UNIVERSIDADE PAULISTA - UNIP

CIÊNCIA DA COMPUTAÇÃO

Allan de Sousa Braga

Leticia Gomes

PEDRO HENRIQUE SILVA DIAS

Rhuan Alves Meira Santos

A melhoria de atendimento com base na detecção de emoções via reconhecimento facial para a evolução das empresas

SÃO paulo

2022

Allan de Sousa Braga

Leticia Gomes

PEDRO HENRIQUE SILVA DIAS

Rhuan Alves Meira Santos

A melhoria de atendimento com base na detecção de emoções via reconhecimento facial para a evolução das empresas

Trabalho de Conclusão de Curso apresentado ao Curso de Ciência da Computação da Universidade Paulista, como requisito parcial para obtenção do grau de Graduação em Ciência da Computação.

Orientador: Marco Gomes

SÃO PAULO

2022

RESUMO

Dado o mercado atual, é perceptível a crescente necessidade, por parte das pequenas, médias e grandes empresas, do aperfeiçoamento contínuo no serviço de atendimento ao consumidor, tendo isso em vista, é de suma importância que seja realizada uma retenção e ampla análise de todos os dados gerados pelos atendimentos prestados aos clientes, dessa maneira, buscando obter meios de inicialização à otimização de serviços prestados à consumidores em instituições públicas e privadas que possuam atendimentos presenciais.

Dado o exposto, com o auxílio de Inteligência Artificial integrada à um componente físico de Internet Das Coisas (IoT), temos como proposta realizar a captura de imagens e interpretação de micro expressões faciais expostas pelo usuário no decorrer de seu atendimento, atingindo o objetivo de ter a capacidade de análise de microexpressões faciais para determinação de emoções do consumidor em relação à seu atendimento com maior acurácia, dados estes, que podem ser utilizados para fins de qualidade.

**Palavras-Chaves:** Emoções, Microexpressões, Inteligência Artificial, IOT, Reconhecimento Emocional.

ABSTRACT

Given the current market, it is noticeable the growing need, on the part of small, medium, and large companies, for continuous improvement in customer service. the data generated by the services provided to customers, thus seeking to obtain means of initiating the optimization of services provided to consumers in public and private institutions that have face-to-face services.

Given the above, with the help of Artificial Intelligence integrated into a physical component of the Internet of Things (IoT), we propose to capture images and interpret micro facial expressions exposed by the user during their service, achieving the objective of have the ability to analyze facial microexpressions to determine consumer emotions in relation to their service with greater accuracy, data that can be used for quality purposes.

**Keywords**: Emotions, Microexpressions, Artificial Intelligence, IoT, Emotional Recognition.

LISTA DE SIGLAS

FACS: *Facial Action Coding System*;

IA: Inteligência Artificial;

ADA: *American Psychological Association*;

EMFACS: *Emotion Facial Action Coding System*;

AU: *Action Units*;

UA: Unidades de Ação;

RNA: Redes Neurais Artificiais;

CNN: *Convolutional Neural Network*;

RNC: Rede Neural Convolucional;

CPU: *Central Processing Unit*;

GPU: *Graphics Processing Units*;

ASIC: *Application-Specific Integrated Circuit*;

TPU: *Tensor Processing Units*;

NN*: Neural networks*;

IDE: *Integrated Development Environment*;

AWS: *Amazon Web Services*;

ECS: *Elastic Container Service*;

ECR: *Elastic Container Registry*;

RDS: *Relational Database Service*;

RAM: *Random Access Memory*;

SGBD: Sistema de Gerenciamento de Banco de Dados;

3D: Tridimensionalidade;

API: *Application Programming Interface*;

REST: *Representational State Transfer*;

LGPD: Lei Geral de Proteção de Dados;

IoT: *Internet of Things*;

UX: *User Experience*;

UI: *User Interface*;

CONTECSI: *International Conference on Information Systems and Technology Management*;

LED: *Light Emitting Diode*;

SD: *Secure Digital*;

Mb: Megabytes;

Kb: Kilobytes;

ROM: *Read Only Memory*;

RAM: *Random Access Memory*;

SRAM: *Static* *Random Access Memory*;

BIT: *Binary Digit*;

PWM: *Pulse Width Modulation*;

EEPROM: *Electrically Erasable Programmable Read-Only Memory*.

LISTA DE figuras

[Figura 1 - Expressões universais por Paul Ekman e Wallace Friesen 21](#_Toc103367983)

[Figura 2 - Características que representam a microexpressão de Surpresa 22](#_Toc103367984)

[Figura 3 - Características que representam a microexpressão de Medo 23](#_Toc103367985)

[Figura 4 - Características que representam a microexpressão de Aversão 24](#_Toc103367986)

[Figura 5 - Características que representam a microexpressão de Raiva 25](#_Toc103367987)

[Figura 6 - Características que representam a microexpressão de Felicidade 26](#_Toc103367988)

[Figura 7 - Características que representam a microexpressão de Tristeza 26](#_Toc103367989)

[Figura 8 - Características que representam a microexpressão de Desprezo 27](#_Toc103367990)

[Figura 9 - Modelo de neurônio 30](#_Toc103367991)

[Figura 10 - Funções de ativação 30](#_Toc103367992)

[Figura 11 - Rede Neural Profunda 31](#_Toc103367993)

[Figura 12 - GoogLeNet 32](#_Toc103367994)

[Figura 13 - Top 5 classificações ImageNet 33](#_Toc103367995)

[Figura 14 - Filtros de Camadas Convolucionais 34](#_Toc103367996)

[Figura 15 - Filtros na Detecção de Objetos 35](#_Toc103367997)

[Figura 16 - Diagrama de integração de componentes do projeto 47](#_Toc103367998)

[Figura 17 - Desempenho e Classificação do Aprendizado de Máquina 50](#_Toc103367999)

LISTA DE tABELAS

[Tabela 1 - Requisitos funcionais 41](#_Toc103368000)

[Tabela 2 - Requisitos não-funcionais 42](#_Toc103368001)

Sumário

[1 INTRODUÇÃO 12](#_Toc113384381)

[1.1 Motivação 12](#_Toc113384382)

[1.2 Objetivos 14](#_Toc113384383)

[1.2.1 Objetivo Geral 14](#_Toc113384384)

[1.2.2 Objetivo Específico 14](#_Toc113384385)

[2 desenvolvimento 15](#_Toc113384386)

[2.1 Procedimentos metodológicos 15](#_Toc113384387)

[2.1.1 Caracterização da pesquisa 15](#_Toc113384388)

[2.1.2 Proposta da solução 15](#_Toc113384389)

[2.1.3 Delimitações 15](#_Toc113384390)

[2.1.4 TRABALHOS RELACIONADOS 15](#_Toc113384391)

[2.2 Fundamentação teórica 17](#_Toc113384392)

[2.2.1 O estudo das emoções 17](#_Toc113384397)

[2.2.2 Expressões faciais 18](#_Toc113384398)

[2.2.3 Microexpressões e FACS 19](#_Toc113384399)

[2.2.3.1 FACS 19](#_Toc113384400)

[2.2.3.2 Microexpressões 20](#_Toc113384401)

[2.2.4 Inteligência Artificial 27](#_Toc113384402)

[2.2.4.1 Machine Learning 28](#_Toc113384403)

[2.2.4.2 Rede NeuraL Artificial 28](#_Toc113384404)

[2.2.4.3 Redes Neurais Convolucionais 31](#_Toc113384405)

[2.2.5 OpenCV 36](#_Toc113384406)

[2.2.6 TensorFlow 36](#_Toc113384407)

[2.2.7 Keras 37](#_Toc113384408)

[2.3 Materiais e Métodos 38](#_Toc113384409)

[2.3.1 Ferramentas Utilizadas 38](#_Toc113384410)

[2.3.1.1 Microsoft Visual Studio 38](#_Toc113384411)

[2.3.1.2 .Net core 3.1 38](#_Toc113384412)

[2.3.1.3 C# (CSharp) 38](#_Toc113384413)

[2.3.1.4 Amazon Web Services 39](#_Toc113384414)

[2.3.1.5 Docker 39](#_Toc113384415)

[2.3.1.6 Google Colab 40](#_Toc113384416)

[2.3.1.7 Jupyter notebooks 40](#_Toc113384417)

[2.3.1.8 MySql 40](#_Toc113384418)

[2.3.1.9 TinkerCad 40](#_Toc113384419)

[2.3.2 Requisitos de Software 40](#_Toc113384420)

[2.3.2.1 Requisitos Funcionais 41](#_Toc113384421)

[2.3.2.2 Requisitos Não-Funcionais 42](#_Toc113384422)

[2.3.3 Desenvolvimento 43](#_Toc113384423)

[2.3.3.1 Dataset 43](#_Toc113384424)

[2.3.3.2 Hardware 43](#_Toc113384425)

[2.3.3.3 API 45](#_Toc113384426)

[2.3.3.4 Diagrama de Integração 47](#_Toc113384427)

[2.4 Resultados 50](#_Toc113384428)

[3 CONSIDERAÇÕES FINAIS 51](#_Toc113384429)

[REFERÊNCIAS BIBLIOGRáFICAS 52](#_Toc113384430)

# 1 INTRODUÇÃO

É perceptível que a qualidade de atendimento das empresas perante seus consumidores é uma pauta importante a ser discutida. Bom atendimento é mais importante para consumidor do que qualidade, revela pesquisa. Revista Pequenas Empresas & Grandes Negócios, 08 dez. 2009. Com base nisso, pode-se afirmar que a qualificação profissional de colaboradores é o caminho de sucesso para as empresas, pois a fidelização do cliente se torna uma consequência, a partir do momento em que sua experiência consiste em um bom atendimento, afinal, o contágio emocional, fenômeno que ocorre quando as emoções e o comportamento de uma pessoa afetam as emoções e o comportamento de outra, é uma manifestação multiplicadora que envolve atributos psicológicos, fisiológicos, comportamentais e social de forma inconsciente e automática. Dessa maneira, desencadeando a sincronia emocional, de atenção e comportamento (HATFIELD, CACIOPPE E RAPSON, 1994).

Através da constante busca de excelência nos processos de atendimento ao público, foi observado que as expressões faciais inconscientes durante o decorrer de um atendimento, são ótimas fontes de feedbacks que os clientes fornecem. Sob esse viés, foi idealizado um sistema para análise de expressões, baseado em reconhecimento facial com propósito de detecção de emoções capaz de definir a experiência do usuário em um atendimento.

## 1.1 Motivação

A capacidade de processamento dos computadores nos últimos anos aumentou a um nível que possibilitou criar e aprimorar técnicas para auxiliar na resolução de problemas, uma destas técnicas é a aprendizagem de máquina, Machine Learning, mas outro fator que também colocou a aprendizagem de máquina em evidência foi a facilidade de utilizar de ambientes remotos para processar uma grande quantidade de dados sem precisar ter uma estrutura própria. Esses ambientes remotos facilitaram o acesso de pessoas com algum interesse ou de outras áreas que não a computação. Empresas de menor porte também utilizam esse tipo de ambiente para criar e testar novas soluções, facilitando o acesso que antes necessitava de máquinas com grande poder de processamento para iniciar em certos assuntos que exigem um pouco mais.

O atendimento ao cliente sempre traz inúmeras maneiras de avaliar os serviços prestados e buscar indicadores de problemas em produtos e serviços, comumente sendo representado pela pesquisa de satisfação no final do atendimento, o que se mostra não sendo muito prático e confiável, trazendo muitas vezes uma burocracia extra indesejada. Com a implementação de um sistema de análise e reconhecimento de emoções existe a possibilidade de evitar alguns destes problemas.

As emoções expressadas em certas situações mudam conforme o contexto social em que o indivíduo está inserido. Porém, algumas expressões faciais podem ser consideradas como sendo universais (EKMAN, 2003). Devido a este aspecto as emoções expressadas são uma fonte bem confiável de dados, podendo até serem utilizadas em conjunto com outras formas de avaliação, dando uma ótima base para criar relatórios para análise de atendimento, produtos e serviços.

## 1.2 Objetivos

### 1.2.1 Objetivo Geral

Avaliar o desempenho de atendimentos presenciais ao público com a identificação de emoções baseado no reconhecimento facial por meio de Inteligência Artificial, sobre uma arquitetura de rede neural treinada para a identificação de microexpressões da face humanas, visando o aperfeiçoamento do atendimento e satisfação do consumidor com base em sua experiência prévia.

### 1.2.2 Objetivo Específico

Analisar as características que permitem identificar qual modelagem apresenta melhor resultado em relação às emoções abordadas.

Considerando o desenvolvimento do trabalho e o objetivo geral apresentado, destacam-se os seguintes objetivos específicos:

* Utilizar o reconhecimento facial para detecção das emoções expressadas fundamentado nas imagens da base. As emoções presentes na base são: neutro, felicidade, tristeza, raiva, medo, aversão e surpresa;
* Realizar uma melhoria nos atendimentos ao consumidor utilizando microexpressões e Facial Action Coding System (FACS), que permite identificar características importantes nas expressões faciais humanas;
* Compreender qual emoção foi expressa pelo usuário durante seu atendimento;
* Treinar uma inteligência artificial em Python com Colab Notebook para reconhecer emoções;

# 2 desenvolvimento

## 2.1 Procedimentos metodológicos

Metodologia de pesquisa é o estudo de métodos e técnicas, e linha de seguimento a ser percorrida dentro de um estudo. O presente artigo tem como abordagem o método de pesquisa qualitativa, pois tem como base a aplicação de dados não mensuráveis e para fins descritivos.

### 2.1.1 Caracterização da pesquisa

Essa pesquisa qualitativa procura entender o comportamento do consumidor, com base em suas propriedades e experiências individuais.

O estudo foi produzido para conhecimento teórico a respeito da relevância dos sentimentos e emoções expressadas diante de uma situação, mas também para propiciar uma solução de unificação para alinhar esses sentimentos aos resultados de satisfação do público ao atendimento prestado por um estabelecimento.

### 2.1.2 Proposta da solução

A proposta para o presente artigo é o uso de uma câmera para captar as expressões faciais de um consumidor após a prestação de um atendimento, em distintas iluminações e ângulos onde serão encaminhadas para a IA (Inteligência Artificial) analisar os sentimentos expressos em sua face, detectando suas emoções como fonte de feedback ao atendimento realizado.

### 2.1.3 Delimitações

O artigo delimitou-se a análise de identificação facial de emoções com a utilização de uma IA (Inteligência Artificial) e idealização de projeto de hardware e software.

### 2.1.4 TRABALHOS RELACIONADOS

A mais de uma década, a tecnologia de análise de expressões faciais é de interesse das empresas e pesquisadores, em 2012 a rede de supermercados Walmart registrou uma patente referente a reconhecimento de emoções para otimizar a experiencia dos clientes durante as compras, analisando se há insatisfação no caixa e caso ocorra, entrar em contato para analisar os problemas relacionados.

Tendo visto que empresas do porte do Walmart estariam entrando neste mercado, apareceram empresas que fornecem serviços relacionados ao mercado de análise facial, dentre elas a NVISO, que fornece uma solução de analisar em tempo real as reações do usuário enquanto ele usa algum aplicativo através da câmera do dispositivo utilizado para o acesso.

Outra abordagem interessante é o protótipo publicado na revista H-TEC que busca utilizar a análise facial para ajudar familiares e cuidadores de pessoas com mobilidade ou fala reduzida a identificar problemas como desconforto, dor e afins utilizando uma webcam.

Um trabalho publicado na CONTECSI de 2020, tem como objetivo capturar e analisar as expressões faciais dos colaboradores em tempo real durante a jornada de trabalho. para facilitar decisões nas avaliações e embasar possíveis encaminhamentos para profissionais especializados.

## Fundamentação teórica



### 2.2.1 O estudo das emoções

As emoções são comportamentos que os seres humanos vivenciam em resposta de acontecimentos e contextos identificadas pelo rosto por meio das expressões faciais. Segundo a ADA, American Psychological Association, o tipo de emoção que uma pessoa transmite é determinado pela condição que desencadeia a emoção, por exemplo, uma pessoa se sente contente ao ser bem recebida ou se sente triste ao não ter atenção. O desencadeio de emoções se dá a experiencia de ter uma reação emocional em mídia ou ambiente social. As emoções governam diariamente as vidas das pessoas, visto que decisões são tomadas baseadas em emoções momentâneas, estando felizes, tristes, entediados, zangados ou frustrados. Entender as emoções pode ajudar a ter uma vida com maior facilidade e estabilidade, pois hobbies e atividades, por exemplo, são escolhidos com base nas emoções que eles provocam.

Segundo o livro *Discovering Psychology* (1990), a emoção é um estado psicológico complexo que envolve três componentes distintos: em primeiro lugar a experiência subjetiva, que é a essência de uma emoção, e que as manifestações objetivas no comportamento e na fisiologia do corpo ou do cérebro são, na melhor das hipóteses, indicadores indiretos dessas experiências internas; além dessa, a resposta fisiológica, que se trata de reações do corpo experimentadas durante uma situação, como taquicardia ou mãos suadas; e pôr fim a resposta comportamental ou expressiva, a expressão real da emoção, como um sorriso para indicar felicidade ou um olhar baixo para indicar tristeza.

Neste sentindo, existem dois tipos de comunicação: a verbal, que talvez seja o modo de comunicação mais óbvio e poderoso, a partir do compartilhamento de informações entre indivíduos usando palavras, sejam ditas ou escritas; por outro lado, a comunicação não verbal inclui expressões faciais, gestos, paralinguística, como volume, velocidade ou tom de voz, linguagem corporal, olhar, toque e aparência. Essa última pode gerar diversas interpretações, indicando a individualidade de cada pessoa envolvida e suas distintas experiências. Sendo assim, ao avaliar a comunicação não-verbal, especificamente à linguagem corporal, é interessante atentar-se a alguns sinais que a fisionomia, em geral, possa transmitir.

### 2.2.2 Expressões faciais

A Expressão facial se trata de um dos mais importantes aspectos para eventual entendimento do ser humano. A série “Lie to Me” é um exemplo de como pode-se utilizar esse meio para identificar uma possível mentira, por meio de investigações juntamente à equipe especialista em identificar essa inverdade, a partir da leitura de gestos e microexpressões faciais de cada suspeito. Isso porque essas expressões são maneiras que o ser humano emprega para revelar suas emoções e sentimentos, de modo voluntário ou involuntário, de acordo com o jeito de cada indivíduo.

A diferença na expressão de uma pessoa que recebeu um presente agradável e, por outro lado, a de um motorista cujo carro tenha sido abatido no trânsito é perceptível, isso porque são maneiras do corpo, consequentemente, o rosto, reagir a determinada situação em que está inserido - juntamente às reações comportamentais da fisiologia de cada um.

Segundo Ekman (2003) e Plutchik (2002), essa situação possa ser explicada pelo simples fato de a face ser a região corporal com maiores mecanismos para expressão do indivíduo, ou seja, o conjunto de todos os músculos dessa área são essenciais e responsáveis pela própria revelação e demonstração das expressões de cada um, conforme a situação em que está inserido.

Em *The Expression of the Emotions in Man and Animals*, se defende a tese de que todos os humanos e até mesmo outros animais, demonstram emoções por meio de comportamentos semelhantes (DARWIN, Charles, 1872). Hoje, muitos psicólogos concordam que certas emoções são universais para todos os seres humanos, independentemente da cultura: raiva, medo, surpresa, aversão, felicidade e tristeza. Tendo isso em vista, é importante partir do pressuposto que as emoções são as principais a serem investigadas, a partir delas, pode-se analisar suas derivadas.

Isso acontece porque existem diversas maneiras de reconhecimento facial que possibilitam a identificação das expressões humanas, tais como: curvatura da boca, presença ou não de sorriso, altura e posição das sobrancelhas, franzimento da testa, direção dos olhos. Na junção de todos esses aspectos, é possível identificar o que uma pessoa pode estar sentindo, permitindo que ela expresse quaisquer emoções - por meio de ações musculares que diferenciem particularmente uma emoção da outra.

A tecnologia é um grande avanço nesse aspecto, já que muitas empresas com influência e expertise utilizam a ferramenta de identificação facial para suas respectivas finalidades - cadastro de perfil na base de clientes, desbloqueio de celular, pagamento em bancos, até mesmo identificação de suspeitos, dentre outras. Nesse sentido, para o andamento do trabalho, para haver uma maior fidelização de clientes em uma determinada empresa cujo serviço fora fornecido, é importante a análise das expressões faciais destes, com a finalidade de reconhecer seu grau de satisfação com base no atendimento que lhe foi prestado - com o respectivo consentimento do cliente e dentro de todas medidas legais.

### 2.2.3 Microexpressões e FACS

Para aprofundarmos sobre microexpressões faciais, é importante antes, entendermos que para ler e compreender rostos e expressões faciais adequadamente, é preciso usar procedimentos de análise científica. Com esse propósito, foi criado o Facial Action Coding System (FACS), uma metodologia científica criada pelo Dr. Paul Ekman e Wallace Friesen, no ano de 1978, para que pudessem descrever as expressões faciais humanas, e a conversão de FACS para Emotion FACS (EMFACS) baseado em pontuações universais do estado emocional para emoções básicas.

#### 2.2.3.1 FACS

FACS é um sistema de análise e codificação de rosto desenvolvido por Paul Ekman e Wallace Friesen. Como o único método reconhecido e usado globalmente, o FACS não apenas reconhece e analisa todos os aspectos dos movimentos faciais, mas também decodifica a codificação de marcadores, até alcançar o EMFACS (Emotional FACS), para identificar 7 emoções básicas e generalizadas de qualquer pessoa, em qualquer lugar do mundo (alegria, raiva, medo, tristeza, aversão, desprezo e surpresa).

Desde a sua criação, este método científico foi tecnicamente aprimorado e implementado de forma a auxiliar profissionais de diversas áreas a obter a maior precisão possível na leitura de microexpressões, pois pode identificar qualquer expressão isolando a unidade de movimentos musculares. Em termos simples, o sistema permite o uso de códigos (Unidades de Ações - AU) para leitura de microexpressões faciais. O EMFACS é um núcleo de códigos usado para realizar a interpretação das pontuações, com base em AUs.

As *Action Units*, são as Unidades de Ações, que tem como sigla o famoso AU, as Unidades de Ações são independentes de quaisquer interpretações, podendo ser utilizadas em conjunto, o que possibilita na composição do reconhecimento facial, indo do básico às mais complexas, onde se torna necessário a utilização de softwares avançados para esse trabalho de identificação. É por meio da codificação facial que animadores de filmes e desenhos animados, criadores de jogos, câmeras de segurança, interrogatórios policiais, as inserções de reconhecimento facial em smartphones, por exemplo, conseguem realizar pesquisas e melhorar a experiência do usuário.

#### 2.2.3.2 Microexpressões

As microexpressões são expressões faciais muito breves e involuntárias que os humanos fazem quando experimentam emoções. Geralmente duram de 0,5 a 4,0 segundos, não podem ser falsificadas, são caracterizadas por movimentos involuntários, de baixa intensidade (em relação aos movimentos de contração muscular), muito rápidas e estão associados ao vazamento de emoções que se tenta esconder.

As microexpressões são a chave para ler rostos, em outras palavras, o povo de qualquer país do mundo fica tão triste quanto tribos isoladas pelo mundo, que nunca viram personagens de TV ou filmes para imitar suas expressões. O Dr. Ekman chegou a descobrir que pessoas cegas de nascença fazem expressões faciais sem nunca ter visto outra pessoa fazer.

Charles Darwin desenvolveu a primeira abordagem moderna do assunto (emoções e expressões faciais) em 1872, quando publicou sua pesquisa em seu livro “The Expression of Emotions in Man and Animals”. Neste trabalho, entre muitas contribuições, ele apresenta evidências de que certas expressões faciais e os músculos que as produzem, podem estar envolvidos no desenvolvimento de certas emoções. Mas essa história de ler expressões do rosto, se popularizou com as pesquisas do Dr. Paul Ekman, que identificou em seus estudos 7 expressões faciais universais.

Figura - Expressões universais por Paul Ekman e Wallace Friesen

Diagrama

Descrição gerada automaticamente

Fonte: Bez et al (2017)

##### 2.2.3.2.1 Surpresa

Embora a expressão seja breve, os movimentos faciais nos permitem observar o ambiente, mudar nosso foco para outro evento possivelmente ameaçador e reagir mais rapidamente. Seja sua causa boa ou ruim, a reação expressada pelo rosto será a mesma. As características faciais demonstradas ao sentirmos surpresa são as sobrancelhas inteiramente e pálpebras levantadas, boca aberta e pupilas dilatadas.

Figura - Características que representam a microexpressão de Surpresa

Diagrama

Descrição gerada automaticamente

Fonte: https://www.thejournal.ie/lying-2308154-Sep2015/ (2015)

##### 2.2.3.2.2 Medo

A microexpressão do medo está intimamente ligada ao choque, cada movimento nos prepara para uma resposta de luta ou fuga. Estes movimentos faciais tiram proveito da forma que nossos corpos funcionam. Alargar nossos olhos abre nosso campo de visão, deixando entrar mais luz e permite que vejamos as ameaças ao nosso redor. O mesmo pode ser dito para nossas vias de oxigênio, abrir nossas narinas aumenta nossa ingestão de oxigênio e nos ajuda a nos preparar para fugir ou lutar. As características faciais demonstradas ao sentirmos medo, são as sobrancelhas puxadas para cima e juntas, pálpebras superiores puxadas para cima e a boca esticada.

Figura - Características que representam a microexpressão de Medo

Diagrama

Descrição gerada automaticamente

Fonte: https://www.thejournal.ie/lying-2308154-Sep2015/ (2015)

##### 2.2.3.2.3 Aversão

Quando apertamos os olhos com desgosto, nossa acuidade visual aumenta, ajudando-nos a encontrar a origem de nossa aversão. A cara de nojo, não mostra apenas nossa aversão, também funciona como proteção. Enrugar o nariz fecha a passagem nasal, dessa maneira nos protegendo de gases perigosos e apertar os olhos protege de danos. As características faciais demonstradas ao sentirmos aversão/desgosto são as sobrancelhas puxadas para baixo, nariz enrugado, lábio superior puxado para cima e os lábios inferiores soltos.

Figura - Características que representam a microexpressão de Aversão

Homem de barba e bigode

Descrição gerada automaticamente

Fonte: https://www.thejournal.ie/lying-2308154-Sep2015/ (2015)

##### 2.2.3.2.4 Raiva

O rosto de raiva funciona muito bem, cada movimento desta expressão faz uma pessoa parecer mais forte, de acordo com pesquisadores. Esse rosto permite que a ameaça saiba que estamos falando sério. É uma das emoções mais poderosas e mostra o quão expressivo um rosto humano pode ser. Esse rosto serve como um aviso, seja simplesmente para intimidar ou para mostrar que um conflito começou. As características faciais demonstradas ao sentirmos raiva são as sobrancelhas puxadas para baixo, pálpebras superiores puxadas para cima, pálpebras inferiores tensionadas, margens dos lábios enroladas e os lábios podem ser apertados.

Figura - Características que representam a microexpressão de Raiva

Diagrama

Descrição gerada automaticamente com confiança média

Fonte: https://www.thejournal.ie/lying-2308154-Sep2015/ (2015)

##### 2.2.3.2.5 Felicidade

Apesar da conotação amigável, muitos pesquisadores têm a convicção de que o movimento por trás do sorriso possui uma origem sinistra. Geralmente, primatas exibem seus dentes para afirmar sua superioridade e garantir seu status. Alguns pesquisadores creem que de certa forma, esta pode ser uma forma ancestral do sorriso. As características faciais demonstradas ao ficarmos felizes são os músculos à volta dos olhos apertarem, rugas perto da parte externa dos olhos, bochechas levantadas e os cantos dos lábios levantados na diagonal.

Figura - Características que representam a microexpressão de Felicidade

Interface gráfica do usuário

Descrição gerada automaticamente com confiança média

Fonte: https://www.thejournal.ie/lying-2308154-Sep2015/ (2015)

##### 2.2.3.2.6 Tristeza

A tristeza é uma microexpressão difícil de se simular, de acordo com pesquisadores. Um dos indicadores de tristeza é a elevação da sobrancelha, que pouquíssimas pessoas podem forjar a pôr vontade própria. As características faciais demonstradas ao ficarmos tristes são os cantos internos das sobrancelhas levantados, pálpebras soltas e os cantos dos lábios puxados para baixo.

Figura - Características que representam a microexpressão de Tristeza

Diagrama

Descrição gerada automaticamente

Fonte: https://www.thejournal.ie/lying-2308154-Sep2015/ (2015)

##### 2.2.3.2.7 Desprezo

Embora o desprezo possa esconder a raiva e desconfiança, a microexpressão é singular. Podendo ser mais ou menos intensa, é a única expressão que se manifesta em apenas um lado do rosto. No seu ponto mais forte, uma sobrancelha pode abaixar enquanto a pálpebra inferior e o canto do lábio sobem do mesmo lado. Na sua forma mais secreta, o canto do lábio pode subir apenas brevemente. As características faciais demonstradas ao sentimos desprezo são os olhos neutros com o canto do lábio puxado para cima e para trás de um lado.

Figura - Características que representam a microexpressão de Desprezo

Foto preta e branca de rosto de homem visto de perto

Descrição gerada automaticamente

Fonte: https://www.thejournal.ie/lying-2308154-Sep2015/ (2015)

### 2.2.4 Inteligência Artificial

A inteligência artificial é um subconjunto do vasto domínio da ciência da computação que se dedica a resolver problemas em tempo real que estão relacionados às atividades humanas. Embasada na inteligência humana, a inteligência artificial é a atividade de máquina que consiste no comportamento inteligente de softwares baseados em uma estrutura de algoritmos que são capazes de racionalizar, absorver conhecimento, modificar, reconhecer padrões, memorizar e comunicar de maneira absolutamente lógica e não natural com o objetivo de pensar, tomar decisões e resolver problemas práticos do mundo real.

As máquinas inteligentes estão se tornando ferramentas fundamentais que promovem a existência e constante evolução das atividades sociais cotidianas. Atualmente, a inteligência artificial pode auxiliar em diagnósticos médicos, cirurgias, resgates, vigilância, segurança, entre outros cenários e a tendência é que evolua progressivamente, resolvendo problemas cada vez mais complexos.

O Machine Learning, traduzido como Aprendizado de Máquina, e Redes Neurais são domínios fundamentais para o desenvolvimento de softwares e hardwares que utilizam inteligência artificial.

#### 2.2.4.1 Machine Learning

O Aprendizado de Máquina é um dos domínios da Inteligência Artificial que promove a evolução dos algoritmos computacionais projetados para simular habilidades humanas e aprende com o ambiente ao redor, melhorando gradualmente sua precisão.

Técnicas baseadas em Aprendizado de Máquina como planejar, aprender, ouvir e falar utilizam uma ampla estrutura de algoritmos (redes neurais), ferramentas de ciência de dados para a vasta gama de tarefas de processamento e uma determinada fonte de parâmetros para analisar dados e fornecer resultados como classificações, previsões e insights com base nas informações providas pelos usuários, dessa maneira, impulsionando a tomada de decisão para determinado problema. Existem incontáveis algoritmos de Machine Learning, mas cada um é focado para solução de um problema específico.

Os algoritmos de aprendizado de máquina testam várias hipóteses imparciais com base nos recursos que educamos, dessa forma podendo realizar análises e processamentos de grande volume de dados de maneira extremamente precisa, algo intangível para o cérebro humano.

#### 2.2.4.2 Rede NeuraL Artificial

As Redes Neurais Artificiais (RNAs), também identificadas como nós interrelacionados de inteligência artificial, é um modelo computacional composto por neurônios artificiais, relativamente baseado na rede neural biológica do cérebro humano e de animais. As RNAs, buscam processar informações da mesma maneira que o cérebro processa, identificando relações e padrões entre informações e dados integrais, associando-os e classificando-os. A rede neural é o principal componente do Machine Learning, aprendizado de máquina, pois é a estrutura de algoritmos utilizada para processamento de entrada de dados, aprendendo a realizar tarefas com base em exemplos do ambiente, sem a necessidade de programação para uma tarefa em específico. Por exemplo, no reconhecimento de imagens, as redes neurais artificiais podem aprender identificar um objeto 'X' apenas analisando parâmetros de dados denominados como 'X' e 'Y', em outras palavras analisando dados divididos por classe de reconhecimento. As redes neurais aprendem de maneira autônoma, sem ao menos ter o conhecimento de alguma característica do objeto, pois isso é realizado de maneira automática conforme vão identificando novos objetos ou novas características de objetos já identificados, de maneira similar ao processo de aprendizado por experiência do cérebro biológico humano e animal.

##### 2.2.4.2.1 Neurônio Artificial

Pela razão das estruturas de RNAs terem sido desenvolvidas fundamentadas no modelo neurológio humano, os elementos computacionais ou unidades processadoras denominados como “neurônio artificial”, são modelos sintetizados dos neurônios biológicos. Os neurônios artificiais usados nas estruturas de redes neurais artificial são executados de maneira não linear, proveem saídas contínuas e realizam execuções de processamentos simples.

A abstração artificial de um neurônio biológico para representar o mundo real foi proposta por McCulloch e Pitts (1943). Este modelo de neurônio tenta simular as realidades biológicas que ocorrem

dentro de uma célula do sistema nervoso, é o mais simples e envolve de forma descomplicada os componentes e o funcionamento de um neurônio biológico, em outras palavras, de forma paralela e conectiva. Em termos práticos de execução, neste modelo, o neurônio receptor é estimulado de maneira variável pelos sinais de entrada (𝑋𝑚), que podem ser derivados de outros neurônios, onde dentre todos as incitações, algumas podem estimular mais e outras menos. A unidade de medida do nível de estímulo do neurônio receptor é representada pelo denominado, peso sináptico (𝑊𝑘𝑚), que se refere força ou amplitude de uma conexão entre dois nós, onde quanto maior o peso, maior será o estímulo.

O modelo de neurônio artificial é representado abaixo, onde há elementos mencionados previamente.

Figura - Modelo de neurônio

Diagrama

Descrição gerada automaticamente

Fonte: Adaptada de Haykin (2008)

A função de ativação orquestra o comportamento do neurônio e no modelo acima, projetado por McCulloch e Pitts (1943), não é a única forma de gerar o valor de saída do neurônio artificial, ao todo, há 3 diferentes tipos de funções de ativação.

Figura - Funções de ativação

Gráfico, Gráfico de caixa estreita

Descrição gerada automaticamente

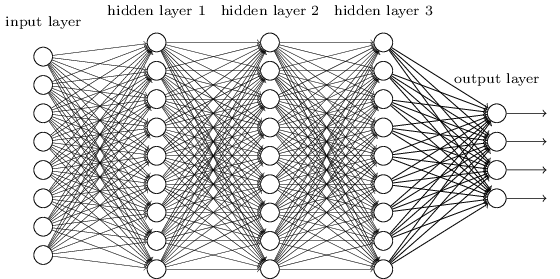
Fonte: Neural Network Toolbox (2009)

A função linear produz uma saída linear contínua, a função de escada uma saída não-linear e discreta de forma binária e a função sigmoidal produz uma saída não-linear contínua.

#### 2.2.4.3 Redes Neurais Convolucionais

Para entrarmos a fundo no tema Redes Neurais Convolucionais, é relevante que previamente tenhamos o entendimento de onde foi retirada a ideia com base biológica para a elaboração dessa arquitetura. Nos anos 60, Wiesel e Hubel realizaram uma experiência que apresentava alguns neurônios que são ativados simultaneamente quando submetidos a curvas e linhas, buscando assim conceber o reconhecimento de imagem. De uma forma simples, esse é o principal conceito de Rede Neural Convolucional: filtrar curvas, bordas e linhas e em toda camada acrescida modificar essa filtragem para uma imagem com maior complexidade.

Figura - Rede Neural Profunda



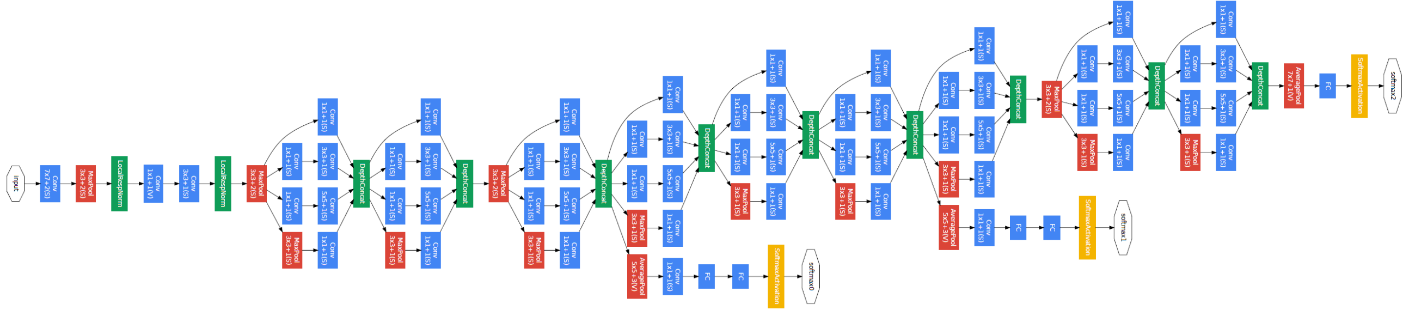
Fonte: Nielsen (2015)

Uma antiga dificuldade na categorização de imagens se trata do próprio reconhecimento delas, e as CNN (*Convolutional Neural Network*) têm um ótimo antecedente de solução dessa dificuldade. Se tem datado que a primeira rede neural convolucional que obteve êxito teve seu desenvolvimento realizado por Yann LeCun, nos anos 90, possuindo sete camadas divididas entre convolução e totalmente conectada. Posteriormente, as redes neurais convolucionais se desenvolveram, em versões mais complexas e robustas, como por exemplo, AlexNet, que detém 60 milhões de parâmetros mesmo possuindo somente 8 camadas (5 convoluções e 3 totalmente conectadas), enquanto o GoogleNet detém 22 camadas e 4 milhões de parâmetros.

##### 2.2.4.3.1 GoogleNet

Em 2015 foi proposta a GoogleNet, que serve de exemplar de uma arquitetura de rede neural convolucional profunda. O principal objetivo desta rede é utilizar um número mínimo de parâmetros sem perda de profundidade, módulo Inception proposto á isso reduz significativamente a quantidade de parâmetros de rede.

Figura - GoogLeNet



Fonte: Quoc Le & Barret Zoph (2017)

##### 2.2.4.3.2 AlexNet

AlexNet é uma arquitetura de rede neural convolucional, projetada por Alex Krizhevsky, Ilya Sutskever e Geoffrey Hinton. A Arquitetura AlexNet tem seu nome por conta de Alex Krizhevsky, um dos principais autores do artigo. AlexNet competiu e venceu o Desafio de Reconhecimento Visual em Grande Escala da ImageNet (ILSVRC) 2012 com uma taxa de erro entre os 5 primeiros de 16,4%.

Figura - Top 5 classificações ImageNet

Gráfico, Gráfico de barras

Descrição gerada automaticamente

Fonte: https://tech-pt.netlify.app/articles/pt511372/index.html

Como o modelo tinha a necessidade de treinar 60 milhões de parâmetros (um número considerável), ele é propenso a overfitting. Em geral, uma das melhores maneiras é utilizar dropout e aumento de dados que ajuda significativamente a reduzir o overfitting. Portanto, a primeira e a segunda camada completamente conectada na arquitetura usam um dropout de 0,5 para isso. Aumentar artificialmente o número de imagens por meio do aumento de dados ajuda a expandir de uma forma dinâmica o conglomerado de dados em tempo de execução, o que ajuda o modelo a generalizar melhor.

##### 2.2.4.3.3 Camadas

Pode-se desmembrar as Redes Neurais Convolucionais em dois módulos fundamentais, são eles: convolução e classificação. A fase de convulação extrai componentes de uma imagem que retratam seu conteúdo. A classificação é encarregada da função de como o próprio nome diz, classificar as informações retiradas da fase de convolução.

Uma RNC (Rede Neural Convolucional) é fragmentada em camadas, onde cada uma delas se encarrega de uma exclusiva subtarefa.

###### 2.2.4.3.3.1 Camada de Convulação

Esta camada extrai e mapeia o conteúdo de uma imagem e converte em dados. O procedimento ocorre com a aplicação de filtros que possibilitam coletar elementos de sub-blocos de uma imagem.

###### 2.2.4.3.3.2 Camada de Pooling

Esta camada obtém os blocos extraídos na camada anterior e descomplica a informação, reunindo os dados da imagem em um só valor e transferindo para a próxima camada.

##### 2.2.4.3.3.3 Camada Totalmente Conectada

Totalmente conectada é a camada onde se inicia o procedimento de classificação das informações retiradas nas demais camadas. Ela transforma o bloco em uma linha única com todos os dados fornecidos.

###### 2.2.4.3.3.4 Percepção Convolucional

Durante uma aplicação de filtros dentro da camada de convolução, os filtros permitem identificar padrões na imagem, dentre elas: texturas, bordas, formas, linhas, cantos, cores, curvas e partes de objetos contidos na imagem.

Figura - Filtros de Camadas Convolucionais

Tela de jogo de vídeo game

Descrição gerada automaticamente com confiança média

Fonte: Blog Ateliware (2021)

Uma camada de rede mais profunda permite aplicar um filtro em grau aprimorado, possibilitando em camadas mais profundas a detecção de objetos como pássaros, cães e gatos em seus filtros.

Figura - Filtros na Detecção de Objetos

Uma imagem contendo Texto

Descrição gerada automaticamente

Fonte: Blog Ateliware (2021)

##### 2.2.4.3.4 Função de Ativação

Um elemento matemático incluso na infraestrutura das redes neurais decifrando problemas. Conforme a evolução da Inteligência Artificial além das inúmeras funções de ativação, novas são criadas, baseando-se em pesquisas constantemente ampliadas.

###### 2.2.4.3.4.1 Função Linear

Esta função é constante, não tem dependência do valor de entrada (x). Esta função é a melhor estratégia para trabalhos simples, quando a interpretação é pretendida.

Sua fórmula consiste na seguinte função: F(X) = AX

###### 2.2.4.3.4.2 Função Sigmóide

Esta função é constantemente consumida, é leve e diferenciável. Quando se tem muitos neurônios, a função tem variação de 0 a 1 em um molde de S. A função leva os valores (Y) para as extremidades, o que resulta em uma qualidade pretendida ao categorizar os valores em uma classe.

Sua fórmula consiste na seguinte função: F(X) = 1 / (1 + e ^ -X)

###### 2.2.4.3.4.3 Função Tanh

Esta função é similar a sigmóide, a Tanh nada mais é que uma versão escalonada da sigmóide, contudo é simétrica relacionada à origem. É contínua e diferenciável. Sua variação é de -1 a 1.

Sua fórmula consiste na seguinte função: (X) = 2 / (1+e^(-2X)) – 1

###### 2.2.4.3.4.4 Função ReLU

Não é uma função linear, assim age como facilitadora na ação de solucionar erros e possuir muitas camadas de neurônios. Sua imprescindível vantagem comparada as demais funções é que ela não opera os neurônios simultaneamente, o que torna a rede eficiente e prática dentro da computação.

Sua fórmula consiste na seguinte função: F(X) = MAX(0, X)

### 2.2.5 OpenCV

OpenCV (Open Source Computer Vision) é uma biblioteca de programação de código aberto desenvolvida pela Intel com o objetivo de fornecer uma infraestrutura comum para aplicativos de visão computacional, acelerá-la em produtos comerciais e torná-la mais acessível. Pode ser utilizada em diversas linguagens de programação como C++, Python, Ruby e Java, são usadas para diversos tipos de análises em imagens e vídeos como tracking, reconhecimento facial, edição de fotos e vídeos, detecção e análise de textos, entre outros.

### 2.2.6 TensorFlow

Criado pela Google Brain, TensorFlow é uma biblioteca de código aberto para aprendizado de máquina que pode ser aplicado a uma ampla variedade de finalidades. É um sistema para criação e treinamento de máquina e muito usado para detectar correlações nos dados obtidos, buscando semelhança com a forma como humanos aprendem e raciocinam. Ele tem um ecossistema abrangente de ferramentas, e recursos aos desenvolvedores criar e implantar soluções com Machine Learning. Com suporte ao uso de CPUs multi-core, GPUs de uso geral e ASICs personalizados conhecidos como TPUs (Tensor Processing Units).

Os dados de entrada de uma aplicação utilizando o TensorFlow é um tensor, que se trata de um array multidimensional tipado. Um tensor é um vetor ou matriz com X dimensões que representa todos os tipos de dados. Um tensor pode ser originado a partir de uma extração de dados ou ser o resultado de uma operação.

### 2.2.7 Keras

Keras é uma API de rede neural Python fácil de usar que é executada em bibliotecas de aprendizado profundo como o TensorFlow. Ele foi projetado para permitir a prototipagem rápida de modelos de aprendizado profundo, com foco na facilidade de uso, modularidade e extensibilidade. Ele suporta Redes Neurais Convolucionais (CNNs) para visão computacional e Redes Neurais Recorrentes (NNs) para processamento sequencial de dados e qualquer combinação desses dois modelos. Como o desenvolvimento da biblioteca Keras é apoiado por várias empresas importantes no ecossistema de deep learning, Keras permite que você execute o mesmo código em uma CPU ou GPU.

## 2.3 Materiais e Métodos

Este capítulo tem por objetivo descrever os materiais e métodos utilizados para desenvolvimento, testes e implementação do projeto, além de requisitos funcionais e não-funcionais.

### 2.3.1 Ferramentas Utilizadas

Em seguida, vemos a descrição e função das ferramentas que foram utilizadas no desenvolvimento do projeto

#### 2.3.1.1 Arduino IDE

O Arduino IDE se trata de um ambiente de desenvolvimento integrado no qual possuímos todas as ferramentas necessárias para programar placas Arduino, para escrever seu código de forma satisfatória, rápida e eficiente. No IDE, temos acesso as funções de destaque de sintaxe, correção de erros, adição de bibliotecas, monitoramento serial e envio de código para a placa, para que o programa possa rodar.

#### 2.3.1.2 Microsoft Visual Studio

O Microsoft Visual Studio é uma IDE, Integrated Development Environment ou Ambiente de Desenvolvimento Integrado, desenvolvido pela Microsoft e é usado para diversos tipos de desenvolvimento de software, como programas de computados, sites, aplicações web e aplicações móveis em diversas linguagens de programação. O Microsoft Visual Studio contém recursos como ferramentas para auxílio e compiladores para facilitar o processo de desenvolvimento de software.

Neste projeto, o Microsoft Visual Studio foi utilizado para o desenvolvimento de uma Web API, Application Programming Interface, em .Net Core 3.1, framework oriundo da linguagem de programação C#, também conhecida como CSharp.

#### 2.3.1.3 .Net core 3.1

O .Net Core 3.1 é uma tecnologia ASP.NET Core que utiliza WebAssembly para ser usado como destino de compilação do navegador para linguagens de programação de ordem superior, para que possam ser usadas no desenvolvimento Web ao invés de JavaScript. Dessa maneira, abrindo o desenvolvimento Web para o público de C#.

#### 2.3.1.4 C# (CSharp)

C#, pronunciado como CSharp, é uma linguagem de programação orientada a objetos e fortemente tipada lançada pela Microsoft em 2001. É uma linguagem de programação OpenSource (código aberto), simples, moderna, versátil, flexível, segura, eficiente e está em constante evolução.

#### 2.3.1.5 Docker

O Docker, uma ferramenta OpenSource (código aberto) lançada em 2013 é uma plataforma de software para criar aplicativos baseados em contêineres, ambientes de execução pequenos e leves que fazem uso compartilhado do kernel do sistema operacional, porém, executados isoladamente uns dos outros e neste projeto, foi utilizado para conteinerização da Web API em .Net Core 3.1 e implementação no Amazon ECR. Em nosso projeto, a ferramenta é utilizada para geração de imagens no MySql

#### 2.3.1.6 Google Colab

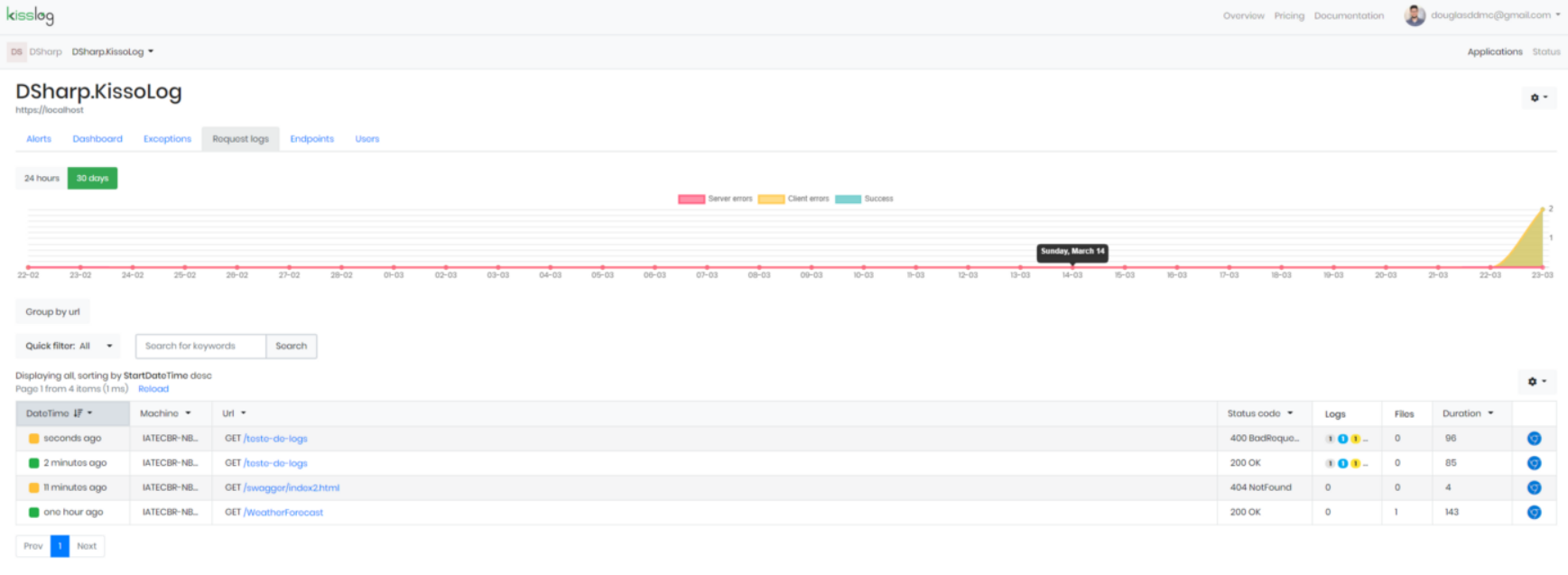
O Google Colaboratory, mais conhecido como Google Colab é um serviço que fornece acesso a máquinas com um grande poder de processamento para a execução de Jupyter notebooks com uma arquitetura totalmente em nuvem possuindo versão gratuita que já permite acesso a ótimas máquinas, com GPUs, discos e memória RAM inferiores aos planos pagos, mas já fornecendo uma ótima base para começar ou executar algoritmos simples.

#### 2.3.1.7 Jupyter notebooks

Jupyter notebooks é uma tecnologia que utiliza a linguagem de programação Python, que tem sido amplamente usado devido a um de seus aspectos chave, e execução por blocos, dando a possibilidade de dividir o código a ser executado em diferentes blocos que podem serem executados de forma independente, ou seja, caso venha a ocorrer algum problema durante a execução de um bloco não é necessário executar todo o código novamente, quebrando um pouco a linearidade do algoritmo.

#### 2.3.1.8 Kisslog

Kisslog, trata-se de uma extensão presente no VS.Code que a partir de uma aplicação criada nos permite o monitoramento dos Logs, gerando um Dashboard com a exibição de todos eles, facilitando a leitura e entendimento da API e seu funcionamento, como erros e sucessos.



#### 2.3.1.9 Heroku

Heroku é uma PaaS (*Platform as Service*), ou seja, uma plataforma na nuvem sendo um serviço que suporta inúmeras linguagens de programação, como: Java, PHP, Ruby on Rails, Node, Scala, Go, Clojure e Phyton. Como o Heroku não é suportado pela linguagem de programação C#, o *deploy* das aplicações ASP.NET pode ser feito via contêiner Docker ou utilizando *BuildPacks.* A plataforma nos permite a hospedagem de códigos e nos ajuda a não se preocupar muito com a disponibilidade, escala e infraestrutura da aplicação, comummente é mais utilizada para aplicações back-end.

#### 2.3.1.10 AWS RDS

O *Amazon Relational Database Service* ou Amazon RDS é um serviço de banco de dados relacional distribuído fornecido pela Amazon Web Services. É um serviço web com hospedagem "na nuvem" projetado para simplificar a configuração, operação e dimensionamento de bancos de dados relacionais para aplicativos. O Serviço é utilizado para gerenciar backups, patches de software, detecção e recuperação automática de falhas, pode ser habilitado backups automatizados ou pode realizar a criação manual de seus próprios snapshots de backup, também pode utilizar os backups para restauração do banco de dados.

#### 2.3.1.8 MySql

O MySQL é uma das principais distribuições de SGBD (Sistema Gerenciador de Bancos de Dados) de padrão SQL disponíveis para uso comercial. É executável em vários Sistemas Operacionais, estável e foi utilizado para persistência de dados deste projeto, implementado no serviço AWS RDS (*Amazon Relational Database Service*).

### 2.3.2 Requisitos de Software

Segundo Pressman (2011, p.126), os requisitos de um sistema estão entre as tarefas mais difíceis que um engenheiro de software pode enfrentar, afinal, os requisitos de um software dependem do tipo do sistema que está sendo desenvolvido, dos usuários que o software será direcionado, entre muitos outros aspectos e cada uma dessas características dever ser definidas como funcionalidades diretas e não diretas do sistema, em outras palavras, definidas como requisitos funcionais e não-funcionais.

#### 2.3.2.1 Requisitos Funcionais

Os requisitos funcionais são referentes ao que o sistema não pode deixar de ter, ou seja, suas principais funcionalidades. A tabela abaixo especifica os requisitos funcionais do projeto em questão.

Tabela - Requisitos funcionais

|  |  |  |
| --- | --- | --- |
| **Requisito** | **Prioridade** | **Descrição** |
| Comunicação entre Arduíno e API | Essencial | Desenvolvimento de lógica para comunicação entre hardware e software em ambas as partes. |
| Integração entre API e Inteligência Artificial | Essencial | Desenvolvimento de lógica para integração de API com a Inteligência Artificial para processamento de reconhecimento de emoções. |
| Persistência de dados pós processamento | Essencial | Persistir dados gerados pela Inteligência Artificial em banco de dados MySql |
| Disponibilização de dados por API | Essencial | Desenvolvimento de lógica de acesso à dados na API para consumo e disponibilização de dados dia rotas baseadas em arquitetura REST. |
| Desenvolvimento de Front-End para consumo de dados persistidos | Essencial | Desenvolvimento de página Web para consumo de dados e exibição. |
| Desenvolvimento de ferramentas para análise de dados | Essencial | Desenvolvimento de ferramentas como gráficos e relatórios para fácil visualização e análise de dados. |
| Implantação de API em ambiente AWS | Essencial | Implantação e integração de API .Net Core 3.1, MySql e Inteligência Artificial em ambiente AWS para disponibilização ao usuário final. |

Fonte: O Autor

#### 2.3.2.2 Requisitos Não-Funcionais

Conhecidos como requisitos de qualidade, os requisitos não-funcionais, são os comportamentos e estrutura do sistema como sua usabilidade, segurança, eficiência, manutenção, arquitetura, entre outras características que não são representados através de funções.

A tabela abaixo especifica e descreve os requisitos de qualidade do projeto em questão.

Tabela - Requisitos não-funcionais

|  |  |  |
| --- | --- | --- |
| **Requisito** | **Prioridade** | **Descrição** |
| Logs de informações da API | Essencial | Desenvolvimento e configuração de logs de informações da API |
| Logs de avisos da API | Essencial | Desenvolvimento e configuração de logs de avisos da API |
| Logs de erros da API | Essencial | Desenvolvimento e configuração de logs de erros da API |
| Usabilidade | Essencial | A interface final para interação do usuário com relatórios e análise de dados deve ser clara, simples e intuitiva. |
| Segurança de Informações | Essencial | As informações de clientes, como as imagens capturadas pelo hardware, devem ser criptografadas pré-processamento e excluídas pós-processamento. |
| Sistema de Gerenciamento de Banco de dados | Essencial | O bando de dados a ser utilizado deve ser o MySql. |
| Linguagem de programação C# e uso de .Net Core 3.1 para desenvolvimento de API | Essencial | A API deve ser desenvolvida em linguagem de programação CSharp incluindo a utilização do Framework .Net Core 3.1. |
| Linguagem de programação Python para desenvolvimento e treino da Inteligência Artificial. | Essencial | A Inteligência Artificial deve ser desenvolvida e testada a partir da linguagem de programação Python. |
| Consistência de dados | Essencial | Os dados não podem ser corrompidos ou excluídos por falha do sistema |
| Observabilidade | Essencial | Implementação de ferramenta que garanta observabilidade do sistema. |

Fonte: O Autor

### 2.3.3 Desenvolvimento

#### 2.3.3.1 Dataset

O banco de imagens utilizado foi o FER2013, esta base originou-se de um desafio de reconhecimento facial disponibilizando 28709 imagens para treinamento, 3589 para teste e a mesma quantidade para validação, são imagens de expressões faciais de 48 por 48 pixels em tons de cinza classificadas em 7 categorias, neutro, felicidade, tristeza, raiva, medo, aversão e surpresa. Classificadas com base nos primeiros estudos de Ekman, por isso, a falta da categoria desprezo. (CARRIER; COURVILLE, 2013).

#### 2.3.3.2 Hardware

A Seguir, uma descrição a respeito da ferramenta hardware utilizada, tendo por objetivo descrever de forma aprofundada e detalhada suas características e detalhes,

##### 2.3.3.2.1 ESP32 Cam

O Módulo ESP32-CAM é um microcontrolador fabricado pela empresa Espressif Systems que possui alta performance, baixo consumo de energia (tensão de 3.3V), capacidade de fornecer comunicação sem fio através de conexão Wireless Fidelity (WiFi) padrão 802.11 b/g/n e através do próprio sistema Bluetooth, 4MB de memória, CPU de dois núcleos, 448Kb de memória ROM, 520KBytres de memória RAM, 36 portas e dentre elas, 12 podem ser utilizadas como conversor analógico digital de 12 bits e Bluetooth. (SYSTEMS, 2019).

O ESP32 Cam é um pequeno módulo de câmera com chip ESP32-S e câmera OV2640 de 2 megapixels que será integrada com a placa Arduino UNO Wifi REV2. Este módulo também comporta um LED de alto luminosidade para atuar como flash e uma entrada para cartões SD. Dado seu alto poder de processamento, o módulo permitirá a captura de imagens e vídeos de clientes durante atendimentos para posteriormente, processamento e resultados dos atendimentos prestados.

Para configuração deste módulo, é necessário a instalação do complemento ESP32 no Arduino IDE. Todas suas funções são programadas utilizando a linguagem de programação C++, assim como a placa eletrônica do Arduino.

O uso do software citado é simples, primeiramente é necessário realizar a configuração do modelo da placa a ser utilizada, depois seleciona-se a porta serial virtual do dispositivo conectado ao computador e por fim transfere-se o programa para o hardware por meio do botão carregar (ARDUINO, 2015).

#### 2.3.3.3 API

A Web API concentrará a maior parte das regras funcionais, afinal, será a peça centralizadora que integrará todos os outros serviços, como o hardware, inteligência artificial, banco de dados e ferramentas front-end para análise de dados.

Baseada em arquitetura REST (Representational State Transfer), a Web API disponibiliza rotas em métodos POST e GET para transferência de dados, ao todo serão dois *endpoints* disponíveis.

1. POST: Envio de dados e imagens do hardware para a Web API com finalidade de processamento da análise das expressões do cliente durante atendimento.

Rota:https://{domínio}/analise-expressoes/classificações/

1. GET: Obter dados que foram gerados a partir do processamento de análise de emoções por parte da inteligência artificial e persistidos na base de dados, onde o parâmetro de busca é a identificação do colaborador e retornará todos os resultados de atendimento de um colaborador específico.

Rota:https://{domínio}/analise-expressoes/classificacoes/colaborador/{id}

Todas as transferências de dados serão realizadas com criptografia E2EE, de ponta-a-ponta.

A API foi desenvolvida baseada no modelo arquitetural de N camadas, também conhecida como arquitetura de distribuidor, onde n representa o número de camadas. A diferença entre a arquitetura de 3 camadas e a arquitetura de N camadas é que há mais de um servidor de aplicativos intermediando entre a camada de interface do usuário e a camada de banco de dados.

Isso é feito para distribuir as lógicas de negócios entre diferentes servidores para que todos possam atuar individualmente. Esta arquitetura é muito flexível com o aumento do número de servidores de banco de dados. O aplicativo baseado em arquitetura de N camadas é mais seguro do que as outras camadas de arquitetura.

Até mesmo fazer alterações ou modificações em qualquer camada pode ser feito sem afetar a outra camada. A qualquer momento, a lógica de negócios pode ser modificada ou o tamanho do banco de dados pode ser aumentado, sem afetar o cliente.

Assim, camadas na arquitetura de software são o número de máquinas nas quais o trabalho de um aplicativo é distribuído para aumentar a segurança, escalabilidade e produtividade do software.

#### 2.3.3.4 Diagrama de Integração

Figura - Diagrama de integração de componentes do projeto

Diagrama

Descrição gerada automaticamente

Fonte: O Autor

##### 2.3.3.4.1 Usabilidade

O Projeto consiste em uma Inteligência Artificial integrada a um hardware que possui um botão físico para iniciar e finalizar o processo de captura de imagens do consumidor. Ao acionar o botão, o responsável dará início ao processo, em que a câmera iniciará com a captura de imagens e armazenamento em memória. Quando o atendimento for finalizado no momento que o botão for novamente pressionado, as imagens serão enviadas para a API e processadas pela Inteligência Artificial, por fim, persistindo os dados do processamento da análise em um banco de dados.

Temos como objetivo principal a melhoria de atendimento para a evolução das empresas e para chegarmos neste resultado precisamos trabalhar com a UX (*User Experience*) e UI (*User Interface*), que nada mais é do que a experiência do usuário e a interface do usuário, respectivamente.

Na atividade, a UX é a relação entre usuário e produto ou serviço. Sua principal função é atender que todas as etapas de interação do usuário sejam fáceis e dinâmicas, elas devem impactar positivamente os usuários e colaborarem com os objetivos do negócio. Já a UI deve ser uma interface funcional e usual, tornando-se a intermediária visual ou não, entre o usuário e a máquina.

Baseando-se em nosso projeto, utilizamos o usuário no centro de decisões do produto, com isso traremos apenas simples ações de botões que irão capturar as expressões de forma descomplicada e automática, sem gerar frustações, isso porque normalmente as avalições de atendimento são realizadas via formulários e são cansativas e pouco simplificadas, o que pode gerar descontentamento com a interface. O nosso objetivo é uma experiência positiva do cliente para a fidelização da empresa.

Com base em expressões faciais podemos medir a qualidade de um atendimento, um dos aspectos essenciais para o sucesso de uma corporação. Propõe-se que a empresa utilize os dados fornecidos por nossa API para montar dashboards para análise interna das expressões coletadas durante os atendimentos realizados em seus estabelecimentos.

##### 2.3.3.4.2 Proteção de dados

Atualmente os debates sobre como os dados dos usuários são capturados sempre chegam a um ponto chave, autorização de uso dos dados para análise, sendo assim, o consumidor deve aceitar o uso de sua imagem antes do atendimento ser iniciado, caso o recuse, o processo não será realizado.

Considerando a Lei Geral de Proteção Dados (LGPD), uma vez que as imagens foram capturadas é de nossa responsabilidade garantir a segurança e integridade, em razão disso a criptografia será utilizada durante todo o processo de transporte e armazenamento temporário das imagens, as imagens capturadas serão armazenadas somente até o momento da análise, após serem tratadas e analisadas pela API serão excluídas para garantir que não sejam utilizadas por terceiros.

## 2.4 Resultados

Após a fase de aprendizado de máquina, a rede neural encontra-se apta para realizar suas classificações e predições, com isso construímos uma matriz de confusão para identificarmos a proporção de acertos da IA após sua análise. Na tabela, as microexpressões da esquerda são as emoções expressas nas imagens utilizadas para a identificação, e as microexpressões na parte inferior são as emoções detectadas através do processamento das RNAs.

Figura - Desempenho e Classificação do Aprendizado de Máquina

Calendário

Descrição gerada automaticamente

Fonte: O Autor

Durante a fase de treinamento, o processo utiliza as imagens separadamente para testar cada passo dado, e através deste teste efetuamos o cálculo para acurácia do modelo, que é a proporção total de acertos e obteve-se o valor de 65% de acurácia.

# 3 CONSIDERAÇÕES FINAIS

Dado o exposto, é possível concluir que a Inteligência Artificial traz inúmeras possibilidades de soluções para problemas que anteriormente não eram possíveis, e uma dessas é a capacidade de análise de expressões baseadas no reconhecimento facial, abrindo um leque de oportunidades e melhorias para processos empresariais, como o atendimento, um dos pilares mais importantes para a fidelização de clientes.

Tendo em vista essa oportunidade, juntamente à finalidade de analisar os procedimentos e eficácia dos serviços prestados nas empresas, o trabalho apresenta a proposta de desenvolvimento de um protótipo que analisa a comunicação não-verbal transmitida pela face de consumidores durante determinado atendimento, para fins de avaliações e feedbacks de seus colaboradores. Dessa forma, retornando o percentual de presença de cada uma das sete emoções básicas, sendo elas 'Raiva', 'Nojo', 'Medo', 'Felicidade', 'Tristeza', 'Surpresa', 'Neutro'.

# REFERÊNCIAS BIBLIOGRáFICAS

Ekman, Paul. **A linguagem das emoções**: revolucionando sua comunicação e seus relacionamentos reconhecendo todas as expressões das pessoas ao redor. São Paulo: Lua de Papel, 2003.

Ferreira, Caio Henrique. **Base de dados CICEM: Microexpressões em Faces Humanas**. São Paulo: CICEM, 2017.

Hatfield, Elaine; Cacioppo, John T; Rapson, Richard L. **Emotional** **Contagion.** United Kingdom: Cambridge University Press, 1993.

Ekman, Paul; Rosenberg, E. L. (Eds.). **What the face reveals**: basic and applied studies of spontaneous expression using the facial action coding system (FACS). New York: Oxford University Press, 2005.

Ferreira, Caio Henrique. **As Faces das Emoções Básicas**. São Paulo: CICEM, 2017-2018.

Ekman, Paul; Friesen, W. V. Hager, J. C. **The Facial Action Coding System**. Salt Lake City, UT: research Nexus ebook, 2002.

Boulogne, Duchenne de. **The mechanism of human facial expression**. Nova York/EUA: Universidade de Cambridge, 1990.

Ferreira, Caio Henrique. **Estudos sobre a mensuração científica da face humana**: vol. 1 – o guia do emocionauta. São Paulo: CICEM, 2018.

Darwin, Charles. **A expressão das emoções no homem e nos animais**. São Paulo: Companhia de Bolso, 2009.

Ferreira, Caio Henrique. **Base de Dados CICEM: Microexpressões em Faces Humanas 2.0**. São Paulo: CICEM, 2020.

Ekman, Paul. **A linguagem das emoções**. São Paulo/SP: Leya, 2011.

Ferreira, Caio Henrique; Magalhães, A. Freitas. **Facial Action Coding System 2.0:** Guia F-M FACS 2.0. São Paulo: F-MGB Lab, 2017.

Grazer, B. Lie to me, **Television Series**. California: 20th Century Fox Television, 2009-2011.

Magalhães, A. Freitas. **O código de Ekman: o cérebro, a face e a emoção**. Porto/PT: Escrytos, 2011.

Bom atendimento é mais importante para consumidor do que qualidade, revela pesquisa. **Revista Pequenas Empresas & Grandes Negócios**, 2009.

Gomes, Fernando; JOHN, Valquiria Michela. Everybody Lies: **O estudo das emoções na série Lie To Me**. Itajaí/SC: Revista estudos em comunicação, v. 18, 2015.

VIANA, Isabel. **Comunicação não verbal e expressão facial das emoções básicas**. Quinta de Prados/PT: Revista de Letras, n°13, 2014.

Ekman, Paul. **Would you lie to me?** EUA: Artigo publicado no jornal The Guardian, 2003.

Novaes, Felipe C. **O legado de Darwin para a compreensão das emoções**. **Instituto Brasileiro de Linguagem Emocional, 2022.**

Pain, Giorglli Soares. **Análise das microexpressões faciais como prova inominada.** Londrina/PR: Instituto de Pós-Graduação – IPOG, 2019.

COVIELLO, L. SOHN; Y. KRAMER; A. MARLOW; C. FRANCESCHETTI, M. **[Detecting Emotional Contagion in Massive Social Networks](http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0090315)**. Jornals Plos One, 2014.

**Jabr, Ferris. The evolution of emotion: Charles Darwin's little-known psychology experiment.** United States: Scientific American, 2010.  
  
CHOI, A. A.; MELO, B. C.; M. KHOOSHABEH, C. P.; WOOD, W; GRATCH, C. J. **Physiological evidence for a dual process model of the social effects of emotion in computers. International Journal of Human-Computer Studies**, 2015.  
  
Magalhães, A. Freitas. **Encyclopedia of Human Behavior**. United States: Second Edition, 2012.

Tran, V. [**The influence of emotions on decision-making processes in management teams**](https://archive-ouverte.unige.ch/unige:236). Geneva, Switzerland: Université de Genève, 2004.

SILVA, Ivan Nunes da; SPATTI, Danilo Hernane; FLAUZINO, Rogério Andrade. **Redes Neurais Artificiais: para engenharia e ciências aplicadas**. 2ª edição. ed. rev. e aum. São Paulo: Artliber Editora Ltda., 2016.

HAYKIN, S. S. **Redes neurais artificiais: princípio e prática**. 2a Edição, Bookman, São Paulo, Brasil, 2000.

FURTADO, Maria Inês Vasconcellos. **Redes Neurais Artificiais: Uma Abordagem Para Sala de Aula**. Ponta Grossa (PR): Atena Editora, 2019. Disponível em: https://www.atenaeditora.com.br/wp-content/uploads/2019/05/e-book-RedesNeurais-Artificiais-uma-Abordagem-para-Sala-de-Aula.pdf.

Roger, PRESSMAN. **Engenharia de Software: uma abordagem profissional.** AMGH, 2011. Disponível em: https://books.google.com.br/books?id=eRIOuQAACAAJ.

COURVILLE, Aaron. **Challenges in Representation Learning: Facial Expression Recognition Challenge**. 2013. Disponível em: <https://www.kaggle.com/c/challenges-in-representation-learning-facialexpression-recognition-challenge>. Acesso em: 25/04/2022

EKMAN, P**. Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life**. Henry Holt and Company, 2003.

ABADI, Martín, **TensorFlow: A System for Large-Scale Machine Learning.** 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), Savannah, GA, 2016. Disponível em: https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi.

OPENCV TEAM. **OpenCV: about**. About. 2020. Disponível em: <https://opencv.org/>. Acesso em: 27/04/2022

BEZ, M. **Reconhecimento Facial e Micro Expressões**. 3. ed. São Paulo, Murof, 2017.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. Deep Learning. MIT Press: 2016.

**KERAS. How convolutional neural networks see the world**, 2016. Disponível em <https://blog.keras.io/how-convolutional-neural-networks-see-the-world.html>. Acesso em 28/04/2022

**DATARISK. O que é Matriz de Confusão**. Disponível em:https://ajuda.datarisk.io/knowledge/o-que-%C3%A9-matriz-de-confus%C3%A3o. Acesso em 27/04/2022

Passos, Bianka. **O mundo do ponto de vista das Redes Neurais Convolucionais**. Blog Ateliware. Disponível em: <https://ateliware.com/blog/redes-neurais-convolucionais>. Acesso em: 24/04/2022

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, **A. Tendências da Inteligência Artificial na educação: uma visão narrativa**. Procedia Computer Science, 2018

**Deep Learning Book**. Capítulo 8 - Função de Ativação. Disponível em: <https://bityli.com/CETQsx>

HOCKENBURY, Don; HOCKENBURY, Sandra E. **Discovering Psychology**. 1990

DARWIN, Charles, **The Expression of the Emotions in Man and Animals**. 1872

AELA, Editorial. **Qual a Diferença Entre UX e UI Design? – Entenda De Uma Vez.** Disponível em: https://aelaschool.com/experienciadousuario/qual-a-diferenca-entre-ux-e-ui-design-entenda-de-uma-vez/. Acesso em: 17 junho 2022.

BASSO, Bianca. **Inteligência Artificial Para Reconhecimento De Emoções Em Um Ambiente Empresarial**. 17th International Conference on Information Systems and Technology Management. Disponível em: http://contecsi.submissao.com.br/arquivos/6526.pdf. Acesso em: 19 jun. 2022.

SYSTEMS, E. ESP32 Series datasheet. 2019.

ARDUINO. 2005, 2015, 2020. Disponível em: http://arduino.cc/