

UNIVERSIDADE FEDERAL DO MARANHÃO
CENTRO DE CIÊNCIAS EXATAS E TECNOLÓGICAS
CURSO DE CIÊNCIA DA COMPUTAÇÃO

LUIZ CARLOS SOARES NETO

**ANÁLISE DE ALGORITMOS DE BIOMETRIA FACIAL PARA CONTROLE DE
PESSOAS EM ESPAÇOS PRIVATIVOS**

São Luís/MA

2019

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Núcleo Integrado de Bibliotecas/UFMA

Soares Neto, Luiz Carlos.

ANÁLISE DE ALGORITMOS DE BIOMETRIA FACIAL PARA CONTROLE
DE PESSOAS EM ESPAÇOS PRIVATIVOS / Luiz Carlos Soares
Neto. - 2019.

60 f.

Orientador(a): Tiago Bonini Borchartt.

Monografia (Graduação) - Curso de Ciência da
Computação, Universidade Federal do Maranhão, SÃO LUÍS,
2019.

1. Análise de algoritmos. 2. Aprendizagem Profunda.
3. Biometria Facial. 4. Detecção Facial. I. Bonini
Borchartt, Tiago. II. Título.

LUIZ CARLOS SOARES NETO

**ANÁLISE DE ALGORITMOS DE BIOMETRIA FACIAL PARA CONTROLE DE
PESSOAS EM ESPAÇOS PRIVATIVOS**

Monografia apresentada ao curso de Ciência da Computação da Universidade Federal do Maranhão, como parte dos requisitos necessários para obtenção do grau de Bacharel em Ciência da Computação. Orientador: Prof. Dr. Tiago Bonini Borchardt

São Luís/MA

2019

LUIZ CARLOS SOARES NETO

**ANÁLISE DE ALGORITMOS DE BIOMETRIA FACIAL PARA CONTROLE DE
PESSOAS EM ESPAÇOS PRIVATIVOS**

Monografia apresentada ao curso de Ciência da Computação da Universidade Federal do Maranhão, como parte dos requisitos necessários para obtenção do grau de Bacharel em Ciência da Computação. Orientador: Prof. Dr. Tiago Bonini Borchardt

Aprovada em ___/___/_____

Prof. Dr. Tiago Bonini Borchardt
(Orientador)
Universidade Federal do Maranhão

Profa. Dra. Simara Vieira da Rocha
Examinador 1

Prof. Dr. Geraldo Braz Junior
Examinador 2

São Luís/MA

2019

Agradecimentos

Agradeço a Deus que me possibilitou vivenciar toda experiência maravilhosa e emocionante de fazer algo que amo, de me proporcionar momentos de alegria, tristeza e principalmente conquistas nestes anos que pude desfrutar dentro do curso. Por colocar em minha vida as pessoas certas para me motivar nos momentos que me senti para baixo, agradeço por sempre me guiar e cuidar de mim.

A Universidade Federal do Maranhão (UFMA), que me proporcionou o desenvolvimento necessário para este momento com profissionais dispostos a ensinar o caminho para o bom profissional da área de ciência da computação, ao trabalhar no departamento de pesquisa da UFMA como bolsista na área de programação, e proporcionar o início da minha carreira, na prática, do desenvolvimento de tecnologia.

Ao meu orientador Prof. Dr. Tiago Bonini Borchardt pela oportunidade de desenvolver este projeto que para mim, foi desafiador, por acreditar em mim e me apoiar, direcionar e orientar, por sempre estar disposto a tirar dúvidas e inspirar a me desenvolver cada vez mais em minha vida profissional.

Agradeço imensamente à minha família, minha mãe Sandra Maria Nunes Soares e ao meu Pai Robert Wagner Costa Soares e a minha irmã Louise Nunes Soares, que me deram forças nos momentos difíceis, que compreenderam quando eu tinha que estar ausente em encontros familiares, que acreditam na realização dos meus sonhos.

A minha namorada e quase noiva Julia Manayra, que nunca hesitou em acreditar em mim mesmo quando eu não acreditava, por me motivar e estar ao meu lado nestes últimos 6 anos, por sempre compreender quando precisei me ausentar no desenvolvimento deste projeto me incentivando sempre a crescer como um profissional de qualidade, agradeço por acreditar nos meus sonhos e me dar forças para correr atrás deles.

Aos meus chefes Seu Francisco e Dr César Castro, que acreditaram e me apoiaram nos dias que precisei desenvolver este projeto, sou grato pela compreensão e incentivo para o meu crescimento profissional e educacional. Aos meus colegas de trabalho Alex Santos, Renato Kerly e Amaury pelo incentivo a realização de mais essa etapa na minha vida.

"Às vezes, quando você inova, você comete erros. É melhor admiti-los rapidamente, e seguir em frente para melhorar suas outras inovações."

(Steve Jobs)

Resumo

Identificar pessoas, seja através do reconhecimento facial ou biométrico, está associado aos serviços que estão começando a ser implantados em vários lugares do mundo, seja em local público, uma grande ou pequena empresa, ou até mesmo em simples condomínio, para ter um maior controle e organização no ambiente atuante. No entanto, o processo para realizar esse tipo de tarefa, requer um envolvimento de novas tecnologias, pessoas qualificadas e investimento financeiro adequado. Com relação aos custos, pode-se ter uma variação de acordo com o objetivo da aplicação e seu local de atuação, desde uma pequena padaria, a um grande estádio de futebol, por exemplo. O principal objetivo deste trabalho é fazer uma análise de diferentes tipos de algoritmos de detecção e reconhecimento facial para a atividade de Biometria Facial, com o objetivo de mostrar a interação e as diferenças desses algoritmos. Utilizou-se alguns algoritmos de reconhecimento facial para comparações e suas vantagens, associam-se os resultados comparativos com outros métodos de reconhecimento facial, como por exemplo, o Haar Cascade com KNN ou até mesmo utilizando SVM, ou árvore de decisão etc, além disso, tem-se disponível, diversos materiais sobre Python e suas bibliotecas de reconhecimento associado ao OpenCV e a biblioteca de reconhecimento facial Dlib e Face Recognition, que possui uma precisão mais acentuada e voltada para o mercado, e que será um dos componentes inseridos nesse trabalho. Para geração dos resultados foram utilizadas imagens gravadas pelo autor e bases públicas disponíveis na internet.

Palavras-chave. Biometria Facial; Aprendizagem Profunda; Redes Neurais; Reconhecimento Facial; Máquina de Vetores de Suporte; Histograma de Gradientes Orientados.

Abstract

Identifying people, whether through facial or biometric recognition, is associated with services that are beginning to be deployed around the world, whether in a public place, a large or small business, or even a simple condominium, for greater control and organization in the active environment. However, the process to accomplish this type of task requires the involvement of new technologies, skilled people and adequate financial investment. Regarding costs, you can have a variation according to the purpose of the application and its place of operation, from a small bakery, to a large football stadium, for example. The main objective of this work is to analyze different types of facial detection and recognition algorithms for Facial Biometrics activity, aiming to show the interaction and differences of these algorithms. Some facial recognition algorithms have been used for comparisons and their advantages, comparative results are combined with other facial recognition methods, such as Haar Cascade with KNN or even using SVM, or decision tree, etc. Python and its associated OpenCV recognition libraries and the Dlib and Face Recognition facial recognition library, which has a sharper, more market-oriented accuracy, will be available as part of this work. . To generate the results, images recorded by the author and public databases available on the internet were used.

Keywords. Facial Biometrics; Deep learning; Neural Networks; Face Recognition; Oriented Gradient Histogram

Lista de ilustrações

Figura 1 – Etapas Biometria Facial	18
Figura 2 – Representação de mudanças classificador binário	19
Figura 3 – Hierarquia aprendizado supervisionado	20
Figura 4 – Explicação do classificador AM	22
Figura 5 – Tabela de características de sistemas AM	22
Figura 6 – Ilustração de classificação do algoritmo SVM	24
Figura 7 – Exemplo SVM não linear	24
Figura 8 – Etapas do algoritmo HOG	26
Figura 9 – Exemplo de uso do gradiente para reconhecimento no algoritmo HOG .	27
Figura 10 – Ilustração de classificação algoritmo KNN	28
Figura 11 – Valor de d no KNN, geração de formatos	29
Figura 12 – Ilustração de comparação no valor de k para o KNN	30
Figura 13 – Primeira ilustração de um neurônio artificial	34
Figura 14 – ilustração de uma Rede Neural multicamada	35
Figura 15 – Passos completos do algoritmo convolucional	37
Figura 16 – Primeira etapa do algoritmo convolucional	37
Figura 17 – Segunda etapa do algoritmo convolucional	38
Figura 18 – Terceira etapa do algoritmo convolucional	38
Figura 19 – Representação da rede neural ResNet	39
Figura 20 – Ilustração de detecção de um ângulo desfavorável	40
Figura 21 – Ilustração de modificações faciais e a dificuldade de detecção	40
Figura 22 – Elementos faciais que dificultam a detecção facial	41
Figura 23 – A sobreposição de rostos na detecção facial	41
Figura 24 – Metodo de detecção facial com agrupamento de pixels	42
Figura 25 – Passos da detecção baseada em redes neurais	44
Figura 26 – Ilustração da rotação de imagens para detecção facial	45
Figura 27 – Metodologia aplicada no projeto	46
Figura 28 – Painel de controle sistema criado	50
Figura 29 – Ilustração de captura de rosto pelo sistema	51
Figura 30 – Conjunto de imagens após transformações em cinza e escala	51
Figura 31 – Amostra do conjunto de características da base de treinamento	52
Figura 32 – Representação do reconhecimento de face feito por <i>Eigenface</i>	52
Figura 33 – Erros reconhecimento	53
Figura 34 – Erro no reconhecimento da utilização do <i>Eigenface</i>	53

Lista de abreviaturas e siglas

AM	Aprendizado de Máquina
ANN	<i>Artificial Neural Network</i>
CNN	<i>Convolutional Neural Network</i>
HOG	Histograma de Gradientes Orientados
IA	Inteligência Artificial
KNN	<i>K-Nearest Neighbors</i>
MLP	<i>Multilayer Perceptron</i>
PCA	<i>Principal Component Analysis</i>
SVM	Máquina de Vetores de Suporte
UFMA	Universidade Federal do Maranhão
XML	Extensible Markup Language

Sumário

1	INTRODUÇÃO	12
1.1	Objetivos	13
1.2	Trabalhos Relacionados	13
1.3	Estrutura do Trabalho	15
2	FUNDAMENTAÇÃO TEÓRICA	16
2.1	Biometria Facial	16
2.2	Aprendizado de Máquina	19
2.3	Máquina de Vetores de Suporte(SVM)	23
2.3.1	Histograma de Gradientes Orientados(HOG)	25
2.4	K-Nearest Neighbors(KNN)	27
2.4.1	Eigenface	30
2.5	Redes Neurais	33
2.5.1	Aprendizagem Profunda	35
2.5.2	Redes Neurais Convolutivas	36
2.5.3	ResNet	38
2.6	Detecção e Reconhecimento Facial	39
3	METODOLOGIA	46
3.1	Escolha do método de reconhecimento	47
3.2	Banco de imagens de entrada	47
3.3	Detecção de faces	47
3.4	Extração de características	47
3.4.1	Algoritmo de pré-processamento	47
3.4.2	Extração no <i>Eigenface</i>	48
3.4.3	Extração na Rede Neural	48
3.5	Cálculo da distância euclidiana	48
3.6	Calculando informações recebidas	49
4	RESULTADOS	50
4.1	Resultado com o <i>Eigenface</i>	50
4.1.1	Resultado da detecção	51
4.1.2	Resultado de reconhecimento	52
4.2	Resultado com a Rede Neural	53
5	CONCLUSÃO	54

5.1	Trabalhos Futuros	54
------------	--------------------------	-----------

	REFERÊNCIAS	55
--	--------------------	-----------

1 INTRODUÇÃO

Conhecida como uma das ferramentas no combate aos problemas de deficit de Segurança Pública conforme se vê em MORAES (2010), a atividade de Biometria Facial, tem como um dos principais objetivos ajudar na identificação de pessoas desconhecidas, trazendo para a população, uma maior segurança e comodidade no seu convívio social. De acordo com Filho, Silva e Neto (2016), uma das maiores inseguranças que podem acontecer na entrada de condomínios e residências, é a falta de conhecimento por parte dos seguranças ou porteiros, em identificar quais são os verdadeiros moradores que adentram tais locais, e muitas vezes o processo de identificação é feito de forma manual através de anotações ou pela memória humana.

A atividade de Biometria Facial, utiliza-se de câmera conectada a uma central de processamento, que tem como objetivo a captura de imagens e seu transporte para análise em algoritmo de detecção e reconhecimento facial, onde é possível um agente humano visualizar o processo e tomar a ação de acordo com o resultado fornecido.

Aliado à atividade de Biometria Facial, a Inteligência Artificial(IA) é a principal ferramenta utilizada no reconhecimento de pessoas, aprimorando e modernizando a atividade de vigilância assistida, e coletando informações de forma segura e com menor custo. De acordo com Filho, Silva e Neto (2016), uma das formas mais utilizadas para a maior segurança na identificação de pessoas em território particular, é a identificação de qualquer pessoa que entre no espaço de uma residência ou condomínio, tal gesto pode gerar uma melhora no ambiente de vigilância residencial, mas podem possuir algumas falhas de identificação no falso fornecimento de informações, além do tempo gasto na execução de todo esse processo.

A área da Inteligência Artificial é largamente utilizada nesse processo tecnológico de segurança, a IA tem como objetivo o desenvolvimento de técnicas computacionais sobre o aprendizado para a criação de modelos computacionais que possui a capacidade de capturar informações de forma automática, tal captura, proverá o acúmulo dessas informações, que serão classificadas e rotuladas, para haver algum sentido na sua existência. Com base em soluções de sucesso na resolução de problemas, foi criado assim o que é chamado de experiência computacional. Distingui-se dos diversos tipos de aprendizado através de sua linguagem de descrição, modo, paradigma e forma de aprendizado, como se observa em Monard e Baranauskas (2003).

Conforme observado por Nunes et al. (2017), ao longo dos anos, muitos algoritmos de detecção facial foram criados, contudo, nenhum deles ainda possui a mesma eficiência de detecção do ser humano, ou seja, dado um ambiente qualquer, o olho humano tem

menor chance de erros para apontar uma face, contudo, com relação ao reconhecimento facial, uma pessoa tem a capacidade com maior limitação na construção de uma memória de faces reconhecidas, ou seja, tem-se uma menor precisão na identificação de uma grande quantidade de rostos possíveis. Mesmo que uma pessoa possua uma memória excelente para o reconhecimento, tal habilidade fica limitada àquela pessoa em questão, causando problemas no cenário de impossibilidade de trabalho por parte desse profissional.

1.1 Objetivos

O objetivo deste trabalho é propor uma análise que utilize técnicas de processamento de imagem, para reconhecimento de elementos em vídeo, bem como faces de pessoas.

Destacam-se como objetivos específicos deste trabalho:

- Analisar algoritmos de detecção e reconhecimento facial que podem ser usados em sistemas de Biometria Facial.
- Utilizar técnicas de Classificação automática de elementos em vídeo.

1.2 Trabalhos Relacionados

Um protótipo foi desenvolvido por Okabe e Carro (2015), com o objetivo de reconhecer pessoas utilizando imagens capturadas por câmeras de segurança, para tanto se utilizou técnicas de processamento para reconhecer e identificar face, utilizando os algoritmos *Haar Cascade*, *Eigenface*, *Fisherface* e LBPH, o protótipo foi desenvolvido utilizando a biblioteca OpenCV em Java, em seus experimentos obtiveram resultados relacionados aos erros ao detectar pessoas em posição frontal, para o algoritmo *Eigenface* teve um erro de 74%, *Fisherface* obteve 81,8% e para o LBPH obteve 81,8%, utilizando-se de pouco custo no processamento computacional.

Um estudo foi desenvolvido por Braga et al. (2013) com o objetivo de desenvolver um sistema de reconhecimento facial automático, para o desenvolvimento da aplicação, o autor utilizou a plataforma MATLAB, em seu teste utilizou 3 bases de dados que contém imagens de rostos de pessoas, utilizou o algoritmo de aprendizagem *Adaboost*, e a detecção através do algoritmo *Viola-Jones*, obtendo como resultados o tempo médio de 1,6 segundos para analisar 200 imagens, com 130 corretamente, 1 incorretamente e 69 rejeitadas pelos métodos de aceitação.

Um estudo de reconhecimento de padrões faciais foi proposto por Silva e Cintra (2015) afim de estudar técnicas biométricas faciais. Para o desenvolvimento de seus experimentos foram utilizados bases de dados públicas, os autores propuseram um algoritmo híbrido denominado como *Principal Component Analysis* (PCA) e *Linear Discriminant*

Analysis (LDA), quando comparado com outros trabalhos encontrados na literatura, seu método teve um desempenho bastante significativo obtendo cerca de 97,75% de acurácia, melhorando em 25% ao LDA-NN e 40% a forma clássica do LDA, para o reconhecimento.

No trabalho desenvolvido por Wong e Cipolla (2007) é apresentada uma metodologia para o reconhecimento de atividade humana em vídeo, bem como expressões faciais, gestos com as mãos e movimentos do corpo. Utilizou a Análise semântica probabilística latente (PLSA), SVM e o classificador de vizinho mais próximo (NNC), deste modo em seus experimentos obtiveram 86,62% de acurácia.

Na dissertação de SANTOS, STEMMER e CASAGRANDE (2018) foi desenvolvido um sistema de rastreamento de múltiplos alvos voltados para técnicas de visão computacional, utilizando *Background Extractor*, que se trata de uma técnica para remoção de plano de fundo, o *Fast Retina Keypoint* e o filtro *Kalm*. Deste modo propõe o RAMUK (Rastreador baseado em filtro de Kalman) e o RAMUK-DL (Rastreador baseado em filtro de Kalman e Descritores Locais) experimentados no rastreamento de pessoas em vídeo, onde se realizou um comparativo com outros trabalhos e obteve resultados de média de comparação dos melhores parâmetros em 63,62% e 69,85%, para a altura mínima de 30 pixels tem a acurácia de 63,08%.

Na tese de Almeida (2017) são avaliadas duas técnicas de pontos de interesse espaço-temporais, aplicadas no reconhecimento de ações humanas, como sentar, andar, acenar, etc., em uma sequência de imagens de vídeo, para isto se utiliza redes neurais e máquina de vetores de suporte como classificadores. Em seus experimentos sua melhor acurácia obteve 76,85%.

Em Tramontin et al. (2016) foi proposto por um sistema embarcado utilizando reconhecimento facial utilizando técnicas de reconhecimento facial e visão computacional, este descreve a formulação matemática da técnica PCA, e a aplicação em sistemas de reconhecimento facial, para validação se utilizou um dispositivo de baixo custo que obteve resultados promissores em seus experimentos uma acurácia de 75% para reconhecimento de pessoas previamente cadastradas.

Uma investigação de trabalhos que abordam os problemas de reconhecimento em vídeo que utilizaram técnicas de *Deep Learning* é realizada por Dias et al. (2018), neste é utilizado uma rede convolutiva para a classificação de uma base de dados com 1.134 imagens adquiridas de duas câmeras. Foram realizados 4 experimentos, com o objetivo de identificar pessoas, onde realizou uma análise sobre a influência do desempenho do pré-processamento das imagens, número de épocas de treinamento da rede e classes, deste modo o resultado de melhor expressão utilizou 16 classes com uma acurácia de 96,99%.

É possível destacar Fontaniva (2014) que se trata de um trabalho encontrado na literatura, que aborda o problema de reconhecimento facial utilizando técnicas de

visão computacional, na construção de um sistema. Este utilizou os algoritmos: HOG, *Adaboost*, SVM, *FisherFaces* e *Eigenface* para identificar rostos de pessoas e reconhecê-las em câmera em uma sala de aula, seu trabalho apresentou como resultado para um número de treinamento de 50 imagens, tem-se uma acurácia de 86,25%.

1.3 Estrutura do Trabalho

Este trabalho está organizado em cinco capítulos, de forma a apresentar o conteúdo mais claramente, conforme os parágrafos a seguir.

No Capítulo 2 será apresentada a fundamentação teórica necessária para compreensão e embasamento deste estudo.

O Capítulo 3 descreve a metodologia utilizada neste trabalho.

No Capítulo 4 serão apresentados e discutidos os Resultados obtidos neste trabalho.

No Capítulo 5 é apresentada a Conclusão e os trabalhos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta a fundamentação teórica utilizada no desenvolvimento deste trabalho, para compreensão das técnicas utilizadas na metodologia proposta de análise da detecção e reconhecimento facial em sistemas de biométrico facial. Assim, ao decorrer do capítulo será discutida a base para o desenvolvimento deste estudo, partindo da compreensão da Biometria Facial e a técnica de Detecção e Reconhecimento Facial.

2.1 Biometria Facial

De acordo com Nunes et al. (2017), desde as civilizações antigas, os pequenos conjuntos populacionais já possuíam as habilidades de reconhecimentos com grande facilidade, contudo, com a expansão territorial das pessoas, em conjunto com a alta taxa de crescimento populacional, esse processo de reconhecer pessoas, tornou-se algo de enorme complexidade, que tem como solução imediata apresentada, a utilização de sofisticados sistemas de identificação de pessoas, essa identificação corresponde a um conjunto de dados relacionados aquele indivíduo.

O termo Biometria é a junção de dois sentidos, bio vem de vida e metria vem de medida, ambos com origem do grego, ou seja, em seu sentido completo, a Biometria Facial é basicamente a medida da face de um ser vivo, utilizando-se de características próprias, em função de autenticar-se de forma mais segura.

Utilizada em diversos campos de segurança, a Biometria tem como umas das suas principais utilidades a autenticidade de pessoas para acessos a determinados locais ou recursos, possui-se dois tipos de autenticação conhecidos: a verificação e identificação, conforme se observa em Costa, Obelheiro e Fraga (2006).

De acordo com Bolle et al. (2013), a verificação é a alegação do usuário em dizer que é tal pessoa ou que tais dados estão associados à pessoa em questão, o usuário fornece alguns dados cadastrais que são previamente registrados, como, por exemplo, seu e-mail e senha, ou seu número de registro etc, então para acessar recursos ou locais, o usuário fornece seus dados biométricos alegados, a comparação para verificar é dita 1:1, pois o conjunto de dados de verificação é comparado a um único conjunto de dados presente em uma base de dados.

Esse método de acesso é utilizado nos dias atuais em diversos sistemas pela internet como redes sociais, lojas online e até mesmo os bancos, de acordo Assunção (2002), diversos tipos de ataques que podem roubar credenciais de verificação e causar vários prejuízos financeiros e sociais. Continuando com as afirmações de Costa, Obelheiro e

Fraga (2006), na identificação, tem-se o fornecimento de dados biométricas do próprio usuário, ou seja, características daquele usuário que são intransferíveis, e que representam padrões codificados não disponibilizados em forma de dados de fácil legibilidade, sendo de responsabilidade do sistema identificar o usuário.

Nesse contexto tem-se uma comparação de acesso 1:N, nomeada de busca aberta, com dados biométricos diversos que possuem maiores ou menores mudanças de características, de acordo com o tipo biométrico utilizado, tem-se aqui um limiar a ser aceito para que julgue a correta rotulação no processo de identificação, usualmente pode-se mesclar a verificação com a identificação, ou pode-se associar diversos tipos de identificação.

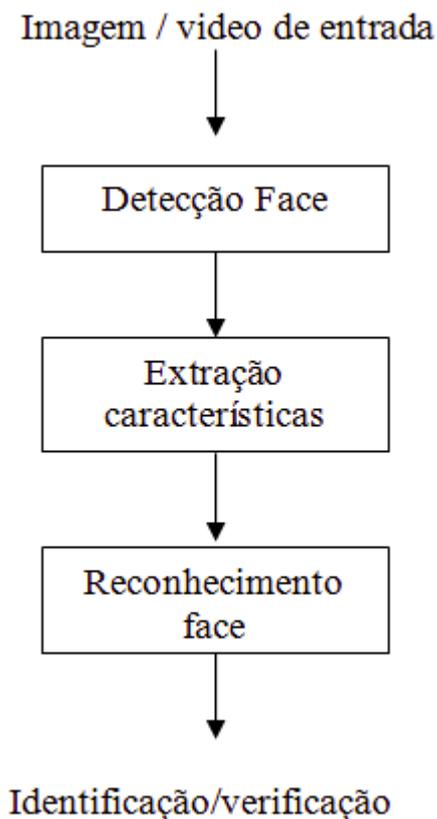
Como se observa em MORAES (2010), a biometria facial é apenas uma entre as diferentes formas de biometrias conhecidas, essa forma de identificação, possui algumas vantagens que estão criando o seu destaque na sua utilização por setores da sociedade, umas dessas vantagens é a forma não invasiva de aparelhos de captura de imagens, o baixo custo e a facilidade para a sua utilização, e a constante evolução no poder de processamento de dispositivos de registro de imagem. Na Figura 1, têm-se os passos para o Reconhecimento Facial, constituído pela detecção facial, ou seja, saber se existe uma face naquela imagem, a extração de características, que são os dados ali presentes e o reconhecimento de face, que consiste em sabe a quem pertencem o rosto identificado.

Um dos conceitos importantes no estudo da Biometria é o de identificação de face e controle de acesso, conforme diz MORAES (2010), tem-se uma base de características e dados associados a determinados indivíduos previamente cadastrados, e uma comparação entre os dados fornecidos pela face da pessoa e os dados conhecidos na base de dados, através de cálculos probabilísticos, uma determinada taxa de aceitação deve ser satisfeita, caso o rosto seja detectado, mas não conhecido pelo sistema, tem-se um caso de falsa identificação, em resumo, no controle de acesso, tem-se como retorno o valor binário de aceitação para detecção de um rosto na imagem, enquanto que na identificação de face, tem-se como retorno negativo para o rosto desconhecido.

Outros motivos para aceitação da Biometria Facial, vem dos valores biológicos, geralmente o primeiro conhecimento consolidado da mente humana, é o reconhecimento entre pessoas, é algo natural que pode ser mais aceitável dentre os sistemas biométricos, por conta dessa identificação entre o sistema e o ser humano, o avanço na forma biométrica facial vem superando em utilização, alguns outros sistemas com maior confiabilidade, como, por exemplo, o de impressão digital e o reconhecedor de Íris, como se pode ver em Costa, Obelheiro e Fraga (2006).

Coincidente com o que fala MORAES (2010), tem-se a noção de que várias outras formas de autenticação, possuem atos falhos e prejudiciais que podem acarretar queda de toda a tentativa de segurança imposta no local, um exemplo disso é a utilização de cartões magnéticos, senhas, passaporte, etc., a falsificação e perda desses materiais reforça

Figura 1 – Imagem representando as etapas do Reconhecimento Facial



Fonte: (MORAES, 2010)

toda a tese apresentada da falta de confiabilidade. Ainda conforme MORAES (2010), o desempenho do sistema biométrico, depende da sua similaridade entre padrões, raramente tem-se uma compatibilidade de 100%.

A porcentagem não tem um valor fixo, pois há uma grande variação no ambiente de aplicabilidade do algoritmo, como, por exemplo, têm-se as mudanças na face humana, onde pode-se encontrar uns óculos, uma barba ou até mesmo uma simples maquiagem que pode influenciar nessa alteração de taxa comparativa. O limiar utilizado na biometria, recebe o nome de limiar de casamento, que consiste em um limiar que determina cada tipo de usuário de entrada de acordo com o resultado do seu reconhecimento, como fator de classificação, têm-se as classes seguintes:

- Verdadeiro Positivo: Tem-se como verdadeiro no padrão resultante, liberando seu acesso para os recursos de interesse.
- Verdadeiro Negativo: Tem-se como falso no padrão resultante, recusando seu acesso para os recursos de interesse.

- Falso Positivo: Tem-se como falso no padrão resultante, mas verdadeiro em classificação, liberando seu acesso para os recursos de interesse.
- Falso Negativo: Tem-se como verdadeiro no padrão resultante, mas como falso em classificação, recusando seu acesso para os recursos de interesse.

Na Figura 2, tem-se uma ilustração dos itens mencionados, assim pode-se realizar a medida da ação de cada tipo de resultado da Biometria Facial.

Figura 2 – Imagem representando as mudanças de um classificador binário com os resultados da biometria

		Classe Real	
		Positiva	Negativa
Classe Sugerida pelo Classificador	Positiva	Verdadeiro Positivo	Falso Positivo
	Negativa	Falso Negativo	Verdadeiro Negativo

Fonte:(MORAES, 2010)

A sociedade atual tem o privilégio de acompanhar grandes avanços da tecnologia, como se destaca em MORAES (2010), esses avanços estão relacionados com a facilidade que o ser humano possui em realizar determinadas tarefas, como exemplo tem-se a transação bancária mais segura, não ter que decorar inúmeras senhas de acesso, ter mais segurança no ambiente que vive e trazer confiabilidade em sistemas de ensino à distância.

2.2 Aprendizado de Máquina

Sendo uma área da IA, conforme Monard e Baranauskas (2003) o Aprendizado de Máquina(AM) tem a finalidade de criar métodos computacionais sobre a forma de aprender ou consumir conhecimento, visando atingir de forma automatizada tomadas de decisões e adquirir experiências. Existem características entre os sistemas de Aprendizado, com peculiaridades semelhantes e distintas, como, por exemplo, sua linguagem de descrição, modo, paradigma e a maneira como o aprendizado é conduzido. Através do processo de indução que é orientado por um modelo de aprendizado baseado em exemplos, defini-se aqui dois tipos de Aprendizados de Máquina, o supervisionado e o não-supervisionado.

No aprendizado supervisionado, o qual é utilizado no presente projeto, tem-se a apresentação ao indutor, um conjunto de casos de treinamento com diversas características interligadas a um rótulo de identificação, cada exemplo utilizado no treinamento, possui uma lista de informações sobre tal peculiaridade. O objetivo desse algoritmo, é catalogar grupos de dados para construir um classificador que caracterize novas informações de entrada do sistema, para assim associar tais entradas com dados pre-existentes. Para rótulos discretos possuímos a classificação, para os contínuos existem a regressão.

O aprendizado não-supervisionado é a tentativa de agrupar exemplos fornecidos, tais conjuntos são denominados de clusters, de acordo com suas características e com o contexto do problema a ser resolvido, esse processo de classificação pode ser repetido inúmeras vezes para ajustar alguns parâmetros, ou até mesmo, adicionar novos atributos. Não existe a figura do indutor ou supervisor no aprendizado não-supervisionado. Na Figura3 a seguir, descreve-se a hierarquia dos aprendizados, as elipses em cinza representar o modelo supervisionado com classificação.

Figura 3 – Imagem representando a hierarquia do aprendizado supervisionado



Conforme é descrito em Michalski et al. (1998), os sistemas de aprendizado são classificados em duas categorias:

- Caixa-preta: o desenvolvimento da sua classificação é algo obscuro, ou seja, não há

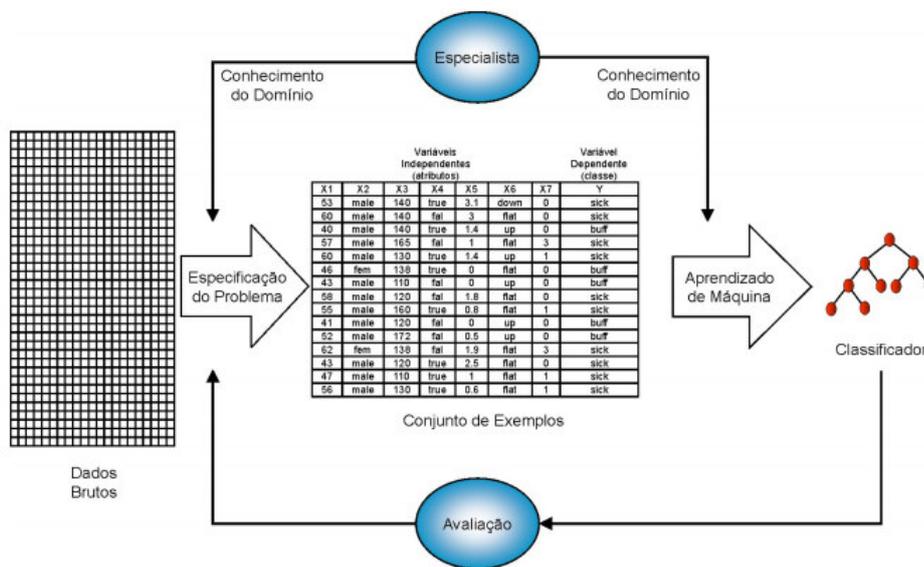
um padrão claro de procedimentos que seja compreensível à análise humana.

- Orientados a conhecimento: tem como uma das suas finalidades, criar estruturas que sejam de fácil compreensão humana.

Destacam-se também os paradigmas do AM, ou seja, alguns padrões que definem diferentes tipos de aprendizados, abaixo exemplos de alguns paradigmas a seguir de acordo com Monard e Baranauskas (2003):

- Simbólico: Nesse sistema, cria-se uma lista de conceitos a partir de exemplos e contra exemplo utilizando símbolos, esses símbolos são representados através de expressões lógicas que envolvem o uso de árvore de decisão, regras ou rede semântica. Como é observado em Monard e Prati (2006), um exemplo de utilização do modelo Simbólico é na mineração de dados, pois ele gera resultados que são de fácil compreensão humana.
- Estatístico: Consiste em utilizar modelos estatísticos para encontrar padrões indutivos. A maioria dos métodos do modelo estatístico, são paramétricos, ou seja, os dados têm origem em um conjunto probabilístico e se afirma verdades sobre parâmetros desse conjunto Zuben et al. (1996), em outras palavras, cria-se modelos estatísticos e então classificam-se as características de acordo com cada modelo. Como exemplo, tem-se a combinação linear que crie combinações lineares com os dados trabalhados, e depois faça uma comparação com os dados de entrada para realizar a distribuição de padrões (classificação). O método estatístico de maior destaque é o Bayesiano, que conforme Pardo e Nunes (2002) consiste em um aprendizado supervisionado, que tem como vantagem o poder de colocar junto às probabilidades, o conjunto de domínio caso o tenha, e também a possibilidade de diminuir ou aumentar a probabilidade das características que se desejam classificar. Na Figura4 é feita uma representação de um modelo de classificação de um Aprendizado de máquina supervisionado.
- Baseado em exemplos: *"Uma forma de classificar um exemplo é lembrar de outro similar cuja classe é conhecida e assumir que o novo exemplo ter a mesma classe"* Monard e Baranauskas (2003), seguindo esse raciocínio, pode-se deduzir um modelo baseado em exemplos, ou seja, utilizam formas pré-classificadas para padronizar as entradas desconhecidas. A denominação desse tipo de sistema é chamada de Lazy, no qual se mantém em memória os exemplos utilizados na supervisão dos padrões para classificadores. Aha, Kibler e Albert (1991) apresenta um exemplo do Nearest Neighbours e Raciocínio Baseado em Casos (RBC), que possuem a maior notoriedade nesse tipo de algoritmo.
- Conexionista: São os modelos que representam as Redes Neurais, que são modelos matemáticos inspirados nas conexões dos modelos biológicos do cérebro humano, as

Figura 4 – O entendimento compactado fornecido pelo classificador presente na direita.



Fonte: (MONARD; BARANAUSKAS, 2003)

unidades dessas redes são altamente conectadas Monard e Baranauskas (2003), e hoje representam o modelo mais utilizado na área de aprendizado de máquina (em incluso o presente projeto), não é por acaso que várias pesquisas relacionadas às Redes Nerais, possuem semelhanças com atividades feitas pelo ser humano, como, por exemplo, a visão computacional (visão humana) e o reconhecimento de voz(a fala humana).

- Genético: Semelhante ao que se vê nas teorias de Darwin Papavero (2003), este paradigma tem como origem o modelo evolucionário de aprendizagem, um conjunto de elementos pré-dispostos realizam competições probabilísticas entre si para fazer uma predição, ou seja, tentar atingir um melhor resultado no futuro com base em combinações em buscas paralelas e adaptativas Pacheco et al. (1999).

Figura 5 – Características de Sistemas de AM

Modos	Paradigmas	Formas	Linguagens de Descrição
Supervisionado	Simbólico	Incremental	Exemplos ou Objetos - \mathcal{L}_E
Não-Supervisionado	Estatístico	Não-Incremental	Hipóteses - \mathcal{L}_h
Semi-Supervisionado	Baseado em Exemplos Conexionista Evolutivo		Conhecimento do domínio - \mathcal{L}_K

Fonte: (MONARD; PRATI, 2006)

Tabela representada na Figura 5, simplificada para o entendimento das características do Sistema de Aprendizado de Máquina citada no início desse tópico.

2.3 Máquina de Vetores de Suporte(SVM)

De acordo com Soares (2008), a máquina de vetores de suporte, é uma parte constituinte de aprendizados estatísticos, que tem como uma de suas características a possibilidade de abranger suas possibilidades de generalização, sendo mais uma metodologia que tem a capacidade de extrair padrões a parte de amostras de dados.

Criada por Vapnik, Guyon e Hastie (1995), a SVM tem como base de suas definições, a teoria do aprendizado estatístico, que de acordo com Henriques (2008), possui como diferencial em relação a outras técnicas de classificação, a possibilidade de evolução no seu vetor de pesos w em meio ao processo de aprendizagem, focando-se em desvios da função final $f(x)$ em relação à função dita real, $F(x, w)$, com um valor de entrada x . Com a definição na teoria de aprendizagem estatística, tem-se a função 2.1 a seguir:

$$y = f(x) \quad (2.1)$$

para problemas de regressão y pertence a R e x pertence a R^N , em problemas de multi-classes, y pertence a N , já para reconhecimento de padrão binário, tem-se que $y \in (\pm 1)$, para que todo o processo aconteça, utiliza-se um grupo de exemplos de classes y_i em função do padrão x_i .

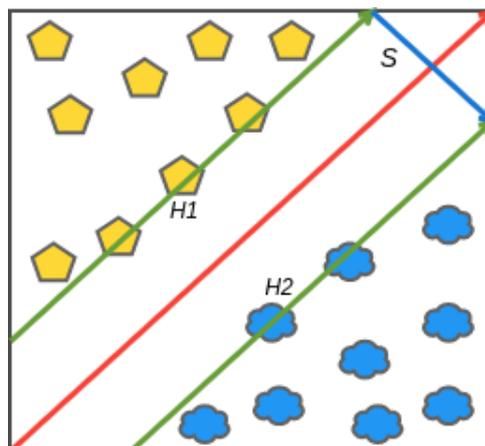
A base de aprendizagem estatístico tem como foco matemático o controle das formas de generalização de uma máquina de aprendizagem, ou seja, tem-se uma base matemática para a criação de funções para classificar dados, assim como para o seu treinamento e teste de novas entradas, evitando-se sempre a ocorrência de duas coisas, tanto a memorização dos dados de treinamento o chamado *overfittin*, quanto a incapacidade de criar novas generalizações, que se chama de *underfitting*, conforme Henriques (2008).

Conforme Soares (2008), os SVMs tem como finalidade, criar soluções através da eficiência em treinamento e representando a não linearidade em funções. Inicialmente, tem-se uma forma de separação de características, na Figura 6 tem-se a representação de dois padrões distintos, que são separados de forma linear por um hiperplano de margem de separação S .

Na linha H_1 tem-se a fronteira de separação para a classe (pentágono), enquanto no H_2 a fronteira da classe nuvem, ambos com 3 e 2 objetos de delimitação de fronteira respectivamente, tais objeto que marcam essa fronteira de características são denominados de Vetores de Suporte.

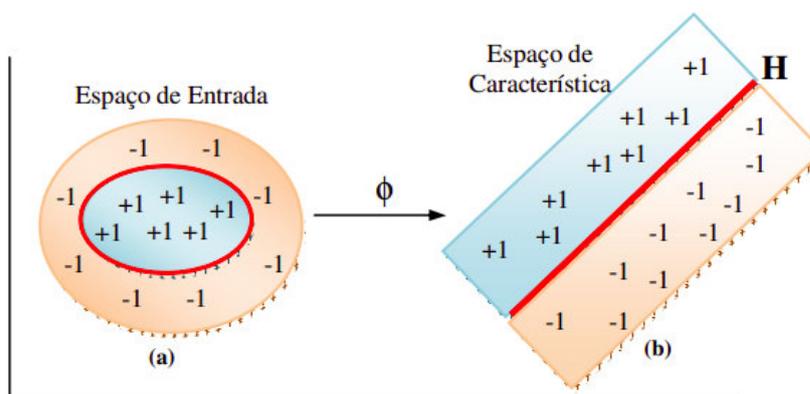
A Determinação do hiperplano é definida pelos próprios Vetores de Suporte, contendo uma limitação máxima, tem-se a representação da classificação, característica principal do sistema SVM. Para problemas onde a classificação linear não produz uma solução satisfatória, tem-se a opção de aumentar o poder dimensional, para separar as

Figura 6 – Imagem de exemplo do SVM, inspirado no modelo apresentado em Soares (2008)



características num plano no espaço, como se pode ver na Figura 7, através das funções de características, usam-se métodos não-lineares afim de evitar a busca exacerbada em uma alta dimensão de valores, essa função com não linearidade se chama de *Kernel*.

Figura 7 – Imagem de exemplo do SVM não separável de forma linear, usando multi-dimensão de separação de características. Soares (2008)



Fonte: (SOARES, 2008)

De acordo com as afirmações de Haykin (2007), operando no meio de entrada, o kernel executa a soma ponderada determinada pelos vetores de suporte, para atingir o objetivo de classificação, ou seja, sem considerar de fato o espaço de análise para separação de dados, o *Kernel* consegue criar um excelente hiperplano para delimitar diferentes características. Estando definido o ambiente multidimensional, o SVM escolhe um hiperplano com base na sua margem máxima, ou seja, a distância de maior amplitude entre os grupos de características presentes de treinamento, sendo assim tem-se duas categorias teóricas de classificação, a Dimensão VC (*Vapnik - Chevonenkis*) e Minimização

de Risco Estrutural.

Na Dimensão VC, tem-se um grupo de funcionalidades para catalogar o número máximo de dados de treinamento sem erro de retorno, para todas as classificações possíveis e suas funções, o objetivo é simplesmente obter o melhor resultado de acordo com a potencialização da SVM. Na minimização, oferece-se a indução como procedimento, conforme dito por Soares (2008), usa a Dimensão VC como termo de controle para obter esse melhor desempenho. Algumas características das SVM podem ser descritas a seguir, conforme em Lorena e Carvalho (2003):

- Melhor capacidade de generalização: os classificadores oriundos da SVM, em sua maioria, atingem resultados satisfatórios de generalização, ou seja, a sua efetividade em catalogar dados não pertencentes a determinado grupo de treinamento, como dito no tópico anterior, classificadores que são especializados, ou seja, do tipo *overfitting*, que obtêm resultados baixos com relação a novas entradas.
- Exuberância em dimensões de maior proporções: objetos de proporções maiores, elevam a estrutura da SVM, exemplificando podem-se citar as entradas de imagens,
- Elevação da função objetivo: utiliza-se da premissa de otimização uma função quadrática, que contem o mínimo global, uma ligeira vantagem em relação às redes neurais, que precisa minimizar suas funções objetivos.
- Eficiente definição da teoria: base teórica bem consolidada no campo da Estatística e conseqüentemente matemática.

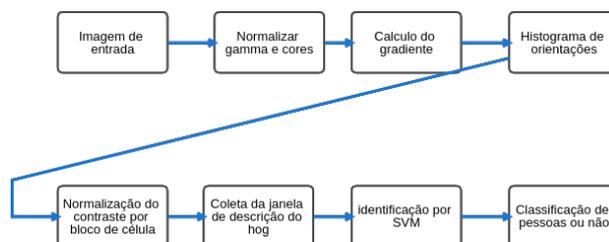
2.3.1 Histograma de Gradientes Orientados(HOG)

Algoritmo que possui como princípios, descrever o formato e a apresentação de dados por meio de distribuição de intensificadores de pixels ou pelas medidas das bordas, os processos feitos para a criação dos seus descritores são os seguintes como se observa em RIGAMONTI (2018):

- Cálculo do gradiente dos pixels.
- Agrupamento de pixels em células.
- Agrupamento de células em Blocos.
- Obtenção do descritor.

Conforme Dalal e Triggs (2005), essa técnica é baseada na geração de histogramas orientados para uma imagem através de gradientes, e realiza a classificação em conformidade

Figura 8 – Imagem de exemplo do algoritmo HOG, mostra-se as etapas para o reconhecimento

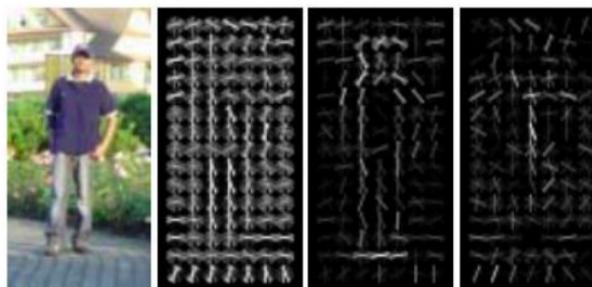


aos objetos treinados que foram definidos. Para haver de fato o reconhecimento, na Figura 8, têm-se as etapas para haver a classificação no algoritmo HOG.

- Normalizar cores: sendo uma etapa opcional, tem como objetivo melhorar o desempenho com base nas combinações de cores RGB e Lab.
- Cálculo do gradiente: é utilizado aqui, a técnica *Sobel ID*, que realiza o cálculo da norma dos pontos na imagem, realizando assim o gradiente, quando a imagem é colorida, cada cor tem seu cálculo efetuado.
- Histograma de orientações: criados na fase de cálculo dos gradientes, sendo do tipo orientado, o qual varia de 0 a 360 graus ou não orientado variando de 0 a 180 graus, basicamente são células em formato geométrico circular ou retangulares, o valor do histograma para reconhecimento de pessoas tem em média 20 graus.
- Normalização do contraste por bloco de células: tem a utilidade de evitar a sobreposição de células na imagem.
- Coleta da janela de descrição do HOG: captura a janela da imagem no qual deve-se realizar o reconhecimento.
- Identificação por SVM: fase de identificação do bloco de células otimizado.

Conforme RIGAMONTI (2018), o cálculo do gradiente é feito aplicando-se a máscara derivada discreta, no eixo horizontal e vertical para atingir todos os pixels da imagem, logo em seguida tem-se o agrupamento dos pixels para criar as células. Em seguida, agrupam-se as células formando vários blocos ao longo da máscara, em seguida são criados os descritores, que são uma listagem com as células dos blocos analisados, para finalizar, realiza-se o treinamento através dos descritores em um aprendizado supervisionado, na Figura 9, tem-se um exemplo desse processo com a formação da máscara de características.

Figura 9 – Imagem de exemplo do algoritmo HOG, descida do gradiente na imagem para treinamento



Fonte: (RIGAMONTI, 2018)

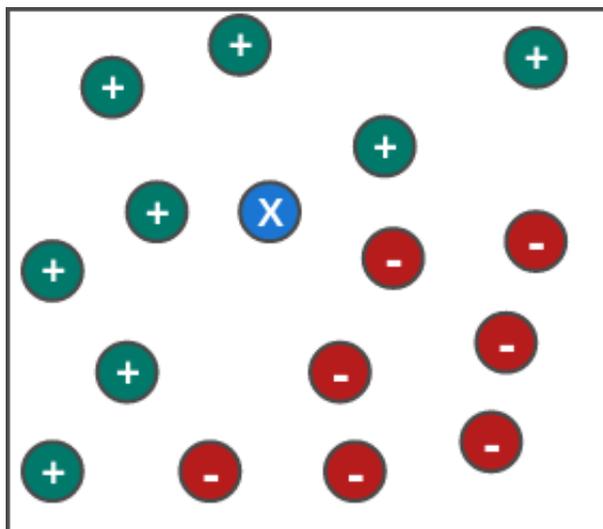
2.4 K-Nearest Neighbors(KNN)

Algoritmo introduzido em meados dos anos de 1991 por Aha, Kibler e Albert (1991), tem como característica principal, a ideia de buscar as K amostras identificadas mais próximas do exemplo não rotulado que são classificados, com base nessa aproximação de características, pode-se citar a classificação do exemplo de entrada. De acordo com Ferrero (2009), os algoritmos KNN exigem pouco recurso computacional durante a fase de treinamento, contudo, tais recursos são elevados quando se trata de classificação e rotulação de um novo exemplo de entrada, pois haverá a sua comparação com todos os outros exemplos pré-existentes. Assim uma ilustração deste processo é representado na Figura 10, um problema para catalogar novos exemplos no pós treinamento, os rótulos descritos como positivos(+) refere-se a pessoa acima do peso, por exemplo, enquanto os negativos(-) descrevem-se uma pessoa com peso normal, utilizando o algoritmo KNN nesse caso, tem-se um $k = 1$, o novo elemento X a ser classificado como positivo, pela sua proximidade com elementos desse tipo.

Conforme Ferrero (2009), determina-se a seguir os parâmetros fundamentais para a execução do algoritmo KNN:

- Conjunto de dados para treinamento: há uma proporcionalidade entre a quantidade de exemplos a serem treinados e o reajuste para encontrar a classificação adequada para a entrada do algoritmo utilizado neste trabalho, pois há a comparação do elemento que desejo classificar, com todos aqueles já treinados de acordo com Alpaydin (2004), caso essa quantidade de exemplos seja muito grande, haverá uma maior lentidão, até chegar um momento de lentidão contínua sem resposta. Para evitar tais problemas, deve-se utilizar apenas os dados mais importantes para cada classe.
- Medida de semelhança entre os exemplos: Existe a importância de determinar quais medidas serão adotadas para comparar aquilo que será rotulado, com o objeto que já

Figura 10 – Imagem de exemplo do KNN, inspirada no modelo apresentado em Ferrero (2009)



é rotulado no grupo de treinamento, algumas dessas medidas podem ser encontradas em Dubes e Jain (1988) e Everitt (1993). Quando as medidas descritas são aplicadas por cardinalidades, existe uma facilidade maior em determinar suas semelhanças e diferenças em relação a todos os dados de treinamento, Ferrero (2009) demonstra, onde os pares (E_i, E_j) , determinados pelas condições:

- I - $dist(E_i, E_j) \geq 0, \forall(i, j)$ (*positividade*)
- II - $dist(E_i, E_j) = 0 \leftrightarrow E_i = E_j$ (*identidade*)
- III - $dist(E_i, E_j) = dist(E_j, E_i)$ (*simetria*)
- IV - $dist(E_i, E_j) = dist(E_i, E_q) + dist(E_q, E_j), \forall(i, j, q)$

Aplicando-se a propriedade de desigualdade triangular, simétrica, identidade e positiva, obtêm-se uma métrica entre os dados a serem comparados.

Usualmente, observa-se esses dados no formato de um plano cartesiano, contendo uma reta ou uma sigmoide, distinguindo as fronteiras dos diferentes dados, de acordo com as métricas aplicadas, sendo assim, deduz-se que além do plano bidimensional, têm-se uma distinção em planos maiores para melhorar a classificação das entradas, é por isso que se apresentar agora a fórmula criada por Bohnenblust (1938) para calcular de forma genérica, a distância entre dois pontos no espaço n-dimensional de acordo com o parâmetro d , as atribuições citadas nas condições de medidas, são n-dimensionais, portanto, cada descrição corresponde a um ponto no universo do conjunto $E_i = (x_{i1}, x_{i2}, \dots, x_{in})$.

$$dist(E_i, E_j) = \left(\sum_{l=1}^n |x_{il} - x_{jl}|^d \right)^{\frac{1}{d}} \quad (2.2)$$

O valor de d na equação anterior, está diretamente relacionado ao formato da figura geométrica formada pelos pontos de igual distância do centro, quando se têm d tendendo ao infinito, a formação exata de um quadrado, $d = 1$ ou $d = 2$, é obtida a distância de Manhattan e uma Euclidiana respectivamente, em Aggarwal, Hinneburg e Keim (2001) representa o valor de d para 1 e 2, gerando valores melhores que $d \geq 3$, de acordo com Ferrero (2009), os resultados gerados, para valores menores em d eram bem melhores, como representado na Figura 11. Através do coeficiente de correlação, determinam-se as semelhanças e diferenças entre os dados que se busca comparar, essas comparações podem ser feitas através da semelhança dos dados, e também pela magnitude dos seus valores aproximados, como as medidas feitas pelo coeficiente de correlação é feita levando em consideração vários tipos padrões de dados, não há uma concordância com a desigualdade triangular, portanto, pode-se batiza-la de medidas semi-métricas, pois podem ou não levar em consideração dados que sejam do grupo de triangular concordância.

Figura 11 – Valores de d variando na medida de Minkowsky: (a) $d = 1$ (b) $d = 2$ (c) $d = 3$ (d) $d = 4$ (e) $d = 20$ (f) $d \mapsto \infty$

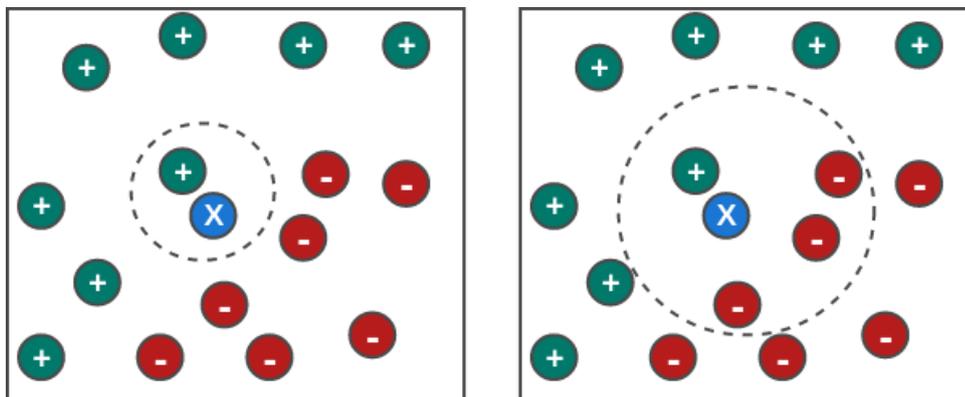


Fonte: (FERRERO, 2009)

Em Liu e Motoda (2007) alguns algoritmos de seleção de atributos foram utilizados em diversos tipos de classificação, tal trabalho é de suma importância, pois nem todo atributo tem as qualidades para serem apresentados de forma coesa, os mesmos, podem inserir atributos pouco importantes, que podem gerar apenas mais custos computacionais desnecessários.

- Número do grupo de vizinhos mais próximos: de acordo com Ferrero (2009), levando em consideração as características principais do algoritmo *KNN*, que seriam padronizar determinados tipos de dados considerando a sua classe ser ou não semelhante aos k vizinhos mais próximos, se possuímos apenas 1 vizinho próximo, então automaticamente, herda-se a classificação desse único vizinho próximo de acordo com a medida de semelhança aplicada, caso se tenha mais de um vizinho próximo, faz-se a comparação entre ambos, através dos atributos escolhidos, o modo mais comum para proceder essa comparação, é escolher a quantidade predominante entre as diversas características dos classificadores.

Figura 12 – Imagem de exemplo com a mudança do valor de k , inspirada no modelo apresentado em Ferrero (2009)



Para melhorar a compreensão, abaixo na Figura 12, a ilustração leva em consideração valores negativos(-) e positivos(+), assumindo o valor de $k = 1$ e $k = 5$ respectivamente, demonstrando que a influência do algoritmo está relacionada diretamente com o valor de k , pois ele determina o equilíbrio de comparação de $k - vizinhos$ e sua adaptação de acordo com o problema aplicado, geralmente prefere-se a quantidade ímpar de elementos a serem comparados, para sempre haver uma superioridade cardinal de uma classe em detrimento de outra.

2.4.1 Eigenface

Trata-se de uma técnica desenvolvida para detecção e identificação de rostos humanos foi proposta por Turk e Pentland (1991), desde então vem sido utilizada por diversos trabalhos na literatura, este aborda o problema de se reconhecer rostos como um problema bidimensional, deste modo são projetados em conjunto de características 2D, deste modo tendo a capacidade de reconhecimento de maneira não supervisionada.

Desta forma o algoritmo procura os principais componentes da distribuição de faces, ou os vetores próprios da matriz de covariância de um dado conjuntos de imagens. Deste modo, esses vetores tornam-se as características que compõe um rosto. Assim na equação 2.3, é representada a média do conjunto, aonde as imagens originais são representadas por uma matriz $N \times N$ que consiste em uma combinação linear, uma vez que se tratam de uma representação, são conhecidos como *Eigenfaces*, pois são subespaços das imagens faciais primárias.

$$\psi = \frac{1}{M} \sum_M^{n=1} \Gamma_n \quad (2.3)$$

Assim, estes vetores são submetidos a uma análise de componentes principais, para determinar um conjunto M com vetores ortogonais que melhor representam a distribuição

das imagens, sendo calculado a média da diferença de cada nova face dado pela equação $\phi = \Gamma_i - \Psi$, de tal forma que o vetor dado pela (ordem) k , u_k é escolhido e seja maximizado pela equação 2.4.

$$\lambda_k = \frac{1}{M} \sum_{n=1}^M (u_k^\tau \phi_n)^2 \quad (2.4)$$

A matriz de covariância é representado na equação 2.5 , e os vetores por u_k e os valores são dados por λ_k , onde a matriz é representada por $A\phi_1\phi_2\dots\phi_M$. Pela matriz C possuir dimensão $N^2 \times N^2$, torna-se complexa a determinação de seus vetores N^2 , no entanto, é possível reduzir a dimensão do número de vetores próprios de N^2 , que se trata da dimensão do número de rostos, para M que se trata do número de imagens do conjunto do treino, desde que $M \ll N^2$, os cálculos tornam-se simples.

$$C = \frac{1}{M} \sum_{n=1}^M \phi_n \phi_n^\tau = AA^\tau \quad (2.5)$$

Desta forma na equação 2.6 é representado a solução para matriz C , onde são realizadas combinações lineares para imagens faciais em ϕ , onde v_i são vetores de A^τ sendo pre-multiplicados.

$$A^\tau A v_i = \mu_i v_i \quad (2.6)$$

Demonstrando-se assim que $A v_i$ são vetores de $C = AA^\tau$ da equação 2.7, a partir desta formula pode-se construir uma matriz tal que $L = A^\tau A$, representado na equação 2.8.

$$AA^\tau A v_i = \mu_i v_i \quad (2.7)$$

$$L_{mn} = \phi_m^\tau \phi_n \quad (2.8)$$

Determinando-se os M vetores v_i de L , que são combinações lineares das M imagens que foram treinadas e dão origem às chamadas *Eigenfaces* u_i representado pela equação 2.9.

$$u_i = \sum_{k=1}^N v_{ik} \phi_k \quad (2.9)$$

Assim as *Eigenfaces* de L que possuem os maiores valores podem ser utilizados para a classificação, para tanto a face é transformada em um componente *Eigenface*, como mostra a equação 2.10.

$$\omega_k = u_k^T(\Gamma - \psi) \quad (2.10)$$

A contribuição de cada *Eigenface* é descrita pelos pesos ω_k , onde o vetor ω^T pode ser utilizado como reconhecedor de padrões faciais. Para tanto calcula-se a menor distância euclidiana dada pela equação 2.11.

$$\epsilon_k^2 = \|(\Omega - \Omega_k)\|^2 \quad (2.11)$$

Onde Ω_k descreve a classe de face, e cada Ω_i é calculado através da média dos resultados da representação dada por *eigenfaces* de um pequeno número de faces, também é possível calcular a distância entre a imagem de teste e o espaço das faces através da equação 2.12 com a equação 2.13.

$$\epsilon^2 = \|(\phi - \phi_f)\|^2 \quad (2.12)$$

$$\begin{cases} \phi = \Gamma - \psi \\ \phi = \sum_{i=1}^M \omega_i u_i \end{cases} \quad (2.13)$$

Assim esta abordagem, é dividida em 6 passos: A criação de uma biblioteca de imagens de características de rosto de indivíduos, devendo-se incluir diferentes imagens por indivíduo, com uma variedade de expressões faciais e iluminação, o cálculo da matriz L , que determina os valores dos vetores e valores próprios, e as M *eigenfaces* u_k , o cálculo e armazenamento do vetor de características por indivíduo da biblioteca, cálculo da margem de erro ϵ_k para definição da distância máxima das classes de faces, por fim cada face identificada, o vetor característico é calculado e comparado a outros vetores já armazenados, caso a comparação obtenha sucesso satisfazendo-se as condições definidas, este é dado como pessoa reconhecida, caso contrário e dado como pessoa desconhecida. Deste modo dando a possibilidade de adicionar a base.

Conforme Bissi et al. (2018), o algoritmo PCA é uma técnica da matemática que define um conjunto de dados, utilizando-se de componentes com melhor representatividade tais conjuntos de dados conhecidos, são usados para a redução da dimensão dos dados e para detecção de padrões, tal algoritmo utiliza a norma euclidiana para realizar as suas classificações.

2.5 Redes Neurais

Nos anos 50 houve um encontro no "*Darhmouth College*", que significou um início de formalização das técnicas de IA, que tinham a capacidade de adquirir experiências e transforma-las em tomadas de decisões automatizadas, que independiam da intervenção direta do ser humano conforme Shannon (1956), neste tem as definições de Redes Neurais, que é um dos paradigmas da Aprendizagem de Máquina que é inspirado em modelos da biologia para o funcionamento do sistema cognitivo humano. A compreensão do cérebro humano, para o estudo do seu funcionamento na neuroanatomia, data-se do ano de 1875, iniciado por Camillo Golgi como visto em Hubel (1979), ao longo dos anos, vários estudos foram feitos e com o avanço da neuroanatomia adquire-se a inspiração para o desenvolvimento da Rede Neural Artificial, que tem como base de semelhança com o cérebro humano, seus envios de informações entre as unidades mínimas (células nervosas) através de reações químicas e como são feitas essas ligações através das sinapses Quillfeldt (2005), que tem como característica o armazenamento de informações, onde impulsos elétricos percorrem os caminhos dos neurônios e deixam rastros que se denomina de memória de acordo com T (2018)

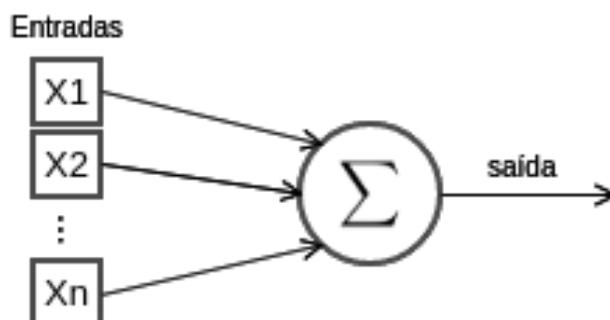
A transmissão de dados entre neurônios é feita através dos neurotransmissores, reações entre moléculas no terminal sináptico e estímulos da membrana dendrítica a estímulos, tais estímulos aumentam ou diminuem os sinais sobre os neurônios vizinhos, reagindo de acordo com os modelos de sinapses presentes. Em analogia a esse processo químico-biológico, descrevem-se as atividades e passagens de informações em uma Rede Neural Artificial, que muda seu comportamento de acordo com a variância de valores entre os neurônios, no qual é chamado de conexionismo, que é discutido em Barreto (2002) como uma técnica aplicada para resolver problemas na IA. A utilização de modelos biológicos nas Redes Neurais, não tem como objetivo criar uma cópia do cérebro humano, tem como finalidade criar modelos que são capazes de aprender e tentar resolver diferentes tipos de problemas, possuindo com seus resultados uma curva de erros, que tendem a diminuir de acordo com a experiência, que se auto-organiza no final de ciclos de aprendizado, ou mesmo apresentar comportamentos inesperados.

A partir do entendimento das Redes Neurais Biológicas, é possível compreender o início do trabalho na criação das Redes Neurais Artificiais. Com o primeiro modelo datado de 1943 criado por WARREN S. MCCULLOCH e WALTER PITTS de acordo com McCulloch e Pitts (1943), os primeiros trabalhos com Redes Neurais Artificiais teve como objetivo simular o funcionamento dos neurônios biológicos, criando um modelo simplificado que se firmou como a base das Redes Neurais que se usa hoje, a preocupação em comparar os modelos cognitivos naturais e os artificiais, foi perdendo força ao longo dos anos. Os focos principais para a criação das Redes Neurais Artificiais de acordo com Barreto (2002) foi modelar o sistema nervoso para estudar o seu comportamento e na

construção de computadores com alto grau de paralelismo.

Na Figura 13, é representado o primeiro modelo visual de neurônio artificial proposto por McCulloch e Pitts (1943), é utilizado uma representação binária como similaridade à presença e ausência de estímulos que são encontrados nos neurônios naturais, o modelo visual abaixo, foi apresentado apenas 19 anos depois em 1962 por Rosenblatt (1962).

Figura 13 – Modelo de Rede Neural proposto por Rosenblatt (1962)



$$F(u_1, u_2, u_3 \dots u_n) = \sum_n^1 w_i u_i \quad (2.14)$$

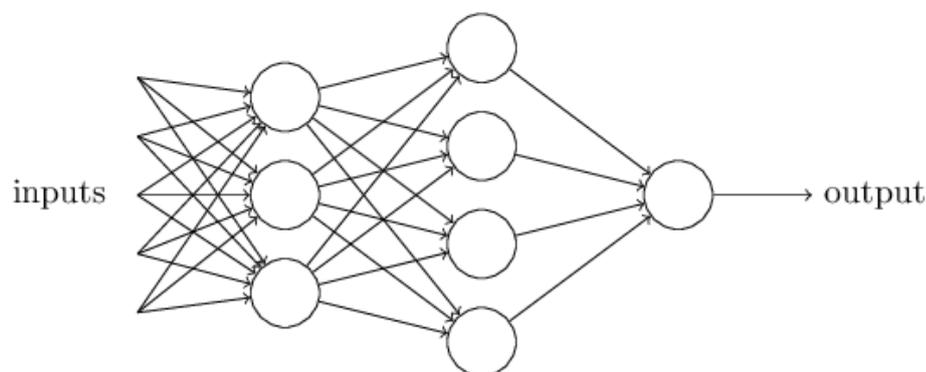
Conforme dito em Nielsen (2015) o modelo visual criado por Frank Rosenblatt na década de 60 que se observa na Figura 13, possui nos dias atuais, uma denominação ou nomenclatura muito conhecida na literatura e estudos sobre Redes Neurais, é o que denomina-se de Perceptron, sua estrutura recebe como entrada 'N' valores, e tem como resposta uma única saída binária, tem-se também a utilização de pesos, que são números reais que enfatizam a importância de cada entrada na única saída, com a saída binária de 0 ou 1, o cálculo do Perceptron realizado através de uma soma ponderada que no final é comparada a um valor limiar que representa o limite entre a presença ou ausência de sinal naquela sinapse, sua representação está descrita na fórmula a seguir:

$$saída = \begin{cases} 0 & se \sum_j w_j x_j \leq limiar \\ 1 & se \sum_j w_j x_j > limiar \end{cases} \quad (2.15)$$

Além do exemplo anterior de acordo com Nielsen (2015), é possível exemplificar com modelos um pouco mais complexos que envolve um número maior de camadas, precisamente a seguir na Figura 14, uma rede com 4 camadas, sendo 1 camada de entrada, uma de saída e 2 camadas ocultas, tais redes são chamadas de Redes Multicamadas, que em inglês chama-se *Multilayer perceptrons*.

Uma maior eficiência nos resultados é alcançada de acordo com o número de amostras a serem analisadas, o que corresponde ao número de entradas da rede adotada como em Sun et al. (2017), em resumo, as Redes Neurais mostram-se possuir importantes

Figura 14 – Modelo multicamada de uma Rede Neural Artificial de 4 camadas.



Fonte: (NIELSEN, 2015)

contribuições no avanço da tecnologia pelo mundo, no processamento de linguagens naturais Collobert e Weston (2008) ou reconhecimento de imagens Simonyan K. e Zisserman (2014), ajudando vários campos de atividades humanas a se modernizarem e criados novos tipos de trabalhos e avanços científicos.

2.5.1 Aprendizagem Profunda

Sendo uma sub-área da Aprendizagem de Máquina, a Aprendizagem Profunda tem como objetivo a criação de métodos que simulam as atividades ocorridas no sistema cognitivo humano, como exemplo disso, pode-se citar o reconhecimento visual, de voz e o processamento de linguagem natural. A nomenclatura, Aprendizado Profundo, tem como conotação o problema de treinamento de Redes Neurais Artificiais, que tenham que aprender características de maneira hierárquica, de modo que esses padrões de níveis mais altos sejam a combinação de padrões de níveis mais baixos, tem-se a resolução de vários problemas utilizando uma única camada intermediária Bezerra (2016).

Hoje a Aprendizagem Profunda possui diversas aplicações na vida das pessoas, seja para catalogar conteúdos em uma rede social, como para indicar produtos para determinados tipos de usuários, ou mesmo em efeitos nas câmeras dos celulares inteligentes, cada vez mais, esses métodos utilizam as técnicas de Aprendizagem Profunda, incentivando cada vez mais o financiamento privado na pesquisa e mercado mundial Hinton, LeCun e Bengio (2015). Essa revolução tem como fator principal, os milhares de dados de textos e imagens produzidos e disponibilizados na internet ou por grandes organizações, e também o aumento da capacidade dos computadores modernos, que possibilitou a diminuição no tempo de processamento e aumento na quantidade de armazenamento.

Os conceitos básicos de interligação com a Aprendizagem Profunda, o Processamento de imagens com o foco em aprendizado supervisionado e não-supervisionado, conceitos de

classificação, Redes Neurais MLP, filtragens e convoluções Shalev-Shwartz e Ben-David (2014). Os métodos que usam a Aprendizagem Profunda, tem como finalidade descobrir regras através de atributos, utilizando um grupo de parâmetros como exemplo e métodos que o guiam para o aprendizado do modelo, no final desse processo têm-se como resultado, métodos que são capazes de receber dados na sua forma mais bruta e devolver como resposta ou saída, a representação de um modelo de resposta para o problema em questão Ponti e Costa (2018).

De acordo com Bezerra (2016), as classes de Rede Neurais de aprendizagem profunda mais popular na atualidade são as seguintes: Autocodificadoras, Redes Convolucionais, Redes Recorrentes. Tais classes estão no topo de utilização para a criação de projetos de aprendizagem profunda.

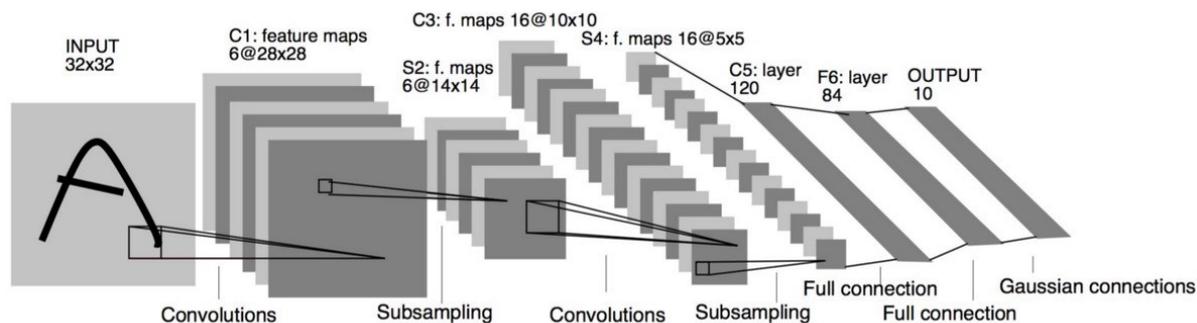
2.5.2 Redes Neurais Convolutivas

Com uma característica multicamada e a necessidade de ter uma base extensa para aumentar a efetividade dos seus resultados, as Redes Neurais acompanharam a evolução tecnológica de softwares e hardware, pois no passado era inviável o uso das mesmas, pela sua característica computacional, ou seja, no seu treinamento, quanto maior o número de camadas e de perceptrons, mais alto é o custo com recursos, é possível citar a utilização de imagens para este projeto em específico, onde é realizado o processamento de várias imagens em um curto intervalo de tempo e cada imagem gera um custo bastante elevado pela presença de um número alto de dados a serem computados e seu acompanhamento o tempo real de resposta.

O surgimento das Redes Neurais Convolucionais, tem seu início em meados dos anos 1988 como descrito em LeCun et al. (1998), tem-se a criação de uma arquitetura cujo nome foi batizado de LetNet, como nos anos 80 não se previa de recursos modernos para facilitar o processamento de figuras complexas, imagens mais simples eram utilizados para o estudo e exemplificação da convolução, os procedimentos utilizados envolviam reconhecer caracteres, dígitos, etc conforme Rocha (2015). Na Figura 15, é apresentado um exemplo do procedimento de funcionamento da Rede Neural Convolutiva, tal Rede é utilizada no presente trabalho, pois suas principais aplicabilidades, estão diretamente envolvidos no processamento de imagens, na Figura 16 observa-se um exemplo da inicialização do processo de extração de características e sua finalização na Figura 18, como em Rocha (2015) tais processamentos são descritos a seguir:

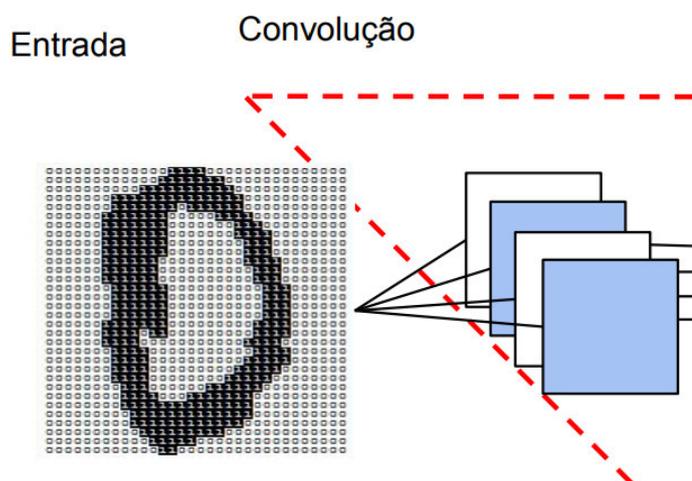
- Convolução: É a própria retirada de características da imagem analisada que pode ser observado na Figura 16, ou seja, é como se estivesse se refinando e diminuindo o grupo de imagem a serem analisadas.

Figura 15 – Representação a rede neural LeNet



Fonte: (LECUN et al., 1998)

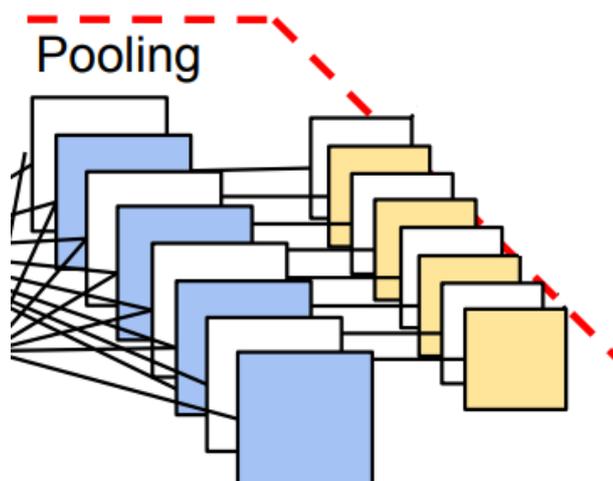
Figura 16 – Imagem retirada da representação da rede LeNet, destacando sua convolução



Fonte: (VARGAS; PAES; VASCONCELOS, 2016)

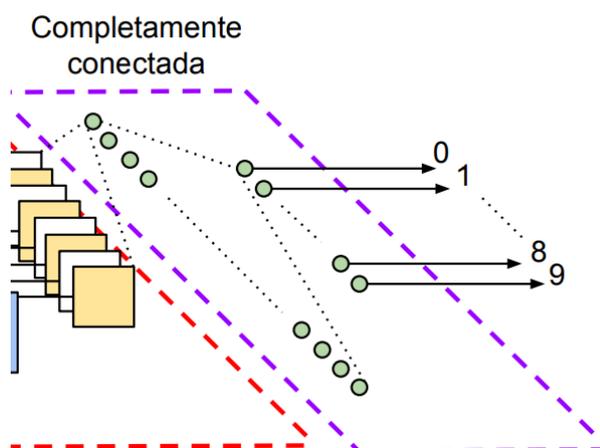
- Não linearidade: sua finalidade é colocar a não linearidade na rede com cálculos matemáticos.
- *Pooling*: na Figura 17 se observa o processo como uma das etapas da convolução. sua finalidade é reduzir a amostra de análise, mantendo as informações com maior relevância para serem analisadas.
- Camada completamente conectada: nada mais é que a Classificação, usa-se a rede MLP com um método para ativar a camada de saída.

Isolando-se uma parte da rede LetNet desenvolvida por LeCun et al. (1998), observam-se algumas características, a primeira é a simplicidade de entrada discutida anteriormente, observa-se apenas dois tipos de entrada, branco ou preto, sendo representadas pelo 0 ou 1, essa simplicidade se reflete no início do processamento, tendo apenas 2

Figura 17 – Imagem retirada da representação da arquitetura LeNet, destacando o *Pooling*

Fonte: (VARGAS; PAES; VASCONCELOS, 2016)

Figura 18 – Imagem da representação da camada completamente conectada



Fonte: (VARGAS; PAES; VASCONCELOS, 2016)

possibilidades de entradas possíveis, assim também a padronização estacionária, ou seja, a uniformidade de resultados na análise de qualquer parte da imagem, divide-se a imagem em pequenos quadrados chamados de pixel, e organiza-se uma série de padrões extraídas dos mesmos, tais características serão as entradas da rede desenvolvida neste trabalho.

2.5.3 ResNet

Trata-se de uma arquitetura de rede convolucional proposta por He K. (2016), o nome ResNet vem da abreviação de Residual Networks, esta é uma arquitetura muito utilizada na literatura em tarefas de visão computacional, possuindo um bom desempenho

Figura 19 – Imagem representando a da representação da rede neural ResNet

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
		3×3 max pool, stride 2				
conv2.x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3.x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4.x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5.x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Fonte: He K. (2016)

quando se trata de classificação de imagens.

Vencedora do ILSVRC (*ImageNet Large Scale Visual Recognition Challenge*) realizado em 2015, que se trata de um concurso de reconhecimento visual de grande escala ImageNet, onde diversas equipes competem para propor arquiteturas de redes capazes de reconhecer grandes números de imagens e categoriza-las.

Estes desafios são de fundamental importância para o desenvolvimento da área, e a Arquitetura proposta na ResNet abriu uma imensa gama de possibilidades com redes neurais extremamente profundas, visto que esta possui 150 camadas como pode ser observado na Figura 19, antes dela, redes tão profundas tinham problemas com o desaparecimento de gradientes, pois como o gradiente é propagado de volta para as camadas anteriores, a multiplicação repetida pode tornar o gradiente infinitamente mínimo, ou seja, a medida que a rede aumenta em profundidade, seu desempenho torna-se saturado ou se degrada rapidamente.

Baseada na rede VGG desenvolvida por Simonyan K. e Zisserman (2014), a ResNet possui suas camadas convolutivas principais com filtros de 3x3 pixel, o bloco residual possui também 3x3 camadas convolucionais com o mesmo numero de canais de saída. Cada camada é seguida por uma camada de normalização e outra de ativação do tipo ReLU.

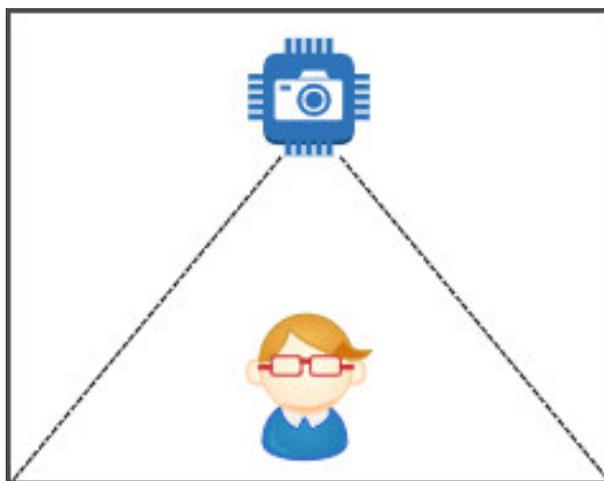
2.6 Detecção e Reconhecimento Facial

De acordo com Lopes, Filho e No (2005), antes de realizar o reconhecimento facial propriamente dito em uma imagem ou em um conjunto de imagens, deve-se aplicar os métodos de detecção facial, tais métodos consistem em delimitar uma área da imagem onde se têm um rosto presente, tal aspecto é de suma importância no reconhecimento facial, pois delimita uma área menor para a extração de características, com isso poupa-se muito trabalho de busca, pois muitos algoritmos buscam tais características em toda a imagem.

Ainda conforme Lopes, Filho e No (2005), alguns problemas podem surgir ao longo do uso das técnicas de detecção facial, pois se trabalha com imagens que podem sofrer alterações, como no caso de vídeos, ou podem conter elementos fixos que dificultem a análise da imagem (fotos), alguns elementos que impedem a detecção facial está a seguir:

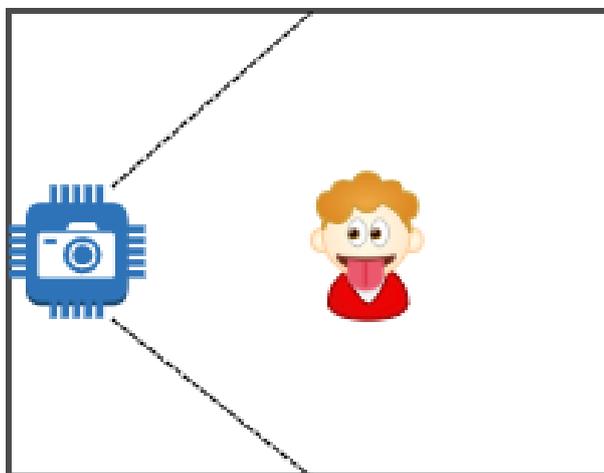
- Pose: a Figura 20 ilustra a posição da câmera ou da pessoa filmada, podem variar de acordo com os ângulos posicionados, dificultado-se o registro de um rosto na imagem.

Figura 20 – Imagem ilustrativa exemplificando a captura posicional desfavorável para a detecção facial.



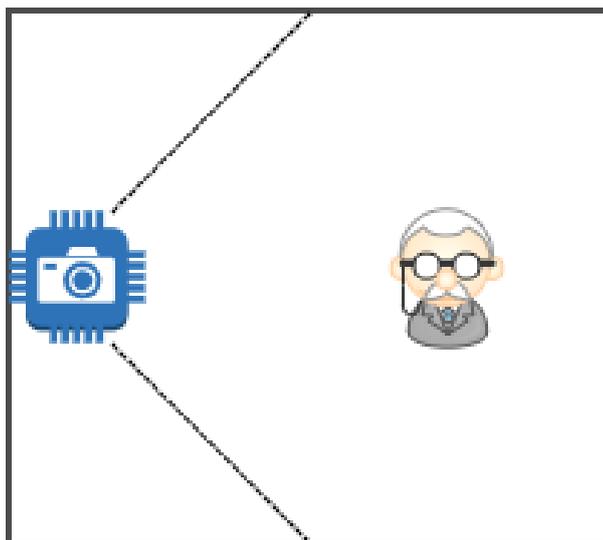
- Expressão Facial: a Figura 21 apresenta a expressão da face na imagem, pode influenciar na detecção do rosto, pois pode-se fugir do padrão de rosto conhecido pelo algoritmo em questão.

Figura 21 – Imagem ilustrativa exemplificando a captura facial fora do padrão, pode dificultar na detecção de faces.



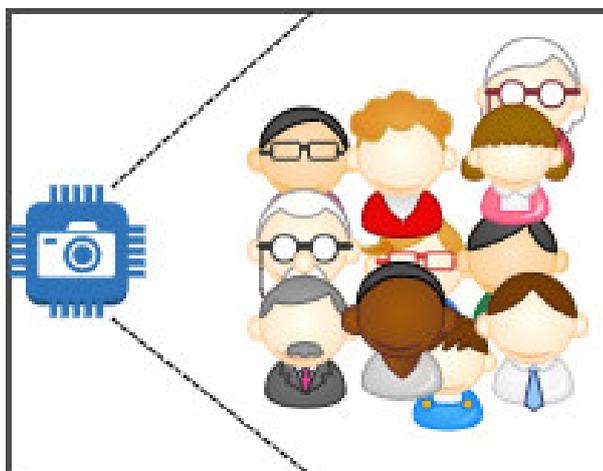
- Presença de elementos estruturais: a Figura 22 representa alguns elementos nos rostos das pessoas, podem também dificultar na detecção de face em uma imagem, tais elementos podem ser óculos, barbas, cabelo sobre o rosto etc, fatores de iluminação ou uma distorção na imagem.

Figura 22 – Imagem ilustrativa exemplificando como elementos presentes na face da pessoa, trazendo uma maior dificuldade na detecção facial.



- Ocultação: a Figura 23 ilustra imagens adquiridas de ambientes de pouco controle, podem apresentar algumas anomalias na presença de faces, uma que se destacam é a sobreposição de faces por outras faces ou por objetos presentes.

Figura 23 – Imagem ilustrativa exemplificando como a sobreposição pode dificultar a detecção de faces em uma imagem.



A seguir abordar-se alguns métodos de detecção facial e alguns exemplos apresentados em Lopes, Filho e No (2005):

- Métodos baseados em conhecimento: São métodos baseados em regras bases estabelecidas, que tem como base o conhecimento antecipado sobre determinado problema, ou seja, uma base que já conhece aquilo que representa uma face, tal padrão é possível, pois existem características fixas que representam uma face, como, por exemplo, organização dos elementos presente em uma face, que na maioria das vezes conta com a presença de dois olhos em paralelo, um nariz mais centralizado na direção vertical juntamente com uma boca no mesmo sentido direcional. Algumas desvantagens desse tipo de algoritmo, está relacionada está na construção do seu conjunto de regras para detectar um rosto em uma nova entrada, pois caso de presença de regras muitos gerais, pode-se facilmente obter uma grande taxa de falsos positivos, ou seja, elementos que pode ser taxados como uma face, e caso as regras forem muito específicas, têm-se a falta de efetividade na detecção de faces em uma imagem.

Como exemplo tem-se a técnica que apresentada em Yang, Kriegman e Ahuja (2002), onde se constrói um grupo de padrões de hierárquico para a detecção de faces, consiste num método onde tenta-se encontrar elementos em uma imagem que aparentam ser uma face, mas que sofrerá um novo conjunto de varredura para confirmar se tal elemento é de fato um rosto, a cada nível de filtro do algoritmo, pode-se tentar encontrar mais detalhes que caracterizam um rosto, afim de confirmar que se detecta de forma correta uma face.

Na Figura 24, é apresentado um exemplo onde se diminui a resolução da imagem para obter uniformidade entre pixels próximos, repare que quanto mais se aplica a média no número de pixels presentes na imagem, menor é a sua resolução, e conjuntos de pixels semelhantes vão se agrupando. É possível então deduzir, que o rosto está contido na região central que possui maior luminosidade na imagem, vale destacar que isso pode aplicado em vários pontos da mesma imagem.

Figura 24 – Imagem com várias resoluções a) $n=1$, b) $n=4$, c) $n=8$, d) $n=16$, os pixels são substituídos pela média da intensidade de valor de cada células de pixels



Fonte: (LOPES; FILHO; NO, 2005)

- Métodos baseados em características invariantes:

Conforme Lopes, Filho e No (2005), diferente do método anterior, aqui têm-se uma inspiração na maneira como os seres humanos detectam faces, independente do ponto de vista de quem observa. Como diz Yang, Kriegman e Ahuja (2002), algumas

características podem influenciar negativamente as condições para a detecção, como, por exemplo, iluminações e ruídos na imagem, não influenciando assim, outras características como as cores de pele ou textura da imagem, pois são características que muitas vezes são retiradas para facilitar o processo de detecção. Com relação à cor, em muitos casos, tal imagem sofre o processo de transformação em escala de cinza para facilitar a acusação de rosto em uma imagem, pois as imagens coloridas se despõe de detalhes acima da aparência para detectar um objeto. Ainda de acordo com Yang, Kriegman e Ahuja (2002), as diferentes cores do ser humano, seja branca, parda, negra, amarela etc, não influenciam na eficiência da detecção, mas sim, a formam um cluster no conjunto de cores, utiliza-se a modelagem baseada na medida Gaussiana.

- Métodos baseados em *templates*

Conforme Lopes, Filho e No (2005), é o modelo de detecção baseados em comparação com padrões geométricos, primeiramente, determinam-se as formas geométricas básicas que formam uma determinada imagem, logo após, para delimitar uma face em uma determinada imagem de entrada, procura-se a melhor similaridade entre esta imagem e seu molde geométrico, formado por quadrados, triângulos, círculos, elipses etc, geralmente utilizam-se algoritmos genéticos para encontrar tais formas geométricas nas imagens.

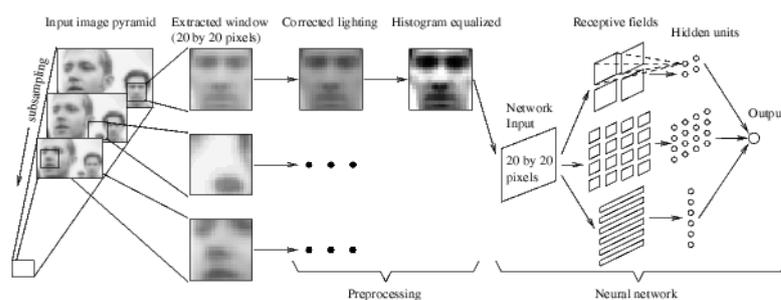
- Métodos baseados na aparência

Como dito por Lopes, Filho e No (2005), são métodos que não possuem conhecimentos prévios sobre o objeto a ser analisado para a detecção, é aqui que se tem maior nitidez do de aprendizado e treinamento de características, pois as informações de treinamento para que haja detecção, retira-se do grupo de imagens de entrada de treinamento, sem a intervenção direta de uma pessoa nessa seleção de características, criando assim para a detecção facial, a representação interna daquilo que seja um rosto.

Uma das principais técnicas utilizadas no método baseado em aparência, são as Redes Neurais Artificiais, pois é um modelo que demonstra ter grande efetividade em encontrar padrões e no agrupamento de características, utilização a desses conceitos para realizar inúmeras tarefas, e uma delas é a detecção facial, onde se mapearam regiões de uma imagem, com o intuito de encontrar uma característica que represente uma face, a função da rede é de determinar aquilo que é uma face e aquilo que não é. Tem-se o exemplo dessa utilização presente em Rowley, Baluja e Kanade (1998), que utiliza um filtro que percorre uma imagem de dimensão 20x20, e busca-se encontrar faces em cada região da imagem analisada, logo após apontar uma possível face presente numa determinada parte da imagem, usam-se espécies de filtros para

melhorar a eficiência de detecção, alguns deles são a equalização de histograma como visto em Filho e Neto (1999) e a normalização da luz como acompanhado em Amaral et al. (2016). Caso a face tenha um tamanho maior do que o da janela analisada, analisa-se subamostragens que tenham potencial de conter uma face, ou seja, divide-se para conquistar a efetividade da detecção. Cada parte de amostra da imagem depois de sofrer o pré-processamento, que consiste em transformar a escala de cores e determinar valores de acordo com um conjunto cardinal que resulta em uma matriz, é enviada para a rede neural, a rede terá como saída a confirmação ou negação da presença de uma face na amostra analisada, utiliza-se na rede várias arquiteturas de análise de dados, com o intuito de melhorar a confiabilidade dos resultados, no exemplo da Figura 25, onde se apresenta, cada camada analisa uma parte da imagem de entrada, como forma de encontrar alguma característica que possa ser uma face presente.

Figura 25 – Passos do algoritmo de detecção baseados em uma Rede Neural Artificial



Fonte: (LOPES; FILHO; NO, 2005)

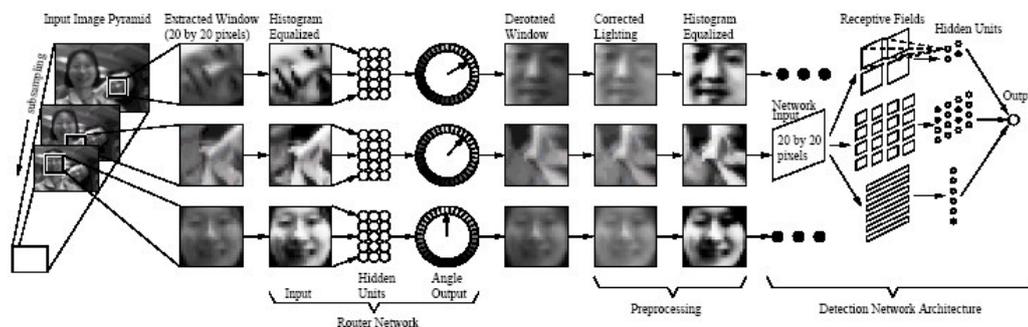
Assim Rowley, Baluja e Kanade (1998) realizou um outro teste que utilizou 78 neurônios na camada mediana e 4357 conexões entre os perceptrons, teve uma taxa de acerto de detecção facial em 92,7%, com imagens de teste retiradas da internet. Entendendo a limitação do seu algoritmo em relação aos elementos que dificultam a detecção de faces, como a iluminação e a rotação de um rosto, Rowley, Baluja e Kanade (1998) aplicou cálculos adicionais na sua técnica, para tratar imagens com tais dificuldades, utilizou-se assim, uma segunda rede neural chamada de roteadora, que possui a finalidade de determinar a perspectiva de rotação. A rede roteadora nesse exemplo possui 400 entradas, 15 unidades na parte central e 36 saídas, onde cada saída presente, representa o equivalente a $i \cdot 10$ graus, ou seja, uma resposta positiva ou negativa para a rotação e cada resposta positiva irá multiplicar por 10 para determinar o ângulo rotacional da face, cada valor responde ao cálculo $\cos(\Theta - i \cdot 10)$, onde θ , corresponde ao ângulo da face apresentada. Como processo final, na equação a seguir, têm-se a soma ponderada do cosseno e seno

das saídas da rede utilizada neste trabalho multiplicadas por 10, determinando assim, o ângulo da face em determinada imagem.

$$\left(\sum_{i=0}^{35} output_i * \cos(i * 10^\circ), \sum_{i=0}^{35} output_i * \sin(i * 10^\circ) \right) \quad (2.16)$$

Logo em seguida, supondo o sucesso na detecção da face e na direção angular que essa face se encontra, têm-se a rotação da imagem no seu sentido contrário, deixando-a em sentido frontal, na Figura 26, é representado o processo de rotação e logo em seguida, o mesmo processo é exibido na Figura 25.

Figura 26 – Representação do algoritmo de rotação facial



Fonte: (LOPES; FILHO; NO, 2005)

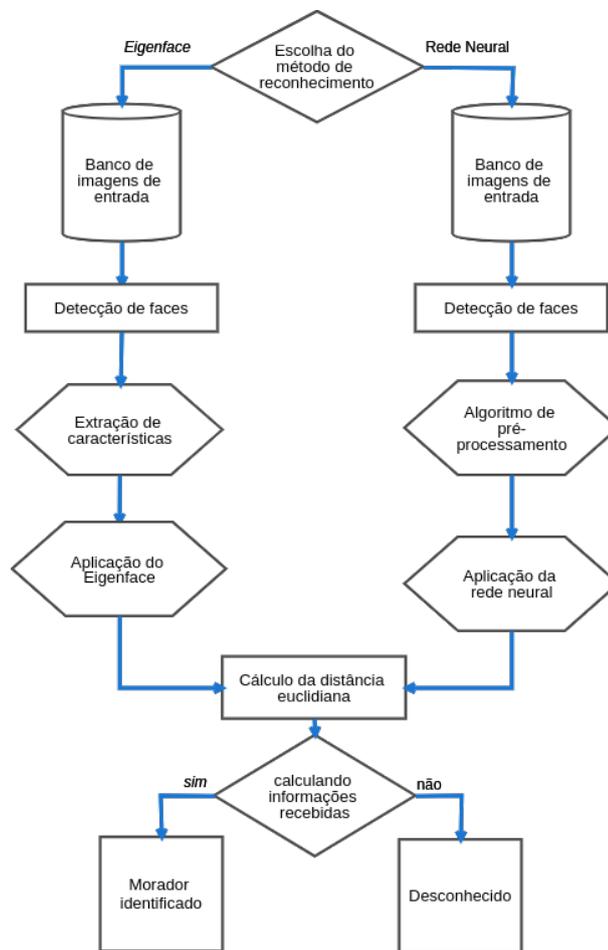
Conforme Braga et al. (2013), tem-se a última etapa para avaliar as características dos dados de entrada, utilizando modelos matemáticos, faz-se a comparação com os dados de faces pré-registradas no banco de dados. Ao contrário do sistema de detecção, onde se tem como saída geralmente apenas duas possibilidades, a de face e não-face, o sistema de reconhecimento facial tem como característica, inúmeras possibilidades de saída, pelo fator numérico de classificação de dados. Pode-se aplicar o reconhecimento de duas formas possíveis, na verificação ou na identificação, na primeira o sistema retorna como positiva ou negativa a verificação de autenticidade de um indivíduo, ou seja, libera os recursos pedidos apenas se a pessoa for aquela que diz ser, esse tipo de comparativo, é do tipo 1:1, ou seja, uma face de entrada é comparada a uma face pré-cadastrada no banco e associada a outros dados. Na identificação tem-se uma comparação 1:N, pois se cria uma busca por um rosto que nos retorne um limiar satisfatório de semelhança.

Ainda segundo Braga et al. (2013), os métodos mais utilizados no reconhecimento facial, são aqueles que usam o conceito da distância do vizinho mais próximo, que normalmente usam a norma Euclidiana, para determinar as proximidades ou distância de características entre as diferentes imagens coletadas, métodos mais avançados usam Redes Neurais e Densidade de probabilidades.

3 METODOLOGIA

Neste capítulo é descrita a metodologia adotada para a realização deste trabalho, no qual foi proposto um sistema de biometria facial utilizando técnicas de visão computacional para detecção e reconhecimento de faces, para tanto utilizou-se, *Eigenface* com sua base PCA e Redes Neurais Convolutivas Juntamente com o método HOG, todos discutidos todos discutidos no Capítulo 2, o diagrama da metodologia está apresentado na Figura 27.

Figura 27 – Etapas da metodologia proposta



Os métodos propostos envolvem a aquisição de dados para a Biometria Facial de pessoas que possuem ou não autorização para adentrar espaços particulares, ou residenciais. Para realizar esse processo têm-se como opções dois sistemas de reconhecimento facial disponível, o primeiro deles o algoritmo de *Eigenface*, conforme discutido na Seção 2.4.1, o segundo em baseado nas redes neurais e o algoritmo HOG conforme se vê em 2.3.1 e 2.5.

3.1 Escolha do método de reconhecimento

Os métodos de reconhecimento escolhidos foram os *Eigenfaces* como se observa em 2.4.1 e as redes neurais definidas em 2.5. Por ser um algoritmo baseado na álgebra linear e estatística, o *Eigenfaces* possui uma menor exigência no seu poder de processamento, contudo, sua eficiência é menor, para sistemas biométricos de face, como possuímos um maior controle no ambiente de coleta dos dados, diminuimos assim essa falha de eficiência, enquanto isso, o algoritmo baseado em redes neurais, exigem maior poder de processamento, contudo detém a ter uma maior efetividade no reconhecimento de pessoas.

3.2 Banco de imagens de entrada

O algoritmo baseado nas redes neurais, possuem uma base pré-treinada para detecção facial de forma empírica na biblioteca Dlib, enquanto que para o reconhecimento, o algoritmo de rede neural tem sua identificação no próprio nome da foto utilizada, armazenada em uma pasta local acessada pelo algoritmo de reconhecimento.

No *Eigenface* tem-se duas pastas de base de dados locais de imagens, na primeira está presente as fotos de pré-processamento de dados, utilizadas para a extração de características, na segunda tem-se a base de características extraídas expressas em um arquivo XML.

3.3 Detecção de faces

Como forma de realizar a detecção para o treinamento da base, utiliza-se o algoritmo PCA, que é uma das bases do *Eigenface*:

Para as redes neurais, utilizam-se funções que estão contidas na Biblioteca Dlib, pertencente à linguagem *Python*, que tem funcionalidade ligada ao algoritmo HOG, ao qual utiliza em uma das etapas do seu processo o classificador SVM referenciado em 2.3.

3.4 Extração de características

A extração é realizada de acordo com o algoritmo utilizado, ambas as formas geram pastas com amostras coletadas no processo de detecção facial, realizada por uma câmera de captura de imagem.

3.4.1 Algoritmo de pré-processamento

Logo após computar as coordenadas das faces encontradas, divide as imagens da maior à menor importância, julga-se como de maior importância, as imagens que possuem

como características, um fundo com características pouco importantes, focando em retirar as faces desses pontos, para ambos os pré-processamentos utiliza-se a biblioteca *OpenCV*.

3.4.2 Extração no *Eigenface*

Método definido em 2.4.1, que busca um conjunto de informações contidas nas pastas de extração, que ignoram as formas geométricas dos rostos nas fotos, utilizando a completa representatividade da face, até porque, tais imagens sofrem uma redução na sua dimensão e complexidade de cores, basicamente é um algoritmo PCA otimizado com redução de covariância dos seus vetores e valores, no ato da classificação, o *Eigenface* utiliza o KNN, que foi definido em 2.4.

3.4.3 Extração na Rede Neural

No presente projeto, foi utilizado um algoritmo convolucional originado da biblioteca *OpenCV* com uma base já treinada e que possui o domínio público para uso. A arquitetura composta é a ResNet-29 discutida em King (2017), que depois de algumas adaptações de camadas que foram removidas e a redução nos números de filtros utilizados, tem como origem a arquitetura de ResNet-34 criada por He et al. (2016). Conforme King (2017) o modelo pré-treinado, utilizou 3 milhões de faces 7485 pessoas distintas, com um resultado de acurácia de 99,38%, com imagens que foram disponíveis em *Labeled Faces in the Wild*, que é uma base destinada para os estudos no campo do reconhecimento facial.

Para geração do campo de informações pelo modelo de varredura de imagem, utilizam-se funções da biblioteca Dlib, tais modelos são comparados aos melhores algoritmos para a detecção e reconhecimento facial, possuem características de interação vetorial com mais de 128 dimensões, para cada uma das faces presentes na imagem, as faces encontradas na imagem geram várias outras imagens com o registro de faces com seus vetores e mapas de informações de cara indivíduo detectado.

3.5 Cálculo da distância euclidiana

Como se pode ver em Linden (2009), uma maneira de comparação entre imagens, para encontrar semelhança entre a entrada e os dados prévios cadastrados, é medir a distância euclidiana entre o agrupamento característico entre às duas faces, na equação a seguir tem-se o cálculo da euclidiana, tais cálculos são embutidos tanto na utilização da *Eigenfaces* quanto da biblioteca *Face Recognition*.

$$d(p, q) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 \dots (q_n - p_n)^2} = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (3.1)$$

3.6 Calculando informações recebidas

Tem-se como resultado, dados de verificação e de identificação, ou seja, além de nos informar se dada pessoa pertence ao nosso conjunto de dados conhecidos, também tem-se a informação sobre a identidade da pessoa conhecida, caso a pessoa não estiver previamente cadastrada, no ato da biometria facial ela será classificada como desconhecida.

4 RESULTADOS

Considerando tudo que já foi exposto nos tópicos anteriores, os resultados envolvem tanto a utilização das Redes Neurais como o método *Eigenface* para a Biometria Facial através do reconhecimento facial. Para facilitar a entrada de dados, foi criada uma simples interface gráfica representada na Figura 28, que recebe os dados para a classificação e o reconhecimento biométrico, com relação à rede neural, optou-se por colocar apenas o botão de reconhecimento e colocar os dados de entrada direto na pasta fonte, utilizou-se a biblioteca *Tkinter*, biblioteca de interface gráfica nativa da linguagem *Python*.

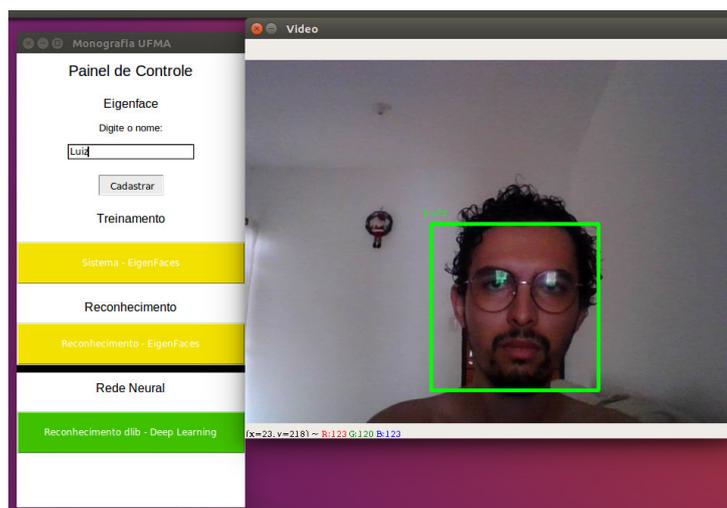
Figura 28 – Painel principal do sistema



4.1 Resultado com o *Eigenface*

Na Figura 29 a entrada do nome da pessoa, o algoritmo cria uma pasta identificada com o nome da pessoa de entrada, ao clicar em cadastrar uma série de fotos são tiradas da pessoa realizando os processos de refinamento diminuindo a dimensão das fotos e colocando-a em escala de cinza.

Figura 29 – Detecção Facial para criar a base de treinamento



4.1.1 Resultado da detecção

Na Figura 30, tem-se uma amostra das imagens após o processo de redimensionamento e redução da complexidade dos canais de cores.

Figura 30 – Painel principal do sistema



Na Figura 31 tem-se o treinamento, cria-se um arquivo em XML, com as características da pessoa para realizar a comparação com a nova entrada e determinar o seu reconhecimento, todos os rostos cadastrados são salvos no mesmo arquivo.

Figura 31 – Arquivo XML com a base treinada

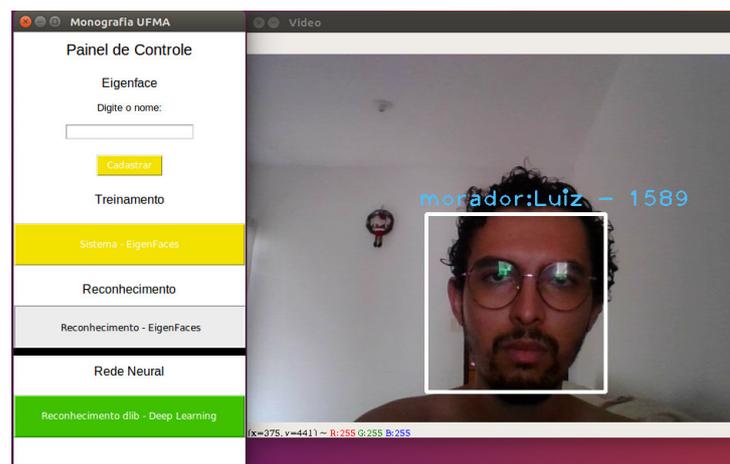
```

<?xml version="1.0"?>
<opencv_storage>
<opencv_eigenfaces>
<threshold>1.7976931348623157e+308</threshold>
<num_components>300</num_components>
<mean_type_id="opencv-matrix">
<rows>1</rows>
<cols>10304</cols>
<dt>d</dt>
<data>
1.3522000000000000e+02 1.3369333333333333e+02
1.3243333333333333e+02 1.2952666666666667e+02
1.2844333333333333e+02 1.2410666666666667e+02
1.2208666666666667e+02 1.1976666666666668e+02
1.1755333333333334e+02 1.1526666666666668e+02
1.1342333333333335e+02 1.1159666666666668e+02
1.0936333333333334e+02 1.0625333333333334e+02
1.0288333333333334e+02 9.842666666666667e+01
9.3310000000000002e+01 8.801666666666666e+01
8.1463333333333338e+01 7.3986666666666665e+01
6.5163333333333341e+01 5.6400000000000006e+01
4.9100000000000001e+01 3.8896666666666668e+01
3.2726666666666667e+01 2.9020000000000003e+01
2.8633333333333336e+01 2.7403333333333336e+01
2.4236666666666668e+01 2.1236666666666668e+01
1.9210000000000001e+01 1.9150000000000002e+01
1.8830000000000002e+01 1.8446666666666669e+01
1.8933333333333334e+01 1.9866666666666667e+01
2.0770000000000000e+01 2.1773333333333333e+01
2.3380000000000003e+01 2.5460000000000001e+01
2.7983333333333334e+01 3.1036666666666669e+01
3.4253333333333337e+01 3.6850000000000001e+01
3.9166666666666671e+01 4.0780000000000001e+01
4.240666666666666e+01 4.3646666666666668e+01
4.4666666666666671e+01 4.5246666666666670e+01
4.578666666666669e+01 4.5689999999999999e+01
4.4690000000000005e+01 4.3810000000000002e+01
4.2896666666666668e+01 4.1596666666666671e+01
4.0413333333333334e+01 3.9486666666666672e+01

```

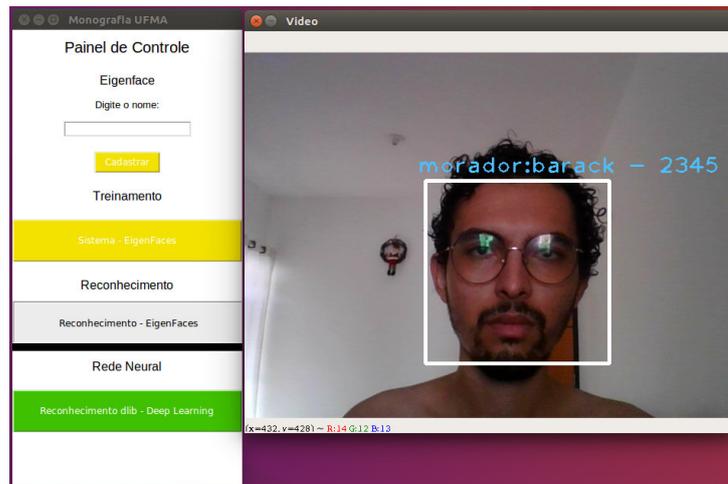
4.1.2 Resultado de reconhecimento

Na Figura 32 mostra-se a pós fase de treinamento, obteve-se a identificação correta da pessoa cadastrada, levando em consideração o cadastro de apenas uma pessoa no sistema.

Figura 32 – Resultado do reconhecimento usando o algoritmo *Eigenface*

Alguns erros ocorreram no reconhecimento, a identificação de errada de pessoas ocorreu em alguns *frames* de captura de vídeo, principalmente quando o rosto da pessoa não estava alinhado com a imagem da câmera, na Figura 33, se observa um exemplo de erro de identificação, há um falso positivo presente, ou seja, detecção da face, com a identificação errada.

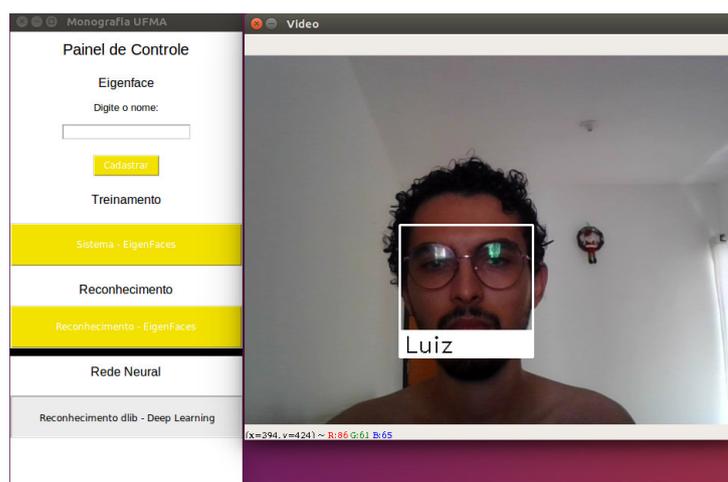
Figura 33 – Resultado mostrando o erro de reconhecimento facial, rotulando outra pessoa



4.2 Resultado com a Rede Neural

Para a utilização do modelo em redes neurais, utilizou-se a biblioteca *Face_recognition*, importa-se a biblioteca para dentro do projeto e cria-se uma ligação entre a interface e os métodos da mesma, coloca-se como entrada apenas uma imagem para reconhecimento e coloca-se no próprio algoritmo o rótulo daquela imagem, após vários testes, os únicos erros apresentados estão relacionados a rotação do rosto, mas em um ambiente controlado de reconhecimento facial, a rede neural não apresentou falhas, pode-se ver a representação na Figura 34.

Figura 34 – Resultado do reconhecimento através da Rede Neural



5 CONCLUSÃO

Como objetivo do trabalho apresentado, foi desenvolvido métodos computacionais que podem ser utilizados em sistema de Biometria Facial, aplicados em diversos tipos de ambientes, para a detecção e reconhecimento do sistema *Eigenface* foram utilizados os algoritmos *Eigenface* e PCA, para a detecção e reconhecimento da Rede Neural, utilizou-se o método HOG juntamente com a Rede Neural Convolutacional ResNet.

Discutiu-se, como a tecnologia pode ajudar em tarefas onde o ser humano apresenta certos limites, em especial o processo de Biometria Facial através das técnicas de Detecção e Reconhecimento Facial, pode-se observar também, a variedade de algoritmos existentes para fazer o processo de reconhecimento, ao longo dos anos a eficiência desses algoritmos obteve sucesso na diminuição de erros, para trazer maior confiabilidade e menos prejuízo no uso de tais técnicas.

Foram expostos aqui dois principais algoritmos que refletem esses anos de pesquisa e avanço nos sistemas de reconhecimento facial, o primeiro deles foi o *Eigenface*, que mostrou-se exigir menor poder de processamento de uma máquina, mas obteve erros acentuados no processo de reconhecimento, e observam-se também as redes neurais, estruturas de inteligência artificial que vem ganhando destaque nos últimos anos, por contam da sua superioridade em eficiência oriunda dos trabalhos feitos com aprendizagem profunda, que apesar de exigir um maior poder de processamento de uma máquina, entrega os resultados esperados. A escolha do *Eigenface* se deu pelo fator de simplicidade envolvendo sua estrutura de processamento, já a escolha da Rede Neural, deu-se pelo fator de difusão dos estudos envolvendo a linguagem *Python* e as bibliotecas *Dlib* e *Face_recognition*.

5.1 Trabalhos Futuros

No projeto apresentado, existem ainda diversos teste e modificações a serem acrescentadas. Na parte da interface gráfica, pode-se construir um módulo mais eficiente para as entradas de dados para a rede neural artificial.

Como forma de melhorar a captura biométrica, pode-se utilizar os algoritmos convolucionais juntamente com o classificador KNN, pois se apresentou muito eficiente quando pesquisado em trabalhos semelhantes.

A utilização de *multithread* poderia ser empregada para uma polarização de captura, processamento e reconhecimento do mesmo sistema em diversos setores diferentes, como, por exemplo, estádios de futebol, ou um condomínio residencial com mais de uma entrada.

Referências

- AGGARWAL, C. C.; HINNEBURG, A.; KEIM, D. A. On the surprising behavior of distance metrics in high dimensional space. In: SPRINGER. *International conference on database theory*. [S.l.], 2001. p. 420–434. Citado na página 29.
- AHA, D. W.; KIBLER, D.; ALBERT, M. K. Instance-based learning algorithms. *Machine learning*, Springer, v. 6, n. 1, p. 37–66, 1991. Citado 2 vezes nas páginas 21 e 27.
- ALMEIDA, A. P. G. S. d. Reconhecimento de ações em vídeo utilizando descritores de pontos de interesse espaço-temporais (stips). 2017. Citado na página 14.
- ALPAYDIN, E. Introduction to machine learning. 2004. *Cover, Copyright Page, Table of Contents for*, p. 1–327, 2004. Citado na página 27.
- AMARAL, V. do; FÍGARO-GARCIA, C.; GATTAS, G. J. F.; THOMAZ, C. E. Normalização espacial de imagens frontais de face em ambientes controlados e não-controlados. *FaSCi-Tech*, v. 1, n. 1, 2016. Citado na página 44.
- ASSUNÇÃO, M. F. A. *Guia do hacker brasileiro*. [S.l.]: Marcos Flávio Araújo Assunção, 2002. Citado na página 16.
- BARRETO, J. M. Introdução as redes neurais artificiais. *V Escola Regional de Informática. Sociedade Brasileira de Computação, Regional Sul, Santa Maria, Florianópolis, Maringá*, p. 5–10, 2002. Citado na página 33.
- BEZERRA, E. Introdução à aprendizagem profunda. *Artigo–31º Simpósio Brasileiro de Banco de Dados–SBBD2016–Salvador*, 2016. Citado 2 vezes nas páginas 35 e 36.
- BISSI, T. et al. Reconhecimento facial com os algoritmos eigenfaces e fisherfaces. Universidade Federal de Uberlândia, 2018. Citado na página 32.
- BOHNENBLUST, F. Convex regions and projections in minkowski spaces. *Annals of Mathematics*, JSTOR, p. 301–308, 1938. Citado na página 28.
- BOLLE, R. M.; CONNELL, J. H.; PANKANTI, S.; RATHA, N. K.; SENIOR, A. W. *Guide to biometrics*. [S.l.]: Springer Science & Business Media, 2013. Citado na página 16.
- BRAGA, L. F. Z. et al. *Sistemas de Reconhecimento Facial*. Tese (Doutorado) — UNIVERSIDADE DE SÃO PAULO, 2013. Citado 2 vezes nas páginas 13 e 45.
- COLLOBERT, R.; WESTON, J. A unified architecture for natural language processing: Deep neural networks with multitask learning. In: ACM. *Proceedings of the 25th international conference on Machine learning*. [S.l.], 2008. p. 160–167. Citado na página 35.
- COSTA, L. R.; OBELHEIRO, R. R.; FRAGA, J. S. Introdução á 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. Citado 2 vezes nas páginas 16 e 17.

- DALAL, N.; TRIGGS, B. Histograms of oriented gradients for human detection. In: . [S.l.: s.n.], 2005. Citado na página 25.
- DIAS, P. F.; PINTO, L. A.; CAVALIERI, D. C.; GARCIA, F. Utilização de técnicas de deep learning para reidentificação de pessoas. 2018. Citado na página 14.
- DUBES, R. C.; JAIN, A. K. *Algorithms for clustering data*. [S.l.]: Prentice hall Englewood Cliffs, 1988. Citado na página 28.
- EVERITT, B. *Cluster Analysis, 3rd edn., Edward Arnold*. [S.l.]: Hodder & Stoughton), London, Melbourne, Auckland and Halsted Press, New York . . . , 1993. Citado na página 28.
- FERRERO, C. A. *Algoritmo kNN para previsão de dados temporais: funções de previsão e critérios de seleção de vizinhos próximos aplicados a variáveis ambientais em limnologia*. Tese (Doutorado) — Universidade de São Paulo, 2009. Citado 4 vezes nas páginas 27, 28, 29 e 30.
- FILHO, F. T. d. M.; SILVA, A. M. d.; NETO, J. N. d. O. Orientação à segurança em condomínios: Normas e procedimentos de segurança. *Revista Técnica em Sistemas Agroindustriais*, v. 2, n. 1, 2016. Citado na página 12.
- FILHO, O. M.; NETO, H. V. *Processamento digital de imagens*. [S.l.]: Brasport, 1999. Citado na página 44.
- FONTANIVA, G. A. R. Monitoramento de ambientes e controle de frequência acadêmica através de visão computacional. Universidade Federal da Fronteira Sul, 2014. Citado na página 14.
- HAYKIN, S. *Redes neurais: princípios e prática*. [S.l.]: Bookman Editora, 2007. Citado na página 24.
- HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2016. p. 770–778. Citado na página 48.
- HE K., Z. X. R. S. e. S. J. Deep residual learning for image recognition. em proceedings of the ieee conference on computer vision and pattern recognition. 2016. Citado 2 vezes nas páginas 38 e 39.
- HENRIQUES, A. d. P. d. M. Classificação de imagens de ambientes coralinos: uma abordagem empregando uma combinação de classificadores e máquina de vetor de suporte. Universidade Federal do Rio Grande do Norte, 2008. Citado na página 23.
- HINTON, G.; LECUN, Y.; BENGIO, Y. Deep learning. *Nature*, v. 521, n. 7553, p. 436–444, 2015. Citado na página 35.
- HUBEL, D. H. *The brain*. [S.l.]: Freeman, San Francisco, 1979. 2–14 p. Citado na página 33.
- KING, D. *High quality face recognition with deep metric learning (2017)*. 2017. Citado na página 48.

- LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, IEEE, v. 86, n. 11, p. 2278–2324, 1998. Citado 2 vezes nas páginas 36 e 37.
- LINDEN, R. Técnicas de agrupamento. *Revista de Sistemas de Informação da FSMA*, v. 4, n. 4, p. 18–36, 2009. Citado na página 48.
- LIU, H.; MOTODA, H. *Computational methods of feature selection*. [S.l.]: CRC Press, 2007. Citado na página 29.
- LOPES, E. C.; FILHO, J. C. B.; NO, R. T. Detecção de faces e características faciais. *Porto Alegre: PUCRS*, 2005. Citado 7 vezes nas páginas 39, 40, 41, 42, 43, 44 e 45.
- LORENA, A. C.; CARVALHO, A. C. de. Introduçãos máquinas de vetores suporte. *Relatório Técnico do Instituto de Ciências Matemáticas e de Computação (USP/Sao Carlos)*, v. 192, 2003. Citado na página 25.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, Springer, v. 5, n. 4, p. 115–133, 1943. Citado 2 vezes nas páginas 33 e 34.
- MICHALSKI, R. S.; BRATKO, I.; KUBAT, M. et al. *Machine learning and data mining: methods and applications*. [S.l.]: wiley New York, 1998. v. 388. Citado na página 20.
- MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. *Sistemas inteligentes-Fundamentos e aplicações*, v. 1, n. 1, p. 32, 2003. Citado 4 vezes nas páginas 12, 19, 21 e 22.
- MONARD, M. C.; PRATI, R. C. Aprendizado de máquina simbólico para mineração de dados. 2006. Citado 2 vezes nas páginas 21 e 22.
- MORAES, J. *Controle de Acesso Baseado em Biometria Facial*. Dissertação (Mestrado) — Universidade Federal do Espírito Santo, 2010. Citado 4 vezes nas páginas 12, 17, 18 e 19.
- NIELSEN, M. A. *Neural networks and deep learning*. [S.l.]: Determination Press, 2015. Citado 2 vezes nas páginas 34 e 35.
- NUNES, F. T. et al. Técnicas de biometria baseadas em padrões faciais e sua utilização na segurança pública. Araranguá, SC, 2017. Citado 2 vezes nas páginas 12 e 16.
- OKABE, R. K.; CARRO, S. A. Reconhecimento facial em imagens capturadas por câmeras digitais de rede. In: *Colloquium Exactarum*. [S.l.: s.n.], 2015. v. 7, n. 1. Citado na página 13.
- PACHECO, M. A. C. et al. Algoritmos genéticos: princípios e aplicações. *ICA: Laboratório de Inteligência Computacional Aplicada. Departamento de Engenharia Elétrica. Pontifícia Universidade Católica do Rio de Janeiro. Fonte desconhecida*, p. 28, 1999. Citado na página 22.
- PAPAVERO, N. Fritz müller e a comprovação da teoria de darwin. *A recepção do darwinismo no Brasil. Rio de Janeiro: Editora Fiocruz*, p. 29–44, 2003. Citado na página 22.

- PARDO, T. A. S.; NUNES, M. d. G. V. Aprendizado bayesiano aplicado ao processamento de línguas naturais. *Série de Relatórios Técnicos do Instituto de Ciências Matemáticas e de Computação-ICMC, Universidade de São Paulo*, n. 180, 2002. Citado na página 21.
- PONTI, M. A.; COSTA, G. B. P. da. Como funciona o deep learning. *arXiv preprint arXiv:1806.07908*, 2018. Citado na página 36.
- QUILLFELDT, J. A. Origem dos potenciais elétricos das células nervosas. *Departamento de Biofísica, IB, URGs*, p. 1–18, 2005. Citado na página 33.
- RIGAMONTI, G. H. Sistema de reconhecimento facial baseado em redes neurais convolucionais profundas com aplicacao no comercio varejista. 2018. Citado 3 vezes nas páginas 25, 26 e 27.
- ROCHA, R. H. S. *Reconhecimento de Objetos por Redes Neurais Convolutivas*. Tese (Doutorado) — PUC-Rio, 2015. Citado na página 36.
- ROSENBLATT, F. Principles of neurodynamics: Perceptrons and the theory of brain mechanisms. *Washington DC: Spartan*, 1962. Citado na página 34.
- ROWLEY, H. A.; BALUJA, S.; KANADE, T. Neural network-based face detection. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 20, n. 1, p. 23–38, 1998. Citado 2 vezes nas páginas 43 e 44.
- SANTOS, T. P. D.; STEMMER, M. R.; CASAGRANDE, J. H. Rastreamento de múltiplos objetos em cenas de videovigilância baseado no algoritmo de extração vibe e filtro de kalman. 2018. Citado na página 14.
- SHALEV-SHWARTZ, S.; BEN-DAVID, S. *Understanding machine learning: From theory to algorithms*. [S.l.]: Cambridge university press, 2014. Citado na página 36.
- SHANNON, C. A universal Turing machine with two internal states automata studies, princeton u. *Press, Princeton, NJ*, 1956. Citado na página 33.
- SILVA, A. L.; CINTRA, M. E. Reconhecimento de padrões faciais: Um estudo. In: *Encontro Nacional de Inteligência Artificial e Computacional, 2015, Proceedings ENIAC*. [S.l.: s.n.], 2015. p. 224–231. Citado na página 13.
- SIMONYAN K. E ZISSERMAN, A. Deep convolutional neural networks for interpretable analysis of eeg sleep stage scoring. *CVPR14*, 2014. Citado 2 vezes nas páginas 35 e 39.
- SOARES, H. B. Análise e classificação de imagens de lesões da pele por atributos de cor, forma e textura utilizando máquina de vetor de suporte. Universidade Federal do Rio Grande do Norte, 2008. Citado 3 vezes nas páginas 23, 24 e 25.
- SUN, C.; SHRIVASTAVA, A.; SINGH, S.; GUPTA, A. Revisiting unreasonable effectiveness of data in deep learning era. In: *IEEE Computer Vision (ICCV), 2017 IEEE International Conference on*. [S.l.], 2017. p. 843–852. Citado na página 34.
- T, L. Disponível em: <http://www.leandroteles.com.br/blog/2015/11/09/memoria-saiba-mais-sobre-essa-funcao-cerebral/>. 2018. Citado na página 33.
- TRAMONTIN, E. D. et al. Análise e aplicação de reconhecimento facial em sistema embarcado. Araranguá, SC, 2016. Citado na página 14.

- TURK, M. A.; PENTLAND, A. P. Face recognition using eigenfaces. In: IEEE. *Proceedings. 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. [S.l.], 1991. p. 586–591. Citado na página 30.
- VAPNIK, V.; GUYON, I.; HASTIE, T. Support vector machines. *Mach. Learn*, v. 20, n. 3, p. 273–297, 1995. Citado na página 23.
- VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: *Proceedings of the XXIX Conference on Graphics, Patterns and Images*. [S.l.: s.n.], 2016. p. 1–4. Citado 2 vezes nas páginas 37 e 38.
- WONG, S.-F.; CIPOLLA, R. Extracting spatiotemporal interest points using global information. In: CITESEER. *2007 IEEE 11th International Conference on Computer Vision*. [S.l.], 2007. p. 1–8. Citado na página 14.
- YANG, M.-H.; KRIEGMAN, D. J.; AHUJA, N. Detecting faces in images: A survey. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE Computer Society, v. 24, n. 1, p. 34–58, 2002. Citado 2 vezes nas páginas 42 e 43.
- ZUBEN, F. J. V. et al. Modelos paramétricos e não-paramétricos de redes neurais artificiais e aplicações. [sn], 1996. Citado na página 21.