



UNIVERSIDADE DO SUL DE SANTA CATARINA

ANTONIO MARCOS FRANCISCO

JÉSSYCA LUIZ ROZAR

**RECONHECIMENTO FACIAL COM
TÉCNICAS DE *MACHINE*
*LEARNING***

Palhoça, SC

2020

ANTONIO MARCOS FRANCISCO
JÉSSYCA LUIZ ROZAR

RECONHECIMENTO FACIAL COM
TECNICAS DE *MACHINE*
LEARNING

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Sistemas de Informação da Universidade do Sul de Santa Catarina, como requisito parcial à obtenção do título de Bacharel em Sistemas de Computação.

Orientador: Prof. Aran Bey Tcholakian Morales, Dr.

Palhoça, SC

2020

ANTONIO MARCOS FRANCISCO
JÉSSYCA LUIZ ROZAR

RECONHECIMENTO FACIAL COM
TECNICAS DE *MACHINE*
LEARNING

Este Trabalho de Conclusão de Curso foi julgado adequado à obtenção do título de Bacharel em Sistemas de Informação e aprovado em sua forma final pelo Curso de Graduação em Sistemas de Computação da Universidade do Sul de Santa Catarina.

Palhoça 15 de julho de 2020.

Professor e orientador Aran Bey Tcholakian Morales, Dr.
Universidade do Sul de Santa Catarina

Prof. Saulo Popov Zambiasi, Dr.
Universidade do Sul de Santa Catarina

Prof. Flávio Ceci, Dr.
Universidade do Sul de Santa Catarina

Dedicamos este trabalho a nossos pais e a todos que de forma direta ou indiretamente nos ajudaram nesta caminhada.

AGRADECIMENTOS

Ao nosso orientador Prof. Dr. Aran Bey Tcholakian Morales, que por sua competência, disponibilidade e valiosas contribuições possibilitaram a finalização deste projeto.

Aos professores do curso de Sistema de Informação que nos forneceram todas as bases necessárias para nossa formação acadêmica.

“Precisamos ser super cuidadosos com Inteligência Artificial. Potencialmente mais perigoso que armas nucleares.” (Elon Musk)

RESUMO

Para o ser humano reconhecer o rosto de pessoas por determinadas características faciais é uma habilidade natural. Porém, implementar a capacidade de reconhecimento facial em máquinas não é uma tarefa fácil. A linguagem de máquinas, conhecida como machine learning, para reconhecimento facial envolve muitos cálculos matemáticos e exige grande poder de processamento. Com a evolução dos algoritmos e o aumento da capacidade de processamento dos computadores o desenvolvimento de máquinas dotados de tecnologia de reconhecimento facial já é possível. Atualmente esta tecnologia está sendo utilizada em diversos setores, como segurança, saúde, setores governamentais entre outros.

Assim, o presente trabalho visa realizar um estudo mais amplo sobre o funcionamento de machine learning. Será abordado também o funcionamento das redes neurais artificiais, técnica desenvolvida para simular o funcionamento da própria rede neural humana. Por fim, será apresentado um comparativo de algoritmos de reconhecimento facial disponibilizados pela biblioteca openCV. Para fins de avaliação do desempenho entre os algoritmos Eigenface, Fisherface e LBPH foi necessário o desenvolvimento de uma aplicação em Python com a capacidade de identificar faces utilizando representações globais da imagem facial.

Palavras-chave: Inteligência artificial, *Machine Learning*, Redes neurais, OpenCV.

ABSTRACT

For the human being to recognize or face people by characteristic characteristics is a natural skill. However, implementing a facial recognition capability on machines is not an easy task. A machine language, known as machine learning, for facial recognition involves many mathematical calculations and requires great processing power. With the evolution of algorithms and an increase in the capacity of processing computers or the development of machines equipped with facial recognition technology, it is already possible. Currently, this technology is being used in several sectors, such as safety, health, and other sectors.

Thus, the present work aims to carry out a broader study on the functioning of machine learning. The functioning of artificial neural networks, techniques developed for simulation or functioning of the human neural network, will also be addressed. Finally, a comparison of facial recognition algorithms provided by the openCV library will be presented. For the purpose of evaluating performance between the Eigenface, Fisherface and LBPH algorithms, it was necessary to develop a Python application with the ability to identify faces, using global representations of the facial image.

Keywords: Artificial intelligence, Machine Learning, Neural networks, OpenCV.

LISTA DE ILUSTRAÇÕES

Figura 1- Linha do tempo da ascensão da IA.....	16
Figura 2 - Representação da IA.....	17
Figura 3 - Resumo de IA.....	18
Figura 4 - Diagrama de uma rede neural.....	19
Figura 5 - Classificação de imagens.....	20
Figura 6 - Componentes do neurônio biológico.....	23
Figura 7 - Neurônio artificial	24
Figura 8 - Camada de uma Rede Neural Artificial (RNA)	25
Figura 9 - Funcionamento do reconhecimento facial.....	27
Figura 10 - Rede neural Convolutacional	28
Figura 11 - Face Média e eigenfaces.....	37
Figura 12 - Aplicação da operação LBP.	38
Figura 13 - Histogramas de Padrões Binários Locais	39
Figura 14 - Pacote de faces Yale.....	40

SUMÁRIO

Sumário

1	INTRODUÇÃO.....	11
1.1	PROBLEMÁTICA	12
1.2	OBJETIVOS.....	12
1.2.1	Objetivo geral	13
1.2.2	Objetivo específico.....	13
1.3	JUSTIFICATIVA	13
1.4	ESTRUTURA	14
2	REVISÃO DA LITERATURA	15
2.1	CONCEITOS E FUNDAMENTOS	15
2.2	REDES NEURAIS	21
2.2.1	RN biológicas	23
2.2.2	RN artificiais	24
2.3	RN Convolucionais	26
3	MÉTODO.....	29
3.1	CARACTERIZAÇÃO DA PESQUISA.....	29
3.2	ETAPAS METODOLÓGICAS.....	30
3.3	PROPOSTA DA SOLUÇÃO	31
3.4	DELIMITAÇÕES.....	32
4	DESENVOLVIMENTO.....	33
4.1	ESTRUTURA DO OPENCV	34
4.2	PROTÓTIPO DESENVOLVIDO	35
4.2.1	Ferramentas utilizadas	35
4.2.2	Linguagem utilizada	36
4.2.3	Algoritmos OpenCV.....	36
4.3	VALIDAÇÃO DO PROTÓTIPO.....	39
4.3.1	Eigenfaces	40
4.3.2	Fisherfaces.....	41
4.3.3	LBPH.....	42
4.4	ANÁLISE DOS RESULTADOS	42
5	CONSIDERAÇÕES FINAIS	43
5.1	TRABALHOS FUTUROS	43
	REFERÊNCIAS	44

1 INTRODUÇÃO

A cada dia cresce o volume de dados e, conseqüentemente a necessidade de processamentos cada vez mais rápido. Dessa forma se torna difícil a realização desse processamento por um ser humano, por esse motivo tornou-se necessário a utilização de ferramentas computacionais para melhorar o processamento de dados, assim obtendo a resolução de problemas de forma mais rápida (FACELI et al.,2015).

Todos nós realizamos várias tarefas diariamente as quais exigem diversos processamentos por parte de nosso cérebro, responsável por administrar tudo o que pensamos e sentimos. Baseado no desempenho do nosso sistema nervoso foi desenvolvido as redes neurais artificiais, que simulam o funcionamento do cérebro humano, sendo capaz de processar e aprender com os dados obtidos (FACELI et al., 2015).

O reconhecimento facial utiliza técnicas de aprendizado de máquinas, chamada *machine learning*, em conjunto com as redes neurais artificiais é uma tarefa desafiadora, pois a cada dia cresce a exigência por uma assertividade maior, principalmente nos sistemas de segurança. Devido a vários tipos de cenários como o ponto de vista da câmera, iluminação e a expressão facial podem dificultar esse reconhecimento. Após a extração das características obtidas no processamento, são analisados se é um padrão de uma face, fazendo parte do padrão de face é possível realizar a diferenciação das características, analisando pelos padrões obtidos e comparando com as demais faces (LI; JAIN, 2004).

O *deep learning* recebe um conjunto grande de dados com um processamento rápido e utilizando um algoritmo sofisticado, dessa forma os computadores podem realizar tarefas que antes eram realizadas somente por seres humanos, como por exemplo, reconhecimento de voz e de imagens. (DEEP... 2019).

A rede neural é composta por diversas unidades, dentre elas os neurônios, que têm como função transmitir e processar dados. A rede neural artificial realiza diversas conexões e se desenvolve através de treinamentos. (DEEP... 2019).

Dessa forma será apresentado neste trabalho o crescimento da area de inteligência artificial no que implica sobre o reconhecimento de padrões, assim como o próprio uso das redes neurais de forma prática para obter o reconhecimento facial.

1.1 PROBLEMÁTICA

O reconhecimento facial tem auxiliado em diversas questões, como a entrada de pessoas autorizadas em determinado local, identificação de possíveis terroristas, fraudes ou pessoas desaparecidas. Sendo assim, um falso reconhecimento facial pode ocasionar muitos problemas como por exemplo, autorizar a entrada de pessoas não permitidas ou transações bancárias não consentidas.

De acordo com Valente Mafei (2019), quanto menos o sistema for treinado e inserido uma menor diversidade dos dados, o sistema pode ser falho, diminuindo a capacidade de uma correta identificação dos tipos de face, como por exemplo as faces de uma mulher negra tem maior quantidade de falhas comparado com uma mulher branca, podendo gerar assim falsos positivos, tomando uma ação que pode prejudicar a vida de uma pessoa inocente.

Dessa forma é necessário ter o conhecimento do algoritmo a ser utilizado de acordo com o cenário, para identificar se o algoritmo X é o melhor para certa aplicação comparado com o algoritmo Y, se o mais importante para o cenário é alta porcentagem de assertividade ou a confiança, dessa forma esse trabalho visa apresentar algumas características importantes sobre o funcionamento de três algoritmos da biblioteca OpenCV para mostrar que deve ser realizado um grande estudo sobre os algoritmos de reconhecimento facial, diminuindo assim os riscos de escolher um algoritmo que não atenda o cenário a ser implementado.

1.2 OBJETIVOS

A seguir, seguem os objetivos gerais e específicos referente a aplicação da técnicas de *machine learning* para reconhecimento facial.

1.2.1 Objetivo geral

O Objetivo deste trabalho é mostrar a aplicação de técnicas de *machine learning* no reconhecimento facial, avaliando a assertividade das diferentes técnicas encapsuladas nos métodos da biblioteca open source OPENCV.

1.2.2 Objetivo específico

Para atingir o objetivo principal será realizado os seguintes procedimentos

- Verificar as técnicas de *machine learning* para reconhecimento facial.
- Conhecer o funcionamento da biblioteca OpenCV, sua estrutura e performance.
- Realizar a implementação com os métodos da biblioteca OpenCV.
- Analisar os resultados de reconhecimento facial e identificar a assertividade e confiança em diferentes tipos de faces.

1.3 JUSTIFICATIVA

Conforme Li e Jain (2004, p. 4), o desempenho de muitos métodos de reconhecimento facial são altamente sensíveis a alterações de iluminação e expressões faciais. Além disso, diante de falsos positivos, ocorreu a necessidade de se aprofundar sobre o reconhecimento de padrões focado no reconhecimento facial para que seja possível identificar a assertividade de cada técnica utilizada no reconhecimento facial.

Dessa forma este trabalho veio para deixar mais claro sobre esse assunto e auxiliar nas futuras escolhas de algoritmos com melhor desempenho de aplicações, voltadas ao reconhecimento de faces.

1.4 ESTRUTURA

O trabalho é organizado em 5 capítulos, sendo que no primeiro será dado uma breve introdução sobre o tema de *machine learning* aplicado em reconhecimento de padrões, assim como as justificativas para o desenvolvimento do trabalho. No capítulo 2 é apresentado sobre as técnicas de inteligência artificial, *machine learning* e *deep learning*, para o reconhecimento de padrões.

No capítulo 3 será apresentado o tipo de pesquisa metodológica que foi realizada no desenvolvimento do trabalho.

No capítulo 4 é apresentado a biblioteca OpenCV que é a base desse trabalho, a implementação de algoritmos de reconhecimento de padrões através do uso dessa biblioteca. O capítulo 4 apresenta também o protótipo desenvolvido e os resultados obtidos, com as considerações finais do trabalho, para compreender as técnicas que podem ser utilizadas para o reconhecimento facial baseando-se nas comparações realizadas.

Para finalizar o capítulo 5, apresenta as conclusões do trabalho e os trabalhos futuros.

2 REVISÃO DA LITERATURA

Neste capítulo serão abordados, os conceitos e fundamentos da inteligência artificial, *machine learning*, *deep learning* e redes neurais que formam a base de conhecimento do trabalho. O estado atual das áreas de conhecimento citadas, as principais aplicações e suas técnicas, também são abordadas.

2.1 CONCEITOS E FUNDAMENTOS

Ao longo da história vários filósofos e cientistas se dedicaram na análise dos vários aspectos constitutivos da inteligência humana. Posteriormente a inteligência, passou a ser estudada por outros campos do saber humano, como pela engenharia, psicologia, pedagogia, ciência cognitiva, neurologia, linguística, computação, entre outros, visando aspectos práticos e comerciais.

Assim como ocorreu com outras ciências que antes pertenciam ao campo de estudo da filosofia e depois se tornaram ciências independentes ou ramo de outras ciências, o mesmo ocorreu com o estudo da inteligência que hoje é alvo do estudo, na computação, pela ciência conhecida como Inteligência Artificial (IA) .

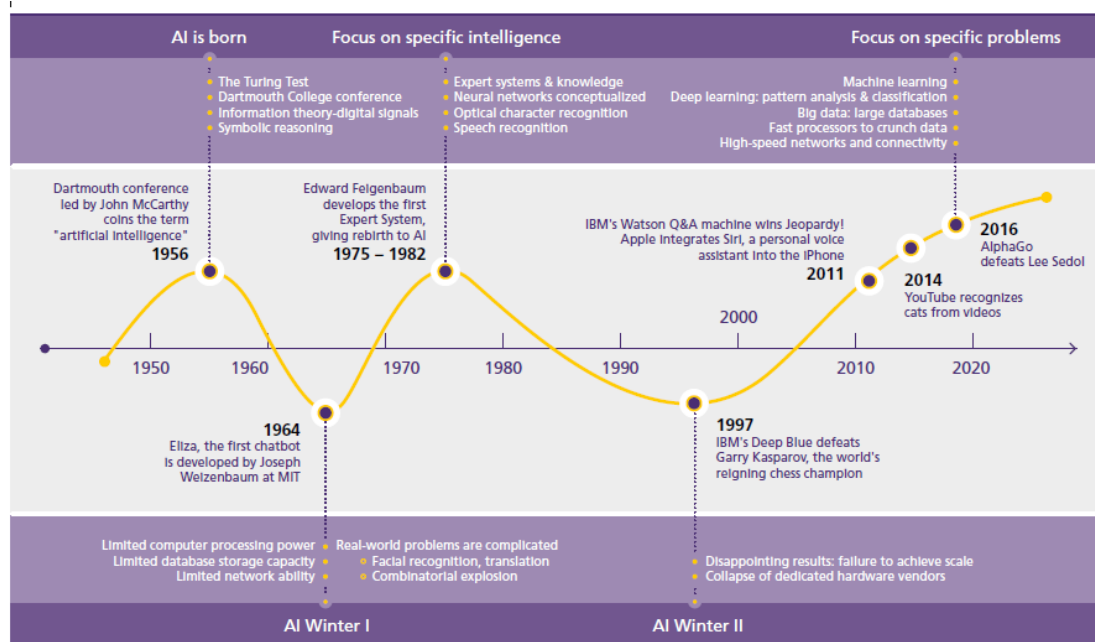
Os primeiros estudos sobre IA surgiram na década de 40, juntamente com os primeiros grandes projetos de construção de computadores, e até o final do século 20, a IA estava restrita as universidades, laboratórios e instituto de pesquisa, e as aplicações práticas existentes eram em contexto muito restritos e específicos e embarcados com hardware (LUGER, 2004).

Em especial nesta última década, e fortemente amparada pela evolução do poder computacional das máquinas, a IA vem sendo aplicada nas mais diversas áreas da atividade econômica, sendo atualmente uma das áreas de maior pesquisa e estudo pelos diversos setores econômicos que descobriram aplicações práticas de grande valor comercial e social (FACELI et al., 2011).

Conforme Figura 1 o termo Inteligência Artificial, surgiu nos Estados Unidos, em 1956, quando John McCarthy reuniu em uma conferência proferida ao Darmouth College, na Universidade de New Hampshire, vários pesquisadores que estudavam como dotar as máquinas de comportamentos inteligentes. Desde seu início, a IA tem atravessado períodos de euforia e de depressão, como mostra a figura a seguir, onde em cada fase desse histórico apresenta evoluções

com novas técnicas e novas aplicações, onde ocorreram diversos pontos interessantes, como o teste de Turing, a Eliza que foi o primeiro chatbot desenvolvido em 1964, e em 1997 a Deep blue desenvolvimento pela IBM ganhou o jogo de Xadrez do melhor jogador a época Kasparov.

Figura 1- Linha do tempo da ascensão da IA



Fonte: IBM

Desde os primórdios, o termo IA foi associado a parte da Ciência da Computação que faz com que os sistemas ou máquinas exibam atributos ou comportamentos considerados como parte da inteligência humana, tais como: aprender, adquirir conhecimentos, entender linguagens, raciocinar e resolver problemas.

Sendo assim, o objetivo central da IA, é a criação de modelos para a inteligência e a construção de sistemas computacionais baseados nesses modelos. Este objetivo é simultaneamente teórico, (a criação de teorias e modelos para a capacidade cognitiva), e prático, a implementação de sistemas computacionais baseados nestes modelos. Em outras palavras, a IA tem se destacado na busca por compreender a inteligência e por englobar diversos campos do conhecimento com o objetivo prático de simular a inteligência (LUGER, 2004).

O caráter dual deste objetivo levou, naturalmente, a pesquisa e desenvolvimento em IA cristalizaram-se em torno de três linhas de pesquisa:

- IA cognitiva: consiste no desenvolvimento de modelos formais para a inteligência humana, tema da ciência cognitiva, também chamada de psicologia computacional;
- IA básica: consiste na exploração e experimentação de técnicas computacionais que apresentem potencial para a simulação do comportamento inteligente;

- IA aplicada: é o desenvolvimento de aplicações nas mais diversas áreas das atividades econômicas (comerciais, industriais, educacionais) utilizando técnicas de IA.

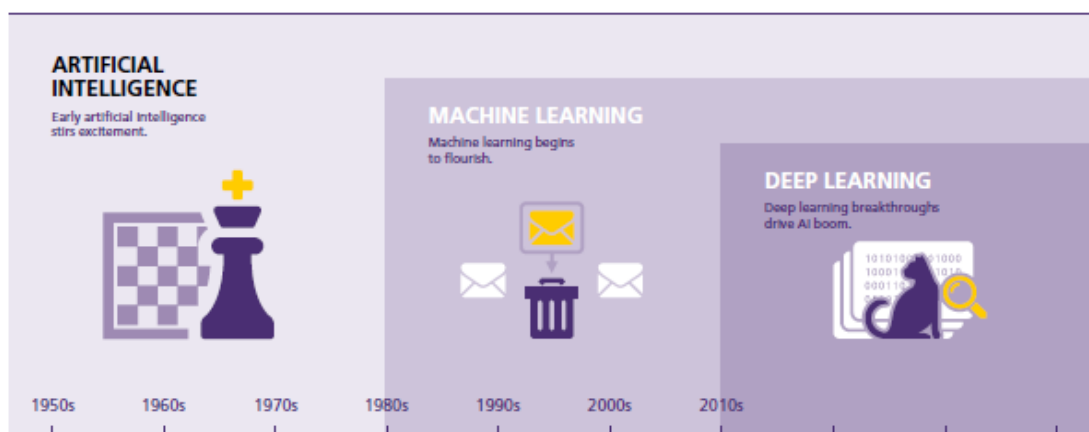
As técnicas da IA básica, desde seus primórdios foram divididas em dois grandes grupos ou paradigmas: IA simbólica e IA biológica.

- Abordagem Simbólica (ou cognitiva): dá ênfase aos processos cognitivos, ou seja, os métodos simbólicos em IA, procuram emular o raciocínio do ser humano na solução de problema.
- Abordagem Biológica: dá ênfase aos modelos naturais (por exemplo, o funcionamento do cérebro, dos neurônios e das suas conexões e o processo de evolução natural), para construir técnicas que tenham potencial de emular o comportamento inteligente.

O presente trabalho está enquadrado na IA básica, na sua abordagem biológica e na IA aplicada. Serão utilizadas técnicas computacionais biológicas, mais precisamente redes neurais, para emular o comportamento inteligente e aplicadas em um estudo de caso (RUSSELL; NORVIG, 2004).

Desde uma perspectiva histórica, a IA básica e aplicada, podem ser abordadas também olhando para as técnicas utilizadas e as aplicações resolvidas com essas técnicas, como mostram as figuras a seguir.

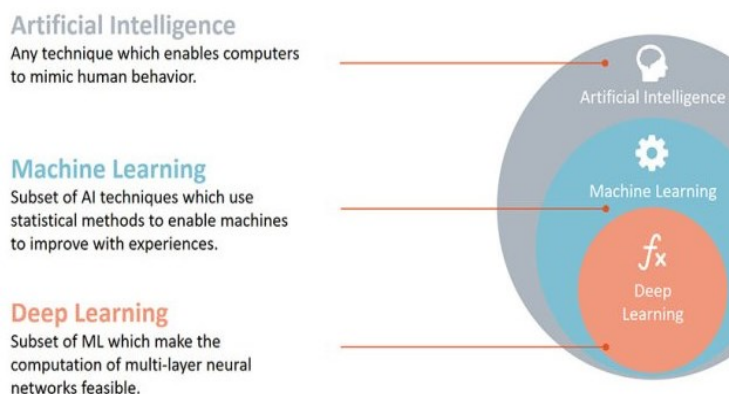
Figura 2 - Representação da IA



Fonte: IBM

Na figura 2 é possível identificar os subconjuntos, Deep Learning é um subconjunto de *Machine Learning* e *Machine Learning* sendo um subconjunto de Inteligência Artificial

Figura 3 - Resumo de IA



Fonte: IBM

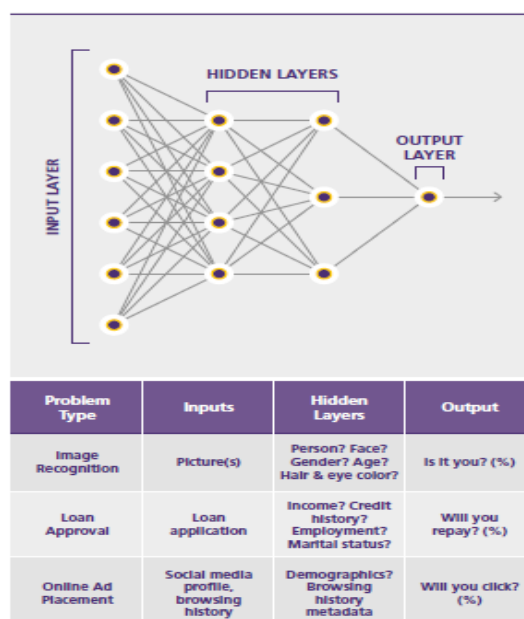
Na Figura 3 apresenta uma breve explicação sobre cada subconjunto de Inteligência Artificial, da qual a Inteligência Artificial é qualquer técnica utilizada em computadores para simular o comportamento humano.

Machine Learning sendo qualquer método de máquina que aprenda com as experiências e por ultimo *Deep Learning* que possui múltiplas redes neurais.

Considerando que IA é um sistema que exibe atributos ou comportamentos inteligentes, aprendizado de máquina é um termo mais específico que se refere a sistemas projetados para receber dados, geralmente em um domínio específico e aprende e descobre padrões ou tendências nos dados. Esses sistemas utilizam a capacidade para avaliar e categorizar os dados recebidos e depois desenhar inferências disso.

Na figura a seguir é apresentado três aplicativos baseados no *machine learning*. Por exemplo, quando a entrada é uma imagem carregada para uma plataforma de mídia social. O software analisa o conteúdo dos pixels da imagem para padrões conhecidos usando algoritmos de aprendizado de máquina, e produz uma saída identificando a foto carregada. É capaz de fazer isso com base na probabilidade estatística de que características da imagem se assemelham às imagens existentes, conjunto de imagens de treino para “calibrar” a rede (IBM, 2018).

Figura 4 - Diagrama de uma rede neural



Fonte: IBM

De acordo com a Data Science Academy (2019) *Machine learning* possui dois principais tipos de aprendizado, o aprendizado supervisionado e o não supervisionado. O objetivo é reproduzir o procedimento que os seres humanos fazem para aprender e tentam replicar em máquinas, o aprendizado supervisionado é aquele que ensina a máquina através de algoritmos o que é uma face e a partir disso ele consegue reconhecer novas faces através das fórmulas criadas.

O não supervisionado é aquele que possui um conjunto de dados e a máquina localiza as imagens que são semelhantes, dessa forma sem saber a saída dessa análise tenta localizar as imagens que possuem um mesmo tipo de padrão, ou também a tentativa e erro, a máquina tenta realizar uma ação, dependendo dessa ação se for uma ação errada, ela faz uma nova tentativa, até encontrar a melhor solução para o problema. Todas essas ações da máquina são realizadas através de fórmulas matemáticas.

Nunca houve na história tantos dados como existe atualmente, são necessários novas técnicas, ferramentas e formas de abordagens para que seja possível processar e analisar essa grande quantidade de dados, gerando valores para mudar a vida das pessoas. Os sistemas de IA estão evoluindo cada dia mais, devido ao bigData que gera muitos dados e em grande velocidade e variedade e ao processamento em GPUs, com esses dois pilares é possível obter ótimos resultados.

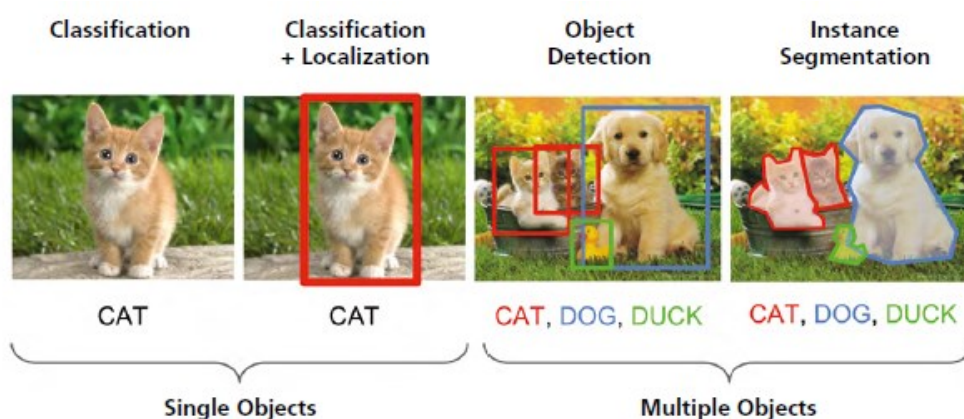
Segundo IBM (2018), o *deep learning* leva o conceito de *machine learning* a um pouco mais longe. *Deep learning* é sobre aprender continuamente, isto é, a intenção do sistema é

aprender com o mundo real e ajustar o modelo de aprendizado à medida que recebe novas informações e cria novas ideias.

Deep learning geralmente é realizado com redes neurais. Redes neurais são uma tentativa de imitar a estrutura e funcionamento do cérebro humano. Com novos dados alimentando uma rede neural, as conexões entre os nós são estabelecidas, fortalecidas ou diminuídas de maneira semelhante as conexões entre neurônios no cérebro humano. Além disso, cada conexão em uma rede neural pode ser sintonizada, atribuindo maior ou menor importância a um atributo, para alcançar a qualidade da saída.

De forma simplificada, a figura a seguir mostra como os algoritmos da aprendizagem profunda podem distinguir o conteúdo de uma imagem, como bem como onde os elementos da imagem estão em relação uns aos outros, analisando dados de pixel. Esse é o tipo de sistema que é mais útil em enfrentar os desafios de dados do mundo real, e é por isso que sistemas de aprendizagem são aqueles direcionados para conjuntos de dados extremamente grandes e de movimentação rápida, normalmente encontrados em plataformas de mídia social e em veículos autônomos.

Figura 5 - Classificação de imagens



Fonte: Stanford

Data Science Academy (2019) afirma que *Deep learning* [...] é atualmente a principal tecnologia utilizada em sistemas de inteligência artificial. São redes neurais profundas que tentam simular em máquinas o funcionamento do cérebro humano, que vem sendo usadas em diversas aplicações de inteligência artificial, com níveis de precisão cada vez maiores.

A técnica do *deep learning* está avançando rapidamente devido ao grande volume de dados que são gerados todo dia, esses dados estão duplicando a cada ano, esse avanço também está relacionado a grande melhoria das máquinas que são utilizadas para a análise dessa técnica, onde houve um aumento de desempenho de hardware, facilitando a coleta de informações e de forma mais eficaz, essa técnica é muito utilizada, pois ela é flexível no uso dos dados, realizando assim a melhor combinação deles, gerando o resultado mais eficaz.

Os dados de forma isolada são apenas dados, porém quando esses dados são analisados e aplicados uma técnica, eles geram informações e as informações são importantes para tomadas de decisão. (DATA SCIENCE ACADEMY, 2019).

De acordo com Russell e Norvig (2004), os modelos tradicionais de *machine learning* sempre foram muito poderosos para lidar com dados estruturados e têm sido amplamente utilizados pelas empresas para classificação de crédito, previsão de rotatividade, segmentação por consumidor e assim por diante. O sucesso desses modelos depende muito do desempenho da fase de engenharia de recursos: quanto mais trabalhamos perto do negócio para extrair conhecimento relevante dos dados estruturados, mais poderoso será o modelo.

Quando se trata de dados não estruturados (imagens, texto, voz, vídeos), os recursos de engenharia manual são demorados, frágeis e não escalonáveis na prática. É por isso que as redes neurais se tornam cada vez mais populares, graças à sua capacidade de descobrir automaticamente as representações necessárias para a detecção ou classificação de recursos a partir de dados brutos. Isso substitui a engenharia de recursos manual e permite que uma máquina aprenda os recursos e os use para executar uma tarefa específica (MIT 6.S191, 2020) A base do *deep learning*, são as redes neurais, em particular a rede neural convolucional(CNN), destaca-se no reconhecimento de padrões e é a principal tecnologia para carros autônomos, reconhecimento facial, transcrição de fala entre outras aplicações.

2.2 REDES NEURAI

Conforme Braga, Carvalho, Ludermir (2000) as Redes Neurais Artificiais – RNAs tentam reproduzir as funções das redes biológicas, buscando implementar seu comportamento básico e sua dinâmica.

Haykin (2001) descreve o sistema nervoso humano como um sistema de três estágios, receptores, rede neural e atuadores. Neste sistema o cérebro, representado pela rede neural, figura como componente central, tendo como função, a análise dos impulsos elétricos

convertidos pelos receptores a partir de estímulos do corpo humano e o envio dessa informação aos atuadores que a converterão em respostas discerníveis ao sistema.

Coppin (2017) afirma que “o cérebro humano possui mais de dez bilhões de neurônios, cada qual conectado, em média, a milhares de outros neurônios”.

Cada um destes neurônios processa e se comunica com milhares de outros continuamente e em paralelo, complementa Braga, Carvalho, Ludermir (2000).

Segundo Lippmann(1997) as Redes Neurais Artificiais são sistemas físicos que podem adquirir, armazenar e utilizar conhecimentos experimentais, que podem alcançar uma boa performance, devido a sua densa interconexão entre os nós da rede. Elas também são conhecidas por modelos conexionistas, modelos de processamento paralelo distribuído e sistemas neuromofológicos.

Gurney (1997) afirma que em uma rede neural artificial, possuem diversas unidades de processamento, cada unidade de processamento possui um peso atrelado a ela e cada uma realiza as operações baseando-se nos seus dados locais adquiridas pelos canais de comunicação, a partir dessas interações dos canais de comunicação é onde são criadas o comportamento inteligente de uma rede neural artificial.

As redes neurais artificiais aprendem através de uma entrada de dados, podendo ser ou não supervisionada. A rede neural profunda, possui muitos neurônios e muitas camadas, se tornando assim a chamada *deep learning*, essas redes são técnicas de computação de forma matemática, se baseando na estrutura neural de organismos inteligentes, adquirindo assim conhecimento através da experiência obtida (DATA SCIENCE ACADEMY, 2019).

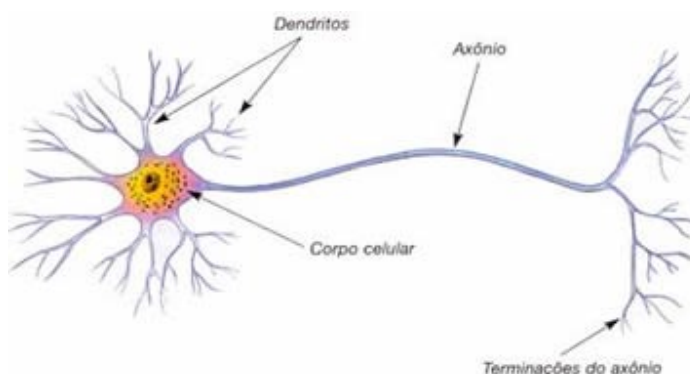
O neurônio é constituído por 3 partes principais: a soma ou corpo celular, do qual emanam algumas ramificações denominadas de dendritos, e por uma outra ramificação descendente da soma, porém mais extensa, chamada de axônio. Nas extremidades dos axônios estão os nervos terminais, pelos quais é realizada a transmissão das informações para outros neurônios. Esta transmissão é conhecida como sinapse (DATA SCIENCE ACADEMY, 2019).

2.2.1 RN biológicas

Os neurônios são divididos em três seções: o corpo da célula, os dendritos e o axônio, cada um com funções específicas, porém complementares. (BRAGA; CARVALHO; LUDERMIR, 2000):

Os dendritos têm por função receber as informações, ou impulsos nervosos, oriundas de outros neurônios e conduzidas até o corpo celular. Aqui a informação é processada, e novos impulsos são gerados. Estes impulsos são transmitidos a outros neurônios, passando através do axônio até os dendritos dos neurônios seguintes. O ponto de contato entre a terminação de um axônica de um de um neurônio e o dendrito de outro é chamado de sinapse. É pelas sinapses que os nodos se unem funcionalmente, formando redes neurais. As sinapses funcionam como válvulas e são capazes de controlar a transmissão de impulsos, isto é, o fluxo da informação entre os nodos na rede neural.

Figura 6 - Componentes do neurônio biológico



Fonte: Fonte: www.datascienceacademy.com.br

Na figura 6 apresenta um exemplo dos components de um neurônio biológico, onde foi a base para desenvolvimento do neurônio artificial. Os pulsos chegam nos dendritos, esses pulsos são processados pelo corpo celular ou núcleo e são passados para o axônio até chegar nas suas terminações e passando para outros neurônios.

Cada parte do cérebro tem uma função determinada e uma arquitetura diferente, variando quantidade de neurônios, sinapses e pesos podendo ser pela criação dos pensamentos, processamento de imagens, áudios, desejos entre outros, esses pensamentos são devidos as redes interligadas da qual realizam um processamento paralelo. Cada peso possui um valor e esse valor é dado através dos treinamentos a chamada memorização (DATA SCIENCE ACADEMY, 2019).

A Data Science Academy (2019) informa que o ponto de contato entre a terminação axônica de um neurônio e o dendrito de outro é chamado sinapse. É pelas sinapses que os neurônios se unem funcionalmente, formando as redes neurais. As sinapses funcionam como válvulas, sendo capazes de controlar a transmissão de impulsos, isto é, o fluxo da informação entre os neurônios na rede neural. O efeito das sinapses é variável e é esta variação que dá ao neurônio capacidade de adaptação.

Um neurônio recebe sinais através de inúmeros dendritos, os quais são ponderados e enviados para o axônio, podendo ou não seguir adiante (threshold). Na passagem por um neurônio, um sinal pode ser amplificado ou atenuado, dependendo do dendrito de origem, pois a cada condutor, está associado um peso pelo qual o sinal é multiplicado. Os pesos são o que chamamos de memória (DATA SCIENCE ACADEMY, 2019).

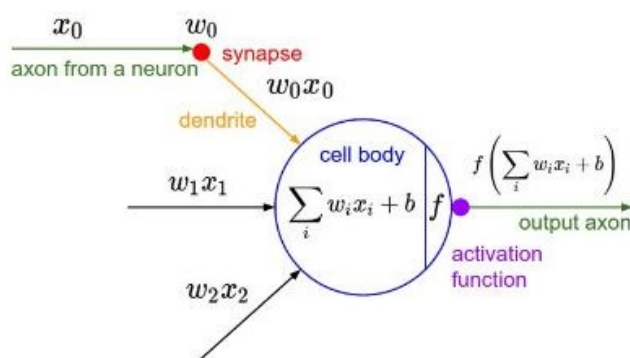
2.2.2 RN artificiais

Em 1943 McCulloch e Pitts propõem um modelo simplificado do funcionamento de um neurônio, conforme Braga, Carvalho e Ludermir (2000):

Sua descrição matemática resultou em um modelo com n terminais de entrada $x_1, x_2, \dots, x_n$ (que representam os dendritos) e apenas um terminal de saída y (representando o axônio). Para emular o comportamento das sinapses, os terminais de entrada do neurônio têm pesos acoplados $w_1, w_2, \dots, w_n$ cujos valores podem ser positivos ou negativos, dependendo de as sinapses correspondentes serem inibitórias ou excitatórias. O efeito de uma sinapse particular i no neurônio pós-sináptico é dado por $x_i w_i$. Os pesos determinam em que “grau” o neurônio deve considerar sinais de disparo que ocorrem naquela conexão.

Uma descrição do modelo está ilustrada na figura abaixo:

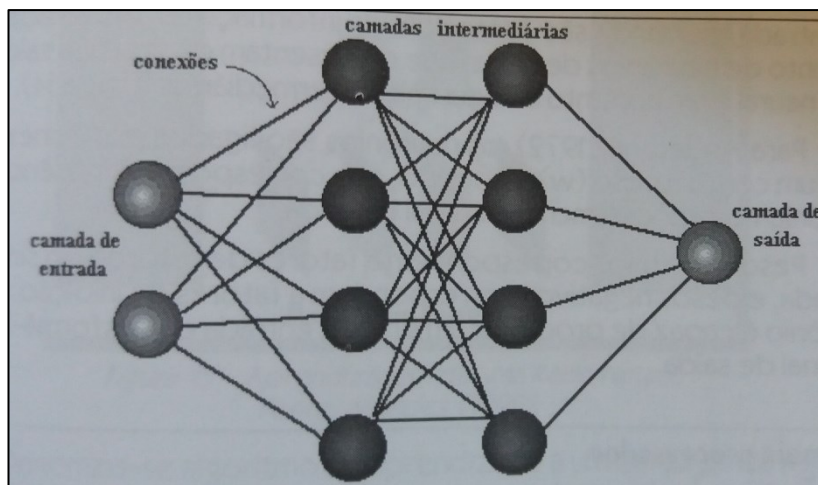
Figura 7 - Neurônio artificial



Fonte: <http://cs231n.github.io/neural-networks-1/>

As redes neurais geralmente são constituídas através de camadas, que podem ser conectadas em outras camadas (FERNANDES, 2003).

Figura 8 - Camada de uma Rede Neural Artificial (RNA)



Fonte: Neves(1998).

Fernandes (2003) afirma que cada camada da rede neural artificial possui uma função específica, as camadas de entrada descrevem os padrões que serão encaminhados para os outros neurônios, a camada intermediária irá extrair as características dos padrões, nessa camada é onde ocorre a maior parte do processamento da rede e por último a camada de saída que irá mostrar o resultado final a partir das camadas anteriores, baseando-se nos padrões de entradas e os pesos atribuídos para cada neurônio.

Segundo Hackley(2001) a aprendizagem ocorre quando a rede neural responde de uma maneira nova ao ambiente no qual está inserida através de um processo de ajustes de parâmetros aplicados a seus pesos sinápticos. O processo de aprendizagem é gradativo e seu desempenho melhora com o conhecimento adquirido do ambiente em que está inserido.

Denomina-se algoritmo de aprendizado a um conjunto de regras bem definidas para solução de um problema de aprendizagem, existem duas maneiras básicas de aprendizagem de uma RNA, o aprendizado supervisionado e não-supervisionado (FERNANDES, 2003).

Segundo (BRAGA; CARVALHO; LUDERMIR, 2000) no aprendizado supervisionado o processo de treinamento é direcionado por um agente externo. Já no aprendizado não-supervisionado não há um supervisor acompanhando o processo de treinamento.

Baseando-se em um neurônio biológico que geram a Inteligência humana, foi criado o neurônio matemático para simular essa inteligência, esse neurônio matemático é a base da Inteligência artificial, da qual realizam vários cálculos, se baseando nas entradas de informações através dos canais de comunicação, a partir desses dados são gerados um resultado através das funções aplicadas sobre os dados recebidos e encaminhando esse resultado para o próximo neurônio (DATA SCIENCE ACADEMY, 2019).

De acordo com a Data Science Academy (2019) a soma ou corpo da célula é representado por uma composição de dois módulos, o primeiro é uma junção aditiva, somatório dos estímulos (sinais de entrada) multiplicado pelo seu fator excitatório (pesos sinápticos), e posteriormente uma função de ativação, que definirá com base nas entradas e pesos sinápticos, qual será a saída do neurônio. Assim como no modelo biológico, o estímulo pode ser excitatório ou inibitório, representado pelo peso sináptico positivo ou negativo respectivamente.

O neurônio matemático é um modelo simplificado do neurônio biológico. Tais modelos inspirados a partir da análise da geração e propagação de impulsos elétricos pela membrana celular dos neurônios. O neurônio matemático recebe um ou mais sinais de entrada e devolve um único sinal de saída, que pode ser distribuído como sinal de saída da rede, ou como sinal de entrada para um ou vários outros neurônios da camada posterior (que formam a rede neural artificial).

2.3 RN Convolucionais

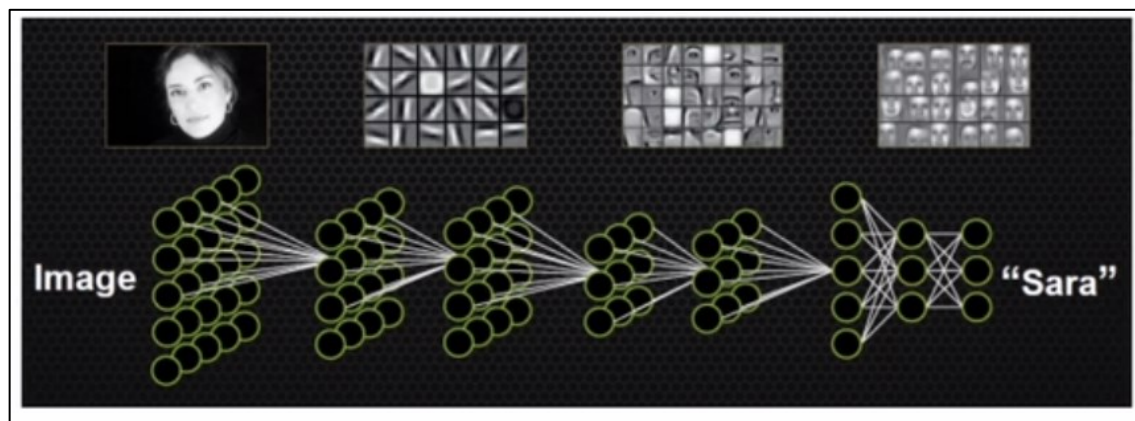
As Redes Neurais Convolucionais são usadas para diversas finalidades, entre elas a classificação de imagens, podendo associá-las por semelhança entre elas. Essas redes artificiais são profundas, são algoritmos que identificam diversos padrões, como placas e faces por exemplo. Esse tipo de rede é muito voltado para o reconhecimento de imagens, onde está surgindo muitos avanços no quesito de visão computacional, podendo ser utilizada em diversos problemas, como na área da segurança em reconhecimento de faces, diagnósticos médicos e carros autônomos por exemplo.

A rede convolucional obtém a imagem com extratos separados de cores, entre elas são o vermelho, verde e azul e com largura e altura que são medidos pela quantidade de pixels, onde a profundidade é de três camadas, uma de cada cor, essas camadas também são chamadas como canais.

Em uma CNN são realizadas várias operações matemáticas, onde ocorrem os filtros, nesses filtros se obtém as informações da imagem e no final como saída, terá uma probabilidade para identificar se a imagem pode ser um animal, pessoa ou objeto por exemplo.

Cada parte da CNN possui funções distintas, cada filtro em uma área dos pixels é gerado um peso para o que foi analisado (DATA SCIENCE ACADEMY, 2019).

Figura 9 - Funcionamento do reconhecimento facial



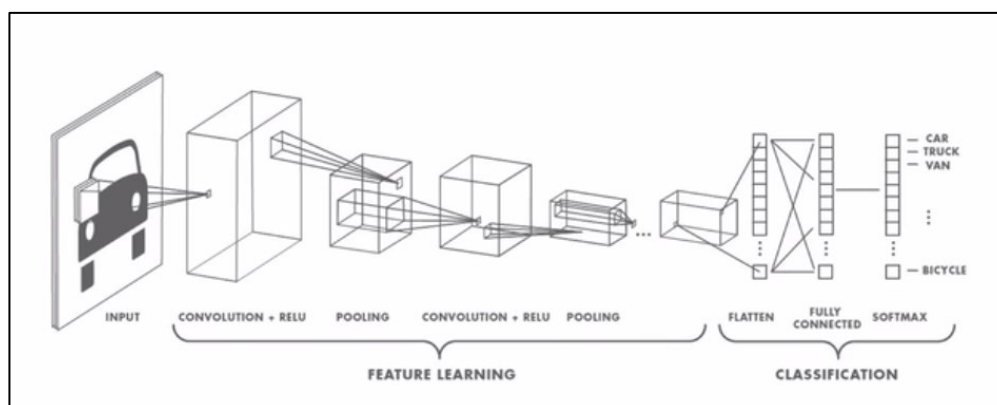
Fonte: Fonte: www.datascienceacademy.com.br

A combinação das entradas de um neurônio utilizamos pesos respectivos de cada uma de suas conexões, produz uma saída, que é passada para a camada seguinte, os pesos atribuídos as conexões de um neurônio podem ser interpretados como uma matriz que representa um filtro de uma convolução de imagens de um domínio espacial, o que também conhecido como kernel ou máscara. Enquanto a formulação clássica de perceptrons, um neurônio é conectado a todos os neurônios da camada anterior, as CNNs apenas um subconjunto de entradas é conectado a cada neurônio, com essa mudança de arquitetura as redes neurais convolucionais passam a realizar a análise de campos receptivos locais. Os neurônios da mesma camada, são agrupados em mapas, uma mapa é produzido pelo agrupamento das saídas de neurônios e juntos cobrem uma parte da imagem que tenha sido processado com filtro em comum (DATA SCIENCE ACADEMY, 2019).

A imagem é transformada em uma matriz numérica, essa matriz é encaminhada para várias camadas, onde são aplicadas primeiro a convolução, coletando assim detalhes da rede, onde irá gerar uma saída, essa saída é passada para a camada posterior que geralmente é a camada de pooling onde é realizado o ajuntamento, esse processo é realizado em várias camadas e no final na terceira transformação essas saídas são passadas para um classificador que é um *machine learning* da qual irá dar a probabilidade do que é a imagem que foi analisada.

A rede neural é treinada para aprender pesos e bias, dessa forma para que um mapa aplique o mesmo filtro em diversas partes da matriz é realizado um compartilhamento desses pesos durante esse processo de treinamento, isso ajuda a diminuir o tempo de treinamento da rede (DATA SCIENCE ACADEMY, 2019).

Figura 10 - Rede neural Convolutacional



Fonte: www.datascienceacademy.com.br

3 MÉTODO

O presente trabalho é uma pesquisa de finalidade básica pura, com objetivos descritivo, com abordagem qualitativa e executada por meio de levantamento bibliográfico.

Inicialmente, buscou-se a base teórica sobre reconhecimento facial, com a realização de fichamentos de obras doutrinárias e trabalhos acadêmicos mais atuais.

Posteriormente, foi realizado um estudo teórico, no que se refere ao uso de *machine learning* que é um subcampo da Inteligência Artificial.

Depois, vislumbra-se um texto dissertativo, em que as informações serão confrontadas, a fim de que seja viabilizado uma análise na biblioteca OpenCV com foco em reconhecimento de padrões faciais e consequentemente uma resposta ao objetivo de avaliar o algoritmo apresentado nessa biblioteca e o índice de assertividade e confiança.

3.1 CARATERIZAÇÃO DA PESQUISA

De fato, o estudo foi realizado para produzir mais do que conhecimento meramente teórico, porém não esteve presente a ambição de desenvolver uma aplicação que resolvesse definitivamente a situação do problema.

Conforme ensina Gil (2010, p. 26), "pesquisas que procuram desenvolver os conhecimentos científicos sem a preocupação direta com suas aplicações e consequências práticas" são classificadas como puras.

Nesse sentido, o presente trabalho visa apresentar uma contribuição para a ciência, proporcionar um aumento do conhecimento por meio de uma pesquisa detalhada em bibliografia específica, caracterizando-se, portanto, como uma pesquisa pura.

Quanto ao objetivo, percebe-se que foi realizado um levantamento bibliográfico, a fim de descrever o conhecimento mais atual já catalogado pela ciência sobre reconhecimento de padrões faciais, que é o assunto principal da pesquisa.

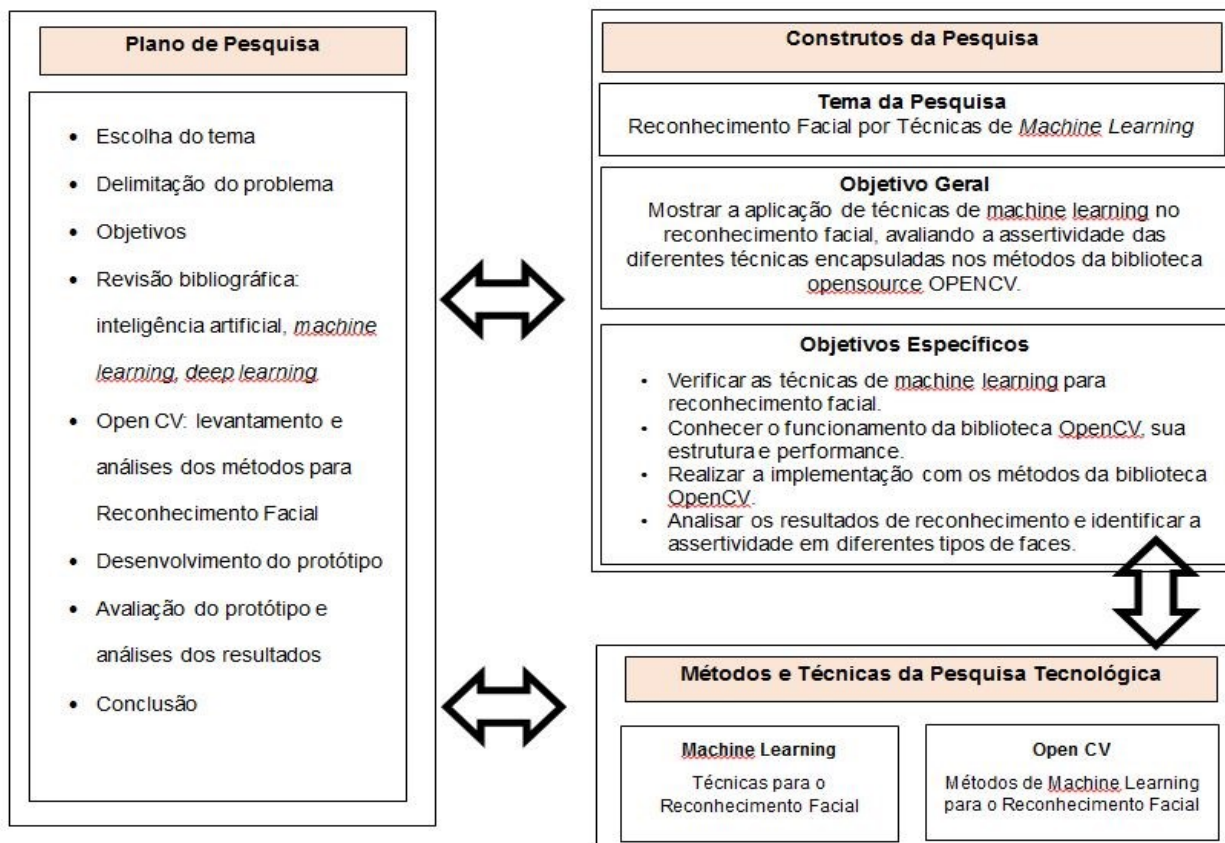
Nesse sentido, pode-se afirmar que a pesquisa se enquadra na caracterização apresentada por Duarte e Furtado (2014, p. 26), quando sustentam que “a pesquisa descritiva restringe-se a constatar o que já existe. Os acontecimentos são narrados. Procura-se conhecer a natureza, as características, a composição e os processos que constituem o fenômeno”.

Por esse motivo, a pesquisa apresentada tem cunho estritamente descritivo.

Silva (2005, p. 86) explica que “a pesquisa qualitativa não apresenta, necessariamente, linguagem matemática ou estatística. Ela utiliza a análise e interpretação da análise estudada, muitas vezes sem fazer uso de amostragem, tabelas e gráficos”.

Portanto, pode-se afirmar que se trata de uma pesquisa de abordagem qualitativa, o que é corroborado pela lição de Marconi e Lakatos (2011, p. 269), quando afirmam que “o método qualitativo difere do quantitativo não só por não empregar instrumentos estatísticos, mas também pela forma de coleta e análise dos dados.

3.2 ETAPAS METODOLÓGICAS



3.3 PROPOSTA DA SOLUÇÃO

A proposta da solução do TCC é a utilização de uma câmera webcam para captação de face em diferentes ângulos e iluminação, onde essas faces detectadas serão analisadas com o banco de dados para verificar a porcentagem de assertividade e confiança, descobrindo os pontos fortes e fracos dessa implementação da qual será utilizada a biblioteca open source chamada OpenCV.

3.4 DELIMITAÇÕES

Esse trabalho delimitou-se a análise de detecção de faces de apenas um software na área de reconhecimento de padrões atualmente e de distribuição gratuita.

4 DESENVOLVIMENTO

O OpenCV é uma biblioteca de software de visão computacional e *machine learning* de código aberto. O OpenCV foi desenvolvido para fornecer uma infraestrutura comum para aplicativos de visão computacional e acelerar o uso da percepção da máquina nos produtos comerciais. Sendo um produto licenciado pela BSD, o OpenCV facilita para as empresas a utilização e modificação do código. O OpenCV é a maior biblioteca do mundo na área de visão computacional, ele possui código aberto, facilitando assim na utilização e alteração do seu código, possuindo uma infraestrutura comum para o desenvolvimento de aplicativos na área da visão computacional e é gratuito para uso comercial (EQUIPE OPENCV, 2020).

O lançamento do OpenCV (Open Source Computer Vision Library) foi em 1999, onde até hoje tem sido utilizado em aplicativos de análise de objetos, sistemas de monitoramento, de inspeção, reconhecimento de padrões entre outras finalizadas (BRADSKI; KAEBLER, 2008).

De acordo com Bradski e Kaebler (2008) o OpenCV fornece as ferramentas necessárias para resolução de problemas na área de visão computacional, tanto para problemas mais simples quanto complexos também.

Essa biblioteca possui tanto algoritmos básicos quanto avançados na área de *machine learning*, usado para diversas finalidades, como a detecção e reconhecimento de faces, rastreamento de movimento em câmeras, classificação de ações humanas, entre outros. Essa biblioteca é muito utilizada em áreas, como universidades, pelos governos e empresas (EQUIPE OPENCV, 2020).

De acordo com a Equipe Opencv (2020) o OpenCV “possui interfaces C ++, Python, Java e MATLAB e suporta Windows, Linux, Android e Mac OS”.

No OpenCV é possível encontrar a documentação, tutoriais, fóruns para auxílio no desenvolvimento, retirando dúvidas sobre implementações, o próprio GitHub para compartilhamento de códigos e também possuem cursos. Sendo assim é possível encontrar no próprio site do OpenCV materiais para iniciar a implementação na área de visão computacional (EQUIPE OPENCV, 2020).

O OpenCV iniciou com a ajuda de grupos de desempenho de software em um laboratório da Intel e foi implementado e otimizado na Rússia onde Vadim Pisarevsky foi quem codificou grande parte do OpenCV, ele era chefe da equipe Rússia onde teve ajuda também de Victor Eruhimov na parte da infraestrutura inicial e Valey Kuriakin que auxiliou dentro do laboratório russo. (BRADSKI; KAEBLER, 2008).

De acordo com Bradski e Kaebler (2008) o OpenCV tinha muitas finalidades para o seu desenvolvimento, dentre eles.

- Avanço na pesquisa de visão, fornecendo código não apenas aberto, mas também otimizado para a infraestrutura de visão traseira. Não há mais reinvenção da roda.
- Disseminar o conhecimento da visão, fornecendo uma infra-estrutura comum em que os desenvolvedores possam construir, para que o código seja mais facilmente legível e transferível.
- Avanço na aplicação comercial baseada em visão, disponibilizando gratuitamente código portátil e otimizado para desempenho - com uma licença que não exigia que a aplicação comercial fosse aberta ou livre.

4.1 ESTRUTURA DO OPENCV

O OpenCV possui em sua estrutura diversos módulos podendo esses serem estáticos ou compartilhados (EQUIPE OPENCV, 2020).

De acordo com a Equipe Opencv (2020) os principais módulos do Open CV são:

- Módulo de estruturas básicas de dados.
- Processamento de imagem, realiza o processamento das imagens separando as imagens lineares e não lineares
- Análise de video que inclui algoritmos para rastrear objetos.
- Calibração de câmera e reconstrução 3D, é um módulo que possui algoritmos de geometria e de reconstrução 3D.
- Detecção de objeto, detecta diversos objetos, como por exemplo pessoas e veículos
- Gui de alto nível, utilizado para interfaces do usuário
- Entrada e saída de vídeo, possui codecs de vídeo e é responsável pela captura do vídeo.
- Módulo de rede neural profunda (DNN) é um módulo processador de imagem que utiliza algoritmos de *deep learning*.
- ML - é a biblioteca de *machine learning*

Aplicações DNN muitas vezes exigem quantidades significativas de treinamento em grandes clusters para determinar os pesos associados com os nós ou neurônios dentro de uma rede neural. O lado positivo é que a rede resultante pode ser implementada em hardware muito mais simples (DEEP...).

4.2 PROTÓTIPO DESENVOLVIDO

Para fins de avaliação do desempenho entre as bibliotecas OpenCV Eigenface, Fisherface e LBPH foi necessário o desenvolvimento de um protótipo de software com a capacidade de identificar faces utilizando representações globais da imagem facial.

O algoritmo foi implementado com a linguagem de programação Python em conjunto com as bibliotecas OpenCV.

A técnica de reconhecimento facial que o software utiliza é o método de correspondência holística, conforme caracterização proposta por (Zhao, et al., 2006), onde através da análise de componentes principais (PCA) e técnicas de *machine learning*, as imagens capturadas serão comparadas com as imagens contidas no banco de dados do software.

4.2.1 Ferramentas utilizadas

Para o desenvolvimento do software foram utilizadas as seguintes ferramentas:

a) PyCharm Community Edition versão 2019.3

Desenvolvida pela empresa JetBrains a ide PyCharm proporciona maior produtividade nos projetos em linguagem de programação Python. Além da edição gratuita Community, a empresa também disponibiliza a edição completa Profissional com suporte para HTML, JS e SQL.

b) Webcam com resolução imagem estática 0,92 megapixel, integrada ao notebook DELL Inspiron modelo 5458.

4.2.2 Linguagem utilizada

Quanto a linguagem, nesse protótipo de software foi utilizado a seguinte linguagem de programação:

a) Python

O Python é uma linguagem de programação orientada a objetos, multiplataforma, versátil e popular. Foi criado no início dos anos 90 por Guido Van Rossum no Centro de Matemática e Ciência da Computação da Holanda (CWI).

4.2.3 Algoritmos OpenCV

Os algoritmos que serão utilizados para extração das características faciais e já implementados na biblioteca OpenCV, serão os métodos Eigenface, Fisherface e LBPH (Local Binary Patterns Histograms).

a) Eigenface

O rosto humano possui partes que não são essenciais no processo de reconhecimento facial. Podemos reconhecer uma pessoa apenas por algumas características do rosto, como formato dos olhos, nariz, testa, etc.

O Eigenface utiliza o mesmo princípio. O algoritmo analisa as imagens de treinamento das faces, mantém todas as características importantes e necessárias para o reconhecimento e descarta as demais. Este processo é conhecido como Análise de Componentes Principais (PCA).

Assim, no método Eigenface o algoritmo irá extrair as características principais do conjunto de imagens de treinamento e de acordo com a variação dos valores dos pixels irá gerar diversas *eigenface*, ou fotos “fantasmas”. A soma dessa variação de valores irá gerar uma face média. Com base na combinação linear dos componentes das diversas *eigenfaces* juntamente com a face média será possível a reconstrução da face original.

A figura abaixo representa as eigenfaces extraídas das faces de um banco de dados, assim como a face média que será a base para a reconstrução das diversas faces.

Figura 11 - Face Média e eigenfaces



Fonte: (DAVE; CHAO; SRIADIBHATLA, 2010)

O algoritmo Eigenface classifica as faces com base no cálculo das distâncias dos pixels vizinhos mais próximos (KNN), da imagem original com as diversas eigenfaces. Faces detectadas com valor de distância fora do limite de confiança (threshold) serão consideradas desconhecidas.

b) Fisherface

O PCA não é um método ideal para realizar a separação entre as classes das faces, pois considera a iluminação como fator importante no reconhecimento. Assim, faces de uma mesma pessoa que possuem alteração na iluminação, podem ser consideradas como a face de uma pessoa diferente. O Fisherface não captura variações de iluminação tão obviamente quanto o método Eigenfaces.

Diferentemente do Eigenface que utiliza o PCA, o Fisherface emprega o método LDA (Linear Discriminant Analysis) para o reconhecimento das faces.

A técnica utilizada no LDA foi desenvolvida pelo matemático R.A. Fisher na década de 1930 e visa encontrar uma combinação linear de recursos que separa duas ou mais classes. O objetivo é maximizar a proporção da matriz de dispersão entre classes e a matriz de dispersão dentro da classe.

c) LBPH

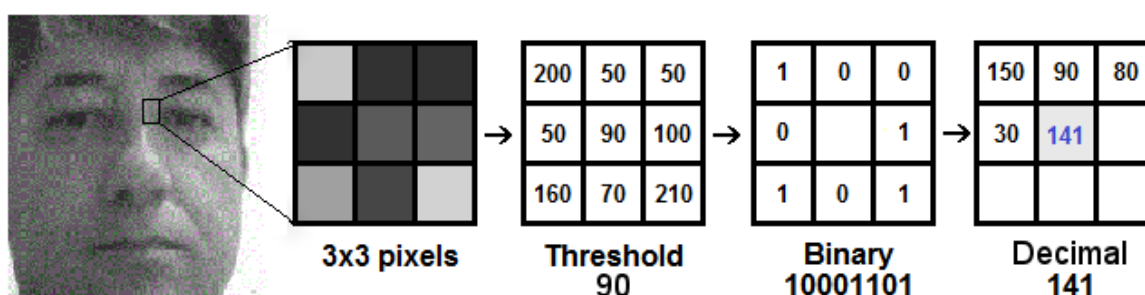
O algoritmo Histograma de Padrões Binários (LBPH) é amplamente utilizado no reconhecimento facial, devido a simplicidade e poder classificativo. Ele é baseado no Operador binário local (LBP) que rotula os pixels de uma imagem limitando a vizinhança de cada pixel e considera o resultado como um número binário.

O LBPH possui quatro parâmetros:

- **Radius:** usado para construir um padrão binário circular local e representa o raio ao redor do pixel central;
- **Neighbors:** o número de pontos de amostra para construir o padrão binário local circular;
- **Grid X:** número de células na direção horizontal;
- **Grid Y:** número de células na direção vertical.

Uma das primeiras etapas computacionais do LBPH é criar uma imagem intermediária que descreva a imagem original de uma maneira melhor, destacando as características faciais. A partir de uma imagem em escala de cinza podemos obter parte dessa imagem como uma janela de 3x3 pixels, conforme imagem abaixo:

Figura 12 - Aplicação da operação LBP.

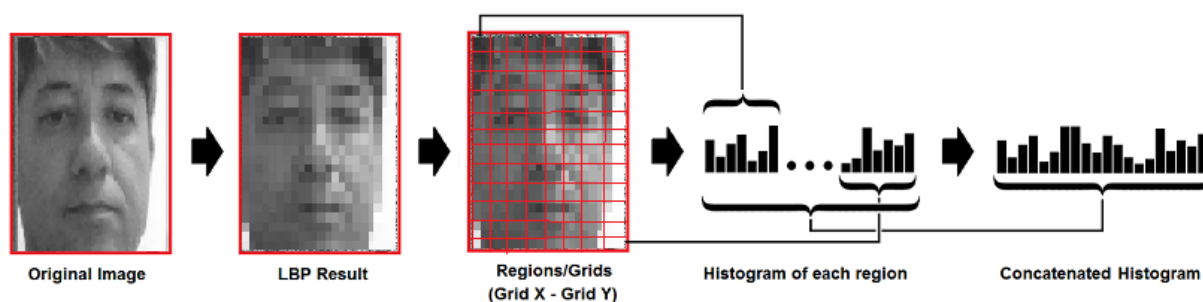


Usando o valor do pixel central como limite, ele compara um pixel com os 8 pixels mais próximos. Se o valor do vizinho for maior ou igual ao valor central, ele será definido como 1, caso contrário, será definido como 0. Assim, obtemos um total de 8 valores binários dos 8

vizinhos. Em seguida, o valor binário é convertido em um valor decimal e o definimos como o valor central da matriz, que na verdade é um pixel da imagem original.

No final deste procedimento, será obtida uma nova imagem que representa melhor as características da imagem original. A partir da imagem gerada na última etapa, podemos usar os parâmetros Grade X e Grade Y para dividir a imagem em várias grades e extrair os histogramas de cada região, como pode ser visto na imagem a seguir:

Figura 13 - Histogramas de Padrões Binários Locais



Após a criação do histograma para cada região, todos os histogramas são mesclados para formar um único histograma. O resultado é então concatenado em um vetor que será utilizado pelo classificador na detecção de faces.

4.3 VALIDAÇÃO DO PROTÓTIPO

Para fins de análise de desempenhos dos algoritmos Eigenfaces, Fisherfaces e LBPH foram utilizados o banco de imagens do pacote Yale Face, amplamente difundido na internet, como mostrado na figura abaixo:

Figura 14 - Pacote de faces Yale



Foram utilizadas para treinamento 135 imagens, sendo 9 imagens de cada indivíduo, totalizando 15 pessoas. Na fase de treinamento foram utilizados os parâmetros PCA, Threshold, Radius, Grid X e Grid Y. Para obtenção de uma melhor acurácia nos resultados, foram alterados os valores para maior quanto para menor do valor padrão dos parâmetros.

4.3.1 Eigenfaces

A tabela abaixo mostra os resultados obtidos nos testes com o algoritmo Eigenfaces. A primeira coluna representa a porcentagem dos acertos obtidos nos testes, representada pela média aritmética simples do número de acertos com o somatório do número das faces. A segunda coluna representa o valor da confiança, representada basicamente pelo valor da distância para encontrar a face mais semelhante no conjunto de dados, ou seja, quanto menor o valor da distância maior é o nível de certeza que o algoritmo possui sobre o reconhecimento. Por fim, a terceira coluna temos os valores dos parâmetros utilizados.

Tabela 1- Resultados do teste do algoritmo Eigenface

Eigenfaces		
Acertos %	Confiança	Parâmetros
70.0	6860.46	Padrão
50.0	3611.65	40, 8000
60.0	4549.91	45, 9000
63.33	4826.14	45, 10000
66.66	5197.90	45, 20000

66.66	5197.90	45, 50000
56.66	3719.85	30, 8000
40.0	3996.87	134, 8000
33.33	3314.64	134, 7000
33.33	3314.64	134, 6000
6.66	1470.06	134, 2000
70.00	6860.46	134, 100000
70.00	6860.46	134, 200000
70.0	6860.46	140, 100000
70.0	6860.46	150, 100000
40.00	3996.87	150, 8000
70.0	6452.80	100, 100000

Conforme tabela acima, nos testes com o algoritmo Eigenfaces o melhor resultado de acertos obtidos foi 70%, com valores dos parâmetros: números de PCA variando entre 134 à 150 e Threshold 100.000.

4.3.2 Fisherfaces

Nos testes com o algoritmo Fisherfaces, o melhor resultado de acertos obtidos foi 96.66%, com valores dos parâmetros: números de PCA 13 e Threshold 8.000.

Tabela 2- Resultados do teste do algoritmo Fisherfaces

Fisherfaces		
Acertos	Confiança	Parâmetros
86.66	2259.16	Padrão
56.66	403.91	3, 2000
70.0	655.68	4, 2000
40.0	847.47	15, 2000
86.66	2259.16	15, 8000
86.66	2259.16	15, 7000
86.66	2259.16	15, 6000
93.33	2148.95	14, 8000
93.33	2148.95	14, 9000
93.33	2148.95	14, 10000
93.33	2148.95	14, 50000
96.66	1973.94	13, 8000
96.66	1894.92	12, 8000
96.66	1810.18	11, 8000
93.33	1645.19	10, 8000
93.33	1645.19	10, 7000
93.33	1645.19	10, 6000

4.3.3 LBPH

Nos testes com o algoritmo LBPH, o melhor resultado de acertos obtidos foi 60%, com valores dos parâmetros: Radius 1, Neighbors 8, Grad X 8, Grad Y 8 e Threshold 50.

Tabela 3- Resultados do teste do algoritmo Eigenface

LBPH		
Acertos %	Confiança	Parâmetros
60.0	11.61	Padrão
60.0	1.25	2, 2, 7, 7, 50
60.0	3.15	3, 3, 7, 7, 50
60.0	11.61	<u>1, 8, 8, 8, 50</u>
60.0	11.61	1, 8, 8, 8, 40
60.0	7.53	1, 8, 7, 7, 50
60.0	4.53	1, 8, 6, 6, 50
60.0	15.71	2, 8, 8, 8, 50
56.66	17.32	3, 8, 8, 8, 50
56.66	18.98	4, 8, 8, 8, 50
<u>56.66</u>	20.78	5, 8, 8, 8, 50
29.99	40.0	50, 8, 8, 8, 50
26.66	30.08	100, 8, 8, 8, 50
60.0	13.84	1, 10, 8, 8, 50
60.0	13.84	1, 10, 8, 8, 54
cv::OutOfMemoryError		<u>1, 20, 8, 8, 50</u>
60.0	20.33	<u>1, 14, 8, 8, 50</u>

4.4 ANÁLISE DOS RESULTADOS

Finalmente diante dos resultados obtidos podemos considerar que o algoritmo com melhor desempenho no reconhecimento das faces foi o Fisherfaces com valor de 96.6% de acertos. Porém, apesar de ter obtido um valor menor que o Fisherfaces no número de acertos, o LBPH com 60% de acertos saiu-se melhor no item confiança. O algoritmo Eigenfaces com valor 70% de assertividade obteve o pior resultado dos três algoritmos no item confiança. Assim, cabe ao usuário avaliar o grau de importância desejada e optar entre um algoritmo com maior assertividade ou maior confiança.

5 CONSIDERAÇÕES FINAIS

No início do trabalho, um dos grandes objetivos era apresentar a relevância do uso das técnicas de *Machine Learning*, *Deep Learning* e Inteligência Artificial aplicadas ao reconhecimento facial. Constata-se, após a realização da pesquisa, que o uso destas novas tecnologias têm proporcionado grandes avanços na área de reconhecimentos de padrões.

Estabeleceu-se também como objetivo, avaliar o grau de acurácia dos algoritmos Eigenface, Fisherface e LBPH da biblioteca OpenCV.

Percebe-se que todos os objetivos foram atingidos pois para tais fins, o presente trabalho apresentou um rico estudo bibliográfico, assim como, o desenvolvimento de um protótipo de software com a capacidade de identificar faces a partir de banco de imagens do pacote Yale Face.

A partir da análise dos resultados obtidos através do protótipo, foi possível comparar os algoritmos e constatar que o algoritmo Fisherfaces obteve maior assertividade no reconhecimento das faces, porém com menor taxa de confiança que o LBPH.

Diante da metodologia definida, ou seja, uma pesquisa com abordagem qualitativa e executada por meio de levantamento bibliográfico, nos deparamos com a dificuldade em encontrar publicações literárias relacionadas ao tema em língua portuguesa. O tempo também foi um grande limitador para o aprofundamento dos trabalhos de pesquisa e desenvolvimento do projeto de software, visto que, as tecnologias abordadas envolvem assuntos complexos e inovadores.

5.1 TRABALHOS FUTUROS

Dado a importância no desenvolvimento de equipamentos com tecnologia de reconhecimento facial cada vez mais confiáveis, faz-se necessário um estudo mais amplo sobre o reconhecimento facial aplicando a tecnologia 3D, assim como, a análise de desempenho de outras bibliotecas openSources e finalmente a utilização de uma câmera externa para testar o funcionamento do software de reconhecimento facial no dia a dia, onde inclui a detecção de faces em condições especiais com variação de ângulos e iluminações.

REFERÊNCIAS

- BRADSKI, Gary; KAEBLER, Adrian. **Aprendendo OpenCV: visão de computador com a biblioteca opencv**. Visão de computador Com a biblioteca openCV. 2008. Disponível em: https://books.google.com.br/books?id=seAgiOfu2EIC&pg=PA16&dq=openCV&hl=pt-BR&sa=X&ved=0ahUKEwiEvM6J2_ToAhVBIrkGHbKVDAQAQ6AEIKzAA#v=onepage&q=openCV&f=false. Acesso em: 19 abr. 2020.
- BRAGA, Antônio de Pádua; LUDERMIR, Teresa Bernarda; CARVALHO, André Carlos Ponce de Leon. **Redes Neurais Artificiais: Teoria e aplicações**. Rio de Janeiro: Ltc, 2000.
- COPPIN, Ben. **Inteligência Artificial**. Rio de Janeiro: Ltc, 2013.
- DATA SCIENCE CADEMY (Org.). **Introdução à Ciência de Dados 2.0**. Disponível em: <https://www.datascienceacademy.com.br/>. Acesso em: 01 nov. 2019.
- DATA SCIENCE CADEMY (Org.). **Python Fundamentos para Análise de Dados**. Disponível em: <https://www.datascienceacademy.com.br/>. Acesso em: 20 out. 2019.
- DEEP Learning Book. Disponível em: <http://deeplearningbook.com.br/o-que-sao-redes-neurais-artificiais-profundas/>. Acesso em: 03 set. 2019.
- DEEP LEARNING BOOK (Org.). **Campos Receptivos Locais em Redes Neurais Convolucionais**. Disponível em: <http://deeplearningbook.com.br/campos-receptivos-locais-em-redes-neurais-convolucionais/>. Acesso em: 17 nov. 2019.
- DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO UNIVERSIDADE DE STANFORD (Stanford). **Redes neurais convolucionais para reconhecimento visual**. 2020. Disponível em: <http://cs231n.github.io/neural-networks-1/>. Acesso em: 24 maio 2020.
- EQUIPE OPENCV (org.). **OpenCV**. Disponível em: <https://opencv.org/>. Acesso em: 18 abr. 2020.
- EQUIPE OPENCV. **OpenCV: Visão computacional de código aberto**. 2020. Disponível em: <https://docs.opencv.org/4.3.0/d1/dfb/intro.html>. Acesso em: 20 maio 2020.
- FACELI, Katti et al. **Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina**. Rio de Janeiro: Ltc, 2011.
- GIL, Antonio Carlos. **Como Elaborar Projetos de Pesquisa**. 5ª ed. São Paulo: Atlas, 2010.
- GOODFIRMS (Org.). **As 8 melhores soluções de software de detecção de rosto de código aberto e gratuito**. Sarah Rose Miller. Disponível em: <https://www.goodfirms.co/blog/best-free-open-source-face-detection-software-solutions>. Acesso em: 07 dez. 2019.
- KLEINA, Nilton. **Como funcionam os sistemas de reconhecimento facial**. 2011. Disponível em: <https://www.tecmundo.com.br/camera-digital/10347-como-funcionam-os-sistemas-de-reconhecimento-facial.htmNil>. Acesso em: 17 nov. 2019.
- KLONTZ, Joshua C. et al. **Reconhecimento Biométrico de Código Aberto**. Disponível em: <http://openbiometrics.org/publications/klontz2013open.pdf>. Acesso em: 08 dez. 2019.

IBM, **Artificial Intelligence in Logistics**. DHL Customer Solutions & Innovation, Troisdorf, Germany, 2018.

LI, Stan Z.; JAIN, Anil K. (Ed.). **Manual de Reconhecimento Facial**. A: Springer, 2004. Disponível em: <<https://epdf.pub/queue/handbook-of-face-recognition6b3f024d0c27687be20b89933aa7d7b780895.html>>. Acesso em: 02 nov. 2019.

LUGER, GEORGE F. - **Inteligência Artificial: Estruturas e Estratégias para a Solução de Problemas Complexos** - 4ª Edição -Bookman Companhia Editora – 2004

MAFEI, Rafael. **Tecnologias de reconhecimento facial são usadas em 37 cidades no país**. 2019. Disponível em: <https://agenciabrasil.ebc.com.br/geral/noticia/2019-09/tecnologias-de-reconhecimento-facial-sao-usadas-em-37-cidades-no-pais>. Acesso em: 21 jun. 2020.

MARCONI, Marina de Andrade; LAKATOS, Eva Maria. **Metodologia Científica**. 5. ed. São Paulo: Atlas, 2011

MIT 6.S191. **Introdução ao Deep Learning**. 2020. Disponível em: <http://introtodeeplearning.com/>. Acesso em: 24 maio 2020.

ROSA, João Luís Garcia. **Fundamentos da Inteligência Artificial**. Rio de Janeiro: Ltc, 2014.

SILVA, Mary Aparecida Ferreira da. **Modelos e técnicas de pesquisa**. 2. ed. Curitiba: Ibepex, 2005. 263 p. Disponível em: <<https://plataforma.bvirtual.com.br/Leitor/Publicacao/1493/pdf>>. Acesso em: 17 nov. 2019.

TECTUDO (Org.). **Facebook melhora reconhecimento facial e chega quase ao nível humano**. 2014. Milena Pereira. Disponível em: <<https://www.techtudo.com.br/noticias/noticia/2014/03/facebook-melhora-reconhecimento-facial-e-chega-quase-ao-nivel-humano.html>>. Acesso em: 07 dez. 2019.