



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

**Influência das Razões Foto Antropométricas no
processo de reconhecimento facial biométrico em
norma frontal em imagens digitais**

Gustavo Carneiro Bicalho

Dissertação apresentada como requisito parcial para
conclusão do Mestrado em Informática

Orientador
Prof. Dr. Flávio de Barros Vidal

Brasília
2018



Universidade de Brasília

Instituto de Ciências Exatas
Departamento de Ciência da Computação

**Influência das Razões Foto Antropométricas no
processo de reconhecimento facial biométrico em
norma frontal em imagens digitais**

Gustavo Carneiro Bicalho

Dissertação apresentada como requisito parcial para
conclusão do Mestrado em Informática

Prof. Dr. Flávio de Barros Vidal (Orientador)
CIC/UnB

Prof. Dr. Eduardo Peixoto Fernandes da Silva Prof. Dr. Bruno L. Macchiavello Espinoza
ENE/UnB CIC/UnB

Prof. Dr. Camilo Chang Dorea (Suplente)
CIC/UnB

Prof. Dr. Bruno Luigi Macchiavello Espinoza
Coordenador do Programa de Pós-graduação em Informática

Brasília, 31 de Julho de 2018

Dedicatória

Dedico este trabalho para a minha família, pais, avós e irmãos, pelo apoio nos momentos de dificuldade. Dedico também aos meus amigos e colegas que sempre se colocaram a disposição para me auxiliar sempre que necessário.

Agradecimentos

Agradeço ao Professor Flávio Vidal pela paciência e conselhos como orientador. Agradeço também aos colegas de projeto Marcelo Cobias e Lucas Porto pelo auxílio na pesquisa e nas matérias do curso. Agradeço à CAPES pela bolsa de estudos que possibilitou que essa pesquisa fosse concluída.

Resumo

Reconhecimento facial é uma área de grande interesse para governos e empresas nos dias de hoje, motivando diversas pesquisas na área. Atualmente, a maioria das pesquisas migrou para a utilização de aprendizado de máquina para apresentar soluções para o reconhecimento facial biométrico. As técnicas utilizando marcos faciais eram a principal abordagem para resolver reconhecimentos faciais biométricos e, mesmo com o avanço de novas técnicas, elas ainda são capazes de obter bons resultados em ambientes controlados. Porém, ainda existem problemas em aberto para serem solucionados, como o tratamento de gêmeos, variação de escala e o crescimento da face. Neste trabalho, inicialmente é feita uma pesquisa sobre os principais algoritmos de reconhecimento facial biométrico e, além disso, é proposto um novo método baseado em valores medidos (razões) de marcos faciais cefalométricos, que utilizam o tamanho da íris como um fator de normalização para solucionar a influência dos efeitos escala da face (crescimento facial) e melhorando os valores de *Equal Error Rate* (EER) para um sistema de reconhecimento facial em cenários específicos em 5%. Também foi feita uma análise para reduzir o número de razões necessárias, reduzindo das iniciais 40 razões para 14, enquanto também aumentando o desempenho do método.

Palavras-chave: Análise foto antropométrica, Visão Computacional, Reconhecimento Facial Biométrico

Abstract

Nowadays, facial recognition arouses interest for governments and companies, motivating various reserches in the area. Presently, most researches have used techniques based on machine learning. Facial landmarks techniques were the first and main approach to solve biometric facial recognition and, even with the rise of newer techniques, they are still capable of achieving great results in controlled environments. However, there are still open problems to be solved, such as how to deal with twins, scale variation and the face growth. In this work, initially a research is made over the most important algorithms for biometric facial recognition and also, we propose a new method based on measured values (ratios) from facial cephalometric landmarks, which uses an iris size as a normalization factor to solve the influence of face scale (face growth) effect and improving Equal Error Rates (EER) scores for a facial recognition system in specifics scenarios under 5%. An analysis to reduce the number of ratios needed was also made, reducing from 40 to 14, while also increasing the performance of the method.

Keywords: Photocephalometric Analysis, Computer Vision, Biometric Facial Recognition

Sumário

1	Introdução	1
1.1	Motivação	2
1.2	Objetivos	2
1.2.1	Objetivo Principal	2
1.2.2	Objetivos Específicos	3
1.3	Organização	3
2	Revisão Bibliográfica	4
2.1	Reconhecimento Biométrico	4
2.1.1	Sistema Biométrico	4
2.1.2	Característica Biométrica	7
2.1.3	Medida de desempenho de sistemas biométricos	8
2.2	Reconhecimento Facial	9
2.2.1	Contextualização	9
2.2.2	Conceitos Básicos	10
2.2.3	Abordagens para o Reconhecimento Facial	11
2.3	Pontos Cefalométricos	15
3	Trabalhos Relacionados	22
3.1	Marcos Faciais	22
3.2	Reconhecimento Facial	24
3.2.1	Reconhecimento Facial por Aprendizado de Máquina	24
3.2.2	Reconhecimento Facial Convencional	28
4	Metodologia	30
4.1	Descrição do Método	30
4.2	Descrição da Base de Dados	31
4.3	Deteção dos pontos	33
4.4	Reconhecimento Facial	34
4.5	Redução do Número de RFAs	37

5 Resultados	39
5.1 Detecção automática dos pontos cefalométricos	40
5.2 Reconhecimento Facial	40
5.2.1 Primeiro cenário	42
5.2.2 Segundo cenário	42
5.2.3 Terceiro cenário	44
5.2.4 Quarto cenário	45
5.3 Discussão	45
6 Conclusão	53
6.1 Publicação Realizada	54
Referências	55

Lista de Figuras

2.1	Diagramas de blocos de inserção, verificação e identificação utilizando os quatro principais módulos de um sistema biométrico: sensor, extrator de características, <i>matcher</i> e a base de dados.	5
2.2	Características biométricas utilizadas. As fisiológicas dependem principalmente da carga genética e as comportamentais dependem do aprendizado e da experiência. Imagem adaptada de	8
2.3	As curvas das taxas de erro FAR e FRR colocadas uma ao lado da outra, em relação ao limiar T configurado pelo sistema. Também é mostrado a relação das taxas com a conveniência e a segurança do sistema biométrico. Além disso, é possível observar o ponto de $EER(T_E)$, onde o $FAR(T_E)=FRR(T_E)$.	9
2.4	pipeline Detection-Alignment-Recognition(DAR)	12
2.5	16 parâmetros faciais utilizados por Kanade para a realização do reconhecimento facial.	13
2.6	35 pontos marcados manualmente para a realização de reconhecimento. . .	15
2.7	Diagrama descrevendo o processo de reconhecimento facial	16
2.8	As quatro razões utilizadas para a realização do reconhecimento	17
2.9	Crescimento de cada região do corpo de acordo com a idade.	19
2.10	28 pontos cefalométricos utilizados visualizados na face e suas abreviações: 1. Glabella (g); 2. Nasion (n); 3. Subnasale (sn); 4. Labiale superius (ls); 5. Stomion (sto); 6. Labiale inferius (li); 7. Gnathion (gn); 8. Midnasal (mid); 9. Entocanthion (en); 10. Exocanthion (ex); 11. Iridion laterale (il); 12. Iridion mediale (im); 13. Pupil (pu); 14. Zygion (zy); 15. Alare (al); 16. Gonion (go); 17. Cheilion (ch); 18. Crista philtri (cph).	20
3.1	Imagem dos pontos marcados na base 300-W, 68 caso inclua os pontos do limite da face do rosto ou 51 caso contrário.	23
3.2	<i>Triplet Loss</i> minimiza a distância entre a âncora e as imagens positivas, e maximiza a distância entre a âncora e as imagens negativas.	26
3.3	Imagem de baixíssima resolução utilizada para reconhecimento facial. . . .	27

3.4	Três subimagens utilizadas para realizar o reconhecimento facial.	29
4.1	Imagem da base de dados utilizada, estando de acordo com as normas previstas pela ICAO.	32
4.2	Exemplo de imagens com variação de escala.	32
4.3	Imagem que ilustra as imagens utilizadas na terceira base dados. O sistema deve lidar com o crescimento do indivíduo durante 5 anos.	33
4.4	4 sub-regiões e os pontos cefalométricos de cada uma delas, a numeração é referente à Tabela 2.1. (a) Região dos olhos; (b) Região do nariz; (c) Região da boca; e (d) Rosto inteiro.	35
4.5	O círculo vermelho é a marcação do especialista, e também é o centro do quadrado preto que determina o limite (definido em píxeis) de onde é considerado um acerto na avaliação da marcação automática. (a) É considerado um acerto quando o círculo verde está dentro do limite estipulado. É considerado um erro se o círculo verde estiver fora do limite estipulado.	36
5.1	Gráfico com os F-scores, variando o pixel da janela de tolerância de 0 a 20.	40
5.2	Gráfico com o F-score do método proposto, utilizando 5 píxeis de tolerância, comparado com os resultados de outras ferramentas reconhecidas na literatura.	41
5.3	Resultados para o primeiro cenário com 20 RFAs.	42
5.4	Resultados para o primeiro cenário com 30 RFAs.	43
5.5	Resultados para o primeiro cenário com 40 RFAs.	43
5.6	Resultados para o segundo cenário com 20 RFAs.	44
5.7	Resultados para o segundo cenário com 30 RFAs.	44
5.8	Resultados para o segundo cenário com 40 RFAs.	45
5.9	Resultados para o terceiro cenário com 15 RFAs.	46
5.10	Resultados para o terceiro cenário com 40 RFAs.	46
5.11	Resultados para o quarto cenário com 14 RFAs.	47
5.12	Resultados para o quarto cenário com 15 RFAs.	48
5.13	Resultados para o quarto cenário com 20 RFAs.	48
5.14	Resultados para o quarto cenário com 30 RFAs.	49
5.15	Gráfico ROC mostrando a comparação entre a metodologia de escolha de RFAs e a escolha por correlação. A imagem mostra as curvas para testes com 15, 20 e 30 RFAs.	51

Lista de Tabelas

2.1	Tabela com os 28 pontos utilizados e suas descrições	18
2.2	Tabela com as 40 razões foto antropométricas, com base nos pontos cefalométricos descritos na tabela 2.1 e visualizados na Figura 2.10. Todas as razões devem ser divididas pela projeção iridiana da equação 2.10	21
5.1	Ordem das RFAs gerada a partir da correlação.	41
5.2	Ordem das RFAs gerada a partir do método proposto na Seção 4.5, com as médias obtidas para cada uma delas.	47
5.3	Área sob a curva para o gráfico da Figura 5.15.	52

Lista de Acrônimos

BPNN	Back-Propagation Neural Network
CLT	Transformadas de Curvelet
CNT	Transformadas de Contourlet
DAR	Detection-Alignment-Recognition
DCT	Transformada discreta dos cossenos
EE	Equal Error
EER	Equal Error Rate
FA	False Accept
FAR	False Accept Rate
FFNN	Feed Forward Neural Network
FLD	Fisher's Linear Discriminant
FM	False Match
FMR	False Match Rate
FNM	False Non-Match
FNMR	False Non-Match Rate
FR	False Reject
FRR	False Reject Rate
GLCM	gray-level co-occurrence matrix
ICAO	Organização da Aviação Civil Internacional

LBPH Histograma de Padrão Binário Local

OpenCV Open Computer Vision

PCA Principal Component Analysis

RFA Razão Foto-Antropométrica

ROC Característica de Operação do Receptor

RRNN Recurrent Regression Neural Network

SVM Support Vector Machine

Capítulo 1

Introdução

Atualmente é comum encontrar aplicações que utilizem reconhecimento facial biométrico, seja nas marcações de fotos no *Facebook* ou ao passarmos pela alfândega para entrar no Brasil[1]. A precisão e acurácia de cada uma dessas ferramentas são de extrema importância. Com isso em mente, apesar de termos ferramentas que atingem mais de 90% de acerto, o reconhecimento facial ainda encontra diversos problemas, como por exemplo: envelhecimento, alteração da face por cirurgias e má qualidade de imagens. A resolução desses problemas possibilitará o crescimento da precisão dos sistemas de reconhecimento facial[2].

Porém, para se compreender o reconhecimento facial, é necessário entender os conceitos de biometria e de sistema biométrico. Em [3], é definido como biometria os traços das pessoas que podem ser medidos, sendo possível obter uma autenticação computando esses traços e gerando um identificador único, difícil de compartilhar, forjar, roubar e de ser alterado. Já um sistema biométrico é um sistema utilizado para autenticação de pessoas, podendo ser uma verificação, quando se deseja verificar uma identidade alegada de um indivíduo, ou uma identificação, quando se deseja buscar a identidade de um indivíduo, de indivíduos, baseado nas características biométricas[4].

O problema de reconhecimento facial, segundo [5], pode ser formulado da seguinte forma: dada uma imagem ou um vídeo de uma cena, identificar ou verificar uma ou mais pessoas nesta cena, utilizando uma base de dados de faces. As soluções mais recentes para a solução deste problema, como o *DeepFace* desenvolvido pelo *Facebook*[6] ou o *FaceNet* desenvolvido pelo *Google*[7], têm utilizado redes neurais treinadas para aumentar a precisão do reconhecimento facial. Essas abordagens têm como requisito um número muito grande de imagens e várias imagens de um mesmo indivíduo para que o treinamento possa ser realizado e obter uma acurácia alta.

Outra abordagem para realização de reconhecimento facial, e também a mais antiga descrita em [8], é baseada em marcos faciais. Nesta abordagem, um conjunto de pontos é

extraído da face e utilizado para obter relações geométricas desta face. Por último essas relações são utilizadas para realizar uma comparação entre as faces a serem identificadas. O conjunto de pontos a ser utilizado pode ser determinante para o sucesso do reconhecimento facial. A vantagem desse método é que não são necessárias diversas imagens de um mesmo indivíduo para realizar o reconhecimento.

Com isso surge a ideia da utilização de um subconjunto de pontos destes marcos faciais utilizados por peritos e papiloscopistas dos serviços de segurança pública e investigação [9] e verificar se com este subconjunto é possível aumentar a acurácia de reconhecimento facial baseados em marcos (ou pontos) faciais (do inglês *landmarks*) já disponíveis atualmente em cenários onde ainda não existem soluções definitivas.

1.1 Motivação

Atualmente, governos e indústrias investem em pesquisas buscando obter níveis humanos no reconhecimento facial biométrico[10]. Com isso, diversos aspectos do reconhecimento devem ser aprimorados para que seja possível aumentar a acurácia sem que o tempo de execução seja extrapolado[2].

Segundo [9], a antropometria é considerada uma das formas de coleta de informações utilizados pela antropologia forense, enquanto que as relações faciais, tanto por índices como por razões, são consideradas ferramentas de análise. Assim, a cefalometria é uma técnica antropométrica que estuda as dimensões da cabeça, com base em pontos específicos, chamados de pontos cefalométricos. Também em [9] foi proposto que pontos cefalométricos podem ser utilizados para estimar o sexo e a idade de pessoas através de 40 razões foto antropométricas. Com essas razões, é possível obter diversas informações sobre o indivíduo como sexo e idade, e ainda inclui a possibilidade de realizar o reconhecimento facial a partir delas.

A utilização dessas 40 razões pode permitir a possibilidade de realizar reconhecimento facial mesmo com uma diferença de idade entre as imagens do indivíduo, possibilitando assim facilitar o reconhecimento em um dos pontos de dificuldade dos sistemas de reconhecimento convencionais[11].

1.2 Objetivos

1.2.1 Objetivo Principal

O principal objetivo deste trabalho é verificar a influência do uso das razões foto antropométricas propostas no trabalho desenvolvido em [9] no reconhecimento facial, verificando

se essas 40 razões podem ser utilizadas para reconhecimento facial e a sua capacidade para realizar reconhecimento facial biométrico mesmo em imagens com escalas diferentes e para imagens com diferença de idade de um mesmo indivíduo. Esta abordagem permitirá a comparação às demais abordagens da literatura atual de reconhecimento facial baseados em marcos faciais (*landmarks*), verificando características de custo computacional, acurácia e viabilidade do seu uso em larga escala em imagens faciais.

1.2.2 Objetivos Específicos

Para que o objetivo principal seja alcançado, será necessário o desenvolvimento de técnicas que combinem processamento de imagens, visão computacional e aprendizado de máquina para a detecção dos pontos cefalométricos, em norma frontal, com precisão para realização da comparação entre eles.

Inicialmente serão obtidos os pontos cefalométricos automaticamente. Em seguida, será possível calcular as razões foto antropométricas proposta em [9] para imagens em norma frontal. Após calculadas, será desenvolvido um método para comparação das razões com as previamente calculadas e armazenadas em uma base de dados, gerando um score que será utilizado para realização do reconhecimento facial. Por último será verificado o desempenho do método proposto em bases com mais indivíduos e em situações adversas. Para isso, serão utilizadas bases de dados próprias, que permitam avaliar a capacidade do método de realizar reconhecimento facial mesmo com o envelhecimento da pessoa, ou em imagens com escalas diferentes.

1.3 Organização

Esta dissertação de mestrado está dividida da seguinte forma: No Capítulo 2 é feita uma revisão bibliográfica abordando os principais conceitos sobre reconhecimento facial e suas aplicações, além disso também serão tratados os conceitos relevantes sobre foto antropometria, principalmente os pontos cefalométricos. No Capítulo 3 serão trazidos trabalhos relacionados recentes nas áreas de detecção automática de marcos faciais e de reconhecimento facial, tanto utilizando machine learning como utilizando técnicas convencionais de reconhecimento. O Capítulo 4 irá abordar a metodologia proposta no projeto, abordando os pontos mais relevantes para que o objetivo seja alcançado. No Capítulo 5, será apresentado os testes realizados para validar a metodologia. E, por fim, no capítulo 6 os resultados serão analisados e então serão propostos trabalhos futuros.

Capítulo 2

Revisão Bibliográfica

Neste capítulo é apresentada uma revisão bibliográfica sobre biometria, reconhecimento facial e sobre os pontos cefalométricos. Na primeira seção serão abordados os principais conceitos sobre biometria e sistemas biométricos. Na seção de reconhecimento facial, serão abordados os principais métodos e o que as abordagens mais recentes trazem como novidades. Na seção seguinte, sobre os pontos cefalométricos, serão abordados o que são esses pontos e quais os pontos e razões foto antropométricas serão utilizadas neste trabalho.

2.1 Reconhecimento Biométrico

2.1.1 Sistema Biométrico

Segundo [12], um sistema biométrico é essencialmente um sistema de reconhecimento de padrões que opera adquirindo dados biométricos de um indivíduo, extraindo um conjunto de características dos dados adquiridos e, por fim compara este conjunto com os já constantes na base de dados. Dependendo do contexto, um sistema biométrico pode operar tanto em modo de verificação como em modo de identificação.

De acordo com [12], no modo de verificação, o sistema deve validar a identidade de uma pessoa comparando os dados de biometria capturados com os dados da própria pessoa, previamente capturados e armazenados em uma base de dados. Neste sistema, um indivíduo que deseja ser reconhecido alega uma identidade utilizando algo que apenas uma pessoa possa possuir (como senhas, tokens, chaves), e o sistema realiza uma comparação um para um para determinar se a alegação é verdadeira ou não. Este método é utilizado para reconhecimento positivo, onde o foco é evitar que diversas pessoas utilizem a mesma identidade.

No modo de identificação, o sistema reconhece um indivíduo procurando por todos da base de dados, procurando por um similar. Assim, o sistema realiza uma comparação um para muitos para determinar a identidade de um indivíduo(ou falha caso o sujeito não tenha sido inserido na base de dados) sem que o sujeito tenha que alegar qualquer identidade. Este método é um componente crítico do reconhecimento negativo, onde o sistema deve determinar se a pessoa é quem ela nega ser. O propósito do reconhecimento negativo é impedir a pessoa de utilizar várias identidades.

A Figura 2.1 mostra o diagrama de blocos da verificação e identificação, além da inscrição de indivíduos na base de dados, que é comum aos dois processos. Para o autor de [12], o termo reconhecimento pode ser utilizado tanto para verificação como para identificação.

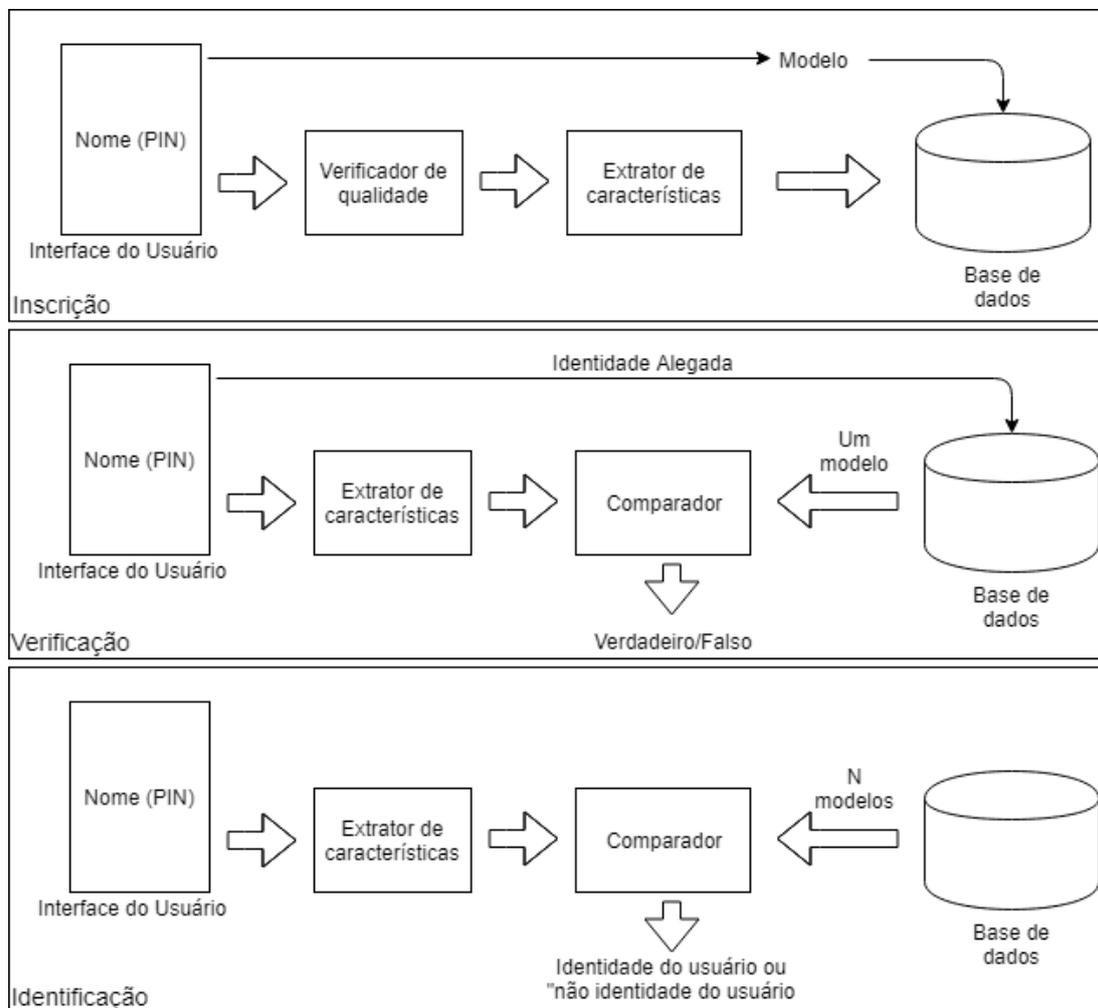


Figura 2.1: Diagramas de blocos de inscrição, verificação e identificação utilizando os quatro principais módulos de um sistema biométrico: sensor, extrator de características, *matcher* e a base de dados [12].

O problema de verificação pode ser definido formalmente como: dado como entrada

um vetor de características X_Q (extraído dos dados biométricos) e uma identidade alegada I , determinar se (I, X_Q) pertence a classe w_1 ou w_2 , onde w_1 indica que a identidade alegada é verdadeira e w_2 que é falsa. X_Q é comparado com X_I , sendo este o vetor de características correspondente ao indivíduo I , para determinar a sua classe. Assim

$$(I, X_Q) \in \begin{cases} w_1, & \text{se } S(X_Q, X_I) \geq t \\ w_2, & \text{caso contrário} \end{cases} \quad (2.1)$$

onde S é a função que mede a similaridade entre os vetores X_Q e X_I , e t é o limiar predefinido. O valor de $S(X_Q, X_I)$ é chamado de similaridade ou score de comparação entre as medidas biométricas do indivíduo e da identidade alegada. Assim, toda identidade alegada é classificada em w_1 ou w_2 com base nas variáveis X_Q , I , X_I , e t e a função S . O limiar t é necessário pois as medidas biométricas de um mesmo indivíduo dificilmente são iguais a cada medição.

Já o problema de identificação pode ser definido como: dado como entrada um vetor de características X_Q , determinar a identidade I_k , $k \in \{1, 2, \dots, N, N + 1\}$. I_1, I_2, \dots, I_N são as identidades inscritas no sistema e I_{N+1} indica o caso rejeitado, onde nenhuma identidade é compatível com a do usuário. Assim

$$X_Q \in \begin{cases} I_k, & \text{se } \max_k \{S(X_Q, X_{I_k})\} \geq t, k = 1, 2, \dots, N \\ I_{N+1}, & \text{caso contrário} \end{cases} \quad (2.2)$$

onde X_{I_k} é o vetor de características correspondente a identidade I_k , e t é um limiar predefinido.

Como a Figura 2.1 mostra, um sistema biométrico é composto por quatro módulos principais, sendo eles: Módulo sensor, módulo extrator de características, módulo comparador e o módulo base de dados.

O módulo sensor é responsável por capturar os dados biométricos de um indivíduo. Um exemplo pode ser uma *webcam* capturando a imagem da face de uma pessoa. O módulo extrator de características é responsável por processar os dados biométricos permitindo a extração de um conjunto de características determinantes para a verificação ou identificação. Por exemplo a extração dos pontos cefalométricos, detalhados na seção 2.3. O módulo comparador é onde as características extraídas são comparadas com as armazenadas previamente na base de dados para gerar o score da comparação. Este módulo também inclui o módulo de tomada de decisão, onde a identidade do indivíduo é confirmada ou rejeitada (verificação) ou a identidade é determinada (identificação) baseado no score de comparação. Como exemplo podemos ter a diferença entre as 15 razões foto antropométricas, que são detalhas na Seção 2.3.

Por último temos o módulo base de dados usado pelo sistema biométrico para armazenar os dados biométricos de usuários previamente inseridos. O módulo de inscrição é responsável por inscrever indivíduos na base de dados do sistema biométrico. Durante a fase de inscrição, as características biométricas de uma pessoa são escaneadas por um leitor de biometria para produzir uma representação digital dessas características. Os dados capturados durante o processo de inscrição podem ou não ser supervisionados por humanos, dependendo da aplicação. Um controle de qualidade é utilizado para tentar garantir que os exemplos adquiridos possam ser seguramente processados pelos demais estágios. Geralmente, diversos conjuntos de uma mesma pessoa são armazenados para tratar as variações observadas durante a extração das características biométricas e o conjunto de dados pode ser atualizado com o tempo.

2.1.2 Característica Biométrica

Na prática, nenhuma característica biométrica atende a todos os requisitos de uma característica biométrica ideal com perfeição[3]. De acordo com [13] e [3], qualquer característica fisiológica ou comportamental humana pode ser usada como uma característica biométrica desde que ela satisfaça alguns requisitos básicos:

- Universalidade: toda a população (a ser autenticada) deve possuir a característica.
- Unicidade: uma característica biométrica deve ser única para cada indivíduo.
- Permanência: a característica deve ser imutável.
- Coleta: a característica tem que ser passível de mensuração por meio de um dispositivo.
- Aceitação: a coleta da característica deve ser tolerada pelo indivíduo em questão.

As tecnologias biométricas existentes são classificadas em dois grupos, as fisiológicas ou estáticas e as comportamentais ou dinâmicas, como é possível visualizar na Figura 2.2.

As características fisiológicas são originárias da carga genética do indivíduo, e apresentam pouca ou nenhuma variação ao longo do tempo. As principais características fisiológicas são a aparência facial, o padrão da íris, a geometria das mãos e as impressões digitais.

As características comportamentais são aprendidas ou desenvolvidas ao longo 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 possuir grandes alterações. As principais características dinâmicas utilizadas são o padrão de voz e a dinâmica da assinatura.

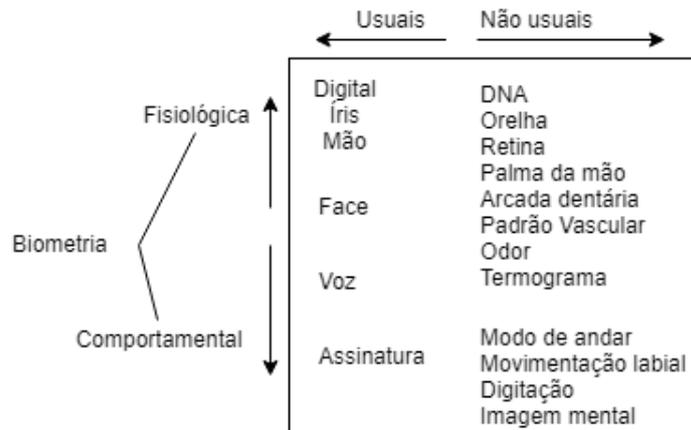


Figura 2.2: Características biométricas utilizadas. As fisiológicas dependem principalmente da carga genética e as comportamentais dependem do aprendizado e da experiência. Imagem adaptada de [3].

2.1.3 Medida de desempenho de sistemas biométricos

Existe uma diferenciação entre os diversos tipos de formas de medir o desempenho para a comunidade biométrica. Cada aplicação biométrica pode ter uma definição diferente dos erros associados. O que é bastante claro e aceito pela comunidade biométrica é que todo sistema biométrico irá cometer erros e que o verdadeiro valor associado às diversas taxas de erro não pode ser estabelecido teoricamente, por cálculo, mas somente por estimativas estatísticas dos erros, que são expressos em taxas de percentagens, utilizadas para medir o desempenho dos sistemas biométricos[14].

Em [14], são citados dois tipos de erros que o comparador pode incorrer. O primeiro tipo de erro é o *False Match*(FM) que é decidir que os exemplares são similares, enquanto na realidade eles pertencem a diferentes indivíduos. A frequência com que este erro ocorre é chamada de *False Match Rate*(FMR). O segundo tipo é o *False Non-Match*(FNM) que é decidir que dois exemplares não são do mesmo indivíduo quando na realidade pertencem ao mesmo indivíduo. A frequência com que este erro ocorre é chamada de *False Non-Match Rate*(FNMR)

Os termos FN e FNM são aplicados geralmente a algoritmos de comparação ou módulos comparadores. Na prática, é utilizada a terminologia convencional de reconhecimento de padrões *False Accept*(FA) e *False Reject*(FR). Onde *False Accept*(FA) é decidir que uma identidade alegada é legítima quando na realidade ela é falsa e *False Reject*(FR) é decidir que uma identidade alegada é falsa quando na realidade ela é legítima. A frequência com que cada um desses erros ocorre são chamadas de *False Accept Rate*(FAR) e *False Reject Rate*(FRR).

Segundo [3], devido à possibilidade de calibrar o sistema por meio do ajuste do limiar, as taxas de erros possuem consequências opostas. FA resulta em brechas na segurança,

com a admissão de usuários não autorizados. Porém FR resulta em problemas de conveniência, já que usuários que deveriam ser autorizados teriam o acesso negado até uma verificação posterior. As taxas de erro FAR e FRR podem ser colocadas uma ao lado da outra como mostra a Figura 2.3. Para a avaliação das qualidades das curvas FAR e FRR e, por consequência, a precisão de operação de um dado sistema, é possível a explicitação de um ponto notável, onde as taxas são iguais, ou seja o limiar $T=T_E$ para o qual $FAR(T)=FRR(T)$. Este ponto é chamado de ponto de operação *Equal Error*(EE), associado a uma taxa *Equal Error Rate*(EER).

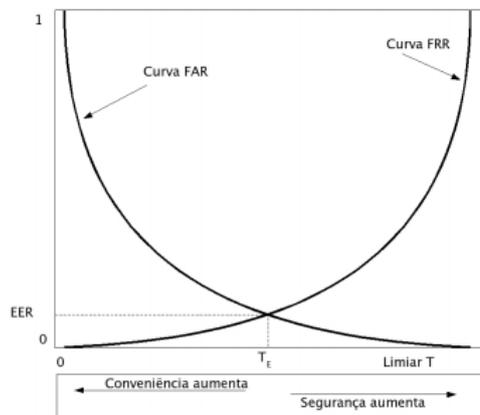


Figura 2.3: As curvas das taxas de erro FAR e FRR colocadas uma ao lado da outra, em relação ao limiar T configurado pelo sistema. Também é mostrado a relação das taxas com a conveniência e a segurança do sistema biométrico. Além disso, é possível observar o ponto de $EER(T_E)$, onde o $FAR(T_E)=FRR(T_E)$. Imagem retirada de [3].

Assim, em um sistema onde é necessário alto nível de segurança, os problemas provenientes de FRs são rotineiros, enquanto os problemas de FAs podem gerar grande prejuízo. Porém, em um sistema que possua maior necessidade de conveniência os FRs são menos aceitáveis, já os FAs podem ser tolerados e validados em uma segunda fase de autenticação.

2.2 Reconhecimento Facial

2.2.1 Contextualização

Reconhecimento facial tem sido uma área ativa de pesquisas desde os anos 70[8]. Dado uma imagem de entrada com várias faces, sistemas de reconhecimento de imagem devem primeiramente detectar e isolar cada uma das faces. Cada face será pré-processada e então uma representação com menos dimensões é obtida, de forma que a variância na imagem seja diminuída. Essa representação é importante para que o classificador seja eficiente.

Os problemas com reconhecimento facial surgem pois a face não é um corpo rígido, e as imagens podem ser tiradas de ângulos diversos. Dessa forma a representação da face deve ser resistente a variações intrapessoais na imagem, como idade e expressões, enquanto mantém uma distinção entre variações interpessoais[15]. A técnica desenvolvida em [16], utilizada para a detecção de faces, é considerada um marco para reconhecimento facial, possibilitando destacar as faces de diversos indivíduos em imagens, facilitando assim o primeiro passo para que seja possível a realização de reconhecimento facial.

Em [11], são descritas as técnicas mais utilizadas para reconhecimento facial até 2009. Este trabalho divide as técnicas de reconhecimento facial em técnicas baseadas em características, abordagens holísticas e que utilizam inteligência artificial. As primeiras técnicas utilizadas foram baseadas em características e buscaram definir uma representação de menos dimensões baseada em razões de distância, áreas e ângulos. Após essa tentativa, surgiram as abordagens holísticas, utilizando estatísticas e inteligência artificial capaz de aprender e obter resultados melhores em conjuntos de imagens de faces.

2.2.2 Conceitos Básicos

Reconhecimento facial é uma tarefa que humanos realizam rotineiramente e sem esforço todos os dias. Com a disponibilidade de computadores com elevada capacidade computacional a um baixo custo, o interesse em processamento de imagens digitais e vídeos aumentou consideravelmente em diversas aplicações, incluindo autenticação biométrica, segurança, interação humano-computador e gerenciamento de multimídias[4]. Assim, o desenvolvimento e pesquisas para a automação do reconhecimento facial surgem naturalmente.

Segundo [4], essas pesquisas em reconhecimento facial são motivadas não apenas pelos desafios fundamentais que este problema apresenta, mas também pelo grande número de aplicações práticas onde identificação humana é necessária. Um sistema de reconhecimento facial é esperado que identifique faces presentes em uma imagem ou vídeo automaticamente, podendo operar nos dois modelos apresentados na seção anterior:

- Verificação facial (ou autenticação),
- Identificação facial.

A performance de um sistema de reconhecimento facial depende de um conjunto de fatores como luminosidade, posição da face, expressões, diferença de idade da imagem na base de dados com a atual, cabelo, óculos e movimento. Assim, podemos dividir as aplicações de reconhecimento facial em duas categorias em termo da cooperatividade do usuário:(1) Cenários com usuários cooperativos e (2) cenários com usuários não cooperativos.

Os casos cooperativos são encontrados em aplicações como login em um computador e controle de acesso físico, onde o usuário está disposto a cooperar apresentado a face de uma maneira apropriada (por exemplo, na norma frontal, sem expressões e com olhos abertos) para ser garantido o acesso ou privilégio. Nos casos não cooperativos, geralmente encontrados em sistemas de segurança, o usuário não tem conhecimento que está sendo identificado, dificultando assim o processo de identificação[4].

2.2.3 Abordagens para o Reconhecimento Facial

Técnicas Holísticas

Em 1933, é proposta a técnica de análise de componente principal, do inglês *Principal Component Analysis* (PCA)[17], isso permitiu que em [18] fosse possível representar uma face como uma combinação de autovetores (*eigenvectors*). Essa é a base de dois dos algoritmos clássicos de reconhecimento faciais, o *Eigenfaces*[19] e o *Fisherfaces*[20], ambos são técnicas que utilizam marcos faciais baseadas em PCA, e ainda possuem aplicação nos dias de hoje.

O *Eigenfaces*[19], tem como principal característica a utilização de marcas significantes na face, essas marcas são chamadas de *eigenfaces*, eles são *eigenvectors* (principais componentes) do conjunto de faces, não sendo necessário corresponderem a características reais como olhos, nariz e boca. A face é representada por uma soma ponderada das características *eigenface*, dessa forma, o reconhecimento se torna uma comparação entre os valores obtidos na imagem, com os computados previamente de indivíduos conhecidos.

Assim como o *Eigenfaces*, o *Fisherfaces*[20] também é baseado em características mais significantes da face, porém como melhoria em relação ao seu antecessor, o *Fisherfaces* trouxe alterações que o tornam mais flexível em relação a variações de luz e expressões faciais. Este método é baseado no Discriminante Linear de *Fisher* (FLD), capaz de produzir classes bem separadas mesmo em condições adversas de luminosidade ou expressões faciais. Dessa forma, os resultados obtidos com o *Fisherface* possui uma taxa de erros menor se comparado com o *Eigenface*.

Inteligência Artificial

As abordagens atuais para reconhecimento facial têm utilizado redes neurais para a realização de reconhecimento facial. O *DeepFace*[6], *FaceNet*[7] e o *OpenFace*[21] são capazes de obter acurácia acima de 90% na base de dados LFW[22].

Esta base de dados[22] apresenta 13233 faces, sendo 5749 indivíduos diferentes. Dessas indivíduos, 1680 pessoas possuem duas ou mais imagens na base de dados e as 4069 pessoas restantes possuem apenas uma imagem na base de dados. As imagens possuem

diferentes posições e qualidades. Todas as imagens tiveram as faces detectadas pela função `cvHaarDetectObjects` do OpenCV, versão 1.0.0, com o classificador Haar Cascade `haarcascade_frontalface_default.xml`.

As três ferramentas citadas realizam a detecção, o alinhamento e o reconhecimento em cada uma das imagens da base de dados. Esse processo é conhecido como *pipeline Detection-Alignment-Recognition* (DAR), como mostra a Figura 2.4. Para completar este *pipeline* são necessários algoritmos que automatizem cada estágio do *pipeline*. Também é necessário que cada estágio aceite as imagens do estágio anterior e preparem a imagens para o estágio seguinte do pipeline.

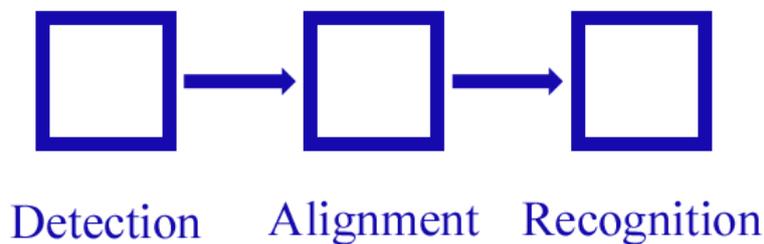


Figura 2.4: pipeline Detection-Alignment-Recognition(DAR), retirado de [22].

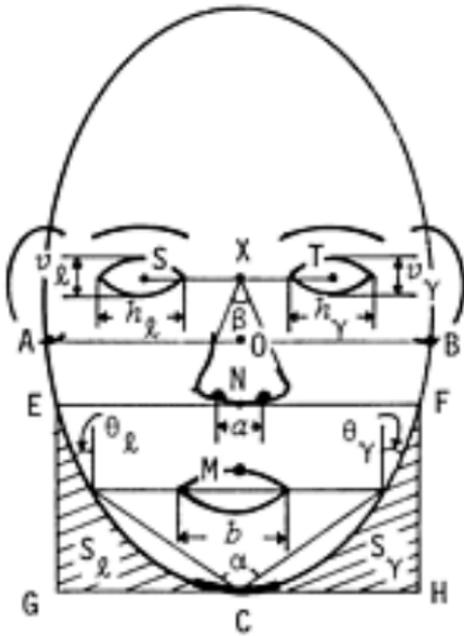
A grande diferença de cada um dos algoritmos é forma como é realizado o alinhamento. No *FaceNet*, o alinhamento é feito através de um treinamento em um grande conjunto de imagens, gerando uma imagem de face padrão. O alinhamento ocorre tentando colocar os olhos, nariz e boca em posições similares em cada imagem. O *DeepFace* gera um modelo 3D da imagem, e faz com que esse novo modelo seja colocado em norma frontal. O *OpenFace* utiliza uma simples transformação afim 2D para deixar os olhos e nariz aparecendo em posições semelhantes.

Com as imagens alinhadas é realizado o treinamento para extrair características e criar uma representação de cada uma das imagens. O *FaceNet* utiliza uma base de dados privada com mais de 100 milhões de imagens, o *DeepFace* também utiliza uma base de dados privada com 4,4 milhões de imagens, já o *OpenFace* utiliza apenas 500 mil imagens de duas base de dados abertas para pesquisas.

Para a base de dados LFW, um ser humano consegue acurácia de 97.53%, o *FaceNet* consegue uma acurácia de 99.64%, o *DeepFace* 97.35% e o *OpenFace* 92.92%. Em comparação com algoritmos que não utilizam redes neurais, o eigenfaces obtêm apenas 60.02% de acurácia. Porém todos eles precisam de diversas imagens de um mesmo indivíduo para obterem altos índices de acurácia.

Reconhecimento Facial Baseado em características

Nos anos 70, Kanade[8] propôs um método para reconhecimento facial automático utilizando métodos simples de reconhecimento facial. Kanade utilizou um vetor com 16 parâmetros da face extraídos automaticamente, como razões de distâncias, áreas e ângulos, e utilizou a distância euclidiana para obter uma acurácia de 75% em uma base de dados com 20 pessoas diferentes e duas imagens por pessoa (uma imagem de referência e outra de teste). A Figura 2.5 mostra quais os 16 parâmetros que foram utilizados.



S_l e S_γ : Áreas sombreadas

$$x_1 = AB/OC$$

$$x_2 = ST/AB$$

$$x_3 = NC/OC$$

$$x_4 = \text{Curvatura no topo do queixo}$$

$$x_5 = (\Delta EGC + \Delta FCH)/(S_l + S_\gamma)$$

$$x_6 = \alpha$$

$$x_7 = \frac{1}{2}(\theta_l + \theta_\gamma)$$

$$x_8 = \beta$$

$$x_9 = XN/XC$$

$$x_{10} = NM/XC$$

$$x_{11} = (h_l + h_\gamma)/ST$$

$$x_{12} = (v_l + v_\gamma)/ST$$

$$x_{13} = a/ST$$

$$x_{14} = b/ST$$

$$x_{15} = ST/XN$$

$$x_{16} = (h_l + h_\gamma)/(v_l + v_\gamma)$$

Figura 2.5: 16 parâmetros faciais utilizados por Kanade para a realização do reconhecimento facial. Imagem retirada de [8].

Os parâmetros são armazenados em um vetor $X = (x_1, x_2, \dots, x_{16})$. Para eliminar as diferenças de dimensões e escala entre os componentes de X , o vetor X é normalizado de

acordo com a equação:

$$y_i = \frac{x_i - m_i}{\sigma_i}, i = 1, 2, \dots, 16, \quad (2.3)$$

onde $m_i = E[x_i]$, $\sigma_i^2 = E[(x_i - m_i)^2]$ e E é a função valor esperado. Isto é, y_i significa quão longe o parâmetro x_i está do valor esperado $m_i = E[x_i]$, normalizado pelo desvio padrão σ_i . Assim uma face é representada pelo vetor $Y = (y_1, y_2, \dots, y_{16})$.

Para obter a similaridade é calculada a distância entre as imagens Y e Y' na forma:

$$d(Y, Y') = \sum_{i=1}^{16} |(y_i - y'_i)|^k. \quad (2.4)$$

Assim, para a imagem de entrada Y , a resposta é a imagem Y'' de um conjunto R de imagens, onde

$$\min_{Y' \in R} d(Y, Y') = d(Y, Y''). \quad (2.5)$$

Em [8], os testes foram realizados para $k = \frac{1}{2}, 1, 2, 3$. Com isso foi possível obter até 75% de acerto no conjunto de 20 imagens. Porém na época a extração automática de características faciais ainda não era muito aprimorada, fazendo com que o resultado fosse semelhante ao obtido manualmente[8].

Segundo [23], esta abordagem é interessante por ser mais tolerante a variações em luminosidade, escala e rotações. Dependendo apenas da marcação das características. Em 95, [23] propôs um método que conseguiu obter 95% de acerto em uma base de dados com 685 imagens(uma de cada indivíduo). Essa abordagem utilizou 35 características da face e 30 realações entre elas, porém só foi possível extraí-las manualmente. A Figura 2.6 mostra quais pontos foram marcados manualmente nesta abordagem para o reconhecimento facial.

Mais recentemente, [24] propôs um método de reconhecimento facial onde a extração das características é automática e possibilitou, além do reconhecimento, a diferenciação de idade e sexo dos indivíduos nas imagens. O processo de reconhecimento é dividido em três partes: Pré-processamento, extração de características e classificação, como é possível verificar na Figura 2.7.

Nesta abordagem, foram utilizadas a distância entre o centro das pupilas, a distância entre o meio dos lábios e a ponta do nariz, a distância entre a ponta do nariz e a linha que junta os dois olhos, o formato da face, largura dos lábios, a razão entre a largura e a altura da face, a cor da pele, a região da barba, a região dos lábios, o final da sobrancelha, a testa, as bochechas, os cílios e a lateral do nariz. Com essas informações são calculadas 4 razões principais:

$$R1 = \frac{d(L_eye, R_eye)}{d(eye, nose)}, \quad (2.6)$$

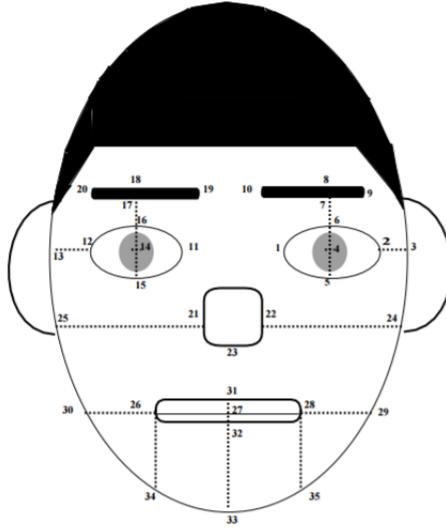


Figura 2.6: 35 pontos marcados manualmente para a realização de reconhecimento. Imagem retirada de [23].

$$R2 = \frac{d(L_eye, R_eye)}{d(eye, lips)}, \quad (2.7)$$

$$R3 = \frac{d(eye, nose)}{d(eye, chin)}, \quad (2.8)$$

$$R4 = \frac{d(eye, nose)}{d(eye, lips)}, \quad (2.9)$$

onde $d(a, b)$ é a distância entre os pontos a e b em uma linha reta, L_eye e R_eye são os centros dos olhos esquerdo e direito, eye é o eixo que corta os dois olhos no centro, $nose$ é a base do nariz, $lips$ é a divisão dos lábios e $chin$ é a parte mais inferior do rosto. A Figura 2.8 permite a visualização destas características.

Com essas 4 razões e as demais características extraídas da face foi possível realizar o reconhecimento facial, acertando 100% dos 14 indivíduos testados. Este método também permitiu que fosse possível determinar o sexo e a idade dos indivíduos testados, conseguindo classificar o gênero da pessoa com 95% de acurácia e a idade com 90%. Porém, não foi realizado testes para um grande volume de imagens e todas as imagens são em norma frontal com luminosidade adequada.

2.3 Pontos Cefalométricos

A extração de características faciais vem sendo utilizada para obter informações de imagens faciais para os mais diversos campos da visão computacional, como determinação do

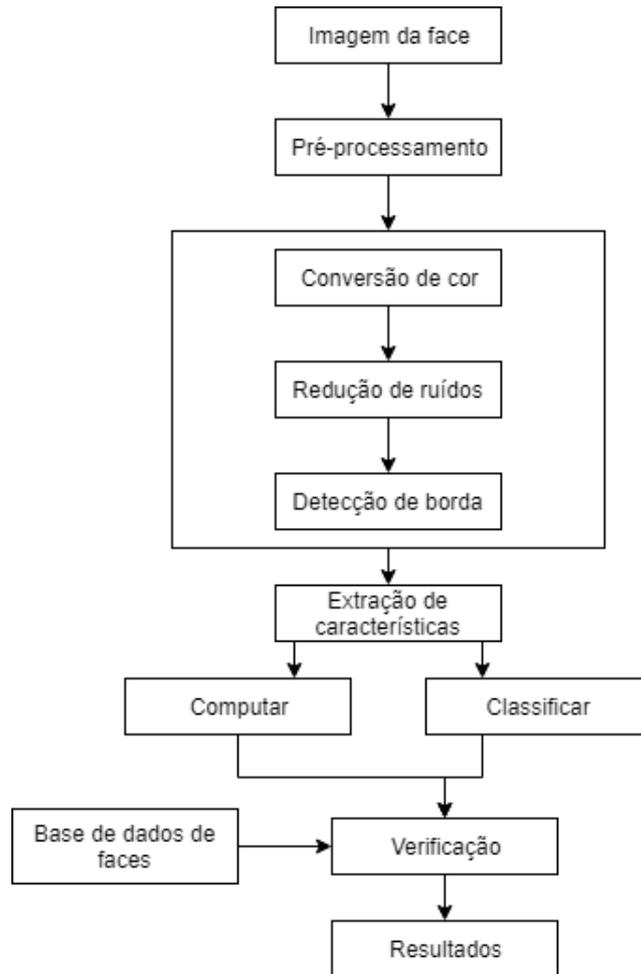


Figura 2.7: Diagrama descrevendo o processo de reconhecimento facial utilizado em [24].

sexo e da idade de pessoas, assim como também podem ser utilizados para reconhecimento facial[25].

De acordo com Byers[26], antropologia forense possui duas abordagens: a primeira é utilizada para coletar a informação e a segunda para analisar os dados coletados. Antropometria é a abordagem a qual fornece métodos para a obtenção de dados, enquanto que os marcos faciais e suas razões por eles obtidas são considerados como uma abordagem de análise. As técnicas antropométricas possuem nomenclaturas diferentes de acordo com a região de análise. Quando se estuda as dimensões de uma cabeça humana (como a distância entre os olhos), essas técnicas são chamadas de cefalometria e os dados métricos extraídos de um conjunto de pontos específicos em uma face são chamados de pontos cefalométricos. As relações entre esses pontos (como a distância entre as pupilas) podem ser chamadas de razões foto-antropométricas (RFAs).

Em [27], foi feito um estudo do crescimento do corpo ao decorrer do tempo. Nele é mostrado quais partes do corpo estabilizam o crescimento durante o desenvolvimento

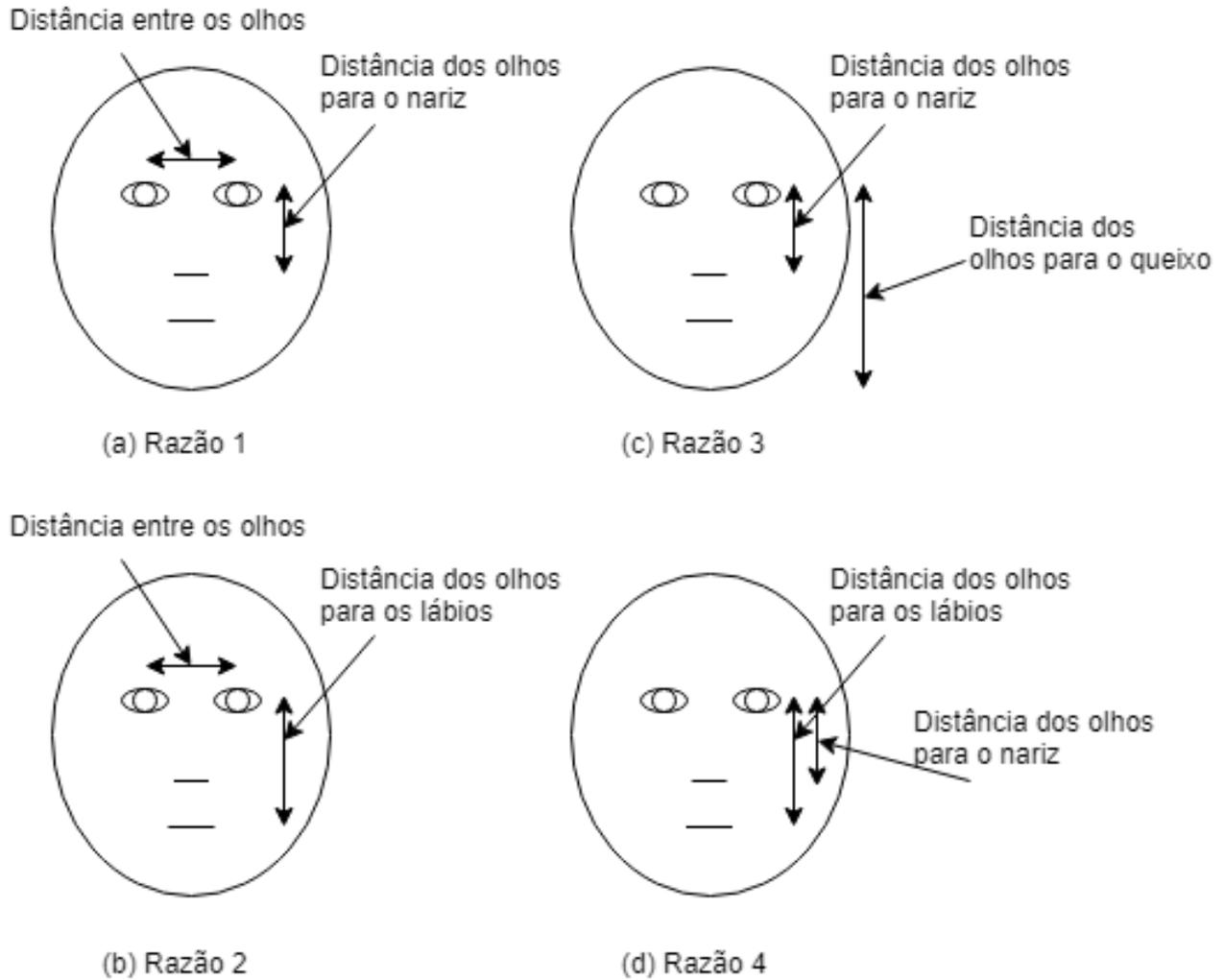


Figura 2.8: As quatro razões utilizadas para a realização do reconhecimento em [24].

de um indivíduo. Com isso foi possível mostrar que a íris está muito próxima de estar completamente desenvolvida aos 7 anos de idade. A Figura 2.9 mostra o crescimento de cada região do corpo separadamente de acordo com a idade.

Em [9] são mapeados 28 pontos cefalométricos, que posteriormente são utilizados para o cálculo de 40 razões foto antropométricas. No trabalho apresentado em [9], essas marcações faciais podem ser identificadas em pares como: Entocanthion, Exocanthion, Iridion laterale, Iridion mediale, Pupil, Zygion, Alare, Gonion, Cheilion, Crista philtri. E outras são identificadas na linha medial da face, tais como: Glabella, Nasion, Subnasale, Labiale superius, Stomion, Labiale inferius, Gnathion, Midnasal. A Tabela 2.1 descreve cada um desses pontos e a Figura 2.10 mostra onde cada um desses pontos se encontra. Ainda em [9], a escolha desses pontos é justificada pela possibilidade de serem marcados em imagens 2D, desde que apresentem certas especificações.

Tabela 2.1: Tabela com os 28 pontos utilizados em [28] e a suas descrições. Os pontos 1 – 8 são identificados na linha medial da face, e os pontos 9 – 18 são identificados em pares.

Marco Cefalométrico	Descrição
1. Glabella (g)	Intersecção entre a linha média facial e a linha horizontal que tangencia o bordo superior da circunferência orbitária.
2. Nasion (n)	Intersecção da linha média facial com a linha horizontal que passa pelo sulco palpebral superior.
3. Subnasale (sn)	Ponto mais inferior da columela (projeção que se encontra entre as narinas).
4. Labiale superius (ls)	Ponto médio da linha branca do lábio superior. Quando presente o arco de cupido, o ponto será marcado em seu ponto mais inferior.
5. Stomion (sto)	Ponto na região de encontro dos lábios superior e inferior, posicionado na metade da distância entre os cheilions.
6. Labiale inferius (li)	Ponto médio da linha branca do lábio inferior.
7. Gnathion (gn)	Ponto mais inferior do mento.
8. Midnasal (mid)	Ponto sobre a linha média da face, tendo por referência à altura dos exocanthions.
9. Entocanthion (en)	Ponto localizado no ângulo medial dos olhos, marcado no encontro das pálpebras superior e inferior.
10. Exocanthion (ex)	Ponto localizado no ângulo lateral dos olhos, marcado no encontro das pálpebras superior e inferior.
11. Iridion laterale (il)	Ponto mais lateral da circunferência iridiana.
12. Iridion mediale (im)	Ponto mais medial da circunferência iridiana.
13. Pupil (pu)	Ponto central da circunferência iridiana.
14. Zygion (zy)	Ponto mais lateral do contorno da face à altura dos ossos zigomáticos.
15. Alare (al)	Ponto mais lateral da asa do nariz.
16. Gonion (go)	Ponto sobre a linha de contorno facial, à altura do stomion.
17. Cheilion (ch)	Ponto localizado na comissura labial, em sua região mais lateral.
18. Crista philtri (cph)	Ponto alto da crista do arco de cupido, no local de chegada das colunas do filtro labial.

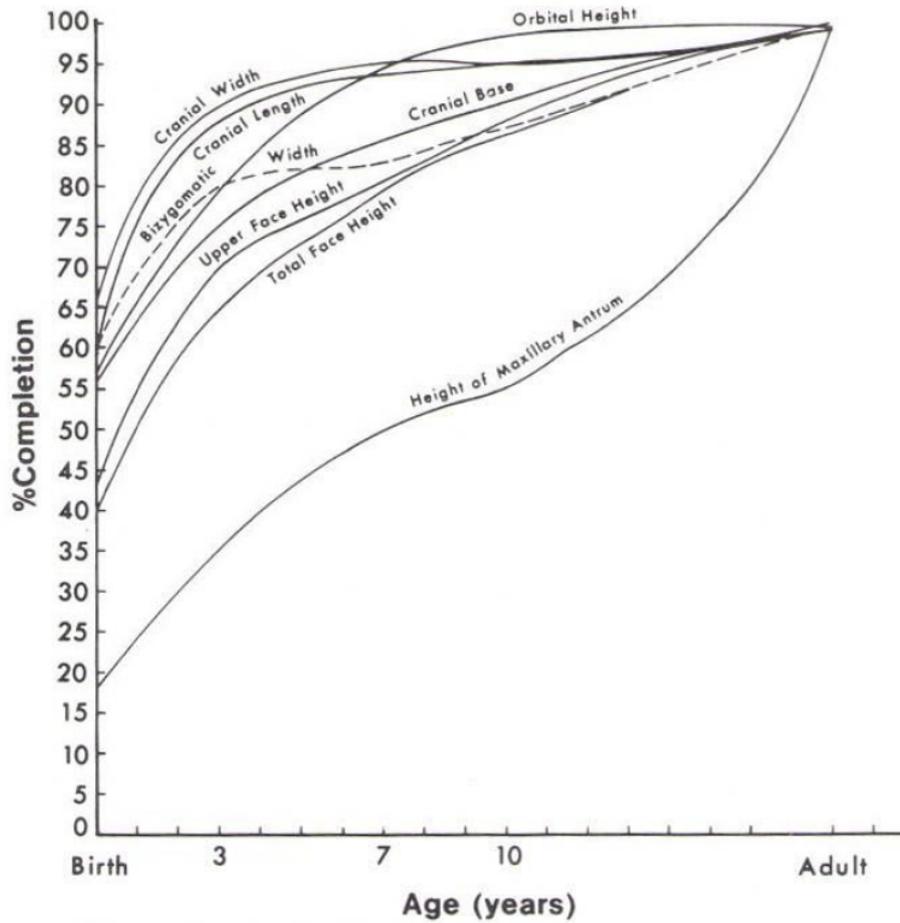


Figura 2.9: Crescimento de cada região do corpo de acordo com a idade[27].

Em [9], além dos pontos, foram propostas 40 razões foto antropométricas. Para isso, foi utilizada a projeção da distância iridiana como parâmetro para o cálculo das razões foto antropométricas. A projeção da distância iridiana (P_{iris}) é calculada como:

$$P_{iris} = \frac{(im_e - il_e) + (im_d - il_d)}{2}. \quad (2.10)$$

Ou seja, é a média das somas das diferenças utilizando os pontos irídion mediale esquerdo (ime), irídion mediale direito (imd), irídion laterale esquerdo (ile) e irídion laterale direito (ild). Com esta medida, é possível definir as 40 razões foto-antropométricas que estão descritas na Tabela 2.2. Com elas, foi possível a verificação da relação entre as razões e o sexo e a idade de crianças e adolescentes, possibilitando assim realizar uma separação tanto por idade quanto por sexo para os indivíduos testados [28]. A escolha dessas razões foi justificadas por serem as razões constantemente utilizadas para construir índices faciais em estudos antropométricos.

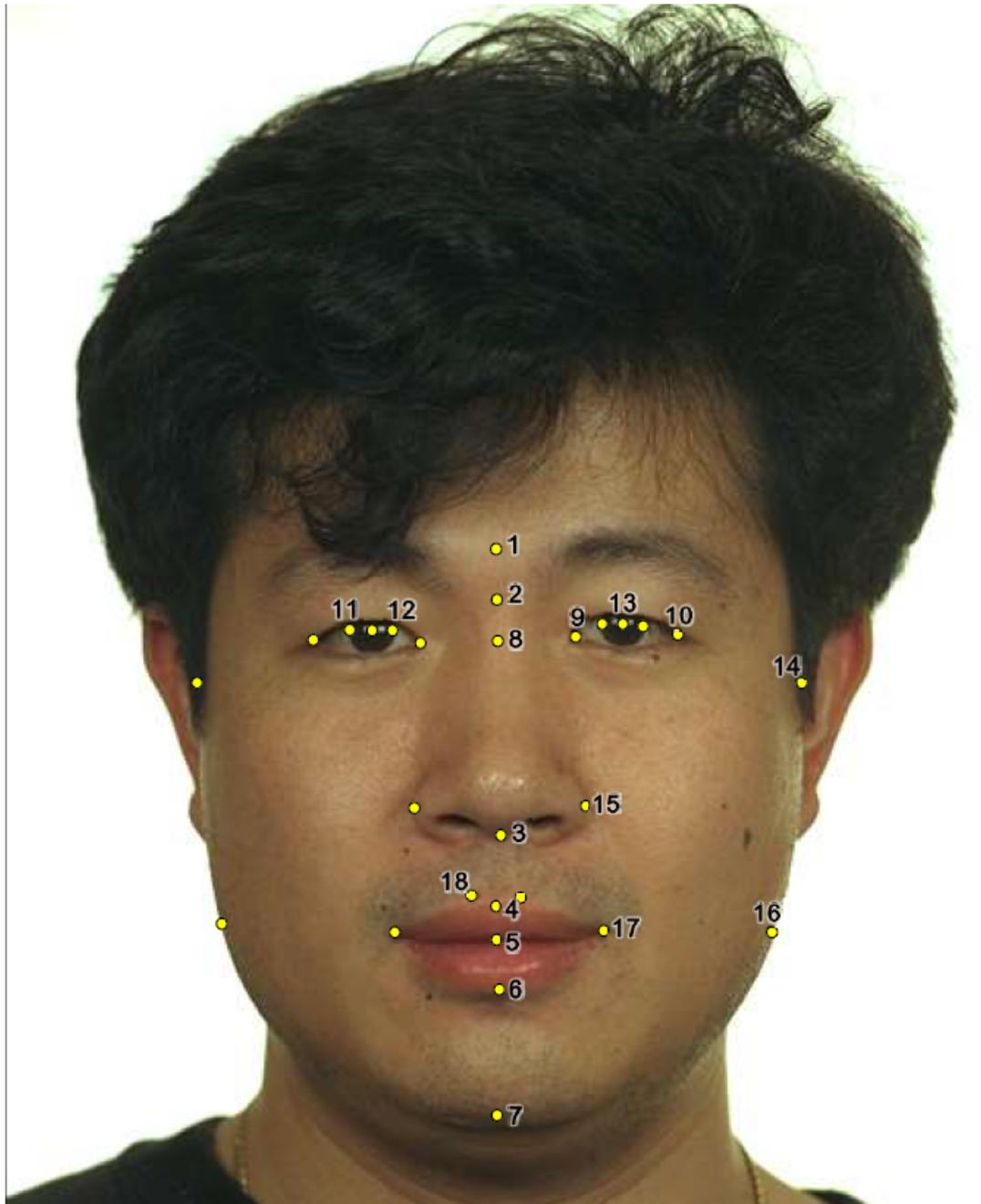


Figura 2.10: 28 pontos cefalométricos utilizados em [28] visualizados na face e suas abreviações: 1. Glabella (g); 2. Nasion (n); 3. Subnasale (sn); 4. Labiale superius (ls); 5. Stomion (sto); 6. Labiale inferius (li); 7. Gnathion (gn); 8. Midnasal (mid); 9. Entocanthion (en); 10. Exocanthion (ex); 11. Iridion laterale (il); 12. Iridion mediale (im); 13. Pupil (pu); 14. Zygion (zy); 15. Alare (al); 16. Gonion (go); 17. Cheilion (ch); 18. Crista philtri (cph).

A íris foi escolhida como fator de normalização, pois segundo mostrado por [29], as dimensões da íris permanecem constantes durante o envelhecimento. Foi mostrado nesse estudo que a íris possui um crescimento durante os estágios pré-natal, e terminam o seu

desenvolvimento durante o primeiro ano de vida. Com isso, a íris tende a manter o mesmo tamanho durante toda a vida de um indivíduo, permitindo que possa ser utilizada para estudos envolvendo o crescimento e envelhecimento da face.

Tabela 2.2: Tabela com as 40 razões foto antropométricas, com base nos pontos cefalométricos descritos na tabela 2.1 e visualizados na Figura 2.10. Todas as razões devem ser divididas pela projeção iridiana da equação 2.10

1. im-im	2. il-il	3. pu-pu	4. mec-gn
5. g-gn	6. mec-ls	7. mec-sto	8. ec-ec
9. n-gn	10. al-al	11. go-go	12. Eye width
13. g-sto	14. n-ls	15. ls-gn	16. g-ls
17. n-sto	18. sn-sto	19. en-en	20. g-sn
21. sto-gn	22. g-li	23. n-li	24. mec-li
25. ls-sn	26. sn-ls	27. ch-ch	28. n-sn
29. sn-gn	30. ls-sto	31. mec-sn	32. cph-cph
33. zy-zy	34. li-gn	35. sn-li	36. g-mec
37. g-n	38. n-mec	39. ls-li	40. sto-li

As informações apresentadas neste capítulo de revisão bibliográfica serão utilizadas para o desenvolvimento de toda a proposta de metodologia que será detalhada no capítulo seguinte.

Capítulo 3

Trabalhos Relacionados

Neste capítulo serão abordados trabalhos recentes relacionados ao trabalho aqui proposto. Serão tratados artigos que tratem sobre marcação de pontos faciais automatizada e reconhecimento facial. Serão abordadas diversas técnicas desde as que utilizem aprendizado de máquina como as que utilizam apenas processamento de imagem e visão computacional. Além disso, será apresentado trabalhos que tratam de reconhecimento facial em imagens com variação de idade e de qualidade das imagens. Ao final, será apresentada uma justificativa para a técnica proposta na Seção 4.

3.1 Marcos Faciais

A detecção de marcos faciais pode ser determinante para o sucesso de um sistema de reconhecimento facial, e por isso tem sido objeto de diversos estudos recentes, como em [30] e [31]. Assim, com estudos para automatização do processo de detecção de marcos faciais, foi possível obter grandes avanços nessa área, auxiliando assim os sistemas de reconhecimento facial.

Para aprimorar a detecção automática de marcos faciais, em [30], foi proposto o primeiro desafio para localização de marcos faciais. O objetivo desse desafio foi comparar o desempenho de diferentes metodologias para este fim, utilizando o mesmo protocolo de avaliação para todas as propostas de solução e, assim, gerar um padrão para a avaliação de marcação de pontos faciais.

Ainda em [30], foi proposto a base de dados 300-W, com foco em examinar as habilidades dos sistemas atuais de lidar com situações sem controle e naturais. Assim, o teste é capaz de cobrir diversas variações de poses, expressões, luminosidade e qualidade de imagem. Tendo foco principalmente em imagens com expressões espontâneas. A base consiste em 300 imagens em ambientes fechados e 300 imagens ao ar livre. O *ground-truth*

da base consiste em 68 pontos faciais, como mostra a Figura 3.1, podendo ser reduzida para 51 caso não sejam incluídos os pontos do limite da face.

Neste desafio foram submetidos 6 trabalhos para a avaliação, e para todas as metodologias, o caso com 51 pontos obteve melhor resultado. Além disso, em imagens ao ar livre, o desempenho foi menor, sendo a luminosidade um dos principais fatores para isso ocorrer. Ainda, o trabalho conclui que a base 300-W não esgota todas as possibilidades e existe espaço para aprimoramentos.



Figura 3.1: Imagem dos pontos marcados na base 300-W, 68 caso inclui os pontos do limite da face do rosto ou 51 caso contrário. Imagem retirada de [30].

Em 2015, [31], propôs um método para adequar um modelo 3D de uma face a uma imagem 2D. Para realizar a extração de marcos faciais, foi utilizado um método de regressão em cascata baseado em aprendizado capaz de aprender a direção do gradiente dos dados. Para este trabalho, é interessante entender a regressão em cascata utilizada.

Dada uma imagem de entrada I e um modelo pré-treinado $\Omega(\Theta)$ correspondente a um vetor de parâmetros Θ , o objetivo de um método baseado em regressão é iterativamente atualizar os parâmetros $\Theta \leftarrow \Theta + \delta\Theta$ para maximizar a probabilidade posterior $p(\Theta|I, \Omega)$. Um método baseado em regressão resolve esse problema de otimização não-linear aprendendo o gradiente de um conjunto de treinamento de uma maneira supervisionada. O objetivo é encontrar um regressor:

$$R : f(I, \Theta) \rightarrow \delta\Theta \quad (3.1)$$

onde $f(I, \Theta)$ é um vetor de características extraídas da imagem de entrada, dados os parâmetros do modelo, e $\delta\Theta$ é a atualização do parâmetro do modelo prevista. Um

método de regressão em cascata gera um regressor forte que consiste de N regressores fracos (Não possuem valor se usados individualmente) em cascata:

$$R = R_1 \circ \dots \circ R_N \quad (3.2)$$

onde R_N é o N -ésimo regressor fraco na cascata. Em [31], foi utilizado um regressor linear simples:

$$R_n : \delta\Theta = A_n f(I, \Theta) + b_n \quad (3.3)$$

onde A_n é a matriz de projeção e b_n é utilizado como compensador do n -ésimo regressor fraco. Mais especificamente, dado um conjunto de amostras de treinamento $\{f(I_i, \Theta_i), \delta\Theta_i\}_{i=1}^M$, primeiro aplica-se o algoritmo de regressão para aprender o primeiro regressor fraco minimizando a perda:

$$\sum_{i=1}^M \|A_1 f(I_i, \Theta_i) + b_1 - \delta\Theta_i\|^2 + \lambda \|A_1\|_F^2, \quad (3.4)$$

onde o λ é um autovalor da matriz A_1 , em seguida atualiza a amostra de treinamento, por exemplo os parâmetros do modelo e o vetor de características correspondentes, utilizando o regressor aprendido para gerar um novo conjunto de treinamento para o aprendizado do segundo regressor fraco. Esse processo se repete até que haja convergência ou exceder um número pré-determinado de regressores em cascata. Dessa forma, foi possível extrair as características de forma extremamente rápida, possibilitando aplicações em tempo real com alta precisão.

3.2 Reconhecimento Facial

3.2.1 Reconhecimento Facial por Aprendizado de Máquina

Atualmente, a maioria das aplicações de reconhecimento facial estão utilizando uma abordagem por aprendizado de máquina[6][7][32][33][34]. Com a grande quantidade de imagens disponíveis essas metodologias têm conseguido bons resultados, porém também necessitam de um grande volume de imagens para realizar os treinamentos.

A primeira abordagem a obter um nível humano de reconhecimento foi a DeepFace[6], desenvolvida pelo Google. Essa metodologia é utilizada para faces em ambientes não controlados. A utilização de *Deep Learning* foi um marco para os algoritmos de reconhecimento facial. Em resumo, o DeepFace contribui em:

- Desenvolvimento de uma arquitetura de rede neural profunda efetiva e método de aprendizado que se aproveita de uma base de dados grande para obter uma representação facial que generalize bem para outras bases de dados;
- Um efetivo sistema de alinhamento facial baseado em modelos 3D explícitos de faces;
- Avanços significativos no estado da arte para a base de dados LFW[22], atingindo níveis quase humanos de performance.

Um dos principais diferenciais apresentados em [6] é o método para realizar o alinhamento da face. O processo de alinhamento se inicia localizando 6 pontos de referência localizados dentro de um corte de detecção da face, sendo marcados os centros dos olhos, ponta do nariz e da boca e os dois cantos da boca. Esses pontos são utilizados para colocar em escala, rotacionar e transladar a imagem em seis posições utilizadas como padrão, gerando assim uma imagem alinhada no plano 2D. Porém ainda falta compensar as rotações fora do plano, que são de extrema importância em condições não controladas.

Assim, para alinhar faces que possuem rotações fora do plano, é utilizado um modelo genérico 3D que é utilizado para distorcer a imagem 2D alinhada para uma imagem modelada em 3D. Gerando uma versão 3D alinhada da imagem. Esse processo ocorre localizando 67 marcos faciais na imagem alinhada em 2D, esses pontos são relacionados com pontos do modelo genérico 3D. Por último, são realizados ajustes para que a perda de informação seja reduzida.

Em [6], diferente do que ocorria em algoritmos de reconhecimento faciais anteriores, é criado uma representação genérica de imagens faciais através de uma grande rede neural profunda. Essa rede é treinada em uma tarefa de reconhecimento facial multi-classe, para classificar a identidade de uma face. Com essa metodologia, o DeepFace conseguiu obter mais de 97% de acerto na base LFW.

Em seguida, surgiu o FaceNet[7], desenvolvido pelo Facebook. Nessa metodologia, a abordagem muda, utilizando uma rede convolucional profunda para diretamente otimizar o processo. A parte mais importante desta abordagem é a utilização de aprendizado de ponta a ponta em todo o sistema. Para atingir isso, é utilizado um *Triplet Loss* que reflete diretamente o que se deseja obter na verificação da face, reconhecimento e clusterização. Isto é, procura-se obter um $f(I)$, para uma imagem I em um espaço \mathbb{R}^d , no qual a distância quadrática entre todas as faces, independentemente das condições da imagem, para um mesmo indivíduo sejam pequenas, enquanto que para indivíduos distintos sejam grandes.

Com isso, deseja garantir que uma imagem I_i^a (âncora) de uma pessoa específica esteja mais próxima de todas as outras imagens I_i^p (positivas) de uma mesma pessoa do que de qualquer imagem I_i^n (negativas) de qualquer outra pessoa. Assim, é possível avaliar todos

os grupos de três imagens, possibilitando o aperfeiçoamento do método. Esse processo é chamado de *Triplet Loss*, a figura 3.2 ilustra essa minimização da distância entre imagens de uma mesma pessoa.

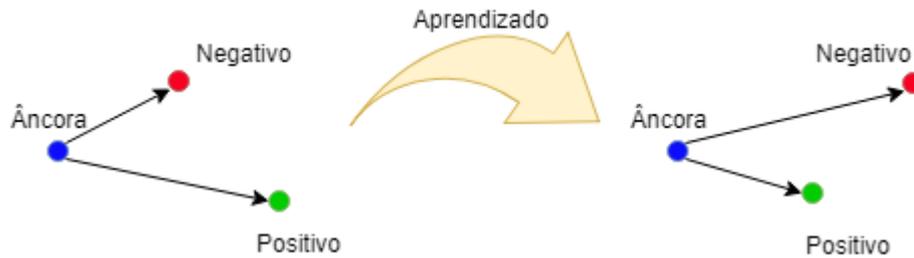


Figura 3.2: *Triplet Loss* minimiza a distância entre a âncora e as imagens positivas, e maximiza a distância entre a âncora e as imagens negativas. Adaptado de [7].

Essa metodologia também foi testada na base de dados LFW[22], e obteve uma taxa de acerto acima de 99%, ultrapassando a performance humana de 97.5%. Esse método não necessita de um alinhamento 3D complexo da face como o do DeepFace, sendo necessário apenas um alinhamento mínimo no recorte da face.

Utilizando a mesma metodologia, [35] propõe que a base de dados LFW[22] gera uma imparcialidade já que as imagens são sempre de celebridades e assim são diferentes de imagens de pessoas comuns capturadas no dia-a-dia. Porém, substituir essas imagens por imagens de pessoas comuns é difícil, pois pode acarretar problemas de privacidade. Em seu trabalho, [35] juntou diversas imagens de diferentes bases de dados para realizar os treinamentos em uma técnica chamada de adaptação de domínio.

Com isso, adaptando a técnica utilizada no FaceNet, e utilizando apenas um modelo de CNN, foi possível obter 99.33% na base LFW, mesmo sem realizar alinhamento das faces e com uma base de dados 44 vezes menor do que a utilizada pelo FaceNet.

Mais recentemente, em 2017, foi proposto em [34] um framework para reconhecimento facial utilizando um descritor de histograma de padrão binário local (LBPH), multi-KNN (k vizinhos mais próximos) e *Back-Propagation Neural Network* (BPNN).

Inicialmente é utilizado o LBPH para reduzir as dimensões da imagem, selecionando características significativas. Com isso, é obtido um robusto T-Dataset, que será usado como entrada para o BPNN ao invés de utilizar o LBPH para cada imagem do treinamento. Esse T-Dataset é obtido com base na correlação entre as novas representações de todas as imagens de treinamento.

Tendo como base a representação LBPH de cada imagem, é calculada a distância entre cada imagem de treinamento e todas as outras imagens de treinamento utilizando cinco métodos de distância (Correlação, euclidiana, canberra, manhattan e mahalanobis).

O método KNN é utilizado para encontrar a saída de cada imagem de treinamento. Dessa forma, as imagens de menor distância dos 5 métodos de distância serão as entradas para o treinamento da BPNN. Assim, utilizando apenas 2200 pares de imagens para o treinamento, a metodologia conseguiu obter 95.71% de acerto na base de dados LFW[22].

Ainda em 2017, visando solucionar problemas de reconhecimento facial em imagens de baixíssima resolução, [36] propõe uma metodologia onde são extraídos tanto características visuais como de textura das imagens. Para extrair as características visuais, é utilizada uma transformada discreta dos cossenos (DCT), e para obter as características de textura é utilizada uma *gray-level co-occurrence matrix* (GLCM).

Essas características visuais e de textura são fundidas para formar uma característica híbrida. Em seguida, uma rede neural do tipo *feed forward neural network* (FFNN) é alimentada com essa característica híbrida para realizar o reconhecimento. Essa metodologia foi testada na base de dados Georgia Tech (GT), onde as imagens foram redimensionadas para simular baixa resolução. A Figura 3.3 ilustra as imagens utilizadas para testes.



Figura 3.3: Imagem de baixíssima resolução utilizada em [36].

Em 2018, [33] propôs um novo método para realizar reconhecimento facial em imagens sequencias onde há mudanças de pose. Foi utilizado uma *Recurrent regression neural network* (RRNN) para solucionar os problemas de reconhecimento facial com mudanças de pose em imagens paradas e vídeos. Na RRNN, três unidades básicas são consideradas para modelar um potencial sequencia de dados. A primeira unidade é a do codificador e decodificador, que é utilizada para modelar a reconstrução sequencial. A segunda unidade é para restringir a abrangência da sequência. E a última é para utilizar a informação discriminada.

Realizando transformações progressivas para as imagens sequencialmente adjacentes, a RRNN pode adaptativamente memorizar e esquecer informações que beneficiem a classificação final. Nos testes realizados, foram utilizados as bases MultiPIE[37], que contém imagens em diferentes posições, iluminação e expressões, e a YouTube celebrities (YTC)[38], que contém 1910 vídeos de faces de 47 pessoas diferentes em diversas varia-

ções de poses, iluminação e expressões. Na base MultiPIE, essa metodologia foi capaz de obter mais de 95% de acerto para todas as variações de pose. Já na base YTC, conseguiu obter 84.6% de acerto.

3.2.2 Reconhecimento Facial Convencional

Ainda existem diversos métodos para resolver reconhecimento facial que não utilizam aprendizado de máquina. Estas metodologias são capazes de obter altos níveis de acerto sem a necessidade de um grande volume de dados para treinamento. Atualmente, possuem taxas de acerto inferiores aos algoritmos que utilizam aprendizado de máquina, porém são capazes de resolver problemas em situações específicas de forma satisfatória e com menos esforço. Aqui serão trazidas metodologias propostas recentemente para solucionar diversos problemas relacionados a reconhecimento facial.

Em 2013, foi proposta uma metodologia em [39] para realizar reconhecimento em imagens de modalidades diferentes, como imagens infravermelho com fotos ou esboços com fotografias. A utilização de similaridades não-lineares é interessante para este tipo de situação pois um conjunto de treinamento com imagens de cada modalidade pode ser utilizado com protótipo e, dependendo da modalidade da nova imagem, a imagem de cada um dos protótipos pode ser selecionada de acordo com modalidade correspondente.

Nesta metodologia, a imagem de teste e as da base de dados são inicialmente filtradas com três tipos diferentes de filtro, e dois descritores de características locais diferentes são extraídos. Um conjunto de treinamento age como um conjunto de protótipos no qual cada tipo de protótipo tenha uma imagem nas modalidades utilizadas na base de dados e na imagem de teste. A similaridade do núcleo não-linear entre uma imagem e a do protótipo é medida em correspondência com a modalidade. Essa metodologia obteve bons resultados utilizando quatro cenários diferentes (próximo do infravermelho, infravermelho térmico, esboço de olho, e esboço forense).

Mais recentemente, em 2017, [40] propôs um método para melhorar o reconhecimento facial em imagens com diferença de idade. A metodologia consiste em três passos:

- Segmentação da face, gerando três subimagens da imagem de entrada;
- Um score de similaridade é calculado utilizando um gradiente de um par que consiste da imagem de entrada e uma imagem registrada para cada um das três subimagens geradas;
- Um classificador de verificação avalia os scores de similaridade e prevê a probabilidade de duas imagens pertencerem a uma mesma pessoa.

Cada uma das subimagens aborda uma região da face, sendo elas: sobrancelhas e olhos, olhos e nariz, e nariz e boca. A figura 3.4 ilustra as três subimagens obtidas.



Figura 3.4: Três subimagens utilizadas para realizar o reconhecimento facial. Imagem retirada de [40].

Ainda em 2017, [41] propõe um método para reconhecimento utilizando as transformadas de contourlet (CNT) e de curvelet (CLT). A CNT permite obter informações suaves do contorno em diferentes direções, enquanto que a CLT, por possuir propriedades multi-escala, multi-direcional e anisotrópica, foi empregado para as bordas mais proeminentes. Imagens de treinamento pré-processadas são decompostas utilizando o CNT e os coeficientes de sub-bandas direcionais são analisados para obter características das imagens. Já o CLT é utilizado para extrair diferentes características estatísticas. Por fim, as características das duas transformadas são integradas.

Para realizar o reconhecimento, foi utilizada o conjunto das características em um classificador SVM (*support vector machine*). Com essa metodologia, foi possível obter acurácia de 98.10% na base de dados FERET[42].

Em 2018, [43] propôs um método que utiliza uma versão aperfeiçoada do PCA em conjunto com uma classificação por regressão linear (LRC). O PCA aperfeiçoado diferencia do tradicional no fato de ao realizar o treinamento, ao invés de usar o *eigenvector* de cada imagem de uma determinada classe, ele calcula a média para a classe toda, como a média de uma classe é uma combinação linear das imagens de própria classe, a média mantém um grande número de variações de uma imagem específica. Além disso, o tempo necessário para realizar o treinamento será drasticamente reduzido.

Tendo as médias de cada vetor de faces, é calculada a matriz de covariância. Em seguida, são calculados os *eigenvectors* da matriz de covariância. Alguns *eigenfaces* serão escolhidos para representar todo o conjunto de dados. E, por último, cada uma das faces será representada como uma combinação linear de todos os *eigenvectors*. Com isso, é utilizado um classificador por regressão linear para determinar a classe de uma nova imagem.

Esse método também se mostrou eficiente para a identificação de expressões faciais, porém ainda não conseguiu se mostrar eficiente em casos de luminosidade extrema ou corrupção de pixels.

Capítulo 4

Metodologia

4.1 Descrição do Método

Neste trabalho, é proposta uma metodologia para realização de reconhecimento facial biométrico onde não são necessárias diversas imagens de um mesmo indivíduo, o que dificulta o uso de uma rede neural capaz de obter bons resultados. Ainda, o método proposto permite que seja realizado reconhecimento facial em imagens afetadas pelo fator de envelhecimento ou então imagens com escalas diferentes.

Dessa forma, utilizando apenas técnicas de processamento de imagem e visão computacional, é possível realizar um reconhecimento facial de forma rápida e eficiente em situações que são consideradas problemáticas para demais sistemas de reconhecimento facial biométricos tradicionais.

Os pontos faciais aqui utilizados já se mostraram capazes de determinar idade e sexo de indivíduos, e neste trabalho verificaremos se também são capazes de individualizar uma pessoa, possibilitando assim que sejam utilizados para realização de um reconhecimento facial biométrico eficiente. Estes pontos foram escolhidos pois, de acordo com o trabalho de [28], eles são capazes de representar o crescimento da face e também porque, de acordo com [27], a íris não cresce drasticamente depois dos 7 anos de idade. Assim, é possível que os impactos do crescimento da face no reconhecimento facial, em decorrência do envelhecimento, sejam reduzidos com a utilização destes pontos, criando assim a possibilidade de obter um sistema de reconhecimento facial biométrico que seja robusto mesmo com a passagem de tempo entre a imagem armazenada na base de dados e uma imagem atual do indivíduo. Dessa forma, essa metodologia pode ser aplicada para identificação de pessoas desaparecidas por vários anos.

A metodologia a ser utilizada nesta proposta será dividida em três partes principais: inicialmente será necessária a obtenção dos pontos cefalométricos e cálculo das razões foto antropométricas de forma automatizada utilizando somente uma imagem facial como

entrada para esta etapa. Com as razões calculadas para cada imagem, será possível realizar a comparação das imagens, sendo computados o número de falsos positivos, falsos negativos, verdadeiros positivos e verdadeiros negativos. Assim, é possível a obtenção da FAR e FRR e, assim, obter a EER da metodologia. Por último será proposta uma metodologia para a redução do número de RFAs necessárias para tornar a metodologia de reconhecimento facial proposta mais rápida, retirando as RFAs que menos influenciam os resultados.

Com esses procedimentos, pretende-se obter um método de reconhecimento facial biométrico que tenha um bom desempenho mesmo em casos em que as imagens possuam uma diferença de idade de até 5 anos ou então para imagens que possuam escalas diferentes. Além disso, também é desejável utilizar o mínimo necessário de razões citadas na Seção 2.3 sem prejudicar o desempenho da metodologia proposta.

Para o desenvolvimento desta metodologia proposta, será utilizada a linguagem de programação C++ em conjunto com a biblioteca para visão computacional OpenCV[44].

4.2 Descrição da Base de Dados

Inicialmente, todas as imagens utilizadas neste trabalho estão de acordo com os padrões descritos nas normas da ICAO 9303[45] e atendem com as seguintes propriedades das imagens: boa luminosidade, visão frontal e sem óculos, como pode ser visualizado na Figura 4.1. O padrão da ICAO é necessário porque, para que seja possível medir o tamanho da íris de um indivíduo, é importante que a íris esteja visível, assim, todas as imagens utilizadas possuem resolução de aproximadamente 640×480 .

Assim, foram utilizadas três bases de dados diferentes, cada uma com o objetivo de verificar parte dos objetivos propostos neste trabalho.

A primeira base utilizada contém 1000 imagens de indivíduos diferentes. A Figura 4.1 exemplifica as imagens utilizadas nesta base de dados. Esta base também possui as posições dos marcos faciais utilizados marcados manualmente. Nesta base será testado se a marcação automática dos marcos faciais pode ser utilizada e também será testado se estes marcos são capazes de diferenciar os indivíduos.

A segunda base utilizada possui 40 imagens de 8 indivíduos diferentes. Nessas imagens não há grande variação de posição, porém há variação na escala das imagens (mantidas as proporções), como é possível ver na Figura 4.2. O objetivo desta base é testar se é possível utilizar a metodologia proposta para realizar reconhecimento facial em imagens que possuem variação de escala.

Na terceira e última base utilizada, estão presentes 6622 imagens, sendo duas de cada indivíduo com um intervalo de 5 anos entre as imagens da mesma pessoa, a Figura 4.3



Figura 4.1: Imagem da base de dados utilizada, estando de acordo com as normas previstas pela ICAO.



Figura 4.2: Exemplo de imagens com variação de escala.

ilustra a situação do crescimento do indivíduo neste período de 5 anos. Nesta base, existe maior variação na expressão facial e na posição da face. Nesta base, pretende-se verificar a capacidade da metodologia proposta em realizar reconhecimento facial em intervalos de 5 anos de crescimento da face, principalmente para usuários ainda em idade de crescimento (de 4 a 18 em média de acordo com o sexo) e variações decorrentes do envelhecimento do indivíduo.

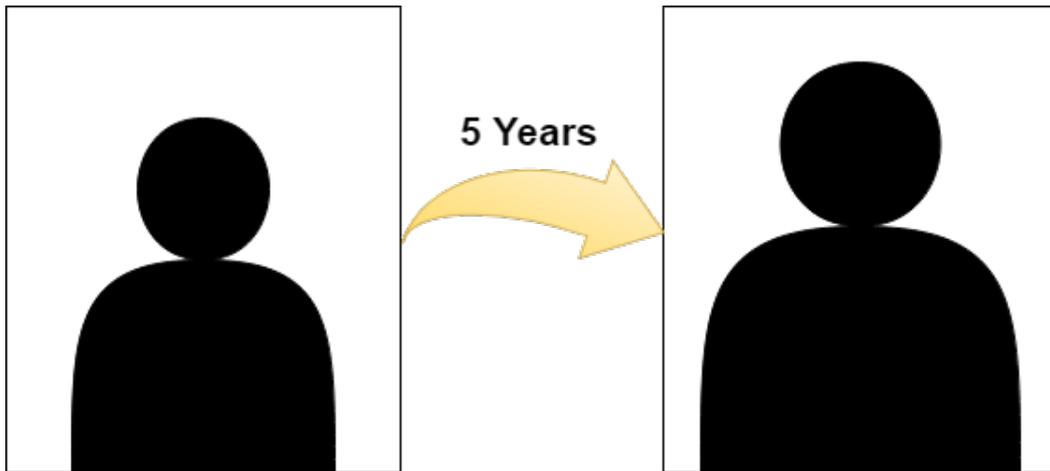


Figura 4.3: Imagem que ilustra as imagens utilizadas na terceira base dados. O sistema deve lidar com o crescimento do indivíduo durante 5 anos.

4.3 Detecção dos pontos

Inicialmente, para detectar os pontos cefalométricos, foram testadas diversas ferramentas disponíveis e reconhecidas na literatura, como: DLIB[46], CLandMark[47], CSIRO FaceAnalysis[48] e o Superviseddescent [31]. Porém, apesar desse reconhecimento, os resultados obtidos para a marcação dos pontos necessários para este trabalho não foram satisfatórios, sendo assim necessário a busca por uma nova abordagem.

Para a detecção dos pontos cefalométricos, foi utilizada uma versão modificada do algoritmo Superviseddescent proposto por [31], que foi detalhado na Seção 3. Nesta versão modificada, a entrada do algoritmo é alterada para receber 4 imagens I de entradas diferentes e o modelo pré-treinado $\Omega(\Theta)$ também é dividido em quatro diferentes, um para cada vetor de parâmetros Θ . Cada uma das imagens I corresponde à uma região diferente do rosto. Para obter essas imagens foi utilizado o *HaarCascade* disponível no OpenCV[49], desenvolvido por [16]. Sendo uma imagem da região dos olhos, uma da região do nariz, uma da região da boca e uma do rosto inteiro. Os vetores de parâmetros Θ estão diretamente associados a uma das imagens I , pois eles serão utilizados para localizar os pontos cefalométricos pertencentes a uma sub-região. Cada vetor Θ possui as

posições X e Y dos marcos faciais de determinada sub-região do rosto marcados por um especialista.

A face é dividida em 4 regiões de treinamento, cada um identificando pontos diversos, totalizando os 28 pontos descritos em [28] e descritos na Tabela 2.1 da Seção 2.3. O treinamento especializado para região dos olhos identifica os pontos Glabella, Nasion, Midnasal, Entocanthion, Exocanthion, Iridion laterale, Iridion mediale e Pupil. O treinamento para a região do nariz identifica os pontos Subnasale e Alare. O treinamento para a região da boca identifica os pontos Labiale superius, Stomion, Labiale inferius, Cheilion, Gonion, Gnathion e Crista philtri. E por último, a região da face como um todo, é utilizada para identificar os pontos Zygion, Gonion e Gnathion. A Figura 4.4 ilustra as 4 sub-regiões e quais os pontos utilizados em cada uma delas.

Para a realização do treinamento serão utilizadas 1000 imagens com as marcações manuais dos 28 pontos cefalométricos realizadas por especialistas, assegurando assim que os pontos marcados manualmente estão corretos. Foi utilizado um procedimento de validação cruzada, descrito em [50], separando aleatoriamente 700 imagens para o conjunto de treinamento e 300 imagens para realização dos testes de validação.

Ao final do treinamento, são gerados quatro arquivos com os dados treinados, um para cada região especializada. Com esses arquivos é possível obter os 28 pontos cefalométricos para uma nova imagem, como mostra a Figura 2.10. Assim, é possível comparar a marcação automática com a marcação manual feita por especialistas, e verificar a acurácia desta metodologia. Com os resultados obtidos, é possível a utilização deste método para a marcação em bases maiores, facilitando assim o restante da metodologia proposta neste trabalho. A acurácia é calculada em comparação com a marcação do especialista. Foi testado o método com uma margem de tolerância variando de 0 a 20 pixels. A Figura 4.5 ilustra a metodologia utilizada para avaliar a marcação automática. Para obter uma medida da eficácia do método, foi calculado o F-score para as 1000 imagens da base de dados.

Tendo esses pontos, pode-se calcular as 40 razões foto-antropométricas descritas na seção 2.3, na tabela 2.2. E com essas razões é possível realizar a comparação entre duas faces e determinar se pertencem ou não a um mesmo indivíduo. Os resultados obtidos para essa metodologia foi comparado com o desempenho das demais ferramentas disponíveis.

4.4 Reconhecimento Facial

Para fins de validação, todos os pontos dos dois primeiros cenários da Seção 5 foram marcados manualmente por peritos forenses. Nos demais cenários de teste, foram utilizados o algoritmo de detecção de pontos descritos na Seção 4.3.

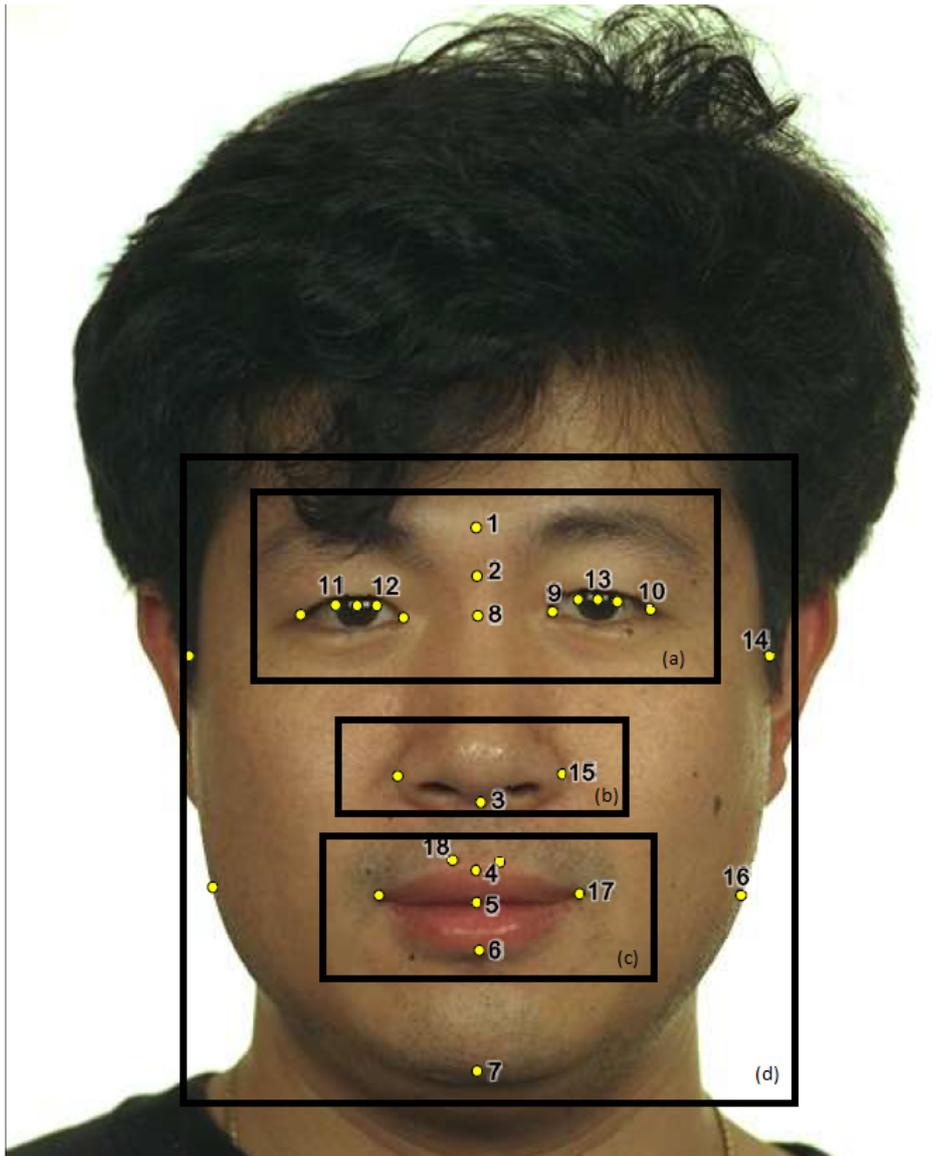


Figura 4.4: 4 sub-regiões e os pontos cefalométricos de cada uma delas, a numeração é referente à Tabela 2.1. (a) Região dos olhos; (b) Região do nariz; (c) Região da boca; e (d) Rosto inteiro. Imagem adaptada de [42].

Tendo todos os pontos marcados ou detectados, as razões foto-antropométricas (RFAs) listadas na tabela 2.2 são calculadas. A partir destas RFAs, é possível comparar uma imagem que queremos verificar com cada uma das faces da base de dados, obtendo assim a distância $D(X, Y)$, calculada como descrito na Equação 4.1.

$$D(X, Y) = \sum_{i=1}^z |(Y_i - X_i)|^k, \quad (4.1)$$

onde $D(X, Y)$ é a distância entre a imagem da base de dados X e o indivíduo com o qual queremos comparar Y , Y_i são os RFAs da imagem Y e X_i são os RFAs da imagem X . A

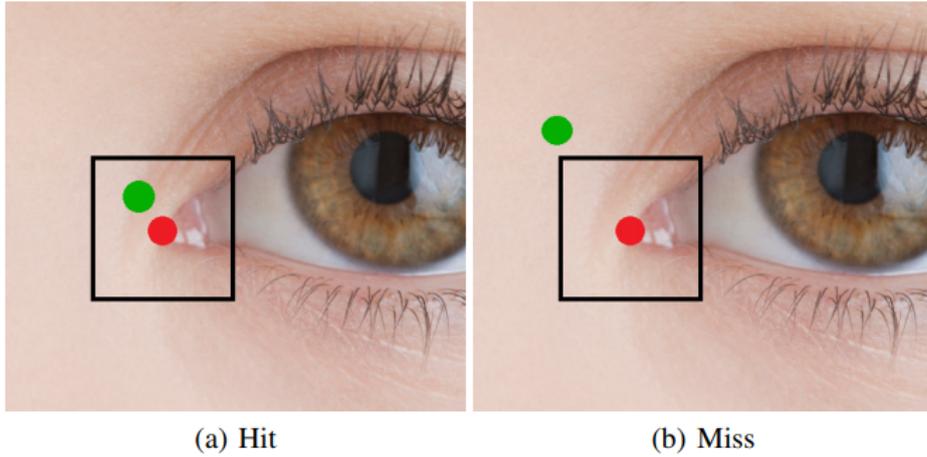


Figura 4.5: O círculo vermelho é a marcação do especialista, e também é o centro do quadrado preto que determina o limite (definido em píxeis) de onde é considerado um acerto na avaliação da marcação automática. (a) É considerado um acerto quando o círculo verde está dentro do limite estipulado. É considerado um erro se o círculo verde estiver fora do limite estipulado.

variável z é o número de razões utilizadas e k é a variável que permite a comparação utilizando Soma das Diferenças Absolutas (SAD), se k for igual a 1, e a Soma das Diferenças Quadradas (SSD), se k for igual a 2. Neste trabalho, para todos os testes realizados da Seção 5, k foi definido como sendo 2.

Sendo possível calcular a distância D para todas as imagens, foi considerado como acerto se a comparação entre as imagens X e Y seguirem a regra da Equação 4.2.

$$(Y, X) \in \begin{cases} \text{acerto,} & \text{Se } D(X, Y) \leq T \\ \text{erro,} & \text{Caso contrário,} \end{cases} \quad (4.2)$$

onde T é o limiar para determinar se as imagens representam a mesma pessoa, seguindo a metodologia de verificação descrita na Seção 2.1.1. O *False Acceptance Rate* (FAR) e o *False Rejection Rate* (FRR), citados na Seção 2.1.3, são as métricas de avaliação de biometria mais utilizadas e devem ser calculadas para cada uma das comparações. Utilizando essas métricas, é possível definir um valor para o limiar, baseado na *Equal Error Rate* (EER), para avaliar o desempenho da metodologia proposta, como descrito na Equação 4.3.

$$EER(T) \leftarrow FAR(T) = FRR(T). \quad (4.3)$$

Assi, utilizando esses valores é possível analisar graficamente os resultados e observar se a metodologia é capaz de realizar um reconhecimento facial biométrico.

4.5 Redução do Número de RFAs

Conseguir reduzir o número de RFAs é de extrema importância para a metodologia por dois motivos. Primeiro, a obtenção de 40 razões para cada imagem e a sua comparação com as demais pode tornar o processo lento. Em segundo, a escolha de um conjunto reduzido de RFAs, mas que sejam capaz de representar melhor uma pessoa, pode inclusive aumentar a acurácia do sistema. Com isso em mente, uma maneira de selecionar quais seriam as melhores RFAs para serem utilizadas em um sistema de reconhecimento facial biométrico ganha importância.

Para a metodologia proposta, foi utilizada uma base de dados com 40 imagens, sendo 5 de cada indivíduo. A repetição de imagens de uma mesma pessoa é necessária pois desejamos obter as razões que sejam mais similares para um mesmo indivíduo.

Foram realizadas duas abordagens para reduzir o número de RFAs. A primeira foi a utilização da correlação entre as RFAs. Inicialmente foi gerada a matriz de correlação e os coeficientes de correlação permitiram que as RFAs fossem ordenadas das menos correlacionadas para as mais correlacionadas. Assim, as últimas RFAs seriam retiradas do processo de reconhecimento, restando apenas as RFAs com menos correlação entre elas.

A segunda abordagem, é a proposta de uma nova solução para decidir qual RFA possui menos influência no processo de reconhecimento facial biométrico. Primeiro, é necessário detectar os pontos cefalométricos e calcular as RFAs das 40 imagens. Em seguida, para cada um dos indivíduos, é escolhida uma imagem de referência, essa imagem será utilizada para comparação com as demais imagens de um mesmo indivíduo, gerando um vetor V de 40 posições, onde serão armazenados o somatório das diferenças entre a imagem de referência e as demais imagens do mesmo indivíduo para cada uma das RFAs. Para a comparação será utilizada a Equação 4.4

$$V(z) = \sum_{i=2}^5 |(Y_{1z} - Y_{iz})| \quad (4.4)$$

onde $V(z)$ é o somatório das diferenças da RFA z para as imagem de referência Y_1 com as demais imagens Y_i do mesmo indivíduo. Dessa forma, em V será possível obter qual RFA possui a menor diferença para um mesmo indivíduo.

Após calcular os vetores V de cada um dos 8 indivíduos da base de dados, é possível calcular a média do valor do vetor V para cada uma das RFAs z , na forma da Equação 4.5, sendo armazenadas no vetor Med 4.4

$$Med(z) = \frac{\sum_{i=1}^8 V(z)_i}{8} \quad (4.5)$$

onde $Med(z)$ é a média dos valores do vetor V do indivíduo i para a RFA z . Dessa forma,

após serem calculados os valores para todos as 40 RFAs e armazenando-os em *Med*, é possível ordenar este vetor de médias em ordem crescente, obtendo-se assim uma ordem para os valores que possuem menor diferença dentro de um mesmo indivíduo. Com isso, é possível realizar os testes de redução do número de RFAs necessários, eliminando sempre as RFAs referentes aos últimos valores do vetor *Med*. Assim, testes foram realizados iniciando com 40 RFAs e reduzindo esse número a cada iteração, em seguida gráficos foram gerados para ilustrar a EER obtida para cada conjunto de RFAs. Ao final, os resultados obtidos com essa metodologia foram comparados com o obtido utilizando a correlação entre as RFAs para reduzir o número necessário.

No próximo capítulo serão apresentados os resultados obtidos para testar a metodologia apresentada.

Capítulo 5

Resultados

Neste capítulo serão apresentados os resultados obtidos nos testes da metodologia apresentada. Inicialmente serão demonstrados os resultados para o módulo de marcação automática dos pontos cefalométricos. Em seguida, serão mostrados os resultados para a metodologia de reconhecimento facial.

Para a validação da metodologia de reconhecimento facial, foram realizados testes em quatro cenários distintos. São eles:

- Primeiro cenário: base de 1000 imagens, sem repetição dos indivíduos.
- Segundo cenário: base com 40 imagens de 8 indivíduos distintos, totalizando 5 imagens por pessoa.
- Terceiro cenário: base com 6622 imagens, sendo duas de cada indivíduo e diferença de no mínimo 5 anos entre as imagens.
- Quarto cenário: utilização do método para redução do número de RFAs utilizadas na mesma base de dados do cenário 3.

Todos os cenários contêm uma visão frontal das faces, sem óculos e boa luminosidade, como descritos na Seção 4. Ainda, em todos eles, cada imagem foi comparada com todas as imagens, incluindo ela mesma, e o valor do limiar (T) iniciando em 0 e sendo incrementando em 0.05 para cada estágio de iteração completado. Esse procedimento foi replicado para um número de RFAs variando de 40 até 10. Para os 3 primeiros cenários, a redução do número de RFAs foi definida pelo resultado de uma análise de correlação nas 1000 imagens da base de dados utilizadas no primeiro cenário. As RFAs foram ordenadas de forma decrescente dos coeficientes de correlação, e as RFAs com menores coeficientes foram descartadas a cada iteração.

5.1 Detecção automática dos pontos cefalométricos

Para testar o método, foi utilizada uma base de dados contendo 1000 imagens e as marcações dos pontos cefalométricos feitas por especialistas. Assim, foi calculado o F-score para cada variação de pixel da janela de tolerância. A Figura 5.1 mostra os resultados obtidos para cada uma das variações de pixel.

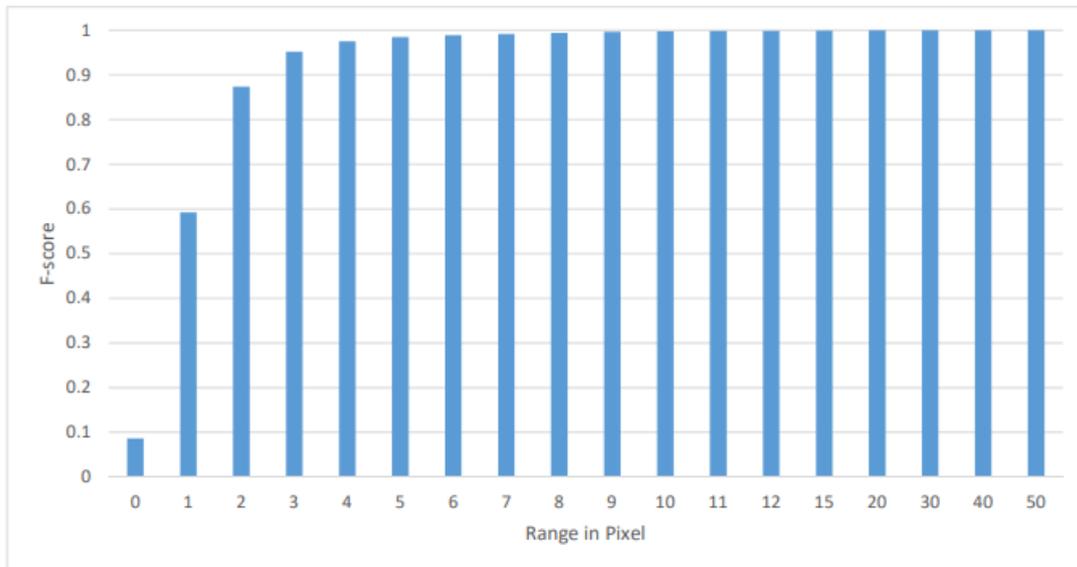


Figura 5.1: Gráfico com os F-scores, variando o pixel da janela de tolerância de 0 a 20.

O método proposto foi comparado com outras ferramentas para detecção de marcos faciais disponíveis e reconhecidas na literatura, como: DLIB[46], CLandMark[47], CSIRO FaceAnalysis[48] e o Superviseddescent [31]. Todos os testes foram realizados nas imagens da base de dados contendo 1000 imagens de indivíduos distintos e com as posições dos marcos faciais marcadas manualmente para efeito de comparação. Além disso, foram utilizados 5 pixels de tolerância para todas as ferramentas. Os resultados dessa comparação podem ser observados na Figura 5.2.

Assim, é possível observar que acima de 5 pixels de tolerância, o método é extremamente eficaz, e pode ser utilizado para o prosseguimento deste trabalho.

5.2 Reconhecimento Facial

Tendo a marcação de pontos obtido resultado satisfatório, foi possível realizar os testes da metodologia de reconhecimento facial biométrico. No primeiro cenário de testes, é desejado observar se as RFAs são capazes de individualizar uma pessoa, não havendo combinação de RFAs que identifiquem duas pessoas como sendo a mesma. No segundo

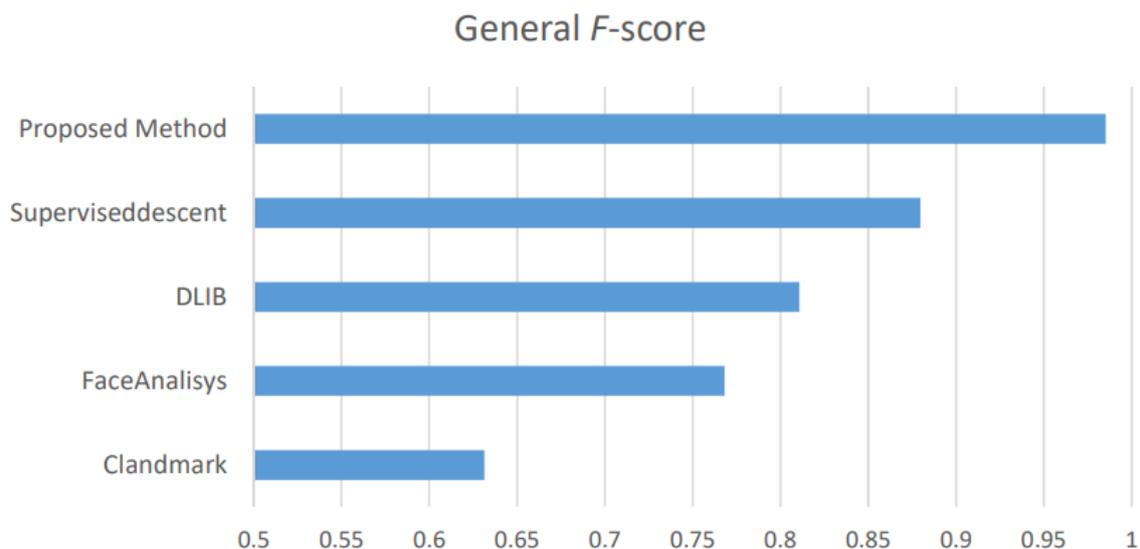


Figura 5.2: Gráfico com o F-score do método proposto, utilizando 5 píxeis de tolerância, comparado com os resultados de outras ferramentas reconhecidas na literatura.

Tabela 5.1: Ordem das RFAs gerada a partir da correlação.

RFAs ordenadas pela correlação				
1. 29	2. 21	3. 26	4. 25	5. 15
6. 5	7. 40	8. 37	9. 4	10. 31
11. 39	12. 34	13. 32	14. 18	15. 30
16. 9	17. 8	18. 28	19. 24	20. 17
21. 6	22. 14	23. 12	24. 23	25. 1
26. 3	27. 2	28. 20	29. 35	30. 38
31. 22	32. 19	33. 7	34. 33	35. 13
36. 16	37. 36	38. 11	39. 10	40. 27

cenário, é observada a capacidade da metodologia de realizar reconhecimento facial biométrico para imagens com variações de escala. No terceiro cenário, o método é avaliado no reconhecimento facial biométrico com 5 anos de diferença entre as imagens de um mesmo indivíduo. No quarto e último cenário, a metodologia para reduzir o número de RFAs necessários é testada.

Nos três primeiros cenários, foi utilizado a correlação entre as RFAs para poder reduzir o número necessário para realizar o reconhecimento facial. Assim, as RFAs que possuem a maior correlação com as demais seriam retiradas primeira, deixando apenas as menos correlatas para o restante do processo. A tabela 5.1 mostra quais são as RFAs utilizadas, onde a numeração é referente à Tabela 2.2.

5.2.1 Primeiro cenário

No primeiro cenário foi utilizada uma base de dados com 1000 imagens, sem repetição dos indivíduos. Neste cenário, é desejado saber se os pontos e razões utilizados são capazes de individualizar uma pessoa ou se ocorreriam muitas redundâncias. Com a aplicação da metodologia, foi possível obter os resultados que podem ser visualizados nos gráficos das Figuras 5.3, 5.4 e 5.5.

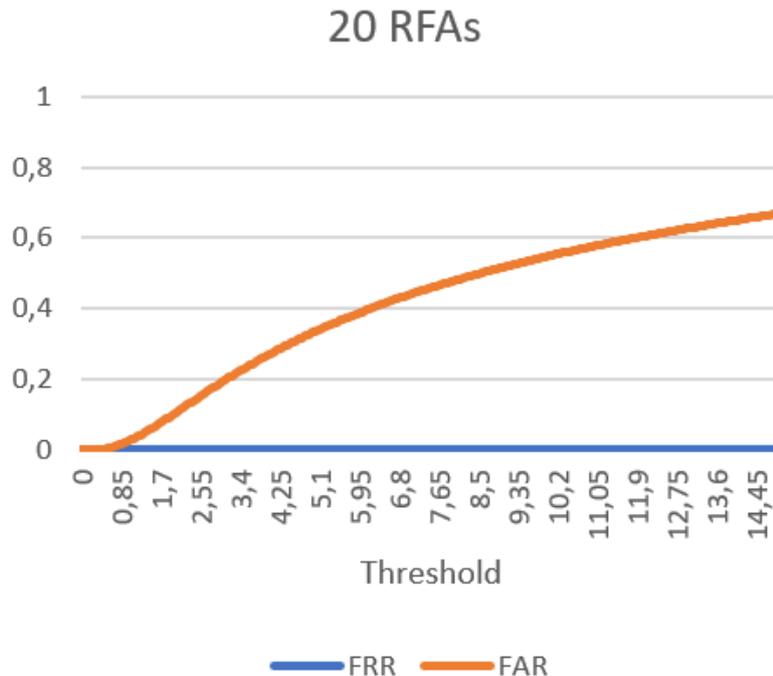


Figura 5.3: Resultados para o primeiro cenário com 20 RFAs.

5.2.2 Segundo cenário

No segundo cenário foi utilizada uma base de dados com 40 imagens de 8 indivíduos distintos, totalizando 5 imagens por pessoa. Neste cenário, as imagens foram obtidas em um ambiente controlado, com as 5 imagens de cada pessoa sendo obtidas sem um intervalo de tempo, resultado em uma pequena variação entre cada imagem, como mostra a figura 4.2. A motivação para este cenário é verificar se a metodologia é capaz de realizar reconhecimento facial biométrico, distinguindo indivíduos e em imagens com alteração de escala. Com a aplicação da metodologia, foi possível obter os resultados que podem ser visualizados nos gráficos das Figuras 5.6, 5.7 e 5.8.

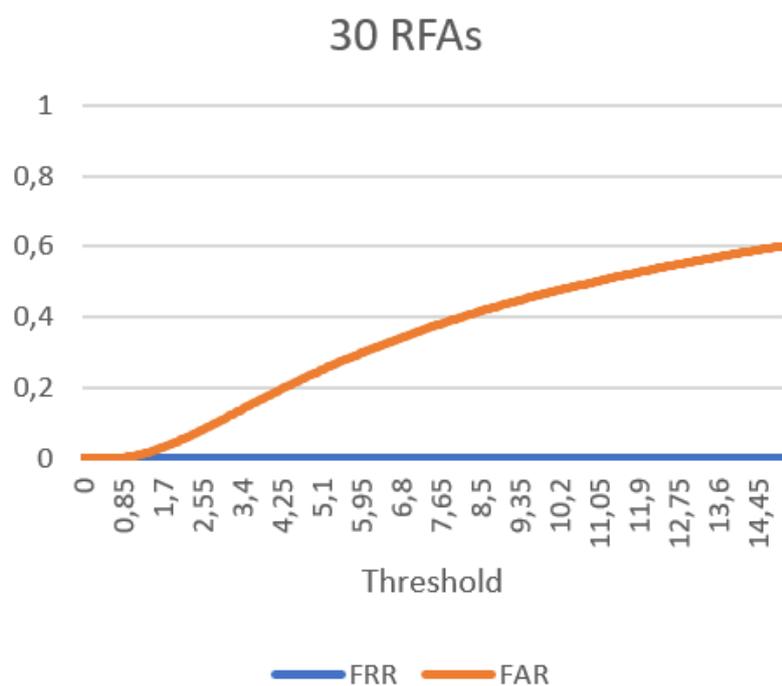


Figura 5.4: Resultados para o primeiro cenário com 30 RFAs.

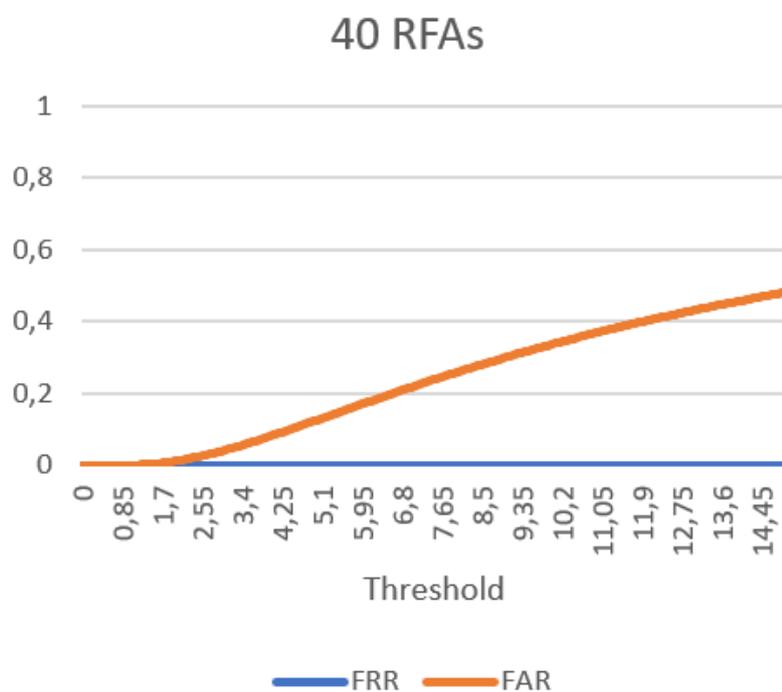


Figura 5.5: Resultados para o primeiro cenário com 40 RFAs.

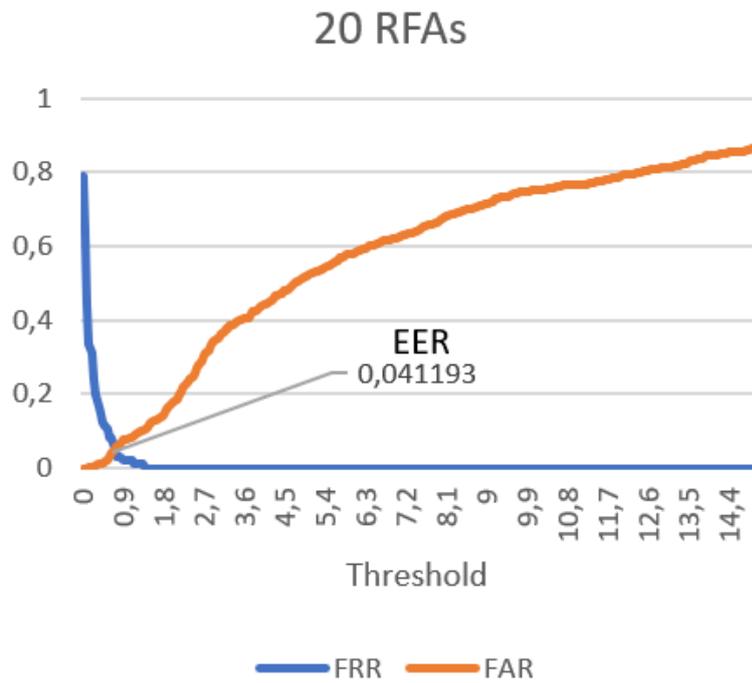


Figura 5.6: Resultados para o segundo cenário com 20 RFAs.

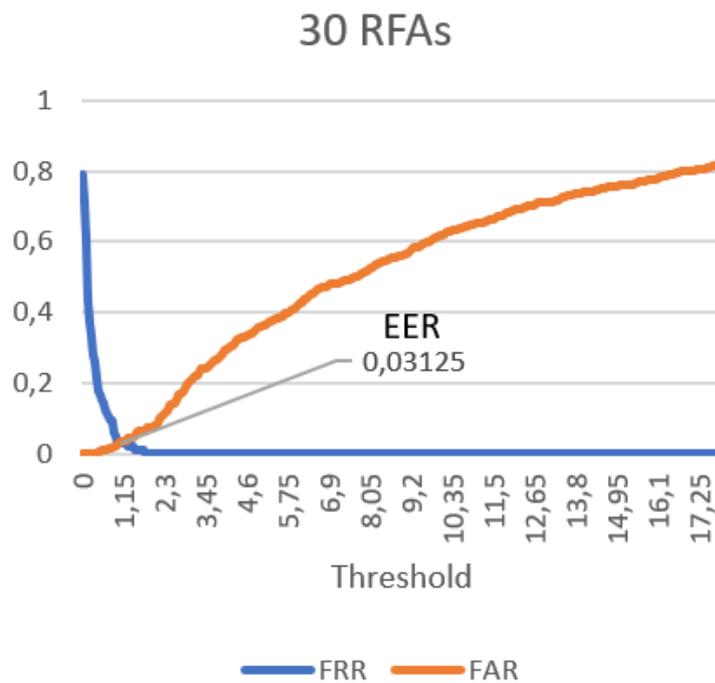


Figura 5.7: Resultados para o segundo cenário com 30 RFAs.

5.2.3 Terceiro cenário

No terceiro cenário foi utilizada uma base de dados contendo 6622 imagens, sendo 2 de cada indivíduo, e um mínimo de 5 anos de intervalo entre as duas fotos. Além disso, foi

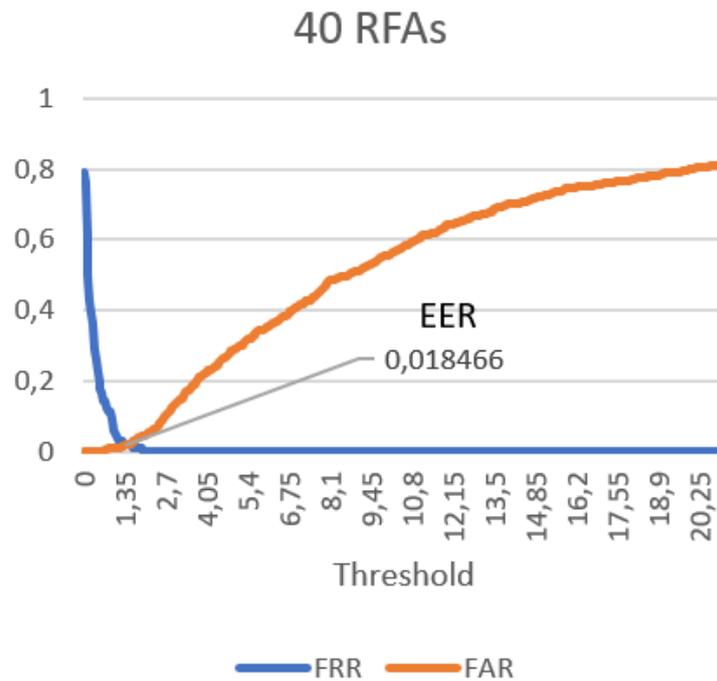


Figura 5.8: Resultados para o segundo cenário com 40 RFAs.

utilizado um ambiente menos controlado, havendo assim imagens com diferentes expressões faciais, como sorrindo ou com a boca aberta. A motivação para este teste é saber como a metodologia se comportaria após 5 anos de crescimento e mudanças faciais. Com a aplicação da metodologia, foi possível obter os resultados que podem ser visualizados nos gráficos das Figuras 5.9 e 5.10.

5.2.4 Quarto cenário

No último cenário, foi utilizado a base de dados com 40 imagens com 8 indivíduos, sendo 5 de cada indivíduo, para ordenar as RFAs mais influentes no processo de reconhecimento. As médias ordenadas e suas respectivas RFAs, com numeração referente à Tabela 2.2, podem ser vistas na Tabela 5.2.

Com isso, foi possível obter os resultados apresentados nas Figuras 5.11, 5.12, 5.13 e 5.14.

5.3 Discussão

No primeiro cenário, como mostra os gráficos nas Figuras 5.3, 5.4 e 5.5, é possível notar que a FRR é sempre zero. Este resultado era esperado, visto que há apenas uma imagem de

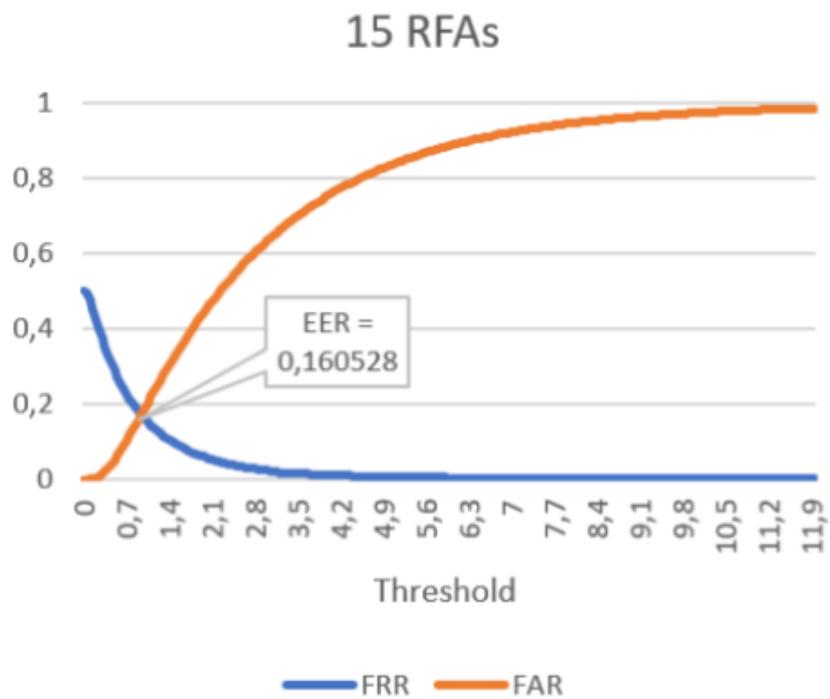


Figura 5.9: Resultados para o terceiro cenário com 15 RFAs.

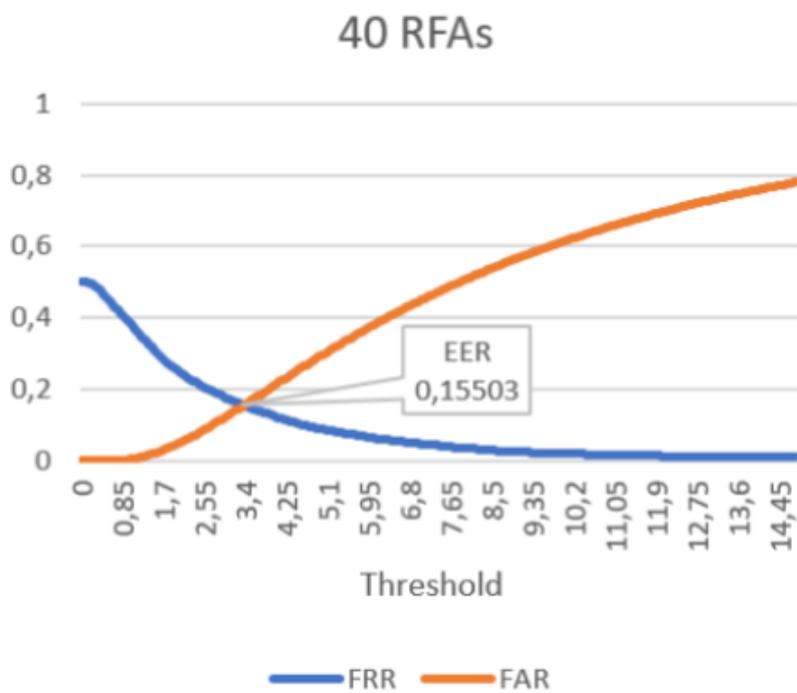


Figura 5.10: Resultados para o terceiro cenário com 40 RFAs.

Tabela 5.2: Ordem das RFAs gerada a partir do método proposto na Seção 4.5, com as médias obtidas para cada uma delas.

RFA	Média	RFA	Média	RFA	Média	RFA	Média
38	0.0750	26	0.1518	7	0.3626	20	0.4232
30	0.0794	18	0.1790	13	0.3756	23	0.4247
36	0.1010	19	0.1837	34	0.3832	33	0.4275
40	0.1081	1	0.1966	17	0.3905	14	0.4293
12	0.1162	3	0.1966	24	0.4041	29	0.4349
37	0.1186	2	0.1966	4	0.4046	5	0.4354
10	0.1242	35	0.2309	6	0.4102	9	0.4418
32	0.1246	8	0.2777	15	0.4128	31	0.45173
39	0.1425	27	0.3326	16	0.4168	28	0.4812
25	0.1518	21	0.3522	22	0.4209	11	0.7234

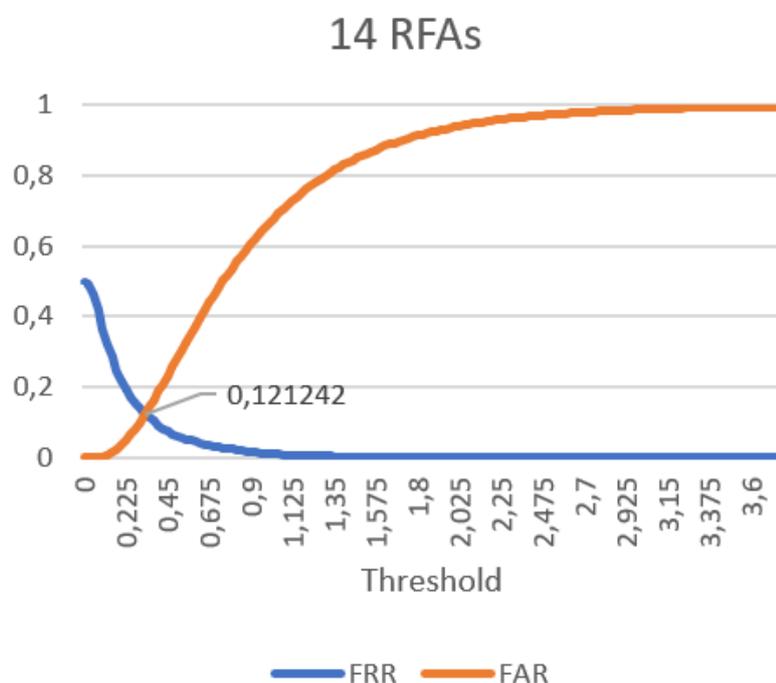


Figura 5.11: Resultados para o quarto cenário com 14 RFAs.

cada indivíduo e o resultado da comparação da imagem com ela mesmo sempre resultará em um acerto (distância entre a imagem e ela mesmo é sempre zero). Para a curva da FAR, é interessante reparar que a curva permanece zero ou próximo de zero para os menores limiares, mostrando que há uma margem para a qual é possível haver uma alta acurácia com uma baixa FAR. Este cenário também nos mostra que a metodologia proposta permite a individualização das pessoas, caso contrário, poderiam haver duas pessoas distintas com distâncias próximas de zero, o que não ocorreu nos testes.

No segundo cenário, como ilustrado pelos gráficos das Figuras 5.6, 5.7 e 5.8, é possível

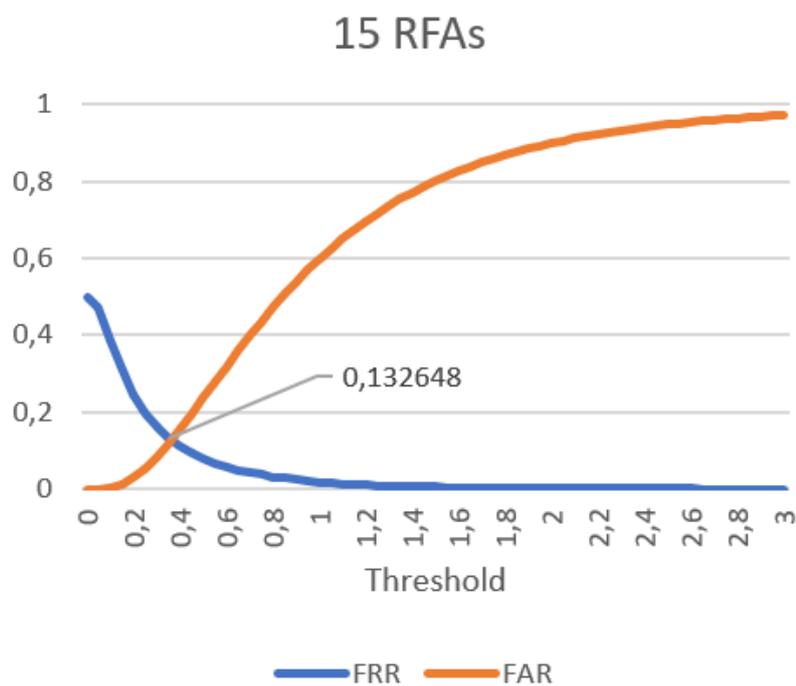


Figura 5.12: Resultados para o quarto cenário com 15 RFAs.

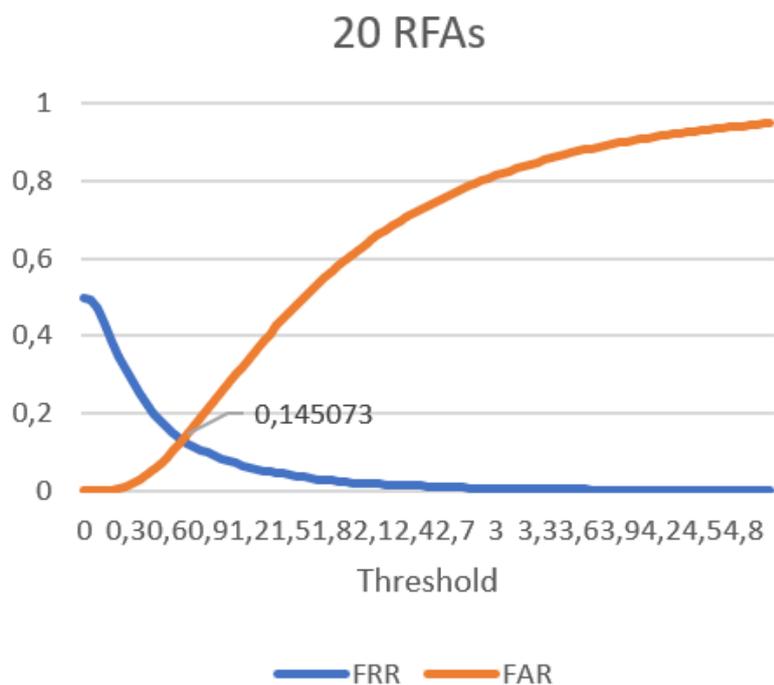


Figura 5.13: Resultados para o quarto cenário com 20 RFAs.

observar que a FRR cai rapidamente quando o valor do limiar aumenta, e isso ocorre para todas as iterações de RFAs. Também é possível notar que há um aumento na EER quando

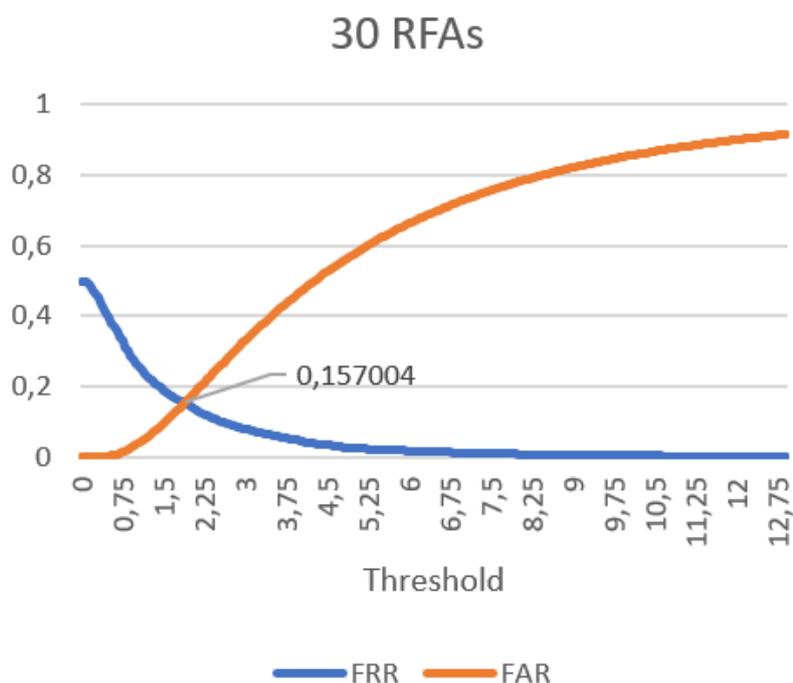


Figura 5.14: Resultados para o quarto cenário com 30 RFAs.

se diminui o número de RFAs, o que é esperado de um sistema biométrico. Neste cenário, assim como nos demais, a FRR nunca inicia em 1, isso ocorre pois as imagens sempre são comparadas com elas mesmas, resultando sempre em uma distância de 0, sendo assim considerada como acerto para qualquer valor de limiar utilizado. Vale notar ainda que a distância entre duas imagens, mesmo que seja de um mesmo indivíduo, nunca é zero, resultando em valores zero para a FAR nos limiares mais baixos. Nestes testes, a EER, para os dois conjuntos de RFAs aqui ilustrados, permanecem com menos de 5% entre eles. Esse resultado mostra que para essa base de dados reduzida, foi possível realizar reconhecimento facial biométrico mesmo em situações em que as imagens possuíam escalas diferentes, atingindo assim um dos objetivos propostos neste trabalho.

Em seguida, no terceiro cenário, presente nos gráficos das Figuras 5.9 e 5.9, a EER sobe de 15.5% para 16% para os casos de 40 RFAs e 15 RFAs, em uma base contendo 6622 imagens. Este resultado já era esperado devido ao fato de se aumentar bruscamente o número de comparações realizadas. Além disso, neste cenário, temos condições menos favoráveis para o reconhecimento, como é o caso de pessoas sorrindo nas imagens e menos controle no posicionamento do rosto, o que dificulta a precisão das RFAs que se encontram abaixo do ponto subnasale. Outro fator que prejudica a acurácia do método, mas não a compromete, é a diferença de idade entre as imagens. Mesmo assim, notamos que a diferença entre a EER nos testes para 40 e 15 RFAs ainda é pequena, havendo apenas

uma variação de 0.5%. Com este teste, podemos concluir que o método pode ser utilizado para realizar um reconhecimento facial biométrico com até 5 anos de diferença entre as imagens, cumprindo assim outro dos objetivos propostos para este trabalho.

No quarto é o último cenário, ilustrado pelos gráficos das Figuras 5.11, 5.12, 5.13 e 5.14, o resultado do método para reduzir o número de RFAs se mostrou muito positivo. É possível observar que a EER diminuiu com um menor número de RFAs utilizadas, se comparado com o teste do terceiro cenário. Isso ocorre pois com o passar do tempo, algumas RFAs menos estáveis podem possuir uma maior variação em relação às outras, o que prejudica o reconhecimento ao invés de facilitá-lo. Com a mudança da ordem de quais RFAs seriam excluídas a cada iteração, a melhor EER foi obtida quando foram utilizadas apenas as 14 RFAs melhores ranqueadas, obtendo-se 12% de EER. Se compararmos os resultados para 15 RFAs dos gráficos das figuras do quarto e do terceiro cenários podemos observar uma melhora de quase 3%, diminuindo a EER de 16% para 13.2%. Com esse resultado, é possível concluir que apesar de poderem ser utilizadas para determinar o sexo e idade das pessoas como mostrado em [9], algumas das RFAs utilizadas em seu trabalho podem atrapalhar no processo de reconhecimento facial biométrico.

Outra forma de comparar o desempenho da metodologia de redução de RFAs é através da curva de Característica de Operação do Receptor (curva ROC), que mede a sensibilidade e a especificidade de um algoritmo de classificação. A sensibilidade é medida pela taxa de verdadeiros positivos (*true positives rate*)(TPR) e mede quantos indivíduos foram classificados corretamente como *match*, já a especificidade pode ser medida através da taxa de verdadeiros negativos (*true negative rate*)(TNR) e mede a capacidade do sistema em prever corretamente a ausência da condição para os que realmente não a têm, que também pode ser expresso pela taxa de falso positivo (*false accept rate*)(FAR), onde $FAR = 1 - TNR$. No gráfico ROC, é feita a relação entre a TPR com a FAR. O Gráfico 5.15 mostra as curvas obtidas para testes realizados com 15, 20 e 30 RFAs, para a nova metodologia de redução de RFAs e para a redução por meio de correlação. Nota-se que há uma grande diferença entre as curvas da nova metodologia de redução de RFAs e das curvas de redução por correlação. Porém, as curvas da mesma metodologia são próximas para 15 e 20 RFAs. Para o caso de 30 RFAs, as curvas das duas metodologias se aproximam.

Na nova metodologia, com um número reduzido de RFAs, as curvas são mais acentuadas no início se comparadas com a metodologia por correlação, significando que a TPR possui um crescimento mais acelerado em relação à FAR, obtendo-se assim uma área sob a curva maior na nova metodologia para qualquer limiar escolhido. É possível notar que quanto maior o número de RFAs utilizados mais as duas metodologias se aproximam, isso se dá pelo fato de que mais RFAs semelhantes são utilizadas nos dois métodos. As áreas

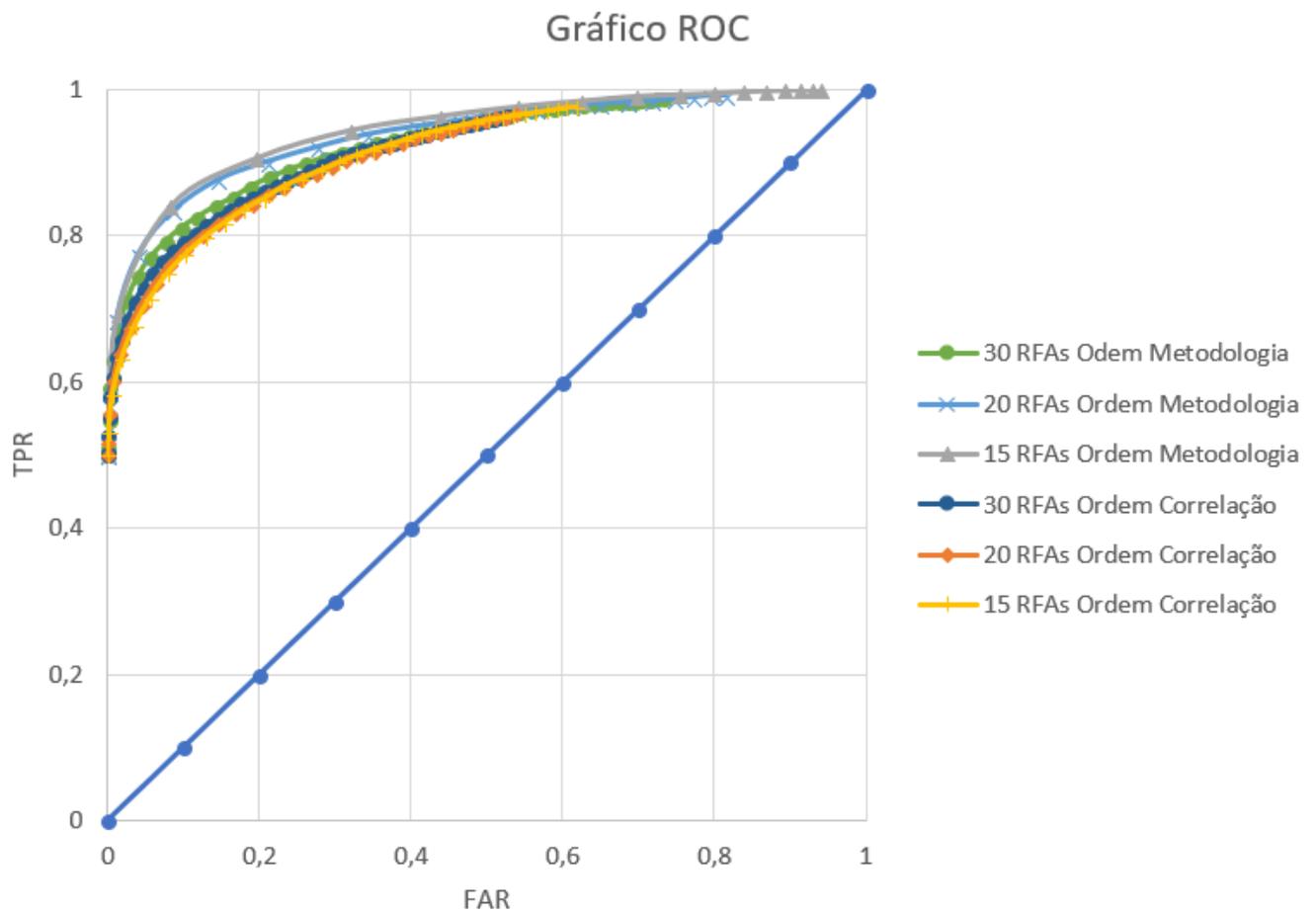


Figura 5.15: Gráfico ROC mostrando a comparação entre a metodologia de escolha de RFAs e a escolha por correlação. A imagem mostra as curvas para testes com 15, 20 e 30 RFAs.

sob a curva aproximadas obtidas podem ser vistas na Tabela 5.3.

Dessa forma é possível afirmar que para qualquer situação, a metodologia de redução por correlação é superada pela novo método para reduzir o número de RFAs necessários.

Tabela 5.3: Área sob a curva para o gráfico da Figura 5.15.

Metodologia	Área sob a curva
30 RFAs Nova Metodologia	0.87718
30 RFAs Correlação	0.87358
20 RFAs Nova Metodologia	0.88109
20 RFAs Correlação	0.86671
15 RFAs Nova Metodologia	0.89455
15 RFAs Correlação	0.86843

No próximo capítulo, serão apresentadas as conclusões do trabalho em relação aos objetivos propostos, e serão apresentados os trabalhos futuros para esta pesquisa.

Capítulo 6

Conclusão

Neste capítulo será realizada uma análise dos resultados obtidos nos testes apresentados no capítulo 5 e também serão apresentadas propostas de trabalhos futuros.

Neste trabalho foi apresentada uma metodologia para reconhecimento facial biométrico, que sofre baixa influência de fatores como escala e o crescimento do rosto, utilizando a íris como fator para normalização dos pontos cefalométricos. Os resultados apresentados neste trabalho permitem confirmar que os pontos cefalométricos podem ser utilizados em sistemas de reconhecimento facial biométrico. A metodologia proposta foi testada em 4 cenários diferentes, cada um com o seu propósito.

A partir dos resultados apresentados na Seção 5, é possível concluir que os objetivos propostos para este trabalho foram atingidos, e que a metodologia aqui proposta mostra uma nova e promissora abordagem utilizando as RFAs para o reconhecimento facial biométrico, o que permite preencher uma lacuna onde diversos sistemas tradicionais falham, como imagens com diferença de idade de um mesmo indivíduo e o problema de escala. Além disso, também foi possível reduzir o número de RFAs necessários para se obter um bom resultado para apenas 14, criando assim um método eficaz que utiliza menos razões que outros métodos tradicionais como proposto por Kanade em [8] ou em [23].

Como trabalhos futuros relacionados, temos a utilização de aprendizado de máquina que permita localizar e estimar os pontos cefalométricos mesmo quando não estejam completamente visíveis, e submeter o sistema para o FRVT[51] para que seja possível uma comparação com demais sistemas que se propõem a resolver o mesmo problema com técnicas baseadas em *landmarks*. Ainda, é possível a utilização de aprendizado de máquina nas RFAs para verificar se o desempenho por esse método supera a metodologia proposta neste trabalho.

6.1 Publicação Realizada

A metodologia de reconhecimento facial e os resultados apresentados nas Seções 5.2.1, 5.2.2 e 5.2.3 deram origem a um artigo que foi publicado no 6th *IAPR/IEEE International Workshop on Biometrics and Forensics* (IWBF 2018), que ocorreu nos dias 7 e 8 de Junho de 2018. Essa conferência possui qualis CAPES B3 em Ciência da Computação no ano 2013-2016.

Referências

- [1] “Reconhecimento facial passa a ser usado em 14 aeroportos,” *Jornal Nacional*. [Online]. Available: <http://g1.globo.com/jornal-nacional/noticia/2016/07/reconhecimento-facial-passa-ser-usado-em-14-aeroportos.html> 1
- [2] A. Azeem, M. Sharif, M. Raza, and M. Murtaza, “A survey: Face recognition techniques under partial occlusion.” *Int. Arab J. Inf. Technol.*, vol. 11, no. 1, pp. 1–10, 2014. 1, 2
- [3] L. R. Costa, R. R. Obelheiro, and J. S. Fraga, “Introdução á biometria,” *Livro texto dos Minicursos do VI Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais (SBSeg2006)*. SBC: Porto Alegre, vol. 1, pp. 103–151, 2006. 1, 7, 8, 9
- [4] A. K. Jain and S. Z. Li, *Handbook of face recognition*. Springer, 2011. 1, 10, 11
- [5] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, “Face recognition: A literature survey,” *ACM computing surveys (CSUR)*, vol. 35, no. 4, pp. 399–458, 2003. 1
- [6] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, “Deepface: Closing the gap to human-level performance in face verification,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1701–1708. 1, 11, 24, 25
- [7] F. Schroff, D. Kalenichenko, and J. Philbin, “Facenet: A unified embedding for face recognition and clustering,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 815–823. 1, 11, 24, 25, 26
- [8] T. Kanade, “Picture processing system by computer complex and recognition of human faces,” in *Doctoral dissertation, Kyoto University*, November 1973. 1, 9, 13, 14, 53
- [9] C. E. P. Machado, M. R. P. Flores, L. N. C. Lima, R. L. R. Tinoco, A. Franco, A. C. B. Bezerra, M. P. Evison, and M. A. Guimarães, “A new approach for the analysis of facial growth and age estimation: Iris ratio,” *PLOS ONE*, vol. 12, no. 7, pp. 1–19, 07 2017. [Online]. Available: <https://doi.org/10.1371/journal.pone.0180330> 2, 3, 17, 19, 50
- [10] “China ‘racing ahead’ in facial recognition investment,” *Biometric Technology Today*, vol. 2018, no. 3, p. 1, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0969476518300286> 2

- [11] R. Jafri and H. R. Arabnia, "A survey of face recognition techniques." *Jips*, vol. 5, no. 2, pp. 41–68, 2009. 2, 10
- [12] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on circuits and systems for video technology*, vol. 14, no. 1, pp. 4–20, 2004. 4, 5
- [13] R. Clarke, "Human identification in information systems: Management challenges and public policy issues," *Information Technology & People*, vol. 7, no. 4, pp. 6–37, 1994. 7
- [14] J. L. Wayman, "Error rate equations for the general biometric system," *IEEE Robotics & Automation Magazine*, vol. 6, no. 1, pp. 35–48, 1999. 8
- [15] T. S. Jebara, "3d pose estimation and normalization for face recognition," Ph.D. dissertation, McGill University, 1995. 10
- [16] P. Viola and M. J. Jones, "Robust real-time face detection," *International Journal of Computer Vision*, vol. 57, no. 2, pp. 137–154, 2004. [Online]. Available: <http://dx.doi.org/10.1023/B:VISI.0000013087.49260.fb> 10, 33
- [17] H. Hotelling, "Analysis of a complex of statistical variables into principal components." *Journal of educational psychology*, vol. 24, no. 6, p. 417, 1933. 11
- [18] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," *Josa a*, vol. 4, no. 3, pp. 519–524, 1987. 11
- [19] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of cognitive neuroscience*, vol. 3, no. 1, pp. 71–86, 1991. 11
- [20] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 19, no. 7, pp. 711–720, 1997. 11
- [21] B. Amos, B. Ludwiczuk, and M. Satyanarayanan, "Openface: A general-purpose face recognition library with mobile applications," Technical report, CMU-CS-16-118, CMU School of Computer Science, Tech. Rep., 2016. 11
- [22] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments," Technical Report 07-49, University of Massachusetts, Amherst, Tech. Rep., 2007. 11, 12, 25, 26, 27
- [23] I. J. Cox, J. Ghosn, and P. N. Yianilos, "Feature-based face recognition using mixture-distance," in *Computer Vision and Pattern Recognition, 1996. Proceedings CVPR'96, 1996 IEEE Computer Society Conference on*. IEEE, 1996, pp. 209–216. 14, 15, 53
- [24] K. Ramesha, K. Raja, K. Venugopal, and L. Patnaik, "Feature extraction based face recognition, gender and age classification," 2010. 14, 16, 17

- [25] J. Shi, A. Samal, and D. Marx, “How effective are landmarks and their geometry for face recognition?” *Computer Vision and Image Understanding*, vol. 102, no. 2, pp. 117–133, 2006. 16
- [26] S. N. Byers, *Introduction to forensic anthropology*. Taylor & Francis, 2016. 16
- [27] D. Ranly, “A synopsis of craniofacial growth. 2nd.” 16, 19, 30
- [28] C. E. P. Machado, “Fotoantropometria para estimativa de idade de crianças e adolescentes com emprego de imagens faciais em norma frontal: relações iridianas,” Ph.D. dissertation, Departamento de Patologia e Medicina Legal da Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, 2015. 18, 19, 20, 30, 34
- [29] R. C. Augusteyn, D. Nankivil, A. Mohamed, B. Maceo, F. Pierre, and J.-M. Parel, “Human ocular biometry,” *Experimental eye research*, vol. 102, pp. 70–75, 2012. 20
- [30] C. Sagonas, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, “300 faces in-the-wild challenge: The first facial landmark localization challenge,” in *Computer Vision Workshops (ICCVW), 2013 IEEE International Conference on*. IEEE, 2013, pp. 397–403. 22, 23
- [31] P. Huber, Z.-H. Feng, W. Christmas, J. Kittler, and M. Rätzsch, “Fitting 3d morphable face models using local features,” in *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 1195–1199. 22, 23, 24, 33, 40
- [32] Z. Yu, F. Liu, R. Liao, Y. Wang, H. Feng, and X. Zhu, “Improvement of face recognition algorithm based on neural network,” in *Measuring Technology and Mechatronics Automation (ICMTMA), 2018 10th International Conference on*. IEEE, 2018, pp. 229–234. 24
- [33] Y. Li, W. Zheng, Z. Cui, and T. Zhang, “Face recognition based on recurrent regression neural network,” *Neurocomputing*, vol. 297, pp. 50–58, 2018. 24, 27
- [34] M. A. Abuzneid and A. Mahmood, “Enhanced human face recognition using lbph descriptor, multi-knn, and back-propagation neural network,” *IEEE Access*, vol. 6, pp. 20 641–20 651, 2018. 24, 26
- [35] G. Wen, H. Chen, D. Cai, and X. He, “Improving face recognition with domain adaptation,” *Neurocomputing*, vol. 287, pp. 45–51, 2018. 26
- [36] K. Arya, A. Rajawat, M. K. Pandey, and S. S. Rajput, “Very low resolution face recognition using fused visual and texture features,” in *Information and Communication Technology (CICT), 2017 Conference on*. IEEE, 2017, pp. 1–5. 27
- [37] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker, “Multi-pie,” *Image and Vision Computing*, vol. 28, no. 5, pp. 807–813, 2010. 27
- [38] M. Kim, S. Kumar, V. Pavlovic, and H. Rowley, “Face tracking and recognition with visual constraints in real-world videos,” in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8. 27

- [39] B. F. Klare and A. K. Jain, “Heterogeneous face recognition using kernel prototype similarities,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 35, no. 6, pp. 1410–1422, 2013. 28
- [40] Y. Somada, W. Ohyama, and T. Wakabayashi, “Segmented face image verification for age-invariant face recognition,” in *Informatics, Electronics and Vision & 2017 7th International Symposium in Computational Medical and Health Technology (ICIEV-ISCMT), 2017 6th International Conference on*. IEEE, 2017, pp. 1–4. 28, 29
- [41] S. Biswas and J. Sil, “An efficient face recognition method using contourlet and curvelet transform,” *Journal of King Saud University-Computer and Information Sciences*, 2017. 29
- [42] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss, “The feret evaluation methodology for face-recognition algorithms,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 10, pp. 1090–1104, 2000. 29, 35
- [43] Y. Zhu, C. Zhu, and X. Li, “Improved principal component analysis and linear regression classification for face recognition,” *Signal Processing*, vol. 145, pp. 175–182, 2018. 29
- [44] G. Bradski, *Dr. Dobb’s Journal of Software Tools*. 31
- [45] B. D. I. Formats-Part, “5: Face image data,” *ISO/IEC JTC1/SC37 N506, ISO/IEC IS 19794*, vol. 5, 2004. 31
- [46] D. E. King, “Dlib-ml: A machine learning toolkit,” *Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009. 33, 40
- [47] M. Uříčář, V. Franc, D. Thomas, S. Akihiro, and V. Hlaváč, “Real-time multi-view facial landmark detector learned by the structured output svm,” in *BWILD 2015: Proceedings of the 11th IEEE International Conference on Automatic Face and Gesture Recognition Conference and Workshops*, 2015. 33, 40
- [48] M. Cox, J. Nuevo-Chiquero, J. Saragih, and S. Lucey, “Csiro face analysis sdk,” *Brisbane, Australia*, 2013. 33, 40
- [49] G. Bradski *et al.*, “The opencv library,” *Doctor Dobbs Journal*, vol. 25, no. 11, pp. 120–126, 2000. 33
- [50] J. Friedman, T. Hastie, and R. Tibshirani, *The elements of statistical learning*. Springer series in statistics New York, 2001, vol. 1. 34
- [51] P. Grother, M. Ngan, I. T. L. N. I. of Standards, and T. I. A. Division, *Face Recognition Vendor Test (FRVT)*, 2014. [Online]. Available: <https://books.google.com.br/books?id=ZU3VAQAACAAJ> 53