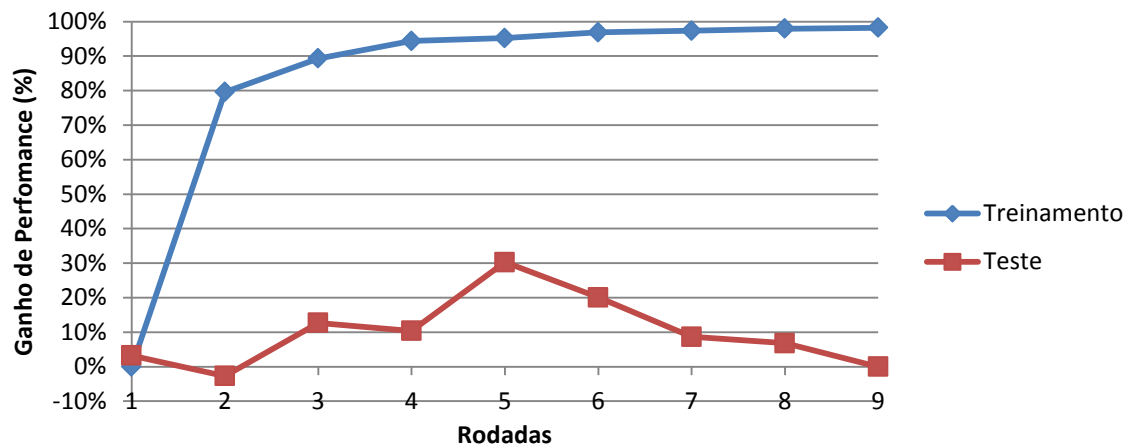


Gráfico 4: Resultado dos ganhos de performance nos conjuntos de treinamento e teste para cada rodada da primeira abordagem da verificação. Cada rodada representa os conjuntos de treinamento e teste conforme o Gráfico 2 e Gráfico 3.

Tabela 13: Média dos resultados obtidos na primeira abordagem da verificação

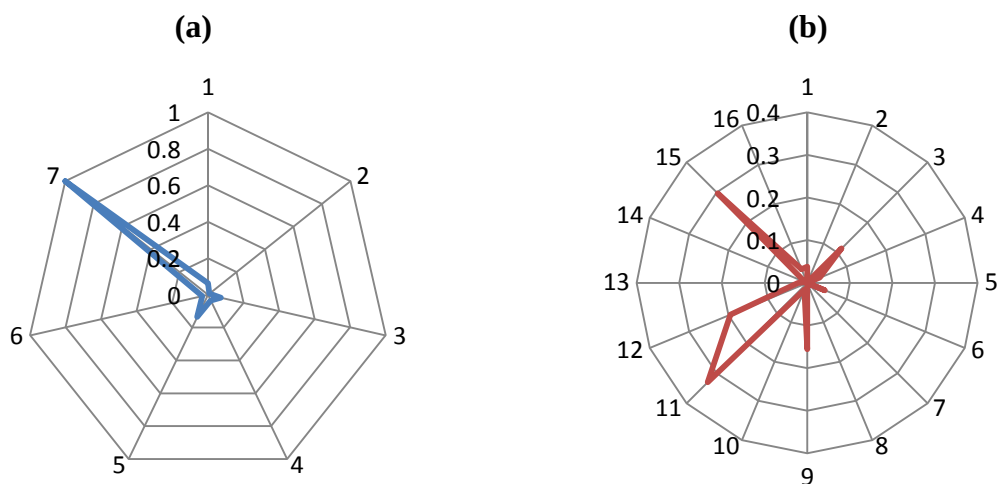
Algoritmo Genético?	Tipo de Banco	EER (Média)	Ganho de Performance Das Médias	Desvio Padrão (ERR)
Não	Treinamento	0,95687%	80,59198%	0,20268%
Sim	Treinamento	0,18571%		0,35603%
Não	Teste	2,55127%	7,41093%	2,20923%
Sim	Teste	2,36220%		2,25549%

Como o objetivo mais importante é melhorar o ganho de desempenho no banco de validação, o melhor resultado encontrado nessa primeira abordagem ocorreu quando o conjunto de dados foi dividido de forma igualitária, isto é, as 50 primeiras classes foram usadas para treinamento do modelo pelo algoritmo genético e as 50 últimas classes foram empregadas na validação (ver Tabela 8). Neste caso, ocorre um ganho relativo de performance superior a 30%, reduzindo a EER de 1,73878% para 1,21224%. Além disso, o GA conseguiu reduzir o número de Falsas Aceitações (FA) de 11 para apenas uma e zerou a quantidade de Falsas Rejeições (FR) que, anteriormente, era igual a quatro. O único caso de FA dada pela aplicação do GA ocorreu entre as amostras das classes 2 e 26, onde o escore

dado pela comparação entre essas amostras (16,195) foi menor que o valor de *threshold* ótimo encontrado pelo GA, 16,800 (ver Gráfico 5).

A partir dos pesos calculados pelo GA no conjunto de treinamento – que foram aplicados no conjunto de validação -, pode se analisar quais atributos foram mais importantes no cálculo do escore dos *templates* e, conseqüentemente, na diminuição da EER. Nos gráficos da Figura 25 é possível observar graficamente a importância de cada um dos atributos que resultou nessa diminuição da EER.

Nos gráficos da Figura 25 é possível observar que o atributo mais importante foi a largura da mão (atributo 7 do gráfico (a)), o qual foi atribuído o valor 30,67343, seguido por uma das medidas de largura extraídas do dedo anelar (gráfico (c)), cujo valor foi aproximadamente 20,37805 (atributo 12). Além disso, também é possível ver que os atributos do dedo indicador foram os que menos influenciaram no cálculo do escore, visto que mesmo o maior peso atribuído a eles não alcança 15% do maior peso atribuído. Na Tabela 14 percebe-se, de forma geral, que o dedo médio foi a parte da mão que mais influenciou no escore, já que a soma dos pesos conferidos ao dedo médio foi maior que nas outras partes. No entanto, vale ressaltar que o dedo médio possui 16 atributos, enquanto a palma da mão, que teve o atributo mais importante para verificação, possui apenas 7 atributos extraídos.



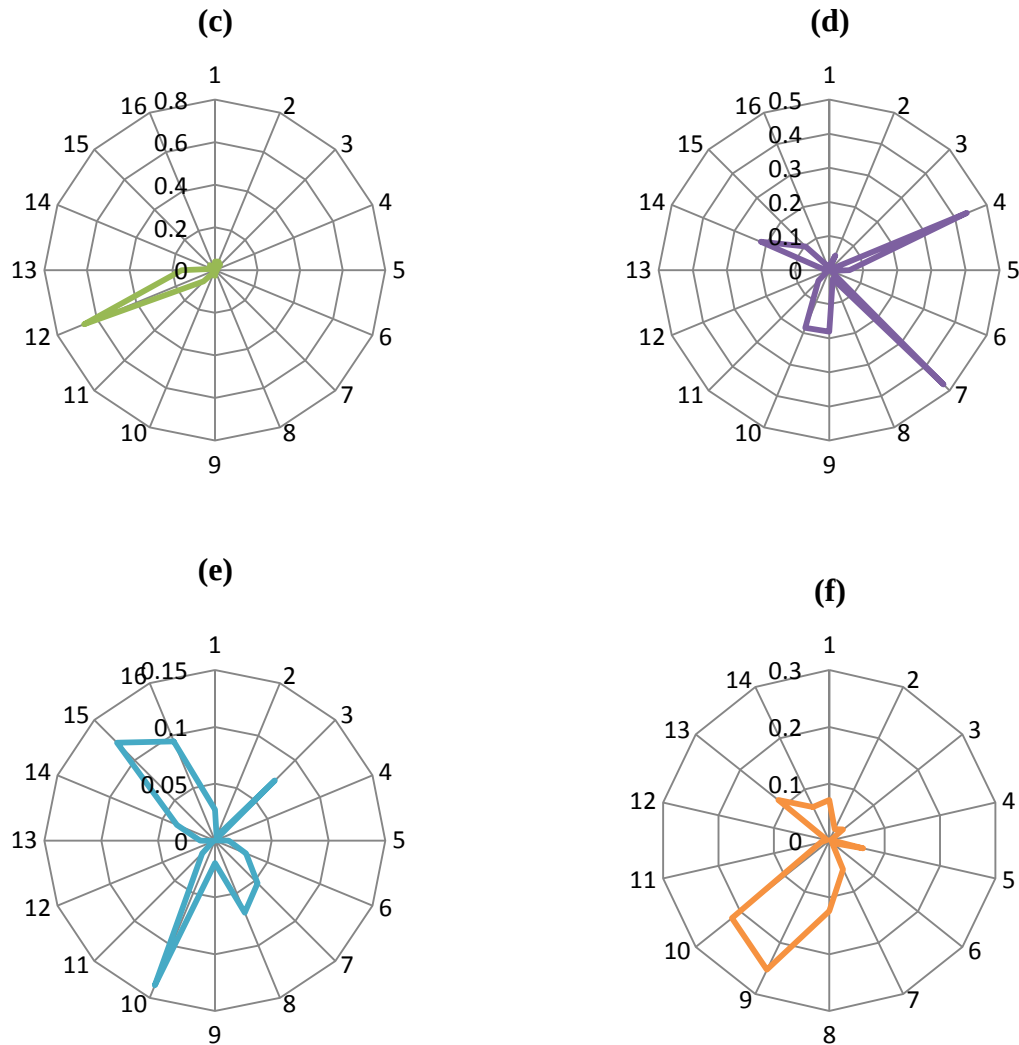


Figura 25: Gráficos do tipo radar que analisam o valor dos pesos atribuídos aos 85 atributos pelo GA. Atributos (a) da palma, (b) do dedo mínimo, (c) do dedo anelar, (d) do dedo médio, (e) do dedo indicador e (f) do polegar. A numeração dos atributos segue a numeração definida na Seção 4.3. Os valores dos atributos foram normalizados na escala 30,67342:1.

Tabela 14: Análise estatística dos pesos atribuídos a cada parte integrante da mão sem normalização.

	Palma	Mínimo	Anelar	Médio	Indicador	Polegar
Menor	0,297723	0,003298	0,020298	0,028736	0,003698	0,086417
Maior	30,67343	10,11899	20,37805	14,54237	4,228203	7,734837
Soma	42,29714	40,52131	36,34039	56,57046	22,7637	33,11522

Por fim, no Gráfico 5 são mostradas as curvas FAR, FRR e ROC do conjunto de teste utilizado na primeira abordagem da verificação que apresentou o melhor ganho de

performance, isto é, quando as 50 últimas classes foram utilizadas. Vale lembrar que a EER é 1,21224%. A área da curva ROC calculada, por sua vez, apresenta o valor de 0,9983.

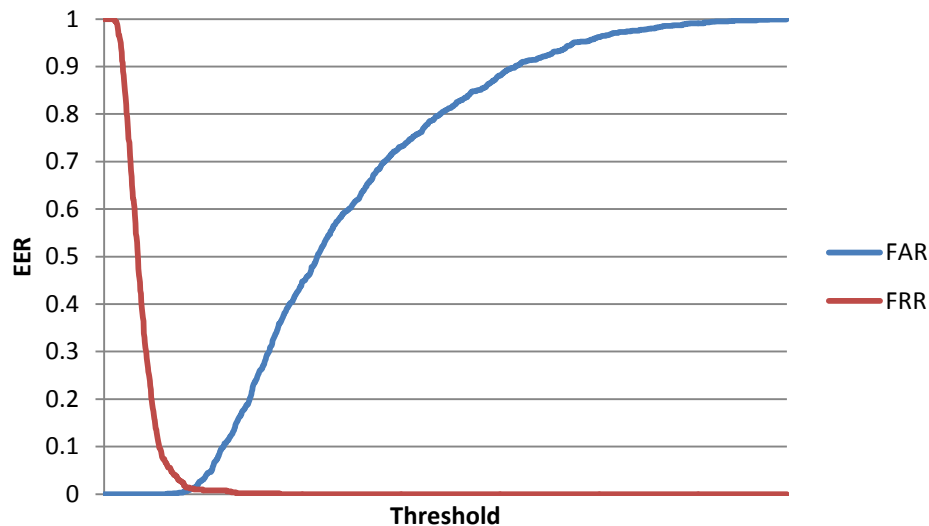


Gráfico 5: Curvas FAR e FRR do conjunto de teste formado pelas 50 últimas classes.

5.2.2 Segunda Abordagem

Na segunda abordagem da verificação, ao invés da variação das classes, as instâncias são variadas através da validação cruzada. Como explicado na Seção 4.5.1, três instâncias de cada classe são usadas para treinamento e duas para teste. No final desta Seção, os resultados são comparados com a primeira abordagem.

A Tabela 15 apresenta o resultado resumido das 10 rodadas da validação cruzada executadas nessas circunstâncias. Como pode ser observada, a utilização do GA no banco de treinamento melhora a média das EER em mais de 71%, reduzindo de 1,23586% para 0,35455%. Além disso, também se pode perceber que o desvio padrão da EER é menor quando o GA é aplicado no banco de treinamento. Da mesma forma, o GA também conseguiu melhorar a EER na média no banco de teste, reduzindo a EER de 1,05909% para 0,955808%. No entanto, o desvio padrão da EER foi maior quando o GA foi aplicado no banco de teste, 0,96773% contra 0,71711% sem aplicação do GA.

Tabela 15: Média dos resultados obtidos na segunda abordagem da verificação

Algoritmo Genético?	Tipo de Banco	EER (Média)	Ganho de Performance Das Médias	Desvio Padrão (ERR)
Não	Treinamento	1,23586%	71,31181%	0,51423%
Sim	Treinamento	0,35455%		0,30112%
Não	Teste	1,05909%	9,53743%	0,71711%
Sim	Teste	0,95808%		0,96773%

O Gráfico 6 mostra o número de vezes em que cada combinação de parâmetros do GA encontrou (ao menos uma vez) a menor EER nos conjuntos de treinamento e teste nesta segunda abordagem da verificação. Assim como na primeira abordagem, a combinação 7 foi a mais importante no conjunto de treinamento, encontrando a menor EER em 9 dos 10 testes da segunda abordagem. Da mesma forma, a combinação 7 também foi a que mais encontrou a menor EER no conjunto de teste, seguida pela combinação 6.

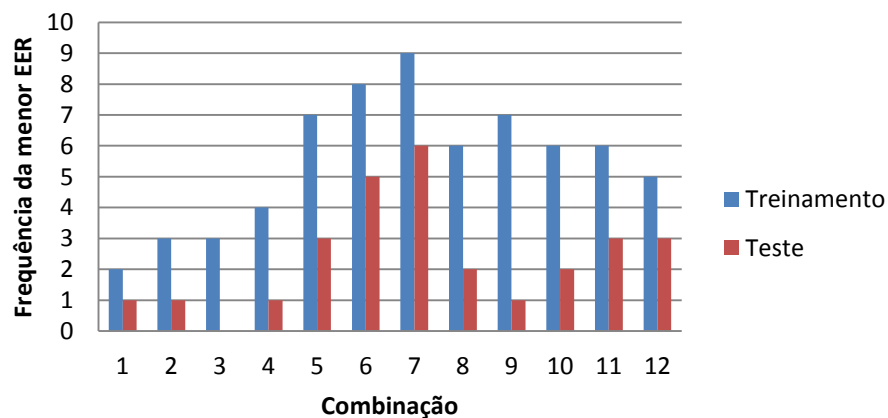


Gráfico 6: Histograma da combinação de parâmetros para a menor EER nos bancos de treinamento e teste na segunda abordagem da verificação

O Gráfico 7 e Gráfico 8 apresentam as EERs obtidas em todos os conjuntos de treinamento e teste da segunda abordagem da verificação. Nos conjuntos de treinamento (Gráfico 7), a diminuição da EER pelo uso do GA ocorreu em todos os casos. Nos conjuntos de teste, por outro lado, o uso do GA piorou a EER no primeiro, no sétimo e no nono conjunto.

No Gráfico 9 é possível analisar o ganho de performance nos conjuntos de treinamento e teste dessa segunda abordagem da verificação separadamente. Percebe-se, nos conjuntos de treinamento, que quase todos os ganhos de performance foram superiores aos 50%. Por sua vez, em alguns casos, os ganhos de performance foram superiores nos conjuntos de validação em relação aos de treinamento. No entanto, nota-se a presença de um possível *outlier* na primeira rodada, onde o ganho de performance foi de -300% no conjunto de teste. É importante destacar que nesse caso, a aplicação do GA piorou a EER de 0,01010% para 0,04040%. Logo, mesmo que o GA tenha piorado tal taxa, esse valor de EER ainda é considerado satisfatório do ponto de vista da segurança biométrica. Porém, esse ganho de performance ocorreu devido ao modo como ele é calculado, já que a fórmula não leva em consideração a piora do resultado, como ocorreu nesse caso.

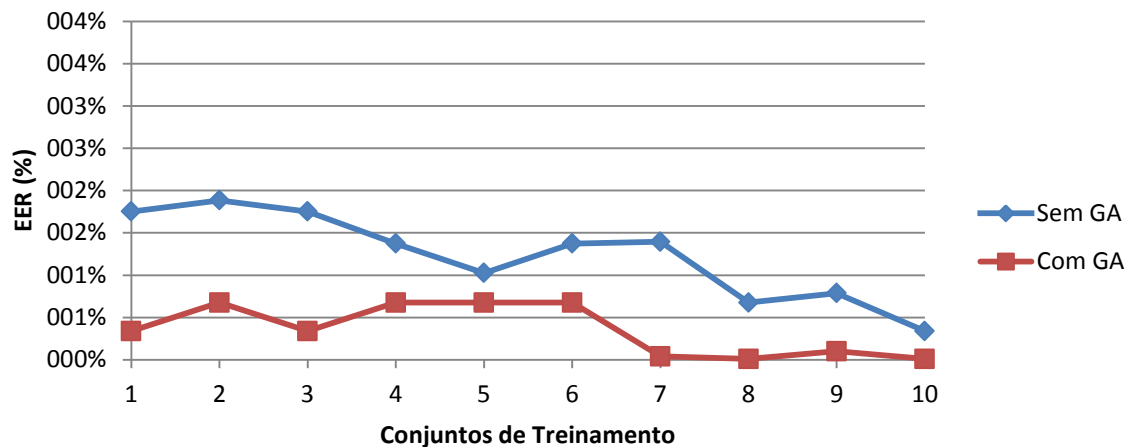


Gráfico 7: Resultados das menores EER nos conjuntos de treinamento na 2ª abordagem da verificação. Cada conjunto de treinamento representa cada permutação possível das instâncias.

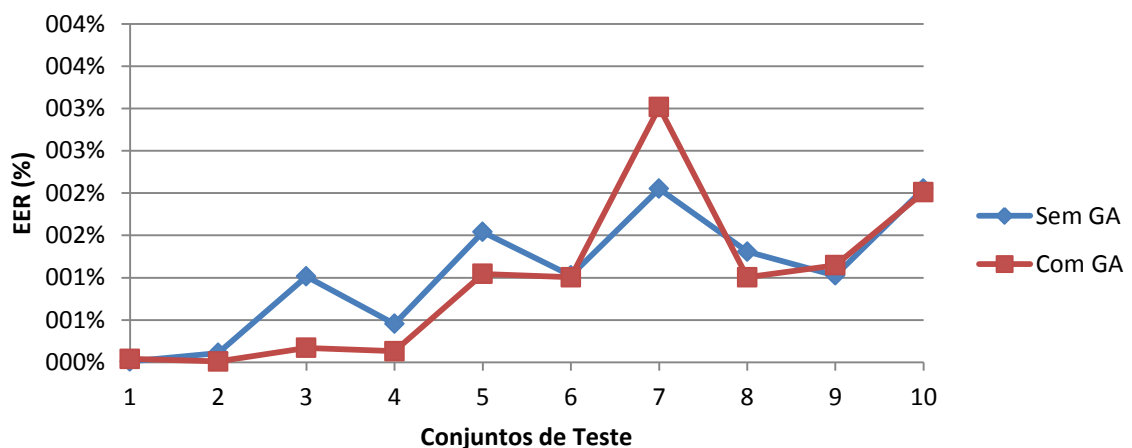


Gráfico 8: Resultados das menores EER nos conjuntos de teste na 2ª abordagem da verificação. Cada conjunto de teste é formado pelas instâncias não utilizadas em cada permutação do conjunto de treinamento.

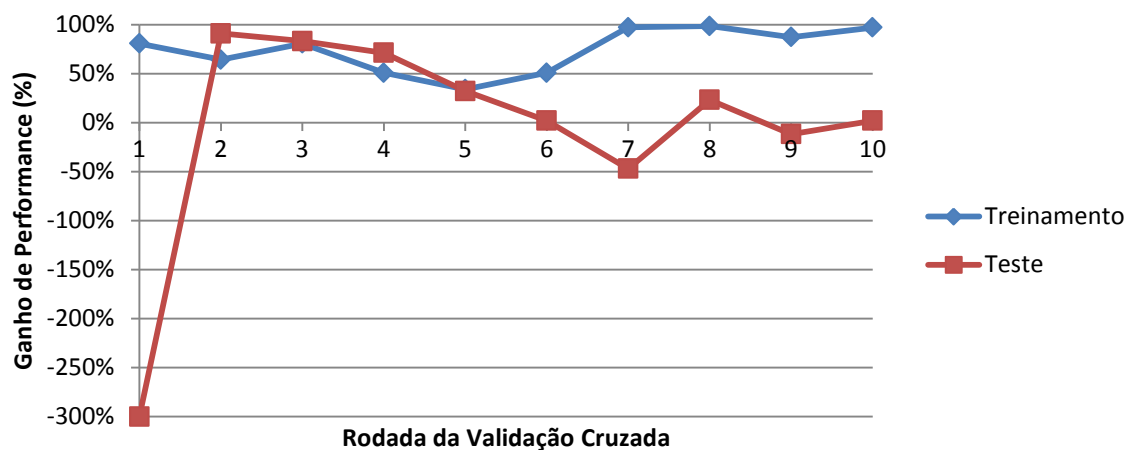


Gráfico 9: Resultado dos ganhos de performance nos conjuntos de treinamento e teste para cada rodada da validação cruzada da segunda abordagem da verificação.

Assim como na primeira abordagem da verificação, foi feita uma análise da importância de cada atributo onde houve o maior ganho de performance no conjunto de teste. Nesse caso, o maior ganho de performance foi encontrado quando as 3ª e 5ª instâncias foram utilizadas no conjunto de teste, alcançando um ganho de performance superior a 90,9% (ver Tabela 48). Essa performance deve-se a redução da quantidade de FA de 104 para apenas 34; enquanto a quantidade de FR foi reduzida de 5 para 2 pela aplicação do GA. Vale ressaltar que as mesmas reduções ocorreram nos dois melhores casos da segunda abordagem, isto é, duas execuções do GA que resultaram na mesma EER. De fato, os mesmos erros aconteceram. Acredita-se que isto veio a ocorrer pela proximidade da soma dos pesos dados ao GA nesses dois casos: 465,3461674 e 455,1874238, respectivamente. No caso da FR, onde 2 amostras foram rejeitadas incorretamente, as duas primeiras amostras da classe 96 foram responsáveis por tais erros.

Como houve duas execuções do GA que encontraram a menor EER (ver Tabela 38), a análise dos atributos será dada em duas figuras diferentes, a constar Figura 26 e Figura 27. Primeiramente, cada figura será analisada individualmente e, ao final, uma análise sobre o resultado das duas figuras será feito em conjunto.

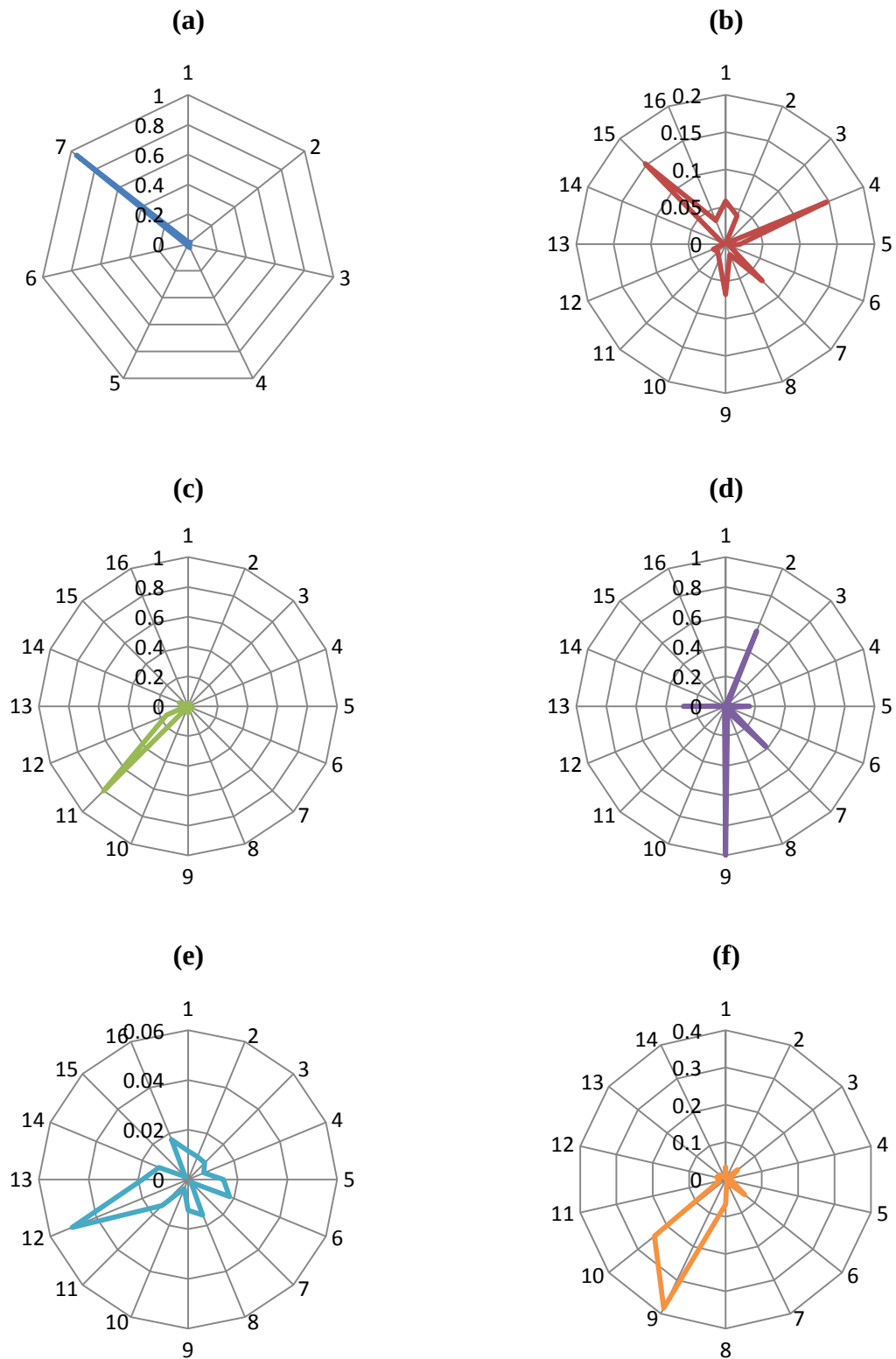


Figura 26: Gráficos do tipo radar que analisam o valor dos pesos atribuídos aos 85 atributos pelo GA na 1ª execução da 6ª combinação mostrada na Tabela 38. Os valores dos atributos foram normalizados na escala 70,61395:1.

Na Figura 26, percebe-se que o atributo mais importante para verificação no conjunto de teste utilizado foi o nono atributo do dedo médio (Gráfico (d)), cujo valor atribuído foi

70,61395, seguido pela largura da mão (atributo 7 do Gráfico (a)), com peso igual a 67,42486. Também é possível ver que o dedo indicador foi o menos importante para verificação, uma vez que os pesos concedidos a seus atributos individualmente não alcançam 6% do valor do maior peso. De forma geral, os atributos do dedo médio foram os que mais influenciaram o resultado da verificação, visto que a soma dos pesos dos atributos do dedo médio foi maior que o dobro da segunda maior soma de pesos (ver Tabela 16). Da mesma forma, os atributos do dedo indicador foram os que tiveram a menor importância, já que a soma dos pesos foi a menor dentre todas as somas, sendo inclusive quase três vezes menor que a segunda menor soma.

Tabela 16: Análise estatística dos pesos atribuídos a cada parte integrante da mão sem normalização. Os valores mostrados são correspondentes a 1ª execução da 6ª combinação mostrada na Tabela 38.

	Palma	Mínimo	Anelar	Médio	Indicador	Polegar
Menor	0,34044	0,052283	0,048667	0,031765	0,107433	0,181299
Maior	67,42487	10,75116	56,65874	70,61396	3,559449	27,02461
Soma	75,17502	48,42427	85,74811	178,0944	16,4487	67,45571

Os gráficos da Figura 27 apresentam os valores normalizados dos pesos calculados pelo GA na outra execução que também gerou o menor valor de EER. Observa-se que novamente a largura da mão se destacou entre os mais importantes atributos, porém o atributo 11 do dedo anelar (que representa a terceira medida de largura extraída desse dedo) foi o que recebeu o maior valor de peso (52,06996). Além disso, o dedo anelar também foi a parte da mão que mais influenciou no cálculo do escore, em razão de apresentar a maior soma dos pesos entre todas as partes (ver Tabela 17). Juntamente com o dedo médio, os pesos concedidos aos atributos desses dedos representaram mais de 53% da soma de todos os pesos. Por outro lado, o dedo indicador foi a parte da mão que teve a menor importância na geração do modelo, porquanto possuiu a menor soma de pesos entre todas as partes da mão.

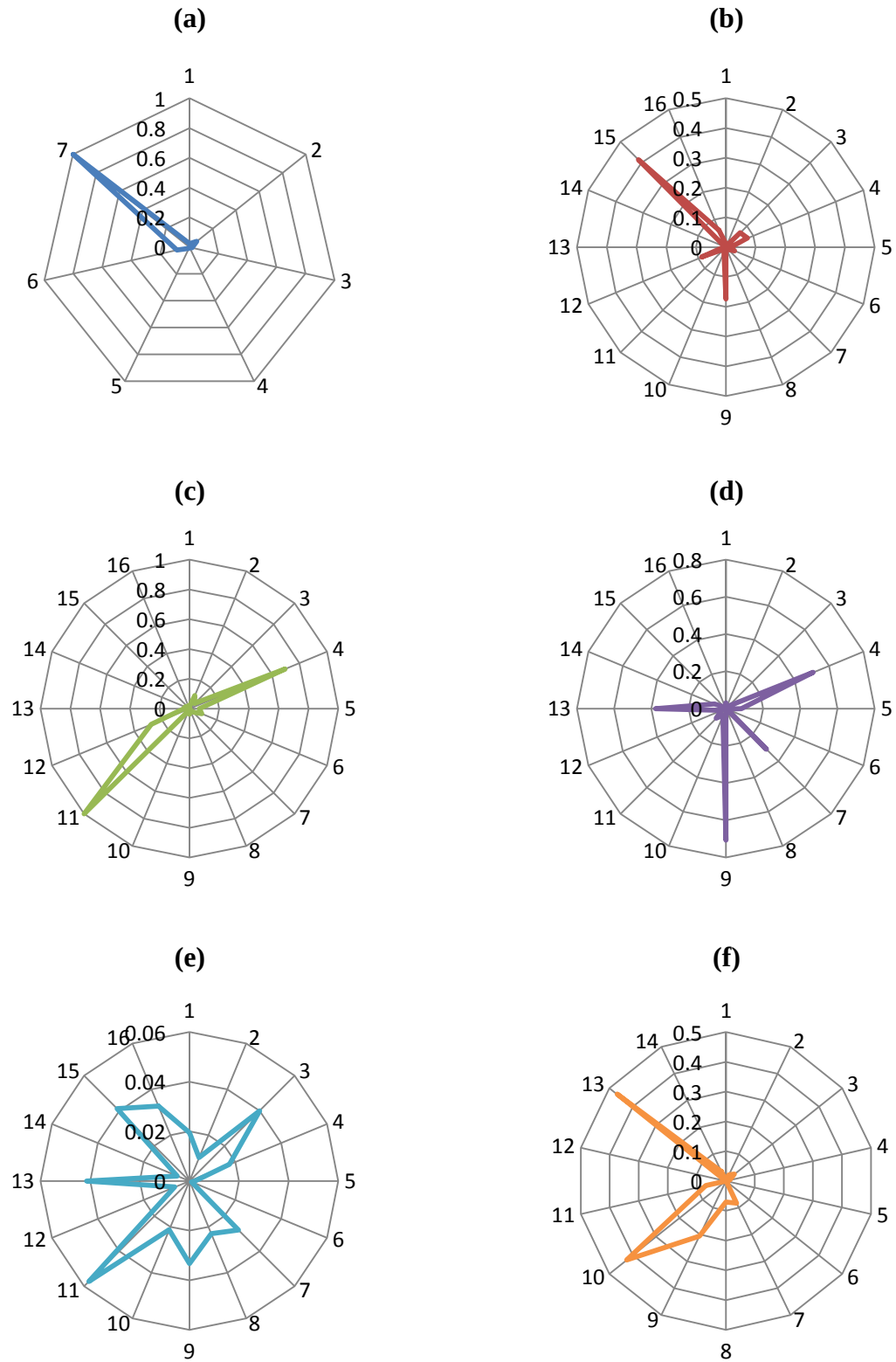


Figura 27: Gráficos do tipo radar que analisam o valor dos pesos atribuídos aos 85 atributos pelo GA na 3ª execução da 6ª combinação mostrada na Tabela 38. Os valores dos atributos foram normalizados na escala 52,06996:1.

Tabela 17: Análise estatística dos pesos atribuídos a cada parte integrante da mão sem normalização. Os valores mostrados são correspondentes a 3ª execução da 6ª combinação mostrada na Tabela 38.

	Palma	Mínimo	Anelar	Médio	Indicador	Polegar
Menor	0,017201	0,016181	0,117773	0,039654	0,046771	0,001497
Maior	51,93787	21,62006	52,06997	36,80822	2,981211	24,31593
Soma	63,28494	52,9527	129,6508	118,8272	20,77649	75,69525

Quando analisados de forma conjunta, os gráficos da Figura 26 e Figura 27 apresentam dados interessantes. Em ambos os gráficos, a largura (atributo 7) da palma da mão foi o atributo que recebeu o segundo maior coeficiente multiplicador. Se levarmos em consideração que esse atributo foi o mais importante na primeira abordagem da verificação (ver Figura 25), isso pode mostrar que a largura da palma da mão pode ser uma característica importante na verificação biométrica da geometria da mão. Além disso, os dedos anelar e médio foram as partes integrantes da mão que receberam a maior soma de pesos em ambos as figuras, como também o dedo indicador recebeu a menor soma de pesos dentre todos os dedos e a palma da mão. Vale ressaltar também que o polegar e a palma da mão, mesmo dispondo de uma quantidade menor de atributos individualmente, foram mais importantes de forma geral que o dedo mínimo, que teve mais atributos extraídos.

Por fim, no Gráfico 10 e são mostradas as curvas FAR, FRR e ROC do conjunto de teste utilizado na segunda abordagem da verificação que apresentou o melhor ganho de performance, isto é, quando as 3ª e 5ª instâncias foram utilizadas. Nesse caso, a EER é 0,010101%. A área da curva ROC, por sua vez, é de 0,99999596.

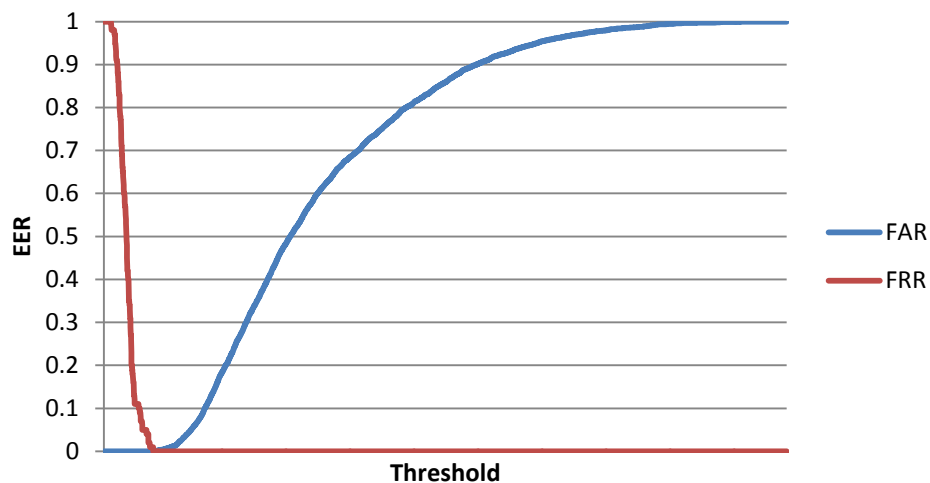


Gráfico 10: Curvas FAR e FRR do conjunto de teste formado pelas 3ª e 5ª instâncias

5.2.3 Terceira Abordagem

A terceira abordagem da verificação aferiu a influência na execução do GA dada pela inclusão dos atributos de textura – tanto de Haralick quanto LBP –, nos melhores casos das duas primeiras abordagens. Somente nessa terceira abordagem que apenas a combinação 7 do GA foi executada para coleta dos resultados, uma vez essa combinação é a mais exaustiva e foi responsável por encontrar o melhor resultado da EER em praticamente todos os testes nos bancos de treinamento utilizados nas duas primeiras abordagens.

Como mostrado anteriormente, a primeira abordagem da verificação, o maior ganho de performance foi obtido para o conjunto de treinamento composto pelas 50 primeiras classes e o conjunto de teste composto pelas 50 classes restantes. Nesse caso, o ganho de performance foi superior a 30% (ver Tabela 8). A Tabela 18, por sua vez, apresenta o resultado obtido pelo GA utilizando 99 atributos (85 geométricos e 14 de textura).

Em comparação com a Tabela 8, pode-se perceber que a inclusão dos atributos de textura influenciou no resultado. Embora o ganho de performance tenha sido menor, é importante destacar dois pontos: o primeiro é que, mesmo com a utilização dos pesos unitários, a EER foi menor – tanto no conjunto de treinamento quanto de validação – que quando os atributos de textura não são usados; e segundo que o GA reduziu a EER para a mesma taxa em ambos os casos nos conjuntos de treinamentos. Isso explica o menor ganho de performance dada pela inserção dos atributos de textura nos conjuntos de treinamento. Por

outro lado, no banco de teste, a aplicação dos pesos calculados pelo GA piorou a EER, aumentando-a de 1,41632% para 2,20204%.

Tabela 18: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (50 primeiras classes) e no banco de testes (50 últimas classes) com a inclusão dos atributos de textura

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	50 primeiras	0,40408%	89,899%
Sim	Treinamento	50 primeiras	0,04082%	
Não	Teste	50 últimas	1,41632%	-55,476%
Sim	Teste	50 últimas	2,20204%	

Ainda em comparação com a primeira abordagem, a inserção dos atributos de textura reduziu a quantidade de FA de 11 para 5 e a quantidade de FR de 4 para 2 no conjunto de treinamento quando os pesos unitários são empregados na verificação. Quando o GA foi aplicado, essas medidas foram reduzidas para 1 e 0, respectivamente - mesmos valores da aplicação do GA no conjunto de treinamento sem os atributos de textura (ver Seção 5.2.1). A falsa aceitação dessa vez ocorreu entre as primeiras amostras das classes 6 e 38.

Na segunda abordagem da verificação, por sua vez, o melhor resultado foi encontrado quando as 1ª, 2ª e 4ª instâncias foram utilizadas para treinamento e as 3ª e 5ª instâncias para teste, onde o ganho de performance foi superior a 90% no conjunto de teste (ver Tabela 48). Em comparação com a Tabela 19, percebe-se um comportamento semelhante da influência dos atributos de textura. Em ambos os conjuntos de treinamento e teste, a aplicação dos pesos unitários, resultou em uma menor EER com os atributos de textura. Com a aplicação do GA, a EER foi reduzida consideravelmente no conjunto de treinamento, de 0,67676% para 0,04040%. Por conta disso, o ganho de performance foi maior no banco de treinamento, 97,260% contra 64,075% sem os atributos de textura. Vale frisar que os pesos calculados pelo GA resultaram no mesmo valor de EER no conjunto de teste (0,01010%).

De maneira mais detalhada, utilizando pesos unitários, a inclusão dos atributos de textura reduziu a quantidade de FA de 104 para 80 e a quantidade de FR de 5 para 4 no conjunto de treinamento. Quando o GA foi aplicado junto aos atributos de textura, a melhora

fica ainda mais evidente. Agora, as 34 falsas aceitações que ocorreram com os 85 atributos, foram reduzidas para apenas 4. Além disso, a quantidade de FR foi reduzida de 4 para 0.

Tabela 19: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (1ª, 2ª e 4ª instâncias) e no banco de testes (3ª e 5ª instâncias) com a inclusão dos atributos de textura

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	50 primeiras	1,47474%	97,260%
Sim	Treinamento	50 primeiras	0,04040%	
Não	Teste	50 últimas	0,02020%	50,000%
Sim	Teste	50 últimas	0,01010%	

Como visto, em relação ao GA, a inclusão dos atributos de textura melhorou a EER no conjunto de treinamento da segunda abordagem da verificação – não havendo melhora de tal taxa na primeira abordagem. Pela Figura 28(a) e Tabela 20, podemos perceber que, apesar dos atributos de textura terem recebido a terceira maior soma de pesos, o maior peso dos atributos de textura não representou 25% do valor do maior peso. Por outro lado, para o melhor caso da segunda abordagem, a soma dos atributos de textura representou a maior soma em comparação ao restante dos outros atributos, o que pode explicar o alto ganho de performance (97,260%), como pode ser visto na Figura 28(b) e Tabela 21.

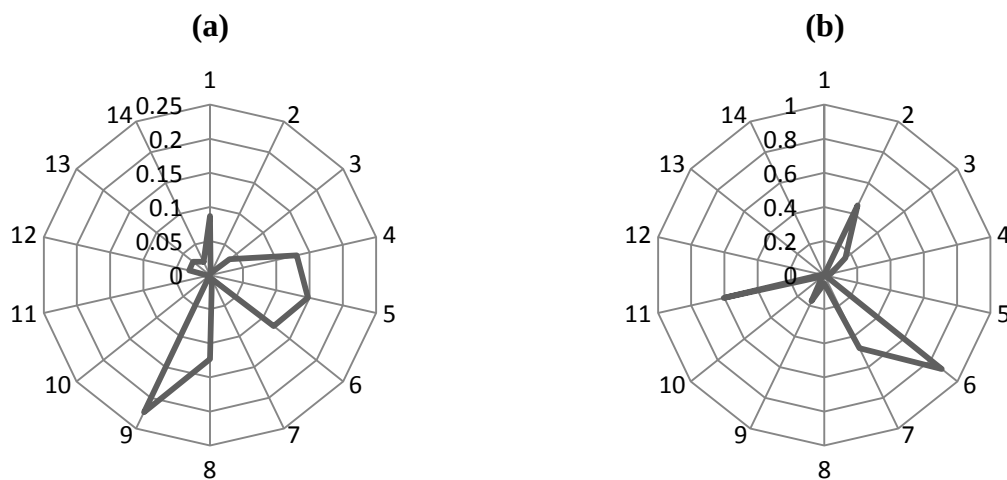


Figura 28: Gráficos do tipo radar que analisam os valores dos pesos dados pelo GA aos atributos de textura nos melhores casos das duas primeiras abordagens da verificação, (a) e (b), respectivamente.

Tabela 20: Análise estatística dos pesos atribuídos pelo GA aos 99 atributos no melhor caso da primeira abordagem da verificação, sem normalização.

	Palma	Mínimo	Anelar	Médio	Indicador	Polegar	Textura
Menor	1,16E-05	0,000294	0,0322	0,000101	1,14E-05	0,000273	0,03938
Maior	35,9285	77,78952	16,02293	18,09729	31,14533	18,33969	17,34941
Soma	95,79665	153,4438	45,72205	43,06314	66,28311	43,53156	76,12116

Tabela 21: Análise estatística dos pesos atribuídos pelo GA aos 99 atributos no melhor caso da segunda abordagem da verificação, sem normalização.

	Palma	Mínimo	Anelar	Médio	Indicador	Polegar	Textura
Menor	0,052942	0,004382	0,000827	0,002244	0,002328	0,000226	0,053339
Maior	30,8947	65,98051	67,34997	111,9171	32,88106	50,71665	98,88116
Soma	52,47432	122,5906	86,88843	267,2549	47,10642	193,9359	323,0811

A próxima Seção desse Capítulo apresenta uma comparação geral entre os resultados obtidos nas três abordagens da verificação até aqui.

5.2.4 Comparação Entre as Abordagens de Verificação

Esta Seção busca comparar o resultado das três abordagens da verificação. A fim de facilitar a compreensão, primeiramente serão comparados os resultados do GA e, em seguida, uma análise a respeito dos atributos, da EER e do ganho de performance será abordada.

Em relação a importância entre as combinações de teste do GA para o encontro do melhor resultado, quando comparamos o Gráfico 6 com o Gráfico 1 – que analisam as duas primeiras abordagens –, podemos observar uma tendência de alta nas combinações de 1 a 7 no conjunto de treinamento em ambos os gráficos. Isso pode mostrar uma correlação forte entre o aumento do número de gerações e cromossomos com a menor EER, ou seja, à medida que se aumenta o número de cromossomos e gerações, aumentam-se também as chances de se encontrar o menor valor da EER. Em ambos os gráficos, também se observa uma tendência neutra no conjunto de treinamento para as combinações de 8 a 12, já que os valores não se distanciam em mais de uma unidade da média dessas combinações. Isso pode mostrar que,

mantendo-se o número de gerações e cromossomos fixo, a alta no aumento da probabilidade de mutação e/ou cruzamento não influencia diretamente na busca pela menor EER. Nos conjuntos de teste de ambos os gráficos, por sua vez, observa-se uma diferença importante em ambos os casos. Enquanto que no Gráfico 1, apenas cinco combinações encontraram a menor EER no conjunto de teste, no Gráfico 6 quase todas as combinações encontraram pelo menos uma vez a menor EER, exceto a combinação 3. Com isso, pode-se deduzir que a presença de todas as classes no conjunto de treinamento, mesmo que em número menor de amostras/classe, é mais importante para geração do modelo do que um número maior de amostras/classe sem a presença de todas as classes. Por fim, não se pode comparar a terceira abordagem com as duas primeiras, já que nem todos dessas duas foram efetuados na terceira abordagem.

A Tabela 22 apresenta o resultado resumido da EER, do ganho de performance e dos atributos mais importantes em cada abordagem. Como pode ser visto, a segunda abordagem apresentou o melhor resultado da EER e do ganho de performance, tanto na média quanto no melhor caso. O mesmo resultado da EER foi encontrado na terceira abordagem, porém com um ganho de performance menor (conforme explicado na Seção 5.2.3). Também é válido destacar que o dedo médio foi a parte integrante da mão mais importante para a verificação nas duas primeiras abordagens, enquanto que os atributos de textura extraídos da ROI foram os mais importantes na terceira abordagem. Além disso, dois atributos do dedo médio (atributos 48 e 53) foram os mais importantes na segunda e terceira abordagem, consecutivamente.

Tabela 22: Visão geral dos resultados obtidos nas três abordagens da verificação. Resultados mostrados para o melhor caso nos bancos de testes.

	EER		Ganho de Performance		Mais Importante	
	Média	Melhor Caso	Média	Melhor Caso	Atributo	Parte da Mão
1ª Abor.	2,36220%	1,21224%	7,41093%	30,282%	7 (Palma)	Dedo médio
2ª Abor.	0,95808%	0,01010%	9,53743%	90.909%	48 (Médio)	Dedo Médio
3ª Abor.	-	0,01010%	-	50,000%	53 (Médio)	ROI

A próxima Seção apresenta os resultados obtidos nas abordagens de classificação aplicadas neste trabalho.

5.3 Classificação

Assim como na Seção 5.2, as subseções a seguir apresentam separadamente os resultados das três abordagens da classificação, agora aplicadas na classificação da base de dados utilizada pelo presente trabalho.

5.3.1 Primeira Abordagem

Na primeira abordagem da classificação, todas as permutações testadas alcançaram 100% na taxa de acerto, utilizando o classificador k -NN com $k = 1$, sem a utilização dos algoritmos genéticos. Isso significa que, para cada amostra de todas as classes, sempre há outra amostra da mesma classe cuja similaridade é a maior em comparação com todas as outras amostras comparadas. Isso também foi verificado quando todas as amostras de todas as classes foram utilizadas. Dessa forma, não houve necessidade da aplicação do GA para tentar melhorar a performance do algoritmo, como também não foi necessário calcular o ganho de performance, já que não há nada a ser melhorado.

5.3.2 Segunda Abordagem

Na segunda abordagem da classificação, conforme explicado na Seção 3.4.2, todas as amostras foram divididas em cinco bancos – um para cada instância – e a classificação 1:N foi aplicada. Nesse caso, uma das permutações era utilizada para treinamento enquanto as demais (sem utilizar nenhuma das amostras de treinamento) eram empregadas no teste. As tabelas desta subseção detalham os resultados obtidos. Primeiramente são mostrados os resultados em todas as permutações utilizando pesos unitários. Logo após são mostrados os resultados nos bancos de treinamento e depois nos bancos de teste.

A Tabela 23 apresenta os resultados da classificação 1:N para todas as permutações dos conjuntos de instâncias utilizando pesos unitários. O menor k significa o menor número de *templates* recuperados pelo sistema (dados pela métrica de similaridade) para que haja pelo menos um *template* de mesma classe. Percebe-se que, nesse caso, o maior dos menores valores de k é 44 (“1x4a”). Logo, caso fosse necessário classificar um usuário pelo sistema seriam necessário retornar 44 amostras para indicar com 100% de certeza que o usuário

estaria entre uma dessas amostras. Percebe-se também que, em alguns casos, o menor número de *templates* recuperados já é um, ou seja, para todas as instâncias utilizadas como *template* sempre a amostra mais próxima pertence à mesma classe.

Tabela 23: Resultados da classificação 1:N utilizando pesos unitários.

Bancos (Temp x Ref)	Menor k
1a x 2a	24
1a x 3a	14
1a x 4a	44
1a x 5a	22
2a x 1a	3
2a x 3a	1
2a x 4a	19
2a x 5a	3
3a x 1a	11
3a x 2a	17
3a x 4a	4
3a x 5a	1
4a x 1a	43
4a x 2a	11
4a x 3a	3
4a x 5a	1
5a x 1a	13
5a x 2a	1
5a x 3a	1
5a x 4a	1

O Gráfico 11 mostra o histograma da combinação de parâmetros do GA que encontraram 100% de taxa de acerto para $k = 1$ (*template* mais similar) nos 20 conjuntos de treinamento possíveis da segunda abordagem da classificação. Vale recordar que, somente na classificação, a execução do GA era interrompida quando $k = 1$. Por esse motivo, pode-se perceber que somente em 2 casos deu-se necessária a execução das 5 primeiras combinações do GA. Na maioria dos casos, já na primeira combinação o valor ótimo de k foi encontrado nos conjuntos de treinamento. Isso mostra que, mesmo para um considerado baixo número de cromossomos e gerações (na primeira combinação esses valores são 100 e 10, respectivamente), o GA conseguiu definir pesos que resultam no valor ótimo de k para a classificação 1:N do banco BioPass-UFPB.

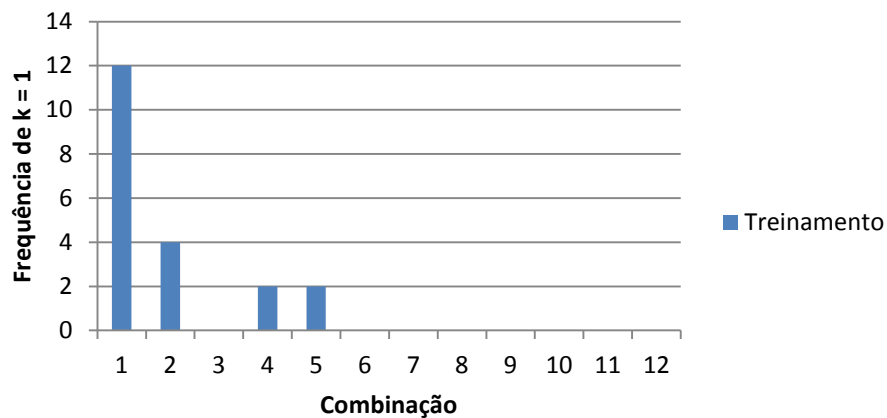


Gráfico 11: Histograma da combinação de parâmetros do GA para o vizinho mais próximo ($k = 1$) nos bancos de treinamento da segunda abordagem da classificação.

A Tabela 24 apresenta as medidas estatísticas dos valores de k nos conjuntos de testes dada pela aplicação dos pesos calculados pelo GA em cada conjunto de bancos de treinamento da segunda abordagem da classificação deste trabalho. É possível perceber que, em alguns casos – como para o conjunto de treinamento “4ax1a” –, a classificação correta pelo *template* mais similar ($k = 1$) também é dada pelos pesos calculados no conjunto de treinamento. Por outro lado, os pesos calculados no conjunto de treinamento “3ax5a”, por exemplo, implicaram em valores considerados altos de k nos conjuntos de teste, como é possível perceber pela média do número de vizinhos nesse caso.

Também é possível observar que os pesos calculados para as instâncias “2ax1a” foram os que geraram a melhor média e o melhor desvio padrão de k nos bancos de testes, enquanto o conjunto de treinamento “3ax5a” resultou nessas piores medidas. Quando analisadas separadamente, a primeira instância gerou os melhores valores de k na média tanto quando utilizada como *template* quando utilizada como referência. Em contrapartida, a quinta instância quando utilizada como *template* foi responsável pela pior média das distâncias.

Tabela 24: Dados estatísticos dos conjuntos de teste para cada conjunto de treinamento da segunda abordagem da classificação.

Bancos de Treinamento (Temp x Ref)	Bancos de Testes				
	Menor <i>k</i>	Maior <i>k</i>	Média de <i>k</i>	Desvio Padrão de <i>k</i>	Média de <i>k</i> por Template
1a x 2a	2	9	4,83	3,06	7,54
1a x 3a	5	22	12,67	5,61	
1a x 4a	2	24	9,17	7,76	
1a x 5a	1	6	3,50	1,87	
2a x 1a	1	4	2,00	1,26	9,79
2a x 3a	1	36	17,17	14,89	
2a x 4a	1	12	5,00	3,69	
2a x 5a	2	33	14,83	13,38	
3a x 1a	1	14	4,00	5,02	13,63
3a x 2a	1	36	16,17	14,63	
3a x 4a	2	21	9,17	7,19	
3a x 5a	5	50	25,17	18,67	
4a x 1a	1	8	3,83	3,06	9,33
4a x 2a	2	7	4,50	1,87	
4a x 3a	3	23	11,17	7,76	
4a x 5a	1	27	17,83	10,65	
5a x 1a	3	26	11,67	8,21	16,83
5a x 2a	4	46	21,67	18,28	
5a x 3a	4	46	22,50	17,28	
5a x 4a	2	25	11,50	8,60	

A Figura 29 apresenta graficamente os resultados dos pesos atribuídos pelo GA no banco de treinamento “2ax1a”. Este foi considerado o melhor banco de treinamento, pois o modelo gerado pelo GA para esse banco apresentou os melhores resultados de *k* nos bancos de testes, como mostrado na Tabela 24. Pode-se observar que o atributo 2 (perímetro) da mão (gráfico (a)) foi o atributo mais importante para classificação, já que lhe foi conferido o maior valor de peso. Se analisarmos cada parte integrante da mão separadamente, o dedo médio foi a parte que mais influenciou no cálculo do score, visto que a soma dos pesos foi a maior em comparação com as outras partes da mão (ver Tabela 25), mesmo embora um de seus atributos teve o menor valor de peso atribuído. Ademais, o polegar foi a segunda parte integrante da mão mais importante para a classificação e também teve o segundo maior valor de peso concedido a um atributo (atributo 9).

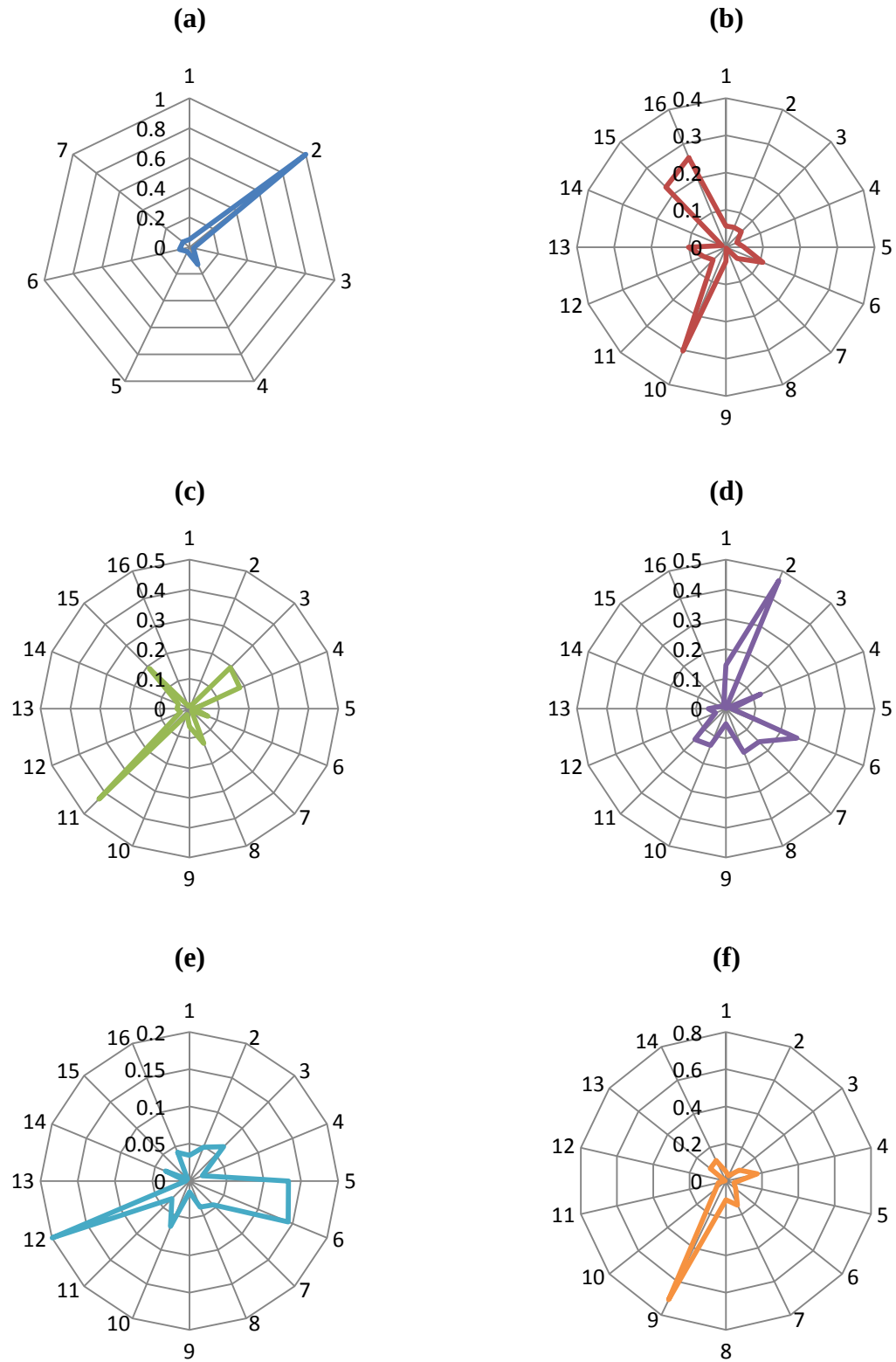


Figura 29: Gráficos do tipo radar que analisam o valor dos pesos atribuídos aos 85 atributos pelo GA no conjunto de treinamento “2ax1a”. Os valores dos atributos foram normalizados na escala 22,77208:1.

Tabela 25: Análise estatística dos pesos atribuídos a cada parte integrante da mão sem normalização. Os valores mostrados são correspondentes ao conjunto de treinamento “2ax1a”.

	Palma	Mínimo	Anelar	Médio	Indicador	Polegar
Menor	0,625866	0,06469	0,182615	0,005766	0,082137	0,107443
Maior	22,77209	6,874084	9,78951	10,59668	4,546845	16,10248
Soma	32,02137	34,2474	33,89375	42,13418	22,06844	41,05721

A próxima Seção apresenta os resultados obtidos na terceira abordagem da classificação empregada neste trabalho.

5.3.3 Terceira Abordagem

A terceira abordagem da classificação mensurou a influência da inserção dos atributos de textura, totalizando 99 atributos, somente na segunda abordagem da classificação, uma vez que não houve necessidade de aplicar o GA na primeira abordagem (ver Seção 5.3.1). Assim como na terceira abordagem da verificação, somente a combinação 7 foi executada, por ser a mais exaustiva.

Como visto anteriormente, na segunda abordagem da classificação, a melhor média de k foi encontrada quando o conjunto de treinamento foi formado pelas segundas instâncias de todas as classes como *template* e as primeiras instâncias como referência (“2ax1a”). Nesse caso, a média de k foi igual a 2. Em comparação com a Tabela 24, a inclusão dos atributos de textura também resultou no mesmo número de *templates* ($k = 4$) para alcançar o mesmo resultado. No entanto, quando aplicados nos conjuntos de testes, os pesos determinados pelo GA diminuíram a média de k de 2 para 1,5. De fato, somente o conjunto de teste “3ax5a” resultou em um valor de k diferente de 1 ($k = 4$, nesse caso). Mesmo que esse conjunto tenha gerado o mesmo maior k , pode-se concluir que a inserção dos atributos de textura conseguiu melhorar o melhor resultado da segunda abordagem da classificação, devido à diminuição da média.

A Figura 30 e a Tabela 26 apresentam os valores dos pesos atribuídos pelo GA aos atributos de textura no melhor caso da segunda abordagem da classificação. Percebe-se que o segundo atributo de Haralick foi o mais importante atributo de textura nesse caso. Além disso, assim como na verificação, a influência dos atributos de textura foi substancial, apesar de que

a soma dos atributos de textura representaram a segunda menor soma dos pesos em comparação as outras partes integrantes da mão.

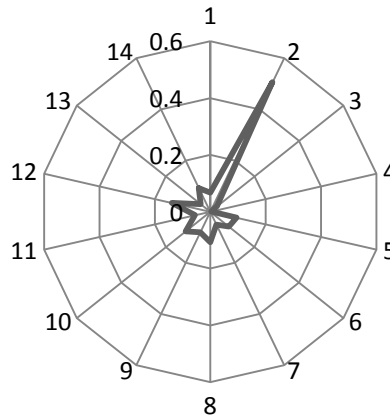


Figura 30: Gráfico do tipo radar que analisa os valores dos pesos dados pelo GA aos atributos de textura no melhor caso da segunda abordagem da classificação. Os valores dos atributos foram normalizados na escala 13,88468:1.

Tabela 26: Análise estatística dos pesos atribuídos pelo GA aos 99 atributos no melhor caso da segunda abordagem da classificação, sem normalização.

	Palma	Mínimo	Anelar	Médio	Indicador	Polegar	Textura
Menor	0,037596	0,023508	0,044706	0,003453	0,051299	0,25101	0,04726
Maior	0,913636	10,77831	8,74779	10,87831	13,88468	10,41967	7,01677
Soma	4,495383	27,26087	36,23672	37,4101	25,63965	31,40666	21,32784

A próxima Seção deste Capítulo compara de forma geral as três abordagens da classificação vistos anteriormente.

5.3.4 Comparação Entre as Abordagens de Classificação

De modo análogo à Seção 5.2.4, uma comparação entre as três abordagens de classificação foi feita. As abordagens são comparadas em relação a aplicação do GA, da taxa de acerto e do número de templates a serem retornados pelo sistema, quando é o caso.

Com relação ao GA, no pior caso em todas as abordagens, apenas cinco combinações (das doze descritas na Seção 4.5) foram necessárias para encontrar o melhor resultado, isto é, 100% de taxa de acerto nos bancos de treinamento. Vale salientar que, diferentemente da verificação, a execução do GA era interrompida quando isso ocorria. Isso pode mostrar que o

GA se comportou melhor para a classificação do que a verificação, posto que menos gerações e cromossomos eram necessários para encontrar o resultado considerado ótimo.

As duas primeiras abordagens da classificação representam tipos distintos de classificações – N:N na primeira e 1:N na segunda -, logo não se pode compará-las diretamente. No entanto, vale lembrar que na primeira abordagem não houve necessidade de aplicação do GA, uma vez que 100% de acerto foi obtido utilizando-se apenas um classificador k -NN com $k = 1$. No entanto, esse tipo de classificação (N:N), é inviável na prática, visto que toda vez que um novo perfil fosse cadastrado no sistema o processo de classificação deveria ser executado. Por outro lado, a segunda abordagem mostrou que, no melhor caso, apenas quatro *templates* seriam necessários ser retornados pelo sistema para garantir que a classe da amostra de referência também estaria presente nos quatro *templates* retornados.

A terceira abordagem da classificação analisou a influência dos atributos de textura somente no melhor caso da segunda abordagem, já que não houve necessidade na primeira, como mostrado. A Tabela 27 compara os resultados dessas duas abordagens. De acordo com os resultados, pode-se concluir que a inclusão dos atributos de textura conseguiu melhorar o melhor caso da classificação 1:N, já que houve uma redução na média do número de templates a serem retornados pelo sistema.

Tabela 27: Visão geral dos resultados obtidos nas duas últimas abordagens da classificação. O valor de k representa o número de *templates* que o sistema deve retornar para garantir que a classe da amostra de referência está presente nos k templates retornados. Resultados mostrados para o melhor caso nos bancos de testes.

	k			Mais Importante	
	Menor	Maior	Média	Atributo	Parte da Mão
2ª Abor.	1	4	2	2 (Palma)	Dedo Médio
3ª Abor.	1	4	1,5	70 (Indicador)	Dedo Médio

5.4 Análise dos Atributos

Essa Seção analisa, de forma conjunta e resumida, a atuação dos atributos nos resultados apresentados da verificação e classificação. Além de compara os atributos individualmente, também foi feita uma comparação a respeito de cada parta integrante da mão.

Ao comparar os resultados da segunda abordagem da classificação com os da verificação, percebe-se que o dedo médio e o dedo anelar foram as partes integrantes da mão mais e menos importante, respectivamente, em todos os resultados. Além disso, a terceira medida de largura extraída do dedo anelar (atributo 11) se destacou recebendo valores expressivos de peso nos três casos. Já o atributo 9 do dedo polegar também foi importante na segunda abordagem da classificação e no primeiro resultado da segunda abordagem da verificação, mas não no segundo. Vale ressaltar que enquanto o dedo polegar não foi uma das partes integrantes da mão mais importantes na verificação, ele representou a segunda parte mais importante na classificação.

Em relação as terceiras abordagens, não houve uma coincidência direta da influência dos atributos, mas pode-se perceber que os atributos de textura foram mais importantes na verificação do que na classificação. Além disso, enquanto na verificação o atributo mais importante foi do dedo médio (atributo 53), na classificação foi do dedo indicador (atributo 70).

De maneira geral, foi possível observar que o dedo médio foi a parte integrante da mão mais importante nos resultados encontrados, já que ele recebeu a maior soma de pesos em quase todos os melhores casos das abordagens de verificação e classificação. De fato, somente na terceira abordagem da verificação que os atributos da ROI extraída da palma da mão foi mais importante que o dedo médio. Além disso, atributos do dedo médio também foram os mais importantes atributos da segunda e terceira abordagem da verificação.

O próximo Capítulo apresenta as considerações finais a respeito de todo o trabalho, resumizando os resultados e mencionando os trabalhos futuros.

Capítulo

6

Considerações Finais

Neste trabalho, uma abordagem baseada em algoritmos genéticos foi utilizada para otimização dos parâmetros utilizados na classificação e identificação do banco palmar na base multibiométrica BioPass-UFPB. Além disso, um novo método de *matching*, semelhante à distância de Manhattan, foi apresentado.

Em primeiro lugar, a base multibiométrica BioPass-UFPB foi construída pela colaboração de 100 voluntários, sendo 50 homens e 50 mulheres. Ao todo, 1200 imagens de mão foram capturadas, com 12 imagens por classe. Cada imagem teve 85 atributos geométricos extraídos que foram utilizados nesse trabalho. Os algoritmos genéticos foram aplicados em diferentes abordagens de classificação e verificação para validar o método proposto; avaliar a importância de cada atributo; e melhorar as taxas de acerto e EER, respectivamente. Além disso, acredita-se que as rotinas de testes empregadas foram suficientes para analisar o comportamento do sistema na prática.

De acordo com os resultados, utilizando somente os atributos geométricos, o uso do GA ocasionou uma melhora de desempenho superior a 30% e 90%, respectivamente, na EER nos melhores casos das duas primeiras abordagens da verificação. Na classificação, por sua vez, a primeira abordagem (classificação N:N) gerou 100% de acerto mesmo sem a utilização do GA, tornando desnecessária sua utilização. No entanto, quando aplicado o GA na segunda abordagem da classificação verificou-se que, no melhor caso, apenas 4 *templates* seriam necessários retornar na prática para que ao menos um desses *templates* fosse da mesma classe para cada amostra de referência na classificação do tipo 1:N.

Com relação aos atributos geométricos extraídos das imagens palmares, foi possível perceber que os atributos do dedo médio foram os mais importantes na classificação e verificação de forma geral. Por outro lado, os atributos do dedo anelar, representaram os menos importantes também em ambos os casos. Por fim, também vale destacar que o dedo polegar foi mais importante para a classificação do que para a verificação.

Em relação aos atributos de textura, a inserção deles ocasionou melhoras na EER dos conjuntos de testes dos melhores resultados da verificação. Na classificação, houve uma melhora na média do número de *templates* a serem retornados pelo sistema no melhor resultado da segunda abordagem. Porém, não se pode dizer que tal inserção melhora o resultado da classificação como um todo, visto que todos os testes executados na segunda abordagem da classificação deverão ser reexecutados com os 99 atributos. Tais testes estão reservados para os trabalhos futuros.

De maneira geral, caso o sistema fosse implementado na prática, e tendo em vista os resultados obtidos, pode-se dizer que 4 amostras de cada usuário seriam necessárias no cadastramento do sistema, onde atributos geométricos e de textura seriam incorporados as informações das imagens. Como mostrado, na classificação, apenas 4 amostras seriam necessárias retornar para garantir com 100% de certeza que o usuário a ser classificado estaria entre as amostras retornadas. Por outro lado, apenas 3 amostras de cada usuário poderiam ser utilizadas na verificação, já que, nessas condições, o sistema obteve uma EER de 0,01010%.

A escolha dos atributos para representação de indivíduos é um passo importante para o reconhecimento biométrico através da geometria da mão e impressão palmar. A partir deles, é possível reconhecer quais atributos são mais importantes que outros tanto na classificação quanto a verificação dos sistemas biométricos. Com isso, pode-se reduzir o tempo do reconhecimento biométrico, melhorar a precisão do sistema e até facilitar o uso do sistema, disseminando o traço biométrico nas aplicações civis e governamentais.

Futuramente, espera-se poder verificar a importância de cada atributo de forma independente utilizando o GA como seletor de atributos. Isto é, ao invés de se atribuir pesos reais para cada atributo, pretende-se atribuir apenas 0 ou 1. De tal maneira, será possível saber, de forma precisa, quais atributos realmente influenciam a verificação e a classificação e se poderá comparar com outros trabalhos da literatura que fazem uso dessa técnica de seleção de atributos. Também se espera poder aplicar outros algoritmos conhecidos de otimização, por exemplo: *Firefly* e Otimização de Enxame de Partículas (do inglês, *Particle Swarm Optimization*, PSO); e comparar seus resultados com os obtidos pelo GA neste trabalho. Por fim, espera-se analisar estatisticamente os resultados e avaliar a relevância do método proposto.

Por fim, resultados parciais deste trabalho foram publicados em Silva et al. (2015). Além disso, a utilização do GA pelo autor deste trabalho também foi abordada em Silva et al. (2015), este voltado para o reconhecimento biométrico através de impressões digitais. Espera-se ainda publicar em breve um outro trabalho apresentando a base BioPass-UFPB, tornando-a pública para a comunidade acadêmica com os resultados publicados até aqui.

Referências

- Ackley, D. & Littman, M., 1991. Interactions between learning and evolution. *Artificial Life II*, 10, pp.487-509.
- Antonelli, A., Cappelli, R., Maio, D. & Maltoni, D., 2006. Fake Finger Detection by Skin Distortion Analysis. *IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY*, Setembro. pp.360-73.
- Back, T. & Schwefel, H.-P., 1993. An overview of evolutionary algorithms for parameter optimization. *Evolutionary computation*, 1(1), pp.1-23.
- Banzhaf, W.e.a., 1999. *Foundations of genetic algorithms*. Morgan Kaufmann Publishers Inc.
- Batista, L.V., 2006. *Notas de Aula da Disciplina Introdução ao Processamento Digital de Imagens*. João Pessoa.
- Behroozmand, R. & Almasganj, F., 2007. Optimal selection of wavelet-packet-based features using genetic algorithm in pathological assessment of patients. *Computers in Biology and Medicine*, 37(4), pp.474-85.
- Bergadano, F., Gunneti, D. & Picardi, C., 2002. User authentication through keystroke dynamics. *ACM Transactions on Information and System Security*, 5(4), pp.367-97.
- BioID, 2005. *Humamscan*. [Online] Available at: <http://www.bioid.com> [Accessed Março 2005].
- BITE, 2005. *Global biometric market and industry report*. Biometric Identification Technology Ethics.
- Bolle, R.M. et al., 2004. *Guide to Biometrics*. 1st ed. Springer Professional Computing.
- Bolle, R.M., Connell, J.H. & Ratha, N.K., 2002. Biometric perils and patches. *Pattern Recognition*, 35, pp.2727-38.
- Booker, L.B., Goldberg, D.E. & Holland, J.H., 1989. Classifier systems and genetic algorithms. *Artificial intelligence*, 40(1), pp.235-82.
- Burge, M. & Burger, W., 2000. Ear biometrics in computer vision. *International Conference on Pattern Recognition*, pp.2822-26.
- Chen, H. & Jain, A.K., 2005. Dental biometrics: Alignment and matching of dental radiographs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8), pp.1319-26.
- Clarke, R., 1994. Human identification in information systems: management challenges and public policy issues. *Information Technology & People*, 7(4), pp.6-37.
- Costa, L.R., Obelheiro, R.R. & Fraga, J.S., 2006. Introdução à Biometria. In *Livro texto dos Minicursos do VI Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais (SBSeg2006)*. Porto Alegre: SBC. pp.103-51.
- Darwin, C., 1951. *The origin of species*. Oxford University Press Oxford.

Dawid, H., 1996. *Adaptive learning by genetic algorithms: Analytical results and applications to economic models*. 2nd ed. New York: Springer-Verlag.

do Nascimento, M.V., Batista, L.V. & Cavalcanti, N., 2014. Comparative study of learning algorithms for recognition by hand geometry. *Systems, Man and Cybernetics (SMC), 2014 IEEE International Conference on. IEEE*, pp.423-28.

Dorigo, M. & Gambardella, L.M., 1997. Ant colony system: a cooperative learning approach to the traveling salesman problem. *Evolutionary Computation, IEEE Transactions on*, 1(1), pp.53-66.

Doyle, S. et al., 2007. Automated grading of prostate cancer using architectural and textural image features. *Biomedical Imaging: From Nano to Macro, 2007. ISBI 2007. 4th IEEE International Symposium on. IEEE*, pp.1284-87.

Du, N., Qi, M., Zhang, Y. & Kong, J., 2009. Palmprint Verification Based on Specific-user. *Electronic Commerce and Security*, 1, pp.314-17.

Fink, G.A., Wienecke, M. & Sagerer, G., 2001. Video-based on-line handwriting recognition. *International Conference on Document Analysis and Recognition*, pp.226-30.

Fix, E. & Jr., J.L.H., 1951. Discriminatory analysis-nonparametric discrimination: consistency properties. *CALIFORNIA UNIV BERKELEY*.

Franklin, S.E., Maudie, A.J. & Lavigne, M.B., 2001. Using spatial co-occurrence texture to increase forest structure and species composition classification accuracy. *Photogrammetric Engineering and Remote Sensing*, 67(7), pp.849-56.

Golberg, D.E., 1989. *Genetic algorithms in search, optimization, and machine learning*. Addison Wesley.

Gonzales, R.C., Woods, R.E. & Eddins, S.L., 2004. *Digital Image Processing Using Matlab*. New Jersey, USA: Pearson Education.

Goranin, N. & Cenys, A., 2010. Evolutionary Algorithms Application Analysis in Biometric Systems. *Journal of Engineering Science and Technology Review*, 3(1), pp.70-79.

Haralick, R.M., Shanmugam, K. & Dinstein, I.H., 1973. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, (6), pp.610-21.

Hill, R.", 2002. Retina Identification. *Biometrics: Personal Identification in Networked Society*, pp.123-41.

Holland, J.H., 1962. Outline for a logical theory of adaptive systems. *Journal of the ACM (JACM)*, 9(3), pp.297-314.

Holland, J.H., 1975. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. 1st ed. U Michigan Press.

- Jain, A., Ross, A. & Prabhakar, S., 2004. An Introduction to Biometric Recognition. *IEEE Transactions on Circuits and Systems for Video Technology, Special Issue on Image-and Video-Based Biometrics*, 14(1).
- Jin, Y. et al., 2008. Palmprint Verification Based on Genetic Algorithm. *Intelligent System and Knowledge Engineering*, 1, pp.471-75.
- Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. *International joint Conference on artificial intelligence*, 14(2), pp.1137-45.
- Kong, A., Zhang, D. & Kamel, M., 2009. A survey of palmprint recognition. *Elsevier*, pp.1408-18.
- Korotkaya, Z., 2003. *Biometric person authentication: Odor*. Inner report. Lappeenranta University of Technology. in "Advanced Topics in Information Processing: Biometric Person Authentication".
- Kulp, A. & Braskamp, H., 2012. *Selecting identification technology*. [Online] Available at: <http://nedap-securitymanagement.com/en/technology-nedapmenu-90/identificationtechnology-nedapmenu-61.html> [Accessed Agosto 2012].
- Leung, F.H., Lam, H.K., Ling, S.H. & Tam, P.K., 2003. Tuning of the structure and parameters of a neural network using an improved genetic algorithm. *Neural Networks, IEEE Transactions on*, 14(1), pp.79-88.
- Ling, L.L., 2003. *Unicamp*. [Online] Available at: http://www.unicamp.br/unicamp/unicamp_hoje/ju/setembro2003/ju229pg8b.html [Accessed Março 2015].
- Lordão, F.A.F., 2009. *Reconhecimento de Formas Utilizando Modelos de Compressão de Dados e Espaço de Escalas de Curvatura*. João Pessoa.
- Luque-Baena, R.M. et al., 2013. Assessment of geometric features for individual identification and verification in biometric hand systems. *Expert systems with applications*, 40(9), pp.3580-94.
- Luque, R.M., Elizondo, D., López-Rubio, E. & Palomo, E.J., 2011. GA-based feature selection approach in biometric hand systems. *Neural Networks (IJCNN), The 2011 International Joint Conference on. IEEE*, pp.246-53.
- Maltoni, D..M.D..J.A.K..&.S., 2009. *Handbook of Fingerprint Recognition*. Springer Publishing Company, Incorporated.
- Manikas, W., Ashenayi, K. & Wainwright, R., 2007. Genetic algorithms for autonomous robot navigation. *Instrumentation & Measurement Magazine, IEEE*, 10(6), pp.26-31.
- Mazumder, P. & Rudnick, E.M., 1999. *Genetic Algorithms for VLSI Design, Layout and Test Automation*. New Jersey: Prentice Hall PTR.
- McLachlan, G.J., Bean, R.W. & Peel, D., 2002. A mixture model-based approach to the clustering of microarray expression data. *Bioinformatics*, 18(3), pp.413-22.
- Mitchell, M., 1995. Genetic algorithms: An overview. *Complexity*, 1(1), pp.31-39.

Mitchell, M., Crutchfield, J.P. & Das, R., 1996. Evolving cellular automata with genetic algorithms: A review of recent work. *Proceedings of the First International Conference on Evolutionary Computation and Its Applications*.

Miyamoto, E. & Merryman, T., 2005. Fast calculation of Haralick texture features. *Human computer interaction institute, Carnegie Mellon University, Pittsburgh, USA. Japanese restaurant office*.

Ojala, T., Pietikäinen, M. & Harwood, D., 1996. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1), pp.51-59.

Ojala, T., Pietikäinen, M. & Mäenpää, T., 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7), pp.971-87.

OpenCV, 2008. *Face Recognition With OpenCV*. [Online] Available at: http://docs.opencv.org/modules/contrib/doc/facerec/facerec_tutorial.html#local-binary-patterns-histograms-in-opencv [Accessed Agosto 2015].

Palubinskas, G., Lucas, R.M., Foody, G.M. & Curran, P.J., 1995. An evaluation of fuzzy and texture-based classification approaches for mapping regenerating tropical forest classes from Landsat-TM data. *International Journal of Remote Sensing*, 16(4), pp.747-59.

Patrick, E.A., 1972. *Fundamentals of pattern recognition*.

Pedersen, J.T. & Moulton, J., 1996. Genetic algorithms for protein structure prediction. *Current Opinion in Structural Biology*, 6(2), pp.227-31.

Phillips, P.J. et al., 2002. The gait identification challenge problem: Data sets and baseline algorithm. *International Conference on Pattern Recognition*, pp.385-88.

Prokoshi, F.J. & Riedel, R., 1999. Infrared identification of faces and body parts. In Jain, A.K., Bolle, R.M. & Pankanti, S. *Biometrics: Personal Identification in Networked Society*. Boston, MA, USA: Kluwer Academic. Ch. 9.

Rahman, A., Anwar, F. & Azad, S., 2008. A simple and effective technique for human verification with hand geometry. *Computer and Communication Engineering, 2008. ICCCE 2008. 8. International Conference on. IEEE*, pp.1177-80.

Richter Jr, C.W. & Sheblé, G.B., 1998. Genetic algorithm evolution of utility bidding strategies for the competitive marketplace. *Power Systems, IEEE Transactions on*, 13(1), pp.256-61.

Roberts, C., 2007. Biometric attack vectors and defenses. *Computers and Security Journal*, pp.14-25.

Ross, A.A., Nandakumar, K. & Jain, A.K., 2006. *Handbook of Multibiometrics (International Series on Biometrics)*. Secaucus, New Jersey, USA: Springer-Verlag.

Sette, P.G.C. & Maillard, P., 2014. Análise de textura de imagem de alta resolução para aprimorar a acurácia da classificação da mata atlântica no sul da Bahia. *XV Simpósio Brasileiro de Sensoriamento Remoto*, pp.2020-27.

Silva, A.G.A. et al., 2015. Analysis of a Genetic Algorithm-based Approach in the Optimization of the SourceAFIS's Matching Algorithm. *International Conference on Scientific Computing*, Julho. pp.23-28.

Silva, A.G.A. et al., 2015. Analysis of the Performance Improvement Obtained by a Genetic Algorithm-based Approach on a Hand Geometry Dataset. *International Conference on Artificial Intelligence*, Julho. pp.125-30.

Smith, J., 2013. *Java GALib*. [Online] Available at: <http://java-galib.sourceforge.net/> [Accessed 2015].

Subcommittee on Biometrics of National Science and Technology Council (NSTC), 2006. *Hand Geometry*. [Online] Available at: <http://www.biometrics.gov/documents/handgeometry.pdf> [Accessed 20 Junho 2012].

Sun, Z., Bebis, G. & Miller, R., 2004. Object detection using feature subset selection. *Pattern recognition*, 37(11), pp.2165-76.

Thorpe, J., van Oorschot, P. & Somayaji, A., 2005. Passthoughts: Authentication with our minds. *Proceedings of the New Security Paradigms Workshop*.

Valid, 2006. *Visual audio lip-motion identification*. [Online] Available at: <http://www.validbiometrics.com> [Accessed Julho 2006].

Whitley, L.D., 1993. *Foundations of genetic algorithms 2*. Morgan Kauffmann.

Wu, X., Wang, K. & Zhang, D., 2004. A Novel Approach of Line Feature Representation and Matching for Palmprint Recognition. *Journal of Software*, pp.869-80.

Yunyong, P. & Yodchanan, W., 2009. Palmprint Image Enhancement Using Phase Congruency. *International Conference on Robotics and Biomimetics*, 21-26 Fevereiro. pp.1643-46.

Zang, D.D., 2004. *Palmprint Authentication*. Springer.

Zhang, D., Kong, W.-K., You, J. & Wong, M., 2003. Online Palmprint Identification. *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, Setembro. pp.1041-50.

APÊNDICE A: RESULTADOS DO GA NA VERIFICAÇÃO

Tabela 28: Resultado da Aplicação do GA utilizando as 10 primeiras classes para treinamento e as 90 últimas classes para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo (h:m:s:ms)	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,011111111	0:00.265	0,029675406
						0,011111111	0:00.234	0,013782772
						0,011111111	0:00.250	0,016697878
2	100	0,5	10	0,5	Dois Pontos	0,011111111	0:00.265	0,015087391
						0,011111111	0:00.234	0,013907615
						0,011111111	0:00.249	0,018127341
3	100	0,5	10	0,5	Uniforme	0,011111111	0:00.484	0,017877653
						0,011111111	0:00.249	0,027191011
						0,011111111	0:00.234	0,018377029
4	100	0,5	100	0,5	Uniforme	0,011111111	0:01.700	0,014712859
						0,011111111	0:01.669	0,016697878
						0,011111111	0:01.826	0,018002497
5	100	0,5	1000	0,5	Uniforme	0,011111111	0:15.804	0,014837703
						0,011111111	0:15.174	0,022471910
						0,011111111	0:16.940	0,044943820
6	1.000	0,5	100	0,5	Uniforme	0,011111111	0:17.609	0,033339576
						0,011111111	0:17.145	0,023776529
						0,011111111	0:17.669	0,014837703
7	1.000	0,5	1000	0,5	Uniforme	0,011111111	3:44.870	0,015642946
						0,011111111	3:47.449	0,015461923
						0,011111111	3:52.171	0,023776529
8	100	0,75	100	0,5	Uniforme	0,011111111	0:01.903	0,023401998
						0,011111111	0:01.716	0,016822722
						0,011111111	0:01.951	0,031285893
9	100	1	100	0,5	Uniforme	0,011111111	0:01.872	0,013533084
						0,011111111	0:01.825	0,022471910
						0,011111111	0:01.781	0,045749064
10	100	0,5	100	0,75	Uniforme	0,011111111	0:01.716	0,021666667
						0,011111111	0:01.562	0,015087391
						0,011111111	0:01.622	0,011173533
11	100	0,5	100	1	Uniforme	0,011111111	0:01.827	0,027690387
						0,011111111	0:01.747	0,018932584
						0,011111111	0:01.716	0,015461923
12	100	1	100	1	Uniforme	0,011111111	0:01.815	0,01893258
						0,011111111	0:01.781	0,02588639
						0,011111111	0:01.685	0,02011236

Tabela 29: Resultado da Aplicação do GA utilizando as 20 primeiras classes para treinamento e as 80 últimas classes para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,00513158	0:00.651	-
						0,00263158	0:00.515	0,01446994
						0,00513158	0:00.521	-
2	100	0,5	10	0,5	Dois Pontos	0,00513158	0:00.546	-
						0,00513158	0:00.530	-
						0,00513158	0:00.527	-
3	100	0,5	10	0,5	Uniforme	0,00513158	0:00.549	-
						0,00263158	0:00.525	0,01651108
						0,00263158	0:00.528	0,01761076
4	100	0,5	100	0,5	Uniforme	0,00263158	0:04.242	0,01257911
						0,00263158	0:04.330	0,01572785
						0,00513158	0:03.818	-
5	100	0,5	1000	0,5	Uniforme	0,00263158	0:41.536	0,02201741
						0,00263158	0:40.616	0,01981804
						0,00263158	0:42.133	0,02579114
6	1.000	0,5	100	0,5	Uniforme	0,00263158	0:44.223	0,01415348
						0,00263158	0:45.404	0,01824367
						0,00263158	0:45.694	0,01541139
7	1.000	0,5	1000	0,5	Uniforme	0,00263158	8:17.861	0,02185918
						0,00263158	9:30.419	0,01509494
						0,00263158	8:37.758	0,01509494
8	100	0,75	100	0,5	Uniforme	0,00263158	0:04.960	0,01195411
						0,00263158	0:04.908	0,01509494
						0,00263158	0:04.247	0,01273734
9	100	1	100	0,5	Uniforme	0,00263158	0:04.987	0,01651108
						0,00263158	0:05.077	0,02264241
						0,00263158	0:04.440	0,01666930
10	100	0,5	100	0,75	Uniforme	0,00263158	0:04.150	0,01635285
						0,00263158	0:04.366	0,01635285
						0,00263158	0:04.327	0,01635285
11	100	0,5	100	1	Uniforme	0,00263158	0:04.922	0,01179589
						0,00263158	0:04.195	0,01289557
						0,00263158	0:04.227	0,02138449
12	100	1	100	1	Uniforme	0,00263158	0:04.636	0,01635285
						0,00263158	0:04.427	0,01761076
						0,00263158	0:04.449	0,01431171

Tabela 30: Resultado da Aplicação do GA utilizando as 30 primeiras classes para treinamento e as 70 últimas classes para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,00626437	0:00.978	-
						0,00396552	0:00.896	-
						0,00678161	0:00.884	-
2	100	0,5	10	0,5	Dois Pontos	0,00114943	0:00.943	0,01726708
						0,00396552	0:00.906	-
						0,00396552	0:00.910	-
3	100	0,5	10	0,5	Uniforme	0,00229885	0:00.934	-
						0,00114943	0:00.914	0,02189441
						0,00396552	0:00.902	-
4	100	0,5	100	0,5	Uniforme	0,00114943	0:07.407	0,01788820
						0,00114943	0:07.827	0,01860248
						0,00396552	0:08.104	-
5	100	0,5	1000	0,5	Uniforme	0,00114943	1:14.677	0,02076605
						0,00114943	1:27.919	0,02168737
						0,00114943	1:15.634	0,02343685
6	1.000	0,5	100	0,5	Uniforme	0,00114943	1:17.569	0,01726708
						0,00114943	1:23.862	0,01480331
						0,00114943	1:20.752	0,01613872
7	1.000	0,5	1000	0,5	Uniforme	0,00114943	14:36.021	0,01726708
						0,00114943	16:04.010	0,01613872
						0,00114943	16:07.562	0,01860248
8	100	0,75	100	0,5	Uniforme	0,00114943	0:07.604	0,02385093
						0,00114943	0:07.594	0,01388199
						0,00114943	0:07.481	0,01326087
9	100	1	100	0,5	Uniforme	0,00114943	0:07.582	0,01726708
						0,00114943	0:08.145	0,01572464
						0,00114943	0:07.459	0,02221532
10	100	0,5	100	0,75	Uniforme	0,00114943	0:07.527	0,01059006
						0,00114943	0:07.082	0,01408903
						0,00114943	0:08.299	0,02148033
11	100	0,5	100	1	Uniforme	0,00114943	0:07.711	0,02014493
						0,00114943	0:07.195	0,01788820
						0,00114943	0:08.252	0,01726708
12	100	1	100	1	Uniforme	0,00114943	0:07.608	0,01583851
						0,00114943	0:07.949	0,02014493
						0,00114943	0:07.649	0,02539337

Tabela 31: Resultado da Aplicação do GA utilizando as 40 primeiras classes para treinamento e as 60 últimas classes para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,005064	0:01.421	-
						0,003814	0:01.404	-
						0,003814	0:01.389	-
2	100	0,5	10	0,5	Dois Pontos	0,005064	0:01.426	-
						0,005064	0:01.452	-
						0,005064	0:01.469	-
3	100	0,5	10	0,5	Uniforme	0,005064	0:01.500	-
						0,003173	0:01.413	-
						0,002532	0:01.411	-
4	100	0,5	100	0,5	Uniforme	0,000641	0:11.140	0,013446
						0,002532	0:11.133	-
						0,002532	0:11.453	-
5	100	0,5	1000	0,5	Uniforme	0,000641	1:51.272	0,015127
						0,000641	1:53.960	0,025212
						0,002532	1:58.994	-
6	1.000	0,5	100	0,5	Uniforme	0,000641	2:03.920	0,012048
						0,000641	2:07.073	0,018771
						0,000641	2:02.372	0,013446
7	1.000	0,5	1000	0,5	Uniforme	0,000641	23:01.718	0,02185
						0,000641	23:33.414	0,02185
						0,000641	22:08.495	0,020169
8	100	0,75	100	0,5	Uniforme	0,002532	0:11.618	-
						0,001282	0:12.036	-
						0,002532	0:11.720	-
9	100	1	100	0,5	Uniforme	0,000641	0:12.563	0,022133
						0,000641	0:12.381	0,015127
						0,002532	0:11.605	-
10	100	0,5	100	0,75	Uniforme	0,002532	0:12.937	-
						0,000641	0:11.069	0,012048
						0,000641	0:11.710	0,020169
11	100	0,5	100	1	Uniforme	0,000641	0:11.839	0,022133
						0,000641	0:11.946	0,022147
						0,000641	0:11.835	0,026342
12	100	1	100	1	Uniforme	0,000641	0:13.167	0,014294
						0,000641	0:11.993	0,016257
						0,000641	0:12.319	0,01541

Tabela 32: Resultado da Aplicação do GA utilizando as 50 primeiras classes para treinamento e as 50 últimas classes para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,004449	0:02.072	-
						0,003041	0:01.968	-
						0,004449	0:02.058	-
2	100	0,5	10	0,5	Dois Pontos	0,003041	0:02.044	-
						0,002224	0:02.014	-
						0,004041	0:01.976	-
3	100	0,5	10	0,5	Uniforme	0,003857	0:02.080	-
						0,002224	0:02.051	-
						0,002633	0:02.052	-
4	100	0,5	100	0,5	Uniforme	0,000408	0:15.952	0,016163
						0,000408	0:16.066	0,03051
						0,000408	0:16.931	0,016571
5	100	0,5	1000	0,5	Uniforme	0,000408	2:49.644	0,022837
						0,000408	3:21.172	0,02202
						0,000408	3:13.854	0,026061
6	1.000	0,5	100	0,5	Uniforme	0,000408	2:53.255	0,02202
						0,000408	2:51.483	0,02102
						0,000408	3:02.035	0,020204
7	1.000	0,5	1000	0,5	Uniforme	0,000408	32:37.429	0,018796
						0,000408	31:41.054	0,026878
						0,000408	30:33.025	0,024245
8	100	0,75	100	0,5	Uniforme	0,000408	0:17.834	0,012122
						0,000408	0:16.104	0,024245
						0,002224	0:18.216	
9	100	1	100	0,5	Uniforme	0,000408	0:17.721	0,01698
						0,000408	0:18.139	0,016163
						0,000408	0:18.714	0,026878
10	100	0,5	100	0,75	Uniforme	0,002224	0:16.916	-
						0,002224	0:16.876	-
						0,000408	0:16.128	0,018388
11	100	0,5	100	1	Uniforme	0,002224	0:16.035	-
						0,000408	0:17.216	0,025061
						0,000408	0:17.340	0,024245
12	100	1	100	1	Uniforme	0,000408	0:18.834	0,018796
						0,000408	0:17.180	0,018388
						0,000816	0:17.385	-

Tabela 33: Resultado da Aplicação do GA utilizando as 60 primeiras classes para treinamento e as 40 últimas classes para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,005042	0:02.888	-
						0,005042	0:02.798	-
						0,005042	0:02.753	-
2	100	0,5	10	0,5	Dois Pontos	0,005042	0:02.985	-
						0,004492	0:02.722	-
						0,005042	0:02.947	-
3	100	0,5	10	0,5	Uniforme	0,003644	0:02.904	-
						0,002811	0:02.818	-
						0,003362	0:02.772	-
4	100	0,5	100	0,5	Uniforme	0,001681	0:23.106	-
						0,001681	0:21.999	-
						0,001963	0:25.477	-
5	100	0,5	1000	0,5	Uniforme	0,000565	3:48.199	-
						0,000847	4:00.952	-
						0,001681	4:10.480	-
6	1.000	0,5	100	0,5	Uniforme	0,000282	4:26.536	0,018365
						0,000282	4:19.815	0,015192
						0,000282	4:10.639	0,020256
7	1.000	0,5	1000	0,5	Uniforme	0,000282	43:56.573	0,015192
						0,000282	43:49.964	0,019006
						0,000282	46:44.940	0,020256
8	100	0,75	100	0,5	Uniforme	0,001681	0:25.213	-
						0,001681	0:24.332	-
						0,001681	0:23.208	-
9	100	1	100	0,5	Uniforme	0,000565	0:23.638	-
						0,001681	0:24.906	-
						0,001681	0:24.095	-
10	100	0,5	100	0,75	Uniforme	0,001681	0:23.081	-
						0,002528	0:23.078	-
						0,000847	0:23.187	-
11	100	0,5	100	1	Uniforme	0,001681	0:26.723	-
						0,001681	0:22.870	-
						0,001681	0:22.530	-
12	100	1	100	1	Uniforme	0,000565	0:24.134	-
						0,001681	0:22.220	-
						0,001681	0:22.589	-

Tabela 34: Resultado da Aplicação do GA utilizando as 70 primeiras classes para treinamento e as 30 últimas classes para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,00442	0:03.694	-
						0,004627	0:03.471	-
						0,00442	0:03.611	-
2	100	0,5	10	0,5	Dois Pontos	0,005756	0:03.544	-
						0,00442	0:03.500	-
						0,005041	0:03.511	-
3	100	0,5	10	0,5	Uniforme	0,00442	0:03.571	-
						0,003499	0:03.611	-
						0,003499	0:03.620	-
4	100	0,5	100	0,5	Uniforme	0,002878	0:31.726	-
						0,002878	0:31.773	-
						0,001542	0:29.409	-
5	100	0,5	1000	0,5	Uniforme	0,000207	4:34.664	0,02431
						0,001035	5:39.231	-
						0,001542	5:01.919	-
6	1.000	0,5	100	0,5	Uniforme	0,000207	5:21.781	0,027126
						0,000414	5:26.154	-
						0,000207	5:13.870	0,02431
7	1.000	0,5	1000	0,5	Uniforme	0,000207	55:24.606	0,02431
						0,000207	1:00:34.626	0,02431
						0,000207	55:36.611	0,027126
8	100	0,75	100	0,5	Uniforme	0,001542	0:29.941	-
						0,000414	0:30.920	-
						0,001542	0:29.216	-
9	100	1	100	0,5	Uniforme	0,001542	0:30.533	-
						0,001542	0:31.548	-
						0,001542	0:32.033	-
10	100	0,5	100	0,75	Uniforme	0,001242	0:32.060	-
						0,001749	0:29.500	-
						0,001542	0:30.623	-
11	100	0,5	100	1	Uniforme	0,001542	0:30.782	-
						0,000828	0:31.721	-
						0,001542	0:31.185	-
12	100	1	100	1	Uniforme	0,000207	0:33.921	0,027126
						0,001542	0:30.037	-
						0,001542	0:32.972	-

Tabela 35: Resultado da Aplicação do GA utilizando as 80 primeiras classes para treinamento e as 20 últimas classes para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,005348	0:04.778	-
						0,005831	0:04.497	-
						0,004723	0:04.512	-
2	100	0,5	10	0,5	Dois Pontos	0,005506	0:04.517	-
						0,003774	0:04.740	-
						0,003774	0:04.440	-
3	100	0,5	10	0,5	Uniforme	0,004407	0:04.607	-
						0,002516	0:04.618	-
						0,003774	0:04.705	-
4	100	0,5	100	0,5	Uniforme	0,002516	0:40.987	-
						0,001258	0:38.649	-
						0,002516	0:39.191	-
5	100	0,5	1000	0,5	Uniforme	0,000475	6:36.359	-
						0,001258	6:44.112	-
						0,001258	6:43.419	-
6	1.000	0,5	100	0,5	Uniforme	0,000633	7:10.327	-
						0,000475	7:04.982	-
						0,000158	6:41.892	0,035921
7	1.000	0,5	1000	0,5	Uniforme	0,000158	1:13:49.872	0,041053
						0,000158	1:14:53.485	0,041053
						0,000158	1:13:07.709	0,035921
8	100	0,75	100	0,5	Uniforme	0,002516	0:38.311	-
						0,001416	0:43.141	-
						0,002516	0:40.834	-
9	100	1	100	0,5	Uniforme	0,000475	0:38.857	-
						0,001258	0:41.819	-
						0,001258	0:39.355	-
10	100	0,5	100	0,75	Uniforme	0,001891	0:49.235	-
						0,001258	0:45.862	-
						0,001258	0:43.768	-
11	100	0,5	100	1	Uniforme	0,002207	0:40.173	-
						0,001258	0:38.580	-
						0,001416	0:37.820	-
12	100	1	100	1	Uniforme	0,001258	0:42.769	-
						0,001258	0:41.468	-
						0,002516	0:39.026	-

Tabela 36: Resultado da Aplicação do GA utilizando as 90 primeiras classes para treinamento e as 10 últimas classes para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,006024	0:05.507	-
						0,005649	0:05.752	-
						0,005094	0:05.413	-
2	100	0,5	10	0,5	Dois Pontos	0,004663	0:05.574	-
						0,004288	0:05.496	-
						0,005649	0:05.586	-
3	100	0,5	10	0,5	Uniforme	0,004719	0:05.776	-
						0,004164	0:05.874	-
						0,003664	0:05.780	-
4	100	0,5	100	0,5	Uniforme	0,002235	0:47.407	-
						0,002235	0:46.094	-
						0,002235	0:43.286	-
5	100	0,5	1000	0,5	Uniforme	0,001305	7:01.044	-
						0,000499	7:03.043	-
						0,002235	8:05.457	-
6	1.000	0,5	100	0,5	Uniforme	0,00118	8:12.257	-
						0,00118	8:31.990	-
						0,00118	8:12.590	-
7	1.000	0,5	1000	0,5	Uniforme	0,000375	0:058542	-
						0,00025	0:06103	-
						0,000125	0:058368	0,079444
8	100	0,75	100	0,5	Uniforme	0,00118	0:52.319	-
						0,001929	0:51.305	-
						0,001305	0:48.986	-
9	100	1	100	0,5	Uniforme	0,00118	0:50.336	-
						0,00118	0:48.950	-
						0,00118	0:53.786	-
10	100	0,5	100	0,75	Uniforme	0,002235	0:58.039	-
						0,003414	0:50.230	-
						0,00118	0:51.496	-
11	100	0,5	100	1	Uniforme	0,00118	0:47.043	-
						0,002235	0:45.904	-
						0,001305	0:50.266	-
12	100	1	100	1	Uniforme	0,000749	0:50.583	-
						0,00118	0:46.161	-
						0,00118	0:53.952	-

Tabela 37: Resultado da Aplicação do GA utilizando as 1ª, 2ª e 3ª instâncias de cada classe para treinamento e as 4ª e 5ª instâncias para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,013434	0:06.107	-
						0,010152	0:06.013	-
						0,013434	0:05.947	-
2	100	0,5	10	0,5	Dois Pontos	0,011566	0:06.059	-
						0,010051	0:06.328	-
						0,013434	0:05.894	-
3	100	0,5	10	0,5	Uniforme	0,010051	0:06.066	-
						0,010051	0:05.949	-
						0,007677	0:06.157	-
4	100	0,5	100	0,5	Uniforme	0,006768	0:51.485	-
						0,006768	0:48.861	-
						0,006768	0:56.794	-
5	100	0,5	1000	0,5	Uniforme	0,006768	9:00.436	-
						0,003384	8:29.502	0,000404
						0,006768	8:13.920	-
6	1.000	0,5	100	0,5	Uniforme	0,004495	9:15.576	-
						0,003485	8:54.564	-
						0,003384	9:31.054	0,000606
7	1.000	0,5	1000	0,5	Uniforme	0,006768	1:46:32.137	-
						0,003384	1:37:13.426	0,000606
						0,003384	1:41:36.778	0,000404
8	100	0,75	100	0,5	Uniforme	0,006768	0:54.858	-
						0,006768	0:53.869	-
						0,006768	0:57.998	-
9	100	1	100	0,5	Uniforme	0,006768	0:53.236	-
						0,003384	0:50.895	0,001212
						0,003384	0:53.829	0,001717
10	100	0,5	100	0,75	Uniforme	0,006768	0:52.172	-
						0,006768	0:51.013	-
						0,006768	0:48.560	-
11	100	0,5	100	1	Uniforme	0,006768	0:52.152	-
						0,010051	1:02.688	-
						0,006768	0:53.687	-
12	100	1	100	1	Uniforme	0,006768	0:53.133	-
						0,00697	0:54.757	-

Tabela 38: Resultado da Aplicação do GA utilizando as 1ª, 2ª e 4ª instâncias de cada classe para treinamento e as 3ª e 5ª instâncias para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,014949	0:06.082	-
						0,013434	0:06.026	-
						0,013434	0:06.240	-
2	100	0,5	10	0,5	Dois Pontos	0,010051	0:06.165	-
						0,011263	0:05.849	-
						0,010051	0:05.944	-
3	100	0,5	10	0,5	Uniforme	0,010051	0:06.043	-
						0,011667	0:06.226	-
						0,010051	0:06.028	-
4	100	0,5	100	0,5	Uniforme	0,010051	0:52.962	-
						0,010051	0:49.013	-
						0,010051	0:55.638	-
5	100	0,5	1000	0,5	Uniforme	0,006768	8:22.466	0,001313
						0,006768	9:18.399	0,00101
						0,010051	0:06.516	-
6	1.000	0,5	100	0,5	Uniforme	0,006768	9:53.730	0,000101
						0,006768	9:34.157	0,000202
						0,006768	9:51.633	0,000101
7	1.000	0,5	1000	0,5	Uniforme	0,006768	1:58:39.079	0,000505
						0,006768	2:00:53.873	0,000707
						0,006768	1:55:30.580	0,000303
8	100	0,75	100	0,5	Uniforme	0,006768	0:52.217	0,0013
						0,010051	0:59.577	-
						0,010051	0:57.197	-
9	100	1	100	0,5	Uniforme	0,006768	0:55.515	0,000909
						0,006768	0:57.609	0,000303
						0,006768	0:53.786	0,001111
10	100	0,5	100	0,75	Uniforme	0,006768	0:53.019	0,000202
						0,010051	0:58.140	-
						0,006768	0:58.684	0,000808
11	100	0,5	100	1	Uniforme	0,006768	0:55.656	0,000808
						0,010051	0:52.154	-
						0,006768	0:56.672	0,001919
12	100	1	100	1	Uniforme	0,010051	1:07.581	-
						0,010051	1:06.886	-
						0,010051	1:08.608	-

Tabela 39: Resultado da Aplicação do GA utilizando as 1ª, 2ª e 5ª instâncias de cada classe para treinamento e as 3ª e 4ª instâncias para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,011263	0:06.237	-
						0,010657	0:05.962	-
						0,012879	0:05.926	-
2	100	0,5	10	0,5	Dois Pontos	0,010051	0:06.103	-
						0,012273	0:05.982	-
						0,007576	0:06.214	-
3	100	0,5	10	0,5	Uniforme	0,008788	0:06.263	-
						0,008687	0:06.067	-
						0,007273	0:06.009	-
4	100	0,5	100	0,5	Uniforme	0,006768	0:50.026	-
						0,006768	0:54.473	-
						0,006768	0:48.766	-
5	100	0,5	1000	0,5	Uniforme	0,006768	8:13.791	-
						0,006768	8:04.319	-
						0,006768	8:27.026	-
6	1.000	0,5	100	0,5	Uniforme	0,006768	9:10.688	-
						0,006768	9:17.370	-
						0,006768	9:21.094	-
7	1.000	0,5	1000	0,5	Uniforme	0,003384	1:35:06.681	0,001717
						0,006768	1:53:08.169	-
						0,006768	2:00:20.159	-
8	100	0,75	100	0,5	Uniforme	0,006768	0:51.463	-
						0,006768	0:51.195	-
						0,006768	0:52.434	-
9	100	1	100	0,5	Uniforme	0,006768	0:50.799	-
						0,006768	0:56.012	-
						0,006768	0:56.679	-
10	100	0,5	100	0,75	Uniforme	0,007879	0:50.213	-
						0,010051	0:50.206	-
						0,006768	0:47.615	-
11	100	0,5	100	1	Uniforme	0,006768	0:49.332	-
						0,006768	0:50.339	-
						0,006768	0:48.253	-
12	100	1	100	1	Uniforme	0,006768	0:51.983	-
						0,006768	0:48.985	-
						0,006768	0:53.128	-

Tabela 40: Resultado da Aplicação do GA utilizando as 1ª, 3ª e 4ª instâncias de cada classe para treinamento e as 2ª e 5ª instâncias para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,010051	0:06.084	-
						0,010051	0:06.323	-
						0,009798	0:05.991	-
2	100	0,5	10	0,5	Dois Pontos	0,00697	0:06.026	-
						0,007677	0:05.927	-
						0,006768	0:05.915	0,010051
3	100	0,5	10	0,5	Uniforme	0,006768	0:05.981	0,010354
						0,006768	0:06.139	0,014596
						0,006768	0:06.111	0,010657
4	100	0,5	100	0,5	Uniforme	0,006768	0:47.911	0,005051
						0,006768	0:52.741	0,011364
						0,006768	0:51.476	0,002323
5	100	0,5	1000	0,5	Uniforme	0,006768	9:04.397	0,010455
						0,006768	8:05.121	0,010253
						0,006768	8:11.879	0,020202
6	1.000	0,5	100	0,5	Uniforme	0,006768	8:54.541	0,010758
						0,006768	9:08.601	0,004646
						0,006768	9:46.776	0,003232
7	1.000	0,5	1000	0,5	Uniforme	0,006768	2:00:53.006	0,013687
						0,006768	1:54:58.380	0,011768
						0,006768	1:50:43.532	0,003939
8	100	0,75	100	0,5	Uniforme	0,006768	0:51.074	0,010354
						0,006768	0:53.605	0,010253
						0,006768	0:55.569	0,005051
9	100	1	100	0,5	Uniforme	0,006768	0:56.643	0,020404
						0,006768	0:56.250	0,010051
						0,006768	0:57.925	0,004646
10	100	0,5	100	0,75	Uniforme	0,006768	0:53.279	0,011364
						0,006768	0:52.182	0,00202
						0,006768	0:58.415	0,013687
11	100	0,5	100	1	Uniforme	0,006768	0:53.247	0,001313
						0,006768	0:49.077	0,010152
						0,006768	0:56.810	0,020101
12	100	1	100	1	Uniforme	0,006768	0:52.764	0,010051
						0,006768	0:51.299	0,002525
						0,006768	0:53.202	0,010051

Tabela 41: Resultado da Aplicação do GA utilizando as 1ª, 3ª e 5ª instâncias de cada classe para treinamento e as 2ª e 4ª instâncias para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,006768	0:05.950	0,010455
						0,010051	0:05.916	-
						0,010051	0:06.066	-
2	100	0,5	10	0,5	Dois Pontos	0,006768	0:05.864	0,014192
						0,006768	0:05.991	0,020505
						0,007273	0:05.993	-
3	100	0,5	10	0,5	Uniforme	0,006768	0:05.970	0,020505
						0,006768	0:05.803	0,01197
						0,006768	0:05.718	0,030152
4	100	0,5	100	0,5	Uniforme	0,006768	0:48.664	0,020202
						0,006768	0:50.373	0,020202
						0,006768	0:49.743	0,022222
5	100	0,5	1000	0,5	Uniforme	0,006768	8:12.084	0,014394
						0,006768	10:33.519	0,020303
						0,006768	7:55.914	0,020101
6	1.000	0,5	100	0,5	Uniforme	0,006768	10:05.947	0,020303
						0,006768	9:15.983	0,020303
						0,006768	9:35.184	0,020101
7	1.000	0,5	1000	0,5	Uniforme	0,006768	1:49:25.111	0,020101
						0,006768	1:50:08.012	0,020303
						0,006768	1:58:11.695	1
8	100	0,75	100	0,5	Uniforme	0,006768	0:48.251	0,020404
						0,006768	0:47.706	0,020101
						0,006768	0:51.628	0,020303
9	100	1	100	0,5	Uniforme	0,006768	0:51.520	0,020404
						0,006768	0:52.178	0,020101
						0,006768	0:53.924	0,020202
10	100	0,5	100	0,75	Uniforme	0,006768	0:54.301	0,020101
						0,006768	0:49.819	0,020303
						0,006768	0:53.045	0,020101
11	100	0,5	100	1	Uniforme	0,006768	0:47.242	0,030556
						0,006768	0:50.306	0,020202
						0,006768	0:49.624	0,020303
12	100	1	100	1	Uniforme	0,006768	0:54.202	0,020404
						0,006768	0:57.568	0,020606
						0,006768	0:51.561	0,020303

Tabela 42: Resultado da Aplicação do GA utilizando as 1ª, 4ª e 5ª instâncias de cada classe para treinamento e as 2ª e 3ª instâncias para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,010051	0:05.913	-
						0,006768	0:05.709	0,010657
						0,010051	0:05.904	-
2	100	0,5	10	0,5	Dois Pontos	0,006768	0:05.767	0,010051
						0,008485	0:06.042	-
						0,007273	0:05.877	-
3	100	0,5	10	0,5	Uniforme	0,006768	0:05.902	0,010758
						0,006768	0:05.944	0,011162
						0,007172	0:05.859	-
4	100	0,5	100	0,5	Uniforme	0,006768	0:50.462	0,010152
						0,006768	0:56.385	0,010152
						0,006768	0:50.117	0,010051
5	100	0,5	1000	0,5	Uniforme	0,006768	8:19.107	0,010152
						0,006768	7:45.404	0,010556
						0,006768	8:10.854	0,010051
6	1.000	0,5	100	0,5	Uniforme	0,006768	9:09.136	0,010051
						0,006768	8:39.706	0,010152
						0,006768	8:56.806	0,010354
7	1.000	0,5	1000	0,5	Uniforme	0,006768	1:50:38.553	0,010051
						0,006768	1:46:15.025	0,010556
						0,006768	1:50:29.791	0,010051
8	100	0,75	100	0,5	Uniforme	0,006768	0:53.738	0,010253
						0,006768	0:53.254	0,010051
						0,006768	0:53.254	0,010051
9	100	1	100	0,5	Uniforme	0,006768	0:55.241	0,010051
						0,006768	0:50.419	0,010051
						0,006768	0:53.891	0,010051
10	100	0,5	100	0,75	Uniforme	0,006768	0:52.143	0,010253
						0,006768	0:57.151	0,010253
						0,006768	0:54.765	0,010051
11	100	0,5	100	1	Uniforme	0,006768	0:48.008	0,010253
						0,006768	0:58.359	0,010051
						0,006768	0:52.577	0,010152
12	100	1	100	1	Uniforme	0,006768	0:50.670	0,010051
						0,006768	0:50.646	0,010051
						0,006768	0:55.489	0,010152

Tabela 43: Resultado da Aplicação do GA utilizando as 2ª, 3ª e 4ª instâncias de cada classe para treinamento e as 1ª e 5ª instâncias para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,003384	0:05.954	-
						0,00399	0:05.988	-
						0,006768	0:06.173	-
2	100	0,5	10	0,5	Dois Pontos	0,003485	0:06.151	-
						0,004697	0:05.951	-
						0,004798	0:05.811	-
3	100	0,5	10	0,5	Uniforme	0,003384	0:05.960	-
						0,003586	0:06.028	-
						0,003384	0:05.866	-
4	100	0,5	100	0,5	Uniforme	0,003384	0:48.938	-
						0,003384	0:47.308	-
						0,003384	0:51.025	-
5	100	0,5	1000	0,5	Uniforme	0,003384	8:26.393	-
						0,003384	8:50.880	-
						0,003384	8:47.029	-
6	1.000	0,5	100	0,5	Uniforme	0,003384	9:39.609	-
						0,003384	9:17.932	-
						0,003384	9:30.643	-
7	1.000	0,5	1000	0,5	Uniforme	0,003384	1:43:03.011	-
						0,003384	1:46:59.599	-
						0,000404	1:50:20.919	0,030152
8	100	0,75	100	0,5	Uniforme	0,003384	0:51.525	-
						0,003384	0:50.868	-
						0,003384	0:50.275	-
9	100	1	100	0,5	Uniforme	0,003384	0:54.980	-
						0,003384	0:56.966	-
						0,003384	0:52.042	-
10	100	0,5	100	0,75	Uniforme	0,003384	0:49.043	-
						0,003384	0:52.073	-
						0,003384	0:59.100	-
11	100	0,5	100	1	Uniforme	0,003384	0:51.500	-
						0,003384	0:50.729	-
						0,003384	0:52.904	-
12	100	1	100	1	Uniforme	0,003384	1:01.869	-
						0,003384	0:59.415	-
						0,003384	1:04.952	-

Tabela 44: Resultado da Aplicação do GA utilizando as 2ª, 3ª e 5ª instâncias de cada classe para treinamento e as 1ª e 4ª instâncias para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,003788	0:06.243	-
						0,003384	0:05.871	-
						0,003384	0:05.669	-
2	100	0,5	10	0,5	Dois Pontos	0,003384	0:05.887	-
						0,003687	0:05.828	-
						0,003384	0:06.140	-
3	100	0,5	10	0,5	Uniforme	0,003384	0:05.823	-
						0,001515	0:05.820	-
						0,003687	0:05.892	-
4	100	0,5	100	0,5	Uniforme	0,000404	0:50.302	-
						0,000303	0:47.526	-
						0,000202	0:49.340	-
5	100	0,5	1000	0,5	Uniforme	0,000101	8:02.555	0,010051
						0,000101	8:19.234	0,020101
						0,000101	8:01.058	0,014697
6	1.000	0,5	100	0,5	Uniforme	0,000101	8:21.143	0,010051
						0,000101	8:27.389	0,012677
						0,000101	8:36.750	0,010051
7	1.000	0,5	1000	0,5	Uniforme	0,000101	1:37:20.480	0,014192
						0,000101	1:30:45.350	0,010051
						0,000101	1:37:15.533	0,010152
8	100	0,75	100	0,5	Uniforme	0,000101	0:48.481	0,010253
						0,000202	0:50.325	-
						0,000101	0:50.464	0,011869
9	100	1	100	0,5	Uniforme	0,000101	0:54.244	0,020606
						0,000101	0:54.639	0,020404
						0,000202	0:53.514	-
10	100	0,5	100	0,75	Uniforme	0,000202	0:48.822	-
						0,000101	0:51.257	0,011061
						0,000101	0:49.263	0,013485
11	100	0,5	100	1	Uniforme	0,000101	0:49.307	0,020101
						0,000303	0:47.494	-
						0,000202	0:50.461	-
12	100	1	100	1	Uniforme	0,000202	0:50.067	-
						0,000101	0:47.682	0,013485
						0,000101	0:55.005	0,010051

Tabela 45: Resultado da Aplicação do GA utilizando as 2ª, 4ª e 5ª instâncias de cada classe para treinamento e as 1ª e 3ª instâncias para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,006414	0:06.126	-
						0,003687	0:05.789	-
						0,006768	0:05.975	-
2	100	0,5	10	0,5	Dois Pontos	0,003384	0:05.907	-
						0,003384	0:05.853	-
						0,005808	0:05.733	-
3	100	0,5	10	0,5	Uniforme	0,003384	0:06.002	-
						0,003384	0:05.788	-
						0,003384	0:05.988	-
4	100	0,5	100	0,5	Uniforme	0,003384	0:51.282	-
						0,003384	0:47.919	-
						0,003384	0:51.213	-
5	100	0,5	1000	0,5	Uniforme	0,003384	8:49.870	-
						0,003384	8:42.423	-
						0,003384	8:33.971	-
6	1.000	0,5	100	0,5	Uniforme	0,003384	8:54.140	-
						0,00101	8:38.882	0,011465
						0,003384	8:34.783	-
7	1.000	0,5	1000	0,5	Uniforme	0,003384	1:45:54.430	-
						0,003384	1:58:53.529	-
						0,003384	1:39:24.470	-
8	100	0,75	100	0,5	Uniforme	0,003384	0:49.280	-
						0,003384	0:53.124	-
						0,003384	0:48.682	-
9	100	1	100	0,5	Uniforme	0,003384	0:50.658	-
						0,003384	0:48.326	-
						0,003384	0:49.943	-
10	100	0,5	100	0,75	Uniforme	0,003384	0:45.718	-
						0,003384	0:51.598	-
						0,003384	0:48.009	-
11	100	0,5	100	1	Uniforme	0,003384	0:50.558	-
						0,003384	0:48.858	-
						0,003384	0:49.605	-
12	100	1	100	1	Uniforme	0,003384	0:50.412	-
						0,003384	0:52.094	-
						0,003384	0:51.678	-

Tabela 46: Resultado da Aplicação do GA utilizando as 3ª, 4ª e 5ª instâncias de cada classe para treinamento e as 1ª e 2ª instâncias para teste.

Comb.	Nº de Crom.	Prob. de Cross.	Nº de Ger.	Prob. de Mut.	Tipo de Cross.	EER (Trein.)	Tempo	EER (Teste)
1	100	0,5	10	0,5	Um Ponto	0,001212	0:08.278	-
						0,000808	0:05.808	-
						0,001313	0:05.810	-
2	100	0,5	10	0,5	Dois Pontos	0,001616	0:06.326	-
						0,000505	0:05.825	-
						0,003384	0:05.625	-
3	100	0,5	10	0,5	Uniforme	0,000707	0:06.167	-
						0,000505	0:05.983	-
						0,000404	0:05.960	-
4	100	0,5	100	0,5	Uniforme	0,000202	0:52.056	-
						0,000101	0:46.212	0,020606
						0,000101	0:45.877	0,020606
5	100	0,5	1000	0,5	Uniforme	0,003384	7:58.194	-
						0,000101	8:07.513	0,023333
						0,000101	7:55.962	0,020505
6	1.000	0,5	100	0,5	Uniforme	0,000101	8:48.346	0,030253
						0,000101	8:31.707	0,020606
						0,000101	8:42.853	0,020101
7	1.000	0,5	1000	0,5	Uniforme	0,000101	1:27:31.110	0,020202
						0,000101	1:23:32.289	0,020202
						0,000101	1:25:01.214	0,020101
8	100	0,75	100	0,5	Uniforme	0,000101	0:49.911	0,020202
						0,000101	0:48.084	0,020303
						0,000101	0:46.692	0,020101
9	100	1	100	0,5	Uniforme	0,000101	0:45.623	0,024848
						0,000101	0:46.077	0,020202
						0,000101	0:48.623	0,020202
10	100	0,5	100	0,75	Uniforme	0,000101	0:47.512	0,020404
						0,000101	0:47.537	0,020303
						0,000101	0:46.821	0,020101
11	100	0,5	100	1	Uniforme	0,000101	0:50.736	0,020101
						0,000101	0:48.981	0,030354
						0,000101	0:47.573	0,020404
12	100	1	100	1	Uniforme	0,000101	0:48.239	0,020101
						0,000101	0:49.319	0,020202
						0,000101	0:46.694	0,020505

Tabela 47: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (1ª, 2ª e 3ª instâncias) e no banco de testes (4ª e 5ª instâncias)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	1a, 2a, e 3a instâncias	1,75253%	80,692%
Sim	Treinamento	1a, 2a, e 3a instâncias	0,33838%	
Não	Teste	4a e 5a instâncias	0,01010%	-300,000%
Sim	Teste	4a e 5a instâncias	0,04040%	

Tabela 48: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (1ª, 2ª e 4ª instâncias) e no banco de testes (3ª e 5ª instâncias)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	1a, 2a, e 4a instâncias	1,88384%	64,075%
Sim	Treinamento	1a, 2a, e 4a instâncias	0,67677%	
Não	Teste	3a e 5a instâncias	0,11111%	90,909%
Sim	Teste	3a e 5a instâncias	0,01010%	

Tabela 49: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (1ª, 2ª e 5ª instâncias) e no banco de testes (3ª e 4ª instâncias)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	1a, 2a, e 5a instâncias	1,75253%	80,692%
Sim	Treinamento	1a, 2a, e 5a instâncias	0,33838%	
Não	Teste	3a e 4a instâncias	1,01515%	83,085%
Sim	Teste	3a e 4a instâncias	0,17172%	

Tabela 50: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (1ª, 3ª e 4ª instâncias) e no banco de testes (2ª e 5ª instâncias)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	1a, 3a, e 4a instâncias	1,37374%	50,735%
Sim	Treinamento	1a, 3a, e 4a instâncias	0,67677%	
Não	Teste	2a e 5a instâncias	0,45455%	71,111%
Sim	Teste	2a e 5a instâncias	0,13131%	

Tabela 51: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (1ª, 3ª e 5ª instâncias) e no banco de testes (2ª e 4ª instâncias)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	1a, 3a, e 5a instâncias	1,02525%	33,990%
Sim	Treinamento	1a, 3a, e 5a instâncias	0,67677%	
Não	Teste	2a e 4a instâncias	1,54040%	32,131%
Sim	Teste	2a e 4a instâncias	1,04545%	

Tabela 52: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (1ª, 4ª e 5ª instâncias) e no banco de testes (2ª e 3ª instâncias)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	1a, 4a, e 5a instâncias	1,37374%	50,735%
Sim	Treinamento	1a, 4a, e 5a instâncias	0,67677%	
Não	Teste	2a e 3a instâncias	1,02525%	1,970%
Sim	Teste	2a e 3a instâncias	1,00505%	

Tabela 53: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (2ª, 3ª e 4ª instâncias) e no banco de testes (1ª e 5ª instâncias)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	2a, 3a, e 4a instâncias	1,39394%	97,101%
Sim	Treinamento	2a, 3a, e 4a instâncias	0,04040%	
Não	Teste	1a e 5a instâncias	2,05051%	-47,044%
Sim	Teste	1a e 5a instâncias	3,01515%	

Tabela 54: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (2ª, 3ª e 5ª instâncias) e no banco de testes (1ª e 4ª instâncias)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	2a, 3a, e 5a instâncias	0,67677%	98,507%
Sim	Treinamento	2a, 3a, e 5a instâncias	0,01010%	
Não	Teste	1a e 4a instâncias	1,30808%	23,166%
Sim	Teste	1a e 4a instâncias	1,00505%	

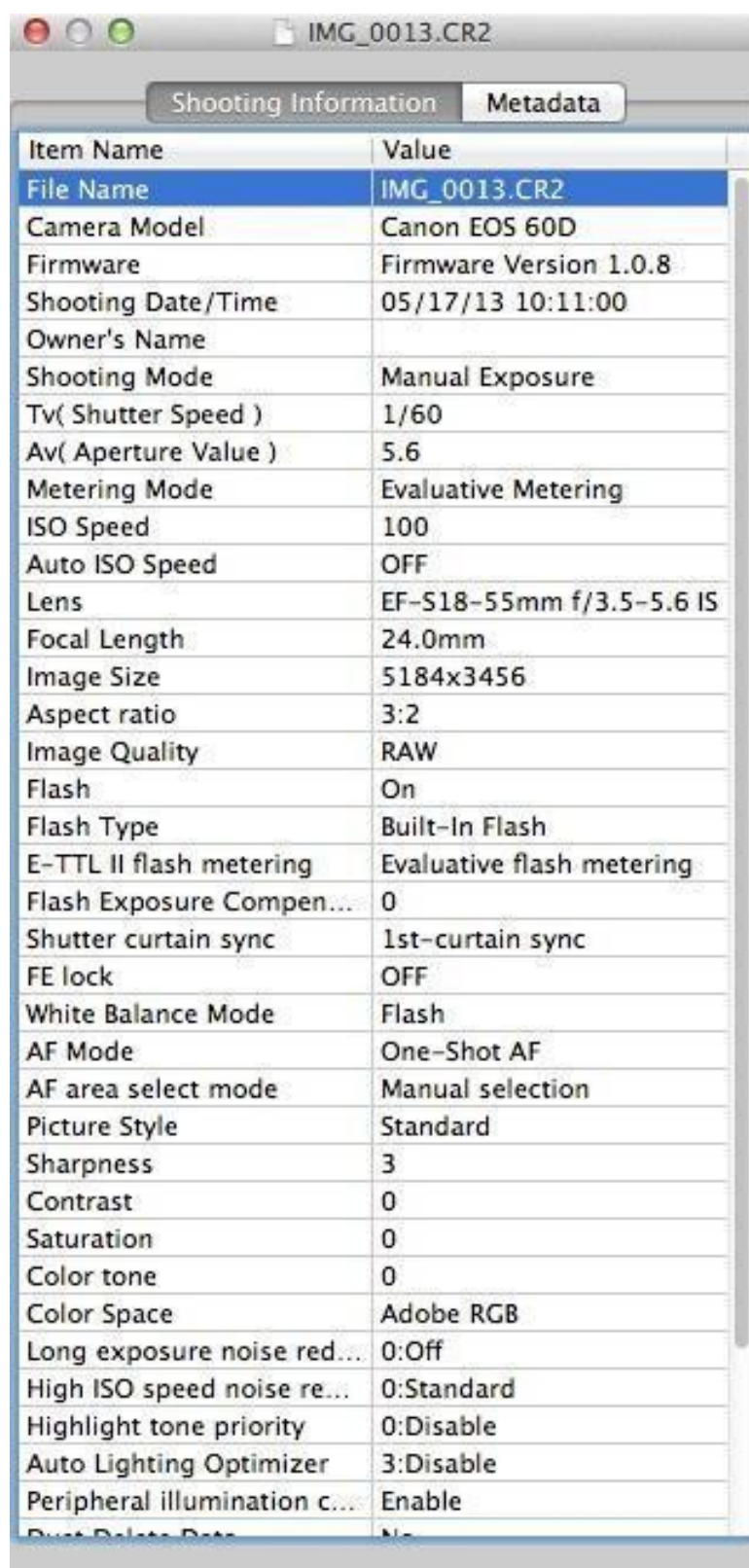
Tabela 55: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (2ª, 4ª e 5ª instâncias) e no banco de testes (1ª e 3ª instâncias)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	2a, 4a, e 5a instâncias	0,78788%	87,179%
Sim	Treinamento	2a, 4a, e 5a instâncias	0,10101%	
Não	Teste	1a e 3a instâncias	1,02525%	-11,823%
Sim	Teste	1a e 3a instâncias	1,14646%	

Tabela 56: Resultado do ganho de desempenho obtido pelo GA no banco de treinamento (3ª, 4ª e 5ª instâncias) e no banco de testes (1ª e 2ª instâncias)

Algoritmo Genético?	Tipo de Banco	Classes	EER	Ganho de Performance
Não	Treinamento	3a, 4a, e 5a instâncias	0,33838%	97,015%
Sim	Treinamento	3a, 4a, e 5a instâncias	0,01010%	
Não	Teste	1a e 2a instâncias	2,05051%	1,970%
Sim	Teste	1a e 2a instâncias	2,01010%	

ANEXO A: PARÂMETROS DA CÂMERA DE CAPTURA



Item Name	Value
File Name	IMG_0013.CR2
Camera Model	Canon EOS 60D
Firmware	Firmware Version 1.0.8
Shooting Date/Time	05/17/13 10:11:00
Owner's Name	
Shooting Mode	Manual Exposure
Tv(Shutter Speed)	1/60
Av(Aperture Value)	5.6
Metering Mode	Evaluative Metering
ISO Speed	100
Auto ISO Speed	OFF
Lens	EF-S18-55mm f/3.5-5.6 IS
Focal Length	24.0mm
Image Size	5184x3456
Aspect ratio	3:2
Image Quality	RAW
Flash	On
Flash Type	Built-In Flash
E-TTL II flash metering	Evaluative flash metering
Flash Exposure Compens...	0
Shutter curtain sync	1st-curtain sync
FE lock	OFF
White Balance Mode	Flash
AF Mode	One-Shot AF
AF area select mode	Manual selection
Picture Style	Standard
Sharpness	3
Contrast	0
Saturation	0
Color tone	0
Color Space	Adobe RGB
Long exposure noise red...	0:Off
High ISO speed noise re...	0:Standard
Highlight tone priority	0:Disable
Auto Lighting Optimizer	3:Disable
Peripheral illumination c...	Enable
Dust Delete Data	No

ANEXO B: TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Prezado(a) Senhor(a),

A pesquisa intitulada **Pesquisa e Desenvolvimento de um Sistema Multifator e Multibiométrico de Identificação e Autenticação de Usuários**, que está sendo desenvolvida por pesquisadores do Programa de Pós-Graduação em Informática, do Centro de Informática da Universidade Federal da Paraíba, tem como objetivo a construção de um banco de informações biométricas e a pesquisa e desenvolvimento de métodos inovadores de reconhecimento multibiométrico. Esta investigação contribuirá para um avanço na fronteira do conhecimento associada à identificação e autenticação por análise de traços multibiométricos, e resultará na disponibilização do primeiro serviço público multifator e multibiométrico de autenticação de usuários para transações via Internet do Brasil.

Nesta etapa da pesquisa, pretende-se construir uma base de dados contendo traços biométricos relativos à geometria da mão, à impressão palmar e à impressão digital.

Solicitamos sua colaboração no sentido de preencher o formulário anexo e permitir a obtenção e armazenamento de imagens de suas mãos e impressões digitais, bem como sua autorização para apresentar os resultados deste estudo em eventos e artigos técnico-científicos. É garantido o sigilo da sua identidade. As imagens capturadas poderão ser apresentadas em eventos ou artigos técnico-científicos, mas sua identidade será mantida em sigilo. Esta pesquisa não oferece riscos previsíveis.

Esclarecemos que sua participação no estudo é voluntária e, portanto, o(a) senhor(a) não é obrigado(a) a fornecer as informações e/ou colaborar com as atividades solicitadas. Caso decida não participar do estudo, ou resolver a qualquer momento desistir do mesmo, não sofrerá nenhum dano. Os pesquisadores estarão a sua disposição para qualquer esclarecimento que considere necessário em qualquer etapa da pesquisa.

Diante do exposto, declaro que fui devidamente esclarecido (a) e dou o meu consentimento para participar da pesquisa e para publicação dos resultados. Estou ciente que receberei uma cópia deste documento.

Assinatura do Participante da Pesquisa

Assinatura do Pesquisador Responsável

ANEXO C: INSTRUMENTO

Número de identificação do participante: _____

Data de Nascimento: _____

Sexo: () Masculino () Feminino

Nacionalidade: _____

Naturalidade: _____

Ocupação principal: _____

ANEXO D: CERTIDÃO DE APROVAÇÃO DE PROJETO DE PESQUISA



UNIVERSIDADE FEDERAL DA PARAIBA - UFPB
HOSPITAL UNIVERSITÁRIO LAURO WANDERLEY - HULW
**COMITÊ DE ÉTICA EM PESQUISA EM SERES
HUMANOS - CEP**

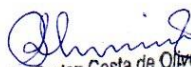
Comitê de Ética em Pesquisa
Hospital Universitário Lauro Wanderley
Universidade Federal da Paraíba

CERTIDÃO

Com base na Resolução nº 196/96 do CNS/MS que regulamenta a ética da pesquisa em seres humanos, o Comitê de Ética em Pesquisa do Hospital Universitário Lauro Wanderley - CEP/HULW, da Universidade Federal da Paraíba, em sua sessão realizada no dia 25/02/2013, após análise do parecer do relator, considerou **APROVADO** o projeto de pesquisa intitulado **PESQUISA E DESENVOLVIMENTO DE UM SISTEMA MULTIFATOR E MULTIBIOMÉTRICO DE IDENTIFICAÇÃO E AUTENTICAÇÃO DE USUÁRIOS**, do pesquisador responsável LEONARDO VIDAL BATISTA. Certificado de Apresentação para Apreciação Ética - CAAE nº: **13007513.0.0000.5183**.

O pesquisador responsável fica, desde já, notificado da obrigatoriedade de no término da pesquisa enviar ao CEP/HULW/Plataforma Brasil o relatório final da pesquisa (*online*).

João Pessoa, 27 de fevereiro de 2013.


Iaponira Cortez Costa de Oliveira
Coordenadora do Comitê de Ética em Pesquisa - CEP/HULW

Profª Drª Iaponira Cortez Costa de Oliveira
Coordenadora do Comitê de Ética em Pesquisa-HULW

Endereço: Hospital Universitário Lauro Wanderley-HULW - 4º andar, Campus I - Cidade Universitária.
Bairro: Castelo Branco - João Pessoa - PB. CEP: 58051-900 CNPJ: 24098477/007-05
Fone: (83) 32167964 — Fone/fax: (083)32167522 E-mail - comitedeetica@hulw.ufpb.br