

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
COORDENAÇÃO DE CIÊNCIA DA COMPUTAÇÃO

VAMBASTER JOSÉ DA COSTA

RECONHECIMENTO DE PADRÕES FACIAIS: UMA SÍNTESE

TRABALHO DE CONCLUSÃO DE CURSO

Santa Helena, Paraná

2019

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
COORDENAÇÃO DE CIÊNCIA DA COMPUTAÇÃO**

VAMBASTER JOSÉ DA COSTA

RECONHECIMENTO DE PADRÕES FACIAIS: UMA SÍNTESE

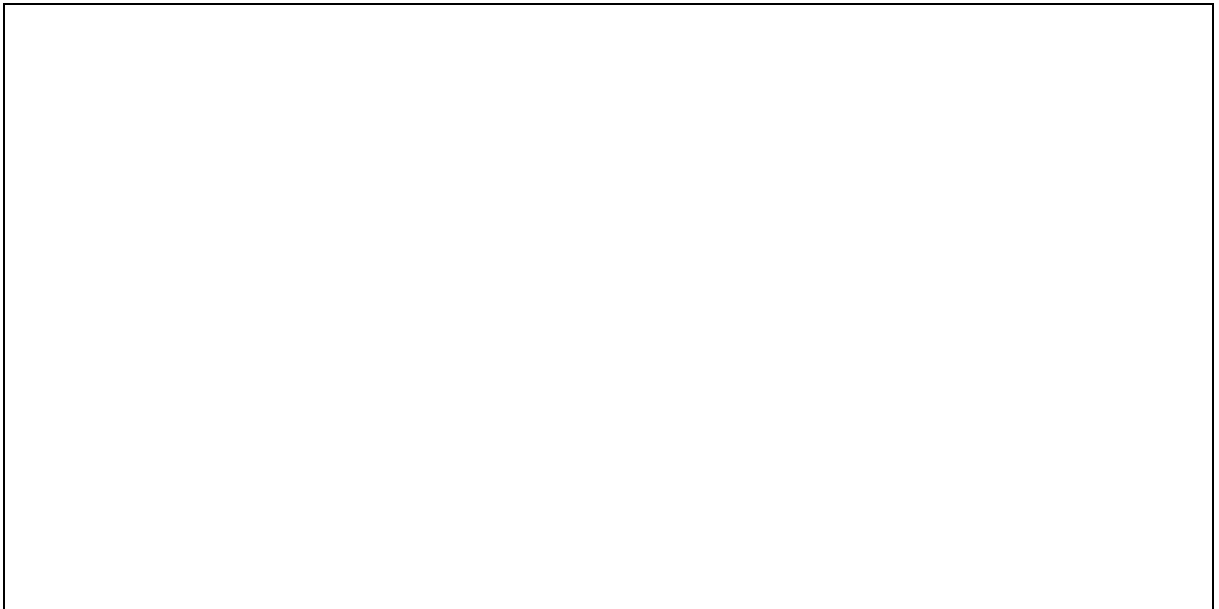
Trabalho de Conclusão apresentado ao Curso de Ciência da Computação da Universidade Tecnológica Federal do Paraná, Câmpus Santa Helena, como requisito parcial à obtenção do título de Bacharel em Ciência da Computação.

Orientador: Prof. Dr. Thiago França Naves

Santa Helena, Paraná

2019

Ficha catalográfica

A large, empty rectangular box with a thin black border, occupying the central portion of the page. It is intended for the entry of cataloging data.

Elaborada pela bibliotecária Cleide Bezerra



APÊNDICE 9

TERMO DE APROVAÇÃO

“RECONHECIMENTO DE PADRÕES FACIAIS: UMA SÍNTESE”

por

“Vambaster José da Costa”

Este Trabalho de Conclusão de Curso foi apresentado às 17:00 do dia 03 de Julho de 2019 na sala 410 como requisito parcial à obtenção do grau de Bacharel em Ciência da Computação na Universidade Tecnológica Federal do Paraná - UTFPR - Câmpus Santa Helena. O(a) aluno(a) foi arguido pela Banca de Avaliação abaixo assinados. Após deliberação, a Banca de Avaliação considerou o trabalho aprovado (aprovado ou reprovado).

Prof. Dr. Thiago França
Naves (Presidente - UTFPR/Santa
Helena)

Prof. Me. Anderson Brilhador
(Avaliador 1 - UTFPR)

Prof. Me. Claudio José Biazus
(Avaliador 2 - UTFPR)

Prof. Dra. Giani Carla Ito
(Professor responsável pelo TCC -
UTFPR/Santa Helena)

Prof. Dra. Arlete Terezinha Beuren,
(Coordenador do curso de Bacharelado em
Ciência da Computação – UTFPR/Santa Helena)

AGRADECIMENTOS

Agradeço primeiramente a Deus por ter me dado sabedoria para até aqui chegar.

Agradeço a minha esposa Diana Kochem da Costa por sempre me incentivar nesta caminhada, me auxiliando e por fazer parte das minhas conquistas.

Agradeço aos meus pais. Minha mãe Adriane Aparecida da Costa, por sempre me ajudar a conseguir todas as minhas conquistas, me ensinar o caminho que deveria andar, sempre me incentivar a ir mais longe. Meu padrasto, Paulo Sergio Alves de Sousa, que sempre foi um ótimo mentor, me orientando nas minhas decisões. Também dedico este trabalho ao meu pai Armindo Adolfo da Costa (in memoriam), que foi um exemplo de caráter e dignidade.

Agradeço a todo o corpo docente da Universidade Tecnológica Federal do Paraná - UTFPR, campus Santa Helena, pela excelência no ensino.

Agradeço ao Professor Dr. Thiago França Naves pela orientação, colaboração e todos os esforços empenhados para que o trabalho fosse concretizado.

Agradeço aos meus colegas, no apoio e no decorrer desta caminhada.

Agradeço também a banca examinadora deste trabalho, que é uma das minhas conquistas.

Nós só podemos ver um pouco do futuro, mas o
suficiente para perceber que há muito a fazer.
Alan Turing

RESUMO

COSTA, Vambaster José da. **Reconhecimento de padrões faciais: uma síntese**. 2019. 97 páginas. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) - Universidade Tecnológica Federal do Paraná. Santa Helena. Paraná. Brasil

Reconhecimento facial é uma tarefa biométrica executada de forma natural, instantânea e diária pelo ser humano, que reconhece pessoas de forma individual ou em grupos através do contato visual. Com o avanço tecnológico, as técnicas de reconhecimento facial estão presentes em diversos ambientes, como por exemplo, nos aplicativos de celulares, redes sociais e também em locais monitorados por câmeras. Assim, a tarefa de reconhecimento facial torna-se um recurso importante e versátil no mercado, sendo utilizada para promover tanto o entretenimento aos usuários quanto a segurança e monitoramento dos indivíduos que frequentam um determinado local. Apesar da ascensão e avanço tecnológico dos reconhecedores faciais, a metodologia para criação e configuração de tal sistema é complexa, sendo necessário o correto desenvolvimento de etapas que são encadeadas em seus resultados formando assim a abordagem final reconhecedora de faces. A atual literatura acerca de reconhecimento facial possui poucos trabalhos com foco no processo completo para geração deste tipo de sistema, com a maior parte destes concentrando-se em tarefas ou etapas específicas do processo de construção. Diante disso, este trabalho tem por objetivo a elaboração de uma síntese acerca da criação de um sistema de reconhecimento facial, apresentando todas as etapas necessárias para tal e suas respectivas técnicas, abordagens conceituais e nuances relacionadas. Serão abordados os algoritmos de reconhecimento facial *Eigenfaces*, *Fisherfaces* e *Local Binary Patterns Histograms* (LBPH), elencando as atividades de detecção de faces; pré-processamento de imagens faciais; extração de características; e classificação final de uma face. Por fim, são apresentados os resultados de diversos experimentos com os algoritmos mencionados, analisando o desempenho das etapas do reconhecimento facial e a qualidade na identificação final dos indivíduos.

Palavras-chave: Reconhecimento facial, Extração de Características, *Eigenfaces*, *Fisherfaces*, *LBPH*.

ABSTRACT

COSTA, Vambaster José da. **Recognition of facial patterns**: a synthesis. 2019. 97 p. Work of Conclusion Course (Graduation in Computer Science) - Federal Technology University - Paraná. Santa Helena. Paraná. Brasil

Facial recognition is a biometric task performed in a natural, instant and daily way by the human being, which recognizes people individually or in groups through visual contact. With the technological advancement, facial recognition techniques are present in a variety of environments, for example, in mobile applications, social networks, and in camera-monitored locations. Thus, the facial recognition task becomes an important and versatile resource in the market, being used to promote both the entertainment to the users as well as the security and monitoring of the individuals who attend a certain place. Despite the technological advancement and rise of facial recognizers, the methodology for creating and configuring such a system is complex, requiring the correct development of tasks that are chained in its results, thus becoming the final recognizing face approach. The current literature on facial recognition has few works focused on the complete process for the generation of this type of system, with most of them concentrating on tasks or specific stages of the whole process. The aim of this work is the elaboration of a synthesis about the creation of a facial recognition system, presenting all the necessary steps for this and their respective techniques, conceptual approaches and related nuances. The facial recognition algorithms Eigenfaces, Fisherfaces and Local Binary Patterns Histograms (LBPH) will be addressed, listing the activities of face detection; pre-processing of facial images; feature extraction; and final face classification. Finally, the results of several experiments with the mentioned algorithms will be presented, analyzing the performance of the tasks of facial recognition and the quality in the individuals final identification.

Keywords: Facial Recognition, Feature Extraction, Eigefaces, Fisherfaces, LBPH.

LISTA DE ILUSTRAÇÕES

Figura 1 - Detecção facial utilizado pelo Facebook.	18
Figura 2 - Categorias de classificação de características biométricas.	22
Figura 3 - Visão computacional para reconhecimento de imagens.	25
Figura 4 - Filtragem gaussiana em imagem com ruídos.	26
Figura 5 - Aplicação de contraste, a esquerda a imagem original, a direita a imagem aplicada o contraste.	26
Figura 6 - Imagem de tomografia destacando contornos da cena.	27
Figura 7 - Fluxograma das etapas de do processamento de imagem.	28
Figura 8 - Processo da tarefa de verificação.	30
Figura 9 - Processo da tarefa de identificação.	31
Figura 10 - Atividades do processo para reconhecimento facial.	32
Figura 11 - Detecção facial realizada pelo aplicativo do Facebook.	34
Figura 12 - Detecção de faces em um grupo de pessoas com ambiente ao fundo e diversos objetos.	34
Figura 13 - Fluxograma das etapas de um método baseado em conhecimento.	36
Figura 14- Funcionamento do método de <i>template</i>	37
Figura 15 - Diagrama da aplicação de técnicas de pré-processamento.	39
Figura 16 - Maior foco sobre a face após a aplicação da técnica de recorte.	40
Figura 17 - Normalização de luminosidade com uso de equalização de histogramas.	40
Figura 18- Filtragem de face com uso do filtro passa-baixa.	41
Figura 19 - Distribuição de cores por <i>pixels</i> de uma imagem e a representação de sua dimensionalidade de atributos.	43
Figura 20 - Geração do espaço de características.	44
Figura 21 - Linearização Aplicada pelo algoritmo LDA.	47
Figura 22 - Processo de detecção e reconhecimento facial.	51
Figura 23 - Exemplo de faces da Yale Database.	53
Figura 24 - Imagem da base <i>Yale Face</i> com 15% de oclusão.	54
Figura 25 - Exemplo de faces <i>AT&T Database of Faces</i>	54
Figura 26 - Exemplo Aplicando recursos de Haar sobre uma imagem facial.	55
Figura 27 – Quatro padrões de configuração de <i>features</i>	56

Figura 28 – Soma dos <i>pixels</i> na região d utilizando a integral da imagem.	57
Figura 29 - Exemplo de funcionamento da cascata dos classificadores.	58
Figura 30 – Histograma dos <i>pixels</i> contidos em uma imagem em escala de cinza.	59
Figura 31 – Imagem em escala de cinza antes da aplicação de contraste linear.	60
Figura 32 - Imagem em escala de cinza após a aplicação de contraste linear.	60
Figura 33 – Aplicação de realce não linear por histogramas.	61
Figura 34 – Aplicação de segmentação por <i>Thresholding</i> , onde a esquerda está a imagem original e a direita a nova imagem a partir da segmentação.	63
Figura 35 - Ilustração da geração da face média.	65
Figura 36 - Subtração da face média	66
Figura 37 - Exemplo de uma operação com LBP.	68
Figura 38 - Exemplo de nova imagem gerada pelo LBP.	69
Figura 39 - Exemplo da extração do histograma da imagem.	70
Figura 40 - Exemplo de LBP circular.	70
Figura 41 – Exemplo do cálculo de distância de Manhattan	72
Figura 42 - Esquema de divisão de amostras para a condução dos experimentos.	74
Figura 43 – Esquema utilizado para a construção do algoritmo de otimização dos testes.	79
Figura 44 - Média dos resultados de confiança dos algoritmos com a base <i>Yale Face</i>	82
Figura 45 - Média dos resultados de confiança dos algoritmos com a base <i>AT&T Faces</i>	85
Figura 46 - Média de confiança dos testes com oclusão	88
Figura 47 – Resultados <i>Eigenfaces</i> com 15% de oclusão na base de teste.	89
Figura 48 - Resultados <i>Fisherfaces</i> com 15% de oclusão na base de teste.	89
Figura 49 - Resultados <i>LBPH</i> com 15% de oclusão na base de teste.	90

LISTA DE QUADROS E TABELAS

Quadro 1 - Comparações de traços biométricos.....	24
Quadro 2 - Algoritmos de extração de características.....	44
Tabela 1 – Resultados <i>Eigenfaces</i> com base <i>Yale Face</i>	79
Tabela 2 - Resultados <i>Fisherfaces</i> com base <i>Yale Face</i>	80
Tabela 3 – Resultados <i>LBPH</i> com base <i>Yale Face</i>	81
Tabela 4 - Resultados <i>Eigenfaces</i> com base <i>AT&T Faces</i>	82
Tabela 5 - Resultados <i>Fisherfaces</i> com base <i>AT&T Faces</i>	83
Tabela 6 – Resultados <i>LBPH</i> com base <i>AT&T Faces</i>	84
Tabela 7 – Resultados do algoritmo <i>Eigenfaces</i> com base <i>Yale Face</i> obstruída.	86
Tabela 8 – Resultados do algoritmo <i>Fisherfaces</i> com base <i>Yale Face</i> obstruída.	86
Tabela 9 – Resultados do algoritmo <i>LBPH</i> com base <i>Yale Face</i> obstruída.	87

LISTA DE ABREVIATURAS E SIGLAS

FDC	Função de Distribuição Cumulativa
FLDA	Fisher Linear Discriminant Analysis
FN	Falso Negativo
FP	Falso Positivo
K-NN	K-Nearest-Neighbor
LBP	Local Binary Patterns
<i>LBPH</i>	Local Binary Patterns Histogram
LDA	Linear Discriminant Analysis
LPP	Locality Preserving Projections
PCA	Principal Component Analysis
RGB	Red, Green e Blue
SURF	Speeded Up Robust Features
SVM	Support Vector Machine
TFA	Taxa de Falsos Aceitos
TFR	Taxa de Falsos Rejeitos
TIC	Taxa de Identificação Corretas
TIF	Taxa de Identificação Falsas
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo
XML	Extensible Markup Language

LISTA DE SÍMBOLOS

σ : Desvio-padrão

n : Número de valores da amostra

Ψ : Vetor médio

SUMÁRIO

1 INTRODUÇÃO	15
1.1 OBJETIVOS	17
1.1.1 Geral	17
1.1.2 Específicos.....	17
1.2 JUSTIFICATIVA	17
1.3 CONTRIBUIÇÕES DO TRABALHO	19
1.4 DELIMITAÇÕES DO TRABALHO	20
2 REFERENCIAL TEÓRICO	21
2.1 BIOMETRIA	21
2.2 VISÃO COMPUTACIONAL.....	25
2.3 HISTORIA DO RECONHECIMENTO FACIAL.....	28
2.4 TAREFAS DO RECONHECIMENTO FACIAL	29
2.5 PROCESSO DE RECONHECIMENTO FACIAL	32
2.5.1 Detecção de Faces	33
2.5.2 Pré-Processamento.....	38
2.5.3 Extração de Características.....	41
2.5.4 Correspondência de Características de Classificação.....	45
2.6 NUANÇAS DO RECONHECIMENTO FACIAL.....	46
2.6.1 Variação na Aparência da Face	46
2.6.2 Análise da Face no Espaço de uma Imagem.....	46
2.6.3 Gerenciar Porções Não Lineares e Convexa da Imagem.	47
2.6.4 Espaço de Alta Dimensionalidade e Baixas Amostras	48
3 TRABALHOS RELACIONADOS	49
4 METODOLOGIA.....	51
4.1 BASES DE IMAGEM.....	52
4.2 DETECÇÃO FACIAL.....	55
4.3 PRÉ PROCESSAMENTO DE IMAGENS FACIAIS.....	57
4.3.1 Realce	59
4.3.2 Suavização	62
4.3.3 Segmentação.....	62
4.4 EXTRAÇÃO DE CARACTERÍSTICAS DA FACE E CLASSIFICAÇÃO	62
4.4.1 Extração de Características.....	63
4.4.1.1 Eigenfaces.....	64

4.4.1.2 Fisherfaces	66
4.4.1.3 LBPH	68
4.4.2 Reconhecimento Facial.....	70
4.4.2.1 Distância Euclidiana	71
4.4.2.2 Distância Cosseno.....	71
4.4.2.3 Distância de Mahalanobis.....	71
4.4.2.4 Distância de Manhattan	72
5 ANÁLISE DOS EXPERIMENTOS	73
5.1 FERRAMENTAS E FASES DO PROCESSO	73
5.1.1 Informações Técnicas	74
5.1.2 Parâmetros dos Algoritmos.....	75
5.1.3 Métricas de Qualidade	76
5.2 RESULTADOS E DISCUSSÕES	78
5.2.1 Resultados <i>Yale Face</i>	78
5.2.2 Resultados <i>AT&T Faces</i>	82
5.2.3 Resultados <i>Yale Face</i> Obstruída.	85
6 CONCLUSÃO.....	91
REFERÊNCIAS.....	93

1 INTRODUÇÃO

Sabe-se que reconhecimento facial é uma tarefa biométrica executada de forma natural e instantânea pelo ser humano, que é possível ser realizada todos os dias, reconhecendo pessoas através do contato visual. Existem diferentes métodos biométricos que analisam características inerentes dos seres humanos com o objetivo de reconhecer essas como padrões, tais como: íris ocular; impressão digital; palma da mão; voz; movimentação corporal; padrões faciais. Dentre esses, o reconhecimento facial se destaca como um dos mais precisos e simples de ser executado, seja por não exigir participação ativa dos indivíduos ou por ser menos suscetível a possíveis fraudes (BISSI, 2018). Por exemplo, técnicas de reconhecimento de mãos e dedos quando a epiderme do usuário está danificada, no momento da leitura, podem apresentar falhas. Métodos de reconhecimento por íris ou retina requerem um equipamento especializado e são sensíveis ao movimento do corpo, já o reconhecimento por voz é suscetível à ruídos do ambiente.

O emprego de técnicas biométricas e principalmente de reconhecimento facial está ligado a garantia e manutenção da segurança, seja através da autenticidade de documentos, protegendo acesso a ambientes ou monitorando espaços específicos ou dos indivíduos que lá estão. A biometria oferece vantagens sobre os métodos tradicionais de identificação como chaves, senhas, documentos ou *tokens* (SILVA e CINTRA, 2015). A necessidade do indivíduo de memorizar sequências alfanuméricas ou ter em mãos algum objeto para identificação e acesso a ambientes não se faz necessária. A garantia de controle e segurança vêm da baixa probabilidade de mudança natural ou temporal nos traços biométricos que os indivíduos possuem.

O reconhecimento facial consiste em identificar padrões em características faciais como formato da boca, face e nariz, distância dos olhos, entre outros (PRADO, 2018). O reconhecimento de faces é feito sem a interação direta com o indivíduo que será analisado, basta uma câmera para captura da imagem e um sistema computacional que processe o reconhecimento e realize a identificação. O reconhecimento identifica indivíduos através de análise facial automática de imagens presentes em fotos ou vídeos, sendo que o sistema pode operar de dois modos (FAGERTUN, 2005):

1. Verificação ou autenticação da face: consiste na comparação de 1 para 1, ou seja, comparar uma face que está sendo consultada com outra face para a qual está sendo requerida a autenticação.

2. Identificação ou reconhecimento de face: consiste na comparação de 1 para N, ou seja, compara uma imagem facial consultada em relação a todas as imagens faciais da base de dados, para determinar a identidade da face consultada.

No desenvolvimento deste trabalho as técnicas mencionadas e os experimentos realizados serão feitos utilizando a identificação ou reconhecimento de face, sendo que para cada face identificada em uma imagem a mesma é comparada com todas as faces que estiverem na base de dados utilizada para determinar a identidade do sujeito.

Para desenvolver um sistema de reconhecimento facial diversas etapas precisam ser consideradas, desde a montagem da base de fotos que será utilizada como arcabouço de faces conhecidas inicialmente pelo sistema, até a escolha do algoritmo que irá classificar e reconhecer uma face capturada de um indivíduo específico. Algumas das dificuldades que podem surgir nesse processo são as poses da face no momento da captura que podem ofuscar partes importantes da face, ou até mesmo a iluminação do ambiente que pode não ser favorável a exposição de alguma característica facial. Assim, tal sistema possui uma construção complexa que depende de muitos fatores e de uma correta parametrização em cada uma das etapas que o compõem.

Sendo assim, este trabalho pretende revisar e apresentar os principais conceitos que envolvem reconhecimento facial, desde a concepção de uma base de faces conhecidas até a escolha do algoritmo de reconhecimento final. As principais técnicas utilizadas, referências bibliográficas mais recentes e as configurações necessárias em cada etapa serão apresentados, bem como serão conduzidos experimentos a fim de verificar todas as informações apresentadas e resultados esperados para as metodologias de reconhecimento facial discutidas.

1.1 OBJETIVOS

Os objetivos desse trabalho são apresentados a seguir, classificados por: objetivo geral e específicos.

1.1.1 Geral

Sintetizar o processo de criação de um sistema de reconhecimento facial, através das diferentes metodologias possíveis de serem utilizadas nas etapas de configuração do mesmo junto com experimentos que corroborem a performance e acurácia dos algoritmos e técnicas consideradas.

1.1.2 Específicos

- 1) Realizar estudo sobre a história, as áreas de conhecimento e tarefas envolvidas no processo de reconhecimento facial;
- 2) Apresentar as etapas necessárias para construção de um reconhecedor de faces, através das diferentes técnicas disponíveis junto ao encadeamento de tarefas e seus respectivos resultados esperados;
- 3) Realizar experimentos com os algoritmos e técnicas escolhidos no estudo analítico;
- 4) Analisar o resultado dos experimentos, ajustar os parâmetros das bases de imagens e incluir os mesmos no documento do trabalho;

1.2 JUSTIFICATIVA

O uso de reconhecimento facial tem recebido muita atenção e destaque em diferentes áreas da computação e do mercado tecnológico. Seu uso para garantia e aumento da segurança já é feito e estudado há mais de uma década com aplicação em situações como:

- Acesso seguro a ambientes e locais de acesso restrito como cofres;
- Monitoramento de pessoas e identificação de suspeitos em locais públicos;
- Estudo do comportamento de pessoas com identificação em gravações de vídeo;

Além do foco em segurança nos últimos anos (FBI; 2014), o uso de reconhecimento facial para serviços de entretenimento têm sido uma vertente de grande exposição, principal-

mente em aplicativos de celular (HUA, 2011). Com a presença cada vez mais comum de câmeras com alta qualidade nos aparelhos telefônicos, os aplicativos de modificação facial e filtros de face são alvo de estudo e investimento para alcançar qualidade e capacidade de reconhecimento cada vez maior (MICROSOFT AZURE, 2019). A Figura 1, mostra exemplos de filtros a partir de reconhecimento facial utilizados no aplicativo para celulares Facebook (FACEBOOK, 2018).



Figura 1 - Detecção facial utilizado pelo Facebook.

Fonte: Facebook (2018).

Mesmo com o avanço no uso de reconhecimento facial, a construção de um sistema que desempenhe com eficiência a tarefa ainda é complexa. Geralmente os ambientes onde ocorre o processo de reconhecimento sofrem com muita interferência de fatores externos, como iluminação e pose dos indivíduos (AZEEM, 2014). Assim, é necessário conhecer as etapas do processo de reconhecimento facial, bem como as opções de técnicas e algoritmos que podem ser utilizados em cada etapa e como as diferentes metodologias combinam essas técnicas para um sistema mais eficiente em determinado cenário de uso.

Na literatura, grande parte dos trabalhos que focam no processo de reconhecimento facial descrevem a tarefa como uma operação singular. Com isso, as diferentes etapas do processo

e suas características são pouco consideradas e geralmente não são feitas inferências sobre como diferentes técnicas podem ser integradas para resolver problemas específicos. Por isso, este trabalho pretende elaborar uma revisão sistêmica e analítica sobre o processo de reconhecimento facial. O foco será dado na descrição das principais técnicas e como é possível considerar a integração de diferentes algoritmos, na construção de uma metodologia que possa aumentar a qualidade e performance dos resultados em um determinado cenário.

1.3 CONTRIBUIÇÕES DO TRABALHO

O trabalho desenvolvido pretende auxiliar o entendimento dos processos para a construção de sistemas de reconhecimento fácil aplicado a identificação de pessoas que possuem acesso a ambientes controlados, contribuindo no contexto de escolha, formas de uso e integração das técnicas abordadas, com foco em obter melhores resultados em cenários específicos com uso de reconhecimento facial, para acadêmicos e pesquisadores que tenham interesse sobre tal área do conhecimento.

Este trabalho propõe um estudo sistêmico sobre reconhecimento facial e pretende contribuir principalmente com uma análise sobre as principais etapas, técnicas e algoritmos envolvidos neste processo. Essa análise no decorrer do trabalho apresenta outras contribuições no contexto de escolha, formas de uso e integração das técnicas abordadas, com foco em obter melhores resultados em cenários específicos com uso de reconhecimento facial. Essas contribuições são:

1. Síntese dos processos de reconhecimento facial: as etapas necessárias para uso do reconhecimento serão apresentadas com detalhes, bem como as principais dificuldades encontradas e as melhores estratégias para contorná-las.
2. Análise dos melhores algoritmos para cada etapa do processo de reconhecimento facial: os mesmos serão apresentados com detalhes de funcionamento e com ênfase nos modelos estruturais e matemáticos que os representam. Também é feita uma análise da integração dos algoritmos, na construção de metodologias específicas com foco no uso em diferentes problemas.
3. Corroboração da qualidade de diferentes metodologias: os experimentos conduzidos e apresentados demonstram a acurácia e eficiência das metodologias propostas, bem como uma análise do uso prático dessas com relação as ferramentas de teste e gerenciamento das bases de imagens.

1.4 DELIMITAÇÕES DO TRABALHO

Com o objetivo de apresentar uma síntese sobre reconhecimento facial junto a uma análise das técnicas e metodologias disponíveis, foram definidas algumas limitações no escopo deste trabalho para tornar o mesmo coeso com seu escopo e escalável em sua construção. Assim, algumas das limitações deste trabalho são:

1. **Trabalhos relacionados:** Este trabalho irá focar nos trabalhos de maior impacto na área de reconhecimento de faces. Prioridade será dada para aqueles que auxiliem no entendimento dos algoritmos e técnicas escolhidos para construir as metodologias propostas neste trabalho.
2. **Base de imagens faciais:** na seção dedicada aos experimentos e testes serão utilizadas bases de imagem disponíveis por grupos de pesquisa em reconhecimento facial. Assim, não será proposta a construção de uma base de imagens com faces capturadas e gerenciadas pelos autores, apenas a modificação de bases conhecidas na literatura com aplicação de algum ruído sobre as faces. O foco da análise dos experimentos será na performance e qualidade das metodologias propostas, uso de bases de imagens disponíveis na literatura contribui para que esse objetivo seja atingido.
3. **Experimentos:** existem diversas métricas na literatura para mensurar a qualidade de um sistema de reconhecimento facial. Contudo, neste trabalho os experimentos serão focados em analisar apenas sete dessas, que são: Verdadeiro Positivo, Verdadeiro Negativo, Falso Positivo, Falso Negativo, Confiança, Precisão, Sensibilidade e Especificidade. Essa escolha está alinhada com o objetivo principal do trabalho, que consiste em demonstrar a qualidade dos algoritmos apresentados em relação a sua eficiência na identificação ou reconhecimento de faces dos algoritmos apresentados.

2 REFERENCIAL TEÓRICO

Este capítulo está dividido em cinco sessões, onde serão abordados os principais tópicos e tarefas necessárias para realização de reconhecimento de faces. Na seção 2.1 os conceitos sobre biometria são apresentados. Seção 2.2 é descrito alguns conceitos básicos de visão computacional. Na Seção 2.3 é apresentado os pontos descritos na literatura como mais importantes da história do reconhecimento facial. Na Seção 2.4 é apresentado as principais tarefas do reconhecimento facial. Na seção 2.5 é descrito os principais processos realizados no reconhecimento facial. Na seção 2.6 é apresentado algumas das nuances que podem ocorrer no reconhecimento facial.

2.1 BIOMETRIA

A palavra biometria tem proveniência do idioma grego - bio (vida) e metron (medida) - e refere-se a todas as técnicas que permitem reconhecer ou autenticar pessoas por meio de suas características comportamentais ou fisiológicas (GUGLINSKI, 2013). Há registros da utilização de métodos biométricos, como a impressão digital, que foram utilizadas há muitos séculos, embora não fossem tão refinadas como nos dias atuais. Os antigos babilônicos apertavam seus dedos na argila para registrar suas obras de arte. Os chineses usavam a impressão digital para negociar e ajudar a identificação de seus filhos (SAEED e NAGASHIMA, 2012).

O reconhecimento biométrico representa todo um conjunto de métodos e procedimentos que visa identificar indivíduos pelas suas medidas e padrões de características fenômeno biológicas. Entre algumas dessas características temos o comportamento, expressões corporais, características físicas, traços imutáveis e peculiaridades do ser humano onde Srivatava (2013), divide a biometria em duas categorias principais que são apresentadas na Figura 2, sendo essas:

Fisiológicas: São as características diretamente atreladas ao corpo humano, como DNA, impressão digital, retina, íris, formato da mão e medidas corporais. As características de ordem fisiológica são as mais confiáveis do ponto de vista de segurança e da capacidade de reconhecimento, uma vez que é alta a complexidade para fraudá-las e são levadas pelo ser humano no decorrer de sua vida.

Comportamentais: São as características relacionadas ao comportamento da pessoa, que pode sofrer variação de acordo com o seu estado psicológico ou estado sentimental. Entre as comportamentais temos: tom de fala, padrão de caminhada, ritmo de batimentos cardíacos, etc. As características comportamentais podem ser falsificadas com maior facilidade que

as fisiológicas, pois os padrões de um indivíduo podem ser imitados por outro e as características podem ser simuladas por dispositivos eletrônicos.

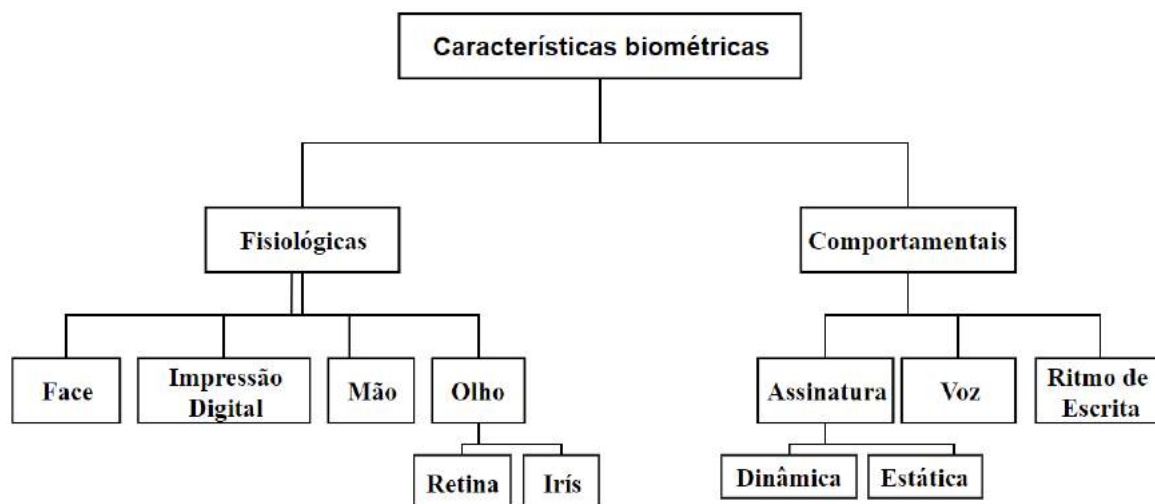


Figura 2 - Categorias de classificação de características biométricas.

Fonte: adaptado de Sinfic (2018).

Todos os sistemas biométricos compartilham do mesmo princípio básico, consistem em etapas predefinidas, sendo basicamente a extração das características do indivíduo; criação de um modelo com os dados; apresentação através de alguma interface; registro dos dados biométricos; resultados correspondentes com os dados capturados (SRIVASTAVA, 2013). Os primeiros sistemas de reconhecimento facial automatizados começaram a ser projetados na década de 60. Eram baseados em informações geométricas da face do indivíduo, mas exigiam que seus administradores localizassem as características na imagem (orelha, boca, olhos) antes de calcular as distâncias para um ponto de referência em comum. Desde então, várias linhas de pesquisa sobre o assunto se iniciaram, permitindo que novas tecnologias transformassem os processos deixando-os cada vez mais robustos e eficazes (FAGERTUN, 2005).

Entre os processos de identificação biométrica o reconhecimento facial destaca-se pela sua eficiência em evitar fraudes por autenticação biométrica. A complexidade em fraudar um método de identificação que utiliza dados exatos, particulares e de fácil captura em ambientes abertos e fechados é complexa (ARAUJO, 2010). O cerne da operação do reconhecimento facial consiste em identificar padrões em características faciais como formato da boca, da face, distância dos olhos, alinhamento do nariz, dentre outros (SILVA, SINTRA, 2015).

Outra vantagem do reconhecimento facial é sua facilidade em capturar os dados biométricos de um indivíduo para análise. Nesse, não se faz necessária a interação direta do indivíduo com um sistema computacional específico, diferentemente de outras abordagens biométricas que exigem o contato direto, bastando uma câmera para captura da imagem de forma indireta e um programa de reconhecimento facial. Uma desvantagem atrelada ao modo simples de captura é a falta de padronização no processo de captura utilizado e no ambiente que foi realizado a obtenção da imagem. Em muitos casos as imagens geralmente são capturadas com grandes espaços abertos com muitas faces presentes, além de condições de iluminação arbitrárias, diversos ângulos de visão, possibilidade de oclusão de algumas faces, que dificultam assim, o processo de reconhecimento facial.

Os diferentes modos de biometria podem ter suas vantagens avaliadas de acordo com a relevância das seguintes métricas de uso:

- **Acessibilidade:** o traço biométrico é acessível à leitura de forma simples ou necessita de um procedimento específico de captura, como por exemplo a leitura de uma impressão digital;
- **Desempenho:** o traço possibilita um bom reconhecimento do mesmo, ou seja, se a biometria garante a eficiência no processamento e correspondência ao indivíduo identificado;
- **Evasão:** o traço biométrico pode ser burlado, como por exemplo, o reconhecimento de um indivíduo com características semelhantes aos dados contidos na base e que correspondem a outro indivíduo;
- **Mensurabilidade:** o traço biométrico é mensurável, ou seja, a dimensionalidade da característica é possível de ser medida ou é complexa demais e exige alto processamento computacional;
- **Permanência:** o traço biométrico mantém suas características principais com o passar do tempo, não sofrendo com modificações tais como, por exemplo, a digital do dedo ou a íris do olho;
- **Unicidade:** o traço biométrico é único em cada indivíduo ou possui um alto fator de exclusividade permitindo a identificação de vários indivíduos com um traço semelhante;
- **Universalidade:** o traço biométrico está presente em todos os indivíduos da espécie que será identificada tais como, os olhos e boca nos seres humanos.

O Quadro 1, resume a qualidade de cada método biométrico a partir dos valores das métricas de uso. Os valores atribuídos podem ser utilizados como um guia para o uso de um determinado método em um problema específico que envolve o uso de biometria. De acordo com a SRIVASTAVA (2013), os métodos biométricos com as melhores avaliações são: face, impressão digital, impressão da palma da mão. Ambos os métodos se destacam em diversas métricas, porém nenhuma delas possui avaliação do item Desempenho como alta, e o reconhecimento facial se destaca sobre as duas outras no quesito Evasão e Universalidade, o que mostra as vantagens já mencionadas nesta seção sobre o uso de reconhecimento de faces frente a outras metodologias.

Quadro 1 - Comparações de traços biométricos

Biometria	Acessibilidade	Desempenho	Evasão	Mensurabilidade	Permanência	Unicidade	Universalidade
Arcada dentária	Baixo	Médio	Alto	Baixo	Baixo	Baixo	Baixo
Assinatura	Alto	Médio	Alto	Médio	Baixo	Alto	Baixo
DNA	Alto	Alto	Baixo	Baixo	Alto	Alto	Alto
Face	Alto	Baixo	Alto	Médio	Médio	Médio	Alto
Geometria da mão	Médio	Médio	Médio	Alto	Baixo	Médio	Alto
Íris	Médio	Alto	Baixo	Médio	Alto	Alto	Alto
Impressão digital	Alto	Médio	Médio	Alto	Alto	Alto	Médio
Impressão da Palma da mão	Alto	Médio	Médio	Alto	Alto	Alto	Médio
Ouvido	Médio	Alto	Baixo	Médio	Médio	Médio	Baixo
Retina	Alto	Alto	Baixo	Baixo	Alto	Baixo	Alto
Voz	Baixo	Médio	Alto	Médio	Baixo	Alto	Baixo
Ritmo de digitação	Baixo	Baixo	Médio	Baixo	Baixo	Baixo	Baixo

Fonte: Srivastava (2013).

2.2 VISÃO COMPUTACIONAL

Visão computacional é o método computacional que estuda, estrutura e criar maneiras de abstrair a realidade para o meio digital, para que os computadores possam interpretar objetos da realidade, para obter informações de imagens ou qualquer dado multidimensional. É a área do conhecimento que busca minimizar a visão humana quanto a capacidade de extrair das imagens obtidas as suas características, permitindo interpretar e descrever os elementos presentes nela (SZELISKI, 2010), imitando o ser humano para nas tomada de decisões de acordo com as informações contidas na imagem observada.

Executar os objetivos da visão computacional não é algo tão trivial, principalmente quando uma imagem possui uma grande quantidade de elementos. Grande parte das aplicações que utilizam visão computacional realiza a etapa de pré-processamento de imagem, para tratar a imagem de entrada antes de identificar ou extrair os elementos desejados por meio de algoritmos (PRINCE, 2012). A Figura 3 apresenta um exemplo de aplicação da visão computacional no processo de reconhecimento de objetos em imagens digitais.

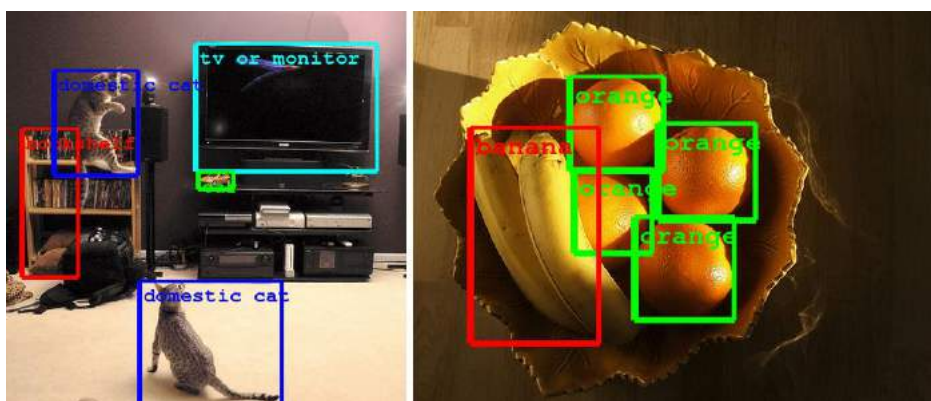


Figura 3 - Visão computacional para reconhecimento de imagens.

Fonte: Google AI (2014).

O processamento digital de imagens pode ser definido como o processo realizado por um computador onde espera-se como entrada e a saída uma imagem digital (GONZALEZ; WOODS, 2006).

A tarefa computacional de processamento de imagem é dividida em 3 níveis: baixo, médio e alto nível. Onde o processo de baixo nível é responsável pelo melhoramento da quali-

No processo de médio nível é realizada extração das regiões que compõem partes de um objeto ou cena, destacando as primitivas como retas, pontos, curvas, polígonos entre outros.

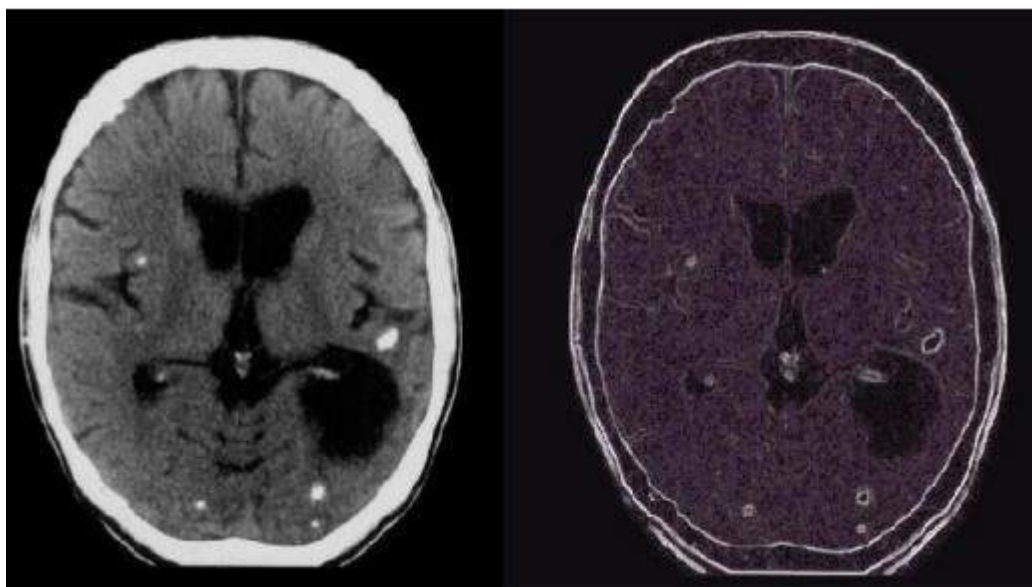


Figura 6 - Imagem de tomografia destacando contornos da cena.

Fonte: Computação Gráfica (2009).

Na Figura 6 é apresentado um exemplo de aplicação de filtro gradiente para a acentuação e detecção dos contornos da imagem. Os pontos se observados isoladamente não aparentem ter nenhum sentido, mas agrupados destacam as propriedades de mais destaques e relevantes de uma imagem dando contorno e forma aos destaques da cena.

O processo de alto nível desempenado pela visão computacional está relacionado a interpretação dos dados de entrada já normalizados pelos processos de baixo e médio nível, onde entram em ação algoritmos para a interpretação dos objetos ou cenas contidas na imagem de entrada processada.

Segundo Gonzalez e Woods (2006) o processamento de imagens é constituído por uma série de etapas como demonstrado na Figura 7 para melhor visualização. Porém não significa que todos os processos devem ser realizados necessariamente.

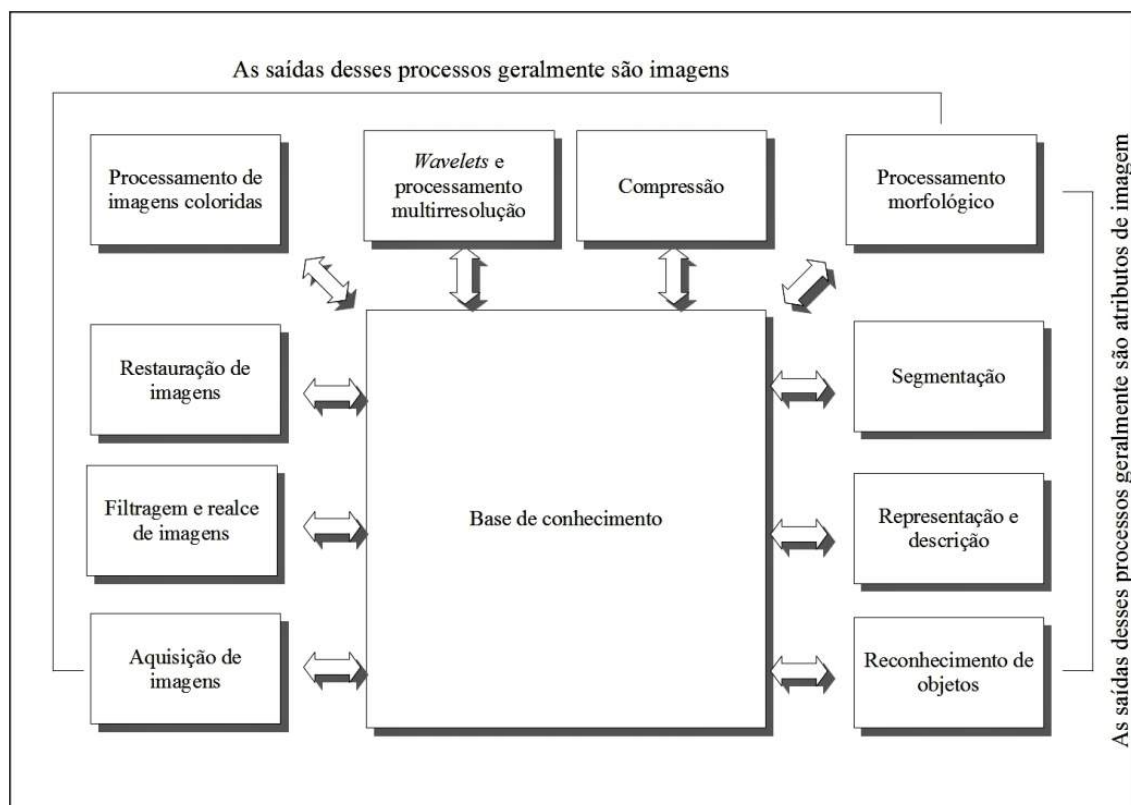


Figura 7 - Fluxograma das etapas de do processamento de imagem.

Fonte: Brilhador (2015).

2.3 HISTORIA DO RECONHECIMENTO FACIAL

Os primeiros sistemas de reconhecimento facial automatizados começaram a ser projetados na década de 60 (BLEDSOE, 1996). Mas somente em 1970 foram propostos sistemas de reconhecimento automatizados que operavam com base em informações geométricas da face do indivíduo, mas exigiam que seus administradores localizassem as características na imagem (orelha, boca, olhos) antes de calcularem as distâncias para um ponto de referência em comum (GALTON, 1888). Durante a maior parte da década de 1980, pesquisas sobre reconhecimento facial permaneceram essencialmente adormecidas.

Uma abordagem mais concreta do reconhecimento facial foi realizada por Fischler et al. (1973) e mais tarde por Yuille et al. (1992), quando se foi voltada a atenção novamente ao processo de reconhecimento facial. Desde então, várias linhas de pesquisa sobre o assunto se iniciaram, permitindo que novas tecnologias transformassem os processos deixando-os cada vez mais robustos e eficazes, com muitos desenvolvimentos, pesquisas e novas aplicações sendo apresentadas (ZHAO et al., 2003).

Algoritmos mais recentes já podem detectar faces e suas poses em ambientes desordenados (VIOLA E JONES, 2001). Viola e Jones (2001), propuseram uma abordagem de um detector baseado na integral da imagem, treinando vários classificadores usando uma cascata impulsionada de simples características, para reconhecimento de objetos, que após encontrar a integral da imagem, cada classificador reconhecia um padrão sobre a imagem, de modo que combinados em modo cascata, entram em acordo sobre o objeto reconhecido, classificando-o com base sobre o voto majoritário.

2.4 TAREFAS DO RECONHECIMENTO FACIAL

O processo de reconhecimento facial é executado para verificar a correspondência entre as características de uma face capturada e outras que possam estar presentes em uma base de dados. Contudo, esse processo pode ser executado de formas distintas para atender requisitos específicos de um problema ou ambiente que necessite de avaliação biométrica. Assim, o procedimento de reconhecer faces pode ser dividido em três tarefas específicas, que alteram a forma como o sistema reconhecedor vai atuar e processar as faces captadas (FAGERTUN, 2005). Essas tarefas específicas são: verificação (autenticação), identificação (reconhecimento) e observação (reconhecimento geral).

A tarefa de verificação (autenticação) trata o processo de reconhecimento facial como uma tarefa de identificação para acesso a um determinado local, onde o indivíduo é testado para comprovar se ele é quem realmente diz ser. O teste de verificação é feito separando os indivíduos em dois grupos distintos, os clientes, que são pessoas tentando obter acesso com sua própria identidade, e os impostores, que são as pessoas tentando obter acesso ao sistema através de uma identidade falsa. Na verificação existe o confronto de um para um entre as imagens para a identificação do usuário, ou seja, sua imagem facial é comparada com o objetivo de encontrar uma única imagem correspondente dentre as armazenadas no sistema. A Figura 8 mostra o processo de verificação facial, bem como o propósito da tarefa de verificar a autenticidade de um indivíduo específico ou a presença de um impostor (FAGERTUN, 2005).

Duas métricas são utilizadas para avaliar a qualidade do sistema na verificação de indivíduos, a Taxa de Falsos Aceitos (TFA) e a Taxa de Falsos Rejeitos (TFR) (YANG, 2002). A TFA mede a porcentagem de impostores que ganham acesso e mostra quando o sistema possui falhas na montagem do banco de dados de faces, que permite que características semelhantes de um cliente e um impostor sejam identificadas como iguais. Já a TRF mede a porcentagem

de clientes que têm o acesso negado, que diz respeito a falhas no modo como o sistema capta os dados no momento do acesso, não permitindo que as características obtidas possam ser identificadas no banco de dados. Na tarefa de verificação, faz-se necessária a interação do usuário com o sistema fornecendo a sua face para captação e comparação imediata em relação a sua identidade, como é feito geralmente em áreas com acesso restrito tais como em prédios e departamentos (YANG, 2002).

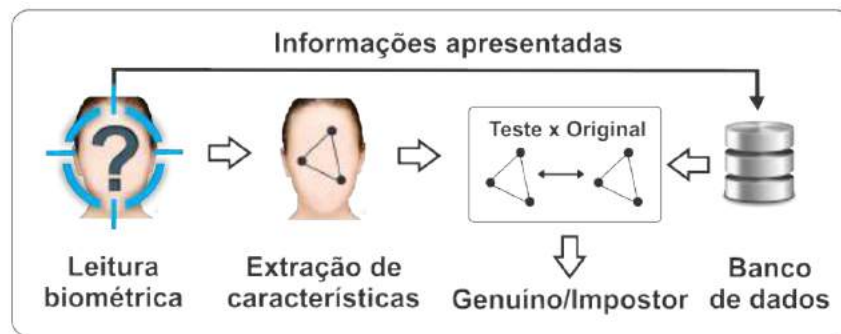


Figura 8 - Processo da tarefa de verificação.

Fonte: Silva (2015).

Entre os requisitos mais complexos em sistemas de reconhecimento facial está a atividade de vigilância, que consiste na verificação contínua de vários indivíduos e sua correspondência com relação as faces já conhecidas e armazenadas (PRADO, 2018). Essa atividade é feita pela tarefa de identificação (reconhecimento), que diferente da tarefa de verificação não exige interação direta dos usuários, capturando de forma automática a face de todos os indivíduos que entram em contato com uma câmera que monitora o ambiente. Aqui temos um confronto de um para muitos, onde a imagem de uma face capturada sem o conhecimento do usuário é comparada com várias armazenadas a fim de identificar o mesmo como conhecido ou desconhecido. Para realizar a identificação é necessário que o banco de dados com imagens contenha não só características faciais de um conjunto específico de pessoas, mas sim o máximo de faces possíveis para que seja possível fazer uma identificação de conhecido para qualquer indivíduo no ambiente (FAGERTUN, 2005). A Figura 9 mostra o procedimento da tarefa de identificação.

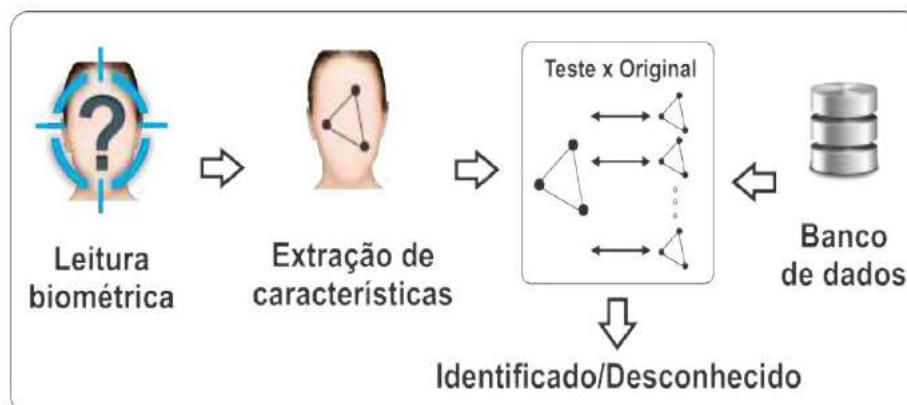


Figura 9 - Processo da tarefa de identificação.

Fonte: Silva (2015).

Na tarefa de identificação, o teste facial é feito com base na premissa de que todas as faces do teste são de pessoas conhecidas. Assim é comum o banco de dados conter imagens providas de diversas bases espalhadas em locais diferentes de onde o sistema se encontra operando. Quanto maior diversidade de amostras nas bases utilizadas maiores as chances de identificar um indivíduo como conhecido em um ambiente aberto e com muitas pessoas. Com isso, duas métricas são utilizadas para medir a porcentagem de faces que são identificadas de forma correta pelo sistema, a Taxa de Identificações Corretas (TIC) e Taxa de Identificações Falsas (TIF) (KRIG, 2014). Essas métricas são utilizadas para avaliar a performance do sistema de reconhecimento e também para mostrar em qual aspecto o mesmo possa estar apresentando menor eficiência ou comportamento incorreto. Como exemplo, uma vez que a TIF está alta o sistema pode estar com dificuldades devido ao baixo número de amostras de faces no banco de dados.

A tarefa de observação (reconhecimento geral) é uma generalização da tarefa de identificação, que nesse caso inclui a possibilidade de considerar pessoas desconhecidas. Na observação existe a possibilidade de montagem de uma lista de observação, contendo faces desconhecidas a serem observadas em um determinado ambiente. Assim, é feita a verificação dos indivíduos que vão surgindo no ambiente a fim de descobrir se algum desses possui correspondência facial para a lista de observação. Nos testes de correspondência da tarefa de observação as métricas de TIC e TIF são utilizadas, mas é possível incluir as métricas de TFA e TRF para mensurar o quanto faces capturadas de indivíduos desconhecidos estão sendo marcadas como correspondentes na lista de observação (KRIG, 2014).

2.5 PROCESSO DE RECONHECIMENTO FACIAL

O uso de reconhecimento facial em qualquer uma de suas possíveis tarefas é uma atividade complexa, cuja execução depende de diversas etapas que são executadas em uma ordem específica. Essa ordem define o processo necessário para construção e uso de um sistema reconhecedor de faces, que inicia com uma base de imagens de faces que será submetida junto os algoritmos necessários a este processo.

A primeira atividade executada em um processo de reconhecimento facial é a detecção de faces, que irá encontrar as faces dos indivíduos presentes nas imagens do banco de dados. Em seguida é executado o pré-processamento, que recebe as faces encontradas na tarefa anterior e aplica técnicas que visam melhorar a qualidade da imagem e consequentemente a performance do reconhecedor (GALTON, 1888). A próxima atividade é a extração de características responsável por encontrar os atributos mais relevantes na separação dos tipos de faces presentes nas imagens pré-processadas, além de reduzir a dimensionalidade do domínio de imagens faciais. Por fim, é executada a Classificação e Correspondência de faces, onde as faces dos indivíduos que foram capturadas por uma câmera têm suas identidades verificadas e validadas, como legítimas ou não, por meio de algoritmos de classificação de padrões (ZHAO et al, 2003). A Figura 10 exibe um diagrama com o fluxo das atividades que integram o processo de reconhecimento facial.

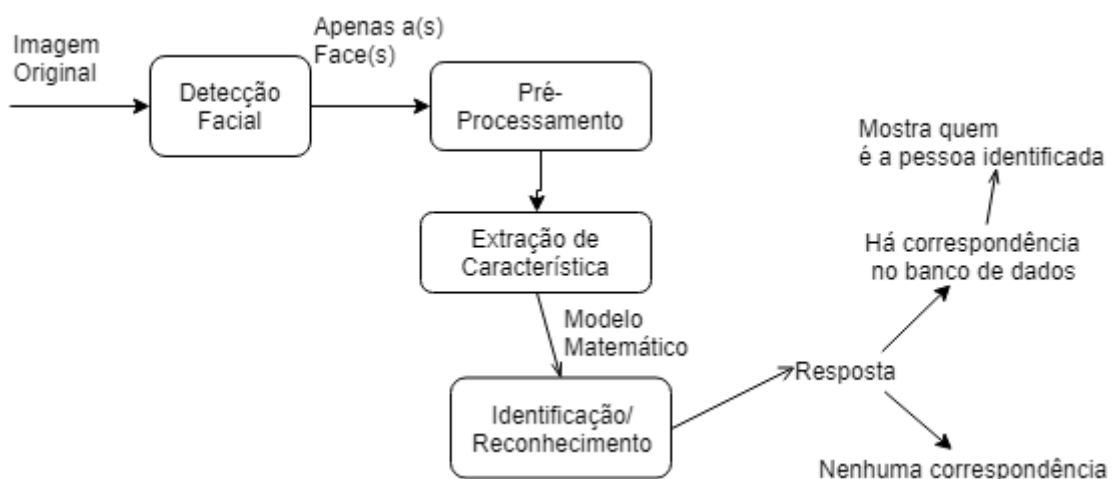


Figura 10 - Atividades do processo para reconhecimento facial.

Fonte: adaptada de Braga (2013).

2.5.1 Detecção de Faces

Detecção de faces é a primeira atividade que deve ser feita em um processo de reconhecimento facial, como mostrado na Figura 10. O objetivo dessa atividade é dentro de uma imagem digital detectar faces que pertençam a faces humanos, extraíndo-os para que sejam utilizados nas etapas posteriores do reconhecimento facial. Como primeira atividade, a detecção facial elimina todos os elementos de uma imagem que não são faces, o que faz com que as próximas etapas não precisem lidar com diversos detalhes irrelevantes que possam existir em uma imagem acelerando o processo de reconhecimento como um todo. YANG et al. (2002) define o processo de reconhecimento facial como: “em uma imagem qualquer, o objetivo da detecção de faces é determinar se existe ou não alguma face na imagem e, se presente, retornar a localização e dimensão de cada face na imagem”.

A detecção facial representa em sua complexidade uma área totalmente independente do reconhecimento facial, sendo utilizada em diversas áreas como robótica, interação homem computador, segurança digital e a própria visão computacional. A aplicação de tal técnica pode ser vista na indústria, setor de propagandas e nos últimos anos no setor de entretenimento. Como exemplo, serviços de mídias sociais, como o Facebook, utilizam a detecção de face para identificar faces presentes em imagens, de modo que o usuário possa notificar essas pessoas aumentando o alcance da imagem dentro da rede social. A Figura 11, mostra um exemplo da ferramenta de detecção facial da rede social Facebook, onde a mesma identifica e isola a face do usuário em relação a imagem em geral.

Os algoritmos de detecção de faces presentes na indústria recebem melhorias constantes para que seja possível a detecção em imagens com diversos tipos de ruídos, como natureza ao fundo, sobreposição de pessoas e presença de objetos do mundo real. A Figura 12 representa as dificuldades em realizar detecção facial em grupo de pessoas com ruídos na imagem e presença de diversos objetos. Na imagem, algumas faces não foram identificadas devido a oclusão parcial da face ou mesmo em função de outra pessoa que se encontrar mais à frente na direção da imagem, ou por estar muito próximo ao ambiente de fundo onde a face não exibe contraste adequado.

Mesmo com todos os avanços que a detecção facial vem recebendo nos últimos anos pela indústria em geral (FBI, 2014), a tarefa ainda permanece complexa em certos tipos de cenários. Isso ocorre, pois em qualquer imagem a ser analisada, o algoritmo não possui conhecimento da localidade das faces das pessoas (AZEEM, 2014). Diversos objetos presentes na

imagem podem ser confundidos com uma face, ou de fato, estar próximo de uma ocluindo essa face. No geral, algoritmos que executam essa tarefa são treinados exaustivamente para obterem um resultado satisfatório. O treino é feito com diversas imagens que possuem faces distintas, bem como com imagens que possuem objetos que não são faces, mas podem confundir o algoritmo de detecção. Assim, o problema de detecção facial pode ser definido como um problema de reconhecimento de padrões, porém com um padrão simples com duas possíveis classes, a primeira sendo a presença de uma face e a segunda a presença de qualquer outro objeto presente na imagem.

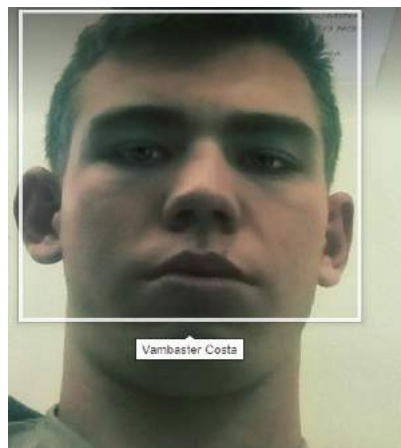


Figura 11 - Detecção facial realizada pelo aplicativo do Facebook.

Fonte: Facebook (2018).



Figura 12 - Detecção de faces em um grupo de pessoas com ambiente ao fundo e diversos objetos.

Fonte: Towards Data Science (2018).

Os algoritmos com foco em detecção de face podem ser divididos em quatro categorias que definem os métodos operacional do mesmo, sendo estas: baseado em conhecimento, baseado em características invariantes, baseado em *templates* e baseado em aparência (PRADO, 2018).

Baseado em Conhecimento: É aquele que busca estabelecer regras de produção utilizando o conhecimento humano sobre como uma face pode ser descrita para o seu encadeamento. As regras geradas com esse tipo de conhecimento focam nas relações entre os elementos de uma face, do ponto de vista que o ser humano possui. Como exemplo, está a distância média que os dois olhos de uma pessoa possuem e o triângulo que geralmente é formado pelo conjunto queixo, olho esquerdo e olho direito. Esse método possui boa performance em localização de faces devido a descrição precisa do formato de uma face humana em uma imagem com muitos elementos. Uma grande desvantagem do método baseado em conhecimento é a possibilidade de generalização das regras, de modo que diversos objetos da cena possam ser enquadrados como face o que gera uma explosão de falsos positivos. Um trabalho pioneiro nessa metodologia foi o de Yang e Huang (1994), nesse os autores propõem etapas para uso de regras baseadas em conhecimento. Cada uma das etapas possui uma prioridade em relação ao nível de especificidade das regras que a compõem, no primeiro nível as regras são as mais gerais possíveis e no decorrer da execução as etapas vão ficando mais específicas até atingir o maior nível na última dessas regras. A Figura 13 mostra os detalhes das etapas, as regras dos níveis 1 e 2 devem ser aplicadas na imagem com uma qualidade de resolução menor que a original, a ideia é conseguir comparar paletas de cores das possíveis faces e o fundo. No nível 3 o objetivo é encontrar as bordas dentro da imagem, como um medidor para verificar se uma possível boca ou olho identificado na etapa anterior não é falsa, uma vez que essas partes do corpo possuem contornos bem definidos e não permitiriam bordas horizontais. Assim, o número de falsos positivos é reduzido e nos experimentos realizados por Yang e Huang (1994) a média de acertos foi de 83%.

Baseado em Características Invariantes: É um método que visa o contraponto ao baseado em conhecimento, onde o objetivo é encontrar características que são invariantes ou imutáveis em qualquer face e priorizar a busca por esse no momento da detecção facial. Como principais características temos a cor da pele, textura da face e posicionamento dos elementos da face. A imagem a ser analisada é convertida para uma escala de cinza onde é possível obser-

var melhor o contraste nas escalas de cinza e eliminar redundâncias de exposição. Assim, independente da pose, iluminação ou ponto de vista da face na imagem ainda é possível localizá-la. Contudo, esse método não obtém bons resultados quando outros tipos de ruídos estão presentes na imagem como presença de cores fracas, objetos na mesma tonalidade que a pele das pessoas e fundos com tons neutros (YANG; KRIEGMAN; AHUJA, 2002).

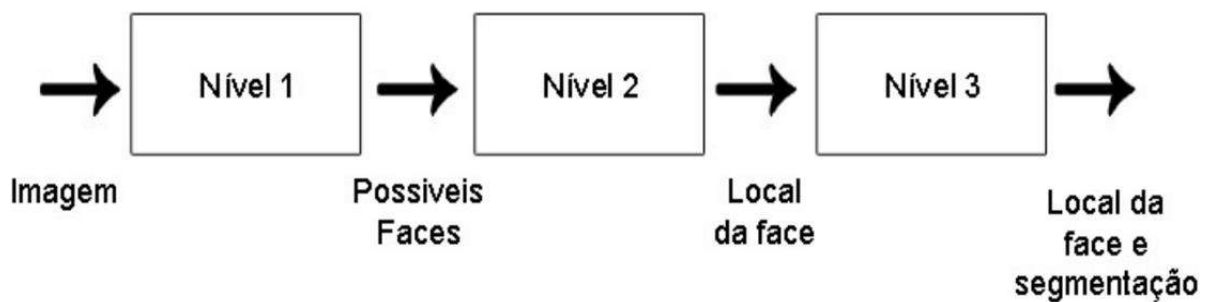


Figura 13 - Fluxograma das etapas de um método baseado em conhecimento.

Fonte: Braga (2013).

Baseado em *Templates*: É um método que busca por padrões definidos a priori para localizar e extrair uma face, mas diferente do baseado em conhecimento este faz uma montagem paramétrica de uma face utilizando essa como um *template* para encontrar outras (DINIZ et al., 2013). A construção de uma face *template* é feita a partir de elementos básicos como retângulos e quadrados posicionados como elementos de uma face, esses formam pontos de controle que vão diminuindo à medida que o algoritmo percorre a imagem para corresponderem ao formato das faces que lá possam estar. A cada iteração em um método de *template*, os pontos de controle tendem a se aproximar de alguma face deixando o encaixe entre elas o mais próximo possível. A Figura 14, mostra o princípio de funcionamento do método de *template*. Na primeira iteração os pontos de controle estão posicionados sobre uma posição aleatória da imagem, nas iterações seguintes estes vão se aproximando do formato final da face detectada.

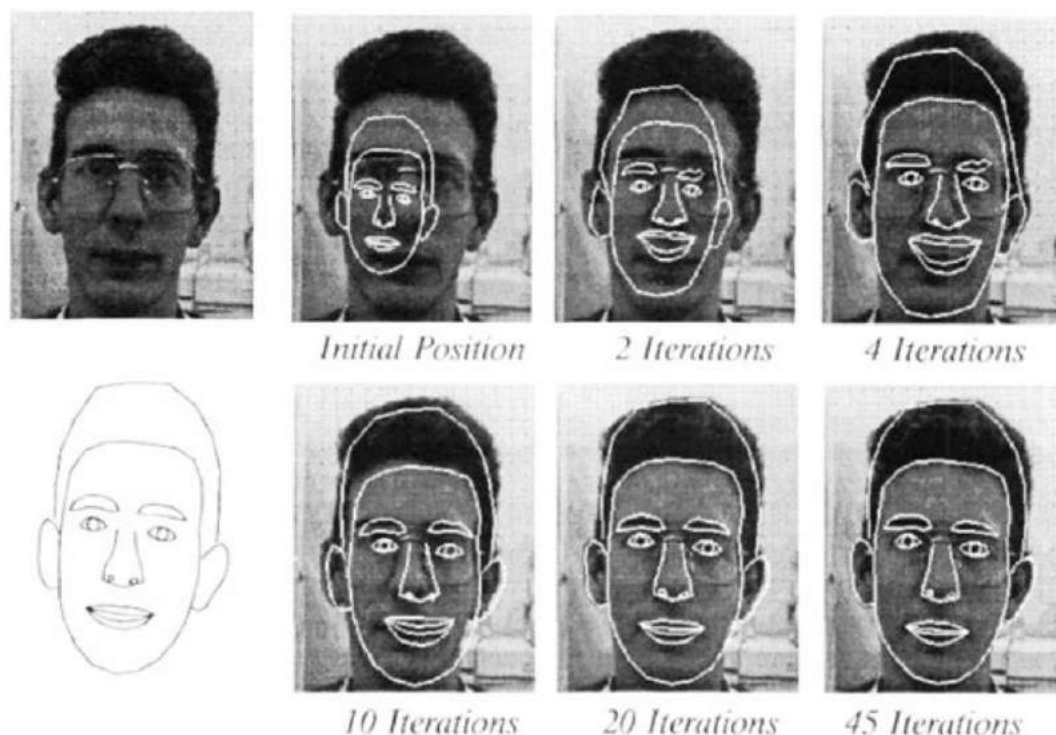


Figura 14- Funcionamento do método de *template*.

Fonte: Cootes, Taylor, Lanitis (1994).

Baseado em Aparência: É o método que atualmente recebe maior atenção da comunidade científica, com novos algoritmos e técnicas que vem sendo aprimoradas (ISAAC, 2018). O objetivo desse é utilizar algoritmos de aprendizado de máquina, que são treinados com grandes bases de dados com imagens contendo faces, bem como daquelas imagens que contém outros tipos de objetos. O método de aparência trata a detecção facial como um problema de reconhecimento de padrões, como já mencionado anteriormente, onde o propósito é aprender sobre a decisão correta entre duas classes, se possui face(s) ou não. Com a possibilidade de utilizar algoritmos de aprendizado de máquina ou propor modificações ao já existentes, o uso do método de aparência é grande e aberto a inúmeras possibilidades de melhores soluções. Entre os algoritmos que são utilizados nesse método temos: *support vector machine* (SVM), *k-nearest-neighbor* (KNN), *k-means*, entre outros. Esses são muitas vezes combinados com outros algoritmos formando um método híbrido, que busca resolver o problema por etapas onde cada uma efetua um processamento específico na imagem.

Por tratar-se de uma tarefa complexa, a detecção facial utiliza duas medidas na avaliação da qualidade dos resultados obtidos. A primeira, a quantidade de objetos que são identificados

erroneamente como sendo uma face, ou seja, a quantidade de falsos positivos. A segunda, considera a quantidade de faces que existem e não foram identificadas como tal, sendo esses valores os falsos negativos. Os valores de ambas as avaliações devem ser o menor possível, sendo utilizados para guiar em qual requisito o algoritmo utilizado na detecção possa estar com problemas. Um dos algoritmos com melhor desempenho e mais utilizados na literatura para detecção de faces é o chamado Viola-Jones, que utiliza o método de aparência como base. Esse algoritmo e os demais utilizados nessa etapa do processo de reconhecimento facial serão apresentados com detalhes no capítulo 4.

2.5.2 Pré-Processamento

O pré-processamento de imagens é a segunda etapa a ser executada no processo de reconhecimento facial. O objetivo é submeter as imagens de faces na base de dados a técnicas que efetuem algum tipo de processamento sob essas, onde a saída desse processo é uma nova imagem com alguma modificação (KRING, 2014). A nova imagem é gerada para aumentar a performance do reconhecimento, fornecendo melhores informações sobre a face, normalização das cores dos *pixels* e aproximação das faces presentes ou remoção de ruídos presentes na cena. Após a etapa de pré-processamento todas as imagens da base estarão no mesmo padrão de cor e dimensão, além disso todas as faces capturadas por câmeras ou qualquer outro dispositivo, devem ser submetidas a essa etapa do processo antes da tentativa de reconhecimento.

Uma das grandes dificuldades de obter altas taxas de reconhecimento facial na tarefa de identificação ou reconhecimento em ambientes abertos é a baixa qualidade das imagens e por consequência a baixa qualidade das faces que podem não serem detectadas e extraídas. Diversos fatores como posição e qualidade da câmera entre outros afetam a captura da face. Assim, o pré-processamento surge como uma forma de minimizar esse problema e aumentar a taxa de acerto dos algoritmos de reconhecimento. Entre as técnicas mais utilizadas nessa tarefa temos: recorte da imagem; redimensionamento; normalização e filtragem (FAGERTUN, 2005). A Figura 15, mostra o diagrama da aplicação de técnicas de pré-processamento que é feito sob uma imagem com as técnicas mencionadas.



Figura 15 - Diagrama da aplicação de técnicas de pré-processamento.

A imagem facial que é dada como entrada para o pré-processamento é convertida para uma escala de cor cinza, a partir de uma transformação na intensidade de seus *pixels*. Essa conversão é para que detalhes de iluminação e cores de objetos não afetem a análise e extração de características feita na próxima etapa, realces de fundo da imagem podem gerar um viés que interfere nas posições com *pixels* da face (ROZARIO et al., 2014). A primeira técnica aplicada sobre a imagem convertida em cinza é o recorte, sendo este aplicado junto com a etapa anterior de detecção facial. Quando a face é detectada é preciso executar o recorte para extraí-la, isso é feito de modo que a face represente a maior área possível da imagem com o máximo de remoção do fundo e de outros objetos. Uma forma eficiente de executar o recorte é copiar a imagem e aplicar uma conversão do padrão de cor espacial RGB para o formato YCbCr (DHARAVATH, 2014). Com os valores de Cb e Cr satisfazendo a seguinte condição, $133 \leq Cr \leq 173$, é possível verificar partes inferiores da imagem para determinar o melhor recorte. A Figura 16, mostra um exemplo de recorte feito em imagens onde foi possível aumentar o espaço ocupado pela face e diminuir interferência do fundo.

A técnica de redimensionamento visa aumentar ou diminuir o tamanho da imagem, com o objetivo de atingir um padrão único para toda a base. Assim, o desempenho do algoritmo de reconhecimento não será afetado por padrões de tamanho diferentes nas imagens de faces. Uma forma de executar um bom redimensionamento é utilizando o algoritmo de classificação KNN, onde o mesmo analisa bases de imagem com redimensionamento ideal e incorpora esse padrão para a base que está sendo pré-processada (SHARIF et al., 2017).



Figura 16 - Maior foco sobre a face após a aplicação da técnica de recorte.

Fonte: Sharif et al. (2017).

A normalização lida com o problema de variação de iluminação nas faces capturadas, sem o controle de iluminação a imagem possui contraste não uniforme e a distribuição de cinza não é semelhante nos *pixels* (VIOLA e JONES, 2001). A equalização de histogramas é uma técnica utilizada para suavizar a luminosidade difusa. Dada uma imagem de tamanho $M \times N$, uma Função de Distribuição Cumulativa (FDC) percorre cada *pixel* e encontra o valor de equalização a ser utilizado para normalização. Com isso, a luminosidade na imagem torna-se padronizada com prioridade na escala para aumentar a visibilidade da face (SHARIF et al., 2017). Um exemplo de uso da equalização de histogramas é apresentado na Figura 17.



Figura 17 - Normalização de luminosidade com uso de equalização de histogramas.

Fonte: Sharif et al. (2017).

Por fim, a filtragem é feita para lidar com ruídos gaussianos que geralmente surgem em imagens que possuem luminosidade difusa, mesmo após a normalização da mesma. Esses ruídos surgem em forma de *pixels* serrilhados ou fusão de *pixels* vizinhos, que faz com partes da imagem fiquem com mesmo formato de apresentação e dificulta a localização de pontos de

extração de características (DHARAVATH et al., 2014). Entre os filtros que podem ser utilizados para eliminar os ruídos, o chamado passa-baixa analisa na composição da imagem onde estão as altas frequências em *pixels* e elimina essas, dando prioridade a frequências baixas. Nessas, a imagem recebe uma sobreposição das menores frequências que deixam um aspecto constante no contraste, facilitando a análise da mesma (SHARIF et al., 2017). A técnica de filtragem com passa-baixa é exibida na Figura 18, a imagem filtrada apresenta uma camada de sobreposição com a replicação das menores frequências nos *pixels*.



Figura 18- Filtragem de face com uso do filtro passa-baixa.

Fonte: Sharif et al. (2017).

2.5.3 Extração de Características

Uma imagem de tamanho $M \times M$, com resolução $N \times N$ *pixels* em uma face, possui um número N^2 de possíveis atributos, que podem ser consideradas por um algoritmo de classificação no momento de reconhecer uma face. Por exemplo, uma imagem com resolução de 128×128 *pixels* teríamos um total de 16.384 atributos a serem utilizados, um valor que inviabiliza o uso de qualquer algoritmo de classificação para reconhecer uma face com base nestes atributos. Essa grande quantidade de características advém do fato que cada *pixel* pode representar uma informação de cor e luminosidade única na imagem. Cada *pixel* pode então ser considerado como um atributo específico para definição da imagem e para reconhecer as diferenças entre elas. Assim, a explosão de características gera a alta dimensionalidade, pelo grande número de dimensionalidade, que torna-se um problema recorrente no uso e processamento de imagens (BELHUMEUR et al., 1997).

Diante da complexidade dos atributos e alta dimensionalidade a tarefa de extração de características possui um papel vital no processo de reconhecimento facial. Seu principal objetivo é de analisar as imagens faciais e encontrar os conjuntos de características que melhor definam e diferenciem as faces da base de imagens, reduzindo a dimensionalidade original dessas para um novo espaço onde os atributos restantes e suas combinações sejam o suficientes para que algoritmos de classificação possam identificar estas faces, tornando a imagem padronizada.

Os algoritmos de extração de características operam com duas vertentes principais (GONZALES; WOODS, 2006). A primeira de ranquear as melhores características individuais que podem ser utilizadas no novo espaço de dimensionalidade reduzida (FAGERTUN, 2005). Na segunda, verificam combinações de características e a relevância dessas na separação de classes das faces da base, os atributos combinados têm prioridade de uso sobre aqueles que possuem apenas relevância com uso individual. Os atributos melhores ranqueados com as devidas prioridades formam o conjunto de características que serão utilizados no processo de reconhecimento facial. Um exemplo de redução de dimensionalidade é apresentado na Figura 19, a esquerda uma distribuição de atributos que descrevem a imagem em um alto espaço dimensional. As retas 1PC e 2PC mostram a combinação dos valores de dois atributos que melhor descrevem as imagens analisadas, esses são extraídos do conjunto original e formam um novo espaço de características do lado esquerdo da Figura 19, do plano *N-dimensional* para o plano 2D. Esse, configura uma dimensionalidade reduzida com atributos que representam a imagem e serão utilizados para o reconhecimento facial.

Uma das metodologias mais utilizadas na extração de características pelos algoritmos que executam essa tarefa é a exploração de covariância dos *pixels* da imagem, onde são gerados auto vetores, pois é a partir desses valores que melhor se definem as características relevantes da face. Os auto vetores descrevem a variância dos atributos considerados mais relevantes, sendo possível reduzir a operação de comparação de faces a avaliação dos auto vetores gerados (SILVA, 2015). Essa transformação é executada de forma linear, a qualidade de um auto vetor é dada pelos seus autovalores associados. Dentre os principais algoritmos que utilizam extração de características por meio de covariância de *pixels* temos: *Eingenfaces*; *Fisherfaces* e *Local Binary Patterns*. A Figura 20, ilustra a estratégia de covariância de *pixels*, os auto vetores selecionados representam os atributos mais relevantes para reconhecimento da face de entrada. Em

seguida, os auto vetores são combinados para aumentar a possibilidade de diferenciação das faces que serão comparadas, consequentemente aumentando também a taxa de acerto.

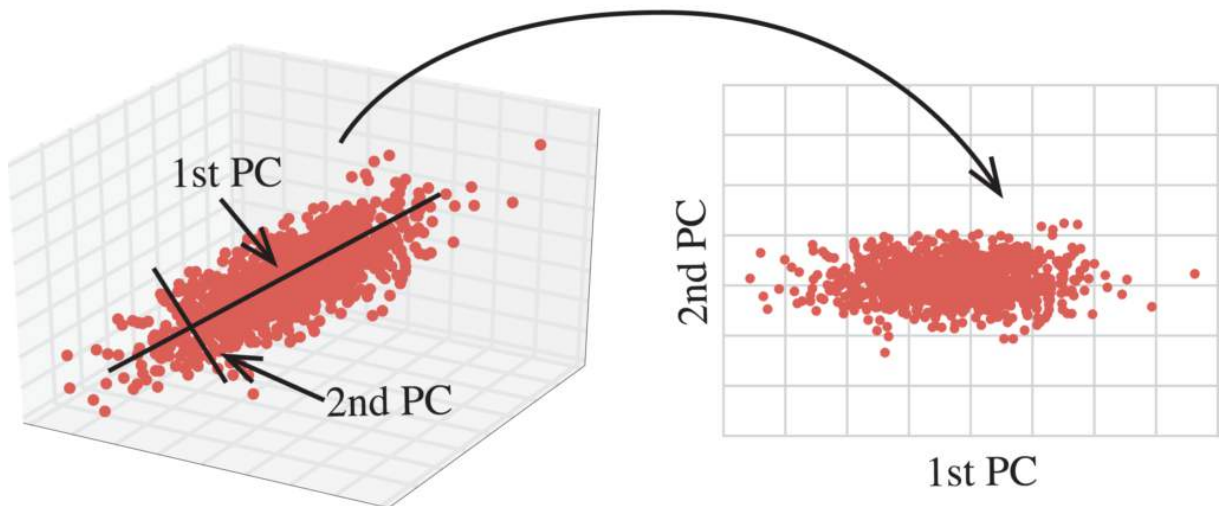


Figura 19 - Distribuição de cores por *pixels* de uma imagem e a representação de sua dimensionalidade de atributos.

Fonte: SaiGayatri Vadali (2018).

Os principais algoritmos de extração de características podem ser divididos em dois grupos. O primeiro grupo é dividido nos algoritmos que fazem análise da imagem e preservam sua estrutura global, ou seja, a distância entre os atributos em relação aos seus valores internos é mantida no novo espaço de dimensão reduzida. O segundo grupo é composto pelos algoritmos que preservam apenas a estrutura local da imagem. Nesses, após definir os atributos mais relevantes da face, apenas os valores de características com distâncias próximas são mantidos. Atributos separados por regiões diferentes não mantêm os valores de correlação entre eles, que significa que estruturas globais deixam de existir e são mantidas apenas aquelas locais e próximas.

Com a separação dos algoritmos em dois grupos mencionados anteriormente, ainda é possível dividi-los entre aqueles que utilizam funções lineares ou não lineares para extração de características e redução de dimensionalidade. Funções lineares efetuam transformações nos vetores de atributos através de operações simples, que deformam esses no espaço em relação a altura e rotação na busca por uma dimensionalidade maior. Já as funções não lineares efetuam operações complexas que deformam os dados e separam os mesmos para definir os atributos

mais relevantes. O Quadro 2, mostra os grupos e subgrupos de alguns dos mais utilizados algoritmos de extração de características. Dentre esses, o *Fisher Linear Discriminant Analysis* (FLDA), *Principal Component Analysis* (PCA) e *Locality Preserving Projections* (LPP) serão apresentados e seus detalhes de funcionamento serão discutidos na Seção 4.4.

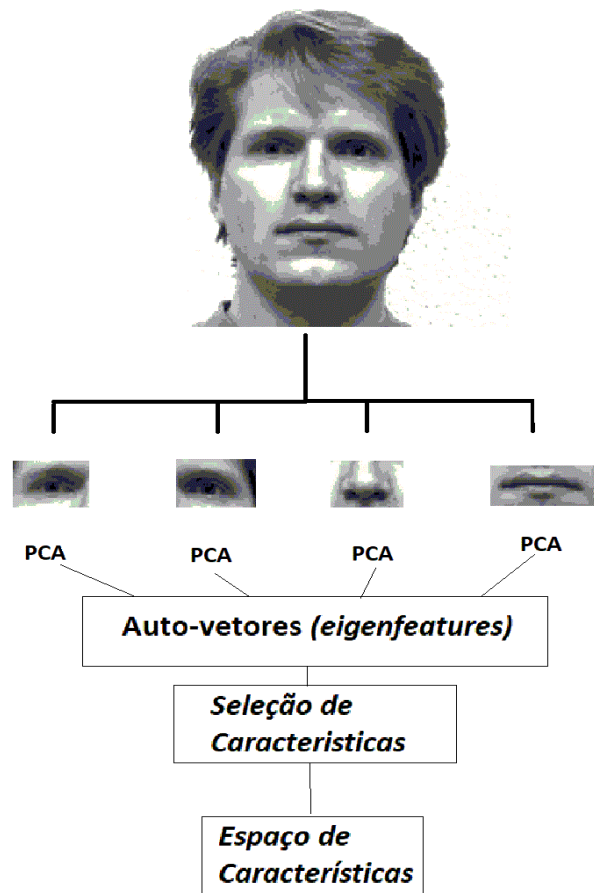


Figura 20 - Geração do espaço de características.

Fonte: Adaptado de Teófilo Emídio de Campos (2000).

Quadro 2 - Algoritmos de extração de características

Preservação	Técnica	Algoritmo
Estrutura Global	Linear	Fisher Linear Discriminant Analysis
		Principal Component Analysis
	Não Linear	Kernel Fisher Linear Discriminant Analysis
		Kernel Principal Component Analysis

Estrutura Local	Linear	Locality Preserving Projections
	Não Linear	Isomap
		Laplacian Eigenmap

Fonte: Rozario et al. (2014).

2.5.4 Correspondência de Características de Classificação

O processo de correspondência consiste em avaliar se as características de uma imagem de entrada com uma face, correspondem a uma imagem da mesma cena contida na base de imagens. Após a extração das características da imagem de entrada é realizado o confronto para estabelecer algumas correspondências preliminares entre essa imagem com as contidas na base de dados. Um problema enfrentado pela correspondência, que exige muita cautela dos algoritmos classificadores, é a classificação de múltiplas faces como faces verdadeiras correspondentes (HASSABALLAH et al., 2016).

O reconhecimento pode ser aplicado de dois modos, verificação (autenticação) onde é realizado o confronto de um-para-um, ou identificação que realiza o confronto de um-para-muitos. O processo de verificação ocorre quando um usuário precisa ser autenticado de acordo com uma segunda validação, onde o sistema responde se o usuário é mesmo quem ele diz ser. Um exemplo de autenticação seria o acesso de uma pessoa a um ambiente com um crachá, onde o sistema responde se a face contida no crachá corresponde face do usuário. Já no processo de identificação o sistema captura uma face, retira as características protuberante e confronta os dados de entrada com os dados contidos na base de imagem, calculando matematicamente as distâncias de proximidade e equivalência das características de todas as imagens contidas no banco de dados. A face é classificada como verdadeira de acordo com a proximidade das características e os critérios de avaliação do algoritmo (BRAGA, 2013).

Os métodos mais básicos de classificação de similaridade entre as faces utilizadas pelos algoritmos classificadores são as fórmulas de distância Euclidiana, distância de Manhattan, distância Cosseno e distância de Mahalonobis, mas não se limitam há estas (SILVA, 2016). Algoritmos mais sofisticados se baseiam em densidade de probabilidade, redes neurais e SVM. Uso de algoritmos de classificação junto a bases de imagens faciais e os detalhes de execução serão apresentados e discutidos no capítulo 4.

2.6 NUANÇAS DO RECONHECIMENTO FACIAL

No processo de reconhecimento facial, pequenas alterações em características e manipulação dos dados de entrada, como ruídos, mudança do equipamento de obtenção de imagem digital e ambiente de captura podem interferir na classificação da mesma, alterando o desempenho do algoritmo e a confiabilidade do resultado.

2.6.1 Variação na Aparência da Face

Diferentes tipos de alterações podem ocorrer em uma imagem facial, como pose facial, expressão facial, iluminação, obstáculos na imagem, tipo de equipamento que fez a captura e a qualidade de captura dessa imagem (FAGERTUN, 2015). Essas características influenciam no resultado dos algoritmos classificadores. Por serem características que demandam mais processamento dos algoritmos, estas podem gerar pontos estáticos e estratégicos para o conhecimento facial (SHARIF et al, 2017).

2.6.2 Análise da Face no Espaço de uma Imagem

Uma imagem de entrada contém dimensão $N \times M$, porém somente um subconjunto de coordenadas contém uma face. Uma imagem é caracterizada pela sua dimensionalidade, que por sua vez é calculada com base no tamanho de suas linhas e colunas em quantidade de *pixels*. Como exemplo, podemos escolher uma imagem com 32×32 , linhas e colunas em *pixels* respectivamente, sua dimensionalidade será 1024 dimensões. Efetuando a redução da dimensionalidade a um nível aceitável, reorganizando os *pixels* na imagem (levando em consideração sua vizinhança), a imagem permanecerá com as mesmas características, desde que todas as outras imagens sejam rearranjadas de forma idêntica. Assim facilitando o processo de extração de características baseados em características tradicionais em reconhecimento (TURK, 2001). Por outro lado, faces podem não ser detectadas caso estejam contidas em um subconjunto pequeno de uma imagem que contenha uma dimensionalidade muito grande, ou seja, ocupem uma pequena porção das dimensões da mesma. Nestes casos os algoritmos classificadores não conseguem aplicar suas técnicas de extração de características para o reconhecimento.

2.6.3 Gerenciar Porções Não Lineares e Convexa da Imagem.

Para que seja feita a redução de uma imagem tronando o seu reconhecimento menos custoso é necessário que seja realizada a normalização da imagem, diminuindo a não linearidade e convexa da imagem. Isso pode ser obtido através das aplicações e dos algoritmos Análise de Componentes Principais (PCA) e Análise Linear Discriminante (LDA). As metodologias podem ser combinadas junto aos motores de classificação como Redes Neurais, *Support Vector Machine* (SVM) entre outros, utilizados para esse fim (FAGERTUN, 2015).

A Figura 21 representa o modelo de aplicação da linearização de uma imagem, onde no plano (a), apresenta todas as características extraídas de uma imagem, já no plano (b) foi realizada a normalização e extração das principais características. O gerenciamento de porções não lineares e a região convexa de uma imagem, consiste no gerenciamento da dispersão dos *pixels* presentes em uma imagem, de acordo com os níveis de intensidade cor, que durante o processamento da imagem esses *pixels* podem conter uma grande variedade de intensidades de cor, que quando plotados em um gráfico (ver Figura 21), esse *pixels* ficam dispersos em relação ao eixo limiar, onde existe a necessidade de transpor os *pixels* mais dispersos da imagem o mais próximos do limiar de características, deixando-os o mais agrupado possível, esta correção ocorre aplicando filtros que manipulam as intensidades das cores dos *pixels*, normalizando a imagem (FAGERTUN, 2015).

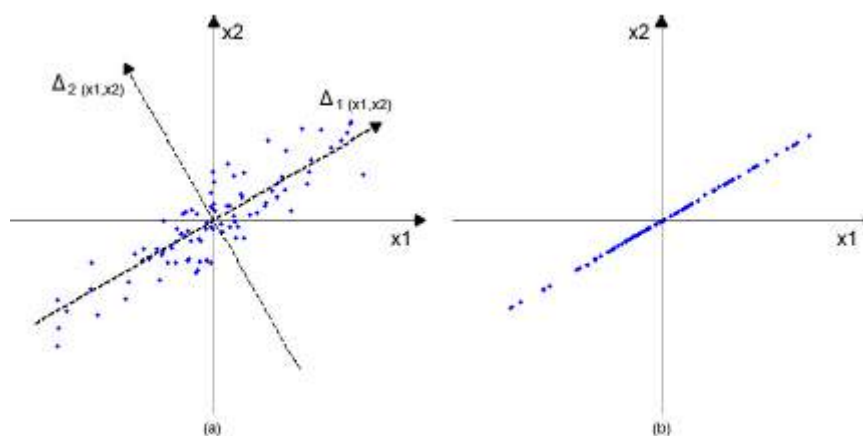


Figura 21 - Linearização Aplicada pelo algoritmo LDA.

Fonte: Braga, (2013)

2.6.4 Espaço de Alta Dimensionalidade e Baixas Amostras

Outro problema associado ao reconhecimento facial é a alta dimensionalidade das imagens de entrada e a quantidade das amostras de um indivíduo armazenadas na base de treinamento, que são geralmente pequenas, para serem comparadas. Com uma imagem de alta dimensionalidade os algoritmos tendem a executar mais tarefas para realizar a normalização das imagens de entrada, deixando-a com os mesmos padrões das imagens da base. Tão como uma pequena amostra de imagens de um indivíduo pode não ser o suficiente para que um algoritmo classificador seja capaz de fazer a sua aproximação com as características contidas na base de dados, aumentando a probabilidade de classificação de falso-positivo, e a diminuição de sua acurácia, onde por exemplo, dependendo da posição da face do indivíduo em relação a captura pode não ser detectada ou identificada (FAGERTUN, 2015).

3 TRABALHOS RELACIONADOS

A metodologia deste trabalho consiste em desenvolver uma síntese dos processos de reconhecimento facial enquanto é realizada uma explicação matemática junto de demonstração prática com os algoritmos descritos no capítulo da metodologia. Para melhor compreensão do desenvolvimento desse trabalho e sua aplicação, foram selecionados alguns trabalhos que são descritos nas próximas sessões, que são relacionados ao trabalho desenvolvido.

Uma das técnicas mais difundida na literatura no processo de detecção facial foi apresentada por Viola e Jones (2001), um algoritmo para detecção de objetos e faces baseado na seleção de recursos de *haar*, na integral da imagem, treinamento de classificadores de aprendizado supervisionado e classificadores em cascata.

A integral da imagem caracteriza-se pelo cálculo de intensidade de cinza sobre um subconjunto de pixels contidos em uma imagem digital. A partir da integral da imagem pode-se identificar padrões de características haar-like, que são máscaras retangulares aplicadas sobre todas as áreas de uma imagem de modo sequencial, comparando a intensidade luminosa dos pixels presentes em cada área. As máscaras aplicadas sobre a imagem digital realiza a varredura utilizando diversos classificadores fracos que realizam uma votação majoritária sobre as características de uma determinada área, se pode ou não existir uma face na região verificada, caso mais de 50% dos classificadores votarem a favor de conter uma face na região verificada, a região que possivelmente possui um uma face é delimitada.

Técnicas de aprimoramento do reconhecimento através da face é alvo de vários estudos para aprimoramento e aplicação em diversas áreas, como segurança e entretenimento. Uma abordagem voltada para a segurança é o alvo do estudo realizado por Silva (2015), onde destaca um estudo comparativo entre o comportamento de algoritmos sobre a identificação de indivíduos, com a intenção de aplicação em sistemas de autenticação de pessoas à ambientes restritos, levando em consideração intensidades de luz, expressões faciais, ângulo e oclusões sobre a face.

O trabalho científico de Zhao et al (2011) é baseado sobre um novo sistema de captura de objetos em vídeo, que utiliza uma estrutura de aprendizagem estatística para o processamento de objetos em vídeo. Utilizando um método de detecção de objetos em vídeo extraindo um conjunto de características em um espaço de tempo, onde são comparados através de um método projetado baseado na estratégia de aprendizagem supervisionada, *Support Vector Machine* (SVM), que classifica os objetos de acordo com os padrões das classes treinadas. Também uma

estrutura de classificação multi-classe para manipular os dados que não estiverem normalizados.

No trabalho de Prado (2018) é realizado uma comparação de técnicas de reconhecimento facial para aplicação em ambientes reais e semi-controlados. O autor realiza a comparação de técnicas de reconhecimento facial conhecidas na literatura, aplicando sobre um ambiente semi-controlado, considerando variações nas iluminações e a qualidade das imagens de entrada, desenvolvendo técnicas de aplicações de ruídos sobre as imagens da base de treinamento, para que fiquem de acordo com as imagens de captura realizadas pela câmera, para identificação de pessoas que acessam uma determinada área restrita.

4 METODOLOGIA

Neste capítulo será abordada a metodologia necessária para construção de um sistema de reconhecimento facial. Este processo foi dividido em cinco principais atividades, que são executadas sequencialmente e que contém suas próprias subatividades. O objetivo é detalhar o processo de forma lógica e com foco no funcionamento dos algoritmos envolvidos, seus conceitos matemáticos e as operações que envolvem os métodos de visão computacional.

As atividades necessárias para desenvolver uma abordagem de reconhecimento facial é construção de uma Base de Imagens; Detecção de Faces; Pré-Processamento; Extração de Características e Reconhecimento de Faces. Essas quatro atividades são representadas na Figura 22, que mostra o processo completo de reconhecimento facial iniciando pela detecção facial e terminando quando uma face é reconhecida ou não pelo sistema.

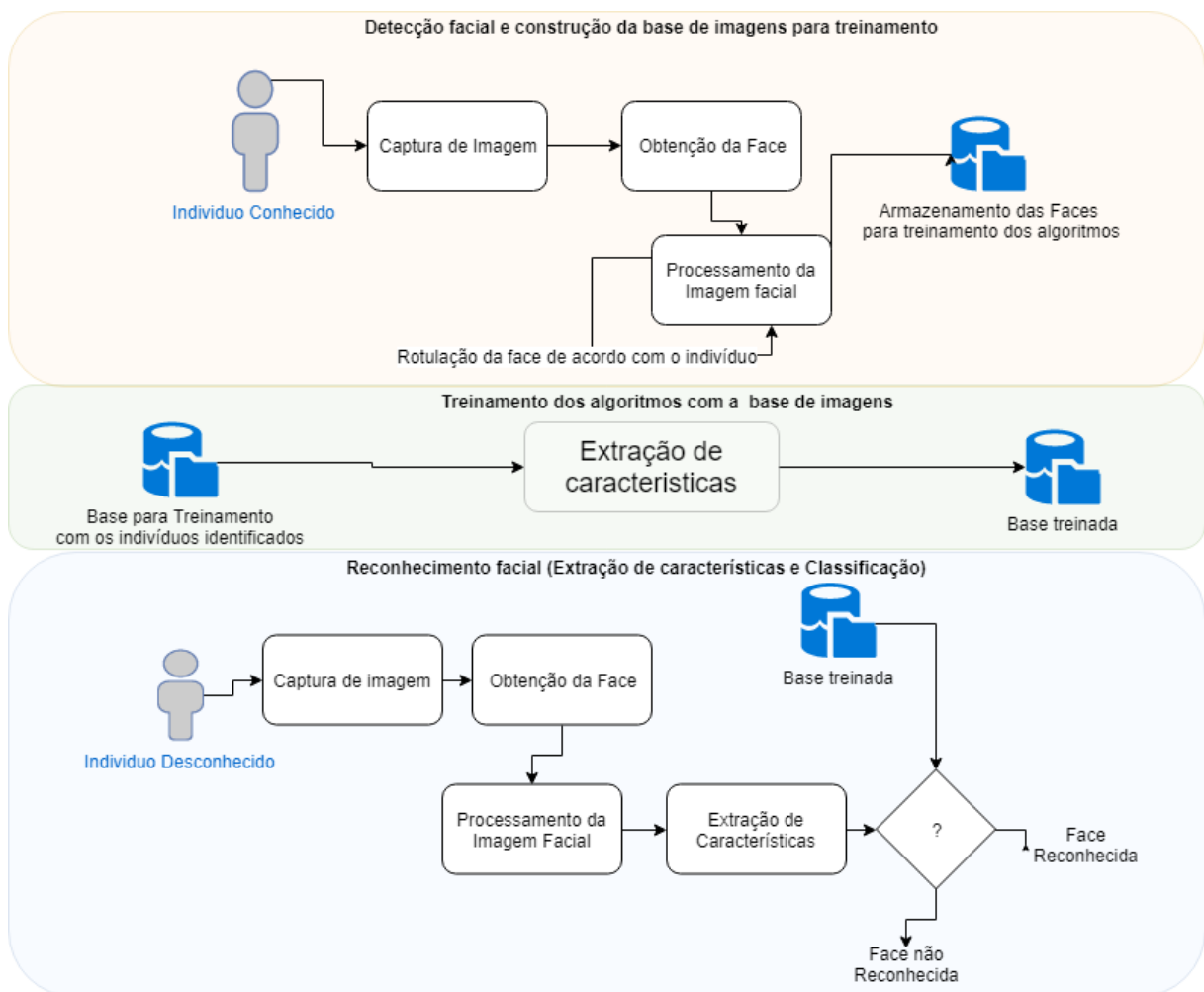


Figura 22 - Processo de detecção e reconhecimento facial.

As atividades de construção de uma Base de Imagens e Detecção de Faces são representadas pela parte posterior e central da Figura 22, a detecção é executada primeiro com o objetivo de colher amostras dos indivíduos que devem ser reconhecidos pelo sistema. Em seguida, as amostras são organizadas e armazenadas gerando a base de imagens, que será utilizada pelos algoritmos de extração de características para escolha dos atributos mais relevantes e treinamento. Na parte inferior da Figura 22 encontram-se as atividades de Pré-Processamento, Extração de Características e Reconhecimento de Faces. O pré-processamento das imagens de faces já pode ser feito na etapa de detecção facial, sendo que as mesmas técnicas geralmente são utilizadas em ambas as fases. A atividade de Pré-Processamento prepara a base de imagens para que a Extração de Características e Reconhecimento seja bem-sucedida. Nessa última atividade ocorre a redução de dimensionalidade das imagens da base, bem como o treinamento do algoritmo que irá reconhecer cada face. Os detalhes de cada atividade e os algoritmos utilizados serão discutidos e apresentados no decorrer deste capítulo.

4.1 BASES DE IMAGEM

Para realizar o processo de reconhecimento facial é necessária uma base de imagens com ao menos uma face, para assim realizar o treinamento dos algoritmos que posteriormente executaram as tarefas de comparação entre faces, como descrito no capítulo 2.4.

Para a realização do reconhecimento facial é indispensável a utilização de imagem faciais dos indivíduos que se deseja identificar, para o treinamento dos algoritmos, de preferência que estas imagens que serão utilizadas para a construção das bases de imagens contém apenas imagens faciais, eliminando partes das imagens que não são essenciais para o reconhecimento, como por exemplo o fundo da imagem. O processo de treinamento do algoritmo exige ao menos uma imagem facial de um indivíduo, onde esta imagem seja previamente identificada, para que no momento do treinamento da base o algoritmo possa etiquetar esta imagem com uma identificação. A construção da base de imagem é importante pelo fato de que o algoritmo utilizará as imagens de treinamento para comparar com as novas imagens de entrada (teste), onde serão confrontadas as características e extraídas as semelhanças entre as imagens.

Deve-se ter o cuidado com as imagens de treinamento, que claramente devem ser diferentes das imagens utilizadas para teste, pois não deve ocorrer a interseção entre as imagem para realizar análise de performance dos algoritmos.

Para a metodologia deste trabalho foram selecionadas duas bases de imagens conhecidas na literatura *Yale Database* e *AT&T Database of Faces*. Além destas, uma terceira base foi gerada a partir de modificações com inserção de pontos de oclusão sobre as faces da *Yale Database*. A base de imagens da *Yale* foi escolhida por representar faces sem pose, sem oclusão e com boa localização, representando um teste limpo para todos os algoritmos. A *AT&T Database of Faces* será utilizada para testar os algoritmos em relação a lidar com as complicações de diferentes poses para um mesmo indivíduo. A terceira base é composta por inserção de faixas preta que irão ocluir olhos ou boca de algumas das poses faciais da *Yale Database*, com o objetivo de testar a performance dos algoritmos sob condições de oclusões. A seguir são descritas as bases escolhidas para a metodologia deste trabalho.

- **Yale Database:** Contém 165 imagens em escala de cinza de 15 pessoas diferentes, que foram obtidas de diferentes ângulos, com diferentes expressões faciais e graus de luminosidade. Todas as imagens são padronizadas no tamanho 320 x 243 *pixels* com fundo em tom de branco, como mostra a Figura 23.

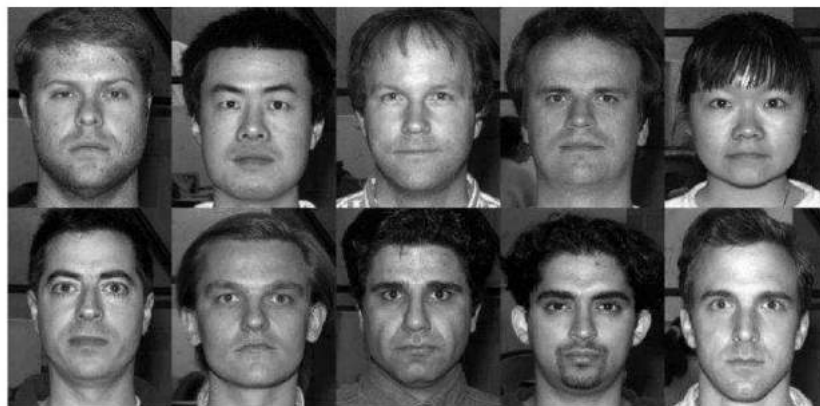


Figura 23 - Exemplo de faces da Yale Database.

Fonte: Bissi et al. (2018).

- **AT&T Database of Faces:** Contém 400 imagens em escala de cinza de 40 pessoas diferentes, que foram obtidas de diferentes ângulos com diferentes expressões faciais. As variações dos graus de luminosidade são de diferentes horários do dia, padronizadas no tamanho 92 x 112 *pixels*, como exibido na Figura 25.

- **Yale Database Modificada:** A principal **modificação** está na aplicação de pontos de oclusão sobre uma região de cada imagem (sobre a boca), com o intuito de verificar o comportamento dos algoritmos perante a presença de alguns obstáculos na imagem facial. Os pontos de oclusão serão aplicados de forma percentual, sobrescrevendo os *pixels* em pontos aleatórios da imagem. A Figura 24 mostra um exemplo da aplicação das oclusões sobre 15% dos *pixels* da imagem facial. Para que o a modificação fosse mais assertiva, o local onde seria aplicado a oclusão foi removido da imagem original e desenhado uma tarja preta no lugar.



Figura 24 - Imagem da base *Yale Face* com 3%, 9% e 15% de oclusão respectivamente.

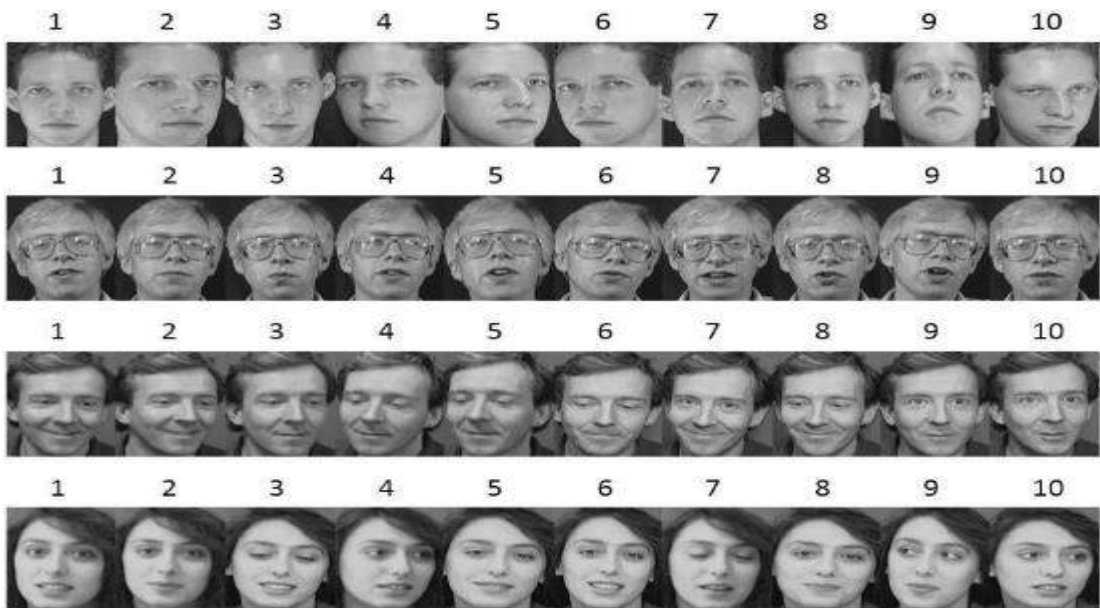


Figura 25 - Exemplo de faces *AT&T Database of Faces*.

Fonte: Braga et al. (2013).

4.2 DETECÇÃO FACIAL

A Detecção facial é um processo básico e fundamental do reconhecimento facial, como descrito anteriormente no capítulo 2. A detecção consiste em encontrar uma ou mais faces em uma imagem ou figura com pessoas e outros objetos em um mesmo ambiente ou cenário. Nos experimentos deste trabalho não se fez necessário a aplicação do reconhecimento facial, pois as bases utilizadas já possuem as imagens faciais separadas e catalogadas. Devido a importância dessa atividade para o processo de reconhecimento facial, será apresentada uma técnica fundamental utilizada na detecção facial, bem como sua parametrização.

O algoritmo mais conhecido para o processo de detecção facial é o de Viola e Jones (VIOLA e JONES, 2001), que está disponível na biblioteca OpenCV. O processo de detecção de Viola e Jones utiliza como principal função a `detectMultiScale`, que é responsável por detectar objetos incluindo faces. O algoritmo realiza uma análise da imagem atrás de padrões que caracterizam um objeto, quando este trata-se de uma face, a função utiliza como parâmetro as características deste objeto tal como a região dos olhos. As regiões precisam apresentar características que a distinguem de outras, como os olhos por exemplo, que geralmente estão em uma posição do rosto que é mais escura que as bochechas. Estes tipos de características são conhecidas como recursos Haar, como demonstrado na Figura 26.

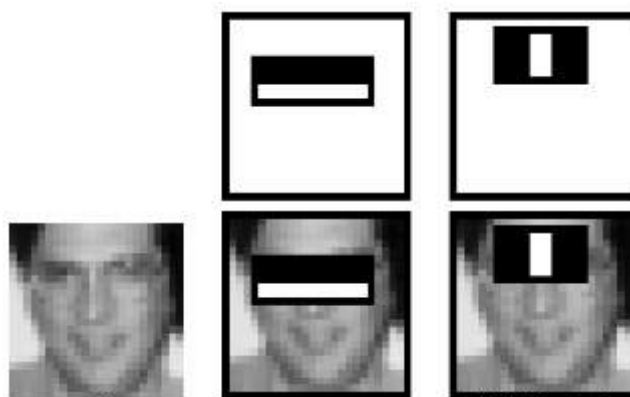


Figura 26 - Exemplo Aplicando recursos de Haar sobre uma imagem facial.

Fonte: Opencv (2018).

Desta forma mapeando dentro de uma imagem digital se existe uma face, de acordo com os padrões de *Harr*, delimitando esta região. O algoritmo de Viola e Jones (2001) é baseado em

três conceitos, a integral da imagem treinamento de classificadores em cascata usando *boosting* e o uso destes em cascata. As *features* retangulares como são denominadas as unidades básicas dado métodos de Viola e Jones (2001), possuem formatos específicos que caracterizam os possíveis padrões de intensidade dos *pixels* em uma imagem. A Figura 27 representa as quatro configurações possíveis de uma *feature*, verificada dentro de uma imagem W , possuem características padrões, mas alteram sua escala durante a verificação.

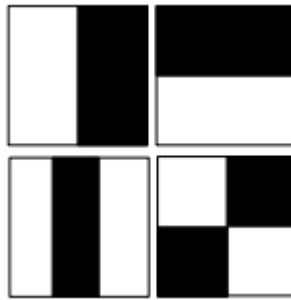


Figura 27 – Quatro padrões de configuração de *features*.

Fonte: Opencv (2018).

O valor de uma *feature* em uma imagem W é calculada de forma onde o valor de uma *feature* em uma imagem W é igual ao somatório de todos os (P) *pixels* na região preta, subtraído o somatório de todos os (P) *pixels* na região branca de uma imagem. O cálculo é realizada sobre todos os *pixels* de forma linear de acordo com Equação 1.

$$f(W) = \sum^W P_{preto} - \sum^W P_{branco} \quad (1)$$

Para acelerar o cálculo das *features*, o método utiliza o cálculo da integral das imagem definida pela Equação 2:

$$ii(x, y) = \sum_{x \leq x', y \leq y'} i(x', y') \quad (2)$$

Onde $i(x', y')$ é o valor do *pixel* da imagem na coluna x' e linha y' e $ii(x, y)$ é o valor da integral da imagem até suas colunas x e linha y . Uma representação da soma dos valores dos *pixels* é demonstrado na Figura 28, onde a região D é calculada utilizando a integral da imagem.

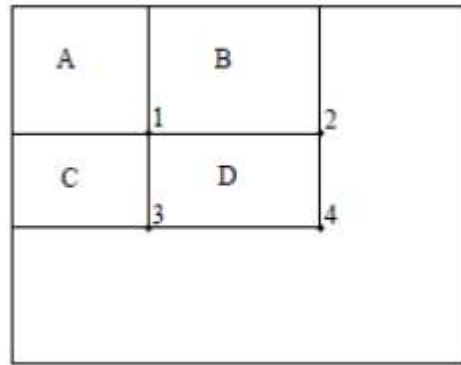


Figura 28 – Soma dos *pixels* na região *D* utilizando a integral da imagem.

Fonte: Viola e Jones (2001).

A representação matemática do cálculo da região *D* é de acordo com a Equação 3:

$$\sum^D pixels = ii(4) + ii(1) - (ii(2) + ii(3)) \quad (3)$$

A classificação de uma *feature* dentro de uma imagem se dá pela classificação pelos chamados classificadores fracos. Representados pela função representada na Equação 4:

$$h(w, p, f, \theta) = \begin{cases} 1, & \text{se } pf(w) < p\theta \\ 0, & \text{se } pf(w) \geq p\theta \end{cases} \quad (4)$$

onde, *w* é a sub-janela de 24×24 *pixels*; *f* é o número de *features*; *p* é a intensidade do *pixel*; θ é limite (*threshold*) de distância para estimar a verossimilhança.

Para a classificação o algoritmo utiliza os classificadores fracos em conjunto formando classificadores fortes. Por fim, constrói-se uma cascata de classificadores onde, cada nível possui um conjunto distinto de classificadores fortes. A Figura 29 esquematiza o fluxograma realizado pelo algoritmo de Viola e Jones (2001), no processo de classificação. Para que o algoritmo tenha mais eficiência ele realiza o treinamento dos classificadores fracos utilizando o algoritmo AdaBoost, que realiza a combinação de classificadores baseados em *boosting* que é a combinação de vários classificadores fracos com acerto superior a 50%, que além de aprender a classificação escolhe quais características são mais adequadas para o uso.

4.3 PRÉ PROCESSAMENTO DE IMAGENS FACIAIS

A etapa de pré-processamento da face visa a normalização da mesma, isto é, a redução das variações das imagens obtidas no processo de detecção facial. Nessa etapa são configurados

os padrões das imagens que estarão na base de treinamento e nas imagens de teste. As etapas mais básicas do pré-processamento são: a conversão das imagens em escala de cinza; o redimensionamento das imagens para que estejam no mesmo padrão da base de treinamento.

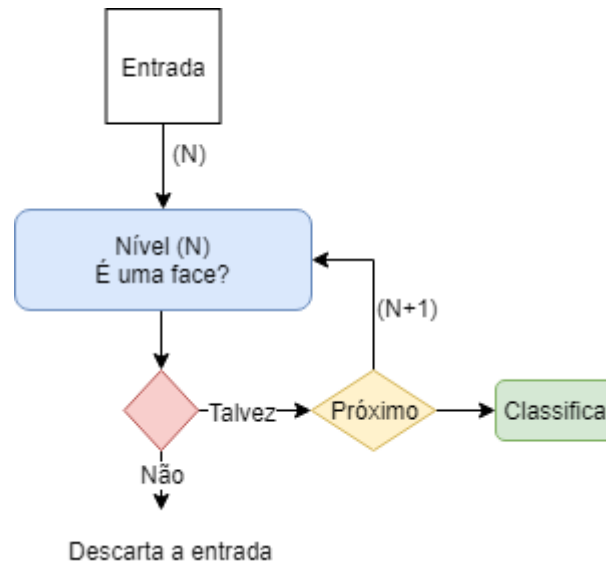


Figura 29 - Exemplo de funcionamento da cascata dos classificadores.

- **Transformação em escala de cinza:** A biblioteca Opencv disponibiliza de uma função de conversão da imagem em escala de cinza, chamada *cvtColor*, que leva por parâmetro “COLOR_RGB2GRAY”, que realiza a conversão dos valores RGB de um *pixel* i para uma escala de cinza correspondente (0 a 255). A conversão para a escala de cinza é realizada de acordo com a Equação 5 a seguir (OPENCV, 2000).

$$Y_i = 0.299 * R + 0.587 * G + 0.114 * B \quad (5)$$

- **Redimensionamento da imagem:** Para que as imagens sigam um padrão é necessário o redimensionamento das imagens para as mesmas dimensões, a biblioteca opencv também disponibiliza de uma função chama “*resize*”, que redimensiona uma imagem para um tamanho X_i e Y_i especificados (OPENCV, 2000).

Como mencionado anteriormente estes procedimentos, não serão aplicados pois as bases de imagens utilizadas já estão dimensionadas e em escala de cinza, com o mesmo padrão. Em uma aplicação onde se realiza o processo de detecção facial, o pré-processamento de uma imagem é indispensável, pois a imagem resultante do processo de detecção pode apresentar diversas imperfeições, tais como: a presença de *pixels* com ruído, contraste ou brilho inadequados,

entre outros problemas que pode sofrer uma imagem. A função do pré-processamento é aprimorar a qualidade da imagem para as etapas subsequentes, estas operações são ditas de baixo nível pois trabalham diretamente sobre os valores dos *pixels*, sem nenhum conhecimento sobre a imagem como todo (GONZALEZ; WOODS, 2006). As técnicas de pré-processamento possuem diversas aplicações, dentre elas estão aplicações como filtros de realce, suavização e segmentação.

4.3.1 Realce

A técnica de realce é responsável por melhorar a qualidade visual de uma imagem, aumentando o contraste entre os elementos da imagem. Existem diversas técnicas de realce de imagem, uma das mais conhecidas na literatura é o realce por histograma, que consiste em um histograma diretamente relacionado a frequência que cada valor de escala de cinza que se apresenta sobre a imagem digital, como na Figura 30.

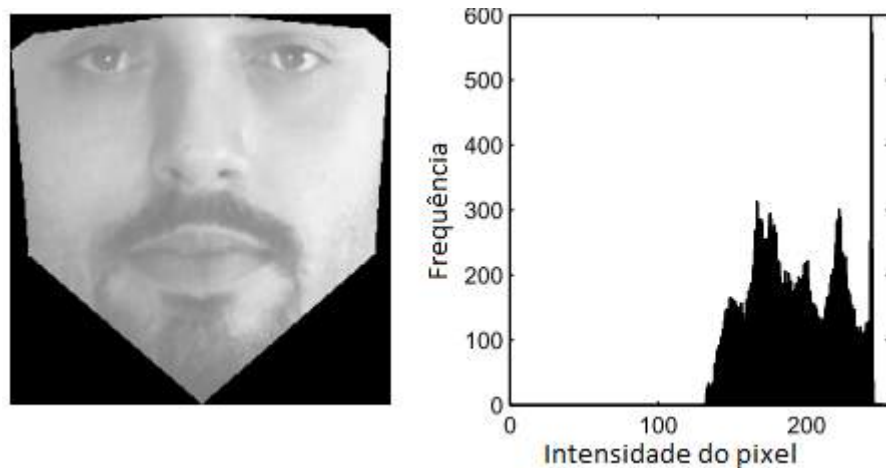


Figura 30 – Histograma dos *pixels* contidos em uma imagem em escala de cinza.

Fonte: Adaptada de Fagertun (2005).

O contraste por histogramas é realizado através de uma função de transformação radiométrica, que mapeia a faixa de valores dos *pixels* de uma imagem original para uma nova faixa de valores radiométricos mais bem distribuídos em um novo domínio, utilizado para aumentar o contraste da imagem. A técnica de realce por histogramas pode ser aplicado de duas formas, lineares ou não lineares. A Figura 31 é representada por uma imagem digital obtida através de um dispositivo de captura, porém sem nenhum filtro sobre a imagem.

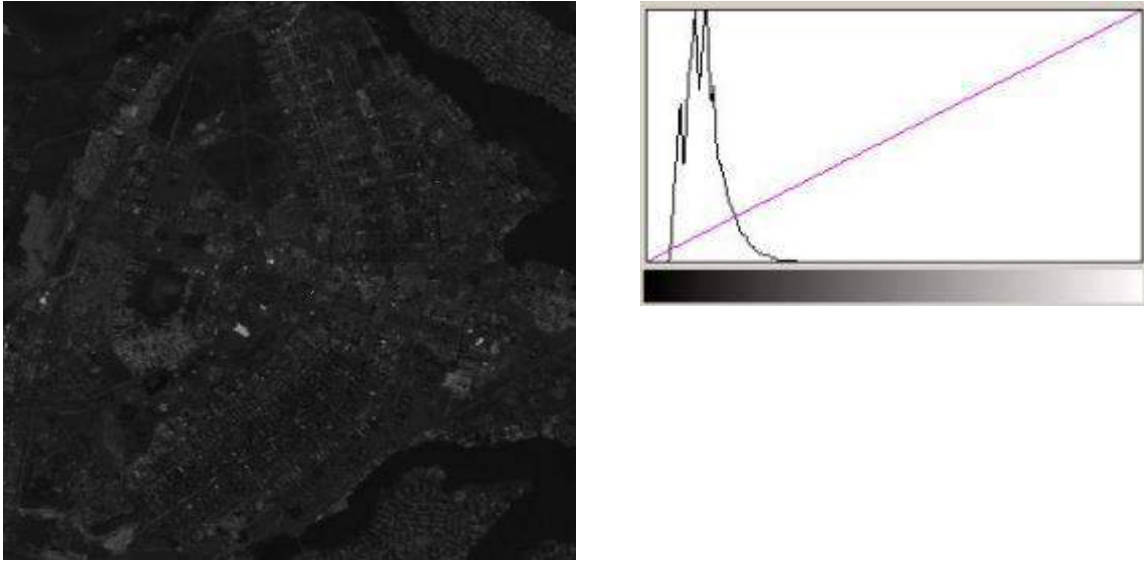


Figura 31 – Imagem em escala de cinza antes da aplicação de contraste linear.

Fonte: adaptado de DPI (2019).

A Figura 32 demonstra uma aplicação de realce linear por manipulação de histogramas onde uma imagem com radiometria de 8 *bits*, tem os seus *pixels* alterado para uma nova faixa radiométrica com valor entre 0 e 255 em níveis de cinza.

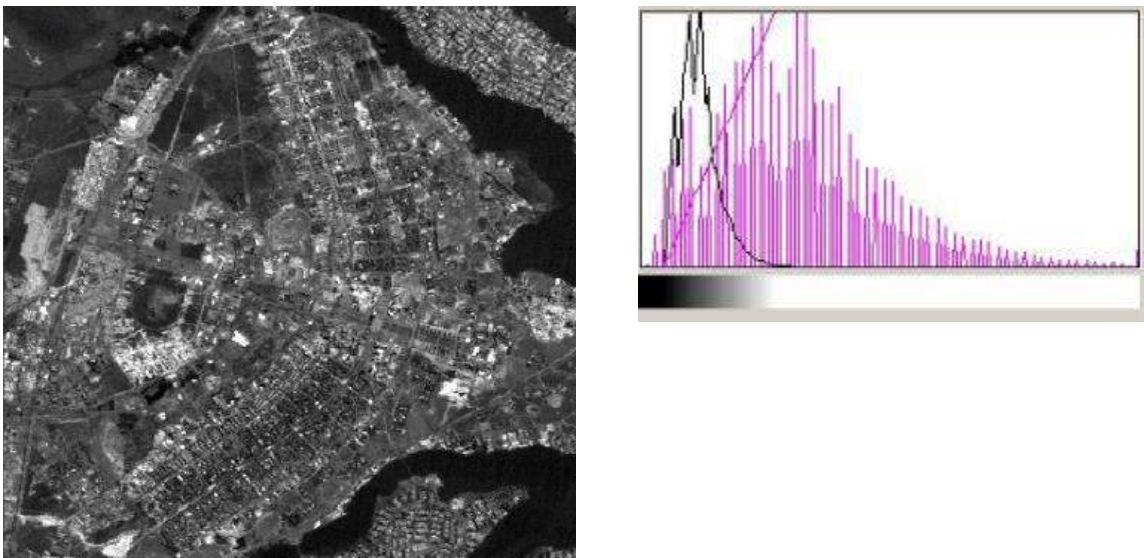


Figura 32 - Imagem em escala de cinza após a aplicação de contraste linear.

Fonte: adaptado de DPI (2019).

A realização do realce linear, se dá pela Equação 6:

$$Y = a * X + b \quad (6)$$

Onde para determinar a transformação digital da imagem para um novo domínio, o processo de transformação é calculado da seguinte forma:

- Encontra-se o valor de X , que é representado pelo $Xmin$ e $Xmax$ dos valores digitais baseados no histograma da imagem.
- Calcula-se o valor de a , pela fórmula: $a = (255/(Xmax-Xmin))$;
- O parâmetro b é encontrado pela relação: $b = -a*Xmin$.

Assim aplica-se a equação (6) para cada um dos *pixels* digitais da imagem de entrada, gerando uma nova imagem de saída.

Os realces não lineares aplicados através de histogramas, tem como objetivo realçar valores mais escuros de uma imagem de acordo com a Equação 7.

$$Y = f(x) = a * x^{\frac{1}{2}} \quad (7)$$

A Figura 31 representa a aplicação de uma transformação não linear onde são destacados os pontos mais escuros em uma imagem digital.

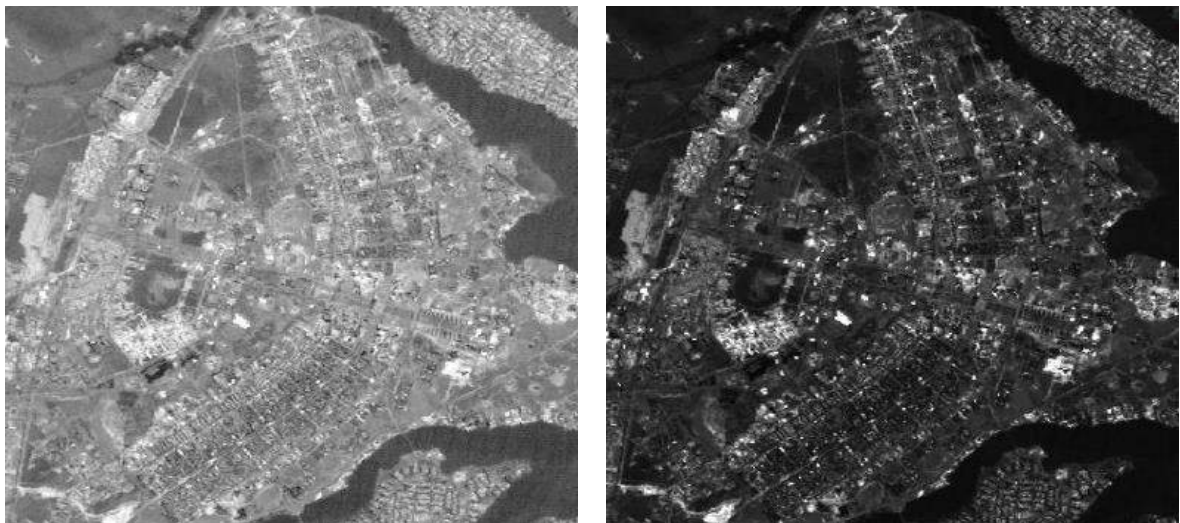


Figura 33 – Aplicação de realce não linear por histogramas.

Fonte: adaptado de DPI (2019).

4.3.2 Suavização

As técnicas de suavização são responsáveis por realizar uma varredura sobre a imagem digital a procura de extração de informações, possibilitando a correção de imperfeições presentes dentro de uma imagem digital como por exemplo ruídos ou borrões que foram criados no momento da captura da imagem. Os filtros de suavização estão divididos em dois grupos, lineares e não lineares.

Os filtros de suavização lineares são classificados como filtros passa-baixa e passa-banda. A aplicação de filtros de suavização, tem a finalidade de dar um aspecto desfocado sobre a imagem, assim se consegue reduzir a diferença entre um *pixel* central em relação a seus vizinhos. A aplicação de suavização tem como resultado sobre o *pixel* processado a média da intensidade da cor de cinza do *pixel* em relação ao *pixel* vizinho, assim reduzindo o nível de ruído, mas se aplicada a suavização exageradamente pode ocorrer a distorção da imagem, alterando-a muito em relação a original (GONZALEZ e WOODS, 2002).

4.3.3 Segmentação

Segmentação é um dos processos mais importantes do processamento de imagem pois o mesmo é responsável pela separação de características dentro de uma imagem digital, separando os objetos contidos em uma imagem em diferentes conjuntos de *pixels*. A segmentação é responsável por encontrar objetos e formas em uma imagem digital. Um dos processos mais comuns de segmentação é o *Thresholding*, que analisa cada *pixel* dentro de uma imagem digital transformando a imagem de entrada em uma imagem binária, composta apenas por *pixels* da cor branca e preta, de acordo com um limiar proposto. A representação matemática do processo é dada pela Equação 8:

$$P = \begin{cases} 1, & \text{se } P_{i,j} \geq T \\ 0, & \text{se } P_{i,j} < T \end{cases} \quad (8)$$

Onde T é uma constante previamente determinada no início da execução do algoritmo e $P(i,j)$ é o *pixel* analisado e P é no novo *pixel*. A Figura 34 demonstra um exemplo de aplicação dos filtros de segmentação por *Thresholding*.

4.4 EXTRAÇÃO DE CARACTERÍSTICAS DA FACE E CLASSIFICAÇÃO

O processo de extração de características e a classificação são duas atividades diretamente relacionadas, que executam o reconhecimento de uma face. Isso ocorre, pois existem

abordagens que durante a extração já desenvolvem um modelo capaz de reconhecer faces, desde que estas sejam relacionadas a indivíduos da base e sejam submetidos ao mesmo processo de extração. Além desses, existem os algoritmos que após a extração utilizam as características reduzidas como parâmetro para outro algoritmo de classificação. Este por sua vez é treinado com os parâmetros extraídos e utilizando uma função de similaridade consegue estimar as diferenças da face de um novo indivíduo com aqueles que estavam na base, reconhecendo ou não a face. Em ambos os casos as atividades de extração de características e a classificação possuem forte correlação, onde seus respectivos resultados de saída são utilizados em conjunto para reconhecer faces.



Figura 34 – Aplicação de segmentação por *Thresholding*, onde a esquerda está a imagem original e a direita a nova imagem a partir da segmentação.

Fonte: Mangayarkarasi e Naveena, 2014.

A seguir serão descritos três algoritmos de grande relevância na literatura e que possuem estratégias distintas para a atividade de extração de características, são eles *Eigenfaces*, *Fisher-Faces* e *LBPH*. Em seguida, serão apresentados as principais funções utilizadas para o cálculo de similaridade entre as características extraídas de uma base de imagens e uma nova face para a classificação e reconhecimento da mesma.

4.4.1 Extração de Características

Nas seções seguintes são apresentados os processos realizados pelos algoritmos *Eigenfaces*, *Fisherfaces* e *LBPH* para a extração de características, detalhando cada um dos processos matematicamente representados. O *Eigenfaces* e *Fisherfaces* utilizam técnicas específicas para extrair as características mais importantes de uma imagem e ao mesmo tempo reduzir sua di-

mensionalidade. Essas técnicas são o *Principal Component Analysis* (PCA) e o *Linear Discriminant Analysis* (LDA). O PCA é um procedimento estatístico, utilizado na criação de métodos preditivos, que permite realizar a análise de vários fatores juntos, projetando um conjunto de vários dados contidos em uma matriz de dados originais multidimensional em um novo espaço dimensional menor, selecionando apenas as principais características para realizar a análise. PCA é baseado nas características de autovetores para explicar o grande número de variáveis lineares não correlacionadas, onde são separados os primeiros principais componentes para análise, variáveis chamadas de eixos principais (PEARSON, 1901).

O LDA é um método utilizado na estatística para aprendizado de máquina e reconhecimento de padrões, onde procura-se encontrar as combinações lineares das características ou realizar a separação de classes. Comumente é utilizado como classificador linear ou na redução de dimensionalidade para um classificador (FISHER, 1936). LDA se assemelha com os métodos estatísticos de análise de variância e análise de regressão, também está relacionado ao método PCA, que buscam variáveis que melhor explicam um conjunto de dados. LDA trabalha sobre um conjunto de dados rastreando as características mais importantes de um conjunto de dados reorganizando-os em novos subconjuntos, conhecidos como classes, tornando os dados mais próximos na nível intraclasses e mais distantes entre as classes (JOHNSON; WICHERN, 1999).

4.4.1.1 Eigenfaces

Eigenfaces é uma técnica melhorada do algoritmo PCA. A técnica desenvolvida por Sirovich e Kirby (1987), funciona de forma análoga ao PCA, porém é utilizado uma leve otimização para reduzir a matriz de covariância, consequentemente reduzindo o processamento para calcular os auto vetores e autovalores (DINIZ et al., 2011). Este método consiste na extração de características mais importantes de uma imagem facial, para posteriormente compará-la com as imagens já classificadas anteriormente, que estão na base de treinamento. O algoritmo *Eigenfaces* tem o intuito de diminuir o número de características presentes em uma imagem para um subespaço de menor dimensionalidade.

O método *Eigenfaces* é formado por 7 etapas:

- Cálculo da face média;
- Subtração pela face média;
- Transformação das imagens em vetores;

- Construção da matriz de covariância;
- Cálculo das *Eigenfaces*;
- Cálculo dos vetores de características;
- Cálculo de similaridade.

Matematicamente, o algoritmo *Eigenfaces* realiza a extração das principais características (PCA), gerando uma face média com a combinação linear de diversas *Eigenfaces*. Em uma imagem de tamanho $N \times N$, o algoritmo considera a imagem como sendo um vetor de dimensão N^2 , onde busca os autovetores da matriz de covariância do conjunto de imagens. Conforme ocorre o aumento dos números de imagens de um mesmo indivíduo, aumenta a exatidão do processo, pois eleva-se o número de informações contidas nos *Eigenvectors*.

Para a geração da face média, no processo de treinamento, considera-se a Equação 9:

$$\Psi = \frac{1}{M} \sum_{i=1}^M \Gamma_i \quad (9)$$

Considerando M o conjunto de imagens faciais na base de treinamento supervisionada, e Γ como vetor da imagem analisada, tem como resultado um vetor médio Ψ , que representa a “face média” das imagens utilizadas para treinamento do algoritmo, como na Figura 35.



Figura 35 - Ilustração da geração da face média.

Após o cálculo da face média na Equação 9, é gerado um novo vetor Φ , subtraindo Ψ de cada imagem Γ como mostra a Equação 10 e esquematizado na Figura 36.

$$\Phi = \Gamma - \Psi \quad (10)$$



Figura 36 - Subtração da face média

Realiza-se o cálculo da matriz de covariância (C) Equação 11, pois a dimensionalidade de M é maior que a dimensionalidade de Φ .

$$C = \Phi^T \Phi \quad (11)$$

Calcula-se os autovetores v e autovalores λ da matriz C , e em sequência cria-se a matriz de transformação utilizando a Equação 12.

$$\mu = v\Phi \quad (12)$$

A matriz μ conterá $M-1$ autovetores significativos. Entretanto caso $M-1$ ainda seja muito grande, é possível realizar a eliminação de alguns desses auto vetores pela ordem de importância de seus correspondentes autovalores λ , gerando M' auto vetores selecionados. Para finalizar o cálculo de treinamento, é realizada a normalização de M' vetores da matriz μ .

Para a realização do reconhecimento é realizado o seguinte, dada uma imagem de teste Y , projeta-se a imagem junto ao auto espaço, centralizando-a com a face média gerada no treinamento, então é realizada a classificação usando classificadores definidos, ou N classificadores combinados. A Face é classificada calculando a distância da imagem matriz μ_Y de teste com os valores da matriz μ de treinamento, quando a distância é mínima ou abaixo do limiar pré-definido. Caso contrário a classe é classificada como não conhecida (TURK e PENTLAND, 1991).

4.4.1.2 Fisherfaces

O algoritmo *Fisherfaces* foi baseado nas técnicas do LDA, um método estatístico frequentemente usado para a redução e classificação de dimensionalidade. Assim como o *Eigenfaces*, o *Fisherfaces* se baseia na redução de dimensionalidade no espaço de características (Souza, 2014), porem consegue trabalhar melhor com em relação a luminosidade sobre as faces, segundo os estudos de Belhumeur, Hespanha e Kriegman (1997), o método *Fisherfaces* foi reconhecido como sendo mais eficiente que o *Eigenfaces* por ter taxas menores de erros, porem

apresenta alguns pontos negativos por ter dificuldades em tratar as variações de poses, que é um fator crucial em um ambiente de reconhecimento real.

O LDA realiza uma redução da dimensionalidade específicas das classes, maximizando a variância entre as classes e minimizando a variância dentro delas, maximizando a separação entre as classes. Após a rotulação da face do indivíduo o método agrupa as imagens em classes, tentando modelar a dispersão dos pontos visando maior confiabilidade (BELHUMEUR, 1997).

Para o funcionamento do algoritmo *Fisherfaces*, primeiramente é gerada um conjunto de imagem de baixa dimensionalidade utilizando o PCA, então aplica a técnica do LDA, aumentando o espaço dimensional de dispersão das classes e minimizando a dimensionalidade intraclasse. Por meio disto dada uma imagem Y de entrada, o reconhecimento facial é realizado com a projeção no subespaço das classes, utilizando uma medida de semelhança para comparar com as demais faces projetadas no subespaço (FIGUEREDO, 2011 apud SOUZA, 2014).

O processo de execução do algoritmo *Fisherfaces* são descritas matematicamente na sequência.

A matriz de dispersão intraclasse é representada pela formula (6), onde x_i^j é o i -ésimo exemplo da classe j , μ_j é a média da classe j , c é o número de classes e $|T_j|$ o número de exemplos na classe j , de acordo com a Equação 13.

$$S_w = \sum_{j=1}^c \sum_{i=1}^{|T_j|} (x_i^j - \mu_j) * (x_i^j - \mu_j)^T \quad (13)$$

A matriz de dispersão intraclasse é representada pela formula (7), onde μ representa a média de todas as classes, de acordo com a Equação 13.

$$S_b = \sum_{j=1}^c (\mu_j - \mu) * (\mu_j - \mu)^T \quad (14)$$

A maximização da média das interclasses e a minimização da intraclasse são obtidas ao maximizar a taxa $(\det(S_b) / \det(S_x))$. Com W sendo a matriz com auto vetores associados a matriz diagonal dos autovalores λ , calcula-se a projeção do no espaço dimensional, de acordo com a Equação 15:

$$S_b W = \lambda S_w W \quad (15)$$

onde, limita-se as matrizes a ordem de $c-1$, devido a comparação ser realizada entre duas classes diferentes, onde c é i número da classe.

O processo de identificação das imagens de teste, é realizado da mesma forma que o método *Eigenfaces*. A imagem de teste é projetada comparando-a com cada uma das imagens de treinamento também projetadas, identificando-a com a distância mínima entre as classes e abaixo do limiar definido. A comparação é feita utilizando um ou N classificadores combinados.

4.4.1.3 LBPH

O **Padrão Binário Local (LBP)** é um operador de textura simples, que identifica e seleciona os *pixels* de uma imagem limitando e convertendo sua vizinhança em um padrão numérico binário. Originalmente o LBP rotula os *pixels* de uma imagem no formato de matrizes, geralmente de tamanho 3×3 , onde o *pixel* central da matriz é o limiar para cálculo dos *pixels* vizinhos, os quais resultam em números binários.

Considerando um subconjunto de *pixels* de uma imagem com uma matriz de 3×3 *pixels* em escala de cinza (Figura 37a), representa-se a matriz de acordo com a intensidade da cor cinza sobre o *pixel* (Figura 37b), seleciona-se o *pixel* do interior da imagem rotulado como limiar. Então é realizado o cálculo convertendo os *pixels* vizinhos ao limiar, de acordo com a Equação 16:

$$X = \begin{cases} X < y, & X = 0 \\ X > y, & X = 1 \end{cases} \quad (16)$$

Onde, X é o *pixel* analisado e y o valor do *pixel* limiar, convertendo a vizinhança para uma sequência binária (figura 25c). Os valores binários resultantes, agora são concatenados formando uma nova sequência binária, valor este que convertido para o sistema decimal é atribuído ao *pixel* central da matriz, representando uma nova imagem em escala de cinza (Figura 37).

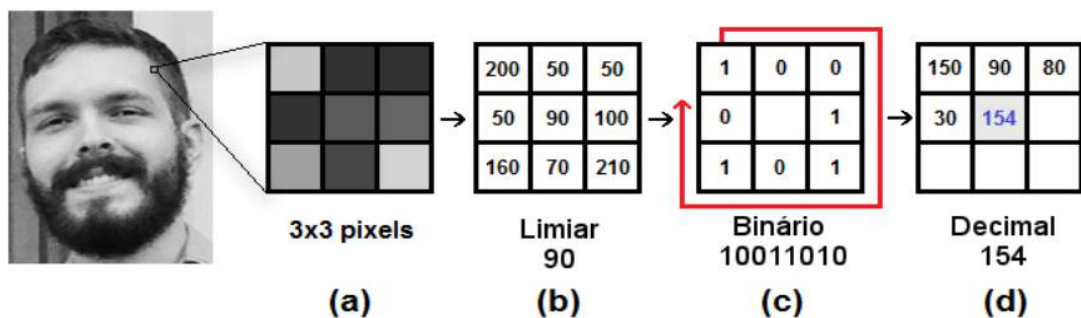


Figura 37 - Exemplo de operação com LBP.

Fonte: Salton (2016).

A Figura 38 representa o resultado da operação do LBP, onde a direita está a imagem original já normalizada e convertida para escala de cinza e a esquerda a nova imagem gerada a partir das operações realizadas pelo *LBP*.



Figura 38 - Exemplo de nova imagem gerada pelo LBP.

Fonte: Salton (2016).

Além de gerar uma nova imagem, ela é projetada em um histograma gerado a partir dos cálculos gerados, gerando um padrão de ocorrência dos resultados dos *pixels* em um histograma de cada operação, ao final é gerado um histograma completo da imagem concatenando estes histogramas locais, chamados histograma da imagem (AHONEN; HADID; PIETIKAINEN, 2006). Após a extração do histograma, a classificação é realizada pelo cálculo de similaridade dos histogramas.

O *LBPH* é o processo que é capaz de representar N histogramas a partir de imagens com maior dimensionalidade, reduzindo as características, em vetores de características relativamente menores. Capaz de realiza operações com diferentes tamanhos de vizinhança, como também, vizinhas circulares permitindo qualquer raio e número de vizinhança, aprimorando o operador LBP (AHONEN; HADID; PIETIKAINEN, 2004). A Figura 40, representa um exemplo do LBP, com vizinhança circular.

Para cada *pixel* de uma imagem, é realizado ao calculo dos seus autovalores de LBP. Então, esses autovalores podem formar o espectro de recursos de LBP. O algoritmo *LBPH* utiliza o histograma do espectro característico LBP como vetor de característica para

classificação. Assim é realizada a comparação entre o histograma gerado a partir da imagem de teste e os histogramas das imagens da base de treinamento. A Figura 39 representa um exemplo de extração de características através de histogramas.

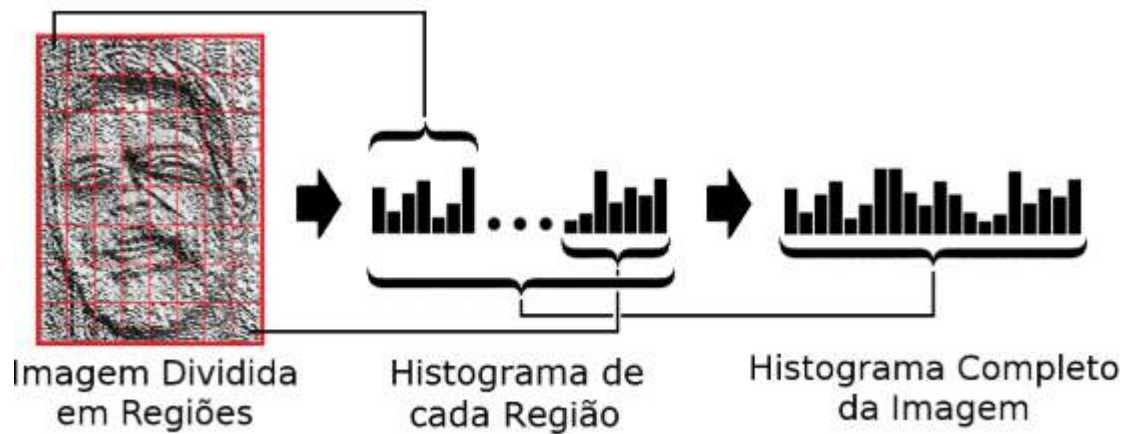


Figura 39 - Exemplo da extração do histograma da imagem.

Fonte: Adaptada de Salton (2016).

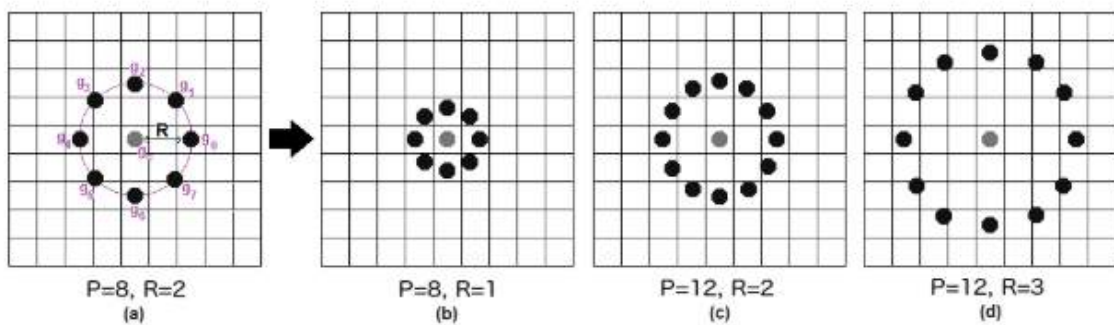


Figura 40 - Exemplo de LBP circular.

Fonte: Salton (2016).

4.4.2 Reconhecimento Facial

O reconhecimento facial pode ser feito utilizando classificadores ou métricas de similaridade. Que realiza calcula a distância entre o vetor de características a imagem de teste dos demais vetores de características de treinamento. Uma das métricas de similaridade mais difundida na literatura é o cálculo de distância euclidiana, representada pela Equação 17.

$$D(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (17)$$

Nas seções seguintes serão apresentadas algumas métricas de similaridades mais comuns em trabalhos conhecidos na literatura. Que são as distancias Euclidiana, distância Cosseno, distância de Mahalanobis e distância de Manhattan.

4.4.2.1 Distância Euclidiana

A distância euclidiana determina a similaridade entre dois pontos em um espaço n-dimensional, onde é calculado a distância entre dois pontos em linha reta. Dado dois pontos x e y encontra-se a distância D entre os pontos, de acordo com a Equação 18.

$$D(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (18)$$

4.4.2.2 Distância Cosseno

A distância Cosseno calcula o cosseno da medida do ângulo formado pelos pontos x e y . O cálculo da distância cosseno é dada pela Equação 19:

$$D(x, y) = \frac{\sum_{i=1}^n (x_i - y_i)}{\sqrt{\sum_{i=1}^n (x_i)^2} * \sqrt{\sum_{i=1}^n (y_i)^2}} \quad (19)$$

4.4.2.3 Distância de Mahalanobis

A Distância de Mahalanobis é baseada na correlação entre as variáveis baseados em grupos dados, onde os grupos devem ter o mesmo número de atributos, mas não necessariamente o mesmo número de dados. A equação da distância de Mahalanobis é representada pela Equação 20:

$$D(x, y) = \sqrt{\sum_{i=1}^n \frac{(x_i - y_i)^2}{\sigma_i^2}} \quad (20)$$

Onde, σ_i^2 é o desvio padrão da amostra. Se observarmos a fórmula da distância de Mahalanobis é a mesmo que da distância Euclidiana, porém normalizada.

4.4.2.4 Distância de Manhattan

A distância de Manhattan é definida como a somatória dos módulos das diferenças entre dois vetores X e Y , que dependo do sistema de coordenadas, mas não depende de sua reflexão ou translação. A equação é representada pela Equação 21:

$$D(X, Y) = \left(\sum_{i=1}^n |X_i - Y_i| \right)^1 = \quad (21)$$

$$|X_1 - Y_1| + |X_2 - Y_2| + \dots + |X_n - Y_n|$$

Onde, por exemplo dado duas coordenadas $X [5,2]$ e $Y [1,4]$, a distância de Manhattan seria D , onde $D(X, Y) = |1-5| + |4-2|$ $D(X, Y) = 6$. Representado pela Figura 41.

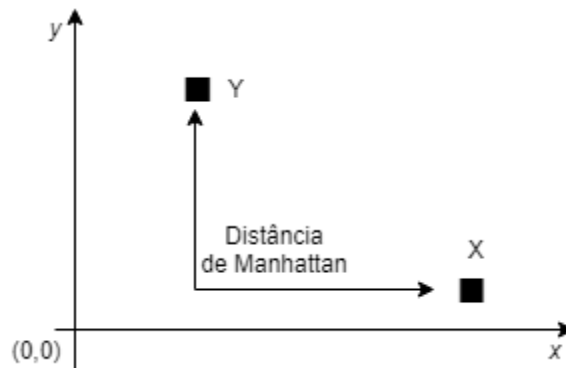


Figura 41 – Exemplo do cálculo de distância de Manhattan

5 ANÁLISE DOS EXPERIMENTOS

Neste capítulo serão abordados os procedimentos utilizados para a condução dos experimentos e os resultados obtidos, utilizando os algoritmos e processos descritos até aqui neste trabalho. O objetivo da análise dos experimentos é demonstrar o comportamento e qualidade dos algoritmos abordados junto às bases de imagens faciais escolhidas. Além dos resultados serão apresentados os principais parâmetros de cada algoritmo e como estes afetam os resultados obtidos, executando testes com diferentes valores e cenários para estes.

O restante do capítulo está dividido em duas seções, sendo estas: Ferramentas e Fases do Processo, onde serão descritas as ferramentas utilizadas para executar os algoritmos junto com seus parâmetros e a descrição do processo de testes com coleta dos resultados; Resultados e Discussões, onde serão apresentados os resultados dos algoritmos divididos pelas bases de imagens e valor das métricas de avaliação da qualidade, além das análises sobre os valores obtidos e comportamento das técnicas.

5.1 FERRAMENTAS E FASES DO PROCESSO

Para o desenvolvimento dos experimentos, de acordo com a escolhas das bases descritas no capítulo 4.1, é necessário dividir os dados de amostras em dois sub conjuntos de dados, um grupo de teste e um grupo de treinamento. Normalmente na literatura se encontra diversas formas de divisão de amostras, sem uma definição de regra a respeito da quantidade de amostras deve estarem cada grupo de dados, em nosso desenvolvimento foi utilizado o método de validação cruzada (*holdout*) com a separação de parte das imagens faciais da base de imagem original para treinamento e parte para teste (ZEVIAN E FERREIRA, 2018). Do conjunto de amostras total das imagens faciais, foram retirados 10% das imagens do conjunto aleatoriamente, para que desse dois novos conjuntos, não ocorresse a intersecção das imagens nos novos conjuntos. Todas as amostras da base de treinamento, são previamente rotuladas com uma identificação a qual indivíduo pertence, tornando o treinamento supervisionado, onde o algoritmo sabe a quem pertence cada imagem facial.

Após a divisão das amostras, foi gerado a nova base de treinamento, onde foi passado por parâmetro a base de treinamento para os algoritmos *Eigenfaces*, *Fisherfaces* e *LBPH*, que geraram 3 novas bases de dados, porém agora treinadas.

Para a realização do treinamento de cada um dos algoritmos, é referenciado o conjunto de amostras que o algoritmo utilizará para o treinamento, que ao executar, percorre cada umas

das imagens contidas no conjunto de treinamento, extraindo as características de cada indivíduo. Cada um dos algoritmos executou a extração de características de cada uma das amostras contida na base de treinamento, rotulando essas características em um novo subconjunto de dados, treinamento do algoritmo. Após o treinamento do algoritmo, a base de teste fora passada como parâmetro novamente para os algoritmos, que realiza a extração das características, porém agora confrontando as novas características, com os dados que estão na base de treinamento, classificando as imagens de teste como autêntica ou não de acordo com os parâmetros da subseção 5.1.2, retornando um relatório dos resultados encontrados de cada umas das imagens. A Figura 42, esquematiza o método utilizados da divisão das amostras para a condução dos experimentos executados neste trabalho.

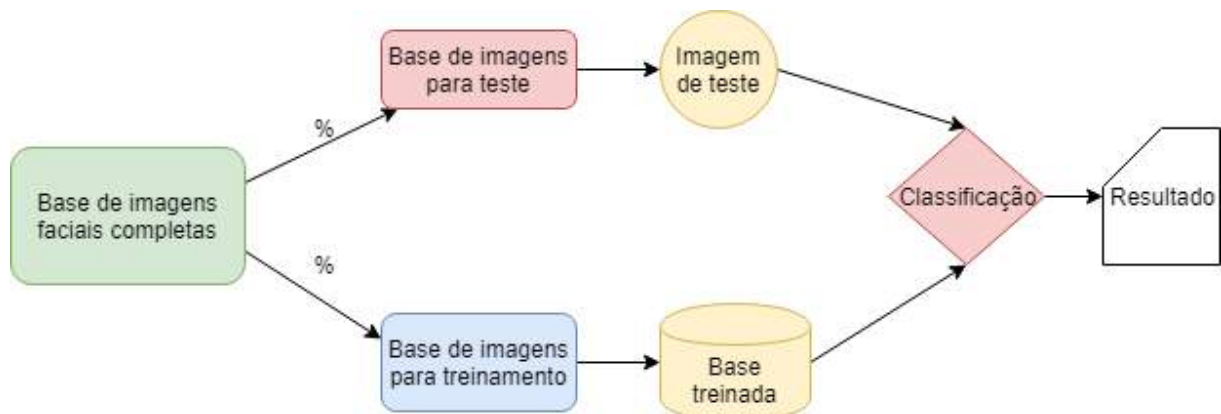


Figura 42 - Esquema de divisão de amostras para a condução dos experimentos.

Nas seções seguintes estão descritas as informações técnicas utilizadas para o desenvolvimento do presente trabalho, assim como a descrição dos parâmetros utilizados pelos algoritmos *Eigenfaces*, *Fisherfaces* e *LBPH*.

5.1.1 Informações Técnicas

Para a condução dos experimentos foram utilizados os seguintes equipamentos e *softwares*:

- Intel© Core™ i3-2375M CPU 1.50GHz × 2, memória de acesso aleatória de 3.7 GiB, e placa de vídeo Intel Corporation 2nd *Generation Core Processor Family Integrated Graphics Controller*.
- Os *softwares* utilizados foram: Sistema operacional Linux Mint 19.1 com versão Cinnamom 4.0.9, Kernel Linux – 4.15.0-46-generic, Python versão 2.7.15, biblioteca OpenCV na versão 4.0.0 e biblioteca OpenCV contrib.

OpenCV é uma biblioteca de código aberto e multiplataforma, desenvolvida originalmente pela Intel, que foi publicada sua versão estável no ano de 2000. É totalmente livre para o uso acadêmico ou comercial, direcionada para o desenvolvimento de aplicações para a área de visão computacional, tais como: reconhecimento de objetos, reconhecimento facial, reconhecimento de movimento, reconstrução 3D, realidade virtual entre outras aplicações (OPENCV, 2000), (BRADSKI; KAEHLER, 2008). A biblioteca possui suporte para diferentes linguagens de programação. O OpenCV contrib é um repositório com módulos extras da biblioteca OpenCV, mantida pela comunidade de código aberto livre.

A linguagem de programação Python, foi escolhida por sua versatilidade, uma grande comunidade de contribuidores e diversos projetos relacionados ao reconhecimento de facial desenvolvido nesta linguagem. Python é uma linguagem de programação de alto nível e fácil aprendizagem, com tipagem dinâmica e forte e multiparadigma, foi lançada em 1991 pelo programador Guido van Rossum. Python é disponibilizada gratuitamente pelo seu site oficial e é compatível com vários sistemas operacionais (PYTHON, 1991).

5.1.2 Parâmetros dos Algoritmos

Os algoritmos utilizados nos experimentos possuem parâmetros que influenciam diretamente em suas estratégias de extração de características e classificação. Assim, esses principais parâmetros serão descritos a seguir, sendo que diferentes valores serão atribuídos a estes durante os experimentos e farão parte do objeto de análise dos resultados obtidos com suas variações. Os principais parâmetros são:

- **Eigenfaces**
 - **num_components**: representa o número de componentes que serão mantidos para análise pelo PCA, no momento de reduzir a dimensionalidade da base. Por padrão este possui o valor 0, utilizando todas os componentes (Opencv, 2000);

- **threshold:** Representa o limiar aplicado no momento de classificar uma face. Por padrão o parâmetro utilizado é `DBL_MAX`, que significa que utilizará o valor máximo como limiar.

- **Fisherfaces**

- **num_components:** representa o número de componentes que serão mantidos para o LDA na redução de dimensionalidade. Por padrão seu valor é 0, utilizando todas os componentes (OpenCV, 2000);

- **threshold:** Representa o limiar aplicado na classificação de uma determinada face. Assim como *Eigenfaces*, por padrão o parâmetro utilizado é `DBL_MAX`, que significa que utilizará o valor máximo como limiar.

- **LBPH**

- **radius:** Representa o raio utilizado pelo LBP circular. Por padrão utiliza-se o valor 1.

- **neighbors:** Representa o número de vizinhos utilizados para construir o LBP circular. Por padrão a biblioteca OpenCV utiliza o valor 8.

- **grid_x:** O número de células horizontais que serão utilizadas para compor o histograma. Por Padrão é utilizado o valor 8.

- **grid_y:** O número de células verticais que serão utilizadas para compor o histograma. Por Padrão é utilizado o valor 8.

- **threshold:** Representa o limiar aplicado na classificação. Assim como *Fisherfaces*, por padrão o parâmetro utilizado é `DBL_MAX`, que significa que utilizará o valor máximo como limiar.

Cabe salientar que na condução dos experimentos foram mantidos tanto os valores padrões dos parâmetros, como também estes foram alterados para analisar os resultados obtidos de acordo com estas alterações. Todos os parâmetros descritos foram obtidos diretamente da documentação da biblioteca OpenCV (2000).

5.1.3 Métricas de Qualidade

Para a condução dos experimentos foram estabelecidas alguns critérios de organização para realização dos testes e escolha das métricas para avaliar a qualidade dos resultados obtidos. Para a realização dos experimentos foi definido:

- Para a base de treinamento *Yale Database*, foram separadas em diferentes diretórios 45, 75 e 150 imagens faciais, contendo 3, 5 e 10 faces com diferentes expressões, de cada indivíduo respectivamente.
- Para a base de treinamento *AT&T Database of Faces*, foi separada em diferentes diretórios 120, 200 e 360 imagens faciais, contendo 3, 5 e 9 faces de cada indivíduo com diferentes expressões, respectivamente.
- Para a Base de treinamento *Yale Database* modificada, foi utilizada a mesma organização da base original acima.
- Para as bases de teste, foi realizada a separação de uma imagem de cada indivíduo em uma pasta diferente, para que ela não fizesse parte do conjunto de treinamento.

As métricas de qualidade selecionadas são Confiança, Precisão, Sensibilidade, Especificidade e as quantidades de Verdadeiro Positivo (VP), Verdadeiro Negativo (VN), Falso Positivo (FP) e Falso Negativo (FN), sendo estes:

- **Verdadeiro Positivo:** Imagem facial de teste corresponde a uma imagem contida na base de treinamento, classificação executada corretamente com grau de confiança abaixo do limiar definido.
- **Verdadeiro Negativo:** A imagem facial de teste corresponde a uma imagem contida na base de treinamento, classificação executada incorretamente com grau de confiança acima do limiar definido.
- **Falso Positivo:** A imagem facial de teste não corresponde a uma imagem contida na base de treinamento, classificação executada incorretamente com grau de confiança abaixo do limiar definido.
- **Falso Negativo:** A imagem facial de teste não corresponde a uma imagem contida na base de treinamento, classificação executada corretamente com grau de confiança acima do limiar definido.
- **Confiança:** Trata-se da medida de distância entre a imagem facial de teste para as imagens treinadas e classificadas na base de treinamento.

$$Confiança = 100 - \frac{Pontos\ que\ combinam * 100}{Total\ de\ pontos\ Caracteristicos} \quad (22)$$

O valor da confiança é obtido diretamente pela função e pela soma dos resultados dos pontos que combinam entre si. Sabe-se que quanto mais próximo de 0 o valor da confiança, mas correta é a classificação.

- **Precisão:** Ou eficiência, é a porcentagem dos resultados corretamente identificada pelo teste, onde:

$$Precisão = \frac{VP + VN}{VP + VN + FP + FN} \quad (23)$$

- **Sensibilidade:** Refere-se ao quão provável o algoritmo é capaz de reconhecer cada uma das faces corretamente, baseado na seguinte formula:

$$Sensibilidade = \frac{VP}{VP + FN} \quad (24)$$

- **Especificidade:** Refere-se ao quão provável o algoritmo é capaz de identificar uma das faces corretamente que não pertence a base de treinamento, baseado na seguinte formula:

$$Especificidade = \frac{VN}{VN + FP} \quad (25)$$

5.2 RESULTADOS E DISCUSSÕES

Nesta seção serão apresentados os resultados obtidos através dos experimentos realizados com os algoritmos analisados. A comparação utiliza as métricas descritas na subseção 5.2.2, junto com análise do desempenho geral dos algoritmos *Eigensfaces*, *Fisherfaces* e *LBPH*. Como já descrito anteriormente, foram usadas bases de treinamento de tamanhos diferentes e com variações nos valores dos parâmetros para uma ampla análise.

Para a geração dos resultados, foi necessário o desenvolvimento de um algoritmo para otimizar a construção das características das imagens contidas nas bases e a realização dos testes. Esse é capaz de realizar os treinamentos com ambos os algoritmos de reconhecimento facial escolhidos e configurar diferentes parâmetros, alterando os tamanhos nas bases de treinamento e de testes, gerando relatórios organizados de acordo com o algoritmo e base utilizados. A Figura 43 mostra o esquema do algoritmo de otimização para realização dos testes.

Os parâmetros utilizados nos algoritmos foram gerado aleatoriamente de acordo com o intervalo aceito pela biblioteca Opencv.

5.2.1 Resultados *Yale Face*

Nas amostras da base *Yale Face* a divisão de imagens faciais utilizada foi de 15 para a base de teste e 45, 75 e 150 para as bases de treinamento. Essa divisão de quantidades está especificada sempre na primeira coluna das tabelas que contém os resultados dos experimentos, as demais informações contidas nessas são os valores de VP, VN, FP, FN, as métricas de qualidade e os valores atribuídos aos parâmetros do algoritmo, respectivamente.

A Tabela 1 mostra os resultados dos testes com o algoritmo *Eigenfaces*, onde foram efetuados quatro testes com cada tamanho específico de base de treinamento. Os parâmetros utilizados para a construção das bases e quantidade de características a serem analisadas por cada indivíduo foram de 40, 45 e 50. Já os limiares de classificação foram de 6000 a 9000. Na Tabela 1 é possível observar que conforme aumenta o número de características principais e o limiar utilizadas pelo algoritmo, os valores de VP também aumentam. Mas um dado que chama atenção é que o classificador teve maior acertos na base com 150 imagens para treinamento com os parâmetros de 45 características principais e limiar de 9000. Nessa configuração o *Eigenfaces* obteve resultado superior aquele utilizando os parâmetros padrões que vêm configurados na biblioteca Opencv.

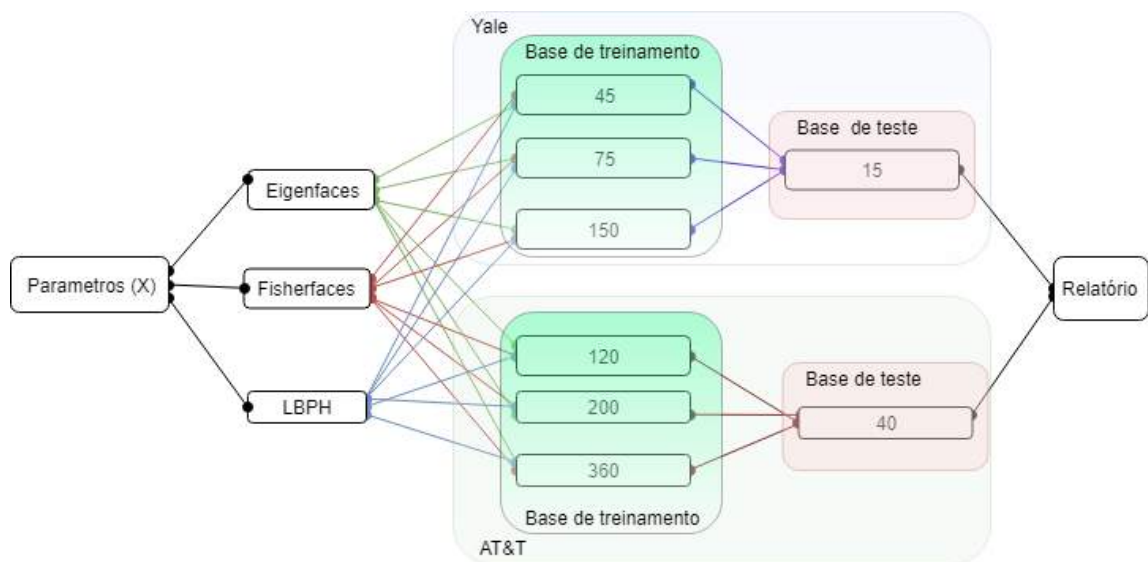


Figura 43 – Esquema utilizado para a construção do algoritmo de otimização dos testes.

Tabela 1 – Resultados *Eigenfaces* com base *Yale Face*.

Eigenfaces/Yale Face										
Teste/Trein	VP	FP	VN	FN	Acertos %	Confiança	Precisão %	Sensibili. %	Especif. %	Parâmetros
45/15	4	0	0	11	26,67	21845,65	100	26,67	0	[40, 8000]

45/15	6	1	0	8	40,00	40514,21	85,71	42,86	0	[45, 9000]
45/15	6	1	0	8	40,00	40514,21	85,71	42,86	0	[50, 9000]
45/15	4	0	6	5	26,67	23392,55	100	44,4	100	Padrão
75/15	12	0	0	3	80,0	40910,8	100	80,0	0	[40, 8000]
75/15	12	0	0	3	80,0	42718,3	100	80,0	0	[45, 9000]
75/15	12	0	0	3	80,0	44427,3	100	80,0	0	[50, 9000]
75/15	12	0	3	0	80	49990,97	100	92,3	100	Padrão
150/15	14	0	0	1	93,3	40975,8	100	93,3	0	[40, 8000]
150/15	15	0	0	0	100,0	52369,9	100	100,0	0	[45, 9000]
150/15	14	0	0	1	93,3	44863,3	100	93,3	0	[50, 9000]
150/15	13	0	2	0	89,6	55785,92	100	100	100	Padrão

Na Tabela 2 o algoritmo *Fisherfaces* também foi treinado com 45, 50 e 150 imagens de treinamento e 15 imagens faciais de teste. Os parâmetros utilizados para a construção das bases treinamento utilizaram os valores 2, 3 e 4 de características a serem utilizadas pelo algoritmo LDA para redução de dimensionalidade e extração de características. Já os limiares de classificação foram de 1000 a 4000. Observa-se que os testes realizados com 150 imagens de treinamento com parâmetros customizados obtiveram melhores resultados frente aqueles utilizando os valores padrão da biblioteca Opencv.

Tabela 2 - Resultados *Fisherfaces* com base *Yale Face*.

<i>Fisherfaces/Yale Face</i>										
Trein/Teste	VP	FP	VN	FN	Acertos %	Confiança	Precisão %	Sensibili. %	Especif. %	Parâmetros
45/15	3	1	0	11	20,0	2153,4	75,0	21,4	0	[4, 1000]
45/15	7	8	0	0	46,7	5563,9	46,7	100	0	[3, 2000]
45/15	8	7	0	0	53,3	10588,2	53,3	100	0	[4, 4000]
45/15	1	0	10	4	6,66	1370,48	100	20	100	Padrão
75/15	11	1	0	3	73,3	6227,6	91,7	78,6	0	[4, 1000]
75/15	12	3	0	0	80,0	6702,8	80,0	100	0	[3, 2000]
75/15	12	3	0	0	80,0	8005,4	80,0	100	0	[4, 4000]
75/15	9	0	6	0	60	9785,48	100	100	100	Padrão
150/15	15	0	0	0	100	5972,56	100	100	0	[4, 1000]
150/15	15	0	0	0	100	5025,29	100	100	0	[3, 3000]
150/15	15	0	0	0	100	5972,56	100	100	0	[4, 4000]
150/15	9	0	5	1	60	12054,78	100	90	100	Padrão

Na Tabela 3 estão os resultados com o algoritmo LBPH, que foi treinado com as mesmas configurações do *Eigenfaces* e *Fisherfaces*. Os parâmetros de quantidade de *pixels* vizinhos a serem tratados foram alternados entre 1 e 4, já os limiares de classificação foram de 35 a 50. Assim como nos outros experimentos os testes realizados com 150 imagens de treinamento e

parâmetros customizados superaram aqueles com valores padronizados da biblioteca Opencv. Destaque para um teste específico que sobressaiu perante aos demais com configuração de parâmetros [1, 1, 4, 4, 40] e a base de treinamento com 150 imagens faciais, que classificou todas as faces corretamente e teve seu nível de confiança muito próximo a zero e com 100% de acerto.

Tabela 3 – Resultados *LBPH* com base *Yale Face*.

<i>LBPH/Yale Face</i>										
Trein/Teste	VP	FP	VN	FN	Acertos %	Confiança	Precisão %	Sensibili. %	Especif. %	Parâmetros
45/15	2	13	0	0	13,3	0,30665307	13,33	100	0	[1, 1, 4, 4, 40]
45/15	9	6	0	0	60,0	110,9801208	60,00	100	0	[3, 3, 7, 7, 45]
45/15	10	5	0	0	66,7	86,92461436	66,67	100	0	[4, 4, 6, 6, 50]
45/15	6	3	3	3	40	137,387122	66,67	66,67	50	Padrão
75/15	14	1	0	0	93,3	0,717562863	93,3	100	0	[1, 1, 4, 4, 40]
75/15	14	1	0	0	93,3	37,94937124	93,3	100	0	[3, 3, 7, 7, 45]
75/15	14	1	0	0	93,3	30,37933745	93,3	100	0	[4, 4, 6, 6, 50]
75/15	14	0	0	1	93,3	155,862994	100	93,33	0	Padrão
150/15	15	0	0	0	100	0,492437755	100	100	0	[1, 1, 4, 4, 40]
150/15	15	0	0	0	100	29,55553295	100	100	0	[3, 3, 7, 7, 45]
150/15	15	0	0	0	100	24,06149228	100	100	0	[4, 4, 6, 6, 50]
150/15	15	0	0	0	100	159,161506	100	100	0	Padrão

Observando os resultados obtidos pelos algoritmos nas Tabelas 1, 2 e 3 é possível observar o aumento no valor de precisão obtido a medida em que é aumentada a quantidade de imagens das bases de treinamento. Com 45 imagens na base treinamento a tendência a FP é grande, demonstrando que uma quantidade maior de imagens faciais na base de treino é importante para o sucesso dos algoritmos.

Analisando o desempenho dos algoritmos, baseados na execução dos algoritmos com os parâmetros padrões, nota-se que de forma geral o algoritmo *LBPH*, teve o desempenho melhor, com todos os tamanhos de base de treinamento, superando os algoritmos *Eingenfaces* e *Fisherfaces*. Um dos motivos do melhor desempenho ter sido do algoritmo *LBPH* sobre a base *YaleFaces*, é pela base conter imagens faciais, com diferente intensidades de escala de cinza, onde o algoritmo *LBPH* por trabalhar sobre os histogramas criados a partir da intensidade de cor dos *pixels*, não tem dificuldade sobre a extração das características, obtendo várias características importantes no momento da classificação.

A Figura 44 mostra os gráficos com a média dos valores de confiança obtidos nos experimentos anteriores. Os valores de confiança estão separados pelos diferentes tamanhos das

bases de treinamento fornecidas pela *Yale Face*, onde para a geração dos gráficos foram utilizados todos os resultados obtidos pelos algoritmos. Cada um dos algoritmos teve sua particularidade. O gráfico do algoritmo *Eigenfaces* (ver Figura 44), juntamente com a tabela de amostras (ver Tabela 1), nota-se que os valores de confiança aumentaram, onde o resultado pode não ser desejado se observador apenas o gráfico, porém o algoritmo passou também a reconhecer mais faces corretamente, onde esse aumento dos confiança passa a ser aceitável. Já no caso do gráfico do algoritmo *LBPH* (ver Figura 44) e a sua tabela correspondente (ver Tabela 3), conforme aumentado as amostras de treinamento, diminuíram os valores medianos de confiança e o algoritmo também passou a reconhecer mais faces, o que se torna um resultado bom.

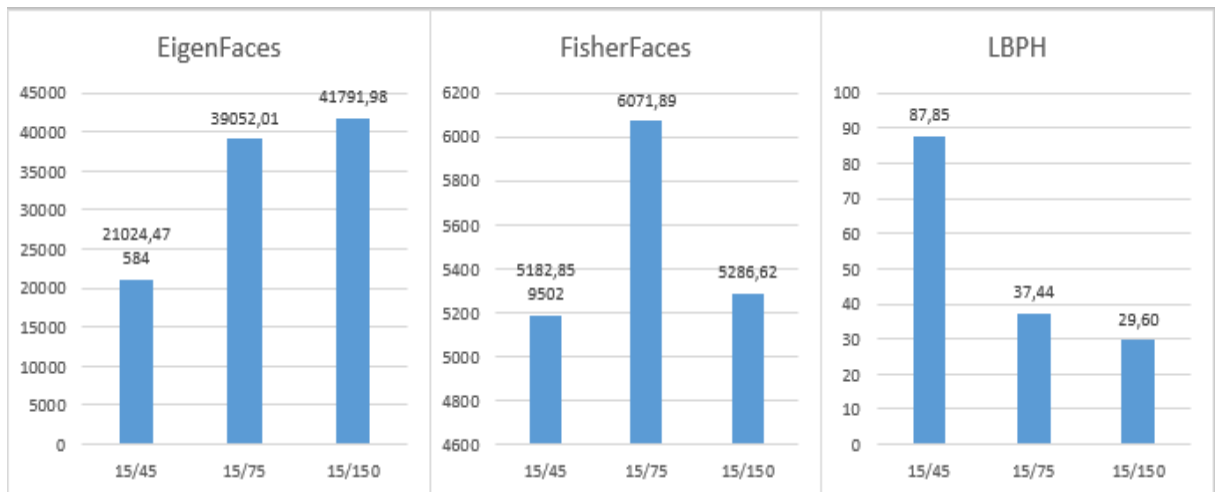


Figura 44 - Média dos resultados de confiança dos algoritmos com a base *Yale Face*.

5.2.2 Resultados *AT&T Faces*

Nas amostras da base *AT&T Faces* a divisão de imagens faciais utilizada foi de 40 para a base de teste e 120, 200 e 360 para as bases de treinamento. A especificação dessas quantidades bem como os valores das demais métrica e parâmetros seguem o mesmo modelo das tabelas anteriores.

Tabela 4 - Resultados *Eigenfaces* com base *AT&T Faces*.

<i>Eigenfaces/AT&T</i>										
Trein/Teste	VP	FP	VN	FN	Acertos %	Confiança	Precisão %	Sensibili. %	Especif. %	Parâmetros
120/40	34	6	0	0	84,21	53807,18	84,21	100	0	[40, 8000]
120/40	34	6	0	0	84,21	55758,90	84,21	100	0	[45, 9000]
120/40	34	6	0	0	84,21	57274,14	84,21	100	0	[50, 9000]

120/40	35	5	0	0	86,84	73261,02	86,84	100	0	Padrão
200/40	38	2	0	0	94,7	48539,2	94,7	100	0	[40, 8000]
200/40	38	2	0	0	94,7	50491,5	94,7	100	0	[45, 9000]
200/40	38	2	0	0	94,7	52085,6	94,7	100	0	[50, 9000]
200/40	38	2	0	0	94,73	73725,65	94,73	100	0	Padrão
360/40	39	1	0	0	97,4	45283,7	97,4	100	0	[40, 8000]
360/40	39	1	0	0	97,4	47270,9	97,4	100	0	[45, 9000]
360/40	39	1	0	0	97,4	48830,4	97,4	100	0	[50, 9000]
360/40	39	1	0	0	97,36	77230,62	97,36	100	0	Padrão

Analisando os resultados da Tabela 4, com os dados do algoritmo *Eigenfaces*, nota-se que o algoritmo teve um bom desempenho, em relação a reconhecimento da faces, pois teve uma grande quantidade de acertos no reconhecimento das faces utilizadas para teste, reconhecendo mais de 80% das faces, como legitima. Nota-se também que o algoritmo mesmo utilizando parâmetros diferentes dos padrões, teve praticamente o mesmo desempenho. Em nossos testes também notamos que mesmo a base da *AT&T* possuindo imagens faciais com diferentes intensidades de cor cinza, teve um bom desempenho sobre eles.

Tabela 5 - Resultados *Fisherfaces* com base *AT&T* Faces.

<i>Fisherfaces/AT&T</i>										
Trein/Teste	VP	FP	VN	FN	Acertos %	Confiança	Precisão %	Sensibili. %	Especif. %	Parâmetros
120/40	21	19	0	0	47,5	4723,00	50	100	0	[4, 1000]
120/40	14	26	0	0	30	2488,01	31,6	100	0	[3, 2000]
120/40	21	19	0	0	47,5	4723,00	50	100	0	[4, 4000]
120/40	36	3	0	1	85	40405,58	91,89	97,14	0	Padrão
200/40	27	13	0	0	62,5	3975,20	65,79	100	0	[4, 1000]
200/40	17	23	0	0	37,5	1448,04	39,47	100	0	[3, 2000]
200/40	27	13	0	0	62,5	3975,20	65,79	100	0	[4, 4000]
200/40	39	1	0	0	97,36	28384,15	97,36	100	0	Padrão
360/40	28	12	0	0	65,0	2938,8	68,4	100	0	[4, 1000]
360/40	19	21	0	0	42,5	1445,3	44,7	100	0	[3, 2000]
360/40	28	12	0	0	65,0	2938,8	68,4	100	0	[4, 4000]
360/40	39	1	0	0	97,36	19869,81	97,36	100	0	Padrão

Observando os dados da Tabela 5, do algoritmo *Fisherfaces*, nota-se uma dificuldade no reconhecimento facial quando utilizados os parâmetros diferentes dos padrões oferecidos pela documentação do algoritmo, onde o mesmo teve melhor desempenho quando foi utilizado os parâmetros padrões disponível na biblioteca *Opencv* na execução dos treinamento e teste. Comparando os resultados das tabelas Tabela 4 e Tabela 5, nota-se que mesmo o algoritmo

Fisherfaces sendo superior ao *Eigenfaces* em relação a extração de características com diferentes iluminações, teve um desempenho melhor na identificação dos indivíduos com os parâmetros padrões oferecidos pela biblioteca Opencv, mas teve um desempenho reduzido quando passado parâmetros diferentes dos padrões na execução do algoritmo.

Tabela 6 – Resultados *LBPH* com base *AT&T* Faces.

<i>LBPH/AT&T</i>										
Trein/Teste.	VP	FP	VN	FN	Acertos %	Confiança	Precisão %	Sensibili. %	Especif. %	Parâmetros
120/40	24	16	0	0	55	1,19	57,89	100	0	[1, 1, 4, 4, 40]
120/40	34	6	0	0	80	359,20	84,21	100	0	[3, 3, 7, 7, 45]
120/40	34	6	0	0	80	341,64	84,21	100	0	[4, 4, 6, 6, 50]
120/40	34	6	0	0	84,21	2282,91	84,21	100	0	Padrão
200/40	32	8	0	0	75	1,77	78,95	100	0	[1, 1, 4, 4, 40]
200/40	37	3	0	0	87,5	325,22	92,11	100	0	[3, 3, 7, 7, 45]
200/40	37	3	0	0	87,5	309,96	92,11	100	0	[4, 4, 6, 6, 50]
200/40	37	3	0	0	92,10	2408,53	92,10	100	0	Padrão
360/40	37	3	0	0	87,5	1,79	92,11	100	0	[1, 1, 4, 4, 40]
360/40	39	1	0	0	92,5	323,87	97,37	100	0	[3, 3, 7, 7, 45]
360/40	39	1	0	0	92,5	307,61	97,37	100	0	[4, 4, 6, 6, 50]
360/40	40	0	0	0	100	2593,25	100	100	0	Padrão

Analisando os dados da Tabela 6, com o algoritmo *LBPH*, nota-se que foi possível gerar resultados com valores de confiança se aproximados de 0 quando utilizado os parâmetros [1, 1, 4, 4, 40], mostrando que o algoritmo teve maior confiança com estes parâmetros, porém a quantidade de acertos foi reduzido comparado aos outros parâmetros, mesmo assim retornou resultados melhores que com os parâmetros padrões da biblioteca Opencv.

Analisando todas as tabelas dos resultados obtidos utilizando a base de imagens *AT&T*, nota-se uma grande diferença nos resultados dos algoritmos, principalmente na quantidade de VP e FP que podem ser caracterizados como os parâmetros mais importantes quando se trata de aplicações em um ambiente de monitoramento real, que o sistema está permitindo o acesso de um indivíduo a uma área restrita, onde os melhores resultados foram obtidos pelo algoritmo *LBPH*, independentemente do tamanho das bases de treinamento, que teve as maiores taxas de VP, confiança e precisão em relação aos outros algoritmos. Porém dependendo do ambiente de aplicação do algoritmo *LBPH*, as verificações com os parâmetros que verificam um maior número de características, podem gerar um certo atraso na resposta do algoritmo, pois depende que ele exerça uma maior esforço computacional durante o treinamento do algoritmo, onde os

detalhes do histograma são aumentados promovendo uma análise mais precisa dos *pixels*, porém o custo computacional também aumenta, o que pode gerar um atraso para obtenção dos resultados.

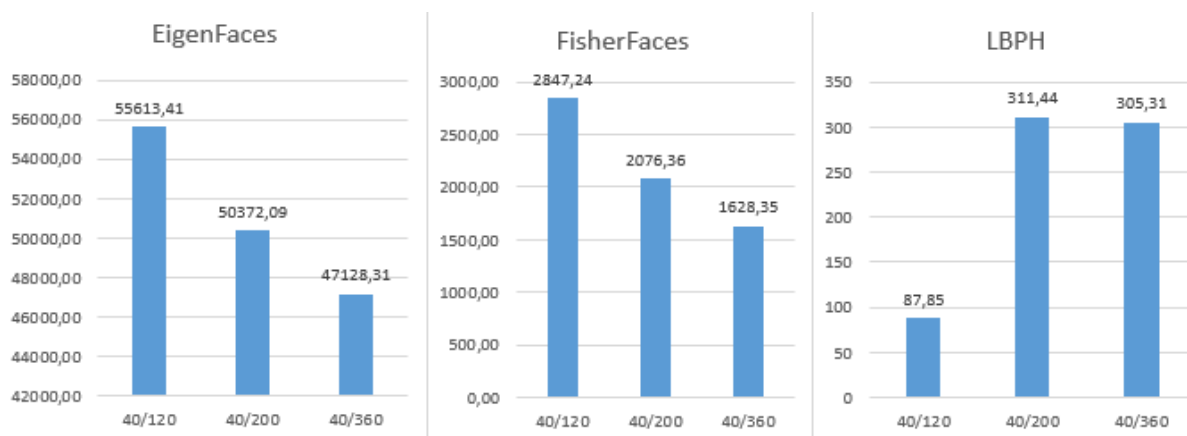


Figura 45 - Média dos resultados de confiança dos algoritmos com a base *AT&T Faces*.

Assim como os resultados da média das confianças da base *Yale Face*, a base *AT&T* teve o mesmo comportamento, como demonstrado na Figura 45, onde a confiança não representa maior precisão necessariamente, necessitando a observação da quantidade de acertos o algoritmo realizou. A particularidade dos resultados obtidos pelos algoritmos utilizando a base *AT&T*, é que os algoritmos *Eigenfaces* e *LBPH* tiveram resultados tanto de confiança, quanto de acertos das faces, melhor que o algoritmo *Fisherfaces*, que apresentou algumas dificuldades no reconhecimento das faces mesmo aumentando a quantidade de amostras durante o treinamento.

Apesar de alguns dos resultados dos algoritmos serem satisfatórios, existe a necessidade de melhor calibração dos parâmetros, pois ainda existem as possibilidade de resultarem em acesso ao áreas restritas que indivíduos não possuem tal direito de acesso.

5.2.3 Resultados *Yale Face* Obstruída.

Nesta subseção serão apresentados os resultados de experimentos com a base *Yale Face* modificada, onde foram adicionadas obstruções nas imagens da base de teste. O objetivo é verificar a capacidade dos algoritmos realizarem o reconhecimento facial dos indivíduos mesmo com obstrução sobre a face e verificar como estes se comportam sob situações de estresse. Na

condução dos experimentos foram utilizados os mesmos processos dos experimentos da base original. Assim nas imagens de teste foram removidos 3%, 9% e 15% dos *pixels* localizados sobre a boca e preenchendo este espaço com a cor sólida preta.

Tabela 7 – Resultados do algoritmo *Eigenfaces* com base *Yale Face* obstruída.

<i>Eigenfaces 15/150</i>										
Oclusão	VP	FP	VN	FN	Acertos %	Confiança	Precisão %	Sensibili. %	Especif. %	Parâmetros
3%	14	0	0	1	93,3	40783,03	100	93,33	0	[40, 8000]
3%	15	0	0	0	100,0	52260,88	100	100,00	0	[45, 9000]
3%	14	0	0	1	93,3	44813,75	100	93,33	0	[50, 9000]
9%	14	0	0	1	93,3	42503,37427	100	93,3	0	[40, 8000]
9%	15	0	0	0	100,0	54030,74575	100	100,0	0	[45, 9000]
9%	14	0	0	1	93,3	46980,75591	100	93,3	0	[50, 9000]
15%	14	0	0	1	93,3	48480,74082	100	93,3	0	[40, 8000]
15%	14	0	0	1	93,3	51214,26669	100	93,3	0	[45, 8000]
15%	14	0	0	1	93,3	53423,23397	100	93,3	0	[50, 8000]

Nos resultados do algoritmo *Eigenfaces* na Tabela 7, pode-se observar que a sua classificação não teve grandes mudanças entre o número de acertos, mesmo com os diferentes tamanhos de obstruções, apenas nos níveis de confiança, que conforme aumentava o valor de obstrução sobre a imagem maior foi o valor de confiança, o que dificultou o reconhecimento facial realizado pelo algoritmo, pois as partes obstruídas pela cor preta poderiam conter características importantes utilizadas no treinamento do algoritmo, mas como a imagem de teste não continham essas características fez com que a imagem de teste tivesse um linear de características pouco distante do linear do conjunto das imagens de treinamento, aumentando o valor das distancias dos classificadores, explicando o aumento da confiança conforme o aumento das obstruções.

Tabela 8 – Resultados do algoritmo *Fisherfaces* com base *Yale Face* obstruída.

<i>Fisherfaces 15/150</i>										
Oclusão	VP	FP	VN	FN	Acertos %	Confiança	Precisão %	Sensibili. %	Especif. %	Parâmetros
3%	15	0	0	0	100	4897,699783	100	100	0	[3, 2000]
3%	15	0	0	0	100	5858,739582	100	100	0	[4, 1000]
3%	15	0	0	0	100	5858,739582	100	100	0	[4, 4000]
9%	15	0	0	0	100	5310,193557	100	100	0	[3, 2000]
9%	15	0	0	0	100	6206,933586	100	100	0	[4, 1000]
9%	15	0	0	0	100	6206,933586	100	100	0	[4, 4000]
15%	15	0	0	0	100	6017,588887	100	100	0	[3, 2000]
15%	15	0	0	0	100	6837,989821	100	100	0	[4, 1000]
15%	15	0	0	0	100	6837,989821	100	100	0	[4, 4000]

Nos resultados do algoritmo *Fisherfaces*, podemos observar que sua única alteração foi no nível de confiança, onde com a obstrução sobre a face do indivíduo de teste, o algoritmo teve resultados iguais aos testes sem obstrução (comparação dos resultados das tabelas Tabela 8 e Tabela 2, com os mesmos parâmetros), diferenciando apenas na confiança. Mas levando em consideração que as faces de teste estavam com uma obstrução os resultados foram bons, onde o algoritmo ainda assim foi capaz de reconhecer as faces, permanecendo com 100% de acertos.

Tabela 9 – Resultados do algoritmo *LBPH* com base *Yale Face* obstruída.

LBPH 15/150										
Oclusão	VP	FP	VN	FN	Acertos %	Confiança	Precisão %	Sensibili. %	Especif. %	Parâmetros
3%	8	7	0	0	53,3	2,716771854	53,3	100	0	[1, 1, 4, 4, 40]
3%	15	0	0	0	100,0	153,8844188	100,0	100	0	[3, 3, 7, 7, 45]
3%	13	2	0	0	86,7	104,7159628	86,7	100	0	[4, 4, 6, 6, 50]
9%	8	7	0	0	53,3	2,926416763	53,3	100	0	[1, 1, 4, 4, 40]
9%	15	0	0	0	100,0	162,5220068	100,0	100	0	[3, 3, 7, 7, 45]
9%	13	2	0	0	86,7	109,8511706	86,7	100	0	[4, 4, 6, 6, 50]
15%	9	6	0	0	60,0	4,034145814	60,0	100	0	[1, 1, 4, 4, 40]
15%	15	0	0	0	100	184,2309173	100	100	0	[3, 3, 7, 7, 45]
15%	13	2	0	0	86,7	124,656836	86,7	100	0	[4, 4, 6, 6, 50]

Observando os resultados dos testes de oclusão com o algoritmo LBPH, nota-se uma certa oscilação nestes, tanto na quantidade de acertos quanto nos resultados de confiança. Analisando esses resultados, pode-se inferir que o algoritmo deixa de realizar alguns reconhecimentos por não conseguir, interpretar as obstruções, que em nosso caso foi aplicada uma tarja preta sobre a imagem, fazendo com que o algoritmo interprete como uma característica, que pode ser confundida com outro indivíduo que está na base, assim acaba não reconhecendo o indivíduo como autêntico, resultando em baixo VP, mas com FP maior pois o seu *threshold* foi abaixo do limiar declarado na execução do algoritmo.

A Figura 46 mostra a média dos valores de confiança nos experimentos anteriores realizados com a base *Yale Face* com oclusão. Analizando a Figura 43, nota-se graficamente a variância dos resultados de confiança obtidos pelos algoritmos. Mesmos as imagens de teste possuindo diferentes tamanos de obstrução sobre a face, os algoritmos continuaram reconhecendo as faces.

Observando o desempenho dos algoritmos Eigenfaces e LBPH, tiveram uma grande variância da confiança entre a obstrução de 9% e 15% da imagem, onde se declarou um valor de *threshold* próximo ao valor médio da confiança com 15% de obstrução, os algoritmos passaria a não reconhecer metade das faces de teste como VP passando a serem FN ou VN.

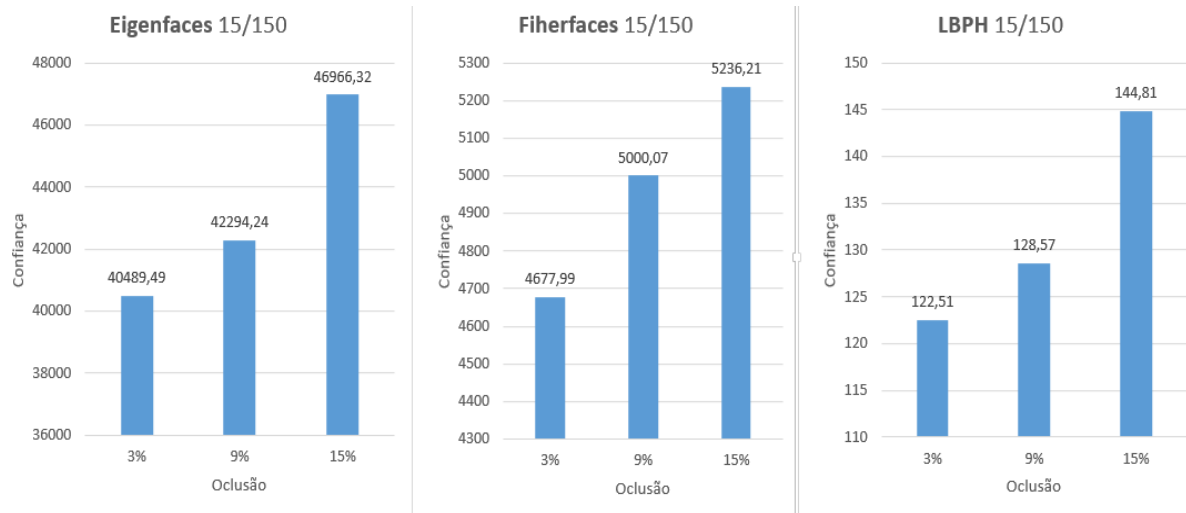


Figura 46 - Média de confiança dos testes com oclusão

A seguir são apresentados os gráficos com valores referentes a quantidade de VP, FP, VN e FN de cada algoritmo em relação a diferentes configurações de parâmetros, com as imagens de treinamento da *Yale Face* com 15% de oclusão sobre a face.

No teste apresentado na Figura 44, os valores de FP e VN do algoritmo *Eigenfaces* foram 0, sendo que a melhor classificação realizada pelo algoritmo foi utilizando os parâmetro de 40 características principais e 9000 de limiar, sobre a base de treinamento com 150 imagens faciais. Conforme foi aumentado o valor do *threshold* utilizado pelo algoritmo, o mesmo teve melhores resultados, comparado com a variação da quantidade de características utilizada pelo algoritmo. Analisando os dados da Figura 47, podemos deduzir que, mesmo as faces de teste tendo obstrução de 15%, se for escolhido os parâmetros corretos na hora do treinamento, mesmo a base de treinamento tendo diferentes graus de iluminação, ainda assim o algoritmo será capaz de identificar as faces corretamente.

A Figura 48 mostra os resultados obtidos pelo algoritmo *Fisherfaces*. Este apresentou alguns resultados de FP, porém conforme foram aumentadas as quantidades de características e limiar o algoritmo passou a reconhecer todos os resultados corretamente. Diferentemente do *Eigenfaces* o algoritmo teve melhores resultados quando aumentou o número de componentes

utilizados pelo algoritmo, quando maiores que 3, o algoritmo foi capaz de separar bem as classes de características e agrupar melhor as características na interclasse, os resultados foram satisfatórios, onde o algoritmo foi capaz de reconhecer todas as faces.

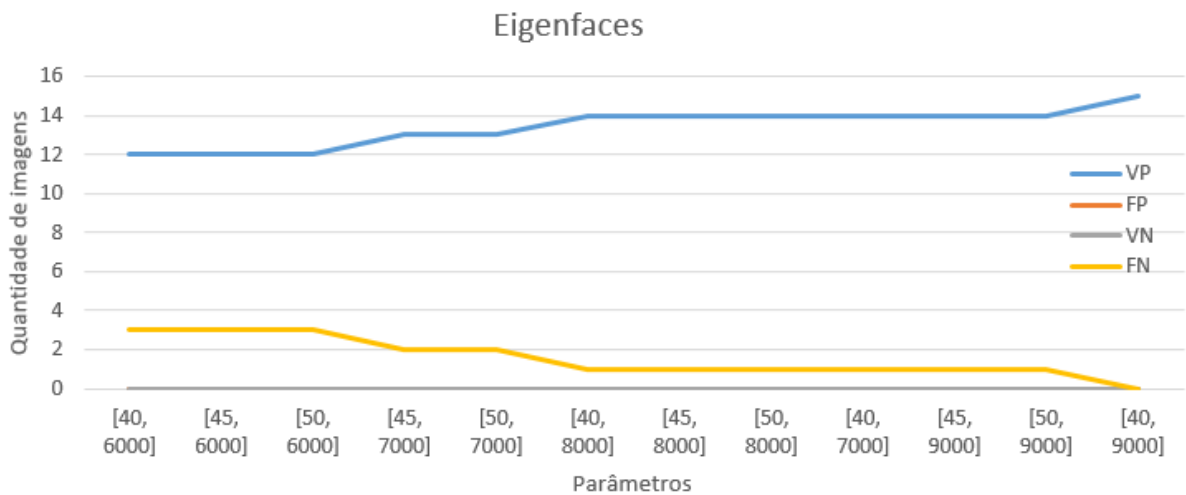


Figura 47 – Resultados *Eigenfaces* com 15% de oclusão na base de teste.

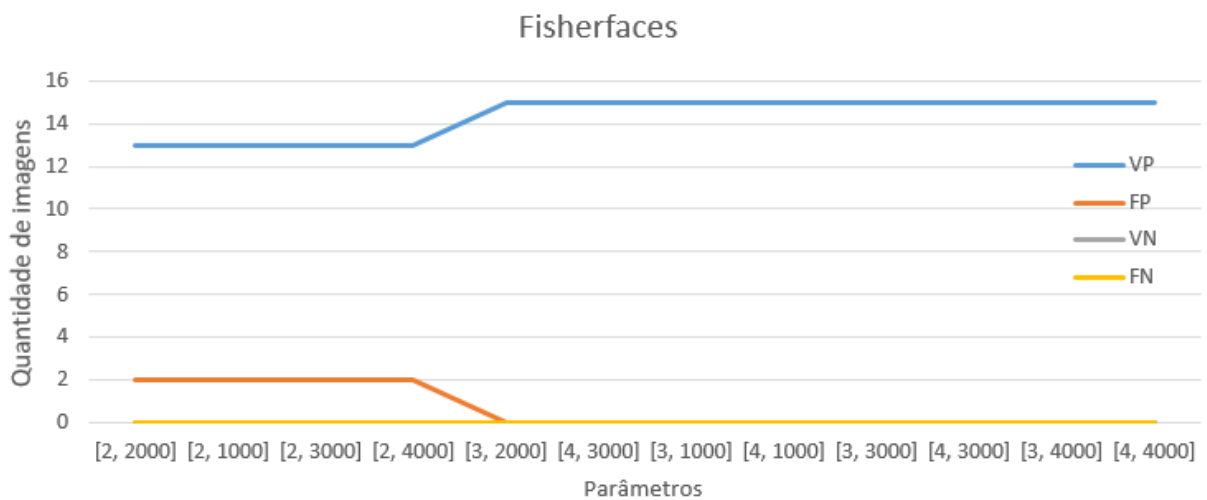


Figura 48 - Resultados *Fisherfaces* com 15% de oclusão na base de teste.

Com uma taxa de 15% de obstrução o algoritmo LBPH passou a ter maior dificuldade de acertos durante as classificações, como mostra o gráfico na Figura 49. O algoritmo conseguiu classificar corretamente o maior número de faces quando foram aumentados valor de

vizinhança analisada para a construção dos histogramas, porem o rocesso de aumentar o numero de visinhança, pode acarretar no erro do algoritmo, pois o mesmo passa a não identificar algumas bordas ou cantos, onde é reduzido os numeros de *pixels*, podendo ficar características importantes sem serem identificadas.

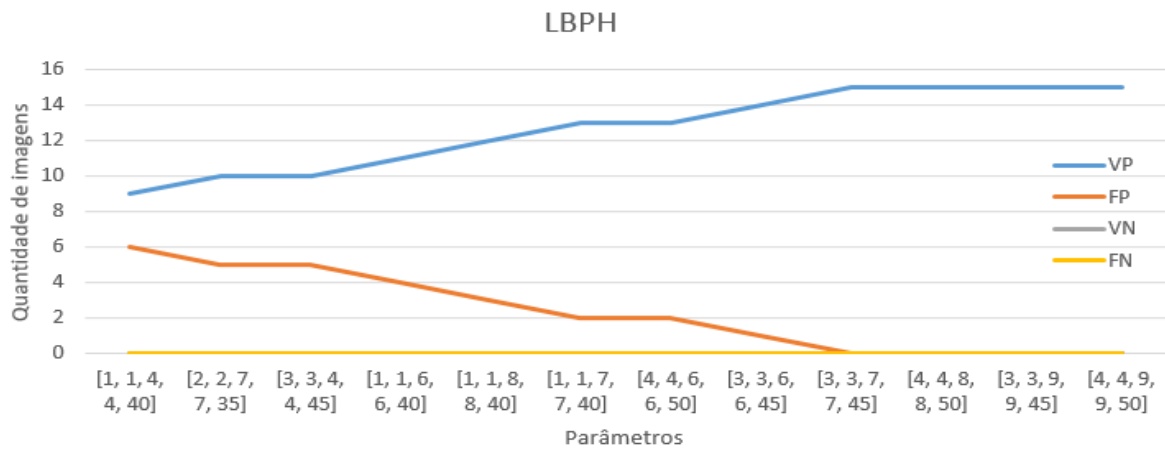


Figura 49 - Resultados *LBPH* com 15% de oclusão na base de teste.

Analizando os 3 graficos dos resultados das oclusões dos algoritmos, nota-se que o algoritmo *Fisherfaces*, teve melhores resultados de reconhecimento facial, mantendo se a maior parte dos testes estavel, mesmo as faces estando obstruidas parcialmente por uma tarja preta.

6 CONCLUSÃO

O trabalho reuniu e sintetizou o processo de reconhecimento facial, organizando e delimitando a base teórica necessária para a atividade apresentando as principais técnicas disponíveis dentro de um processo cronológico e com definição das entradas e saídas esperadas em cada parte deste. Foi dada ênfase a metodologia de reconhecimento facial no que tange a abordagem matemática e conceitual, junto as formas de se utilizar os algoritmos citados em um processo dividido por etapas. Com isso, espera-se que este trabalho seja uma referência adequada para aqueles que irão adentrar na área de reconhecimento facial, uma vez que na literatura atual a maioria dos trabalhos investigados não apresenta todo o processo com suas particularidades, focando em descrever tarefas e técnicas específicas do reconhecimento facial.

Foram abordados com detalhes de configuração, parametrização e algoritmos disponíveis as etapas de criação da base de imagens; detecção facial; extração de características; e reconhecimento facial. Os algoritmos Eigenfaces, Fischerfaces e LBPH foram apresentados com suas estruturas internas e estratégias para reconhecimento facial. A eficácia e comportamento desses foi medida e apresentada com diversos experimentos utilizando variação paramétrica e métricas de controle da qualidade, com análise dos resultados e testes com simulação de ambientes reais através de ruídos.

Nos experimentos foram utilizadas diferentes bases de treinamento. Inicialmente os algoritmos foram analisados frente a métrica de qualidade como acurácia e precisão utilizando a base Yale Face, com variações na quantidade de amostras na base de imagens. Nesses testes, o algoritmo LBPH conseguiu melhor desempenho utilizando a estratégia de histogramas da face. Em seguida, foram executados os experimentos com a base AT&T Faces com as mesmas configurações dos testes anteriores, com o algoritmo Eigenfaces obtendo o melhor desempenho utilizando a estratégia de face média. Por fim, foram analisados os resultados dos experimentos com a base Yale Face modificada, onde o algoritmo Fisherfaces apresentou melhor desempenho com oclusões sobre a face. No geral, o reconhecimento facial obteve seus melhores resultados com maiores quantidades de imagens compondo a base de treinamento.

6.1 TRABALHOS FUTUROS

Propõe-se, como trabalhos futuros, os indicados a seguir:

- Construção de uma síntese das técnicas de reconhecimento facial com abordagens mais atuais, como a utilização de redes neurais, *machine learning* ou *Deep learning*;
- A construção de um sistema de reconhecimento facial baseado em um ambiente real, com a utilização dos algoritmos em conjunto, a fim de realizar os testes aplicados neste trabalho.
- Realizar experimentos com outros tipos de ruídos aplicados sobre as bases de teste.

6.2 MATERIAIS ELABORADOS

Os materiais técnicos elaborados durante a execução deste projeto estão disponíveis em um repositório público disponível em: <https://github.com/VambasterCosta/RECONHECIMENTO_DE_PADROES_FACIAIS_UMA_SINTESE.git>.

REFERÊNCIAS

ARAÚJO, Gabriel Matos. **Algoritmo para reconhecimento de características faciais baseado em filtros de correlação**. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Rio de Janeiro, UFRJ, Rio de Janeiro, RJ, 2010.

AZEEM, Aisha et al. A survey: Face recognition techniques under partial occlusion. **Int. Arab J. Inf. Technol.**, v. 11, n. 1, p. 1-10, 2014.

BELHUMEUR, P. N.; HESPANHA, J. & KRIEGMAN, D. J. **Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection**. IEEE Transactions on Pattern Analysis and Machine Intelligence 19, 7 (July 1997), 711-720.

BISSI, Thelry David. **Reconhecimento Facial com os algoritmos Eigenfaces e Fisherfaces**. 2018. Trabalho de conclusão de Curso (Bacharel em Ciência da Computação) - Universidade Federal de Uberlândia, [S. l.], 2018.

BLEDSON, Woodrow W. **The model method in facial recognition**. Panoramic Research Inc., Palo Alto, CA, Rep. PR1, v. 15, n. 47, p. 2, 1966.

BRADSKI, Gary; KAEHLER, Adrian. **Learning OpenCV: Computer vision with the OpenCV library**. " O'Reilly Media, Inc.", 2008.

BRAGA, Luiz Filipe Zenicola. **Reconhecimento Facial com os algoritmos Eigenfaces e Fisherfaces**. 2013. Trabalho de conclusão de Curso (Curso de Engenharia Elétrica com ênfase em Eletrônica) - Universidade de São Paulo, [S. l.], 2013.

BRILHADOR, Anderson. **Análise semi-automática do arranjo espacial de plantas de milho utilizando visão computacional**. 2015. Dissertação de mestrado (Mestrado em Informática) - Informática da Universidade Tecnológica Federal do Paraná – UTFPR, Câmpus Cornélio Procopio, 2015

CHICAGO TRIBUNE. **In China, facial recognition is sharp end of a drive for total surveillance**. 07 de Janeiro de 2018. Disponível em: <<https://www.chicagotribune.com/nation-world/ct-china-facial-recognition-surveillance-20180107-story.html>> Acesso em: mai. 2019.

COMPUTAÇÃO GRÁFICA. **Processamento e Análise de Imagens Digitais**. Disponível em: <<http://computacaografica.ic.uff.br/transparenciasvol2cap6.pdf>>. Acesso em: Jan 2019.

COOTES, Timothy F.; TAYLOR, Christopher J.; LANITIS, Andreas. **Active Shape Models: Evaluation of a Multi-Resolution Method for Improving Image Search**. In: British Machine Vision Conference. 1994. p. 327-336.

COX, Ingemar J.; GHOSH, Joumana; YIANILLOS, Peter N. **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. p. 209-216.

DHARAVATH, Krishna; TALUKDAR, Fazal Ahmed; LASKAR, Rabul Hussain. **Improving face recognition rate with image preprocessing**. Indian Journal of Science and Technology, v. 7, n. 8, p. 1170-1175, 2014.

DINIZ, F. A. et al. **Redface: Um Sistema De Reconhecimento Facial Baseado Em Técnicas De Análise De Componentes Principais E Autofaces**. Revista Brasileira de Computação Aplicada, v. 5, n. 1, p. 42-54, 2013.

DPI. **Processamento Digital de Imagens Realce de Imagens Digitais**. Disponível em: <http://www.dpi.inpe.br/~carlos/Academicos/Cursos/Pdi/pdi_realce.html>. Acesso em: jun. 2019.

FACEBOOK. **Facebook Inc**. Disponível em: <www.facebook.com>. Acesso em out 2018.

FAGERTUN, Jens. **Face Recognition**. 2005. Dissertação de Mestrado. Technical University of Denmark, DTU, DK-2800 Kgs. Lyngby, Denmark.

FBI. **FBI Announces Full Operational Capability of the Next Generation Identification System**. Disponível em: <<https://www.fbi.gov/news/pressrel/press-releases/fbi-announces-full-operational-capability-of-the-next-generation-identification-system>> Acesso em: mai. 2019.

FISHER, R.A. **The use of multiple measurements in taxonomic problems**. Annals of Eugenics, v.7, p.179-188, 1936.

GALTON, F. **Personal identification and description**. Nature, pages 173-188, 1888. Disponível em: <<http://galton.org/essays/1880-1889/galton-1888-nature-personal-id.pdf>>. Acesso em: 05 nov. 2018.

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing (3rd Edition)**. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006. ISBN 013168728X.

GOOGLE AI: **Building a deeper understanding of images**. Disponível em:<<https://ai.googleblog.com/2014/09/building-deeper-understanding-of-images.html> >. Acesso em: out. 2018.

HASSABALLAH, M; ABDELMGEID, A; ALSHAZLY, H.; **Image features detection, description and matching**. In: Image Feature Detectors and Descriptors. Springer, Cham, 2016. p. 11-45.

HUA, Gang et al. **Introduction To The Special Section On Real-World Face Recognition**. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 33, n. 10, p. 1921-1924, 2011.

ISAAC, R. Augustian; AGARWAL, Abhinav; SINGH, Priyanshi. **Face Recognition Security Module using Deep Learning**. Journal of Network Communications and Emerging Technologies (JNCET) www.jncet.org, v. 8, n. 10, 2018.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 4th ed. Upper Saddle River, New Jersey: Prentice-Hall, 1999, 815 p

KHALID, S; KHALIL, T.; NASREEN, S. **A survey of feature selection and feature extraction techniques in machine learning**. In: Science and Information Conference (SAI), 2014. IEEE, 2014. p. 372-378.

KRIG, S. **Computer Vision Metrics: Survey, Taxonomy, and Analysis**.: Apress, 2014.

KSHIRSAGAR, V. P.; BAVISKAR, M. R.; GAIKWAD, M. E. Face recognition using Eigenfaces. In: **Computer Research and Development (ICCRD), 2011 3rd International Conference on**. IEEE, 2011. p. 302-306.

MANGAYARKARASI, S.; NAVEENA, E. **MATLAB Implementation of Image Segmentation Using Drop Fall Algorithm**. Journal of NanoScience and NanoTechnology, v. 2, n. 1, p. 305-308, 2014.

MICROSOFT AZURE. **Face**. Disponível em: <<https://azure.microsoft.com/pt-br/services/cognitive-services/face/>> Acesso em: mai. 2019.

MOLZ, Rolf Fredi; WINK, Rodrigo. Sistema De Reconhecimento De Faces Em Vídeo. **Anais do Salão de Ensino e de Extensão**, p. 377.

NUNES, Fernanda Todesco et al. **Técnicas de biometria baseadas em padrões faciais e sua utilização na segurança pública**. 2017.

OPENCV. **OpenCV Documentation**. [S.l.], 2000. Disponível em: <<http://docs.opencv.org/2.4/index.html>>. Acesso em: mai. 2019.

PEARSON, K. (1901) **On Lines and Planes of Closest Fit to Systems of Points in Space**. Philosophical Magazine, 2, 559-572

PRADO, Kelvin Salton do. **Comparação de técnicas de reconhecimento facial para identificação de presença em um ambiente real e semicontrolado**. Tese de Doutorado. Universidade de São Paulo (2018).

PRINCE, S. J. D. **Computer vision: models, learning and inference**. : Cap. Intr. e 13. Cambridge University Press, 2012.

PYTHON. **PYTHON Documentation**. [S.l.], 1991. Disponível em: <<https://docs.python.org/3/>>. Acesso em: mai. 2019.

ROZARIO, Liton Jude et al. **Quantitative Analysis Of Pca, Ica, Lda And Svm In Face Recognition**. World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering, v. 8, n. 9, p. 1613-1616, 2014.

SAEED, K.; NAGASHIMA, T. **Biometrics and Kansei Engineering**. New York: Springer, 2012.

SAIGAYATRI VADALI. **Dimensionality Reduction with PCA and t-SNE in R**. Disponível em: < <https://medium.com/@TheDataGyan/dimensionality-reduction-with-pca-and-t-sne-in-r-2715683819>>. Acesso em: nov. 2018.

SHARIF, Muhammad et al. **Face Recognition: A Survey**. Journal of Engineering Science & Technology Review, v. 10, n. 2, 2017.

SILVA, Alex Lima; CINTRA, Marcos Evandro. **Reconhecimento de padrões faciais: Um estudo**. In: Encontro Nacional De Inteligência Artificial E Computacional, 2015, Proceedings ENIAC. 2015. p. 224-231.

SINFIC. **Características biométricas e tecnologias mais comuns**. Disponível em: <<http://www.sinfic.pt/SinficWeb/displayconteudo.do?numero=25032>>. Acesso em: out 2018.

STAN Z.; ANIL, K. Jain. **Handbook of Face Recognition**. New York, SI. 2005.

SZELISKI, R. **Computer Vision: Algorithms and Applications**. : Cap. 1, 3, 5 e 14. Springer, 2010.

TEÓFILO EMÍDIO DE CAMPOS. **Técnicas de Seleção de Atributos e de Classificação para Reconhecimento de Faces**. Disponível em: <<http://www.vision.ime.usp.br/~teo/publications/qualificacao/node43.html>>. Acesso em: out 2018.

THE DATABASE OF FACES. **The database of faces**. Disponível em: <<https://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>>. Acesso em: out 2018.

THE YALE FACE DATABASE B. **Yale face database**. Disponível em: <<http://vision.ucsd.edu/~iskwak/ExtYaleDatabase/Yale%20Face%20Database.htm>>. Acesso em: out 2018.

TOWARDS DATA SCIENCE. **Face detection for beginners**. Disponível em: <<https://towardsdatascience.com/face-detection-for-beginners-e58e8f21aad9>>. Acesso em: out 2018.

VIOLA, Paul; JONES, Michael. **Rapid object detection using a boosted cascade of simple features**. In: Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. IEEE, 2001. p. I-I.

YANG, M. H.; AHUJA, N. & KRIEGMAN, D. **Detecting Faces in Images: A Survey**. IEEE Transactions on Pattern Analysis and Machine Intelligence, v. 24, n. 1, pp. 34-58, 2002.

YANG, MING-HSUAN; KRIEGMAN, DAVID J. ; AHUJA, NARENDA. **Detecting Faces**

in Images: A Survey, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.24,No.1,Janeiro 2002

ZEVIANI WALMES M.; FERREIRA EDUARDO V. **Machine Learning: Validação Cruzada**. 24-08-2018. Disponível em: < <http://leg.ufpr.br/~walmes/ensino/ML/tutorials/01-cross-validation.html>>. Acesso em: jun 2019.

ZHAO, S.; PRECIOSO, F.; CORD, M. **Spatio-temporal tube data representation and kernel design for svm-based video object retrieval system**. Multimedia Tools Appl. 55, 1 (2011), 105 - 125.

ZHAO, WENYI ET AL. **Reconhecimento facial: uma pesquisa bibliográfica**. Pesquisas de computação ACM (CSUR), v. 35, n. 4, p. 399-458, 2003.