

**UNIVERSIDADE DO OESTE DE SANTA CATARINA**

**ANGELINE ZANELA**

**RECONHECIMENTO DE EMOÇÕES COM REDES NEURAIS CONVOLUCIONAIS  
PARA O ENSINO DE LIBRAS**

**CHAPECÓ, SC  
2024**

ANGELINE ZANELA

**RECONHECIMENTO DE EMOÇÕES COM REDES NEURAIS CONVOLUCIONAIS  
PARA O ENSINO DE LIBRAS**

Projeto de Trabalho de Conclusão de Curso apresentado ao Curso de Engenharia da Computação, Área das Ciências Exatas e Tecnológicas, da Universidade do Oeste de Santa Catarina – UNOESC Campus de Chapecó como requisito parcial à obtenção do grau de bacharel em Engenharia da Computação.

Orientador: Prof. Jacson Luiz Matte

Chapecó, SC  
2024

## LISTA DE ILUSTRAÇÕES

Ilustração 1 - Expressões Faciais Básicas .....	10
Ilustração 2 - Principais ramos de estudo da IA .....	16
Ilustração 3 - Classificação de imagens .....	18
Ilustração 4 - Aplicando Métodos de Suavização .....	22
Ilustração 5 - Aplicando alguns Métodos de Limiarização .....	23
Ilustração 6 - Aplicação do Método de Erosão e Dilatação.....	25
Ilustração 7 - Aplicação do Método de Detecção de Bordas de Canny .....	25
Ilustração 8 - Fluxograma do Projeto na Prática .....	28
Ilustração 9 - Etapas do Modelo em Cascata .....	31
Ilustração 10 - Diagrama de atividades .....	35
Ilustração 11 - Tela inicial de captura em tempo real .....	37

## **LISTA DE QUADROS**

Quadro 1 - Descrição das Expressões Faciais.....	10
Quadro 2 - Descrição das Expressões Gestuais .....	11
Quadro 3 - Descrição dos trabalhos .....	15
Quadro 4 - Requisitos Funcionais .....	34
Quadro 5 - Requisitos Não Funcionais .....	34
Quadro 6 - Cronograma do Trabalho de Conclusão de Curso (Quinzenal) .....	38

## **LISTA DE TABELAS**

Tabela 1 - Quantitativo de obras usadas por assunto .....	30
Tabela 2 - Quantitativo Total de obras usadas .....	30
Tabela 3 - Orçamento do projeto.....	39

## **LISTA DE ABREVIACÕES E SIGLAS**

- |            |   |
|------------|---|
| 1. API     | - Application Programming Interface               |
| 2. CNN     | - Convolutional Neural Network                    |
| 3. CNN-2D  | - Convolutional Neural Network - 2 Dimensional    |
| 4. CNNs    | - Convolutional Neural Networks                   |
| 5. FF      | - Feedforward Neural Network                      |
| 6. GPU     | - Graphics Processing Unit                        |
| 7. IA      | - Inteligência Artificial                         |
| 8. IBGE    | - Instituto Brasileiro de Geografia e Estatística |
| 9. LIBRAS  | - Língua Brasileira de Sinais                     |
| 10. LSTM   | - Long Short-Term Memory                          |
| 11. PDI    | - Processamento Digital de Imagens                |
| 12. RNN    | - Recurrent Neural Network                        |
| 13. SVM    | - Support Vector Machine                          |
| 14. TCC    | - Trabalho de Conclusão de Curso                  |
| 15. UNOESC | - Universidade do Oeste de Santa Catarina         |

## SUMÁRIO

<b>LISTA DE ILUSTRAÇÕES.....</b>	<b>3</b>
<b>LISTA DE QUADROS.....</b>	<b>3</b>
<b>LISTA DE TABELAS.....</b>	<b>4</b>
<b>LISTA DE ABREVIACÕES E SIGLAS.....</b>	<b>5</b>
<b>SUMÁRIO.....</b>	<b>6</b>
<b>1 INTRODUÇÃO.....</b>	<b>6</b>
1.1 DELIMITAÇÃO DO TEMA DE PESQUISA.....	6
1.2 PROBLEMA DE PESQUISA E JUSTIFICATIVA.....	7
1.3 OBJETIVOS.....	8
<b>2 REVISÃO DA LITERATURA.....</b>	<b>9</b>
2.1 LIBRAS.....	9
2.1.1 Expressões faciais e gestuais na LIBRAS.....	9
2.2 TRABALHOS RELACIONADOS.....	12
2.3 INTELIGÊNCIA ARTIFICIAL.....	16
2.3.1 Visão computacional e Aprendizado de Máquina.....	17
2.3.2 Redes Neurais Convolucionais.....	18
2.4 TECNOLOGIAS.....	20
2.4.1 TensorFlow.....	20
2.4.2 OpenCV.....	21
2.5 PROCESSAMENTO DE IMAGENS DIGITAIS.....	21
2.5.1 Suavização.....	22
2.5.2 Limiarização.....	23
2.5.3 Erosão e Dilatação.....	24
2.5.4 Detecção de Borda.....	25
<b>3 PROCEDIMENTOS METODOLÓGICOS E TÉCNICOS.....</b>	<b>26</b>
3.1 CARACTERIZAÇÃO DA METODOLOGIA DE PESQUISA.....	26
3.2 DELIMITAÇÃO DO ESTUDO.....	26
3.2.1 População e Amostra.....	27
3.3 QUESTÕES DE PESQUISA.....	27
3.4 ESBOÇO DO PROJETO NA PRÁTICA.....	27
3.5 APLICAÇÃO DA METODOLOGIA.....	29
3.5.1 Construção do Referencial Teórico.....	29
3.5.2 Desenvolvimento do Sistema Computacional.....	30
<b>4 APRESENTAÇÃO DA SOLUÇÃO.....</b>	<b>33</b>
4.1 REQUERIMENTOS.....	34
4.2 PROJETO.....	34
4.2.1 Diagrama de Atividades.....	35
4.2.2 Wireframe.....	36
4.3 VALIDAÇÃO SISTEMA.....	37

<b>5 CRONOGRAMA.....</b>	<b>39</b>
<b>6 ORÇAMENTO.....</b>	<b>39</b>
<b>REFERÊNCIAS.....</b>	<b>41</b>



## 1 INTRODUÇÃO

A Língua Brasileira de Sinais (LIBRAS) é o idioma principal da comunidade surda brasileira, utilizando gestos, posturas corporais e expressões faciais para comunicação. Assim como em qualquer língua, é capaz de expressar conceitos, ideias, sentimentos e opiniões, sendo independente da língua portuguesa.

Segundo o Instituto Brasileiro de Geografia e Estatística (IBGE), na Pesquisa Nacional de Saúde de 2019, apenas cerca de 22,4% das pessoas com algum grau de surdez sabem utilizar a LIBRAS. Muitos deficientes auditivos têm baixa escolaridade e dificuldades com a língua escrita, o que os exclui socialmente, apesar das leis de igualdade. A difusão da LIBRAS é limitada, o que contribui para essa exclusão e reforça a necessidade de métodos de ensino inovadores e acessíveis.

Dessa forma, a LIBRAS desempenha um papel vital, no entanto, a falta de recursos e métodos de ensino inovadores podem dificultar o aprendizado eficaz especialmente para crianças. Nesse contexto, o avanço das tecnologias digitais têm modificado significativamente a forma como nos comunicamos e compartilhamos informações, tanto dentro quanto fora das escolas. Diante dessas transformações, é essencial considerar como podemos aproveitar essas novas tecnologias para melhorar a educação, tornando-a mais acessível e eficaz (OLIVEIRA; MOURA, 2015, p. 77).

O presente trabalho propõe o desenvolvimento de um modelo com visão computacional que ajude na sincronização das expressões faciais e gestuais, para auxiliar no ensino de LIBRAS. O foco seria facilitar o aprendizado da LIBRAS, e também promover a compreensão e o respeito pela diversidade linguística e cultural.

### 1.1 DELIMITAÇÃO DO TEMA DE PESQUISA

O propósito desta pesquisa é explorar o potencial da visão computacional como uma ferramenta pedagógica para o ensino de expressões faciais e gestuais, com foco especial na Língua Brasileira de Sinais (LIBRAS), a fim de facilitar a comunicação e o aprendizado de crianças em idade escolar.

O principal objetivo é desenvolver um modelo utilizando Redes Neurais Convolucionais (CNNs) para detectar e interpretar expressões estáticas, abrangendo tanto o reconhecimento facial quanto o gestual, com ênfase em quatro emoções primárias: felicidade, tristeza, raiva e surpresa. Para alcançar este objetivo, essas emoções foram selecionadas, pois

são consideradas centrais no espectro emocional humano e desempenham um papel fundamental na comunicação não verbal.

Com este modelo para ensino em sala de aula, os educadores poderiam criar atividades interativas e dinâmicas, nas quais os alunos seriam incentivados a expressar suas emoções por meio de sinais, promovendo não apenas o aprendizado, mas também o desenvolvimento da inteligência emocional e da empatia entre os estudantes.

## 1.2 PROBLEMA DE PESQUISA E JUSTIFICATIVA

Segundo TEIXEIRA (2022), a inclusão de pessoas com deficiência auditiva no cenário educacional é um desafio persistente. A pesquisa destaca a importância da tecnologia no ensino da LIBRAS para eliminar barreiras comunicativas. Desse modo, promover a compreensão e o respeito pela diversidade linguística e cultural enriquece a experiência educacional e promove a inclusão e os direitos linguísticos (PEDROSO; PAIVA, 2018).

O ensino das emoções em LIBRAS para crianças com deficiência auditiva apresenta desafios específicos, como a falta de sincronização entre gestos e expressões faciais. Conforme a professora Marineiva, especialista em pedagogia, a dificuldade dos professores em ensinar esses elementos simultaneamente pode resultar em uma compreensão incompleta das emoções pelas crianças. A integração harmoniosa de gestos e expressões faciais é essencial para a comunicação eficaz.

Alves e Frassetto (2015) ressaltam a importância de introduzir LIBRAS desde a infância, pois isso desempenha um papel crucial no desenvolvimento da interação social e da expressão motora das crianças. Lacerda e Turetta (2018) acrescentam que o avanço linguístico para crianças com deficiência auditiva ocorre de maneira distinta do que para crianças ouvintes, destacando a necessidade de criar novos métodos de ensino adaptados a essas diferenças.

Desse modo, o presente trabalho busca apresentar uma solução através do desenvolvimento de um modelo de visão computacional utilizando Redes Neurais Convolucionais (CNNs) para identificar e interpretar expressões faciais e gestuais em LIBRAS. A visão computacional é uma ferramenta poderosa que pode analisar e sincronizar gestos e expressões faciais em tempo real. E as CNNs são altamente eficazes na extração de características complexas e na classificação de padrões visuais em imagens. A proposta é criar um sistema que integre essas tecnologias para proporcionar um ensino mais eficaz e inclusivo

para crianças surdas, abordando os desafios identificados e promovendo um ambiente de aprendizagem mais integrado.

### 1.3 OBJETIVOS

Neste módulo serão apresentados os objetivos que norteiam o desenvolvimento deste projeto. Os objetivos apresentados servem como base para uma compreensão clara do que se pretende alcançar.

#### 1.3.1 Objetivo Geral

Desenvolver um modelo utilizando Redes Neurais Convolucionais (CNNs) para identificar emoções em LIBRAS, integrando reconhecimento facial e gestual para auxiliar no ensino infantil em sala de aula. O sistema proposto é um estudo de caso, que busca facilitar a execução com sincronia das expressões pelas crianças, capacitando-as a se expressarem corretamente e a compreenderem as expressões dos colegas de forma mais eficaz.

#### 1.3.2 Objetivos Específicos

Para alcançar o objetivo geral proposto, é essencial atender aos seguintes objetivos específicos:

- Elaborar um conjunto de dados contendo as quatro emoções selecionadas para o estudo de caso, sendo elas: felicidade, tristeza, raiva e surpresa;
- Realizar o treinamento do modelo, com as tecnologias escolhidas, utilizando o conjunto de dados criado;
- Implementar o modelo treinado, permitindo a identificação das emoções em LIBRAS e interação em tempo real com a câmera;
- Avaliar as métricas do modelo desenvolvido na identificação das emoções.

## **2 REVISÃO DA LITERATURA**

### **2.1 LIBRAS**

A surdez consiste na perda maior ou menor da percepção dos sons. Ela é ocasionada por um déficit auditivo que prejudica a aquisição da linguagem e da fala do indivíduo. Apesar de os idiomas serem predominantemente orais e auditivos, as pessoas com deficiência auditiva se comunicam por meio de gestos e expressões faciais, tendo a LIBRAS como língua natural e oficial (BRASIL, MEC, 2005, p. 19) Com isso, uma dificuldade enfrentada pelas crianças é a escassez de exposição à língua de sinais em suas interações sociais, uma vez que a sociedade prioriza o uso exclusivo do português oral (GARRUTTI; MOREIRA, 2022)

Para entender a Língua Brasileira de Sinais (LIBRAS), é importante reconhecer que ela possui uma estrutura gramatical própria, organizada a partir de parâmetros específicos que formam os níveis linguísticos. Assim como nas línguas orais, onde se pode criar uma infinidade de mensagens combinando um número limitado de fonemas, na LIBRAS é possível gerar inúmeras mensagens através da combinação de elementos visuais limitados (FERRAZ, 2019, p. 15)

Tais elementos visuais são chamados de parâmetros e, de acordo com Ferraz (2019), incluem: a configuração de mãos, que é a forma que a mão assume ao fazer o sinal. O ponto de articulação, que é o local onde o sinal é feito, pode ser no corpo ou no espaço. O movimento da mão durante o sinal. A orientação da palma da mão, que é a direção em que a palma está voltada. E a expressão facial e corporal, ajudando a transmitir a mensagem.

#### **2.1.1 Expressões faciais e gestuais na LIBRAS**

Nos anos de 1960 e 1970, os estudos de Paul Ekman e Izard em diversas culturas demonstraram a universalidade das expressões faciais da emoção, desafiando a opinião predominante de que as emoções, como a linguagem, eram culturalmente específicas e divergentes (KELTNER; EKMAN, 2002). Junto com outros pesquisadores, esses autores mostraram ao longo do tempo que, apesar das diferenças culturais, os seres humanos têm a capacidade de construir, identificar e reconhecer um pequeno grupo de emoções, denominadas emoções básicas e universais (LEDERMAN et al., 2007). A seguir, será apresentado na Ilustração 01 como as expressões faciais são identificadas.

Ilustração 01: Expressões Faciais Básicas.



Fonte: Thaís Guimarães (2017)

No Quadro 01 abaixo, será mostrado de uma forma descritiva como as emoções são identificadas em LIBRAS.

Quadro 01: Descrição das Expressões Faciais.

<b>Expressões Emocionais</b>	<b>Descrição de como são identificadas em LIBRAS</b>
Raiva	“A raiva é manifestada através de sobrancelhas juntas e baixas, olhos fixos e intensos, e lábios firmemente comprimidos ou levemente abertos. A mandíbula pode se tensionar e as narinas podem se alargar. Essa expressão reflete um estado de prontidão para ação ou confronto”.
Nojo	“A expressão facial de nojo envolve o encurtamento do nariz, lábios levantados e puxados para os lados, e uma elevação das bochechas. A pessoa pode também contrair o lábio superior, expondo os dentes superiores, indicando aversão ou repulsa a algo desagradável”.
Medo	“O medo é identificado por olhos arregalados, com as pálpebras superiores levantadas e as sobrancelhas levantadas e achatadas. A boca pode estar aberta com os lábios retraídos para trás, e a expressão facial geral transmite uma prontidão para escapar ou se proteger”
Alegria	"A alegria é reconhecida por movimentos faciais específicos: franzir horizontal da face, franzir da testa,

	elevação da pele e das sobrancelhas, subir ligeiro das pálpebras, contração das pálpebras inferiores, olhos dilatados e semicerrados, contração das têmporas, elevação da raiz e das zonas laterais do nariz, contração das bochechas, alargamento e elevação dos cantos da boca, e esticar do queixo".
Tristeza	“A tristeza é geralmente caracterizada por sobrancelhas levantadas no centro, olhos levemente fechados, com cantos internos das sobrancelhas levantados, o que dá uma aparência de arqueamento para cima. A boca pode ser ligeiramente puxada para baixo nos cantos, criando uma expressão caída. As bochechas tendem a relaxar e os lábios podem tremer levemente”.
Surpresa	“A surpresa é demonstrada através de olhos bem abertos, sobrancelhas levantadas e arqueadas, e a boca pode estar ligeiramente aberta. Esta expressão reflete uma reação súbita a um evento inesperado, preparando a pessoa para uma resposta rápida”.

Fonte: Freitas-Magalhães (2011), Izard (1996), Ekman (1999)

Essas descrições mostram como diferentes emoções são comunicadas através de expressões faciais em LIBRAS, destacando a riqueza e a complexidade da linguagem facial. A capacidade de reconhecer e interpretar essas expressões faciais é fundamental para a comunicação eficaz entre surdos e ouvintes. Para complementar a demonstração de cada emoção, é importante identificar como os sinais devem ser executados. No Quadro 02 será demonstrado a execução desses sinais/gestos.

Quadro 02: Descrição das Expressões Gestuais.

<b>Expressões Emocionais</b>	<b>Descrição de como são identificadas em LIBRAS</b>
Raiva	Para sinalizar “Raiva” em Língua Brasileira de Sinais (Libras), é realizado o fechamento da mão com o polegar direcionado ao peito.
Nojo	Para sinalizar “Nojo” em Língua Brasileira de Sinais (Libras), é posicionado em frente ao peito a mão direita aberta, com os dedos espaçados, e o polegar junto ao peito com a mão parcialmente inclinada.

Medo	Para sinalizar “Medo” em Língua Brasileira de Sinais (Libras), abra sua mão, curve o dedo médio o dedão, tocando-os. Encoste a ponta dos dedos no peito.
Alegria	Para sinalizar “Alegria” em Língua Brasileira de Sinais (Libras), é feito um gesto de pinça com os dedos, indicador e polegar, com o restante dos dedos posicionados para cima.
Tristeza	Para sinalizar “Triste” em Língua Brasileira de Sinais (Libras), é realizado o fechamento da mão, deixando apenas o polegar com o dedo mínimo, o dedo mínimo é apontado para baixo e o polegar é posicionado no queixo.
Surpresa	Para sinalizar “Surpresa” em Língua Brasileira de Sinais (Libras), é colocado as duas mãos abertas, dedos juntos e polegar afastado, um pouco abaixo do ombro de forma vertical.

Fonte: BAIXESOFT (2023)

No presente trabalho, foram escolhidas quatro emoções principais: Alegria, Tristeza, Raiva e Surpresa. Essas emoções são amplamente reconhecidas como universais e essenciais para a comunicação humana. A escolha dessas emoções visa facilitar a identificação e interpretação de expressões faciais e gestuais em LIBRAS, promovendo um aprendizado mais eficaz e inclusivo para crianças em idade escolar. Contudo, focar nessas emoções ajudará a desenvolver e aprimorar a inteligência emocional através do ensino de LIBRAS.

## 2.2 TRABALHOS RELACIONADOS

O primeiro trabalho foi desenvolvido por Romário Pereira da Silva, com o seguinte título "Visão computacional: um estudo de caso aplicado à língua brasileira de sinais (LIBRAS)". Este trabalho aborda a aplicação da visão computacional na interpretação e reconhecimento da Língua Brasileira de Sinais (LIBRAS). Ele evidencia a relevância da tecnologia na promoção da acessibilidade e inclusão de pessoas surdas ou com deficiência auditiva. Esse estudo foi orientado pelo Prof. Me. Alan Klinger Sousa Alves e apresentado em 2022 no Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte.

O objetivo do trabalho foi desenvolver um tradutor de LIBRAS para o português utilizando visão computacional. Para isso, foram utilizadas tecnologias como o framework

MediaPipe e OpenCV, juntamente com a linguagem de programação Python. O processo de desenvolvimento seguiu uma abordagem que identifica individualmente os parâmetros formadores de cada sinal da LIBRAS para posteriormente combiná-los e realizar a tradução.

O resultado final do trabalho foi a criação de um protótipo funcional de tradutor de LIBRAS para o português, capaz de traduzir quatro palavras da LIBRAS (cunhado, Deus, direito e muro) e as letras A, B, C e D. O desenvolvimento do tradutor destacou a importância de dividir o processo de tradução em módulos individuais para explorar tecnologias com melhor desempenho em cada etapa.

O intuito deste projeto foi conceber uma solução que simplifica a interação entre surdos e ouvintes, visando a transpor os obstáculos presentes nesse domínio. Mesmo diante dos progressos tecnológicos em outras esferas, a capacidade de tradução eficaz da linguagem gestual permanece como um desafio a ser enfrentado e superado. Como trabalhos futuros, sugere-se o desenvolvimento de uma versão para tradução de diálogos simples, uma versão para plataformas mobile e a combinação do tradutor de LIBRAS para português com um tradutor de português para LIBRAS. Desse modo, este estudo demonstra a aplicação prática da visão computacional na área da comunicação inclusiva, contribuindo para a acessibilidade e a interação entre pessoas surdas e ouvintes.

O segundo trabalho, se refere à autora Karolyne Pereira da Silva, estudante de Engenharia de Controle e Automação da Universidade Federal do Rio Grande do Sul. Com isso o trabalho "Análise de Aplicação de Visão Computacional e Redes Neurais, em Conjunto com o Uso de Técnicas de Aumento de Dados, na Tradução Automática de LIBRAS" foi conduzido e realizado em 2023 a motivos de investigar a viabilidade do emprego de redes neurais em conjunto com visão computacional para identificar sinais da Língua Brasileira de Sinais (LIBRAS).

Este estudo foi dedicado à área de tecnologia assistiva e tradução automática de LIBRAS. Para o treinamento das redes neurais, foram utilizadas as bibliotecas TensorFlow e Keras, que são, respectivamente, uma plataforma de aprendizado de máquina de código aberto desenvolvida pela Google e uma API de alto nível para redes neurais, simplificando a criação e treinamento de modelos sobre o TensorFlow. A linguagem de programação Python foi adotada para programar todas as bibliotecas e pacotes necessários.

O principal objetivo deste trabalho foi explorar e aplicar técnicas de visão computacional e redes neurais com unidades de memória de longo e curto prazo (*LSTM*) na tradução automática de LIBRAS, com o intuito de aprimorar a comunicação entre surdos e ouvintes. Os resultados obtidos demonstraram promissores avanços no treinamento das redes



neurais. Dentre os conjuntos de dados analisados, destacaram-se os números 2, 4, 7 e 9, evidenciando métricas superiores e indicando a viabilidade do uso dessas tecnologias não somente na tradução automática de LIBRAS, mas também em outras línguas de sinais.

Em resumo, este estudo teve como objetivo avançar na área de tradução automática de LIBRAS, explorando a aplicação de visão computacional e redes neurais LSTM. Os resultados obtidos foram promissores, sugerindo que essa abordagem pode contribuir significativamente para a melhoria da acessibilidade e comunicação para a comunidade surda. A utilização das bibliotecas TensorFlow e Keras facilitou o desenvolvimento e treinamento dos modelos de redes neurais, permitindo a obtenção de resultados satisfatórios.

Como terceiro trabalho, intitulado "Reconhecimento de Libras em frames estáticos de vídeos utilizando CNN e técnicas de pré-processamento de imagens" foi desenvolvido por Matheus Macêdo Claudino como parte de seu Trabalho de Conclusão de Curso (TCC) para a obtenção do título de Bacharel em Ciência da Computação na Universidade Federal de Campina Grande. Realizado em 2022.

O objetivo principal do trabalho foi desenvolver um modelo de aprendizado de máquina capaz de reconhecer um conjunto limitado e pré-determinado de símbolos da Língua Brasileira de Sinais (Libras) a partir de frames estáticos de vídeos. Para alcançar esse objetivo, foi adotada a abordagem de redes neurais convolucionais em 2D (*CNN-2D*) juntamente com técnicas de pré-processamento de imagens. O estudo teve como foco avaliar o impacto dessas técnicas no desempenho do modelo de reconhecimento de Libras.

O desenvolvimento do trabalho foi realizado utilizando a plataforma Google Collaboratory, que oferece suporte de GPU para a execução dos scripts. A implementação da arquitetura de redes neurais convolucionais foi feita utilizando as bibliotecas Keras e TensorFlow. O pré-processamento de imagens, por sua vez, contou com o auxílio da biblioteca OpenCV. A arquitetura da *CNN-2D* projetada consistiu em três sequências de convolução 2D com camadas de ativação ReLU, seguidas por camadas de Max Pooling e camadas totalmente conectadas, além de uma camada de dropout para regularização. O otimizador utilizado foi o Adam, com uma taxa de aprendizado de 0.001, e a função de perda adotada foi a Sparse Categorical Cross Entropy.

O modelo desenvolvido alcançou uma acurácia média de 90.08% e um f1-score aceitável em torno de 0.5 na classificação de 20 classes pré-definidas de Libras, utilizando dados originais do conjunto de dados MINDS-Libras normalizados. Esses resultados demonstraram a eficácia da abordagem proposta para o reconhecimento de sinais de Libras em frames estáticos de vídeos. O trabalho foi desenvolvido com o intuito de contribuir para o

avanço da tecnologia de reconhecimento de Libras, possibilitando a criação de sistemas mais eficientes e precisos para auxiliar pessoas com deficiência auditiva na comunicação e interação social. Ao explorar novas técnicas e metodologias, Matheus Macêdo Claudino buscou não apenas melhorar o desempenho dos modelos de reconhecimento de Libras, mas também promover uma maior inclusão social através da tecnologia. No Quadro 03 podem ser observados os principais pontos de cada trabalho.

Quadro 03: Descrição dos trabalhos.

<b>Trabalho</b>	<b>Tecnologias usadas</b>	<b>Resultados finais/taxa de assertividade</b>	<b>Principais pontos</b>
Romário Pereira da Silva	MediaPipe OpenCV Python	Protótipo funcional traduzindo 4 palavras e 4 letras de LIBRAS para português, com assertividade de: Cunhado: 44,4%, Muro: 55,5%, Direito: 44,4%, Deus: 44,4%	Desenvolvimento de tradutor de LIBRAS para português. Divisão do processo de tradução em módulos individuais. Testes realizados sem controle de iluminação
Karolyne Pereira da Silva	TensorFlow, Keras, Python, Redes Neurais LSTM	Promissores avanços no treinamento das redes neurais, especialmente em conjuntos de dados 2, 4, 7 e 9	Exploração de técnicas de visão computacional e redes neurais LSTM para tradução automática de LIBRAS. Uso de técnicas de aumento de dados.
Matheus Macêdo Claudino	TensorFlow, Keras, OpenCV, Google Collaboratory , CNN-2D	Acurácia média de 90.08%, f1-score de 0.5 na classificação de 20 classes de Libras	Uso de CNN-2D para reconhecimento de sinais de Libras em frames estáticos; Avaliação do impacto de técnicas de pré-processamento de imagens

Fonte: A autora (2024).

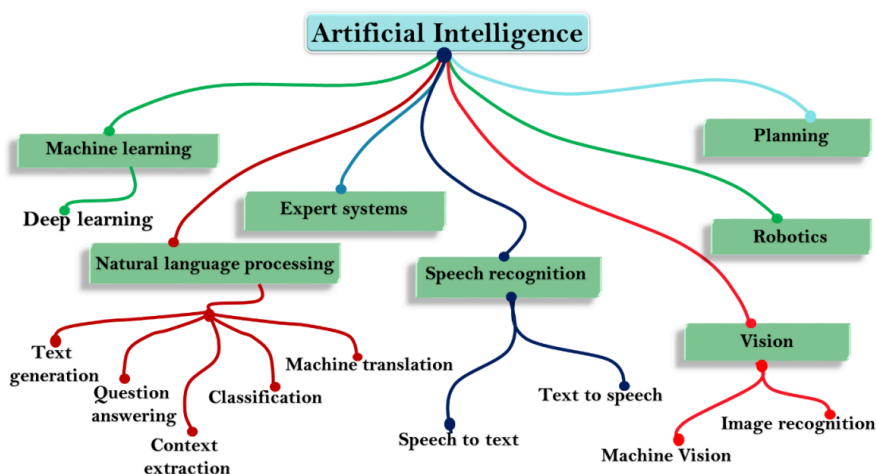
## 2.3 INTELIGÊNCIA ARTIFICIAL

Cozman, Plonski e Neri (2021), definem Inteligência Artificial como programas computacionais capazes de raciocinar sobre conhecimentos, tomar decisões e aprender a partir desse raciocínio. Esses programas são conhecidos como agentes inteligentes, os quais operam de forma autônoma, percebendo e se adaptando ao ambiente ao qual estão inseridos. Atualmente, a IA abrange uma série de campos, desde o aprendizado e percepção até tarefas mais específicas, por exemplo: resolução de jogos, demonstração de teoremas matemáticos, carros autônomos e diagnóstico de doenças (RUSSELL; NORVIG, 2013; LUGER, 2013).

Nesse contexto, o presente trabalho está localizado nos campos de Reconhecimento de Imagem e no subcampo de Machine Learning conhecido como Deep Learning, sendo esse um campo de pesquisa mais recente que aborda técnicas de aprendizado baseadas em redes neurais, como é o caso das Redes Neurais Convolucionais (CNNs). Deep Learning, ou aprendizado profundo, refere-se a um conjunto de algoritmos de aprendizado de máquina que modelam dados em camadas sucessivas de abstração, permitindo a criação de modelos que podem aprender representações complexas dos dados (RUSSELL; NORVIG, 2013).

Essas técnicas são eficazes em tarefas como reconhecimento de imagem, processamento de linguagem natural e outras áreas que envolvem grandes volumes de dados não estruturados. A Ilustração 02 apresenta os principais subcampos da Inteligência Artificial (IA), posteriormente serão explicados alguns conceitos fundamentais de IA para aplicação neste trabalho.

Ilustração 02: Principais ramos de estudo da IA.



Fonte: JavaPoint (2024)

### **2.3.1 Visão computacional e Aprendizado de Máquina**

Segundo Barelli (2018), a visão computacional, enquanto ciência, pode ser entendida como aquela que estuda e desenvolve tecnologias permitindo que máquinas extraiam dados do ambiente ao redor através de imagens capturadas por sensores. Sendo assim, a visão computacional busca emular a visão humana no computador, ou seja, a partir da coleta de imagens de uma câmera ou outro sistema de aquisição, o computador pode analisar esses dados e obter características que auxiliem na tomada de decisões (BUDIHARTO; GROUP, 2014).

O aprendizado de máquina é um ramo da Inteligência Artificial focado em criar técnicas para que sistemas aprendam automaticamente a partir de dados. Em visão computacional, ele é crucial, pois permite que algoritmos aprendam e façam previsões com base nos dados. Existem diferentes tipos de aprendizado de máquina: o supervisionado, o não supervisionado, e o aprendizado profundo (APPMASER, 2023).

O aprendizado supervisionado é útil para resolver problemas como classificação (atribuindo categorias a dados) e regressão (entendendo relações entre variáveis). Ele utiliza conjuntos de treinamento com entradas e saídas corretas para aprender e ajustar seus parâmetros. Algoritmos como redes neurais, Naive Bayes e SVM são comuns nessa abordagem (IBM, 2024). Por outro lado, o aprendizado não supervisionado é usado quando não há rótulos nos dados e o objetivo é encontrar padrões ou agrupamentos. O clustering, por exemplo, é uma técnica importante que agrupa itens semelhantes (FILHO, 2017).

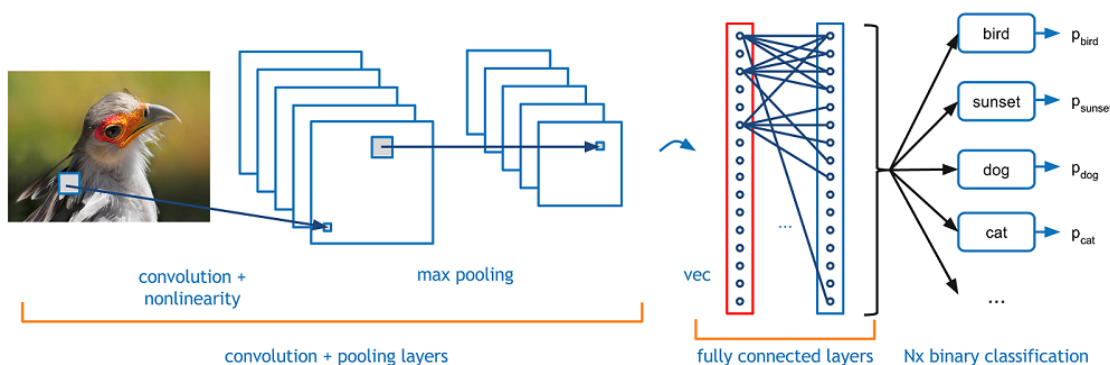
A aprendizagem profunda é uma técnica avançada de aprendizado de máquina que utiliza redes neurais artificiais para imitar a capacidade do cérebro humano de processar informações. Essa abordagem organiza algoritmos em camadas de neurônios artificiais, facilitando o tratamento de dados como imagens ou linguagem. Para ser considerada profunda, uma rede neural precisa ter múltiplas camadas entre suas camadas de entrada e saída. Essas camadas intermediárias são essenciais para a complexidade do modelo e sua eficácia. Existem três tipos principais de arquiteturas de aprendizagem profunda: as Redes Neurais Feedforward (FF), as Redes Neurais Recorrentes (RNN) e as Redes Neurais Convolucionais (CNN) (PURESTORAGE, 2024).

### 2.3.2 Redes Neurais Convolucionais

De acordo com o Deep Learning Book (2022), em 1998, Yann LeCun e seus colaboradores desenvolveram um reconhecedor eficiente para dígitos manuscritos chamado LeNet. Este modelo utiliza várias camadas ocultas, com múltiplos mapas de unidades replicadas em cada camada, agrupando as saídas de unidades próximas, formando uma rede ampla lidando com vários caracteres simultaneamente. Posteriormente, essa arquitetura foi formalizada sob o nome de redes neurais convolucionais.

As Redes Neurais Convolucionais (CNNs) são redes neurais artificiais profundas usadas para classificar imagens, agrupá-las por similaridade (busca de fotos) e realizar reconhecimento de objetos em cenas. Esses algoritmos podem identificar rostos, indivíduos, sinais de rua, cenouras, e muitos outros aspectos dos dados visuais. Na prática, uma Rede Neural Convolucional pode ser representada classificando imagens, conforme a Ilustração 03, referenciando o processo que será desenvolvido no projeto para identificar as quatro emoções escolhidas: felicidade, tristeza, raiva e surpresa.

Ilustração 03: Classificação de imagens.



Fonte: STROSKI (2018)

À medida que a imagem passa pela rede convolucional, ela é processada em volumes de entrada e saída, representados como matrizes tridimensionais, como 30x30x3. De camada em camada, suas dimensões mudam até que a rede produza uma série de probabilidades na camada de saída. Por exemplo, a classe com maior probabilidade é a saída final da rede para a imagem de entrada, como "pássaro" (Deep Learning Book, 2022).

Na Ilustração 03, é perceptível que a rede neural começa com uma imagem como entrada. Conforme Stroski (2018), na camada de Convolução e Não-Linearidade, a imagem

passa por várias camadas onde filtros são aplicados para extrair características básicas como bordas e texturas. Depois de cada camada de convolução, uma função de ativação não-linear (como *ReLU*) é usada para introduzir complexidade no modelo. Quanto mais camadas de convolução, mais detalhados são os traços capturados. Esses filtros, chamados de kernels, começam com valores aleatórios e são ajustados durante o treinamento. Cada filtro é aplicado a uma pequena área da imagem, chamada de campo receptivo, para capturar padrões locais.

Segundo Lacerda (2019), na camada de Pooling, após a convolução, a camada de pooling reduz a dimensão dos dados, mantendo as características mais importantes e diminuindo o número de parâmetros e o custo computacional. No "max pooling", o valor máximo de uma região específica da imagem é selecionado. Nas camadas totalmente conectadas (*fully connected layers*), cada neurônio é conectado a todos os neurônios da camada anterior, permitindo que a rede aprenda combinações complexas das características extraídas.

Finalmente, na classificação, a saída das camadas totalmente conectadas é usada para classificar a imagem em uma ou mais categorias. No exemplo, a rede pode classificar a imagem como "pássaro", "pôr do sol", "cachorro" ou "gato". Cada saída mostra a probabilidade de a imagem pertencer a uma dessas classes.

#### 2.3.2.1 Arquiteturas de Redes Neurais Convolucionais (CNN)

As Redes Neurais Convolucionais (CNNs) são fundamentais para avanços no reconhecimento de imagens e outras tarefas de visão computacional. Ao longo dos anos, diversas arquiteturas foram desenvolvidas, cada uma trazendo inovações que aprimoram a eficiência e a precisão dos modelos (AI Summer, 2024; Manning, 2024).

MobileNetV2, desenvolvida por Sandler et al. em 2018, é uma arquitetura eficiente em termos de recursos, projetada para dispositivos móveis e aplicações com restrições de hardware. MobileNetV2 utiliza convoluções separáveis por profundidade, reduzindo significativamente o número de parâmetros e operações computacionais, enquanto mantém alta precisão na classificação de imagens.

VGGNet, desenvolvida por Simonyan e Zisserman em 2014, é conhecida pela simplicidade e uso de filtros de convolução de 3x3. Consiste em 16 a 19 camadas, enfatizando a profundidade na rede. GoogLeNet (*Inception*), desenvolvida por Szegedy et al. em 2014, introduziu o módulo Inception, que realiza convoluções com múltiplos tamanhos de filtro.

Reduziu significativamente os parâmetros através de técnicas de redução dimensional, vencendo a competição ImageNet em 2014.

ResNet, desenvolvida por He et al. em 2015, resolve o problema do gradiente desaparecido com blocos residuais e conexões de atalho, permitindo redes mais profundas. Possui versões com 50, 101 ou 152 camadas. AlexNet, criada por Alex Krizhevsky em 2012, possui oito camadas: cinco de convolução e três totalmente conectadas, introduzindo a ativação ReLU e o dropout para evitar o overfitting.

Essas redes pré-treinadas representam marcos importantes no desenvolvimento das CNNs, cada uma contribuindo com inovações que permitiram avanços significativos em tarefas de visão computacional, destacando-se na precisão e eficiência dos modelos de aprendizado profundo.

## 2.4 TECNOLOGIAS

### 2.4.1 TensorFlow

De acordo com o site Tech (2022), em 2011, o Google Brain desenvolveu um sistema chamado DistBelief para redes neurais e deep learning. Esse sistema foi amplamente utilizado em várias aplicações comerciais e pesquisas dentro do Google. Para simplificar e acelerar o DistBelief, o Google formou uma equipe de cientistas da computação que redesenhou o sistema, resultando no TensorFlow.

O TensorFlow foi lançado como código aberto em novembro de 2015, com a versão 1.0.0 lançada em fevereiro de 2017 e a versão 2.0 em janeiro de 2018. Quando um software é código aberto, seu código-fonte está disponível para qualquer pessoa visualizar, modificar e distribuir gratuitamente. Isso fomenta uma comunidade ativa de colaboradores que continuamente melhoram o software. O TensorFlow, por exemplo, tem mais de 2.000 contribuidores no GitHub.

Sendo assim, o TensorFlow é uma biblioteca para machine learning, similar a uma biblioteca Python como Numpy, Scikit-learn e Pandas. Programar em TensorFlow é muito semelhante a programar em Python, facilitando a integração e a utilização das mesmas ferramentas e interfaces. Uma vez instalado, o TensorFlow pode ser usado na mesma interface de desenvolvimento que Python, permitindo a execução conjunta de códigos Python e TensorFlow.

Juntamente com o TensorFlow, é relevante abordar o Kera, sendo uma API que abstrai a complexidade da programação de outras API's destinadas ao aprendizado de máquina, como

é o caso do TensorFlow. Tem o intuito de ser uma ferramenta de prototipagem rápida, como destaca em seu site: “Ser capaz de ir da ideia ao resultado com o menor atraso possível é a chave para fazer uma boa pesquisa” (KERAS, 2018). Keras e TensorFlow são duas ferramentas populares no campo de aprendizado profundo (deep learning) e frequentemente são usadas juntas para construir e treinar redes neurais.

Dessa maneira, para trabalhar com redes neurais, utilizar o TensorFlow é uma escolha inteligente facilitando a implementação de conceitos teóricos em deep learning, processamento de linguagem natural e visão computacional.

### 2.4.2 OpenCV

Desenvolvida inicialmente pela Intel em 1999, a OpenCV é uma biblioteca de código aberto com um conjunto abrangente de rotinas e implementações técnicas para processamento digital de imagens (PDI). Escrita em C/C++ otimizado, a OpenCV possui interfaces para C++, Python e Java, o que a torna extremamente versátil e acessível para desenvolvedores de diferentes linguagens de programação (OPENCV, 2018).

Além disso, a OpenCV é reconhecida como a maior e mais utilizada biblioteca para PDI, sendo amplamente adotada em aplicações acadêmicas e comerciais devido à sua eficiência e ampla gama de funcionalidades (BERNART, 2014). Seu uso abrange desde simples manipulações de imagens até complexos algoritmos de visão computacional, contribuindo significativamente para avanços em áreas como reconhecimento de padrões, realidade aumentada e robótica.

## 2.5 PROCESSAMENTO DE IMAGENS DIGITAIS

De acordo com Gonzalez e Woods (2010), uma imagem pode ser definida pelo agrupamento de vários pixels, os quais podem ser representados pela função  $f(x,y)$ , onde  $x$  e  $y$  são as coordenadas espaciais e  $f(x,y)$  é o nível de cinza presente nas coordenadas especificadas da imagem. O processamento de imagens consiste de um conjunto de técnicas computacionais que tem por objetivo fornecer às máquinas a capacidade de obter, reconhecer e interpretar as imagens digitais. Essas podem ser obtidas a partir de diversas fontes, por exemplo, através do espectro eletromagnético, ultrassônico, acústica e eletrônica (GONZALEZ; WOODS, 2010).

Sendo assim, segundo Marques Filho e Vieira Neto (1999), o processamento de imagem pode ser aplicado em duas áreas distintas. Uma delas é o aprimoramento das imagens para facilitar a interpretação humana e a outra área é a análise automática feita por



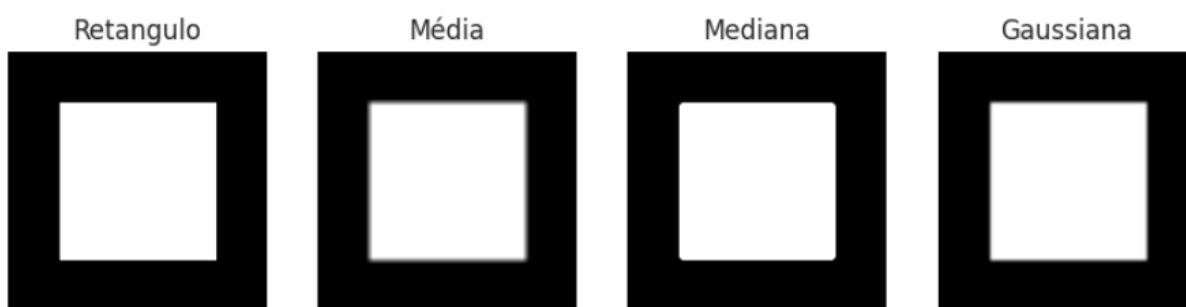
computadores através de informações obtidas em imagens. Neste projeto, o processamento de imagem será feito com a biblioteca OpenCV para Python, a qual possui diversos métodos que podem ser aplicados a este trabalho. Nas próximas subseções, serão apresentadas algumas técnicas de PDI que fundamentam o presente trabalho.

### 2.5.1 Suavização

Este tipo de filtro é utilizado quando se tem a intenção de remover ruídos ou aplicar borramento na imagem. O borramento pode ser necessário quando deseja-se remover pequenos detalhes da imagem antes da extração de objetos grandes, ou quando deseja-se corrigir pequenas discontinuidades em linhas ou curvas. A suavização consiste em alterar os pixels da imagem de forma a remover mudanças abruptas de tonalidades (GONZALEZ; WOODS, 2010).

A OpenCV disponibiliza três funções para aplicar suavização em imagens, sendo elas *blur*, *medianBlur* e *gaussianBlur*. A diferença entre elas é o cálculo que é feito para determinar a nova cor de cada pixel. O exemplo de aplicação dos métodos de suavização podem ser vistos na Ilustração 05.

Ilustração 05: Aplicando Métodos de Suavização.



Fonte: A autora (2024).

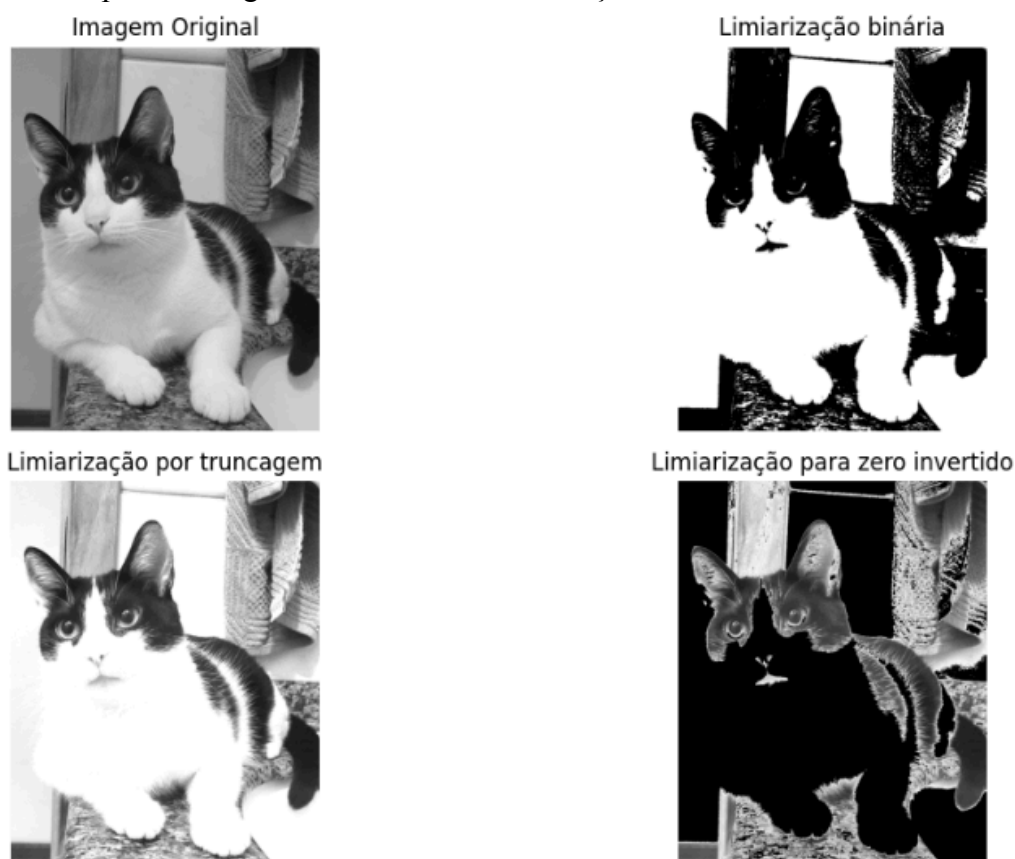
Como mostrado na Ilustração 05, temos que a suavização média, implementada pela função *cv2.blur*, calcula a média dos pixels vizinhos em uma janela definida, resultando em uma imagem onde cada pixel é a média dos seus vizinhos. A suavização por mediana, implementada pela função *cv2.medianBlur*, substitui cada pixel pela mediana dos pixels vizinhos, sendo eficaz na remoção de ruídos. A suavização Gaussiana, implementada pela função *cv2.GaussianBlur*, utiliza uma função gaussiana para calcular a média ponderada dos pixels vizinhos, preservando melhor as bordas e suavizando de maneira mais natural. Cada

método é escolhido conforme o tipo de ruído e as características da imagem (ARTERO; TOMASSELLI, 2018).

### 2.5.2 Limiarização

Conforme Marques Filho e Vieira Neto, (1999), a limiarização é amplamente utilizada para separar o objeto do fundo da imagem. Este processo pode ser feito pela separação do histograma da imagem, onde pixels com tons de cinza acima de um valor específico são convertidos para branco, e os demais para preto. Assim, os pixels com valor 0 (preto) representam o fundo da imagem, e os com valor 1 (branco) representam o objeto. Alguns exemplos de aplicação dos métodos de Limiarização podem ser vistos na Ilustração 06.

Ilustração 06: Aplicando alguns Métodos de Limiarização.



Fonte: A autora (2024).

A Ilustração 06 mostra diferentes tipos de limiarização aplicados usando a função *cv2.threshold* da biblioteca OpenCV. Esses métodos incluem a limiarização binária, por truncagem, e para zero invertida. Cada método transforma os pixels da imagem de maneiras específicas com base em um valor de limiar definido. Esses métodos são úteis para diferentes tipos de processamento de imagens, dependendo das características da imagem e do resultado

desejado (ARTERO; TOMASSELLI, 2018). A seguir, os tipos de limiarização são apresentados em detalhes.

Limiarização Binária (*cv2.THRESH\_BINARY*):

- Transforma pixels acima do limiar em branco (255).
- Transforma pixels abaixo do limiar em preto (0).
- Limiarização por Truncagem (*cv2.THRESH\_TRUNC*):
  - Mantém os pixels acima do limiar no valor do limiar.
  - Mantém os pixels abaixo do limiar inalterados.
- Limiarização para Zero Invertida (*cv2.THRESH\_TOZERO\_INV*):
  - Transforma pixels abaixo do limiar em zero.
  - Mantém os pixels acima do limiar inalterados.

### 2.5.3 Erosão e Dilatação

Segundo Gonzalez e Woods (2010), a erosão é usada para remover ou reduzir o tamanho de objetos em uma imagem. Esse processo requer a definição de um elemento estruturante, cujo tamanho determina a extensão da redução dos objetos na imagem. Objetos menores que o elemento estruturante são completamente removidos. Sendo assim, em oposto se tem a dilatação, que tem o intuito de aumentar os objetos em uma imagem binária. O formato e o tamanho do aumento são controlados pelo elemento estruturante, assim como na erosão.

A principal aplicação da dilatação é a união de lacunas. A OpenCV oferece os métodos *erode* e *dilate*, usados respectivamente para aplicar a erosão e a dilatação em imagens. O exemplo de aplicação dos métodos de erosão e dilatação podem ser vistos na Ilustração 07.

Ilustração 07: Aplicação do Método de Erosão e Dilatação.



Fonte: A autora (2024).

### 2.5.4 Detecção de Borda

Os algoritmos de detecção de bordas são usados para identificar as bordas dos objetos presentes em uma imagem. Entre esses algoritmos, o de *Canny* é um dos mais complexos, mas oferece desempenho superior aos demais. O algoritmo de *Canny* é centrado em 3 objetivos, sendo eles a baixa taxa de erro, deve haver uma resposta de um único ponto de borda e os pontos de borda devem estar bem localizados (GONZALEZ; WOODS, 2010). A biblioteca OpenCV oferece um método que aplica todas as etapas do algoritmo de detecção de bordas de *Canny*. Para utilizá-lo, basta passar a imagem desejada e os valores mínimo e máximo para que um pixel seja considerado como borda. O método retorna uma nova imagem contendo apenas as bordas detectadas. O exemplo de aplicação do método de *Canny* pode ser visto na Ilustração 08.

Ilustração 08: Aplicação do Método de Detecção de Bordas de *Canny*.



Fonte: A autora (2024).

### **3 PROCEDIMENTOS METODOLÓGICOS E TÉCNICOS**

Este módulo aborda os procedimentos metodológicos e técnicos adotados neste projeto, detalhando as etapas e métodos empregados para alcançar os objetivos propostos. A metodologia é estruturada para garantir a aplicação eficaz e a validação das técnicas desenvolvidas, com foco no reconhecimento de emoções em crianças utilizando a Língua Brasileira de Sinais (LIBRAS).

#### **3.1 CARACTERIZAÇÃO DA METODOLOGIA DE PESQUISA**

Este estudo é categorizado como aplicado, dado que, conforme Zanella (2011), sua natureza visa propor soluções para desafios enfrentados pela sociedade. Fleury e Werlang (2016) definem a pesquisa aplicada como um conjunto de atividades que emprega conhecimentos pré-existentes para coletar, selecionar e processar fatos e dados, visando confirmar resultados e gerar impacto.

Para abordar o problema em questão, será adotada uma abordagem com métodos quantitativos. Segundo Fonseca (2002), a pesquisa quantitativa se caracteriza pela quantificação dos resultados, refletindo um retrato fiel da população-alvo devido ao amplo número de amostras. Sendo assim, essa pesquisa é quantitativa no que diz respeito às métricas do modelo. Para alcançar os objetivos propostos, será adotada uma pesquisa exploratória. Gil (2002) destaca que esse tipo de pesquisa busca proporcionar uma maior familiaridade com o tema, estimulando o desenvolvimento e o refinamento de ideias. Frequentemente, a pesquisa exploratória é conduzida por meio de levantamentos bibliográficos, entrevistas com especialistas e análise de casos exemplares.

Os procedimentos técnicos adotados para a elaboração deste projeto envolvem pesquisas bibliográficas, aliadas à orientação da professora de pedagogia, visando consolidar uma compreensão mais abrangente da teoria sobre LIBRAS. A pesquisa bibliográfica consiste na identificação e análise de referências teóricas previamente examinadas e publicadas, tanto em formato impresso quanto eletrônico (Fonseca, 2002).

#### **3.2 DELIMITAÇÃO DO ESTUDO**

O presente trabalho tem como objetivo desenvolver um modelo para auxiliar o ensino de Libras na identificação de emoções em crianças, especificamente felicidade, tristeza, raiva

e surpresa, utilizando gestos e expressões faciais. O objetivo é criar um modelo preciso para o diagnóstico emocional e para apoiar o aprendizado das crianças na escola. O desenvolvimento do modelo será dividido em duas etapas principais: pesquisa bibliográfica no primeiro semestre de 2024 e implementação prática no segundo semestre do mesmo ano. Durante a implementação, o modelo será testado utilizando métricas de validação, para que posteriormente possa ser utilizado pela professora Marineiva Moro Campos de Oliveira na UNOESC. A intenção é avaliar a eficácia do modelo no ensino de Libras e na compreensão das emoções. O projeto é liderado por Angeline Zanela, sob a orientação de Jacson Luiz Matte.

### 3.2.1 População e Amostra

A população da pesquisa consiste em todo o *dataset* utilizado no desenvolvimento da aplicação de reconhecimento de emoções em LIBRAS, que é dividido em dois conjuntos de dados, um para treino e outro para teste. Esse *dataset* é composto por aproximadamente 600 imagens, divididas em 4 classes de emoções. Cada classe, "Felicidade", "Tristeza", "Raiva" e "Surpresa", contém cerca de 150 imagens. Essas imagens são destinadas à identificação de expressões faciais e gestuais.

O conjunto de dados de teste é composto por aproximadamente 250 imagens, representando as mesmas emoções. Essas imagens são capturadas em diferentes ângulos e níveis de iluminação para avaliar a precisão do modelo em condições variadas.

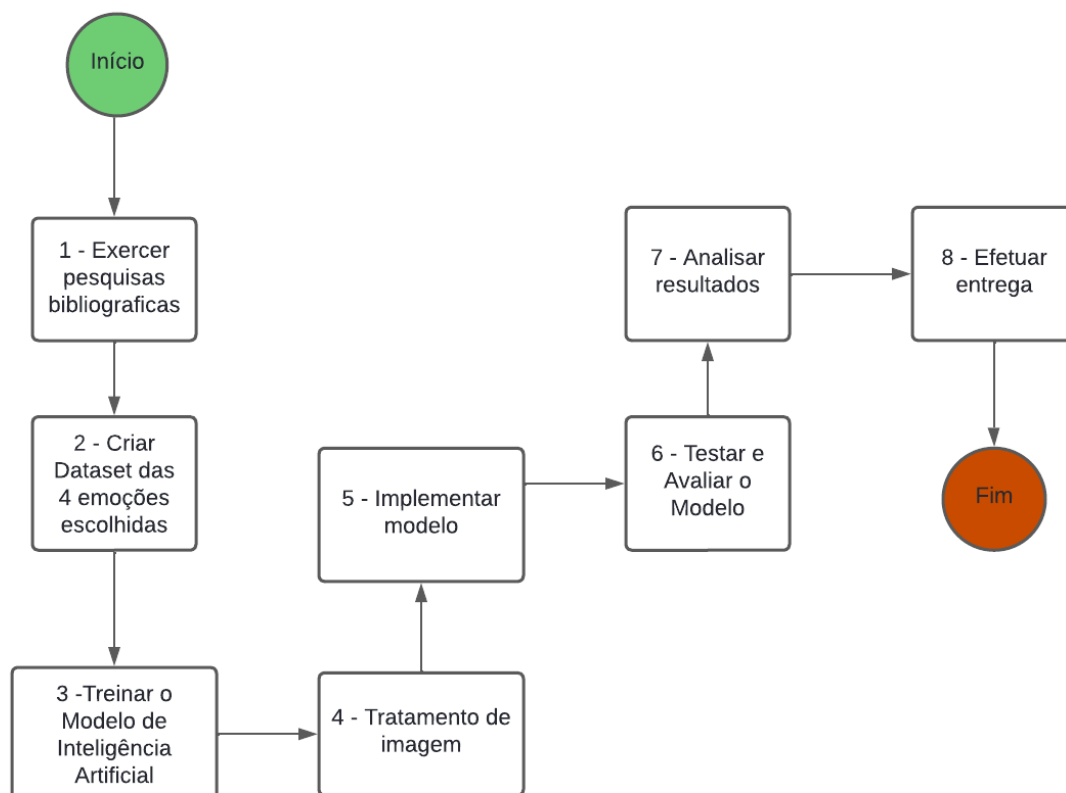
## 3.3 QUESTÕES DE PESQUISA

- É possível criar um modelo com visão computacional para identificação de emoções em Libras por meio do reconhecimento facial e gestual, esse modelo criado conseguirá ter métricas aceitáveis para essas 4 classes de emoções?

## 3.4 ESBOÇO DO PROJETO NA PRÁTICA

Este projeto tem como objetivo identificar e testar quatro emoções principais em crianças para facilitar o entendimento da Língua Brasileira de Sinais (LIBRAS). Para atingir esse propósito, foi elaborado um fluxograma que descreve as etapas necessárias para alcançar os objetivos estabelecidos.

Ilustração 09: Fluxograma do Projeto na Prática



Fonte: A autora (2024).

O fluxograma ilustra as etapas do desenvolvimento do projeto, detalhando o processo desde a pesquisa inicial até a entrega final do modelo de reconhecimento de emoções em LIBRAS. A primeira etapa é exercer pesquisas bibliográficas, onde é realizada uma revisão abrangente da literatura existente sobre LIBRAS, reconhecimento de emoções e técnicas de visão computacional. O objetivo dessa etapa é fornecer uma base teórica sólida e identificar métodos e tecnologias que podem ser aplicados ao projeto.

A segunda etapa é criar um *dataset* das quatro emoções escolhidas. Isso envolve coletar e preparar um conjunto de dados contendo imagens que representam felicidade, tristeza, raiva e surpresa. O objetivo é garantir que o conjunto de dados seja robusto e diversificado, adequado para treinar o modelo de Inteligência Artificial (IA). Em seguida, na terceira etapa, o modelo de IA é treinado utilizando bibliotecas de aprendizado de máquina, como TensorFlow. O objetivo aqui é desenvolver um modelo de rede neural convolucional (CNN) capaz de reconhecer e classificar as emoções com alta precisão.

Na quarta etapa, as imagens passam por tratamento, aplicando técnicas de pré-processamento como redimensionamento, normalização e remoção de ruídos. Isso

melhora a qualidade dos dados utilizados no treinamento, garantindo que as imagens estejam em um formato adequado para o modelo. A quinta etapa envolve a implementação do modelo treinado em um ambiente de execução. O objetivo é facilitar a aplicação prática do modelo, permitindo seu uso no ambiente educacional para identificar emoções em tempo real.

A sexta etapa é testar e avaliar o modelo. Nesta fase, o modelo é submetido a testes rigorosos em diversas situações para avaliar sua precisão na identificação das emoções. As métricas de desempenho, são utilizadas para validar o modelo. Em seguida, na sétima etapa, os resultados dos testes são analisados para identificar pontos fortes e áreas que necessitam de melhorias. O objetivo é ajustar e otimizar o modelo para maximizar sua eficácia.

Finalmente, a oitava etapa é efetuar a entrega do projeto. Isso envolve documentar todo o processo e os resultados obtidos, e realizar a entrega final do modelo, incluindo sua implementação no ambiente educacional. O objetivo final é concluir o projeto com um modelo funcional e validado, pronto para ser utilizado na prática, melhorando o ensino de LIBRAS e a inclusão de crianças surdas. Este fluxograma representa um plano detalhado e estruturado, garantindo que todas as etapas necessárias para o desenvolvimento do modelo sejam seguidas de forma lógica e eficiente.

### 3.5 APLICAÇÃO DA METODOLOGIA

Neste módulo será discutido como esta pesquisa será desenvolvida, trazendo informações sobre o referencial teórico e a coleta de dados, as técnicas específicas utilizadas, e como será a construção de toda a parte prática relacionada ao trabalho em questão.

#### 3.5.1 Construção do Referencial Teórico

Para a construção do referencial teórico, foram realizadas pesquisas abrangentes em diversas áreas pertinentes ao projeto. A quantidade de obras utilizadas para cada assunto específico está detalhada na Tabela 01, assegurando uma base sólida e diversificada para a fundamentação do estudo.



Tabela 01: Quantitativo de obras usadas por assunto

<b>Assunto</b>	<b>Tipo</b>	<b>Qtd</b>
LIBRAS	Livros	4
	Artigos	6
	Sites	1
	Documentos	5
TRABALHOS RELACIONADOS	Livros	0
	Artigos	3
	Sites	0
	Documentos	0
INTELIGENCIA ARTIFICIAL	Livros	5
	Artigos	3
	Sites	9
	Documentos	1
TECNOLOGIAS	Livros	2
	Artigos	1
	Sites	3
	Documentos	1
PROCESSAMENTO DE IMAGENS DIGITAIS	Livros	2
	Artigos	1
	Sites	0
	Documentos	0

Fonte: A autora (2024).

Na Tabela 02 é possível observar o quantitativo de obras por tipo de obra, sendo que os tipos utilizados foram: livros, artigos, sites e documentos.

Tabela 2: Quantitativo Total de obras usadas

<b>Tipo</b>	<b>Qtd</b>
Livros	13
Artigos	14
Sites	13
Documentos	7
Total	47

Fonte: A autora (2024).

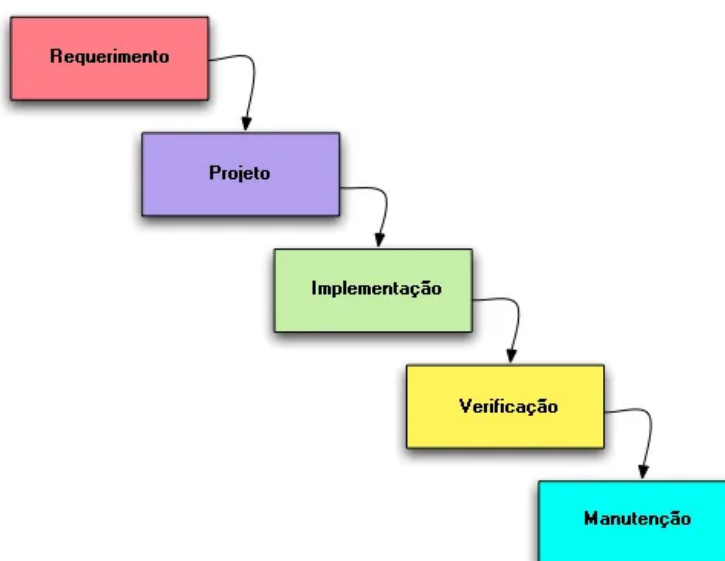
### 3.5.2 Desenvolvimento do Sistema Computacional

Segundo Sommerville (2011), o modelo em cascata se baseia em um processo dirigido a planos, sendo necessário que já exista um planejamento prévio de todas as etapas do desenvolvimento para ele possa ser aplicado, apresenta cada fase fundamental do processo como sendo distinta. Também afirma que o estágio seguinte de desenvolvimento só deve ser

iniciado quando o estágio anterior estiver completo. O fluxo do diagrama pode ser visto na ilustração abaixo.

O desenvolvimento deste projeto seguirá a metodologia em cascata. Esse modelo foi escolhido devido à natureza dos requisitos do projeto, que não devem sofrer grandes alterações ao longo do desenvolvimento. Essa estabilidade nos requisitos possibilita que cada etapa do projeto seja cuidadosamente planejada e definida antes mesmo de iniciar o desenvolvimento.

Ilustração 10: Etapas do Modelo em Cascata



Fonte: Diego Macêdo (2012).

#### 3.5.2.1 Fase de Requerimentos

Na fase de Requerimentos, todas as funcionalidades, serviços e limitações do sistema são identificados em colaboração com os usuários. O objetivo principal é compreender as necessidades do usuário e delinear o escopo, a usabilidade e o modelo de negócio do sistema. Sommerville (2011) afirma que as consultas com os usuários irão definir quais serão os serviços oferecidos, metas do sistema e suas restrições. Também é necessário fazer um levantamento de requisitos funcionais e não funcionais.

Durante essa fase, os requisitos são categorizados em funcionais e não funcionais, obtidos através de conversas com o produtor. Segundo Alff (2022), um requisito funcional descreve uma ação realizada pelo sistema, ou seja, o que o sistema deve fazer para atender a uma determinada funcionalidade.

Por outro lado, um requisito não funcional específica como o sistema deve realizar uma ação e deve ser mensurável para determinar se está sendo atendido pelo software (ALFF, 2022). Desse modo a definição dos requisitos deste projeto se dará através de debates entre a aurora e a professora Marineiva, especialista em pedagogia e ensino de LIBRAS da UNOESC Chapecó.

### 3.5.2.2 Fase de projetos

Na fase de projetos, para Sommerville (2011), ocorre a definição da arquitetura do projeto, juntamente com a identificação das abstrações. São feitas descrições de tudo que será construído através de diagrama de atividades e wireframes de interfaces.

Nesta fase, ocorre a abstração do sistema, onde todos os requisitos são agrupados e estudados para organizar o modelo que será gerado. Além disso, são estabelecidos os passos que o sistema seguirá, incluindo a definição de escopo, a elaboração do cronograma e a verificação das etapas para o desenvolvimento (STANKIEWICZ, 2017, pg. 17).

Desse modo, abordando a definição, o diagrama de atividades representa os estados de uma atividade, diferentemente dos diagramas de objetos, estes são orientados a fluxos de controle, onde cada estado representa um dos passos que o sistema está realizando. Podem ser considerados uma extensão dos fluxogramas, muito conhecidos e utilizados pelos profissionais de tecnologia da informação (BEZERRA, 2007).

Já os Wireframes, que seria a prototipação das telas. Sommerville (2011) afirma que um protótipo serve para mostrar os conceitos e as funcionalidades principais de um sistema ou software, já que este diagrama mostra a versão inicial do desenvolvimento. E conforme Azenha e Fleury (2018), wireframes são desenhos, geralmente das principais telas do sistema, que contém um certo nível de detalhamento já apresentando os padrões, textos e conteúdo que serão exibidos ao usuário final, de forma a facilitar a visualização de um produto antes mesmo do início do seu desenvolvimento.

### 3.5.2.3 Fase de Implementação

Na fase de Implementação, os programadores e desenvolvedores se unem para criar tanto o frontend quanto o backend do sistema, ou seja, é o momento da codificação efetiva (STANKIEWICZ, 2017, pg. 17). Segundo Sommerville (2011), é nesta fase que se inicia o

desenvolvimento da solução. Sendo assim a implementação começa quando o projeto de software é traduzido em um conjunto de programas ou unidades.

Durante esta fase, a solução proposta será desenvolvida, utilizando a linguagem de programação Python para o processamento de imagens, com o suporte das bibliotecas TensorFlow e OpenCV. Paralelamente, serão realizados testes para garantir que o desenvolvimento esteja de acordo com os requisitos estabelecidos.

#### 3.5.2.4 Fase de verificação

Na fase de verificação ocorre a integração e teste completos de todas as partes individuais do projeto, com o objetivo de assegurar que todos os requisitos de software foram cumpridos da forma esperada, e se comportam eficientemente como um sistema completo, capaz de apresentar resultados ao usuário (SOMMERVILLE. 2011).

O sistema é submetido a testes abrangentes para assegurar que todos os requisitos foram atendidos e que o modelo está funcionando conforme o esperado. Esses testes podem incluir testes de funcionalidade, desempenho, segurança e usabilidade, entre outros, dependendo das especificações do projeto.

Uma vez concluídos os testes e confirmada a conformidade do sistema com os requisitos, ocorre a entrega do sistema aos usuários. Esta entrega marca o final do desenvolvimento do modelo e o início de sua utilização pelo usuário para atender às suas necessidades específicas.

#### 3.5.2.5 Fase de manutenção

Na fase de manutenção, o foco está na identificação e correção de erros que surgem durante a utilização do sistema pelos usuários finais. Este estágio é crucial para garantir a estabilidade e eficiência contínuas do sistema em produção. Após a detecção desses erros, são realizadas as correções necessárias. Desse modo a manutenção do sistema é pensada para corrigir erros que não foram detectados nas fases anteriores (SOMMERVILLE, 2011).

4 APRESENTAÇÃO DA SOLUÇÃO

Neste módulo, o presente trabalho apresentará a solução de desenvolvimento do sistema computacional, levando em consideração o que foi mencionado no item 3.5.2.

4.1 REQUERIMENTOS

Foi realizada a coleta de requisitos e especificação dos mesmos no formato de requisitos funcionais e não funcionais. Dessa forma, os requisitos funcionais relacionados a este projeto estão listados no Quadro 04, e os requisitos não funcionais são apresentados no Quadro 05.

Quadro 04: Requisitos Funcionais.

RF01	O sistema deve capturar imagens em tempo real.
RF02	O sistema deve integrar processamento de imagens.
RF03	O sistema deve classificar as emoções.
RF04	O sistema deve reconhecer expressões faciais e gestuais.

Fonte: A autora (2024).

Quadro 05: Requisitos Não Funcionais.

RNF01	O sistema deve possuir uma interface simples.
RNF02	O sistema deve ter um bom desempenho.
RNF03	O sistema deve ser documentado.

Fonte: A autora (2024).

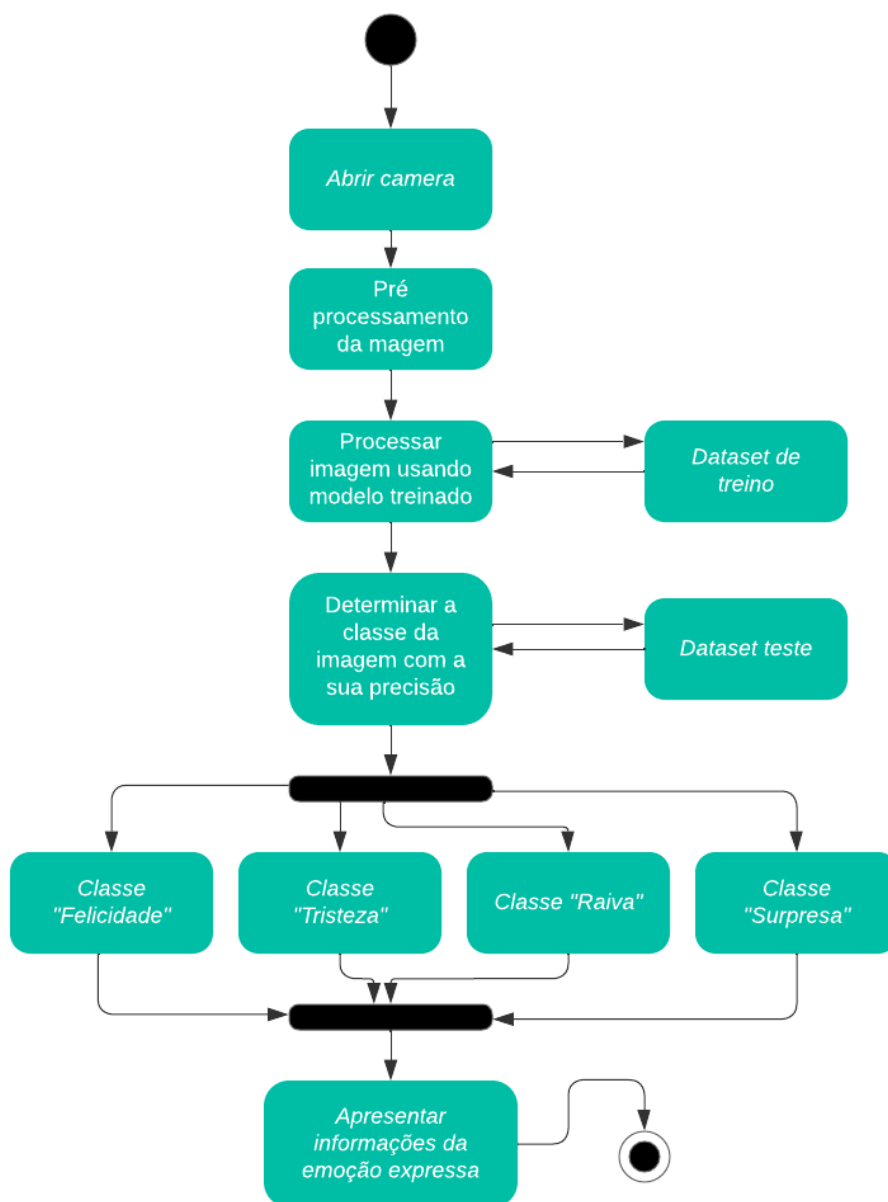
4.2 PROJETO

Neste módulo, foram desenvolvidos os diagramas mencionados acima para representar a estrutura geral de construção do sistema.

### 4.2.1 Diagrama de Atividades

Neste tópico é apresentado o diagrama de atividades na Ilustração 11, que foi desenvolvido para ilustrar o comportamento do usuário e do sistema na utilização dele.

Ilustração 11: Diagrama de atividades.



Fonte: A autora (2024).

O diagrama de atividades visualiza o fluxo do processo de identificação de emoções em LIBRAS usando visão computacional. Cada etapa é crucial para garantir a correta identificação das emoções e a eficiência do sistema. A primeira etapa é abrir a câmera, onde a câmera é ativada pelo usuário para capturar imagens, iniciando assim a captura de dados

visuais que serão processados pelo sistema. Em seguida, a imagem capturada passa pelo pré-processamento, que envolve ajustes de processamento de imagens digitais para melhorar a qualidade e adequação dos dados para o processamento subsequente.

A terceira etapa é o processamento da imagem usando um modelo treinado. Nesta etapa, a imagem pré-processada é analisada utilizando um modelo de Rede Neural Convolutacional (CNN) previamente treinado. O objetivo é aplicar o modelo para identificar padrões relevantes na imagem que correspondem a diferentes expressões emocionais. O modelo de CNN é treinado utilizando um *dataset* de treino, que contém diversas imagens com expressões faciais e gestuais associadas a emoções.

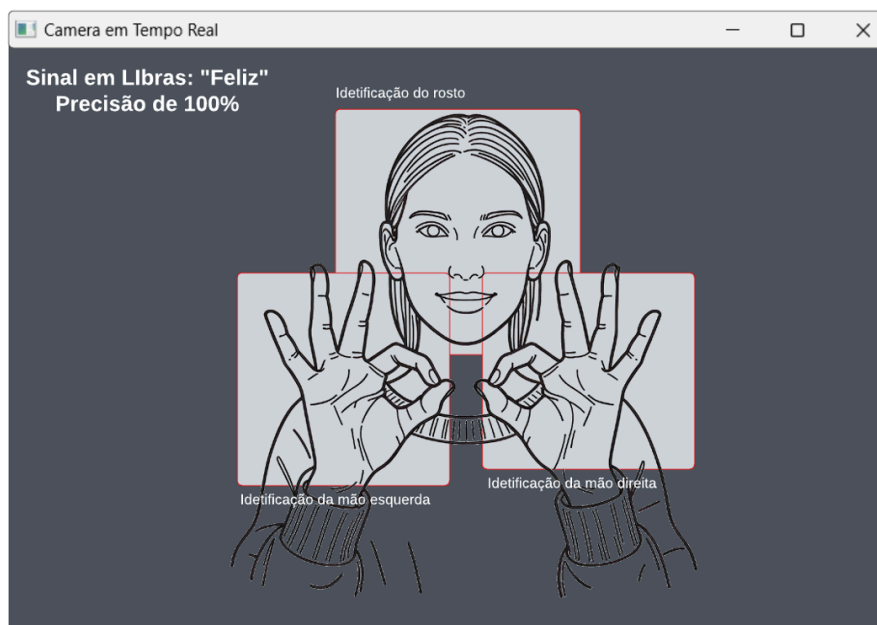
Após o processamento, o sistema determina a classe da imagem com sua precisão. Isso é feito comparando a imagem analisada com um conjunto de dados de teste para verificar a precisão da classificação. Nesta fase, o sistema identifica a emoção expressa na imagem, como felicidade, tristeza, raiva ou surpresa, e assegura que a classificação é precisa.

Finalmente, o sistema apresenta as informações da emoção expressa ao usuário. O sistema exibe as informações sobre a emoção identificada, comunicando ao usuário qual emoção foi detectada pela análise da imagem. Este fluxo linear ajuda a planejar, entender e comunicar as operações do modelo de forma clara e eficiente, garantindo que todas as etapas sejam corretamente implementadas. A implementação do trabalho utilizará a biblioteca TensorFlow para o desenvolvimento do modelo de CNN, que será treinado com um conjunto de dados de expressões faciais e gestuais capturados por uma câmera. O sistema será desenvolvido em Python e integrará tecnologias de visão computacional para realizar a identificação das emoções em tempo real.

#### **4.2.2 Wireframe**

Nesta seção, será apresentada uma tela de exemplo desenvolvida durante a fase de prototipação do modelo, fornecendo um esboço da interface do usuário. O principal objetivo é demonstrar como a câmera será inicializada e como os resultados das emoções identificadas serão exibidos na tela para o usuário.

Ilustração 12: Tela inicial de captura em tempo real:



Fonte: A autora (2024).

Na ilustração 12 é demonstrado um exemplo de como será realizada a identificação da face e do gesto em LIBRAS para o reconhecimento da emoção desejada. A ilustração proporciona ao usuário uma visão clara do que será desenvolvido, com o objetivo de classificar as quatro emoções propostas neste trabalho. Sendo assim, a Ilustração 12 mostra a expressão de “Alegria/Felicidade”, o intuito é o mesmo para as demais emoções, tristeza, raiva e surpresa.

#### 4.3 VALIDAÇÃO SISTEMA

De acordo com Analytics Vidhya (2021), Just into Data (2021), Machine Learning Plus (2021) e KDnuggets (2020), As métricas de validação são usadas para avaliar o desempenho de modelos de aprendizado de máquina e são fundamentais para entender como um modelo está se comportando com dados que não foram usados durante o treinamento. Aqui estão algumas das métricas mais comuns:

A Acurácia (*Accuracy*) mede a proporção de previsões corretas (tanto verdadeiros positivos quanto verdadeiros negativos) em relação ao total de previsões. É uma métrica simples e intuitiva, mas pode ser enganosa em conjuntos de dados desbalanceados.



A Precisão (*Precision*), calcula a proporção de verdadeiros positivos em relação ao total de positivos previstos (verdadeiros positivos + falsos positivos). É útil quando o custo de falsos positivos é alto.

A Revocação (*Recall*) ou Sensibilidade (*Sensitivity*), mede a proporção de verdadeiros positivos em relação ao total de positivos reais (verdadeiros positivos + falsos negativos). É importante em situações onde a detecção de positivos é crucial.

O F1-Score, seria a média harmônica entre precisão e revocação. Oferecendo um balanço entre as duas métricas, sendo útil quando há um trade-off entre precisão e revocação.

E a Matriz de Confusão, mostra a contagem de verdadeiros positivos, falsos positivos, verdadeiros negativos e falsos negativos. É útil para entender os tipos de erros cometidos pelo modelo e ajustar estratégias de mitigação de erros.

Essas métricas ajudam a compreender diferentes aspectos do desempenho do modelo, permitindo uma avaliação mais completa e precisa. Portanto, neste projeto, essas métricas serão usadas para validar o modelo, proporcionando uma análise quantitativa da pesquisa. Essa abordagem quantitativa garantirá que o desempenho do modelo seja rigorosamente avaliado, permitindo ajustes e melhorias contínuas conforme necessário.

5 CRONOGRAMA

Nesta seção, será detalhada a metodologia para a condução do trabalho e do projeto de pesquisa, levando em consideração o tempo disponível e o cumprimento das atividades planejadas. O quadro a seguir apresenta o cronograma utilizado para organizar as tarefas durante o segundo semestre do Trabalho de Conclusão de Curso, visando uma melhor gestão do tempo e eficiência na execução das atividades. É de extrema importância construir um cronograma para o desenvolvimento do projeto. Assim, logo abaixo está o cronograma quinzenal proposto para ser seguido durante o processo de desenvolvimento do TCC2.

Quadro 05. Cronograma do Trabalho de Conclusão de Curso (Quinzenal)

Atividades	Ago	Set	Out	Nov	Dez
Revisão das sugestões da banca					
Validação das técnicas de processamento de imagens					
Testes do modelo de IA					
Ajustes no modelo de IA					
Testes e manutenção do modelo					
Documentação da implementação na monografia					
Entrega da Monografia a banca					
Apresentação da Monografia					

Fonte: A autora (2024).

## 6 ORÇAMENTO

Este projeto foi desenvolvido inteiramente pela autora, assim como os recursos bancados para que a realização fosse possível. Certos recursos foram adquiridos antes do desenvolvimento, portanto não serão computados nos gastos. Todos os gastos referentes à construção deste projeto podem ser observados na Tabela 03. Os softwares utilizados para o desenvolvimento deste trabalho são totalmente gratuitos.

Tabela 03: Orçamento do projeto.

<b>Descrição</b>	<b>Valor Total</b>
Internet(7 meses)	R\$ 210,00
WebCam WB	R\$ 149,90
Material bibliográfico (Disponibilizado pela universidade)	R\$ 0,00
TOTAL	R\$ 359,90

## REFERÊNCIAS

Analytics Vidhya (2021): "**12 Essential Evaluation Metrics for Evaluating ML Models**". Disponível em: Analytics Vidhya. Acesso em: 15 abril 2024.

ALFF, Chico. **O que são Requisitos Funcionais e Não Funcionais?**. 2022. Disponível em: <https://analisederequisitos.com.br/requisitos-funcionais-e-naofuncionais/>. Acesso em: 30 abr. 2024.

ALVES, Elizabete Gonçalves; FRASSETTO, Silvana Soriano. **Libras e o desenvolvimento de pessoas surdas**. Aletheia, Canoas, RS, n. 46, p. 211-221, jan./abr. 2015. Disponível em: [https://repositorio.ifgoiano.edu.br/bitstream/prefix/4056/1/tcc\\_Railson%20Ferreira%20de%20Souza.pdf](https://repositorio.ifgoiano.edu.br/bitstream/prefix/4056/1/tcc_Railson%20Ferreira%20de%20Souza.pdf). Acesso em: 08 abr. 2024.

ARTERO, Almir Olivette; TOMMASELLI, Antonio Maria Garcia. **LIMIARIZAÇÃO AUTOMÁTICA DE IMAGENS DIGITAIS**. Bulletin of Geodetic Sciences, [S.l.], nov. 2018. ISSN 1982-2170. Disponível em: <https://revistas.ufpr.br/bcg/article/view/63040>. Acesso em: 20 abr. 2024.

AZENHA, Flávio Copola; FLEURY, Andre Leme. **Análise do design thinking como uma abordagem de elicitação de requisitos para o desenvolvimento de aplicações web/mobile**. Disponível em: [https://www.researchgate.net/profile/Flavio-Azenha/publication/330637430\\_ANALISE\\_DO\\_DESIGN\\_THINKING\\_COMO\\_UMA\\_ABORDAGEM\\_DE\\_ELICIACAO\\_DE\\_REQUISITOS\\_PARA\\_O\\_DESENVOLVIMENTO\\_DE\\_APLICACOES\\_WEBMOBILE/links/5c4b4697299bf12be3e308ce/ANALISE-DO-DESIGN-THINKING-COMO-UMA-ABORDAGEM-DE-ELICIACAO-DE-REQUISITOS-PARA-O-DESENVOLVIMENTO-DE-APLICACOES-WEB-MOBILE.pdf](https://www.researchgate.net/profile/Flavio-Azenha/publication/330637430_ANALISE_DO_DESIGN_THINKING_COMO_UMA_ABORDAGEM_DE_ELICIACAO_DE_REQUISITOS_PARA_O_DESENVOLVIMENTO_DE_APLICACOES_WEBMOBILE/links/5c4b4697299bf12be3e308ce/ANALISE-DO-DESIGN-THINKING-COMO-UMA-ABORDAGEM-DE-ELICIACAO-DE-REQUISITOS-PARA-O-DESENVOLVIMENTO-DE-APLICACOES-WEB-MOBILE.pdf). Acesso em: 01 maio 2024.

BARELLI, F. **Introdução à visão computacional: Uma abordagem prática com Python e OpenCV**. Casa do código, 2018.

BEZERRA, Eduarda. **Princípios de Análise e Projeto de Sistemas com UML**. 3. ed. Rio de Janeiro: Elsevier Editora Ltda, 2007. 386 p.

BERNART, Eliezer Emanuel. **Proposta de Metodologia para Rastrear Jogadores em Quadras Esportivas Utilizando Visão Computacional**. Joaçaba, 2014. 113 p. Acesso em: 23 abr. 2024.

BRASIL. **Ministério da Educação. Secretaria de Educação Especial. Saberes e Práticas da Inclusão: dificuldades de comunicação e sinalização – Surdez**. 3º Ed. Brasília: MEC/SEESP, 2005.

BUDIHARTO, W.; GROUP, S. P. **Modern Robotics with OpenCV**. Science Publishing Group, 2014. ISBN 9781940366128.

Disponível em: <http://books.google.com.br/books?id=mbFUAwAAQBAJ>. Acesso em: 21 abr. 2024.

CLAUDINO, M. M. **Desenvolvimento de um Sistema de Reconhecimento de Emoções utilizando CNNs**. 2022. Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de Campina Grande. Disponível em: <http://dspace.sti.ufcg.edu.br:8080/xmlui/bitstream/handle/riufcg/29270/MATHEUS%20MAC%20C3%8aDO%20CLAUDINO%20-%20TCC%20ARTIGO%20CI%20C3%8ANCIA%20DA%20COMPUTA%20C3%87%20C3%83O%20CEEI%202022.pdf?sequence=1&isAllowed=y>. Acesso em: 20 maio 2024.

CLORE, G. Why emotions are felt. In: EKMAN, P.; DAVIDSON, R. (Eds.). **The nature of emotions: fundamental questions**. New York: Oxford University Press, 1994. p. 103-111.

CORREIA, Ana Sofia. **Reconhecimento de Emoções através de Expressões Faciais**. 2016. Dissertação (Mestrado) - Universidade Fernando Pessoa. Disponível em: [https://bdigital.ufp.pt/bitstream/10284/4467/1/Tese\\_Ana\\_Sofia\\_Correia.pdf](https://bdigital.ufp.pt/bitstream/10284/4467/1/Tese_Ana_Sofia_Correia.pdf). Acesso em: 20 mar. 2024.

COZMAN, Fabio G.; PLONSKI, Guilherme Ary; NERI, Hugo. **Inteligência artificial: avanços e tendências**. São Paulo: Instituto de Estudos Avançados, 2021.

EKMAN, P. All **Emotions are basic**. In: EKMAN, P.; DAVIDSON, R. (Eds.). **The nature of emotions: fundamental questions**. New York: Oxford University Press, 1994. p. 15-19.

EKMAN, P. **Facial expression and emotion**. *American Psychologist*, v. 48, p. 384-392, 1993.

EKMAN, P. Facial expressions. In: DALGLEISH, T.; POWER, M. (Eds.). **Handbook of cognition and emotion**. New York: John Wiley & Sons Ltd, 1999. p. 301-320.

EKMAN, P. **The Nature of Emotions**. In: **Handbook of cognition and emotion**. New York: John Wiley & Sons Ltd, 1999. p. 15-19.

FERRAZ, C. L. M. **Dicionário de Configuração de Mãos em Libras**. [S.l.]: Editora UFRB, 2019.

FLEURY, Maria Tereza Leme; DA COSTA WERLANG, Sergio Ribeiro. **Pesquisa aplicada: conceitos e abordagens**. Anuário de Pesquisa GVPesquisa, 2016.

FONSECA, João José Saraiva da. **Metodologia da pesquisa científica**. Fortaleza: UEC, 2002. 127 p. Disponível em: <http://www.ia.ufrj.br/ppgea/conteudo/conteudo-2012-1/1SF/Sandra/apostilaMetodologia.pdf>. Acesso em: 30 abr. 2024.

GARRUTTI, É. A.; MOREIRA, T. N. D. A. **A criança surda na educação infantil bilíngue: a importância do social para a construção da linguagem**. Educação e Pesquisa, v. 48, 2022. Disponível em: <https://www.revistas.usp.br/ep/article/view/198807>.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.

GRAFÓLOGOS ONLINE. **Linguagem corporal: emoções primárias**. Disponível em: <https://grafologosonline.blogspot.com/2017/03/linguagem-corporal-emocoes-primarias.html>. Acesso em: 20 mar. 2024.

IBM. **Artificial Intelligence**. 2024. Disponível em: <https://www.ibm.com/br-pt/topics/artificial-intelligence>. Acesso em: 30 mar. 2024.

IBGE - **Instituto Brasileiro de Geografia e Estatística**. PNS 2019: País tem 17,3 milhões de pessoas com algum tipo de deficiência. Agência de Notícias, 2020. Disponível em: <https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/31445-pns-2019-pais-tem-17-3-milhoes-de-pessoas-com-algum-tipo-de-deficiencia>. Acesso em: 20 maio 2024.

IZARD, C.; FINE, S.; SCHULTZ, D.; MOSTOW, A.; ACKERMAN, B.; YOUNGSTROM, E. **Emotion knowledge as a predictor of social behavior and academic competence in children at risk**. Psychological Science, v. 12, p. 18-23, 2001.

Just into Data (2021): **"8 Popular Evaluation Metrics for Machine Learning Models"**. Disponível em: Just into Data. Acesso em: 22 março 2024.

KDnuggets (2020): **"Model Evaluation Metrics in Machine Learning"**. Disponível em: KDnuggets. Acesso em: 20 março 2024.

KELTNER, D.; EKMAN, P. **Emotion: an overview**. Encyclopedia of Psychology, v. 3, p. 162-166, 2002.

LACERDA, Cristina Broglia Feitosa de; TURETTA, Beatriz Aparecida dos Reis. **Representação simbólica por crianças surdas na educação infantil**. Horizontes, v. 36, n. 3, p. 24-35, set./dez. 2018. Disponível em: [https://repositorio.ifgoiano.edu.br/bitstream/prefix/4056/1/tcc\\_Railson%20Ferreira%20de%20Souza.pdf](https://repositorio.ifgoiano.edu.br/bitstream/prefix/4056/1/tcc_Railson%20Ferreira%20de%20Souza.pdf). Acesso em: 08 abr. 2024.

LEDERMAN, S.; KLATZKY, R.; ABRAMOWICZ, A.; SALSMAN, K.; KITADA, R.; HAMILTON, C. **Haptic recognition of static and dynamic expressions of emotion in the live faces**. Psychological Science, v. 18, p. 158-164, 2007.

MACEDO, Diego. **Modelos de Ciclo de Vida**. Disponível em: <https://www.diegomacedo.com.br/modelos-de-ciclo-de-vida/>. Acesso em: 30 abr. 2024.

Machine Learning Plus (2021): "**Top 15 Evaluation Metrics for Machine Learning with Examples**". Disponível em: Machine Learning Plus. Acesso em: 05 maio 2024.

OLIVEIRA, Cláudio; MOURA, Samuel Pedrosa. TIC's na educação: **a utilização das tecnologias da informação e comunicação na aprendizagem do aluno**. Revista Pedagogia em Ação: PUC Minas, Belo Horizonte, v. 7, n. 1, 2015.

PEDROSO, E. C.; DE PAIVA, M. A. **Ensino de Libras no contexto da escola inclusiva e na escola bilíngue: Relato de Experiência**. In: Congresso Brasileiro de Educação Especial. 2018.

PEREIRA, K. **Análise de Expressões Faciais para Identificação de Emoções**. 2022. Trabalho de Conclusão de Curso (Graduação) - Universidade Federal do Rio Grande do Sul. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/267940/001188511.pdf?sequence=1&isAllowed=y>. Acesso em: 16 abr. 2024.

ROMÁRIO. **Reconhecimento Facial de Emoções em Tempo Real**. 2020. Trabalho de Conclusão de Curso (Graduação) - Instituto Federal do Rio Grande do Norte. Disponível em: <https://memoria.ifrn.edu.br/bitstream/handle/1044/2300/tcc-romario.pdf?sequence=1&isAllowed=y>. Acesso em: 20 mar. 2024.

RUSSELL, Stuart; NORVIG, Peter. **Inteligência artificial**. 3. ed. Rio de Janeiro: Elsevier, 2013. 1324 p.

SOMMERVILLE, Ian. **Engenharia de Software**. 9. ed. São Paulo: Pearson, 2011. 548 p.

SOUZA, R. F. de. **Reconhecimento de Emoções em Libras utilizando Redes Neurais Convolucionais**. 2020. Trabalho de Conclusão de Curso (Graduação) - Instituto Federal Goiano. Disponível em: [https://repositorio.ifgoiano.edu.br/bitstream/prefix/4056/1/tcc\\_Railson%20Ferreira%20de%20Souza.pdf](https://repositorio.ifgoiano.edu.br/bitstream/prefix/4056/1/tcc_Railson%20Ferreira%20de%20Souza.pdf). Acesso em: 20 mar. 2024.

STANKIEWICZ, Alessandro. **Modelo de interação ágil: uma adaptação do modelo cascata à organização de pequenas e médias empresas**. 2017. Trabalho de Conclusão de Curso. Universidade Tecnológica Federal do Paraná.

TECH, D. **O que é TensorFlow?** Para que serve? 2022. Disponível em: <https://didatica.tech/o-que-e-tensorflow-para-que-serve/>. Acesso em: 23 abr. 2024.

TEIXEIRA, A. C. L. Dança e educação inclusiva: **processos educacionais formativos de uma artista-professora-pesquisadora com deficiência auditiva na licenciatura em dança**. 2022.

ZANELLA, Liane Carly Hermes. **Metodologia de Pesquisa**. 2. ed. Florianópolis: Universidade Federal de Santa Catarina, 2011. 134 p.



## APÊNDICES



Universidade do Oeste de Santa Catarina

### TERMO DE AUTORIZAÇÃO DE USO DE IMAGEM

Eu Marineiva Moro de Oliveira, portador(a) do CPF número 065.523.379-22, e RG número 5.075.340, depois de conhecer e entender os objetivos, procedimentos metodológicos, riscos e benefícios da pesquisa, bem como de estar ciente da necessidade do uso de minha imagem, **AUTORIZO**, através do presente termo, os pesquisadores e autores do trabalho da pesquisa a fazerem uso de minha imagem, sem quaisquer ônus financeiros a nenhuma das partes.

Esta **AUTORIZAÇÃO** é concedida mediante o compromisso dos pesquisadores acima citados em garantir-me os seguintes direitos:

- 1) Poderei ter acesso de minhas imagens utilizadas no referido trabalho de pesquisa;
- 2) Os dados coletados serão usados para fins científicos e de estudos (livros, artigos, slides e transparências), em favor dos pesquisadores supracitados, mediante o atendimento da legislação vigente do Conselho Nacional Saúde (CNS) que resguarda os aspectos éticos em pesquisas com seres humanos;
- 3) Minha identificação não poderá ser revelada em nenhuma das vias de publicações;
- 4) Os dados coletados serão guardados por 5 anos, sob a responsabilidade do(a) pesquisador(a) da pesquisa, e após esse período, serão destruídos;
- 5) Serei livre para interromper minha participação na pesquisa a qualquer momento e/ou solicitar a posse das minhas imagens.

Por ser a expressão da minha vontade assino a presente autorização, cedendo, a título gratuito, todos os direitos autorais decorrentes dos depoimentos, artigos e entrevistas por mim fornecidos, abdicando do direito de reclamar de todo e qualquer direito conexo à minha imagem e/ou som da minha voz, e qualquer outro direito decorrente dos direitos abrangidos pela Lei 9160/98 (Lei dos Direitos Autorais).

Chapecó, SC, 27 de março de 2024

Assinatura do Pesquisador Responsável

Fone:

E-mail:

Assinatura do participante

Fone: (49) 9154-7751

E-mail: marineiva.oliveira@unoesc.edu.br

*Para atingir os objetivos propostos, o trabalho justifica-se pela necessidade de superar os desafios presentes no ensino de LIBRAS, propondo uma solução inovadora que não apenas facilite o aprendizado, mas também contribua significativamente para o desenvolvimento da inteligência emocional dos estudantes. O objetivo geral é desenvolver um modelo que integre o reconhecimento facial e gestual, com o intuito de apoiar o ensino infantil em sala de aula. As etapas específicas incluem a criação de um conjunto de dados com imagens estáticas representando as emoções, o treinamento do modelo, sua implementação em tempo real, e a posterior avaliação das métricas de desempenho para garantir sua eficácia.*