

**UNIVERSIDADE DO EXTREMO SUL CATARINENSE - UNESC
CURSO DE CIÊNCIA DA COMPUTAÇÃO**

TIAGO ALEFF DA SILVA

**PROTÓTIPO DE RECONHECIMENTO FACIAL COM ÊNFASE EM EXTRAÇÃO DE
CARACTERÍSTICAS DA FACE**

**CRICIÚMA
2018**

TIAGO ALEFF DA SILVA

**PROTÓTIPO DE RECONHECIMENTO FACIAL COM ÊNFASE EM EXTRAÇÃO DE
CARACTERÍSTICAS DA FACE**

Trabalho de Conclusão de Curso, apresentado para obtenção do grau de Bacharel no curso de Ciência da Computação da Universidade do Extremo Sul Catarinense, UNESC.

Orientador: Prof. Esp. Sérgio Coral

CRICIÚMA

2018

TIAGO ALEFF DA SILVA

**PROTÓTIPO DE RECONHECIMENTO FACIAL COM ÊNFASE EM
EXTRAÇÃO DE CARACTERÍSTICA DA FACE**

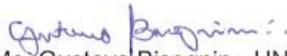
Trabalho de Conclusão de Curso aprovado pela Banca Examinadora para obtenção do Grau de Bacharel, no Curso de Ciência da Computação da Universidade do Extremo Sul Catarinense, UNESC, com Linha de Pesquisa em Teoria da Computação.

Criciúma, 29 de junho de 2018.

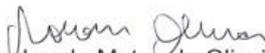
BANCA EXAMINADORA



Prof. Esp. Sérgio Coral - UNESC - Orientador



Prof. Me. Gustavo Bisognin - UNESC



Prof. Dr. Marlon de Matos de Oliveira - UFSC

A todos que estiveram presentes na minha caminhada, sem vocês nada seria possível

AGRADECIMENTOS

Primeiramente, agradeço aos meus pais, Marcio Rosa e Rosani Silva, pelo apoio e conselhos que me encaminharam na vida.

Agradeço também a o meu orientador Sérgio Coral que me guiou durante este trabalho com muita paciência.

A todo corpo docente do curso de Ciência da Computação da Unesc pelos conhecimentos que até aqui adquiridos.

Agradeço também a minha namorada Morgana Roos pela compreensão nos vários momentos que me ausentei.

Agradeço também aos meus colegas da faculdade pelo apoio e ajuda, em especial a Leandro Justin.

Também gostaria de agradecer aos colegas de trabalhos que também contribuíram para o desenvolvimento deste trabalho e a todas as pessoas que contribuíram diretamente e indiretamente na minha jornada.

“Aqueles que se sentem satisfeitos sentam-se e nada fazem. Os insatisfeitos são os únicos benfeitores do mundo.”

Walter S. Landor

RESUMO

Com o intuito de dar continuidade ao trabalho iniciado por Minatto (2016) o propósito abordado nesta pesquisa é o desenvolvimento de um protótipo de software capaz de realizar a identificação e verificação da identidade de indivíduos. A detecção facial é caracterizada pela capacidade do programa em encontrar as faces quando existem em uma imagem, enquanto que, o reconhecimento facial se define pela verificação e ou identificação. A biblioteca de Visão Computacional OpenCV foi utilizada para atingir os objetivos desta pesquisa possibilitando a utilização de algoritmos otimizados. O reconhecimento facial deve ser feito através de um algoritmo de extração de característica da face e neste projeto foi escolhido o algoritmo *Local Binary Patterns* cujo a principal vantagem em relação as técnicas baseadas em aparência é sua menor sensibilidade a iluminação do ambiente. Como resultado foi desenvolvido um protótipo que realiza o a detecção facial e reconhecimento facial observando os requisitos do reconhecimento facial. Este protótipo foi validado em uma base de imagens obtendo o melhor resultado uma acurácia de 81,8% e em seu pior resultado com uma acurácia de 54,1%. Apresentando queda no seu desempenho quando existem algumas variações, como o fundo do ambiente e presença de expressões faciais, caracterizando por meios destes resultados a sua utilização para ambientes controlados.

Palavras-chave: Biometria. Extração de Características da Face, Reconhecimento Facial

ABSTRACT

Aiming to continue started by Minatto (2016), the purpose approached in this research is the development of a software prototype capable of performing the identification and verification of a individual's identity. Facial detection is described by the program's capability of finding existing faces in an image, while facial recognition is defined by the verification or identification. The OpenCV Computational Vision Library was used to achieve the objectives of this research, making possible the use of optimized algorithms. The facial recognition must be made through an algorithm that extracts the characteristics of a face, and on this Project the Local Binary Patterns algorithm was chosen, whose main benefit, regarding the techniques based on appearance, is its smaller sensibility to ambient light. As a result, a prototype that attains facial detection and facial recognition, observing the requirements for facial requirements, was developed. This prototype was evaluated on an image basis, which best obtained result had an accuracy of 82,8%, and, as a worse result, 54,1% accuracy. It presented decay in its performance when some variations are present, as the ambient background or facial expressions, being then characterized by these results to usage in controlled environments.

Key Words: Biometry. Facial Characteristics Extraction. Facial Recognition.

LISTA DE ILUSTRAÇÕES

Figura 1 - Etapas para o Reconhecimento Facial	19
Figura 2 - Ramificações de Técnicas Utilizadas para Reconhecimento Facial	20
Figura 3 – Face Média	23
Figura 4 - Projeção de clusters em espaço bidimensional. Quatro clusters correspondente a variações das faces de diferentes sujeitos	26
Figura 5 - Problema do Coquetel, em Inglês <i>cocktail-party problem</i>	28
Figura 6 - Arquitetura I: procurando imagens bases estaticamente independente....	29
Figura 7 - Encontrando coeficientes estatisticamente independente	29
Figura 8 - Representação do grafo sobre os pontos de referências.....	30
Figura 9 - Nós rotuladas com <i>jets</i>	31
Figura 10 - Face Bunch Graph como representação gera das faces	32
Figura 11 - Aplicação do limiar sobre a matriz	34
Figura 12 - Resultado do LBPH em diferentes tipos de iluminação	34
Figura 13 - Amostra do primeiro grupo de imagens	40
Figura 14 - Amostra do segundo grupo de imagens	40
Figura 15 - Amostra do terceiro grupo de imagens	41
Figura 16 - Amostra do quarto grupo de imagens	41
Figura 17 - Modelagem do protótipo	42
Figura 18 - Tela principal do protótipo.....	43
Figura 19 - Tela para adicionar novos indivíduos a base de bases.....	44
Figura 20 - Estrutura do Protótipo	45
Figura 21 - Detecção de faces	46
Figura 22 - Estrutura das Imagens Salvas no SO	48
Figura 23 - Classe Responsável por Realizar o Treinamento de Imagens e Identificadores	49
Figura 24 - Verificando a Identidade de um Indivíduo	50

LISTA DE TABELAS

Tabela 1 - Tabela de confusão sobre o reconhecimento facial	51
Tabela 2 - Grupo A. Tabela de confusão sobre o reconhecimento facial	52
Tabela 3 - Grupo B. Tabela de confusão sobre o reconhecimento facial	53
Tabela 4 - Grupo C. Tabela de confusão sobre o reconhecimento facial.....	53
Tabela 5 - Grupo D. Tabela de confusão sobre o reconhecimento facial.....	53

LISTA DE ABREVIATURAS E SIGLAS

AAM	Active Apparence Model
EBGM	Elastic Bunch Graph Matching
FBG	Face Bunch Graphs
FN	Falso Negativo
FP	Falso Positivo
ICA	Independent Component Analysis
LDA	Linear Discriminant Analysis
LBP	Local Binary Patterns
OpenCV	Open Source Computer Vision Library
PC	Principais Componentes
PCA	Principal Component Analysis
VP	Verdadeiro Positivo
VN	Verdadeiro Negativo

SUMÁRIO

1 INTRODUÇÃO	9
1.1 OBJETIVO GERAL	10
1.2 OBJETIVOS ESPECÍFICOS	11
1.3 JUSTIFICATIVA	11
1.4 ESTRUTURA DO TRABALHO	12
2 BIOMETRIA	14
2.1 HISTÓRIA	14
2.2 TÉCNICAS BIOMÉTRICAS.....	15
2.2.1 Impressões Digitais.....	15
2.2.2 Íris.....	16
2.2.3 Voz.....	17
3 RECONHECIMENTO FACIAL.....	18
3.1 EXTRAÇÃO DE CARACTERÍSTICAS DE IMAGENS.....	19
3.2 ANÁLISE DE COMPONENTES PRINCIPAIS	21
3.3 ANÁLISES DISCRIMINANTE LINEAR.....	25
3.4 ANÁLISE DE COMPONENTES INDEPENDENTES	28
3.5 COMBINAÇÃO DE GRAFOS EM GRUPOS ELÁSTICOS	31
3.6 PADRÃO BINÁRIO LOCAL.....	35
4 TRABALHOS CORRELATOS.....	38
4.1 DESENVOLVIMENTO DE UM PROTÓTIPO DE SOFTWARE DE RECONHECIMENTO FACIAL DE TEMPO REAL PARA REGISTRO ELETRÔNICO DE PONTO EM AMBIENTES INDOOR COM UTILIZAÇÃO DO DISPOSITIVO KINECT	38
4.2 ANÁLISE DE CARACTERÍSTICAS EM REDES NEURAS ARTIFICIAIS APLICADAS EM VISÃO COMPUTACIONAL COM ÊNFASE EM RECONHECIMENTO FACIAL	38
4.3 RECONHECIMENTO AUTOMÁTICO DE PLACAS VEICULARES UTILIZANDO A TECNOLOGIA OCR E A PLATAFORMA RASPBERRY PI APLICADA NA FISCALIZAÇÃO ELETRÔNICA DE RODOVIAS	39
5 SOFTWARE DE RECONHECIMENTO FACIAL COM ÊNFASE EM EXTRAÇÃO DE CARACTERÍSTICAS DA FACE	41
5.1 COMPONENTES DO PROTÓTIPO.....	41

5.1.1 Framework PyQt5	41
5.1.2 FrameWork OpenCv	42
5.2 BASE DE DADOS	42
5.3 DESENVOLVIMENTO DO PROTÓTIPO DE SOFTWARE	45
5.3.1 Tela Inicial	45
5.3.2. Adicionar novo indivíduo	46
5.3.3 Estrutura do protótipo	47
5.3.4 Entrada de Imagens	48
5.3.5 Detecção Facial	49
5.3.6 Adição Novas Identidades e Reconhecimento Facial	50
5.4 ANÁLISE DOS RESULTADOS OBTIDOS	53
6 CONCLUSÃO	58
REFERÊNCIAS	59

1 INTRODUÇÃO

As implicações sociais e culturais das tecnologias de reconhecimento facial são de grande alcance, mas a atual lacuna de desempenho nessa questão entre máquinas e o sistema visual humano serve de amortecedor ao ter que lidar com essas implicações. (TAIGMAN, 2014, tradução nossa).

O significado para a palavra biometria vem da língua grega, onde *bio* significa vida e *metron* medida. O seu significado designa um método automático de reconhecimento individual por meio das medidas biológicas, anatômicas e fisiológicas, e também características comportamentais do indivíduo. (BIOMETRIA, 2016).

A biometria existe a milhares de anos. Não é uma ideia recente como várias pessoas acreditam. Os antigos egípcios, sumérios e outros povos já usavam técnicas de biometria. Os egípcios frequentemente escreviam diversas características-chaves as quais seriam registradas e conseqüentemente referenciada em ordem para verificar a identificação pessoal em negócios importantes. (ASHBOURN, 2014, tradução nossa). Para Vigliuzzi (2006), um exemplo das anotações utilizadas em negócios: Douglas, pele clara, cabelos castanhos, olhos castanhos, altura mediana, dedos médios, cicatriz na testa do lado esquerdo, entre outros. Dessa forma, eles conseguiam identificar quase todas as pessoas quando necessário.

Para Cavalcanti (2005), é possível dividir as técnicas biométricas em físicas e comportamentais. Dentre os métodos físicos mais estudadas podemos citar: face, impressão digital, retina, íris, formato da orelha, impressão da mão e DNA. E os métodos comportamentais são assinaturas, voz e dinâmica de digitação, são alguns métodos disponíveis.

Os sistemas biométricos podem realizar a identificação e verificação. A identificação ocorre quando os dados obtidos a partir de um sujeito externo ao sistema são comparados com todos os dados já armazenado no sistema. Enquanto, a verificação ocorre quando um indivíduo fornece algum identificador de algum registro que existe no banco de dados e seus dados são comparados especificamente com aquele registro. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa). Os sistemas de reconhecimento facial podem ser classificados em duas categorias. Primeiramente, há os sistemas que podem ser agrupados por cenário controlado, onde o sujeito a ser

testado está localizado em uma posição conhecida com o mínimo de interferências possíveis. O segundo grupo, são classificados por ambientes de teste não controlado, porque podem haver muitas variáveis no ambiente a serem consideradas. (CAVALCANTI, 2005). Tendo em vista estes conceitos, este trabalho irá considerar somente os ambientes de testes controlados.

De acordo com Face... (2016, tradução nossa), para implementar o algoritmo pela extração de características da face existem algumas abordagens predominantes para a resolução deste problema, sendo por exemplo a abordagem geométrica e a fotométrica. Como há interesse por estas áreas de pesquisa, vários algoritmos de extração de características são desenvolvidos para que o ocorra a validação. Alguns dos algoritmos estudados são a Análise de Componentes Principais (PCA), Análise discriminante Linear (LDA) e Combinação de Grafos em Grupos Elásticos (EBGM).

Em seu Trabalho de Conclusão de Curso, Minatto (2016) propôs um protótipo para reconhecimento facial utilizando o algoritmo Viola-Jones e Rede Neural implantado no ambiente Raspberry Pi. Entretanto, nesse protótipo não é desenvolvido o identificador e nem o verificador que classifica a identidade do sujeito como conhecida ou não. Agora, este novo trabalho vem apresentar uma continuidade à os estudos iniciados por Minatto (2016) implementando, além dos princípios de detecção facial, as técnicas necessárias para que ocorra a verificação e identificação do indivíduo.

Deste modo, é proposto como tema de estudo deste trabalho, a aplicação das técnicas necessárias para implementar o identificador possibilitando a identificação e verificação da identidade de indivíduos, por meio do algoritmo de extração de características Padrão Binários Locais (LBP) do inglês, *Local Binary Patterns*.

1.1 OBJETIVO GERAL

Este trabalho tem como objetivo desenvolver um protótipo que seja capaz de realizar a verificação e identificação facial.

1.2 OBJETIVOS ESPECÍFICOS

Dentre os principais objetivos que este estudo pretende alcançar, podemos destacar:

- a) compreender os algoritmos responsáveis pelo reconhecimento e verificação facial;
- b) identificar as etapas necessárias para o desenvolvimento do protótipo;
- c) implementação do protótipo em um computador pessoal;
- d) aplicação do algoritmo de extração de características em uma base de dados.

1.3 JUSTIFICATIVA

Fortalecido pela ampla difusão das câmeras digitais presente em vários computadores pessoais e outros dispositivos móveis. Um programa de reconhecimento facial nos dias atuais tem um potencial muito grande em usabilidade atualmente sendo a escolha óbvia em muitas situações. Outro porquê dessa escolha, é a fácil obtenção de uma foto de um indivíduo, estando presentes em carteiras de motoristas, identidades e passaportes, podendo a partir delas realizar a identificação ou verificação do mesmo. (ASHBOURN, 2014, tradução nossa).

O reconhecimento facial tem sido usado em várias situações delicadas, como o combate de passaporte fraudados, suporte para fortalecer a legislação, identificar crianças desaparecidas, minimizar fraudes em falsificação de identidade e entre outros. (FACE..., 2016). Conforme o Babcock (2015), um exemplo de um sistema de reconhecimento facial, o projeto *Next Generation Identification*, cujo objetivo é a busca por faces automaticamente em um repositório e traz uma lista de candidatos potenciais com base em uma foto submetida por um agente autorizado.

Mesmo havendo pessoas contrárias ao uso de sistema de identificação de faces, ela é muito aceita pela sociedade quanto a sua utilização, porque para analisar os dados de uma imagem este método não utiliza nenhum dispositivo invasivo. (CAVALCANTI, 2015). As técnicas invasivas são aquelas que necessitam de interação com o usuário, enquanto não há interação são conhecidas como técnicas

não invasivas. (TEIXEIRA, 2011). Por exemplo, se em uma investigação o sujeito que não concordar em colaborar com a investigação poderá ser utilizada uma técnica não invasiva, sendo o reconhecimento de face uma possível escolha, para realizar a verificação de identidade.

O LBP tem sido um dos descritores mais largamente usado por causa da sua resistência a mudanças de iluminação, baixa complexidade de implementação, e capacidade para codificar detalhes finos. A aplicação do LBP para o reconhecimento facial se deve ao fato que uma imagem é composta por micro padrões. (AHONEN; HADID; PIETIKAINEN, 2006, tradução nossa; BRAHNAM et al., 2016, tradução nossa). Segundo Delbiaggio et al. (2017) em uma comparação de algoritmos de extração de características o LBP obteve bons resultados na maioria dos testes realizados.

A escolha pela continuação da pesquisa sobre reconhecimento facial, levou-se em conta o processo que se levou até chegar ao resultado atual do protótipo de Minatto (2016). Considerando os estudos já realizado na etapa de detecção facial, será possível avançar para as próximas etapas e concluir o protótipo realizando os estudos e a implementação necessária para o reconhecimento facial.

1.4 ESTRUTURA DO TRABALHO

A seguir a estrutura básica deste trabalho, sendo constituído por cinco capítulos. Onde o primeiro capítulo é formado pela introdução, que abordam assuntos sobre o objetivo geral e específico, seguido pela justificativa do trabalho.

O segundo capítulo, aborda a história e conceitos básicos sobre a biometria. Destacando em subcapítulos seguintes diferentes abordagens para identificação do indivíduo por meio de características com maior potencial de identificação do indivíduo.

O capítulo 2 é apresentado os conceitos específicos à abordagem do reconhecimento facial como a técnica biométrica. Além disso, também são abordados diferentes algoritmos para extração de características de face.

O capítulo 4 é apresentado os trabalhos correlacionados com esta pesquisa, mostrando diferentes abordagens adotadas pelos autores.

O quinto capítulo expõe o trabalho proposto por esta pesquisa, apresentado a metodologia que será utilizada ao decorrer da pesquisa, bem como os recursos necessários e o cronograma para a continuação desta pesquisa.

Sexto e último capítulo apresenta a conclusão desta pesquisa e aborda sugestões para trabalhos futuros.

2 BIOMETRIA

O termo biometria para Jain, Ross e Nandakumar (2011, tradução nossa), é definido pelo reconhecimento automatizado de indivíduos baseados em suas características comportamentais e individuais.

2.1 HISTÓRIA

Em 1870 Bertillon desenvolveu na França um sistema manual que seria capaz de identificar qualquer um pelas medidas do corpo. Em 1882 foi nomeado chefe de uma agência de identificação, estabelecido por meio de uma conexão com a Prefeitura de Polícia de Paris. Então neste período, o sistema começou a ser utilizado por todas as estações de polícia, prisões e penitenciárias da França. (BERTILLON, 1896, tradução nossa). Porém, a Bertillonage não tinha muita automação nas realizações dos vários testes, o que aumentava as chances de erro e não podia garantir variações entre indivíduos. Por este motivo, foi utilizada outra técnica mais simples e precisa, que envolve a comparação manual de impressões digitais. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa). O Bertillonage não pode ser considerado um sistema biométrico porque carecia de automação, mas lançou base para os sistemas de reconhecimentos modernos. Os sistemas totalmente automáticos para o reconhecimento de pessoas. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa).

Em 1892, Frances Galton desenvolve um sistema detalhado de classificação por impressão digitais nas quais ele apresentou um novo sistema usando a impressão de todos os dez dedos. As características (minúcias) que Galton usou para identificar indivíduos ainda são usadas até hoje. E estas minúcias são frequentemente chamados detalhes de Galton. (NSTC, 2009, tradução nossa). Graças aos estudos de Galton, a técnica de Bertillonage foi sendo substituída pelos avanços das técnicas de análise florenses para impressões digitais. (JAIN; ROSS; NANDAKUMAR, tradução nossa).

Trauring (1963, tradução nossa), propôs no seu artigo um sistema de identificação, descentralizado e automático que poderia ser utilizado em bancos ou

como motivo de segurança, sendo possível o uso através da comparação com as minúcias dos dedos. O rápido avanço da tecnologia utilizado em reconhecimento de impressões digitais nos últimos anos acabou impulsionando a popularidade dos sistemas de reconhecimento automatizados desta técnica foi. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa).

Outras modalidades de biometria como voz, geometria da mão e assinatura foram também desenvolvidas depois da metade de 1960 e no começo de 1970. Por volta da década de 1990, o avanço dos hardwares e avanços consideráveis no reconhecimento estatísticos de padrões e na visão de computador, culminou em um avanço sofisticado para a biometria. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa).

2.2 TÉCNICAS BIOMÉTRICAS

O sistema biométrico mede uma ou mais características físicas ou comportamentais, incluindo impressões digitais, palma, rosto, íris, retina, orelha, voz, assinatura, veias da mão, odor ou a informação do DNA de um indivíduo para determinar e verificar sua identidade. Essas características são referidas por diferentes termos, tais como traços, indicadores, identificadores ou modalidades. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa)

Segundo Ashbourn (2000, tradução nossa), uma questão é constantemente levantada quando o assunto é biometria, cujo questionamento busca obter a resposta de qual a técnica biométrica é a melhor. Porém, a resposta depende do que estamos tentando buscar, com quem e em que condições o sistema estará submetido. Uma solução que funciona bem em um ambiente controlado e fechado pode não ser eficiente se implantado em um aeroporto onde o ambiente é aberto e movimentado.

2.2.1 Impressões Digitais

A pele nas mãos e das solas dos pés não possuem cabelo e nem glândulas de óleos, a pele da palma da mão e dos pés possuem um padrão de fluxo de cume e

vales (minúcias) também conhecida como cumes de fricção, ajudam a mão a agarrar objetos aumentando o atrito entre a mão e o objeto (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa). Estes cumes podem ser utilizados como identificadores biométricos, se padrão de formação é requerido como único e imutável. Em casos de gêmeos idênticos os padrões dos cumes não se repetem entre eles. No entanto, quando essas estruturas são danificadas superficialmente acabam alterando o padrão dos cumes de fricção apenas temporariamente. Porém, se em decorrência da lesão ocasionar em uma cicatriz permanente o padrão anterior não ocorrerá novamente. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa).

Existem outros vários detalhes nas impressões digitais que fornecem informações importantes, entretanto os mais utilizados em sistemas automatizados são as terminações das linhas e as suas bifurcações. As extrações das características por meios destas estruturas dependem da qualidade da amostra que podem variar com um tamanho aproximado de 250 bytes a 700 bytes. (COSTA et al., 2006).

O padrão dos sulcos intercalados e vales na ponta de um dedo é referido como impressão digital. O uso da impressão digital para identificação pessoal começou a ser promovida no final do século XIX ao início do século XX, embora as descobertas sobre impressões digitais em artefatos arqueológicos indicam que mesmo naquela época já se tinha noção do potencial do uso dessa técnica. Originalmente, as impressões digitais eram obtidas rolando a ponta de um dedo sobre uma superfície de papel, os avanços da tecnologia permitiram que por meio de sensores e equipamentos de baixo custo a possibilidade de renderização das digitais. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa).

2.2.2 Íris

Para Jain, Ross e Nandakumar (2011, tradução nossa) o uso do globo ocular começou a ser explorada com mais entusiasmo em 1993 com os avanços no conhecimento de reconhecimento de íris. Vários componentes do olho podem ser utilizados como indicadores biométricos, mas é a íris que vem sendo mais utilizada entre os indicadores.

O reconhecimento de íris é um processo de reconhecimento pessoal que analisa padrões aleatórios da íris. O método automatizado desta técnica é relativamente nova, existindo em patentes desde 1994. (NSTC, 2009).

Para realizar a identificação é necessária uma câmera digital de alta qualidade. As câmeras comerciais digitais atuais utilizam luz infravermelha para iluminar a íris sem causar desconforto ou causar algum dano para o indivíduo. (NSTC, 2009). Segundo Daugman (2003), considerando a íris é um órgão interno e extremamente visível, protegido, imutável e possui uma morfogênese aleatória de alta complexidade estatística.

2.2.3 Voz

Os sistemas biométricos que utilizam a voz são mais indicados para casos de serviços em que somente a voz está disponível para análise, como o tele comércio. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa).

A identificação de um indivíduo utilizando a voz é comportamental e fisiológica, onde os componentes fisiológicos responsáveis pela vocalização possuem tamanho e forma definida e são invariantes, mas as características comportamentais mudam com a idade, emocional, condições médicas, etc. Estes sistemas de identificação podem adotar duas direções para a identificação, onde a análise se dará na pronúncia de uma frase predeterminada ou por meio de uma frase indeterminada. A diferença, no sistema determinado o desenvolvimento do mesmo é mais simples, porém menos seguro para a identificação. Enquanto que, no sistema indeterminado é mais complexo o seu desenvolvimento, porém é mais segura em relação a identificação. (JAIN; ROSS; NANDAKUMAR, 2004, tradução nossa).

3 RECONHECIMENTO FACIAL

A face é a característica biométrica largamente utilizada pelo ser humano. Exercendo um papel importante, tal que, é colocada em vários tipos de cartões de identificação. Sendo a parte frontal da cabeça que abriga os principais órgãos que permite o ser humano perceber o mundo ao seu redor. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa).

Por motivos de segurança, a demanda por sistemas automáticos de reconhecimentos de faces vem aumentando nas áreas privadas e públicas. Imagens obtidas por câmeras de monitoramento e até imagens estáticas, esta possibilidade advém da evolução da Tecnologia da Informação. O Reconhecimento Facial faz uso de técnicas e métodos científicos de diversas áreas, como visão computacional, reconhecimento de padrão e inteligência computacional. (PRODOSSIMO; CHIDAMBARAM; LOPES, 2011).

O reconhecimento facial vem sendo utilizado como um recurso a favor da segurança. Utilizado na área de segurança pública para realizar a identificação de pessoas principalmente as que já estão cadastradas na base de imagens da mesma. Este recurso também pode ser utilizado para controle de acesso a ambientes privados onde o acesso deve ser controlado por algum mecanismo. (PRODOSSIMO; CHIDAMBARAM; LOPES, 2011).

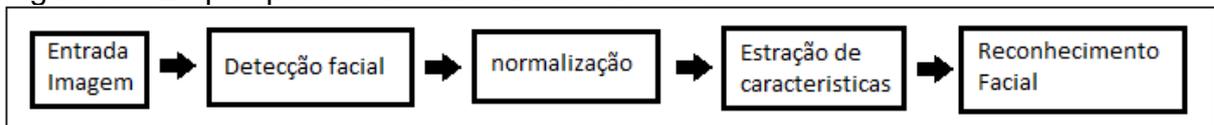
Para Jain, Ross e Nandakumar (2011, tradução nossa) é muito importante o uso da face como identificador biométrico na aplicação da lei e interação homem computador. Na visão computacional, a identificação pessoal baseada em faces é um problema clássico que ganhou muita atenção. Os seres humanos são excelentes em reconhecer rostos familiares, mesmo que os processos cognitivos que estão envolvidos nessa tarefa não sejam muito bem compreendidos. Isto faz que treinamentos de reconhecimento facial em como o ser humano faz são árduos de serem feitos.

Uma imagem pode ser definida por uma função bidimensional, $f(x, y)$, em que x e y representam coordenadas cartesianas, e a amplitude de f em qualquer ponto das coordenadas (x, y) é chamada de intensidade ou nível de cinza da imagem nesse ponto. (GONZALES; WOODS, 2010).

O reconhecimento facial pode ser definido no processo de estabelecimento da identidade da pessoa com base em suas características faciais. Basicamente, o problema da identificação facial envolve a comparação de duas imagens da face e determinar se elas são a mesma pessoa, classificando-a como conhecida ou desconhecida. (KSHIRSAGAR; BAVISKAR; GAIKWAD, 2011, tradução nossa). Enquanto os humanos parecem estar aptos em determinar a similaridade entre duas imagens faciais adquiridos sob diversas condições. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa).

A figura 1 mostra um diagrama representado um funcionamento básico de um sistema de reconhecimento facial. No momento que uma imagem é enviada como entrada para análise no sistema será iniciado a detecção e marcação da face na imagem. Em seguida, a imagem será normalizada, nesta etapa serão executadas geralmente três tarefas: primeiro, normalização espacial, alinhando os centros dos olhos para um local predefinido e ajustando a quantidade de pixels entre os olhos (distância interocular) através de rotação e transformações de escalas. Segundo, é extraída somente a região do rosto para o reconhecimento não seja afetado por fatores que não sejam referentes da face em si, como o estilo do cabelo. Terceiro, a região extraída da face é convertida em um vetor normalizando os pixels que serão armazenados nesses vetores. Deste modo, é extraída as características que possuem um alto potencial de discriminação para que ocorra o reconhecimento facial na fase seguinte. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa)

Figura 1 – Etapas para o Reconhecimento Facial



Fonte: Adaptado de Jain, Ross e Nandakumar (2011, tradução nossa).

3.1 EXTRAÇÃO DE CARACTERÍSTICAS DE IMAGENS

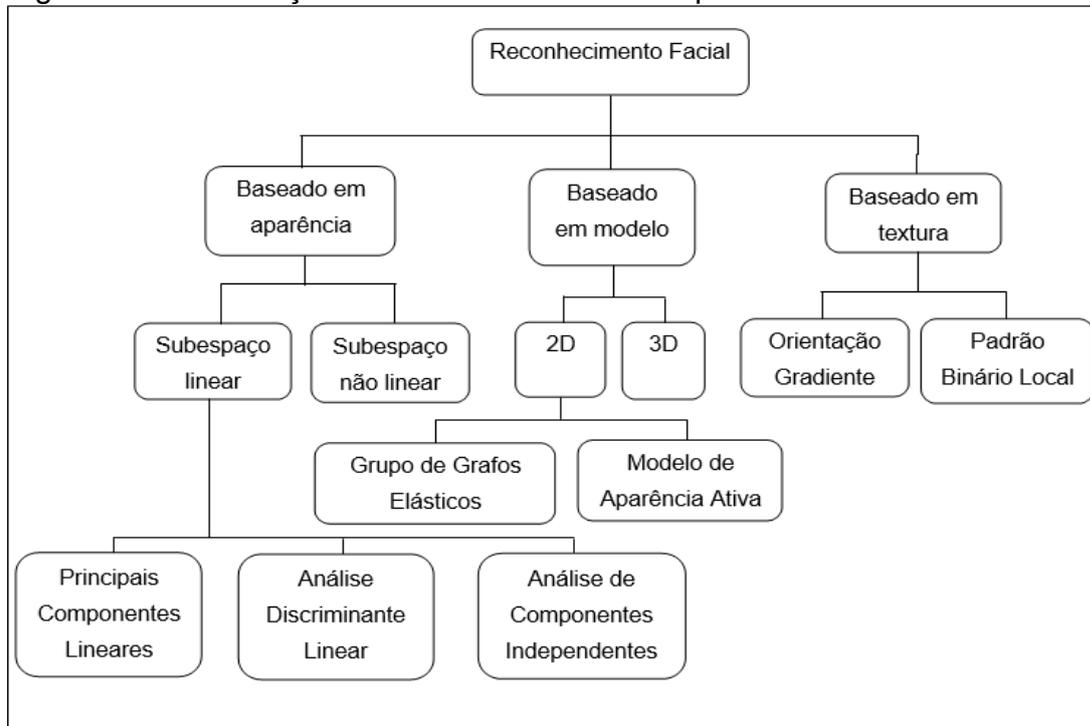
Existem três principal abordagem comumente utilizada para o reconhecimento da face, são eles as técnicas baseadas em aparência, baseada em

modelos e aproximação baseada em textura (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa):

- a) baseada em aparência: geram uma representação compacta de toda a região da imagem, mapeando a imagem que possui uma alta dimensão em um espaço de dimensão inferior. Este espaço de dimensão inferior é definido por um conjunto de vetores. O mapeamento pode se caracterizar como linear ou não linear, os sistemas mais utilizados são *Principal Component Analysis (PCA)*, *Linear Discriminant Analysis (LDA)* e *Independent Component Analysis (ICA)*, implicam em projeções lineares;
- b) baseado em modelo: constroem um modelo 2D ou 3D da face que facilita a comparação da imagem da face mesmo em variações de poses. Exemplo de modelos 2D são o *Face Bunch Graphs (FBG)* e *Active Appearance Model (AAM)*, enquanto que, *Morphable Model* é um modelo 3D;
- c) aproximação baseada em textura: buscam encontrar locais com características importantes que são invariantes na pose ou variações de iluminação. Exemplos de tais características incluem orientações gradientes e *Local Binary Patterns (LBP)*.

Na figura 2 tem-se um modelo básico das técnicas empregadas para o reconhecimento facial. No topo, possuímos o objetivo principal que é a detecção da face. Nos níveis abaixo são mostradas as principais técnicas disponíveis para o reconhecimento facial.

Figura 2 – Ramificações de Técnicas Utilizadas para Reconhecimento Facial



Fonte: Adaptado de Jain, Ross e Nandakumar (2011, tradução nossa).

Análise de Componentes Principais, em inglês como é conhecida como *Principal Component Analysis* (PCA), e Análise Discriminante Linear, em inglês como também é conhecida *Linear Discriminant Analysis* (LDA), são poderosas ferramentas para redução de dimensionalidade e extração de características muito utilizada para aplicações de reconhecimento de padrões. (SAHOOLIZADEH; HEIDARI; DEGHAN, 2008, tradução nossa).

3.2 ANÁLISE DE COMPONENTES PRINCIPAIS

As imagens capturadas possuem uma determinada resolução geralmente superiores a 1000 *pixels*. Se a tarefa de identificação facial usar a imagem de dimensão original da imagem capturada para realizar a identificação de identidade ocorrerá perda de desempenho, isto se, a técnica é baseada em reconhecimento baseado em aparência. (DELAC, 2008, tradução nossa).

A Análise de Componentes Principais (PCA), também é conhecida como a transformada de *Karhunen-Lòeve*, é uma técnica utilizada para a extração e representação de dados muito utilizada nas áreas de reconhecimento de padrões e

visão computacional (Yang et al., 2004, tradução nossa). O PCA é uma técnica de pré-processamento comumente usada no processamento de dados e com alto potencial de redução de dimensionalidade e aprendizado de máquina. (GOODALL; JOLLIFFE, 1988, tradução nossa; HOTELLING, 1933, tradução nossa)

O principal objetivo do PCA é reduzir a dimensionalidade do conjunto de dados consistindo de um largo número de variáveis inter-relacionadas, mantendo o máximo possível da variação presente no conjunto de dados. Isto é adquirido transformando em um novo conjunto de variáveis, os principais componentes (PC's), que não são correlacionados, e como são ordenados para que os primeiros conservem a maior parte da variação presente em todas as variáveis originais da determinada imagem. (GOODALL; JOLLIFFE, 1988, tradução nossa).

Para calcular os componentes principais, seguimos os seguintes passos (PANCERI et al., 2015):

- a) obter a matriz M que representa os dados o conjunto de dados;
- b) calcular o valor médio destes dados;
- c) calcular a matriz A gerada a partir da subtração da média M de todo o conjunto de dado;
- d) calcular a matriz de covariância L , a partir do produto entre a matriz A e sua transposta;
- e) calcular os autovalores e autovetores da matriz de covariância;
- f) formar uma matriz diagonal a partir dos autovetores da matriz de covariância arranjados de modo que a primeira linha, o elemento $(0,0)$, seja o autovetor correspondente ao autovetor de maior valor, e assim sucessivamente até a última linha correndo ao menor autovalor. Essa matriz será chamada de matriz da Transformada de Hotelling;
- g) no último passo, cada vetor de treinamento é projetado no espaço reduzido através da matriz da Transformada de Hotelling. O descritor PCA é calculado por uma combinação linear de autovetores com os vetores originais.

O conjunto de imagem de entrada é normalizada e então as *eigenfaces* são calculadas a partir desse conjunto que serão utilizados para treinamento. (AGARWAL et al, 2010, tradução nossa). O conjunto de imagem são transformados em vetores

bases a partir destes vetores bases é calculado a “face média”. Os vetores bases são representações matemática da imagem que foram obtidas através do treinamento sobre o conjunto. A partir da “face média” é calculada a covariância que forma os autovetores que possuem maiores valores de variação em relação ao conjunto de treinamento são denominados de “componentes principais”, estes componentes principais também podem ser chamados de *eigenfaces*. (GONG; MCKENNA; PSARROU, 2000, tradução nossa). Levando em consideração a maior variação dentro do conjunto de imagens de rosto as eigenfaces mais adequadas compreendem um subespaço M-dimensional que é chamado de “espaço de face”. (AGARWAL et al., 2010, tradução nossa).

O algoritmo da *eigenface* PCA é obtido através do cálculo da “face média”, que é formado pelo autovetores, no “espaço da face” e então é calculado o quão distante cada face está do conjunto da média. A diferença é usada para calcular a matriz de covariância para o conjunto de dados. A covariância entre os dois conjuntos de dados mostra o quanto os conjuntos estão correlacionados. (GONG; MCKENNA; PSARROU, 2000, tradução nossa; SIROVICH; KIRBY, 1987, tradução nossa).

Para identificar se uma imagem pertence a um determinado conjunto a imagem será normalizada e então projetada em uma *eigenface* que será possível obter os pesos do vetor, em seguida será calculada o intervalo entre o peso do vetor de entrada e todos os pesos de vetores obtidos pelo treinamento, se o intervalo obtido for menor que todos pesos do treinamento então a imagem é conhecida, porém se o intervalo for maior a imagem é considerada desconhecida. (KIRBY; SIROVICH, 1990, tradução nossa).

O principal objetivo do PCA é encontrar os vetores que melhores contabilizam a distribuição das imagens dentro do espaço da imagem. Por isso, estes vetores são os eigenvectors da matriz de covariância que corresponde as imagens faciais originais, e por causa que eles são parecidos em aparência, são denominados de eigenfaces.

Para o treinamento do conjunto de imagens se $\Gamma_1, \Gamma_2, \dots, \Gamma_M$. A média do conjunto pode ser definida por:

$$\Psi = \frac{1}{M} \sum \Gamma_n \quad (1)$$

Cada face se diferencia da média pelo vetor:

$$\Phi = \Gamma_i - \psi \quad (2)$$

Figura 3 – Face Média



Fonte: Jain, Ross e Nandakumar (2011, tradução nossa)

3.3 ANÁLISES DISCRIMINANTE LINEAR

A Análise Discriminante Linear (LDA), em inglês como é conhecida *Fisherfaces* se mostrou muito útil em diversos campos, incluindo ciências físicas, biológicas e sociais, engenharia e medicina. Sendo pertinente a problemas associadas a separação estatística entre classes ou grupos distintos e com a alocação de entidades a grupos, sendo finito a existência dos grupos é conhecida *a priori*. (MCLACHLAN, 2004, tradução nossa)

Assim como o PCA, o LDA busca as melhores características que representam as características da face. (SUSHMA; ADGAONKAR, 2011, tradução nossa). O LDA apresenta uma abordagem de classificação de modelos baseada na técnica *Fisher's Linear Discriminat*. Possui como característica a baixa sensibilidade as variações de luz, direções de iluminação e expressões faciais. Permitindo a análise de imagens de uma face particular, sob a variação de iluminação, direções diferentes da fonte de luz, mas com poses fixas, posicionadas em um subespaço linear 3D da alta dimensionalidade da imagem facial. Para minimizar o problema que as imagens possuem com relação a sombras sobre as faces, conseqüentemente provocando desvios de seu subespaço linear, então é projetada linearmente a imagem dentro de um subespaço e é desconsiderada essa região do espaço de baixa dimensão. (BELHUMEUR; HESPNHA; KRIEGMAN, 1997, tradução nossa).

O LDA está fortemente relacionado com a análises de variância e análise de regressão, que assim como o LDA tentam expressar uma variável dependente através da combinação linear de outras características ou medidas. (MCLACHLAN, 2004, tradução nossa)

No contexto de paradigmas de reconhecimento baseados em modelos, acredita-se que geralmente os algoritmos baseados em LDA apresentam resultados superiores aos algoritmos baseados em PCA. Porém, o PCA se mostra menos sensível a diferentes bases de dados de treinamento e também apresenta melhores resultados quando a base de dados possui uma dimensão reduzida. (MARTÍNEZ; KAK, 2001, tradução nossa).

Segundo Webb e Copsey (2011, tradução nossa), no critério de Fisher considerando duas classes distintas de variáveis a abordagem adotada por Fisher,

que deu origem a técnica (*Fisher's Linear Discriminat*), foi determinar a melhor combinação linear das variáveis que melhor separam as duas classes tanto quanto possível. Ou seja, buscando a direção ao longo do qual as duas classes estão melhor separadas em algum sentido. O critério proposto por Fisher é a razão entre a variação dentro da classe e entre classes. Formalmente, a direção que se deseja é representada pela direção w , tal que (fórmula 3):

$$J_F = \frac{|w^T(m_1 - m_2)|^2}{w^T S_w w} \quad (3)$$

Onde m_1 e m_2 representam as classes e S_w é a matriz de covariância da amostra dentro da própria classe agrupada. O critério de *Fisher* fornece um mapeamento para uma dimensão reduzida considerando a discriminação entre duas classes. Para obtermos uma regra de alocação, deve ser definido um limite para atribuímos x a classe W_1 . (WEBB; COPSEY, 2011, tradução nossa).

Na equação 4, a matriz de dispersão de dentro da própria classe (S_w) representa as variações na aparência de um mesmo indivíduo, variações estas que podem ser causadas por variação de iluminação, expressão faciais e obstruções (uso de óculos, por exemplo). Enquanto que, na equação 5, a matriz de dispersão entre classes (S_b) representa variações de identidade de diferentes indivíduos. (SUSHMA; ADGAONKAR, 2011, tradução nossa). Estes conceitos apresentados respectivamente podem ser expressos da seguinte forma:

$$S_w = \sum_{i=1}^c \sum_{j=1}^{N_i} (V_j - \mu_i)(V_j - \mu_i)^T \quad (4)$$

$$S_b = \sum_{j=1}^c N_j (\mu_j - \mu_0)(\mu_j - \mu_0)^T \quad (5)$$

Onde V_j é a i -ésima amostra (imagem), μ_j é a média da classe j , C é o número de classes (quantidade de indivíduos do banco de dados), N_i é o número de amostras na classe j e μ representa a média de todas as classes. (HESPNHA, KRIEGMAN, 1997, tradução nossa)

O objetivo é maximizar o S_b enquanto minimiza o S_w , isto é alcançado quando as colunas de vetores da matriz de projeção (W_{lda}) são os autovalores de $S_w^{-1} S_b$, ou seja (fórmula 6):

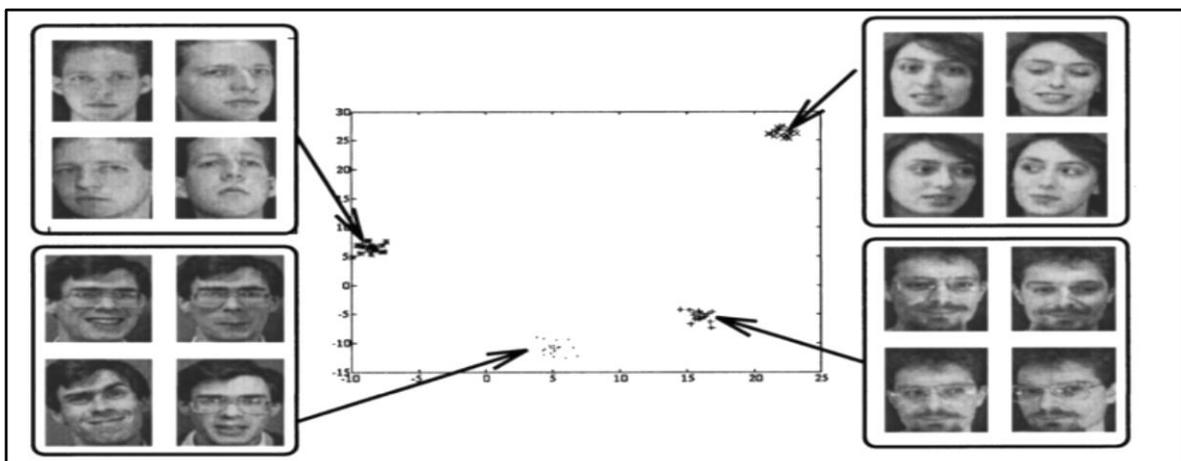
$$w_{opt} = [w_1 w_2 \dots w_m] \quad (6)$$

Onde W i-énésimo conjunto de autovalores generalizados de S_b e S_w correspondentes a um conjunto de autovalores (*eigenvectors*) generalizados decrescentes. Para evitar que S_w se torne singular, pode-se utilizar o PCA para reduzir a dimensão do espaço de recurso para $M - C$, em seguida, a utilização do modelo de LDA é definido pela fórmula anterior para reduzir a dimensão entre $C - 1$, onde C é o número de classes, podemos obter o W_{opt} pela fórmula 7. (SUSHMA; ADGAONKAR, 2011, tradução nossa)

$$w_{opt} = W_{lda} W_{pca} \quad (7)$$

Quando as imagens são projetadas no espaço bidimensional, resultará em uma projeção onde as características das faces do mesmo indivíduo tendem a ficar agrupadas em um *cluster*, enquanto que, características de indivíduos diferentes fiquem agrupados em *cluster* bem separados de outros indivíduos. Veja a figura 4. (ETEMAD; CHELLAPPA, 1997, tradução nossa).

Figura 4 – Projeção de clusters em espaço bidimensional. Quatro clusters correspondente a variações das faces de diferentes sujeitos



Fonte: Etemad e Chellappa (1997, tradução nossa).

3.4 ANÁLISE DE COMPONENTES INDEPENDENTES

A Análise de Componentes Independentes, do inglês Independent Component Analysis (ICA), é a versão mais difundida do PCA. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa). O objetivo do ICA é decompor um sinal observado em uma combinação linear de sinais independentes desconhecidos, com o fim de resolver problemas de separação de fonte cega. Enquanto o PCA minimiza a projeção de erro de dados de entrada comprimidos por meio de estatística de segunda ordem, o ICA busca reduzir a dependência entre segunda ordem e a alta ordem em dados de entrada. (BHELE; MANKAR, 2012, tradução nossa; DRAPER et al., 2003, tradução nossa).

Buscando resolver problemas como o *cocktail-party problem*, por exemplo: onde em uma sala há pelo menos dois locutores discursando simultaneamente. E para cada locutor existe um microfone gravando os discursos da sala. Na equação 8, cada sinal gravado pelos microfones pode ser representado por $x_1(t)$ e $x_2(t)$, x_1 e x_2 são a amplitude e t o índice de tempo. Estes sinais gravados é a soma ponderada dos sinais de fala emitidos pelos alto-falantes, que pode ser denotado por $s_1(t)$ e $s_2(t)$. Isto pode ser representado com a seguinte equação linear. (HYVÄRINEN; OJA, 2000, tradução nossa):

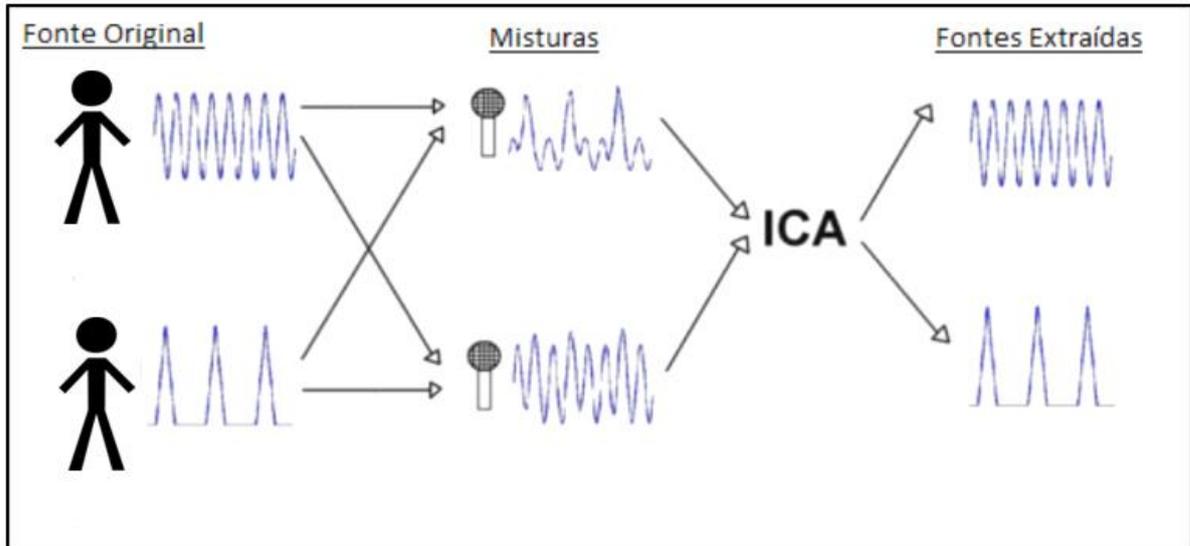
$$x_1(t) = a_{11} + a_{12} + s_2 \quad (8)$$

$$x_2(t) = a_{21}S_1 + a_{22} + s_2 \quad (9)$$

Onde, na equação 9, a_{ij} são parâmetros que dependem das distâncias para os locutores. Para se obter os valores de S_1 e S_2 usando somente o sinal de entrada gravado pelos microfones, sem conhecer a_{ij} torna o problema mais complexo para se resolver utilizando métodos clássicos. Então, utilizando-se ICA é possível usar o S_1 e S_2 e obter algumas informações estatísticas para estimar a_{ij} . O ICA pode estimar a_{ij} considerando que são independentes, assim podendo ser utilizada para separar os sinais originais $s_1(t)$ e $s_2(t)$ de $x_1(t)$ e $x_2(t)$ que são os sinais combinado de ambos.

(HYVÄRINEN; OJA, 2000, tradução nossa). Veja o exemplo da figura 5 quando os sinais são computados pelo ICA:

Figura 5 – Problema do Coquetel, em Inglês *cocktail-party problem*



Fonte: Adaptado de Langlois, Chartier e Gosselin (2010, tradução nossa).

O ICA é um método de aprendizado não supervisionado baseado em estatística de ordem superior. O sistema de modelo do ICA é dado como (EKENEL; SANKUR, 2004, tradução nossa):

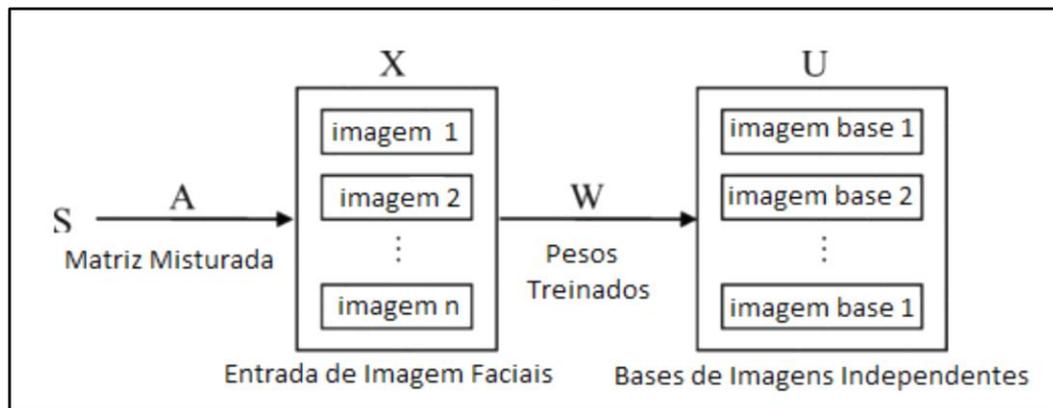
$$X = AS \quad (10)$$

Logo, equação 10, A é definida como sendo a matriz de mistura desconhecida, S denota os vetores de origem desconhecidas do sinal e X indica a matriz de dados. (CARDOSO, 1997, tradução nossa). No método ICA, a única informação que possuímos é as observações, nem a matriz mesclada (A) e nem a distribuição das fontes são conhecidas. Sob os pressupostos de que as fontes são estaticamente independentes e não gaussianas, encontramos a matriz não misturada W, maximizando a medida da independência. Portanto, estima-se uma matriz de separação W, que, em condições ideais é o inverso da matriz de mistura A. (EKENEL; SANKUR, 2004, tradução nossa).

No contexto do reconhecimento facial, há duas formas principais para se aplicar o ICA. Na arquitetura I, conforme mostrado a figura 6, as imagens de entrada

em X são consideradas uma mistura estatisticamente independente de base de imagens S combinadas em uma matriz misturada desconhecida A . O algoritmo do ICA aprende os pesos da matriz W , que será usado para recuperar um conjunto de bases de imagens independentes nas fileiras de U . Nesta arquitetura, as imagens faciais são variáveis e os *pixels* são valores que fornecem as observações para as variáveis. A separação da fonte é consequência da projeção das imagens de entrada no peso de vetores aprendidos, produzindo imagens de base independente, processo executado no “espaço da face”. (DRAPER et al., 2003, tradução nossa).

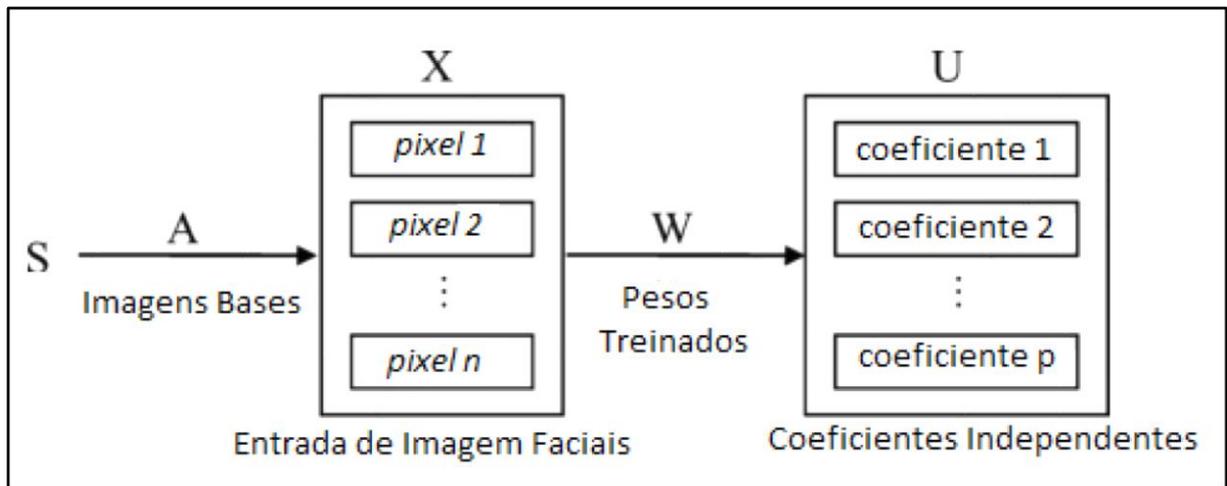
Figura 6 – Arquitetura I: procurando imagens bases estaticamente independente



Fonte: Adaptado de Draper et al., (2003, tradução nossa).

Na arquitetura II, o objetivo do ICA é encontrar o coeficiente estatisticamente independente para os dados de entrada. Diferentemente da arquitetura I, a separação de fonte é executada nos *pixels*, e para cada linha aprendida de pesos de matriz W corresponde a uma imagem. Sendo o A , a matriz inversa de W , que contém as imagens bases em suas colunas. Os coeficientes de fonte estatisticamente independentes em S compreendem a entrada de imagens e são obtidas na coluna U . (DRAPER et al., 2003, tradução nossa).

Figura 7 – Encontrando coeficientes estatisticamente independente



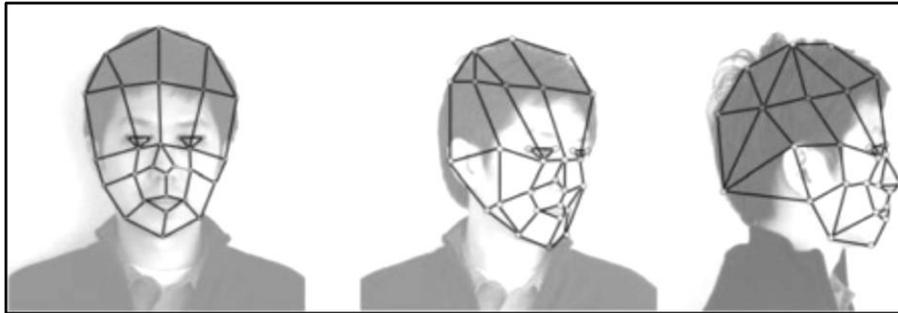
Fonte: Adaptado de Draper et al., (2003, tradução nossa)

3.5 COMBINAÇÃO DE GRAFOS EM GRUPOS ELÁSTICOS

A Combinação de Grafos em Grupos Elásticos, inglês como é conhecido *Elastic Bunch Graph Matching* (EBGM). É uma técnica de reconhecimento facial baseada em modelo, que por meio de uma variação da representação da imagem original pode permitir a identificação de indivíduos em poses diferentes. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa).

Desenvolvido para manipular grandes quantidades de galerias e variações de poses, melhorando a acurácia em relação a comparação entre as imagens no reconhecimento facial. Uma face é representada por um grafo G composta por N vértices e conectados por E arestas. Os nós ficam localizados em pontos de referências da face também chamados de *fiducial points*, podendo ser representado pela seguinte fórmula: $\vec{x}_n, n = 1, \dots, N$. Os pontos de referências podem ser exemplificados pelas pupilas, extremidades da boca, ponta do nariz, e parte superior e inferior das orelhas, etc. O grafo resultante é um objeto adaptável desde que sua estrutura geométrica seja adaptada para a estrutura do objeto. (WISKOTT, 1997, tradução nossa). A figura 8 mostra os pontos referência sobre a face.

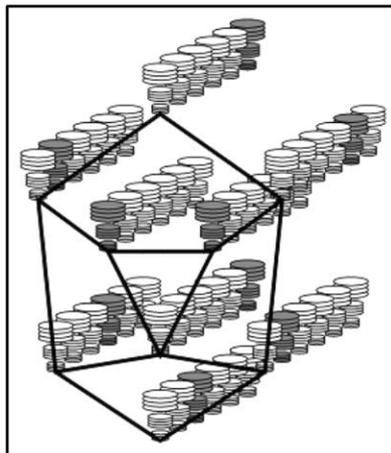
Figura 8 – Representação do grafo sobre os pontos de referências



Fonte: Wiskott (1997)

Cada nó do grafo é nomeado com um conjunto de coeficientes de Garbor, frequentemente chamados de *jet* (figura 8). O *jet* caracteriza a informação da textura local ao longo dos pontos de referências, as arestas que conectam quaisquer vértices são rotuladas com base na distância entre estes pontos de referências. O coeficiente de Garbor pode se obtido através de uma posição da imagem e por meio de convolução da imagem utilizando um complexo filtro de *2D Garbor* centralizado nesse local, e por meio de variação a orientação e a frequência do filtro, obtém-se um conjunto de coeficientes. O uso de *fiducial points* permitem que sejam desenvolvidos grafos parciais mesmo que o rosto esteja inclinado e ou obstruído. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa).

Figura 9 – Nós rotuladas com *jets*



Fonte: Wiskott (1997).

O modelo de *Face Bunch Graph* (FGB), em tradução livre para Grupo de Grafos de Face, é similar a uma estrutura de pilha que constituirá em uma

representação dos grafos faciais obtidos através do conjunto de treinamento. O FGB pode ser obtido em dois estágios. (WISKOTT, 1997, tradução nossa).

- a) primeiro estágio, deve-se de forma manual marcar os pontos de referências na face para as imagens iniciais para se obter o desta imagem FGB. Os próximos grafos a partir de imagens serão obtidas semiautomáticamente, somente havendo intervenção se algum ponto de referência for gerado na posição incorreta da imagem. Este processo se repetirá para cada pose, exemplo: frontal, perfil e semi-perfil;
- b) segundo estágio: um FGB de um grafo de novas das novas imagens serão obtidas através de comparação entre uma representação do conjunto de grafos individuais que estavam em uma estrutura semelhante uma pilha. Cada vértice do FGB será rotulado com um conjunto de *Garbor jet* que representam a variação no ponto de referência da face dentro da população de usuário de treinamento.

Um conjunto de *jet* representam o mesmo ponto de referência da face podendo ser chamado de *bunch*. Por exemplo, um *bunch* que representa um ponto de referência da face como o um olho podem apresentar *jet* que representam características de olhos abertos, fechados, gênero masculino ou feminino. (WISKOTT, 1997, tradução nossa).

Os pontos de referências para as novas imagens são encontrados maximizando a similaridade entre um grafo ajustado para uma dada imagem FGB de pose idêntica, este processo é conhecido como EBGM. Na figura 10, é representada o processo de comparação entre uma imagem. A representação de um FGB com as pilhas de *jets* em cada nó do grafo e dada uma imagem de entrada inicial o processo de ajuste da imagem de entrada para um grafo da face. Ao comparar o grafo da imagem de entrada é comparado com os *jets conhecidos* e os pontos de referência correspondentes, o grafo da imagem de entrada é comparada com grafos do modelo e calculada a sua similaridade. Consecutivamente, o EBGM é mais robusto para variações de poses do que abordagens baseadas em aparências. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa).

Figura 10 – Face Bunch Graph como representação gera das faces



Fonte: Adaptado de Jain, Ross e Nandakumar (2011, tradução nossa).

Cada pilha de disco representa um *jet* somente o melhor representante do melhor vértice está destacado com sombreamento cinza.

3.6 PADRÃO BINÁRIO LOCAL

A textura de uma imagem é fácil para identificar e difícil de definir, como resultado na literatura possuem várias definições propostas. A textura é uma importante característica nas imagens, providenciando informações sobre propriedades físicas de objetos, tal como suavidade e dureza, ou diferenças na reflexão da superfície e também na cor do objeto. Ou ainda a textura pode ser definida como um padrão global que surge a partir de repetição de sub padrões locais. (BRAHNAM et al, 2016, tradução nossa).

Densidade da textura, é um fenômeno visual baseia-se no fato de que muitos objetos possuem em sua aparência algum tipo de padrão com uma certa regularidade. Nesse caso, a medida que os padrões aparecem mais densos e menos detalhados, mais distantes estarão do observador. As texturas também auxiliam na percepção do movimento, como, por exemplo, se girarmos uma esfera sem textura, nosso sistema de visão terá alguma dificuldade para perceber o seu movimento. (AZEVEDO; CONCI, 2003).

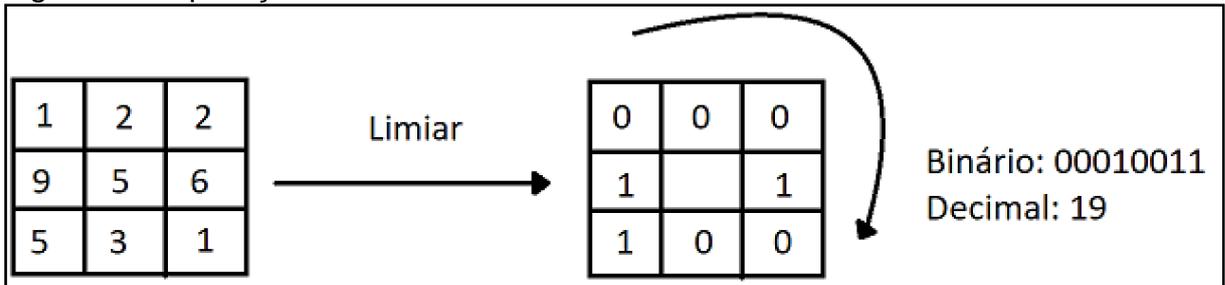
Muitos descritores de textura foram introduzidos, sendo discriminativos e computacionalmente eficientes. A performance oferecida por estes descritores tem conduzido progressos significantes. Sendo uns destes descritores de texturas o Local Binary Patterns (LBP) operador. (OJALA; PIETIKAINEN; MAENPAA, 2002, tradução nossa).

A metodologia do LBP pode ser adaptada e aplicada em tarefas de classificação ao ar livre, reconhecimento facial, detecção facial, reconhecimento de expressões faciais e recuperação de conteúdo de imagem obtendo excelentes resultados. Também pode ser aplicada em remoção de background e detecção de movimento de objetos em tempo real. (AHONEN; HADID; PIETIKAINEN, 2004, tradução nossa).

A ideia básica do LBP é resumir a estrutura local em uma imagem comparando cada pixel com o seu pixel vizinho. Obtendo um pixel central e limitando o valor de seus vizinhos com base no pixel central. Se a intensidade do pixel central é maior ou igual a do seu vizinho, então é atribuído ao pixel vizinho 1 e 0 se não. Estes rótulos binários compõe um histograma que descreve a textura da imagem ou região

em questão. Então com 8 pixel em torno do pixel central, irá permitir 2^8 possíveis combinações, chamados LBP. O primeiro LBP operador descrito na literatura usava 3x3 vizinhos, assim como na figura 11.

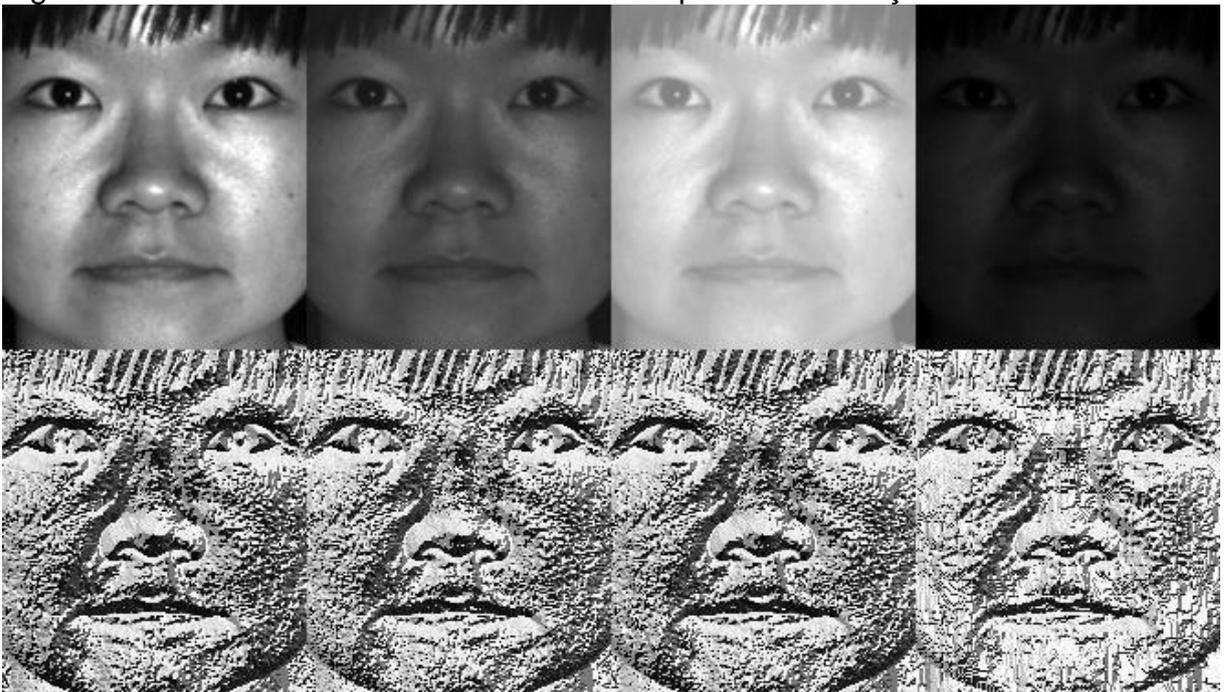
Figura 11 - Aplicação do limiar sobre a matriz



Fonte: Adaptado de OpenCv (2018, tradução nossa).

Posteriormente a publicação do LBP verificou-se que o operador falha em codificar detalhes em diferentes escalas. Então o operador foi estendido para usar uma variação de regiões. A ideia é alinhar um número arbitrário de vizinhos em um círculo com um raio variável, que permiti capturar as regiões com linhas, bordas e cantos.

Figura 12 - Resultado do LBPH em diferentes tipos de iluminação



Fonte: OpenCV (2018, tradução nossa).

Para incorporar as informações espaciais no modelo de reconhecimento facial e dividir a imagem LBP em regiões e extrair o histograma de cada região. Um vetor com as características é então obtido por concatenação dos histogramas locais. Estes histogramas são chamados de *Local Binary Patterns Histograms*. As diferenças entre dois histogramas diferentes indicam se ambos pertencem a mesma classe ou não, resultado no reconhecimento da face. (OPENCV, 2018, tradução nossa).

Do ponto de vista da classificação de padrões, um problema comum no reconhecimento facial é ter uma infinidade de classes e apenas algumas, possivelmente apenas uma, amostra (s) de treinamento por classe. Por esse motivo, classificadores mais sofisticados não são necessários, mas um classificador de vizinho mais próximo é usado (do inglês *nearest-neighbour*). Várias medidas possíveis de dissimilaridade foram propostas para histogramas como por exemplo a intersecção de Histogramas (do inglês *Histogram intersection*), estatística de probabilidade de Log (do inglês *Log-likelihood statistic*) e estatística Chi quadrado (do inglês *Chi square statistic*) (AHONEN, HADID, PIETIKÄINEN, 2004, tradução nossa):

4 TRABALHOS CORRELATOS

As diferentes abordagens para o reconhecimento facial são temas discutidos atualmente, neste capítulo veremos alguns trabalhos relacionados a esta pesquisa.

4.1 DESENVOLVIMENTO DE UM PROTÓTIPO DE SOFTWARE DE RECONHECIMENTO FACIAL DE TEMPO REAL PARA REGISTRO ELETRÔNICO DE PONTO EM AMBIENTES INDOOR COM UTILIZAÇÃO DO DISPOSITIVO KINECT

A dissertação de Guimarães (2015) para o mestrado em Gestão de Sistemas de Informação e Gestão do Conhecimento apresenta um protótipo de software para reconhecimento facial para uso em controle de frequência para horas trabalhadas de uma empresa.

Para o desenvolvimento do protótipo foram utilizados a linguagem de programação CSharp, bando de dados com Mysql, a biblioteca computacional *Open Source Computer Vision Library* (OpenCV) e ambiente de desenvolvimento Visual Studio 2010. E em relação aos *hardwares* utilizados destaca se o uso do *Kinect* da *Microsoft*. As imagens utilizadas no banco de dados foram obtidas através de voluntários e a escolha e ajuste de um ambiente para realização da identificação levaram 8 meses para ser concluído.

Como resultado do protótipo de Guimarães (2015) considera que a proposta se mostrou viável e de fácil utilização. Desde que, seja utilizado em ambientes controlados, pois a solução proposta se mostrou sensível a variações de ambiente e oclusões parciais da face devido ao uso de óculos, por exemplo. O *Kinect* se mostrou fundamental para os resultados finais do trabalho.

4.2 ANÁLISE DE CARACTERÍSTICAS EM REDES NEURAIS ARTIFICIAIS APLICADAS EM VISÃO COMPUTACIONAL COM ÊNFASE EM RECONHECIMENTO FACIAL

Em 2016 Minatto (2016) defendeu seu Trabalho de Conclusão de Curso de Ciência da Computação, onde ela dispõe a desenvolver um protótipo de reconhecimento facial sobre a plataforma *hardware* Raspberry Pi. Utilizando técnicas como o algoritmo Viola-Jones e rede neural artificial para melhorar o desempenho da aplicação quando as imagens possuem iluminação diversificadas causando problemas na detecção das faces.

No desenvolvimento do projeto foram utilizadas como ferramentas o *framework* OpenCV, o Sistema Operacional Raspbian para o Raspberry Pi e linguagem de programação C++, além das tecnologias citadas anteriormente como o algoritmo Viola-Jones e redes neurais.

Para Minatto (2016), o protótipo obteve uma alta taxa de acurácia em relação ao tempo de treinamento das Redes Neurais e mesmo sobre grande quantidade de imagens de treinamento. Mas quando o treinamento foi realizado no Raspberry Pi foi observado que o mesmo não tinha capacidade computacional necessária para realizar o treinamento das Redes Neurais. Por isso, é necessário um hardware mais robusto para que seja realizado o treinamento da rede neural.

4.3 RECONHECIMENTO AUTOMÁTICO DE PLACAS VEICULARES UTILIZANDO A TECNOLOGIA OCR E A PLATAFORMA RASPBERRY PI APLICADA NA FISCALIZAÇÃO ELETRÔNICA DE RODOVIAS

Em seu Trabalho de Conclusão de Curso de Ciência da Computação Zanette (2016) propôs um novo modelo para radares de fiscalização de velocidade eletrônica, que se baseia em ao menos dois radares localizados em posições e distâncias conhecidas entre si. Cada veículo que passa entre um radar será identificado e analisado quando atingir o próximo radar, sabendo a distância entre os radares poderá validado se este veículo excedeu a velocidade permitida de acordo com a distância entre os mesmos.

Para poder possibilitar a identificação automática das placas dos veículos foi utilizado o Raspberry Pi junto com uma câmera para capturar as imagens. Os meios utilizados para realizar tal tarefas se constituem em um circuito autorama com um conjunto de Raspberry Pi e câmera em associação com bibliotecas OpenCV e

Tesseract OCR. Como resultado, foi desenvolvido um protótipo capaz de realizar a identificação dos veículos pelas suas determinadas placas, entretanto, observou-se que o mesmo se mostrou sensível a variações de iluminação do ambiente. (ZANETTE, 2016).

5 SOFTWARE DE RECONHECIMENTO FACIAL COM ÊNFASE EM EXTRAÇÃO DE CARACTERÍSTICAS DA FACE

O projeto consiste em um protótipo que fará o reconhecimento facial através da captura de imagens de uma Webcam. O protótipo de reconhecimento facial terá como foco ambientes controlados, onde a luz, posição do indivíduo serão conhecidos. Dessa forma é minimizado as ocorrências de falsos positivos para a detecção e reconhecimento facial. O levantamento bibliográfico deste projeto permitiu a aplicação dos conceitos usados nesta sessão, bem como o desenvolvimento deste protótipo.

5.1 COMPONENTES DO PROTÓTIPO

Através da pesquisa realizada foi possível uma compreensão mais abrangente dos conhecimentos necessários para o desenvolvimento deste protótipo, possibilitando o uso de ferramentas que auxiliassem no desenvolvimento. O protótipo foi desenvolvido em cima do Sistema Operacional (SO) Windows, possibilitando o uso de diversos componentes que auxiliam no desenvolvimento do mesmo. Nas próximas sessões são abordados os componentes utilizados para o desenvolvimento.

5.1.1 Framework PyQt5

Apesar do OpenCv possuir uma implementação de interface básica, neste trabalho buscou-se a utilização de um *framework* de interface que facilitasse a implementação do protótipo, sendo a escolha o PyQt5. Esta permite que sejam utilizados vários componentes desde interfaces, *Threads* e Interfaces de conexão com o *Bluetooth* e *internet*. (PYQT5... 2018).

O PyQt5 é uma biblioteca que permite utilizar as funções do *framework* QT através da linguagem de desenvolvimento Python. O *framework* Qt é desenvolvido pela The Qt Company's que o disponibiliza sobre duas licenças disponíveis, tal como a versão *Open Source* (GPL) e a sua versão comercial. Na versão comercial há um

custo pela sua aquisição e uma versão para teste, diferentemente da licença GPL que necessita que o código fonte seja disponibilizado em <<https://www.qt.io/download>>. (RIVERBANK... 2018).

5.1.2 FrameWork OpenCV

É uma biblioteca *Open Source* para Visão Computacional licenciado para uso acadêmico e comercial, os produtos gerados através dessa biblioteca não precisam ter o código de fonte aberto. (OPENCV, 2018). Escrita em C e C++ que está disponível nas plataformas de Sistemas Operacionais Linux, Windows, Mac OS X e também nas linguagens de programação Python, Ruby, Matlab, entre outras. O OpenCv foi lançado em 1999, quando Gary Bradski que trabalhava na Intel lançou a biblioteca, com o objetivo para acelerar a Visão Computacional e inteligência artificial, providenciando uma sólida infraestrutura. A biblioteca OpenCv foi desenvolvido para obter um melhor desempenho computacional, devido a linguagem em que ela foi escrita e uso de múltiplos núcleos de processamento. (SHRESTHA; VENKITASUBRAMANIAN; MOENS, 2014, tradução nossa). Esta biblioteca já obteve mais de 14 milhões de *downloads*.

O OpenCv é destinado a resolver problemas de Visão Computacional, ela possui disponíveis algoritmos de baixo e alto nível para a detecção facial, detecção de pedestres, busca de características, e rastreamento de objetos. A biblioteca também conta com aceleração em GPU sem que seja necessário programar especificamente para o uso este recurso. (PULLI et al., 2012).

A biblioteca possui algoritmos específicos para a detecção facial e para o reconhecimento facial. Sendo que as técnicas Eigenfaces, FisherFaces e LBPH estão disponíveis no *framework*. (YOSHINAO, 2012, tradução nossa). Neste protótipo, serão utilizados os algoritmos referentes à os métodos LBPH e detecção facial.

5.2 BASE DE DADOS

Para identificar a eficiência do protótipo foi utilizado uma base de teste de faces disponibilizados gratuitamente pelo *Spacek (2008)*, cujo objetivo da base é ser

utilizado em pesquisas de Visão Computacional, disponibilizando o total de 7900 imagens agrupadas por quatro nível de complexidade em relação a iluminação, ambiente, expressões faciais e posição da face. Não há variação no estilo do cabelo, pois as fotos foram tiradas em uma única sessão. Ambos os grupos possuem imagens de homens, mulheres, raças diferentes e um total de 20 imagens por sujeitos.

O primeiro grupo (figura 13) é composto por um total de 153 indivíduos, com fundo verde, existindo pouca variação de iluminação e variação na posição da cabeça. Este é o grupo com menor complexidade.

Figura 13 – Amostra do primeiro grupo de imagens



Fonte: Spacek (2008)

O próximo grupo (figura 14) possui um total de 72 indivíduos. Possuindo variações da posição do sujeito na imagem, fundo vermelho, sombras, pouca variação na posição da cabeça e expressões faciais.

Figura 14 – Amostra do segundo grupo de imagens



Fonte: Spacek (2008)

O grupo seguinte (figura 15) engloba um total de 152 indivíduos. Possui a principal característica o fundo com cartazes com cores brilhantes, sombras, muita variação de escala e poucas expressões faciais.

Figura 15 – Amostra do terceiro grupo de imagens



Fonte: Spacek (2008)

O último grupo (figura 16) é considerado o mais complexo entre os quatro, composto por um total de 18 pessoas. Dispondo como principal característica as expressões faciais, o fundo não apresenta complexidade relevante, baixa incidência de sombras e há pouca variação de escala entre as fotos.

Figura 16 – Amostra do quarto grupo de imagens



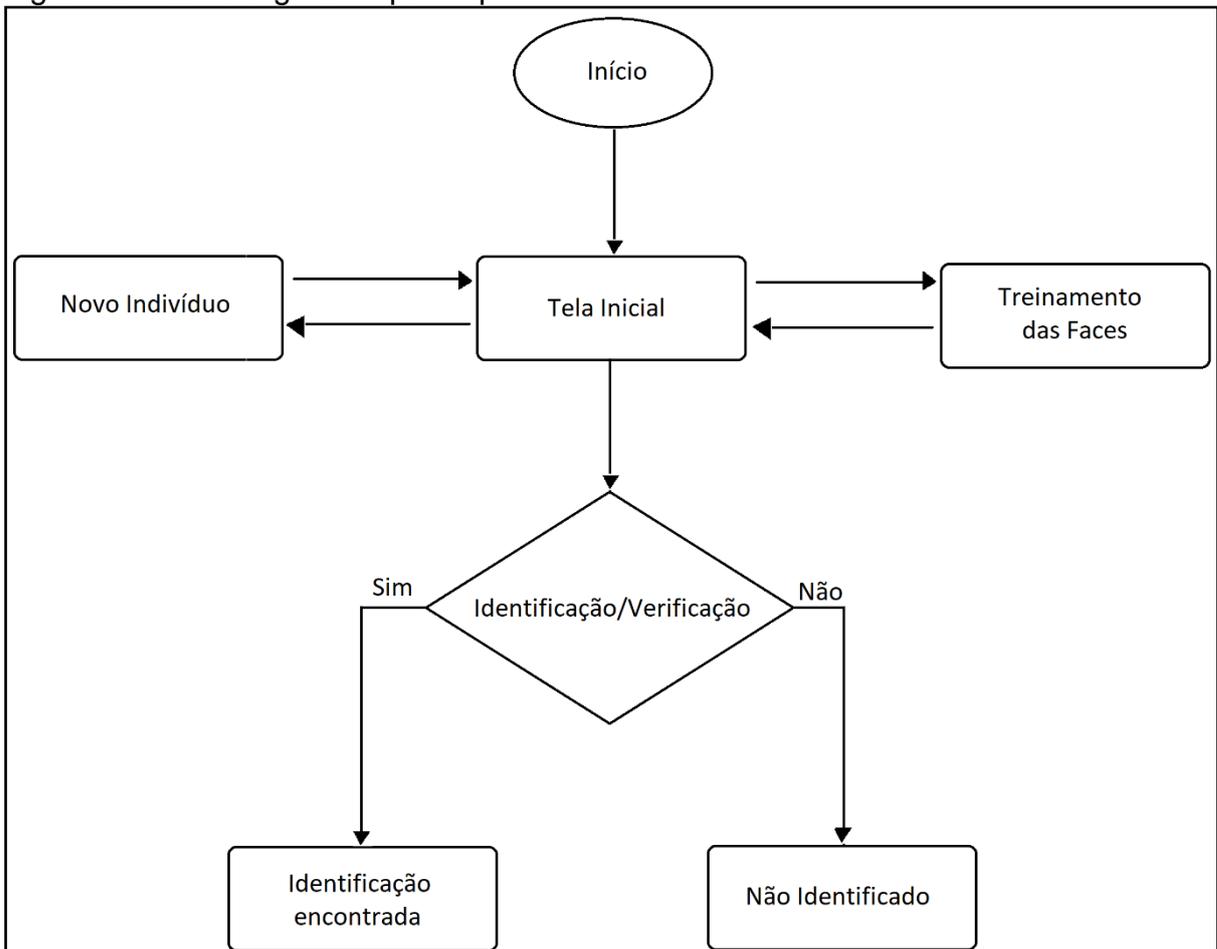
Fonte: Spacek (2008)

Será considerado para treino do nosso classificador o total de 1/3 por indivíduo e os outros 2/3 serão utilizados testes de identidade e os dados serão coletados análise dos dados. (KOHAVI et al., 2008, tradução nossa).

5.3 DESENVOLVIMENTO DO PROTÓTIPO DE SOFTWARE

O projeto desenvolvido pode ser expressado pelo seguinte fluxograma, figura 17. Nas próximas sessões serão abordadas as telas desenvolvidas nesta pesquisa.

Figura 17 - Modelagem do protótipo

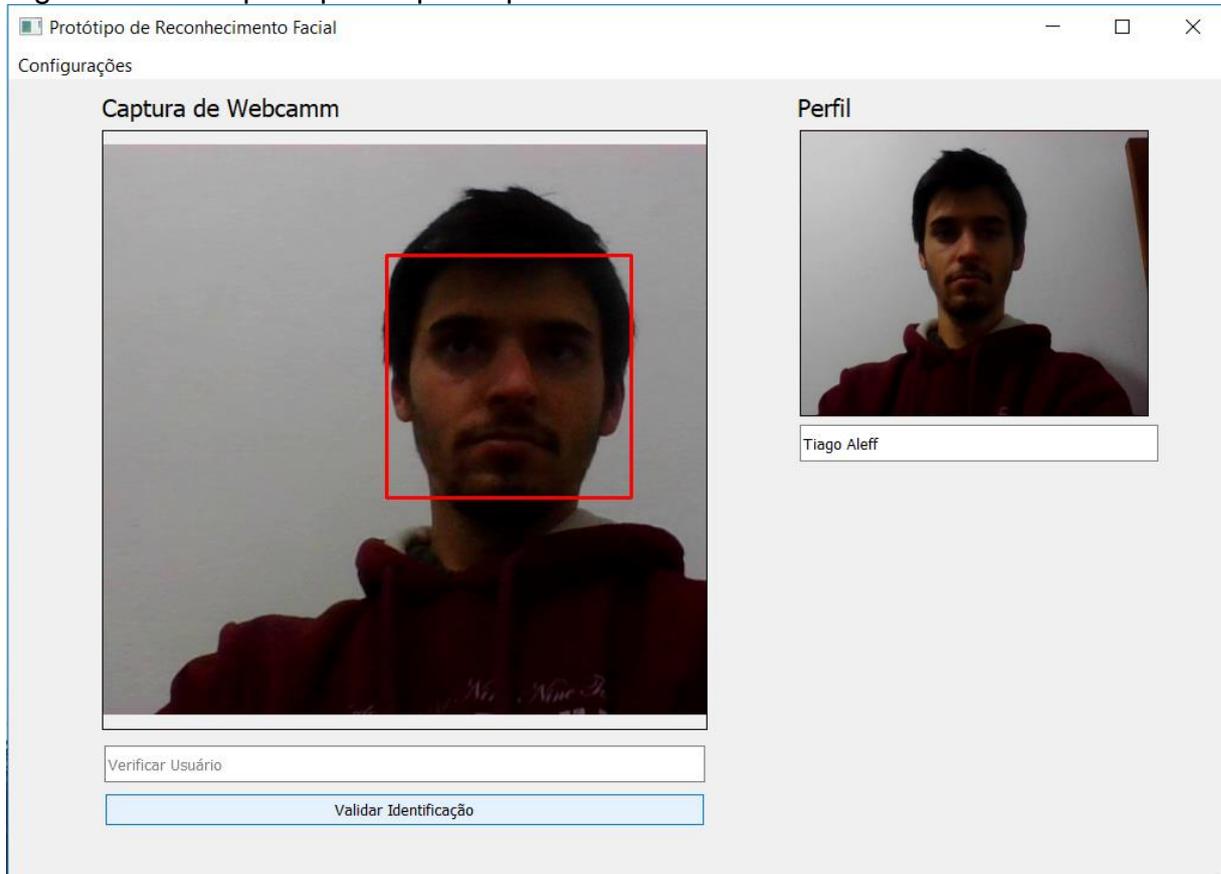


Fonte: Do autor.

5.3.1 Tela Inicial

A tela inicial é a principal tela do protótipo (figura 18), onde podemos acessar todos os recursos disponíveis, assim como treinamento, reconhecimento facial e acessar a funcionalidade para adicionar novos sujeitos a base de face.

Figura 18 - Tela principal do protótipo

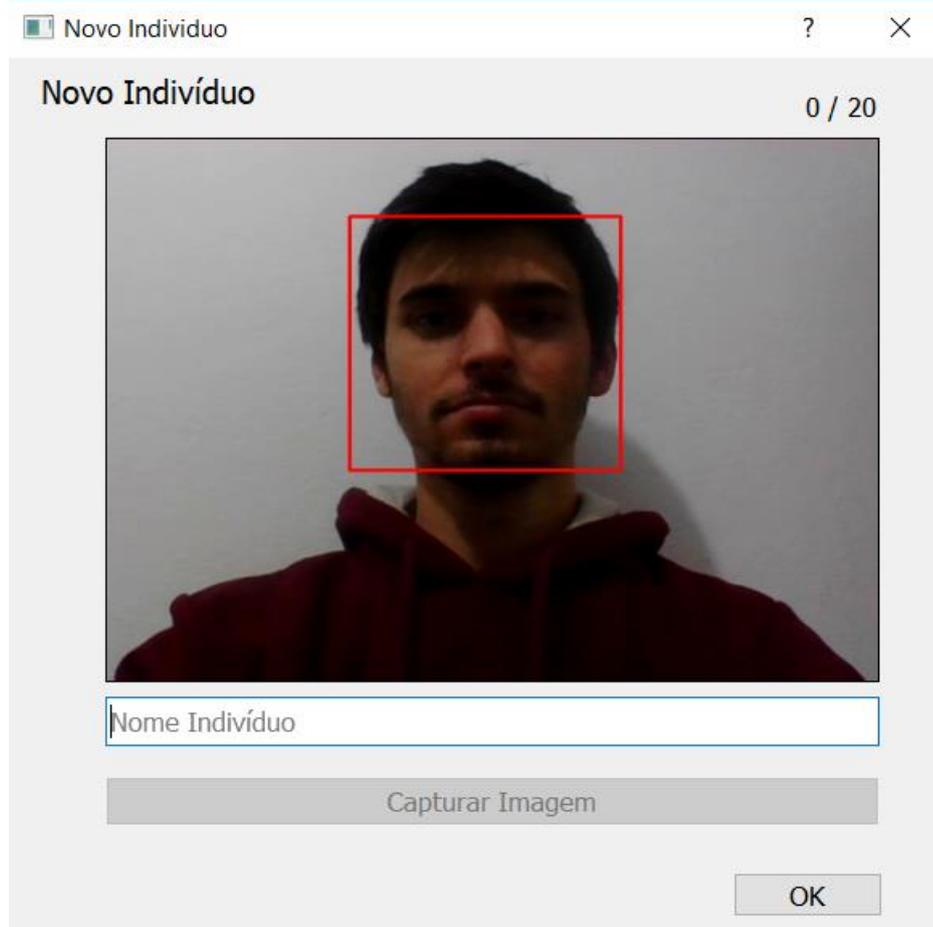


Fonte: Do autor

5.3.2. Adicionar novo indivíduo

A tela de novo usuário (figura 19) pode ser acessada através do item do menu configurações. Nessa tela é possível adicionar novas faces capturando até vinte (20) imagens, este valor foi definido por convenção. E após isso podemos realizar o treinamento destas novas faces utilizando o item do menu configurações e selecionando o item treinamento.

Figura 19 - Tela para adicionar novos indivíduos a base de bases

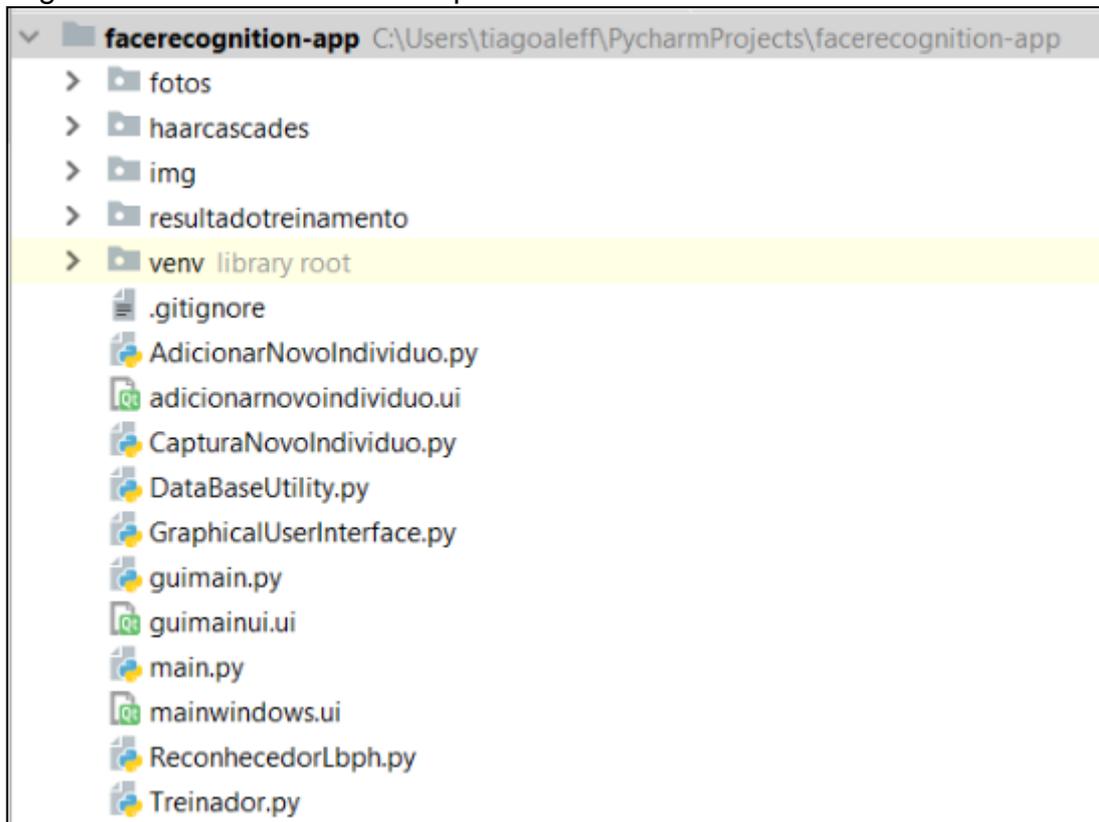


Fonte: Do autor.

5.3.3 Estrutura do protótipo

O protótipo foi desenvolvido através da linguagem de programação Python, na sua versão 3.6, e foi utilizado orientação a objetos para criar os componentes necessários para a modelagem do projeto. Na figura 20 vemos a estrutura final do protótipo.

Figura 20 - Estrutura do Protótipo



Fonte: Do Autor

O diretório *fotos* contém as imagens de faces obtidas pelo protótipo, já o diretório *haarcascades* possuem os arquivos necessários para a detecção facial, o diretório *resultadotreinamento* contém o arquivo gerado no treinamento do reconhecimento facial, enquanto que a pasta *venv* estão todas as bibliotecas e dependências para o funcionamento do protótipo. As classes criadas, arquivos de *interface*, conexão com o banco de dados e de configuração do *git* estão na raiz do projeto.

5.3.4 Entrada de Imagens

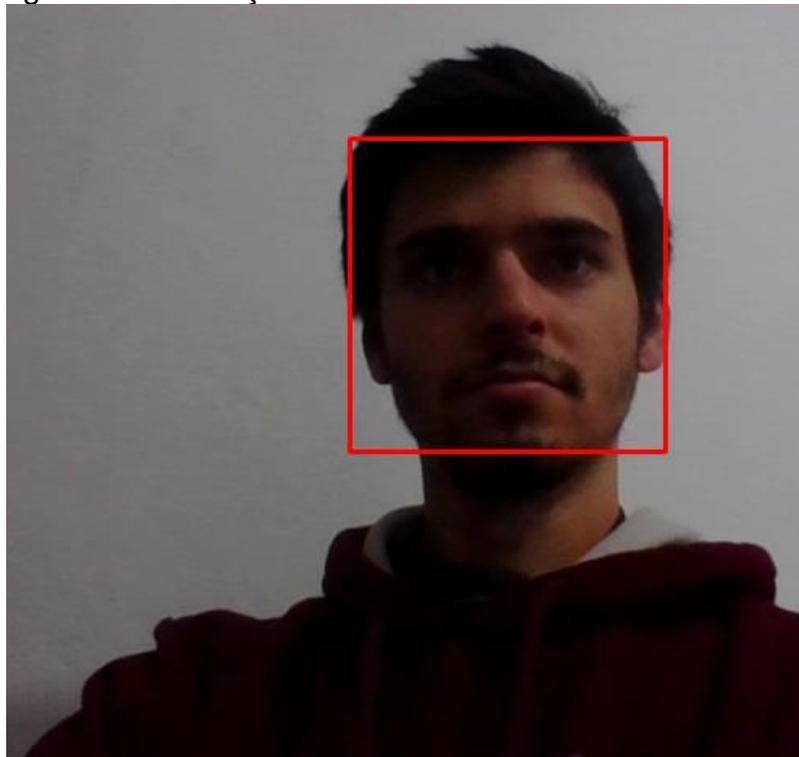
As imagens que são capturadas pelo protótipo são obtidas em tempo real pela Webcam. Elas devem passar por uma normalização antes de serem adicionadas ao banco de faces ou para serem utilizadas no reconhecimento da identidade do indivíduo. Dessa forma, se utiliza de alguns dos recursos disponibilizados pelo *framework* OpenCV.

Tanto para a detecção quanto para o reconhecimento facial, a imagem deve ser transformada para uma escala de cinza e redimensionada para um tamanho padrão reduzindo o processamento computacional posteriormente. A conversão para a escala de cinza ocorrerá a partir da chamada do método *cvtColor()* que irá receber como parâmetro a própria imagem e para que espectro de cor será transformada, neste caso, será utilizado o parâmetro *CV_BGR2GRAY* que aplicará o filtro cinza, alterando a cor da imagem.

5.3.5 Detecção Facial

A detecção facial é uma das etapas essenciais para o reconhecimento facial, que a partir de uma imagem o protótipo deve ser capaz de encontrar as faces existentes, se o protótipo não consegue encontrar as faces, então não será possível realizar o processo de reconhecimento facial (Figura 13).

Figura 21 - Detecção de faces



Fonte: Do autor.

A biblioteca do OpenCV possui uma série de rotinas disponíveis para facilitar o processo de detecção facial, entre estes processos, ele implementa o

método de classificação Haar, possibilitando a detecção de faces ou realizando o treinamento de novos objetos por meio de Aprendizado de Máquina (*FACE... 2018*, tradução nossa). Neste projeto foi utilizado o método de classificação disponível através dos arquivos *haarcascade_frontalface_default.xml* e *haarcascade_eye.xml*. Com estes arquivos podemos fazer a detecção tanto de faces como de olhos. Estes arquivos são resultados do treinamento da detecção facial, implementado através do método Viola-Jones.

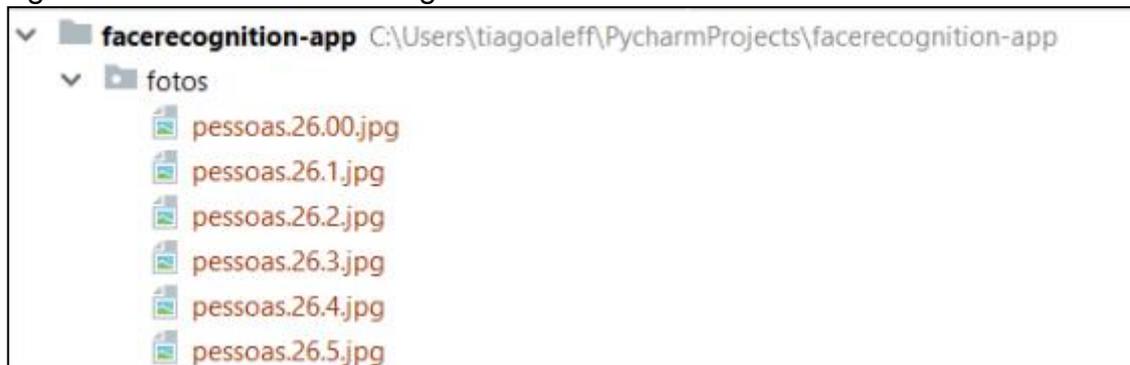
Para carregar os arquivos *haar* é necessário utilizar a classe *CascadeClassifier()*, responsável pela detecção de objetos e ou treinamento de novos objetos, que em seu construtor receberá o caminho relativo ou absoluto do arquivo e o *frame* que será verificado a existência da face. Assim, será possível fazer a predição de faces ou de olhos na imagem conforme o arquivo enviado no seu parâmetro, chamando o método *detectMultiScale()* da classe *CascadeClassifier()*. Para que o método funcione corretamente será necessário que seja aplicado o filtro de cor cinza sobre a face que será passada por parâmetro. (OPENCV, 2018, tradução nossa).

5.3.6 Adição Novas Identidades e Reconhecimento Facial

Para adicionar uma nova identidade no protótipo é fundamental que o processo de detecção facial esteja em execução funcionando adequadamente, ou seja, capturando as faces existentes. Após a captura da face, o frame que contém a face será dividido em dois, uma será processada e renderizada na interface do protótipo para que o usuário veja a detecção e a outra será redimensionada através do método *resize()* para 220 *pixels* de altura e 220 *pixels* de largura e em seguida, após ser normalizada, a imagem será salva em um diretório do SO.

Posteriormente, as imagens das faces salvas serão utilizadas para o treinamento das características da face. Para isso, é necessário a captura de até 20 imagens de um mesmo indivíduo, onde cada imagem capturada será nomeada com um identificador, que referencia uma identidade de uma pessoa, mais o seu enésimo valor dentro do total de faces capturadas de um mesmo indivíduo (Figura 22).

Figura 22 - Estrutura das Imagens Salvas no SO



Fonte: do Autor.

O treinamento destas imagens para realizar o reconhecimento facial será realizado por meio da biblioteca OpenCV que disponibiliza os recursos necessários para esta tarefa. Os recursos relacionados a reconhecimento facial são implementados pela classe *FaceRecognizer* e de suas subclasses *LBPHFaceRecognizer_create()*, *EigenFaceRecognizer_create()* e *FisherFaceRecognizer_create()*. Neste trabalho, para a implementação do protótipo foi utilizado a classe *LBPHFaceRecognizer_create()* que implementa os algoritmos LBPH e possui a sua principal vantagem a sua insensibilidade a variações de iluminação, como já mencionado anteriormente e ilustrado na Figura 12, minimizando assim o tratamento de iluminação da imagem.

Com a finalidade de realizar o treinamento das faces é realizado a leitura da pasta onde está localizada as faces capturadas no processo anterior, para cada arquivo será extraído o nome, normalizando-o para obter os identificadores das faces. Utilizando o método *train()* da classe *LBPHFaceRecognizer_create()*, será realizado o treinamento das faces junto a os identificadores, este método tem como parâmetro obrigatório dois vetores, que devem possuir a mesma quantidade de posições, possuindo as matrizes das faces das imagens e os identificadores de cada indivíduo. De acordo com OpenCV (2018, tradução nossa), para que somente as informações necessárias sejam carregadas de forma encapsulada é necessário utilizar um arquivo criada pela classe *FileStorage* através da chamada do método *save()*, passando como parâmetro o caminho e nome do arquivo, disponível na classe *LBPHFaceRecognizer_create()*. Na Figura 23, temos a representação da classe responsável por realizar o treinamento das características das faces.

Figura 23 - Classe Responsável por Realizar o Treinamento de Imagens e Identificadores

```

1 import os
2
3 import cv2 as cv
4 import numpy as np
5
6
7 class Treinador:
8     lbph = cv.face.LBPHFaceRecognizer_create()
9
10    def __init__(self):
11        pass
12
13    def getIdentificadoresIds(self):
14        caminhos = [os.path.join('fotos', f) for f in os.listdir('fotos')]
15
16        identificadores = []
17        faces = []
18
19        for caminhoImagem in caminhos:
20            imagemFace = cv.cvtColor(cv.imread(caminhoImagem), cv.COLOR_BGR2GRAY)
21            id = int(os.path.split(caminhoImagem)[-1].split('.')[1])
22            identificadores.append(id)
23            faces.append(imagemFace)
24
25        return np.array(identificadores), faces
26
27    def treinar(self):
28        identificadores, faces = self.getIdentificadoresIds()
29        self.lbph.train(faces, identificadores)
30        self.lbph.save('resultadotreinamento/classificadorLBPH.yml')
31        return True

```

Fonte: Do Autor.

Antes de tudo, para realizar o reconhecimento facial obtendo a identificação do indivíduo, ocorrerá após a detecção facial. Deve-se carregar o arquivo gerado no treinamento, *classificadorLBPH.yml*, em memória. Em seguida (figura 24), o reconhecimento será realizado através da classe *LBPHFaceRecognizer_create()* que possui a função *predict()*, que dado uma matriz de uma face irá retornar o identificador do indivíduo e um número de confiança determinando a distância entre o identificador que foi encontrado e o valor zero, quanto mais próximo ao zero, maior será o fator de confiança da predição. Levando em consideração que a predição da face pode retornar um identificador e um valor de confiança com o valor elevado, será considerado que fatores de confiança superiores a 60, portanto possuirão um fator de confiança menor, resultarão em uma predição com resultado não conhecido, informando ao usuário que não foi possível reconhecer o a face.

Figura 24 - Verificando a Identidade de um Indivíduo

```

53     for (x, y, a, l) in facesDetectadas:
54         cv.rectangle(frame, (x, y), (x + l, y + a), (0, 0, 255), 2)
55         frameFaceCinza = cv.resize(frameCinza[y: y + a, x: x + l], (self.largura, self.altura))
56
57         self.id, self.confianca = self.reconhecedor.predict(frameFaceCinza)
58

```

Fonte: Do Autor

A verificação da identidade tem uma pequena diferença no modo como é feita o reconhecimento da face. Porque para fazer a verificação também é considerado um identificador do sujeito que será classificado pelo sistema, neste protótipo utilizaremos o nome do sujeito para fazer a verificação, se o identificador inserido não coincidir com o resultado da classificação da face, então o protótipo retornará falsa a identificação.

5.4 ANÁLISE DOS RESULTADOS OBTIDOS

Para análise dos dados será utilizado a matriz de confusão, também conhecida como matriz de contenção, que é utilizada em muitos trabalhos acadêmicos (MINATTO, 2016; GUIMARÃES, 2015, tradução nossa). A matriz de confusão consiste em dois grupos de imagens, neste caso uma composta por imagens que foram treinadas pelo classificador (positivos), enquanto, o outro grupo não foi submetida ao treino (negativo). Esta divisão resultará em quatro classificações possíveis. O verdadeiro positivo (VP), quando uma imagem que pertence ao conjunto de imagens treinadas é reconhecida pelo classificador corretamente. O falso positivo (FP), quando uma imagem que é positiva não é reconhecida pelo classificador. O verdadeiro negativo (VN), quando uma imagem que não pertence ao conjunto de imagens é reconhecida como tal. E o falso negativo (FN), quando uma imagem que não foi submetida ao treinamento é reconhecida como um indivíduo.

Na matriz de confusão da tabela 1, as colunas representam a classe indicada pelo classificador, enquanto que, as linhas representam a classe real destas imagens. Enquanto que foi adicionado uma nova coluna informando o total de imagens analisadas pelo algoritmo dentro das imagens disponíveis, isso se deve a detecção facial que não detectou todas as faces presentes em cada grupo. São utilizados 2674 imagens que não foram utilizados no treinamento do algoritmo, mas que pertence à

os mesmos sujeitos do treinamento do classificador e acrescentando um grupo de 971 imagens que não pertencem a nenhum indivíduo que não foi submetido ao treinamento.

Tabela 1 - Tabela de confusão sobre o reconhecimento facial

	Classe Positiva	Classe Negativa	Total
Faces compatíveis	1297	406	1703
Faces incompatíveis	238	733	971

Fonte: Do Autor

Podemos verificar que do total de imagens conhecidas pelo classificador (faces compatíveis), foram identificadas corretamente 1297 e 406 foram classificadas incorretamente. Quanto as faces não conhecidas (faces incompatíveis), 238 imagens foram classificadas como sendo false positiva, ou seja, pertencente a algum sujeito já cadastrado, 733 destes foram classificados corretamente.

Através desta matriz da tabela 1 podemos extrair algumas métricas importantes, como a acurácia, sensibilidade, especificidade, eficiência, valor predito positivo e valor predito negativo. A sensibilidade (fórmula 11) é usada para estimar o quão bom é a detecção de amostras positivas, ou seja, reconhecendo que uma face pertence a uma determinada pessoa (Zhu1, 2010). A equação abaixo indica como chegar a este resultado, e obteve-se um resultado de 84,5%.

$$S = \frac{VP}{VP+FN} \quad (11)$$

Especificidade é caracterizada pela probabilidade de faces não conhecidas serem dadas como tal, em vez de serem classificadas como positivas. Este protótipo obteve uma taxa de 64,4%. Este resultado foi obtido através da equação 12 abaixo:

$$E = \frac{VN}{VN+FP} \quad (12)$$

A eficiência é uma média aritmética obtida entre a taxa de sensibilidade e especificidade. Quando um sistema de classificação é sensível a positivos existe a tendência de obter se muitos falsos positivos ou vice e versa. A sensibilidade indica a

relevância do resultado obtido. Foi utilizado a fórmula 13 abaixo e obteve-se a taxa de 74,4%.

$$EF = \frac{\text{sensibilidade} + \text{eficiência}}{2} \quad (13)$$

A acurácia (14) resultou em 75.9%, este valor é obtido através do valor da equação x que é proporcional ao valor VP e VN. Esta medida é muito sensível ao conjunto de dados, se o conjunto de dados for desbalanceado influenciará no resultado deste valor.

$$A = \frac{VP + VN}{VP + VN + FP + FN} \quad (14)$$

Como já mencionado anteriormente o banco de dados utilizado neste projeto possui quatro níveis de complexidade idealizados pelo mantedor desta base, e para obter uma melhor compreensão do desempenho nestes grupos foi empregado similarmente o uso da matriz de confusão. O grupo de imagens (A) possui imagens com fundo, iluminação e escala controlada. Representado na tabela 2 foi classificado pelo proprietário da base de fotos como o grupo com menor complexidade. Ao total foram obtidas um total de 936 imagens positivas e 269 imagens negativas.

Tabela 2 – Grupo A. Tabela de confusão sobre o reconhecimento facial

	Classe Positiva	Classe Negativa	Total
Faces compatíveis	910	26	936
Faces incompatíveis	193	76	269

Fonte: Do Autor

A sensibilidade do classificador neste conjunto de teste foi de 82,5% para identificação de verdadeiros positivos, a especificidade apresentou uma taxa de 74,5% para a capacidade de detectar falsos positivo, a acurácia foi de 81,8% enquanto a taxa de eficiência do protótipo foi de 78,5%

O próximo grupo (B) possui como principal variação as diferenças entre escalas e sombras no fundo da imagem. Possui cerca de 291 imagens positivas e 104 imagens negativas a tabela de confusão (tabela 3) apresenta os resultados obtidos.

Tabela 3 – Grupo B. Tabela de confusão sobre o reconhecimento facial

	Classe Positiva	Classe Negativa	Total
Faces compatíveis	245	46	291
Faces incompatíveis	86	18	104

Fonte: Do Autor

Houve diferenças entre os percentuais do grupo B para o grupo A, a especificidade (28,1%) e eficiência (51,1%) apresentaram diferenças consideráveis, apenas a sensibilidade apresentou uma taxa aproximada ao grupo A.

Fundamentalmente, o subsequente conjunto (C) caracterizado por alterações no fundo das imagem, escala e presença de sombras. Composto por 467 imagens representado na tabela de confusão (tabela 4).

Tabela 4 – Grupo C. Tabela de confusão sobre o reconhecimento facial

	Classe Positiva	Classe Negativa	Total
Faces compatíveis	386	81	467
Faces incompatíveis	100	35	135

Fonte: Do Autor

O percentual em sensibilidade resultou em 69,8%, especificidade em 40,7%, eficiência em 55,3% e acurácia em 54,1%. Apresentando a menor taxa de acurácia dentre os conjuntos analisados, segundo Delbiaggio et al. (2017, tradução nossa) após analisar o desempenho do algoritmo LBPH em diferentes ambientes, ele observou que o mesmo resultava em boas taxas de acurácia quando o ambiente da imagem era conhecido, diferente de quando o reconhecimento acontecia em ambientes completamente diferente ocasionando na queda da acurácia.

O grupo (D) formado por 108 imagens possuindo fundo neutro e presença de sombras. Categorizado como sendo o mais complexo entre os grupos devido as imagens possuírem expressões faciais, na matriz de confusão é caracterizada a classificação (tabela 5)

Tabela 5 – Grupo D. Tabela de confusão sobre o reconhecimento facial

	Classe Positiva	Classe Negativa	Total
Faces compatíveis	108	0	108
Faces incompatíveis	83	23	106

Fonte: Do Autor

Uma característica peculiar a tabela 5 é devido o FP ser zero e como visto na fórmula indicada acaba sendo dividido o próprio VN por ele mesmo, resultando na especificidade de 100% e a eficiência 78.3%. Os percentuais de sensibilidade (51,5%), acurácia (58,0%) se deram mais inferiores.

6 CONCLUSÃO

A implementação deste protótipo de software só foi possível com o entendimento dos requisitos do processo de reconhecimento facial. Desde a aquisição da imagem através da Webcam, a detecção facial, normalização das imagens, extração de características da face e finalmente realizando a classificação da face como conhecida ou não.

As maiores dificuldades encontradas durante o desenvolvimento deste protótipo foi a escolha do algoritmo para a extração de características da face, na área de visão computacional encontram-se diversas técnicas já conceituadas e novos métodos que estão sendo desenvolvidos, aprimorados e estudados. Primariamente, optou-se por três métodos para extração das características da face, sendo as *Eigenfaces* o *Fisherfaces* e o LBPH, contudo analisando as características destas técnicas optou-se somente pelo uso do algoritmo LBPH, optando pela vantagem da menor sensibilidade a iluminação que é uma característica muito forte nestes outros dois algoritmos. Outra dificuldade encontrada neste trabalho foi o tratamento de imagens que possuem uma qualidade baixa interferindo nos resultados do treinamento e nos testes para validação.

O protótipo desenvolvido apresentou resultados satisfatórios. De acordo com os resultados, este protótipo possui características para uso em ambientes controlados, obteve-se um melhor desempenho de acurácia (81,8%), sensibilidade (82,5%), especificidade (64,5%) e eficiência (74,4%), e dessa forma se aproximando de trabalhos semelhantes. Por outro lado, o reconhecedor apresentou redução de desempenho quando as imagens das faces possuem fundo variados reduzindo assim a acurácia (54,1%), sensibilidade (69,8%), especificidade (40,7%) e eficiência (55,3%).

Com base nos resultados obtidos neste trabalho podemos afirmar que os objetivos foram atingidos. Visando a continuidade de trabalhos futuros, propõe-se as seguintes sugestões de trabalhos futuros:

- a) otimizar o protótipo desenvolvido neste trabalho implementando filtros e tratamentos de imagens para otimizar o reconhecimento facial;

- b) implementação de outro algoritmo para extração das características utilizando o paradigma baseado em modelo;
- c) implementar um protótipo de reconhecimento facial para uso em ambientes não controlados;
- d) implementação de um protótipo de reconhecimento facial utilizando Tensorflow da Google;

REFERÊNCIAS

AGARWAL, Mayank et al. Face recognition using principle component analysis, eigenface and neural network. In: Signal Acquisition and Processing, 2010. ICSAP'10. **International Conference on IEEE**, p. 310-314, 2010.

AHONEN, Timo; HADID, Abdenour; PIETIKAINEN, Matti. Face description with local binary patterns: Application to face recognition. **IEEE transactions on pattern analysis and machine intelligence**, v. 28, n. 12, p. 2037-2041, 2006.

AHONEN, Timo; HADID, Abdenour; PIETIKÄINEN, Matti. Face recognition with local binary patterns. In: **European conference on computer vision**. Springer, Berlin, Heidelberg, p. 469-481, 2004.

ASHBOURN, Julian. Biometrics in the New World. Cham: **Springer International Publishing**, 2014. Disponível em: <<http://link.springer.com/10.1007/978-3-319-04159-9>>. Acesso em: 24 out. 2017.

ASHBOURN, Julian. **Biometrics: Advanced Identity Verification**. London: Springer London, 2000. Disponível em: <<http://link.springer.com/10.1007/978-1-4471-0747-7>>. Acesso em: 05 mar. 2018.

AZEVEDO, Eduardo; CONCI, Aura. **Computação Gráfica: Teoria e Prática**. 5. ed. Elsevier: Rio de Janeiro, 2003. 353 p.

BELHUMEUR, P.N.; HESPANHA, J.P.; KRIEGMAN, D.J. Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 19, n. 7, p. 711–720, jul. 1997. Disponível em: <<http://ieeexplore.ieee.org/document/598228/>>. Acesso em: 15 out. 2017.

BERTILLON, Alphonse. **The Bertillon System of Identification**. London and New York: [s.n.], 1896.

BHELE, Sujata G.; MANKAR, V. H. A review paper on face recognition techniques. **International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)**, v. 1, n. 8, p. pp: 339-346, 2012.

BIOMETRIA. 2016. Disponível em: <<http://www.tse.jus.br/eleicoes/biometria-e-urna-eletronica/biometria-1>>. Acesso em: 13 out. 2013.

BRAHNAM, Sheryl et al. **Local binary patterns: new variants and applications**. Springer, 2016.

CARDOSO, J.-F. Infomax and maximum likelihood for blind source separation. **IEEE Signal Processing Letters**, v. 4, n. 4, p. 112–114, abr. 1997. Disponível em: <<http://ieeexplore.ieee.org/document/566704/>>. Acesso em: 17 out. 2017.

CAVALCANTI, Georgi Darmiton da Cunha. **Composição de biometria para sistemas multimodais de verificação de identidade pessoal**. 2005. 213 f. Universidade Federal de Pernambuco Centro de Informática, 2005. Disponível em: <http://repositorio.ufpe.br/bitstream/handle/123456789/2105/arquivo7173_1.pdf?sequence=1&isAllowed=y>. Acesso em: 22 mar. 2018.

COSTA, Luciano R.; OBELHEIRO, Rafael R.; FRAGA, Joni S. Introdução à biometria. **Livro texto dos Minicursos do VI Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais (SBSeg2006)**. SBC: Porto Alegre, v. 1, p. 103-151, 2006.

DAUGMAN, John. The importance of being random: statistical principles of iris recognition. **Pattern recognition**, v. 36, n. 2, p. 279-291, 2003.

DELAC, Kresimir. **Recent Advances in Face Recognition**. [S.l.]: InTech, 2008.

DELBIAGGIO, Nicolas et al. **A comparison of facial recognition's algorithms**. 2017.

EKENEL, H. K.; SANKUR, B. Feature selection in the independent component subspace for face recognition. **Pattern Recognition Letters**, v. 25, n. 12, p. 1377–1388, 2004.

ETEMAD, Kamran; CHELLAPPA, Rama. Discriminant analysis for recognition of human face images. **Journal of the Optical Society of America**, v. 14, n. 8, p. 1724, 1 Ago 1997. Disponível em: <<http://www.opticsinfobase.org/abstract.cfm?URI=josaa-14-8-1724%5Cnhttp://www.opticsinfobase.org/abstract.cfm?&id=1842>>. Acesso em: 15 set. 2017.

FACE Detection using Haar Cascades. 2018. Disponível em: <https://docs.opencv.org/3.4.1/d7/d8b/tutorial_py_face_detection.html>. Acesso em: 09 maio 2018.

GONG, Shaogang; MCKENNA, J Stephen; PSARROU, Alexandra. **Dynamic Vision: From Images to Face Recognition**. Singapore: Imperial College Press, 2000.

GONZALES, Rafael e WOODS, Richard. **Processamento Digitais de Imagens**. . [S.l: s.n.], 2010

GOODALL, Colin; JOLLIFFE, I T. Principal Component Analysis. **Technometrics**, v. 30, n. 3, p. 351, ago. 1988. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1002/0470013192.bsa501/full>>. Acesso em: 12 out. 2017.

GUIMARÃES, Rafael Miranda. **Desenvolvimento de um Protótipo de Software de Reconhecimento Facial de Tempo Real para Registro Eletrônico de Ponto em**

Ambientes Indoor com Utilização do Dispositivo Kinect. 2015. 196 F. Dissertação (Mestrado) - Curso de Sistemas de Informação e Gestão do Conhecimento, Universidade Fumec, Belo Horizonte, 2015. Cap. 7.

HOTELLING, H. Analysis of a complex of statistical variables into principal components. **Journal of Educational Psychology**, v. 24, n. 6, p. 417–441, 1933. Disponível em: <<http://content.apa.org/journals/edu/24/6/417>>. Acesso em: 17 out. 2017.

HYVÄRINEN, Aapo; OJA, Erkki. Independent Component Analysis: Algorithms and Applications. **Neural Networks**, v. 13, n. 45, p. 411–430, 2000.

JAIN, Anil K.; ROSS, Arun A.; NANDAKUMAR, Karthik. **Introduction to Biometrics.** Boston, MA: Springer US, 2011. v. 1. Disponível em: <<http://link.springer.com/10.1007/978-0-387-77326-1>>.

KIRBY, M.; SIROVICH, L. Application of the Karhunen-Loeve procedure for the characterization of human faces. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 12, n. 1, p. 103–108, 1990.

KOHAVI, Ron et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: **Ijcai**. 1995. p. 1137-1145.

KSHIRSAGAR, V.P.; BAVISKAR, M.R.; GAIKWAD, M.E. Face recognition using Eigenfaces. 2011, 3rd **International Conference on Computer Research and Development**, v. 2, p. 302–306, 2011.

LANGLOIS, Dominic; CHARTIER, Sylvain; GOSSELIN, Dominique. An introduction to independent component analysis: InfoMax and FastICA algorithms. **Tutorials in Quantitative Methods for Psychology**, v. 6, n. 1, p. 31-38, 2010.

MARTÍNEZ, Aleix M; KAK, Avinash C. Pca versus Ida. Pattern Analysis and Machine Intelligence, **IEEE Transactions**, v. 23, n. 2, p. 228–233, 2001.

MCLACHLAN, Geoffry. **Discriminant Analysis and Statistical Pattern Recognition.** [S.I.]: John Wiley & Sons, 2004.

MINATTO, Daniele Viera. **Análise de Características em Redes Neurais Artificiais Aplicadas em Visão Computacional com Ênfase em Reconhecimento Facial.** 2016. 97 f. TCC (Graduação) - Curso de Ciência da Computação, Universidade do Extremo Sul Catarinense, Criciúma, 2016. Cap. 6.

OJALA, Timo; PIETIKAINEN, Matti; MAENPAA, Topi. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **IEEE Transactions on pattern analysis and machine intelligence**, v. 24, n. 7, p. 971-987, 2002.

PANCERI, João Antonio Campos et al. RECONHECIMENTO FACIAL BASEADO EM HOG E PCA: UMA COMPARAÇÃO QUANTO À INVARIÂNCIA À

ILUMINAÇÃO. **Revista Ifes Ciência**, ISSN 2359-4799, v. 1, n. 1, 2015.

PRODOSSIMO, F. das C.; CHIDAMBARAM, C.; LOPES, H. S. Otimização Da Detecção De Olhos Em Imagens Faciais Utilizando Algoritmos Colônia De Abelhas Artificiais E Harmony Search. In: Anais do 10. Congresso Brasileiro de Intelig?ncia Computacional, **Anais...SBIC**, 16 mar. 2011. Disponível em: <http://abricom.org.br/eventos/cbic_2011/st_15_6>.

PULLI, K. et al. Real-time computer vision with OpenCV. **Communications of the ACM**, v. 55, n. 6, p. 61, 1 jun. 2012. Disponível em: <<http://dl.acm.org/citation.cfm?doid=2184319.2184337>>. Acesso em: 10 maio 2018.

PYQT5 Reference Guide. 2018. Disponível em: <<http://pyqt.sourceforge.net/Docs/PyQt5/index.html>>. Acesso em: 10 maio 2018.

RIVERBANK Computing Limited. 2018. Disponível em: <<https://riverbankcomputing.com/software/pyqt/intro>>. Acesso em: 09 maio 2018.

SAHOOLIZADEH, A Hossein; HEIDARI, B Zargham; DEHGHANI, C Hamid. A New Face Recognition Method using PCA , **LDA and Neural Network**. v. 2, n. 5, p. 7–12, 2008.

SHRESTHA, Niraj; VENKITASUBRAMANIAN, Aparna N.; MOENS, Marie-Francine. Key Event Detection in Video using ASR and Visual Data. In: **Proceedings of the Third Workshop on Vision and Language**. 2014. p. 46-53.

SIROVICH, L. e KIRBY, M. Low-dimensional procedure for the characterization of human faces. **Journal of the Optical Society of America**, v. 4, n. 3, p. 519, 1987. Disponível em: <<https://www.osapublishing.org/abstract.cfm?URI=josaa-4-3-519>>. Acesso em: 17 mar. 2018.

SPACEK, Libor. **Computer Vision Science Research Projects**. 2018. Disponível em: <<http://cswww.essex.ac.uk/mv/allfaces/index.html>>. Acesso em: 21 abr. 2018.

SUSHMA, Niket Borade; ADGAONKAR, Ramesh P. Comparative analysis of PCA and LDA. Jun 2011, [S.I.]: **IEEE**, p. 203–206. Disponível em: <<http://ieeexplore.ieee.org/document/5994243/>>. Acesso em: 20 abr. 2018

TAHMASEBI, P. e HEZARKHANI, A. e MORTAZAVI, M. Application of discriminant analysis for alteration separation; Sungun copper deposit, East Azerbaijan, Iran. **Australian Journal of Basic and Applied Sciences**, v. 4, n. 4, p. 564–576, 2010.

TAIGMAN, Yaniv. DeepFace: Closing the Gap to Human-Level Performance in Face Verification. Jun 2014, [S.I.]: **IEEE**, p. 1701–1708. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6909616>>. Acesso em: 20 abr. 2018.

TRAURING, Mitchell. Automatic Comparison of Finger-Ridge Patterns. **Nature**, v.

197, n. 4871, p. 938–940, 9 Mar 1963. Disponível em:
<<http://www.nature.com/doi/10.1038/197938a0>>. Acesso em: 14 maio 2018.

VIGLIAZZI, Douglas. **Biometria** - Medidas de Segurança. 2. ed. Rio de Janeiro: Visual Books, 2006.

WEBB, Andrew R; COPSEY, Keith D. **Statistical Pattern Recognition**. Chichester, UK: John Wiley & Sons, Ltd, 2011. Disponível em:
<<http://doi.wiley.com/10.1002/9781119952954>>. Acesso em: 22 maio 2018.

WISKOTT, L. Face recognition by elastic bunch graph matching. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 19, n. 7, p. 775–779, Jul 1997. Disponível em:
<<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=598235>>. Acesso em: 17 set. 2017.

YANG, Dai et al. Exploration of Karhunen-Loeve transform for multichannel audio coding. In: **Digital Cinema and Microdisplays**. International Society for Optics and Photonics, 2000. p. 89-101.

ZANETTE, Jackson Guizzo. **Reconhecimento Automático de Placas Veiculares Utilizando a Tecnologia Ocr e a Plataforma Raspberry Pi Aplicada na Fiscalização Eletrônica De Rodovias**. 2016. 136 F. TCC (Graduação) - Curso De Ciência Da Computação, Universidade Do Extremo Sul Catarinense, Criciúma, 2016. Cap. 7

APÊNDICE (S)

APÊNDICE A – ARTIGO CIENTÍFICO

Protótipo de Reconhecimento Facial com Ênfase em Extração das Características da Face

Tiago Aleff da Silva¹, Sérgio Coral²

¹Curso de Ciência da Computação – Universidade do Extremo Sul Catarinense (UNESC)
Caixa Postal 3167 – 88806-00 – Criciúma– RS – Brasil

²Esp. Professor do Curso de Ciência da Computação – Universidade do Extremo Sul Catarinense (UNESC) Caixa Postal 3167 – 88806-00 – Criciúma– RS – Brasil

tiagoaleff@gmail.com, sergiocoral@unesc.net

Abstract. *Aiming to continue started by Minatto (2016), the purpose approached in this research is the development of a software prototype capable of performing the identification and verification of a individual's identity. Facial detection is described by the program's capability of finding existing faces in an image, while facial recognition is defined by the verification or identification. The OpenCV Computational Vision Library was used to achieve the objectives of this research, making possible the use of optimized algorithms. The facial recognition must be made through an algorithm that extracts the characteristics of a face, and on this Project the Local Binary Patterns algorithm was chosen, whose main benefit, regarding the techniques based on appearance, is its smaller sensibility to ambient light. As a result, a prototype that attains facial detection and facial recognition, observing the requirements for facial requirements, was developed. This prototype was evaluated on an image basis, which best obtained result had an accuracy of 82,8%, and, as a worse result, 54,1% accuracy. It presented decay in its performance when some variations are present, as the ambient background or facial expressions, being then characterized by these results to usage in controlled environments.*

Resumo. *Com o intuito de dar continuidade ao trabalho iniciado por Minatto (2016) o propósito abordado nesta pesquisa é o desenvolvimento de um protótipo de software capaz de realizar a identificação e verificação da identidade de indivíduos. A detecção facial é caracterizada pela capacidade do programa em encontrar as faces quando existem em uma imagem, enquanto que, o reconhecimento facial se define pela verificação e ou identificação. A biblioteca de Visão Computacional OpenCV foi utilizada para atingir os objetivos desta pesquisa possibilitando a utilização de algoritmos otimizados. O reconhecimento facial deve ser feito através de um algoritmo de extração de característica da face e neste projeto foi escolhido o algoritmo Local Binary Patterns cujo a principal vantagem em relação as técnicas baseadas em aparência é sua menor sensibilidade a iluminação do ambiente. Como resultado foi desenvolvido um protótipo que realiza o a detecção facial e reconhecimento facial observando os requisitos do reconhecimento facial. Este protótipo foi validado em uma base de imagens obtendo o melhor resultado uma acurácia de 81,8% e em seu pior resultado com uma acurácia de 54,1%. Apresentando queda no seu desempenho quando existem algumas variações, como o*

fundo do ambiente e presença de expressões faciais, caracterizando por meios destes resultados a sua utilização para ambientes controlados.

1. Introdução

As implicações sociais e culturais das tecnologias de reconhecimento facial são de grande alcance, mas a atual lacuna de desempenho nessa questão entre máquinas e o sistema visual humano serve de amortecedor ao ter que lidar com essas implicações. (TAIGMAN, 2014, tradução nossa).

Para Cavalcanti (2005), é possível dividir as técnicas biométricas em físicas e comportamentais. Dentre os métodos físicos mais estudadas podemos citar: face, impressão digital, retina, íris, formato da orelha, impressão da mão e DNA. E os métodos comportamentais são assinaturas, voz e dinâmica de digitação, são alguns métodos disponíveis.

Os sistemas biométricos podem realizar a identificação e verificação. A identificação ocorre quando os dados obtidos a partir de um sujeito externo ao sistema são comparados com todos os dados já armazenado no sistema. Enquanto, a verificação ocorre quando um indivíduo fornece algum identificador de algum registro que existe no banco de dados e seus dados são comparados especificamente com aquele registro. (JAIN; ROSS; NANDAKUMAR, 2011, tradução nossa). Os sistemas de reconhecimento facial podem ser classificados em duas categorias. Primeiramente, há os sistemas que podem ser agrupados por cenário controlado, onde o sujeito a ser testado está localizado em uma posição conhecida com o mínimo de interferências possíveis. O segundo grupo, são classificados por ambientes de teste não controlado, porque podem haver muitas variáveis no ambiente a serem consideradas. (CAVALCANTI, 2005). Tendo em vista estes conceitos, este trabalho irá considerar somente os ambientes de testes controlados.

De acordo com Face... (2016, tradução nossa), para implementar o algoritmo pela extração de características da face existem algumas abordagens predominantes para a resolução deste problema, sendo por exemplo a abordagem geométrica e a fotométrica. Como há interesse por estas áreas de pesquisa, vários algoritmos de extração de características são desenvolvidos para que o ocorra a validação. Alguns dos algoritmos estudados são a Análise de Componentes Principais (PCA), Análise discriminante Linear (LDA) e Combinação de Grafos em Grupos Elásticos (EBGM).

Em seu Trabalho de Conclusão de Curso, Minatto (2016) propôs um protótipo para reconhecimento facial utilizando o algoritmo Viola-Jones e Rede Neural implantado no ambiente Raspberry Pi. Entretanto, nesse protótipo não é desenvolvido o identificador e nem o verificador que classifica a identidade do sujeito como conhecida ou não. Agora, este novo trabalho vem apresentar uma continuidade à os estudos iniciados por Minatto (2016) implementando, além dos princípios de detecção facial, as técnicas necessárias para que ocorra a verificação e identificação do indivíduo.

Deste modo, é proposto como tema de estudo deste trabalho, a aplicação das técnicas necessárias para implementar o identificador possibilitando a identificação e verificação da identidade de indivíduos, por meio do algoritmo de extração de características Padrão Binários Locais (LBP) do inglês, Local Binary Patterns.

2. Desenvolvimento

O projeto consiste em um protótipo que fará o reconhecimento facial através da captura de imagens de uma Webcam. O protótipo de reconhecimento facial terá como foco ambientes

controlados, onde a luz, posição do indivíduo serão conhecidos. Dessa forma é minimizado as ocorrências de falsos positivos para a detecção e reconhecimento facial. O levantamento bibliográfico deste projeto permitiu a aplicação dos conceitos usados nesta sessão, bem como o desenvolvimento deste protótipo.

2.1 Componentes do Protótipo

Através da pesquisa realizada foi possível uma compreensão mais abrangente dos conhecimentos necessários para o desenvolvimento deste protótipo, possibilitando o uso de ferramentas que auxiliassem no desenvolvimento. O protótipo foi desenvolvido em cima do Sistema Operacional (SO) Windows, possibilitando o uso de diversos componentes que auxiliam no desenvolvimento do mesmo. Nas próximas sessões são abordados os componentes utilizados para o desenvolvimento.

Apesar do OpenCv possuir uma implementação de interface básica, neste trabalho buscou-se a utilização de um framework de interface que facilitasse a implementação do protótipo, sendo a escolha o PyQt5. Esta permite que sejam utilizados vários componentes desde interfaces, Threads e Interfaces de conexão com o Bluetooth e internet. (PYQT5... 2018).

O PyQt5 é uma biblioteca que permite utilizar as funções do framework QT através da linguagem de desenvolvimento Python. O framework Qt é desenvolvido pela The Qt Company's que o disponibiliza sobre duas licenças disponíveis, tal como a versão Open Source (GPL) e a sua versão comercial. Na versão comercial há um custo pela sua aquisição e uma versão para teste, diferentemente da licença GPL que necessita que o código fonte seja disponibilizado em <<https://www.qt.io/download>>. (RIVERBANK... 2018).

O OpenCV é uma biblioteca Open Source para Visão Computacional licenciado para uso acadêmico e comercial, os produtos gerados através dessa biblioteca não precisam ter o código de fonte aberto. (OPENCV, 2018). Escrita em C e C++ está disponível nas plataformas de Sistemas Operacionais Linux, Windows, Mac OS X e também nas linguagens de programações Python, Ruby, Matlab, entre outras. (SHRESTHA; VENKITASUBRAMANIAN; MOENS, 2014, tradução nossa).

O OpenCv é destinado a resolver problemas de Visão Computacional, ela possui disponíveis algoritmos de baixo e alto nível para a detecção facial, detecção de pedestres, busca de características, e rastreamento de objetos. A biblioteca também conta com aceleração em GPU sem que seja necessário programar especificamente para o uso este recurso. (PULLI et al., 2012). A biblioteca possui algoritmos específicos para a detecção facial e para o reconhecimento facial. Sendo que as técnicas Eigenfaces, FisherFaces e LBPH estão disponíveis no framework. (YOSHINAO, 2012, tradução nossa). Neste protótipo, serão utilizados os algoritmos referentes à os métodos LBPH e detecção facial.

2.2 Base de Dados

Para identificar a eficiência do protótipo foi utilizado uma base de teste de faces disponibilizados gratuitamente pelo Spacek (2008), cujo objetivo da base é ser utilizado em pesquisas de Visão Computacional, disponibilizando o total de 7900 imagens agrupadas por quatro nível de complexidade em relação a iluminação, ambiente, expressões faciais e posição da face. Não há variação no estilo do cabelo, pois as fotos foram tiradas em uma única sessão. Ambos os grupos possuem imagens de homens, mulheres, raças diferentes e um total de 20 imagens por sujeitos. As imagens foram agrupadas em quatro grupos de complexidade, desde

imagens com ambiente controlado a alterações de na posição do indivíduo na imagem, variações de fundo e expressões faciais.

Será considerado para treino do nosso classificador o total de 1/3 por indivíduo e os outros 2/3 serão utilizados testes de identidade e os dados serão coletados análise dos dados. (KOHAVI et al., 2008, tradução nossa).

2.3 Desenvolvimento do Protótipo de Software

O projeto desenvolvido pode ser expressado pelo seguinte fluxograma, figura 1. Nas próximas sessões serão abordadas as telas desenvolvidas nesta pesquisa.

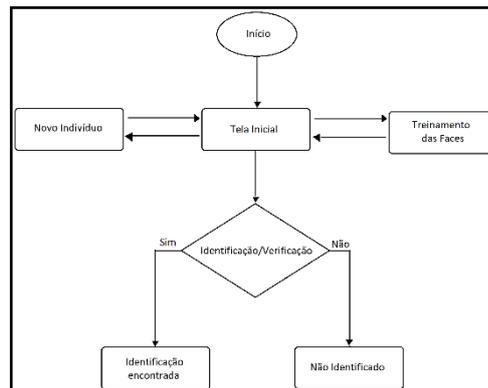


Figure 1. Fluxograma do Protótipo

2.3.1 Tela Inicial

A tela inicial é a principal tela do protótipo, onde podemos acessar todos os recursos disponíveis, assim como treinamento, reconhecimento facial e acessar a funcionalidade para adicionar novos sujeitos a base de face. (Figura 2)



Figure 2. Tela Inicial do Protótipo

2.3.2. Adicionar novo indivíduo

Nessa tela é possível adicionar novas faces capturando até vinte (20) imagens, este valor foi definido por convenção. E após isso podemos realizar o treinamento destas novas faces utilizando o item do menu configurações e selecionando o item treinamento.

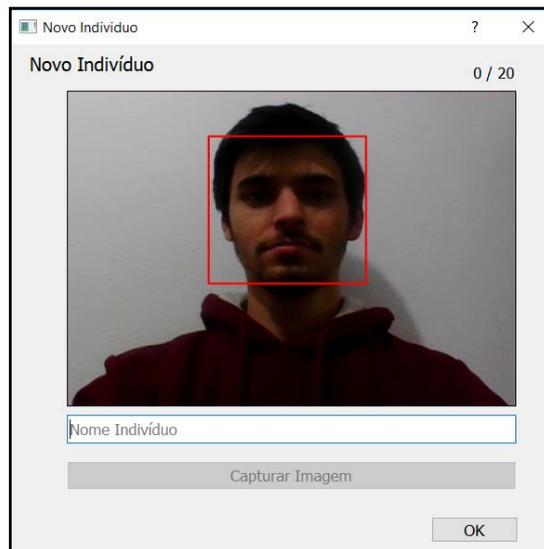


Figure 3. Diálogo para Adicionar Novo Indivíduo

2.3.4 Entrada de Imagens

As imagens que são capturadas pelo protótipo são obtidas em tempo real pela Webcam. Elas devem passar por uma normalização antes de serem adicionadas ao banco de faces ou para serem utilizadas no reconhecimento da identidade do indivíduo. Dessa forma, se utiliza de alguns dos recursos disponibilizados pelo framework OpenCV.

Tanto para a detecção quanto para o reconhecimento facial, a imagem deve ser transformada para uma escala de cinza e redimensionada para um tamanho padrão reduzindo o processamento computacional posteriormente. A conversão para a escala de cinza ocorrerá a partir da chamada do método `cvtColor()` que irá receber como parâmetro a própria imagem e para que espectro de cor será transformada, neste caso, será utilizado o parâmetro `CV_BGR2GRAY` que aplicará o filtro cinza, alterando a cor da imagem.

2.3.5 Detecção Facial

A detecção facial é uma das etapas essenciais para o reconhecimento facial, que a partir de uma imagem o protótipo deve ser capaz de encontrar as faces existentes, se o protótipo não consegue encontrar as faces, então não será possível realizar o processo de reconhecimento facial.

A biblioteca do OpenCV possui uma série de rotinas disponíveis para facilitar o processo de detecção facial, entre estes processos, ele implementa o método de classificação Haar, possibilitando a detecção de faces ou realizando o treinamento de novos objetos por meio de Aprendizado de Máquina (FACE... 2018, tradução nossa). Neste projeto foi utilizado o método de classificação disponível através dos arquivos `haarcascade_frontalface_default.xml` e `haarcascade_eye.xml`. Com estes arquivos podemos fazer a detecção tanto de faces como de olhos. Estes arquivos são resultados do treinamento da detecção facial, implementado através do método Viola-Jones.

Para carregar os arquivos haar é necessário utilizar a classe `CascadeClassifier()`, responsável pela detecção de objetos e ou treinamento de novos objetos, que em seu construtor receberá o caminho relativo ou absoluto do arquivo e o frame que será verificado a existência da face. Assim, será possível fazer a predição de faces ou de olhos na imagem conforme o arquivo enviado no seu parâmetro, chamando o método `detectMultiScale()` da classe

CascadeClassifier(). Para que o método funcione corretamente será necessário que seja aplicado o filtro de cor cinza sobre a face que será passada por parâmetro. (OPENCV, 2018, tradução nossa).

2.3.6 Adição Novas Identidades e Reconhecimento Facial

Para adicionar uma nova identidade no protótipo é fundamental que o processo de detecção facial esteja em execução funcionando adequadamente, ou seja, capturando as faces existentes. Após a captura da face, o frame que contém a face será dividido em dois, uma será processada e renderizada na interface do protótipo para que o usuário veja a detecção e a outra será redimensionada através do método `resize()` para 220 pixels de altura e 220 pixels de largura e em seguida, após ser normalizada, a imagem será salva em um diretório do Sistema Operacional.

Posteriormente, as imagens das faces salvas serão utilizadas para o treinamento das características da face. Para isso, é necessário a captura de até 20 imagens de um mesmo indivíduo, onde cada imagem capturada será nomeada com um identificador, que referencia uma identidade de uma pessoa, mais o seu enésimo valor dentro do total de faces capturadas de um mesmo indivíduo.

O treinamento destas imagens para realizar o reconhecimento facial será realizado por meio da biblioteca OpenCV que disponibiliza os recursos necessários para esta tarefa. Os recursos relacionados a reconhecimento facial são implementados pela classe `FaceRecognizer` e de suas subclasses `LBPHFaceRecognizer_create()`, `EigenFaceRecognizer_create()` e `FisherFaceRecognizer_create()`. Neste trabalho, para a implementação do protótipo foi utilizado a classe `LBPHFaceRecognizer_create()` que implementa os algoritmos LBPH e possui a sua principal vantagem a sua insensibilidade a variações de iluminação, minimizando assim o tratamento de iluminação da imagem.

Com a finalidade de realizar o treinamento das faces é realizado a leitura da pasta onde está localizada as faces capturadas no processo anterior, para cada arquivo será extraído o nome, normalizando-o para obter os identificadores das faces. Utilizando o método `train()` da classe `LBPHFaceRecognizer_create()`, será realizado o treinamento das faces junto a os identificadores, este método tem como parâmetro obrigatório dois vetores, que devem possuir a mesma quantidade de posições, possuindo as matrizes das faces das imagens e os identificadores de cada indivíduo. De acordo com OpenCV (2018, tradução nossa), para que somente as informações necessárias sejam carregadas de forma encapsulada é necessário utilizar um arquivo criada pela classe `FileStorage` através da chamada do método `save()`, passando como parâmetro o caminho e nome do arquivo, disponível na classe `LBPHFaceRecognizer_create()`.

Antes de tudo, para realizar o reconhecimento facial obtendo a identificação do indivíduo, ocorrerá após a detecção facial. Deve-se carregar o arquivo gerado no treinamento, `classificadorLBPH.yml`, em memória. Em seguida, o reconhecimento será realizado através da classe `LBPHFaceRecognizer_create()` que possui a função `predict()`, que dado uma matriz de uma face irá retornar o identificador do indivíduo e um número de confiança determinando a distância entre o identificador que foi encontrado e o valor zero, quanto mais próximo ao zero, maior será o fator de confiança da predição. Levando em consideração que a predição da face pode retornar um identificador e um valor de confiança com o valor elevado, será considerado que fatores de confiança superiores a 60, portanto possuirão um fator de confiança menor,

resultarão em uma predição com resultado não conhecido, informando ao usuário que não foi possível reconhecer o a face.

A verificação da identidade tem uma pequena diferença no modo como é feita o reconhecimento da face. Porque para fazer a verificação também é considerado um identificador do sujeito que será classificado pelo sistema, neste protótipo utilizaremos o nome do sujeito para fazer a verificação, se o identificador inserido não coincidir com o resultado da classificação da face, então o protótipo retornará falsa a identificação.

2.4 Análise dos Resultados Obtidos

Para análise dos dados será utilizado a matriz de confusão, também conhecida como matriz de contenção, que é utilizada em muitos trabalhos acadêmicos (MINATTO, 2016; GUIMARÃES, 2015, tradução nossa). A matriz de confusão consiste em dois grupos de imagens, neste caso uma composta por imagens que foram treinadas pelo classificador (positivos), enquanto, o outro grupo não foi submetida ao treino (negativo). Esta divisão resultará em quatro classificações possíveis. O verdadeiro positivo (VP), quando uma imagem que pertence ao conjunto de imagens treinadas é reconhecida pelo classificador corretamente. O falso positivo (FP), quando uma imagem que é positiva não é reconhecida pelo classificador. O verdadeiro negativo (VN), quando uma imagem que não pertence ao conjunto de imagens é reconhecida como tal. E o falso negativo (FN), quando uma imagem que não foi submetida ao treinamento é reconhecida como um indivíduo.

Na matriz de confusão, as colunas representam a classe indicada pelo classificador, enquanto que, as linhas representam a classe real destas imagens. Enquanto que foi adicionado uma coluna adicional informando o total de imagens analisadas pelo algoritmo dentro das imagens disponíveis, isso se deve a detecção facial que não detectou todas as faces presentes em cada grupo.

Como já mencionado anteriormente o banco de dados utilizado neste projeto possui quatro níveis de complexidade idealizados pelo mantedor desta base, e para obter uma melhor compreensão do desempenho nestes grupos foi empregado o uso da matriz de confusão. O grupo de imagens (A) possui imagens com fundo, iluminação e escala controlada. Representado na tabela 1 foi classificado pelo proprietário da base de fotos como o grupo com menor complexidade. Ao total foram obtidas um total de 936 imagens positivas e 269 imagens negativas.

Tabela 1 – Grupo A. Tabela de confusão sobre o reconhecimento facial

	Classe Positiva	Classe Negativa	Total
Faces compatíveis	910	26	936
Faces incompatíveis	193	76	269

A sensibilidade do classificador neste conjunto de teste foi de 82,5% para identificação de verdadeiros positivos, a especificidade apresentou uma taxa de 74,5% para a capacidade de detectar falsos positivo, a acurácia foi de 81,8% enquanto a taxa de eficiência do protótipo foi de 78,5%

O próximo grupo (B) possui como principal variação as diferenças entre escalas e sombras no fundo da imagem. Possui cerca de 291 imagens positivas e 104 imagens negativas a tabela de confusão (tabela 2) apresenta os resultados obtidos.

Tabela 2 – Grupo B. Tabela de confusão sobre o reconhecimento facial

	Classe Positiva	Classe Negativa	Total
Faces compatíveis	245	46	291
Faces incompatíveis	86	18	104

Houve diferenças entre os percentuais do grupo B para o grupo A, a especificidade (28,1%) e eficiência (51,1%) apresentaram diferenças consideráveis, apenas a sensibilidade apresentou uma taxa aproximada ao grupo A.

Fundamentalmente, o subsequente conjunto (C) caracterizado por alterações no fundo das imagem, escala e presença de sombras. Composto por 467 imagens representado na tabela de confusão (tabela 3).

Tabela 3 – Grupo C. Tabela de confusão sobre o reconhecimento facial

	Classe Positiva	Classe Negativa	Total
Faces compatíveis	386	81	467
Faces incompatíveis	100	35	135

O percentual em sensibilidade resultou em 69,8%, especificidade em 40,7%, eficiência em 55,3% e acurácia em 54,1%. Apresentando a menor taxa de acurácia dentre os conjuntos analisados, segundo Delbiaggio et al. (2017, tradução nossa) após analisar o desempenho do algoritmo LBPH em diferentes ambientes, ele observou que o mesmo resultava em boas taxas de acurácia quando o ambiente da imagem era conhecido, diferente de quando o reconhecimento acontecia em ambientes completamente diferente ocasionando na queda da acurácia.

O grupo (D) formado por 108 imagens possuindo fundo neutro e presença de sombras. Categorizado como sendo o mais complexo entre os grupos devido as imagens possuírem expressões faciais, na matriz de confusão é caracterizada a classificação (tabela 4).

Tabela 4 – Grupo D. Tabela de confusão sobre o reconhecimento facial

	Classe Positiva	Classe Negativa	Total
Faces compatíveis	108	0	108
Faces incompatíveis	83	23	106

Um resultado notório, a tabela 4 possui o valor do Falso Positivo zero e segundo a equação da especificidade resulta em um valor de 100% e eficiência de 78,3%. Os percentuais de sensibilidade (51,5%), acurácia (58,0%) se deram mais inferiores.

3. Conclusão

A implementação deste protótipo de software só foi possível com o entendimento dos requisitos do processo de reconhecimento facial. Desde a aquisição da imagem através da Webcam, a detecção facial, normalização das imagens, extração de características da face e finalmente realizando a classificação da face como conhecida ou não.

As maiores dificuldades encontradas durante o desenvolvimento deste protótipo foi a escolha do algoritmo para a extração de características da face, na área de visão computacional encontram-se diversas técnicas já conceituadas e novos métodos que estão sendo

desenvolvidos, aprimorados e estudados. Primariamente, optou-se por três métodos para extração das características da face, sendo as Eigenfaces o Fisherfaces e o LBPH, contudo analisando as características destas técnicas optou-se somente pelo uso do algoritmo LBPH, optando pela vantagem da menor sensibilidade a iluminação que é uma característica muito forte nestes outros dois algoritmos. Outra dificuldade encontrada neste trabalho foi o tratamento de imagens que possuem uma qualidade baixa interferindo nos resultados do treinamento e nos testes para validação.

O protótipo desenvolvido apresentou resultados satisfatórios. De acordo com os resultados, este protótipo possui características para uso em ambientes controlados, obteve-se um melhor desempenho de acurácia (81,8%), sensibilidade (82,5%), especificidade (64,5%) e eficiência (74,4%), e dessa forma se aproximando de trabalhos semelhantes. Por outro lado, o reconhecedor apresentou redução de desempenho quando as imagens das faces possuem fundo variados reduzindo assim a acurácia (54,1%), sensibilidade (69,8%), especificidade (40,7%) e eficiência (55,3%).

Com base nos resultados obtidos neste trabalho podemos afirmar que os objetivos foram atingidos. Visando a continuidade de trabalhos futuros, propõe-se as seguintes sugestões de trabalhos futuros:

Referências

- Biometria. 2016. Disponível em: <<http://www.tse.jus.br/eleicoes/biometria-e-urna-eletronica/biometria-1>>. Acesso em: 13 out. 2013.
- Cavalcante, Georgi Darmiton da Cunha. Composição de biometria para sistemas multimodais de verificação de identidade pessoal. 2005. 213 f. Universidade Federal de Pernambuco Centro de Informática, 2005. Disponível em: <http://repositorio.ufpe.br/bitstream/handle/123456789/2105/arquivo7173_1.pdf?sequence=1&isAllowed=y>. Acesso em: 22 mar. 2018.
- Face Detection using Haar Cascades. 2018. Disponível em: <https://docs.opencv.org/3.4.1/d7/d8b/tutorial_py_face_detection.html>. Acesso em: 09 maio 2018.
- Kohavi, Ron et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Ijcai. 1995. p. 1137-1145.
- Minatto, Daniele Viera. Análise de Características em Redes Neurais Artificiais Aplicadas em Visão Computacional com Ênfase em Reconhecimento Facial. 2016. 97 f. TCC (Graduação) - Curso de Ciência da Computação, Universidade do Extremo Sul Catarinense, Criciúma, 2016. Cap. 6.
- Pulli, K. et al. Real-time computer vision with OpenCV. Communications of the ACM, v. 55, n. 6, p. 61, 1 jun. 2012. Disponível em: <<http://dl.acm.org/citation.cfm?doid=2184319.2184337>>. Acesso em: 10 maio 2018.
- PYQT5 Reference Guide. 2018. Disponível em: <<http://pyqt.sourceforge.net/Docs/PyQt5/index.html>>. Acesso em: 10 maio 2018.
- Riverbank Computing Limited. 2018. Disponível em: <<https://riverbankcomputing.com/software/pyqt/intro>>. Acesso em: 09 maio 2018.

Shretha, Niraj; Venkitasubramanian, Aparna N.; Moens, Marie-Francine. Key Event Detection in Video using ASR and Visual Data. In: Proceedings of the Third Workshop on Vision and Language. 2014. p. 46-53.

Spacek, Libor. Computer Vision Science Research Projects. 2018. Disponível em: <<http://cswww.essex.ac.uk/mv/allfaces/index.html>>. Acesso em: 21 abr. 2018.