



INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO PIAUÍ
CAMPUS FLORIANO
CURSO TECNOLOGIA EM ANÁLISE E DESENVOLVIMENTO DE
SISTEMAS

HIAN CLAUDIO DE SOUSA COSTA

IMPLEMENTAÇÃO DE TÉCNICAS DE VISÃO COMPUTACIONAL
APLICADAS NO ENSINO DE EXPRESSÕES EM LIBRAS: UM ESTUDO DE
CASO IFPI - CAMPUS FLORIANO

FLORIANO
2023

HLIAN CLAUDIO DE SOUSA COSTA

**IMPLEMENTAÇÃO DE TÉCNICAS DE VISÃO COMPUTACIONAL
APLICADAS NO ENSINO DE EXPRESSÕES EM LIBRAS: UM ESTUDO DE
CASO IFPI - CAMPUS FLORIANO**

Trabalho de Conclusão de Curso (artigo científico) apresentado como exigência parcial para obtenção do diploma do Curso de Tecnologia em Análise e Desenvolvimento de Sistemas do Instituto Federal de Educação, Ciência e Tecnologia do Piauí, Campus Floriano.

Orientador(a): Prof. Dr. Willamys Rangel Nunes de Sousa

FLORIANO

2023

IMPLEMENTAÇÃO DE TÉCNICAS DE VISÃO COMPUTACIONAL APLICADAS NO ENSINO DE EXPRESSÕES EM LIBRAS: UM ESTUDO DE CASO IFPI - CAMPUS FLORIANO

Hian Claudio de Sousa Costa¹
Willamys Rangel Nunes de Sousa²

RESUMO

O objetivo deste estudo era oferecer uma abordagem didática para o ensino de Libras, visando aprimorar estratégias de ensino-aprendizagem e promover a inclusão e acessibilidade de pessoas surdas na sociedade. Isso foi realizado por meio do desenvolvimento de uma ferramenta computacional que emprega técnicas de visão computacional para reconhecer expressões em Libras. Para abordar esses desafios, vários estudos têm explorado o uso de técnicas de visão computacional para auxiliar no ensino de Libras, aproveitando o processamento de imagens e o reconhecimento de padrões para analisar e interpretar gestos da língua de sinais. O processo foi dividido em quatro etapas: coleta de dados, treinamento do modelo, teste do modelo e construção da interface. Após o desenvolvimento e validação da ferramenta, constatou-se que ela foi capaz de reconhecer a maioria dos gestos inseridos durante o treinamento. Este projeto não apenas alcançou seu objetivo principal, mas também abriu caminho para futuras pesquisas e implementações na convergência entre tecnologia e educação em língua de sinais.

Palavras-chave: linguagem de sinais; libras; visão computacional; machine learning.

ABSTRACT

This study aimed to provide a didactic approach to teaching Brazilian Sign Language (Libras) with the goal of enhancing teaching-learning strategies and fostering the inclusion and accessibility of deaf individuals in society. This was accomplished through the development of a computational tool employing computer vision techniques to recognize Libras expressions. Addressing these challenges, numerous studies have explored the use of computer vision techniques to aid in Libras education, leveraging image processing and pattern recognition to analyze and interpret sign language gestures. The process was divided into four stages: data collection, model training, model testing, and interface construction. Following the development and validation of the tool, it was found capable of recognizing the majority of gestures introduced during training. This project not only achieved its primary objective but also paved the way for future research and implementations at the intersection of technology and sign language education.

Keywords: sign language; libras; computer vision; machine learning.

Data de aprovação: 08/01/2024 (data de apresentação do TCC)

¹ Graduando no curso de Análise e Desenvolvimento de Sistemas. IFPI - Campus Floriano. caflo.2021114tads02@aluno.ifpi.edu.br.

² Professor do Ensino Básico, Técnico e Tecnológico. IFPI - Campus Floriano. rangelnunes@ifpi.edu.br.

1 INTRODUÇÃO

Superar os obstáculos de comunicação enfrentados por indivíduos surdos representa, incontestavelmente, um desafio de considerável envergadura, já que a comunicação desempenha um papel fundamental na interação entre seres humanos, viabilizando a partilha de informações e a expressão de pensamentos e emoções (Gadelha *et al.*, 2022).

No entanto, as barreiras linguísticas constituem desafios adicionais que amplificam a complexidade da comunicação entre indivíduos que vivenciam a surdez, dentro desse cenário, as linguagens de sinais surgem como um meio de comunicação fundamental para as comunidades surdas.

No Brasil a comunicação entre essas pessoas é feita através da Língua Brasileira de Sinais (Libras), que é uma língua natural e visualmente gestual que se desenvolveu organicamente ao longo da história da comunidade surda brasileira, reconhecida oficialmente como a língua de sinais do Brasil desde a Lei nº 10.436/2002 (Brasil, 2002). A Libras desempenha um papel crucial na comunicação e inclusão de pessoas surdas, constituindo-se como um meio essencial para a expressão de ideias, emoções e conceitos complexos.

Apesar da importância da Libras, o ensino e a aprendizagem dessa língua enfrentam desafios notáveis. Métodos convencionais de ensino, que se baseiam em aulas presenciais e materiais impressos, não são capazes de proporcionar uma experiência de aprendizado eficaz e acessível a todos os estudantes. Ademais, a carência de recursos e ferramentas específicas para instruir em Libras pode prejudicar o aprimoramento da fluência nessa língua (Ferreira *et al.*, 2021).

Para enfrentar esses desafios, em diversos estudos têm explorado o uso de técnicas de visão computacional para auxiliar o ensino de Libras. Essas técnicas buscam aproveitar o poder do processamento de imagens e do reconhecimento de padrões para analisar e interpretar os gestos e as expressões presentes na língua de sinais. Dessa forma, é possível desenvolver estratégias inovadoras e interativas para o ensino de expressões em Libras, tornando-o mais acessível e eficaz (Garimella, 2022).

Portanto, este projeto tem como objetivo contribuir com uma abordagem didática de ensino de Libras que auxilie nas estratégias de ensino-aprendizagem e na inclusão e acessibilidade de pessoas surdas na sociedade, através do desenvolvimento de uma ferramenta computacional, aplicando técnicas de visão computacional para o reconhecimento de expressões em Libras.

2 REFERENCIAL TEÓRICO

Segundo a Organização Mundial da Saúde (OMS), a surdez é uma condição que afeta a audição e pode ser definida como uma perda total ou parcial da capacidade de ouvir. Ela é caracterizada quando há uma perda de audição igual ou superior a 40 decibéis (dB), no melhor ouvido em adultos e igual ou superior a 30 dB, no melhor ouvido em crianças. Ademais, ela pode ser congênita, ou seja, presente desde o nascimento, ou adquirida ao longo da vida, por meio de diversos fatores, como exposição a ruídos excessivos, infecções, doenças, lesões ou envelhecimento (OMS, 2018).

Mais de 5% da população mundial, cerca de 430 milhões de pessoas, necessitam de reabilitação para lidar com sua perda auditiva incapacitante (432 milhões de adultos e 34 milhões de crianças), segundo dados da OMS. Estima-se que até 2050, mais de 700 milhões de pessoas, equivalente a 1 em cada 10 pessoas, terão perda auditiva incapacitante (OMS, 2023).

No Brasil, dados do Instituto Brasileiro de Geografia e Estatística (IBGE, 2010) mostram que, aproximadamente 10% da população brasileira, cerca de 23,9 milhões de pessoas, possuem alguma deficiência auditiva. Deste total, aproximadamente 2,7 milhões têm surdez, sendo que cerca de 344 mil são crianças e adolescentes com até 14 anos de idade.

A comunicação entre pessoas surdas é geralmente realizada por meio da linguagem de sinais, que é um sistema composto por gestos, expressões faciais e movimentos corporais usados para representar palavras e frases. Através desses sinais, as pessoas surdas podem se comunicar entre si e com aqueles que conhecem a linguagem de sinais (SIL *International*, 2023). É importante ressaltar que existem mais de 300 línguas de sinais diferentes em todo o mundo, cada uma com suas próprias regras gramaticais e vocabulários distintos.

No Brasil, a comunicação entre as pessoas com surdez é realizada por meio da Libras (Língua Brasileira de Sinais), que é a língua oficial dos surdos no Brasil. Essa língua tem sua própria gramática e estrutura, sendo uma língua completa e complexa, assim como qualquer outra língua falada (Quadros, 1997).

Apesar de ser reconhecida como língua oficial dos surdos no país desde 2002, a Libras ainda não é amplamente compreendida pela população ouvinte, o que dificulta a inclusão social e a comunicação das pessoas surdas. Nesse contexto, a tecnologia tem sido uma importante aliada na inclusão social de pessoas com surdez, com diversas soluções tecnológicas disponíveis para auxiliar na comunicação (Souza *et al.*, 2019).

2.1 TECNOLOGIAS ASSISTIVAS

Segundo o Comitê de Ajudas Técnicas da Secretaria de Direitos Humanos da República, tecnologia assistiva consiste em uma área do conhecimento, de característica interdisciplinar, que engloba produtos, recursos, metodologias, estratégias, práticas e serviços que objetivam promover a funcionalidade, relacionada à atividade e participação, de pessoas com deficiência, incapacidades ou mobilidade reduzida, visando sua autonomia, independência, qualidade de vida e inclusão social (CAT, 2021).

As tecnologias assistivas podem melhorar a qualidade de vida de pessoas com limitações, incluindo aquelas com limitação auditiva. Dispositivos como alarmes e campainhas com luzes piscantes podem ajudar as pessoas surdas a perceber eventos sonoros importantes, enquanto aparelhos auditivos e implantes cocleares são tecnologias que melhoram a audição e ajudam as pessoas surdas a perceberem os sons a sua volta. Existem também aplicativos de tradução de voz para texto, como o *Google Tradutor*³, e de texto para voz, como o *NaturalReader*⁴, que podem ser usados para facilitar a comunicação com falantes ouvintes de outras línguas (Who, 2021).

³ <https://translate.google.com.br/?hl=pt-BR>

⁴ <https://www.naturalreaders.com>

2.2 MACHINE LEARNING

O aprendizado de máquina ou *machine learning* tem se destacado como uma área crucial da inteligência artificial, permitindo que os computadores possam aprender padrões a partir de dados e realizem tarefas sem serem explicitamente programados (IBM, 2023). Duas abordagens fundamentais nesse campo são o aprendizado supervisionado e não supervisionado, cada uma delas desempenhando papéis distintos em diferentes contextos.

2.2.1 Aprendizado supervisionado

No aprendizado supervisionado, os modelos são treinados usando um conjunto de dados rotulados, onde cada exemplo de entrada está associado a uma saída conhecida. As técnicas de aprendizado supervisionado são amplamente utilizadas em tarefas de classificação e regressão.

No contexto das tarefas de classificação o objetivo é categorizar os dados em classes pré-definidas. Algoritmos como *Support Vector Machines* (SVM), Redes Neurais e Árvores de Decisão são comumente empregados para tarefas como reconhecimento de imagem, detecção de spam e diagnóstico médico (Ribeiro, 2020).

Nas tarefas de regressão o aprendizado de máquina supervisionado também é aplicado, onde o objetivo é prever valores contínuos. Modelos lineares, árvores de decisão regressivas e redes neurais podem ser utilizados para previsões, como por exemplo, estimar preços de imóveis ou prever a demanda por um produto (Ribeiro, 2020).

2.2.2 Aprendizado não supervisionado

O aprendizado não supervisionado lida com conjuntos de dados não rotulados, onde o modelo é desafiado a encontrar padrões, estruturas ou representações úteis sem orientação explícita. Clustering e redução de dimensionalidade são tarefas comuns no aprendizado não supervisionado.

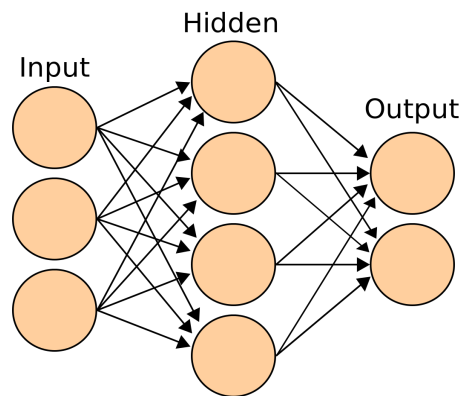
Nas tarefas de *clustering* Algoritmos como *K-Means*, *Hierarchical Clustering* e *DBSCAN* são utilizados para agrupar dados semelhantes, contribuindo para a identificação de padrões intrínsecos. Isso é aplicável em segmentação de mercado, análise de redes sociais e agrupamento de documentos (Coutinho, 2021).

A redução de dimensionalidade, como a Análise de Componentes Principais (PCA) e o *t-SNE*, permite representar dados complexos em espaços de menor dimensão, preservando as informações essenciais. Isso é valioso para visualização de dados e simplificação de modelos (Softtek, 2021).

2.3 REDES NEURAIS ARTIFICIAIS

As Redes Neurais Artificiais (RNA) são modelos matemáticos inspirados pelo funcionamento do cérebro humano, compostos por várias camadas interconectadas de unidades de processamento, conhecidas como neurônios artificiais. Essas redes são treinadas com um grande conjunto de dados, permitindo-lhes aprender padrões e características relevantes para realizar determinadas tarefas (Alves, 2020).

Figura 1 - Arquitetura de uma rede neural



Fonte: Nielsen (2019).

A Figura 1 apresenta a estrutura de uma RNA, que é composta por três partes principais: camada de entrada, camadas ocultas e camada de saída. A camada de entrada recebe os dados brutos ou informações fornecidas ao modelo. As camadas ocultas são responsáveis por realizar cálculos intermediários e aprimorar a representação dos dados. A camada de saída produz os resultados finais ou as respostas desejadas da rede neural (Pregowska; Osial, 2021).

As RNAs possuem uma ampla gama de aplicações em várias áreas. Elas são frequentemente utilizadas em reconhecimento de padrões, processamento de imagens, processamento de linguagem natural, análise de dados, previsão de séries temporais, entre outras tarefas. Por exemplo, elas podem ser usadas para classificar imagens, reconhecer objetos, traduzir idiomas, recomendar produtos com base nas preferências do usuário, entre muitas outras aplicações (Gerstner *et al.*, 2014).

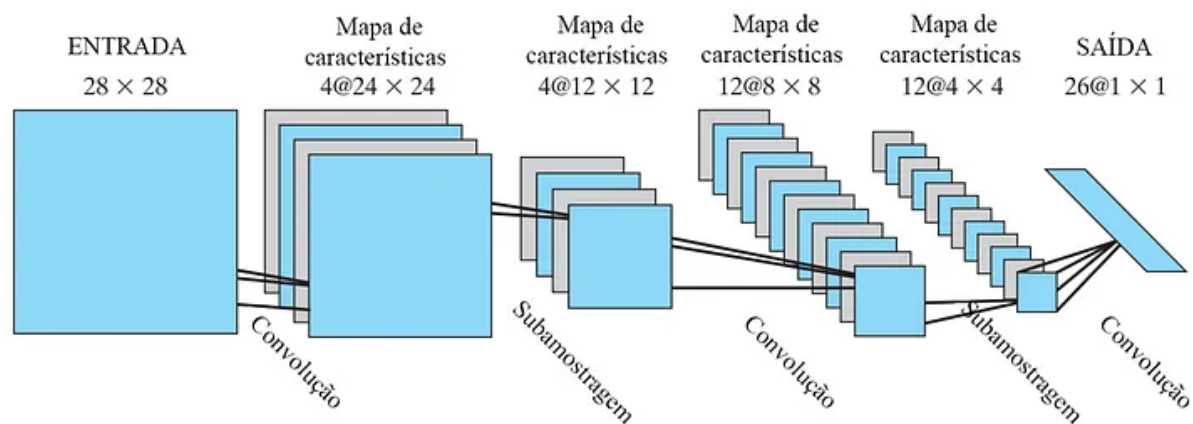
2.4 REDES NEURAIS CONVOLUCIONAIS

As Redes Neurais Convolucionais (CNNs) são um tipo especializado de RNA que se tornou fundamental no campo da visão computacional e processamento de imagens. Essa arquitetura única tem revolucionado a forma como os computadores podem “enxergar” e interpretar o mundo visualmente (Awari, 2023).

Uma vez que a arquitetura das redes neurais convolucionais é inspirada pelo sistema visual do cérebro humano. Assim como o cérebro humano possui neurônios especializados em detectar características específicas em uma imagem, as CNNs são projetadas para extrair automaticamente características relevantes de uma imagem (Ebermam; Krohling, 2018).

Em suma, a rede é formada por uma sequência de camadas, sendo que cada uma delas possui uma função específica na propagação do sinal de entrada, ilustradas pela Figura 2. Segundo Poloni (2022), tradicionalmente existem três tipos principais de camadas: as convolucionais, as de subamostragem (*pooling*) e as totalmente conectadas (*fully-connected*).

Figura 2 - Estrutura de uma cnn



Fonte: Poloni (2022).

As camadas convolucionais são responsáveis por extrair características das camadas anteriores ou da imagem de entrada, caso seja a primeira camada da rede. As de subamostragem, por sua vez, reduzem a dimensionalidade dos dados, além de ajudar a tornar a representação invariante a pequenas translações na entrada. Por fim, a camada totalmente conectada é responsável pela propagação do sinal por meio da multiplicação ponto a ponto e o uso de uma função de ativação. A última camada tem como valor de saída a probabilidade de uma imagem de entrada pertencer a uma das classes para qual a rede foi treinada, a depender da função de ativação utilizada (Poloni, 2022).

2.5 TEACHABLE MACHINE

A ferramenta *Teachable Machine* é um projeto de aprendizado de máquina desenvolvido pelo Google que permite a qualquer pessoa treinar modelos de machine learning sem a necessidade de codificação extensiva ou conhecimento técnico profundo. Essa ferramenta representa um passo significativo no sentido de democratizar o acesso à inteligência artificial e ao aprendizado de máquina, permitindo que um público mais amplo explore e crie modelos de *machine learning* de forma simples e intuitiva (Teachable machine, 2023).

A *Teachable Machine* tem diversas aplicações práticas e educacionais. Ela pode ser usada para criar sistemas de reconhecimento de gestos, classificadores de objetos,

identificação de padrões em imagens, entre outros. Além disso, essa ferramenta pode ser incorporada a projetos interativos e artísticos, ampliando seu escopo de uso. Ela desempenha um papel importante na educação, permitindo que estudantes e entusiastas de tecnologia experimentem conceitos de aprendizado de máquina de forma prática e intuitiva, o que ajuda a popularizar a compreensão e o uso dessa tecnologia.

Um projeto que exemplifica bem a versatilidade do Teachable Machine é o trabalho de Cristina (2021), que visa desenvolver um curso online que tem como objetivo ensinar o desenvolvimento de aplicativos inteligentes para classificação de imagens usando design thinking, alinhado aos currículos de referência, e utilizando ferramentas como o Google Teachable Machine⁵ e o App Inventor⁶. O objetivo é capacitar os estudantes a criar soluções úteis em suas vidas e comunidades, promovendo o conhecimento sobre Machine Learning na educação básica brasileira (Cristina, 2021).

2.6 VISÃO COMPUTACIONAL

Segundo Szeliski (2011), a visão computacional é um campo da inteligência artificial e da ciência da computação que se preocupa com a capacidade dos computadores de interpretar, analisar e compreender imagens e vídeos digitais. É uma área interdisciplinar que combina técnicas de processamento de imagem, aprendizado de máquina, reconhecimento de padrões, geometria computacional e outras áreas relacionadas.

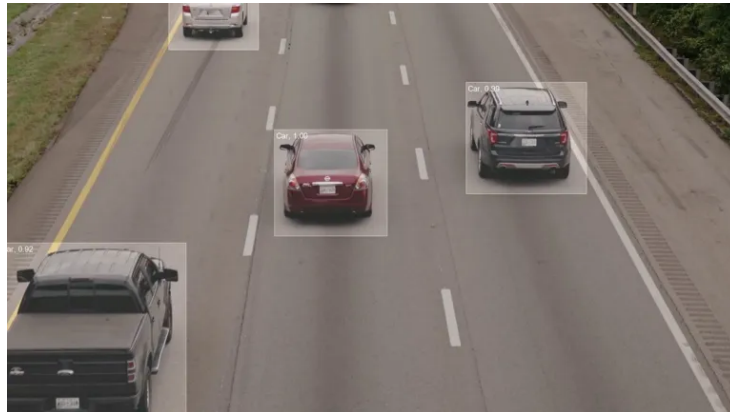
Como a visão computacional amplia o potencial de uso das tecnologias, há uma ampla gama de aplicações práticas para diversas áreas. Assim, alguns exemplos concretos de como essa tecnologia está sendo utilizada são reconhecimento facial, carros autônomos, controle de qualidade industrial, monitoramento de segurança, contagem e rastreamento de objetos, realidade aumentada, entre outras (Salles, 2023).

Uma de suas aplicações mais comuns é a classificação automatizada de veículos, uma vez que com sensores acessíveis em rápido crescimento, como câmeras, os veículos podem ser detectados, rastreados e categorizados em várias faixas simultaneamente, assim como mostra a figura 3 (DSA, 2022).

⁵ <https://teachablemachine.withgoogle.com>

⁶ <https://appinventor.mit.edu>

Figura 3 - Classificação de veículos



Fonte: DSA (2022).

2.7 VISÃO COMPUTACIONAL APLICADA NA LINGUAGEM DE SINAIS

No contexto da comunicação com pessoas surdas, o uso da visão computacional é muito útil no desenvolvimento de ferramentas de reconhecimento de sinais de Libras. É possível utilizar algoritmos de aprendizado de máquina para identificar os movimentos das mãos e expressões faciais utilizadas na língua de sinais, permitindo a criação de softwares capazes de traduzir esses sinais para o português ou outras línguas (Escobedo *et al.*, 2019).

Entre as abordagens que usam visão computacional aplicada neste contexto, destaca-se o trabalho de Guney e Erkus (2021), que descreve a metodologia utilizada para reconhecer a Língua de Sinais Turca em tempo real, por meio de redes neurais convolucionais. A metodologia consistiu em coletar um conjunto de dados de sinais de Língua de Sinais Turca, que foram pré-processados e segmentados em *frames*. Em seguida, esses *frames* foram utilizados para treinar uma rede neural convolucional.

De acordo com Guney e Erkus (2021), o modelo foi validado por meio de testes em tempo real com sinais de Língua de Sinais Turca, utilizando um dispositivo móvel, e obteve uma taxa de acurácia média de 97,5% na identificação dos sinais. Os resultados obtidos demonstram a viabilidade do uso de redes neurais convolucionais para reconhecer a Língua de Sinais Turca em tempo real, o que pode contribuir significativamente para a comunicação e inclusão de pessoas com deficiência auditiva na Turquia.

Outro exemplo do uso dessa tecnologia é o trabalho de Abdulhamied, Nasr e Abdulkader (2023), que propõe um sistema de reconhecimento em tempo real da Língua de Sinais Americana (ASL) utilizando uma rede neural de memória de curto e longo prazo (LSTM) e detecção de mão. A metodologia utilizada consistiu em capturar imagens de vídeo das mãos dos usuários que realizam os sinais da ASL, em seguida, aplicaram um algoritmo de detecção de mão, baseado em histograma de cores para extrair as regiões das mãos. Em seguida, as imagens das regiões das mãos foram pré-processadas e alimentadas em uma rede LSTM treinada com imagens de sinais da ASL.

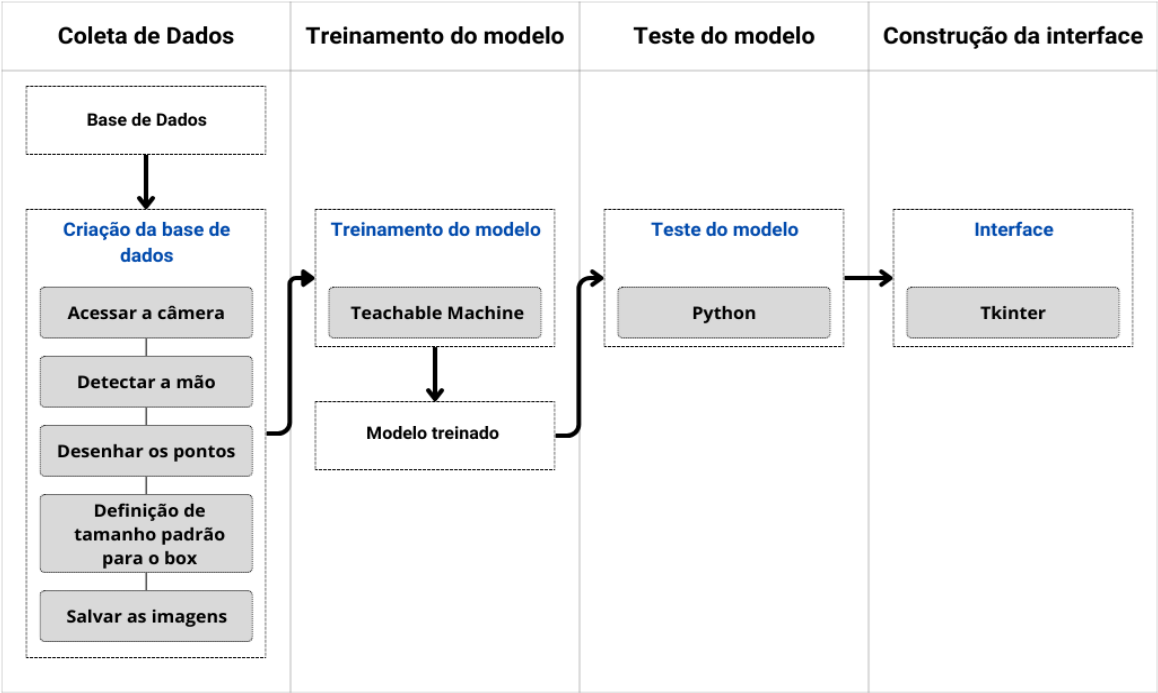
Os resultados do estudo de Abdulhamied, Nasr e Abdulkader (2023) mostraram que o sistema proposto foi capaz de reconhecer corretamente 98,5% dos sinais da ASL em tempo real, com uma taxa de quadros de 25fps (*frames* por segundo), utilizando um conjunto de dados de 29 sinais da ASL realizados por 10 usuários diferentes. Portanto, a combinação de uma rede LSTM com detecção de mão pode ser eficaz para reconhecer sinais da ASL em tempo real, o que pode ter aplicações práticas importantes para pessoas surdas ou com deficiência auditiva.

É importante ressaltar que o desenvolvimento de sistemas de reconhecimento de gestos requer um conjunto de dados abrangente e representativo, bem como algoritmos de aprendizado de máquina eficientes. Com o contínuo avanço da visão computacional e das redes neurais artificiais, espera-se que a tradução entre a linguagem de sinais e a língua oral seja cada vez mais aprimorada, contribuindo para uma comunicação inclusiva e eficiente (Ono *et al.*, 2015).

3 METODOLOGIA

Uma vez que o objetivo do projeto era implementar uma ferramenta computacional baseada em técnicas de visão computacional, para o reconhecimento de gestos de Libras, nesta seção serão abordados os procedimentos metodológicos utilizados no seu desenvolvimento. O processo foi dividido em quatro etapas, sendo elas coleta de dados, treinamento do modelo, teste do modelo e construção da interface, assim como apresentados na Figura 4.

Figura 4 - Fluxograma das etapas do desenvolvimento

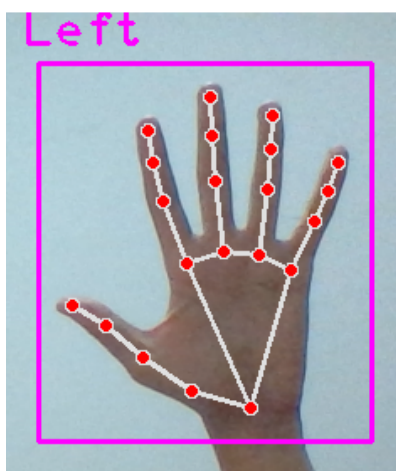


Fonte: De autoria própria.

3.1 COLETA DE DADOS

De acordo com a Figura 4, a primeira etapa da metodologia foi a coleta e construção da base de dados. Para que fosse possível realizar o treinamento de um modelo capaz de reconhecer os gestos da linguagem de sinais, é de extrema importância possuir uma base de dados sólida e padronizada. Desse modo a construção desta base de dados foi realizada utilizando bibliotecas da linguagem de programação Python, algumas das bibliotecas utilizadas foram *OpenCV*⁷ e *Cvzone*⁸.

Figura 5 - Imagem da câmera



Fonte: De autoria própria.

A Figura 5 apresenta a câmera sendo acessada através do OpenCV e sendo capaz de detectar a mão presente na imagem através do módulo *HandDetector* do *Cvzone* (CVZONE, 2023). Além disso, é possível desenhar pontos de interesse que auxiliam no treinamento do modelo.

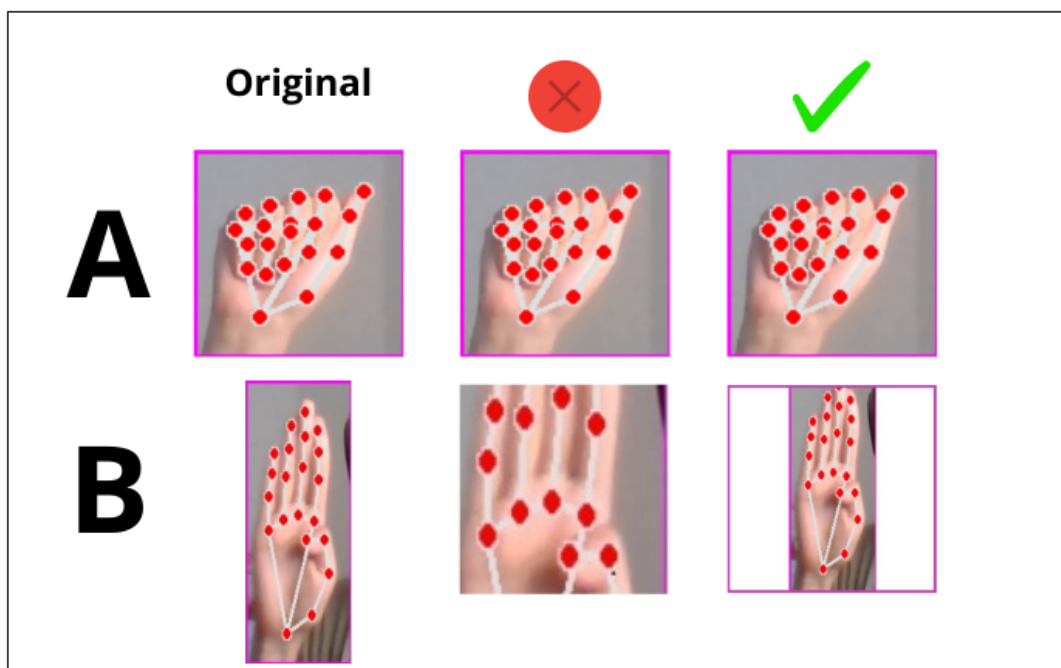
Sendo possível a detecção da mão, o próximo passo foi definir um *box* de captura padrão, uma vez que é importante que as imagens sejam do mesmo tamanho ao treinar uma rede neural. Dessa forma ao padronizar o tamanho das imagens, o pré-processamento dos dados é simplificado.

Desse modo, a Figura 6 apresenta como é realizada a padronização das imagens. Uma vez que os gestos apresentam tamanhos diferentes, a forma de solucionar esse problema é definir um tamanho padrão para o *box* de captura e fazer com que a imagem capturada se ajuste dentro deste, assim como apresentado na imagem. No total foram capturadas 22 mil imagens divididas entre as 26 letras do alfabeto.

⁷ <https://opencv.org/>

⁸ <https://www.computervision.zone/>

Figura 6 - Padronização das imagens capturadas

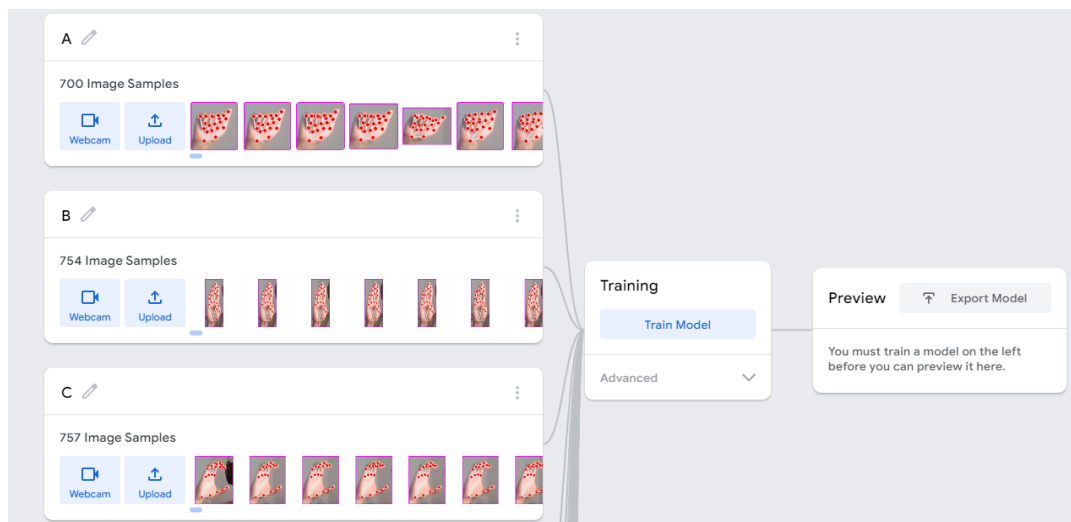


Fonte: De autoria própria.

3.2 TREINAMENTO DO MODELO

Após a coleta de dados, a próxima etapa envolve o treinamento do modelo de reconhecimento de gestos de Libras. As imagens coletadas na etapa anterior são usadas para treinar uma rede neural. Durante o treinamento, o modelo aprende a identificar gestos de Libras com base nas imagens de entrada. A Figura 7 exibe a tela de treinamento do *Teachable Machine*.

Figura 7 - Tela de treinamento do teachable machine

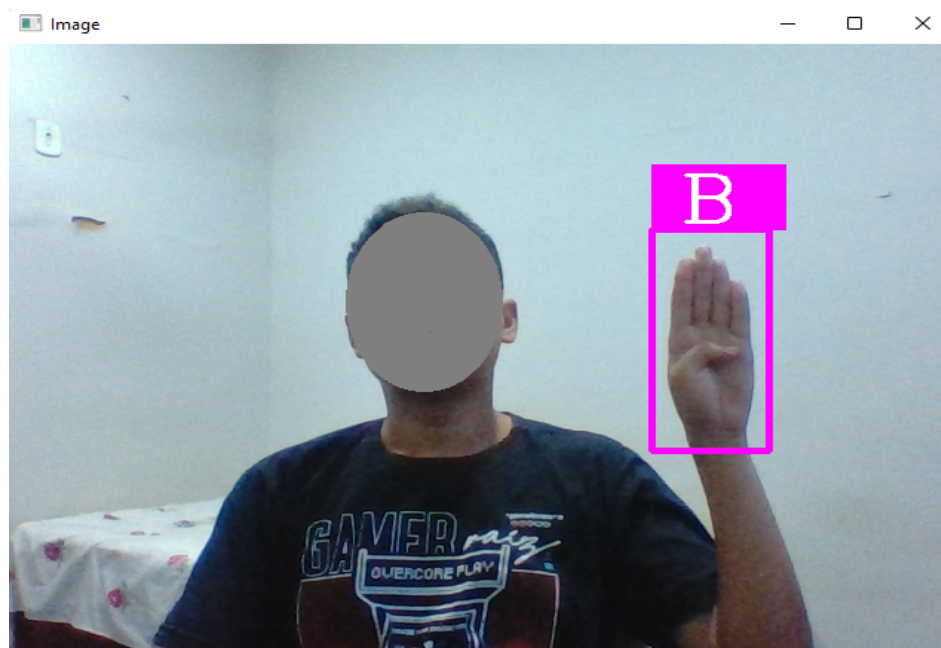


Fonte: De autoria própria.

3.3 TESTE DO MODELO

Após o treinamento, o modelo foi testado para avaliar sua capacidade de reconhecer gestos em Libras. Isso envolve a apresentação de novos gestos a partir da câmera ao modelo, utilizando o arquivo do modelo treinado exportado pelo *Teachable Machine*, e a avaliação de sua capacidade de classificá-los corretamente. O teste desse arquivo é realizado por meio da utilização da biblioteca Classifier do *Cvzone*, contribuindo assim para a determinação da eficiência do modelo.

Figura 8 - Teste do modelo



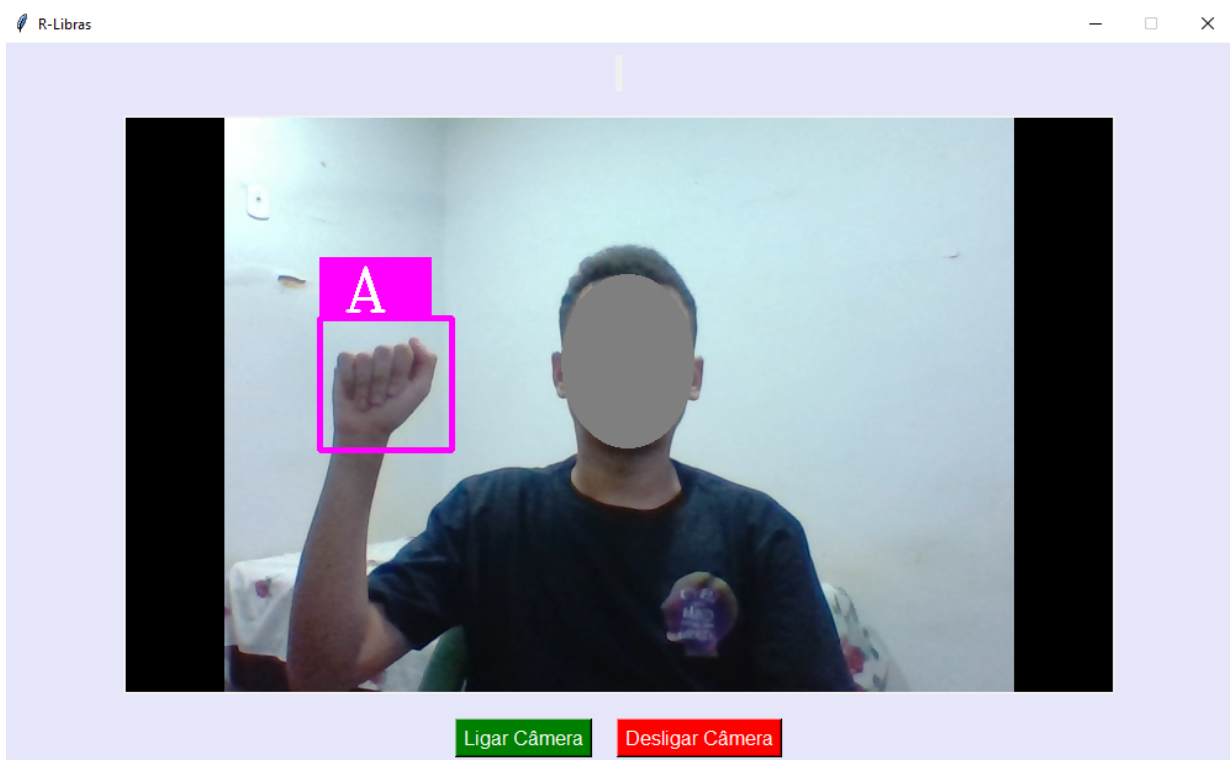
Fonte: De autoria própria.

A Figura 8 exibe o teste realizado e demonstra o modelo sendo capaz de reconhecer gestos de Libras.

3.4 CONSTRUÇÃO DA INTERFACE

Com o modelo se mostrando capaz de reconhecer os gestos, a próxima etapa foi construir uma interface a fim de deixar a ferramenta mais intuitiva, assim como mostrado na Figura 9. A biblioteca *Tkinter*⁹ do Python foi utilizada no desenvolvimento da interface.

Figura 9 - Interface da ferramenta



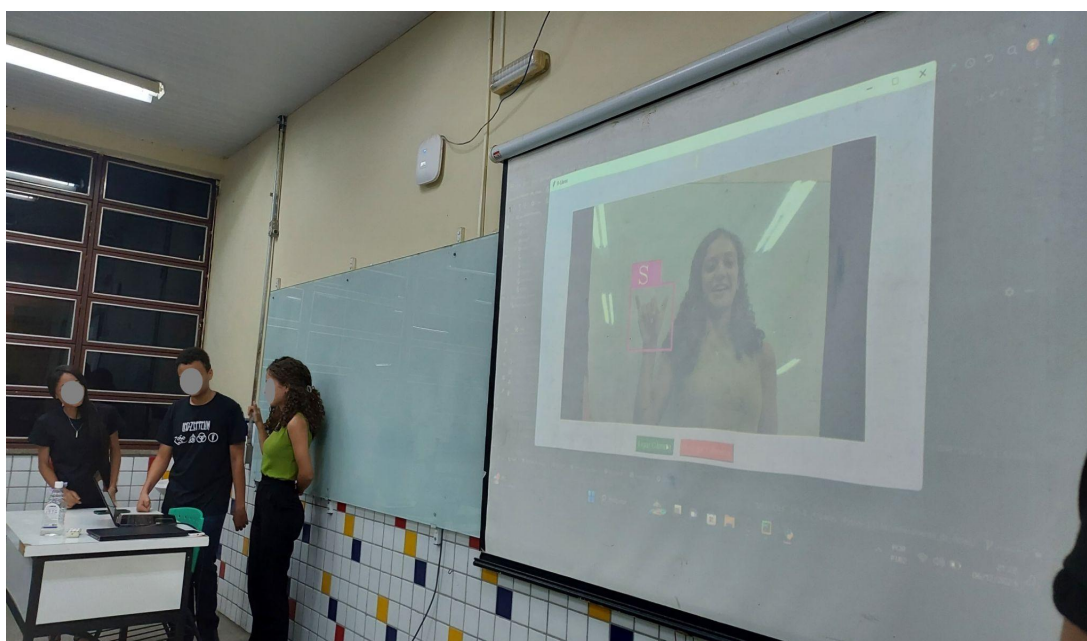
Fonte: De autoria própria.

⁹ <https://docs.python.org/3/library/tkinter.html>

4 VALIDAÇÃO DOS RESULTADOS

Após a conclusão do desenvolvimento da ferramenta, a etapa subsequente consistiu na avaliação da interação dos alunos matriculados na disciplina de Libras com essa tecnologia. Esta avaliação representou um marco crucial, pois permitiu não apenas verificar a eficácia do reconhecimento de gestos, mas também compreender como os estudantes se comportaram com essa nova abordagem tecnológica no aprendizado da língua de sinais. A Figura 10 ilustra essa interação entre os alunos e a ferramenta.

Figura 10 - Interação dos alunos com a ferramenta



Fonte: De autoria própria

Após analisar a interação dos alunos com a ferramenta de reconhecimento de gestos de Libras, a fim de consolidar ainda mais sua eficácia e relevância, foi aplicado um questionário para os alunos. O questionário teve um total de 16 respostas, estão contidas na Tabela 1 as perguntas do questionário.

Tabela 1 - Perguntas do questionário

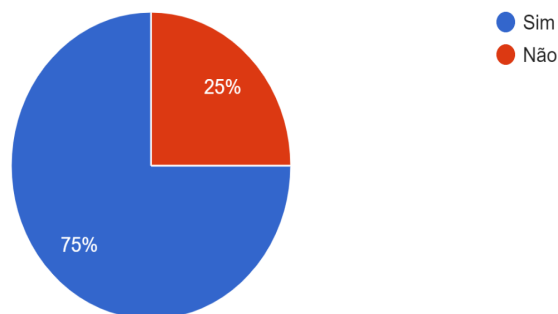
1 - Você possui experiência anterior na língua Brasileira de Sinais (Libras)?
2 - Você já empreendeu esforços para adquirir conhecimento em Língua Brasileira de Sinais (Libras)?
3 - Você enfrentou algum desafio ou considera o processo de aprendizagem da Língua Brasileira de Sinais complexo?
4 - Você considera que esta ferramenta de tradução possui potencial para ser proveitosa no contexto do processo educacional?

Fonte: De autoria própria

A *pergunta 1* visava avaliar se os participantes possuíam experiência anterior na língua. Dos participantes entrevistados, 75% afirmaram possuir experiência prévia em Libras, enquanto 25% declararam não ter experiência anterior na língua, assim como mostrado na Figura 11.

Figura 11

Você possui experiência anterior na língua Brasileira de Sinais (Libras)?
16 respostas



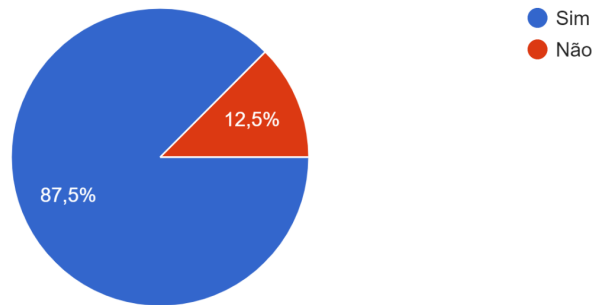
Fonte - De autoria própria

Por sua vez, a *pergunta 2* buscou identificar se os participantes já haviam empreendido esforços para adquirir conhecimento em LIBRAS. Os resultados indicaram que 87,5% dos entrevistados responderam afirmativamente, enquanto 12,5% afirmaram não ter realizado esforços nesse sentido, a figura 12 ilustra esse resultado.

Figura 12

Você já empreendeu esforços para adquirir conhecimento em Língua Brasileira de Sinais (Libras)?

16 respostas



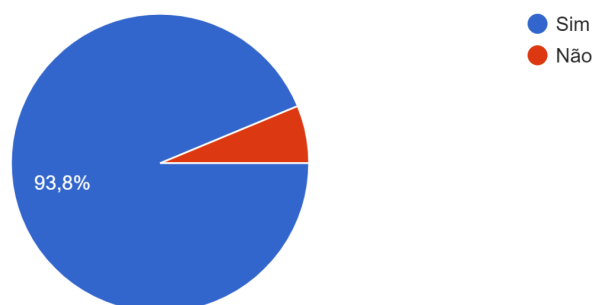
Fonte - De autoria própria

Para compreender a percepção dos participantes sobre a complexidade do processo de aprendizagem em LIBRAS, foi questionado (*pergunta 3*) se eles enfrentaram desafios ou consideram o processo complexo. A maioria expressiva, correspondendo a 93,8%, respondeu afirmativamente, indicando ter enfrentado desafios ou percebido a complexidade no aprendizado, enquanto apenas 6,3% dos participantes responderam negativamente a essa questão, assim como mostra a figura 13.

Figura 13

Você enfrentou algum desafio ou considera o processo de aprendizagem da Língua Brasileira de Sinais complexo?

16 respostas

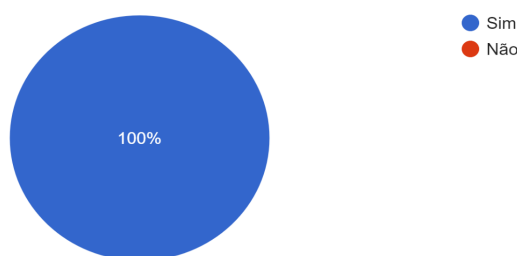


Fonte - De autoria própria

Por fim, os participantes foram questionados (*pergunta 4*) sobre a viabilidade da ferramenta de tradução em LIBRAS no contexto educacional. Todos os entrevistados (100%) consideraram que essa ferramenta possui potencial para ser proveitosa no processo educacional, assim como mostra a Figura 14.

Figura 14

Você considera que esta ferramenta de tradução possui potencial para ser proveitosa no contexto do processo educacional?
16 respostas



Fonte - De autoria própria

5 CONCLUSÃO

Esse projeto tinha como objetivo contribuir com uma abordagem didática de ensino de Libras que auxiliasse nas estratégias de ensino-aprendizagem e na inclusão e acessibilidade de pessoas surdas na sociedade, através do desenvolvimento de uma ferramenta computacional, aplicando técnicas de visão computacional para o reconhecimento de expressões em Libras.

Após o desenvolvimento e validação da ferramenta constatou-se que ela se provou-se capaz de reconhecer a maioria dos gestos inseridos durante o treinamento, contudo provou-se que ocorre uma queda na precisão quando exposta a alguns gestos que não são estáticos. O motivo disso ocorrer se deve ao fato de que a base de dados é constituída por imagens estáticas. Esta limitação revela uma área a ser aprimorada para otimizar a precisão da ferramenta.

Apesar desse obstáculo, a ferramenta representa um avanço significativo na integração de tecnologias inovadoras no ensino da língua de sinais. Contudo, para sua evolução e aprimoramento, é crucial superar suas atuais limitações. Futuras pesquisas podem se concentrar na expansão da base de dados para incluir gestos dinâmicos, possibilitando o aperfeiçoamento do reconhecimento e aumentando a precisão da ferramenta.

Também é importante ressaltar a necessidade de explorar métodos para ampliar a funcionalidade da ferramenta. Estudos adicionais podem focar na implementação de técnicas de aprendizado de máquina mais avançadas para o reconhecimento de gestos dinâmicos e na criação de uma interface mais amigável e interativa para facilitar o aprendizado de Libras.

Assim, este projeto não apenas alcançou seu objetivo central de criar uma ferramenta para o reconhecimento de expressões em Libras, mas também lançou bases sólidas para pesquisas futuras e implementações que visam aprimorar a integração entre tecnologia e educação em língua de sinais.

REFERÊNCIAS

ABDULHAMIED, R.M., Nasr, M.M., & Abdulkader, S.N. (2023). Real-time recognition of American sign language using long-short term memory neural network and hand detection. Indonesian Journal of Electrical Engineering and Computer Science, 30(1), 545-556. ISSN: 2502-4752. DOI: 10.11591/ijeecs.v30.i1.pp545-556.

ALVES, P. Inteligência Artificial e Redes Neurais. IPEA - Centro de Pesquisa em Ciência, Tecnologia e Sociedade, 2020. Disponível em: <https://www.ipea.gov.br/cts/pt/central-de-conteudo/artigos/artigos/106-inteligencia-artificial-e-redes->. Acesso em: 13 jun. 2023.

AWARI. Redes Neurais Convolucionais: Como elas Funcionam. 8 de junho de 2023. Disponível em: https://awari.com.br/redes-neurais-convolucionais/?utm_source=blog&utm_campaign=projeto+blog&utm_medium=Redes%20Neurais%20Convolucionais:%20Como%20elas%20Funcionam&utm_content=design. Acesso em: 20 jun. 2023.

BRASIL. Lei nº 10.436, de 24 de abril de 2002. Dispõe sobre a Língua Brasileira de Sinais - Libras e dá outras providências. Brasília, DF, 2002. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/2002/110436.htm. Acesso em: 10 abr. 2023.

CAPES 2020-2024. DOI: 10-18264/REP. Publicado em 29 de novembro de 2022. Disponível em: <https://educacaopublica.cecierj.edu.br/artigos/22/44/aspectos-linguisticos-da-libras-e-desafios-dos-surdos-na-aquisicao-da-lingua-escrita>. Acesso em: 12 de junho de 2023.

CAT. Tecnologia assistiva: Ferramentas para inclusão de PCD's. Disponível em: <https://www.icom-libras.com.br/2021/10/04/tecnologia-assistiva/>. Acesso em: 09 jun. 2023.

COUTINHO, T. Desvende os benefícios da análise de Clustering para potencializar resultados. Abril de 2021. Disponível em: <https://www.voitto.com.br/blog/artigo/clustering>. Acesso em: 21 de dez. 2023.

CRISTINA, B. Desenvolvimento de um Curso Ensinando a Criação de Apps Inteligentes para a Classificação de Imagens com Machine Learning e Design Thinking. 2021. Trabalho de Conclusão do Curso de Graduação em Sistemas de Informação, Departamento de Informática e Estatística, Centro Tecnológico, Universidade Federal de Santa Catarina, Florianópolis.

CVZONE. Cvzone: Computer Vision Zone. Disponível em: <https://www.computervision.zone/>. Acesso em: 20 nov. 2023.

DBSCAN. DBSCAN Clustering Algorithm in Machine Learning. Disponível em: <https://www.kdnuggets.com/2020/04/dbscan-clustering-algorithm-machine-learning.html>. Acesso em: 20 nov. 2023.

DSA. Data Science Academy. 25 Aplicações Comerciais de Visão Computacional. 2022.

Disponível em:

https://blog.dsacademy.com.br/25-aplicacoes-comerciais-de-visao_computacional/. Acesso em: 21 dez. 2023.

EBERMAM, E; KROHLING, R. Uma Introdução Compreensiva às Redes Neurais Convolucionais: Um Estudo de Caso para Reconhecimento de Caracteres Alfabéticos. Revista de Sistemas de Informação da FSMA, v. 22, pp. 49-59, 2018.

ESCOBEDO, E.; RAMIREZ, L.; CAMARA, G. Dynamic Sign Language Recognition Based on Convolutional Neural Networks and Texture Maps. In: 2019 32nd SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Rio de Janeiro, Brazil, 2019. Proceedings... Rio de Janeiro: IEEE, 2019. p. 265-272. DOI: 10.1109/SIBGRAPI.2019.00043. Disponível em: <https://ieeexplore.ieee.org/document/8919953>. Acesso em: 14 abr. 2023.

FERREIRA, F *et al.* A importância da libras no processo de ensino aprendizagem do aluno surdo. VII CONEDU - Conedu em Casa. Campina Grande: Realize Editora, 2021. Disponível em: <https://editorarealize.com.br/artigo/visualizar/81472>. Acesso em: 12 jun. 2023.

GADELHA, F; PEREIRA, S; MARTINS, V. Aspectos linguísticos da Libras e desafios dos surdos na aquisição da língua escrita. Revista Educação Pública. ISSN 1984-6290. Qualis B1 - avaliação

GERSTNER, W., Kistler, W. M., Naud, R. e Paninski, L. 2014. “Neuronal dynamics: from single neurons to networks and models of cognition.” Cambridge: Cambridge University Press.

GUNEY, S., & Erkus, M. (2021). Real-time recognition of Turkish sign language using convolutional neural networks. Neural Computing and Applications, 34(14), 4069-4079. doi: 10.1007/s00521-021-06664-6.

HIERARCHICAL CLUSTERING. What is Hierarchical Clustering?. Disponível em:

<https://www.displayr.com/what-is-hierarchical-clustering/#:~:text=Hierarchical%20clustering%2C%20also%20known%20as,broadly%20similar%20to%20each%20other>. Acesso em: 23 nov. 2023.

IBM. O que é machine learning?. Disponível em:

<https://www.ibm.com/br-pt/topics/machine-learning>. Acesso em: 19 jun. 2023.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. Censo Demográfico 2010: Características gerais dos moradores e dos domicílios. Rio de Janeiro, 2010. Disponível em: <https://www.ibge.gov.br/estatisticas/sociais/populacao/9662-censo-demografico-2010.html?&t=o-que-e>. Acesso em: 10 abr. 2023.

K-MEANS. K-means: o que é, como funciona, aplicações e exemplo em Python. Disponível em:

<https://medium.com/programadores-ajudando-programadores/k-means-o-que-é-como-funciona-aplicações-e-exemplo-em-python-6021df6e2572>. Acesso em: 20 nov. 2023.

NATURALREADERS. NaturalReader. Disponível em: <https://www.naturalreaders.com>. Acesso em: 08 mai. 2023.

NIELSEN, M. Neural Networks and Deep Learning. Dezembro de 2019. Disponível em: <http://neuralnetworksanddeeplearning.com/chap1.html>. Acesso em: 20 jun. 2023.

ONO, G.; VIEIRA, B. Reconhecimento de Sinais da Libras por Visão Computacional. Mecatrone, 2015. DOI: 10.11606/issn.2526-8260.mecatrone.2015.116344. Disponível em: <https://www.revistas.usp.br/mecatrone/article/view/116344>. Acesso em: 13 jun. 2023.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. Deafness and hearing loss. Geneva, 2023. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss>. Acesso em: 9 abr. 2023.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. Surdez: definição e causas. Genebra, 2018. Disponível em: https://www.who.int/health-topics/hearing-loss#tab=tab_1. Acesso em: 9 abr. 2023.

OPENCV. OpenCV: Open Source Computer Vision Library. Disponível em: <https://opencv.org/>. Acesso em: 20 nov. 2023.

POLONI, K. Redes neurais convolucionais, 2022. Disponível em: <https://medium.com/itau-data/redes-neurais-convolucionais-2206a089c715>. Acesso em 10 abr. 2023.

PREGOWSKA, A. e Osial, M. (2021). “What is an artificial neural network and why do we need it?” Front. Young Minds. 9:560631. DOI: 10.3389/frym.2021.560631. Agnieszka Pregowska.

QUADROS, Ronice M. de. Educação de Surdos: a aquisição da linguagem. Porto Alegre: Artmed, 1997.

RIBEIRO, A. Uma visão geral sobre machine learning – Classificação. Agosto de 2020. Disponível em: [https://statplace.com.br/blog/uma-visao-geral-sobre-machine-learning/#:~:text=Classificação%20no%20Machine%20Learning,\(três%20ou%20mais%20classes\)](https://statplace.com.br/blog/uma-visao-geral-sobre-machine-learning/#:~:text=Classificação%20no%20Machine%20Learning,(três%20ou%20mais%20classes)). Acesso em: 21 Dez. 2023.

RIBEIRO, A. Uma visão geral sobre machine learning – Regressão e Clustering. Agosto de 2020. Disponível em:
<https://statplace.com.br/blog/uma-visao-geral-sobre-machine-learning-regressao-clustering/>. Acesso em: 21 Dez. 2023.

RODRIGUES, A. V-LIBRASIL: uma base de dados com sinais na língua brasileira de sinais (Libras). 2021. Dissertação (Mestrado em Ciência da Computação) - Universidade Federal de Pernambuco, Recife, 2021.

SALLES, A. O que é Visão Computacional e para que serve? 2023. Disponível em:
<https://santodigital.com.br/o-que-e-visao-computacional-e-para-que-serve/>. Acesso em: 21 de Dezembro de 2023.

SANTOS, J. A. *et al.* O que é visão computacional e para que serve? Santo Digital, 2021. Disponível em:
<https://santodigital.com.br/o-que-e-visao-computacional-e-para-que-serve/#:~:text=Visão%20computacional%20é%20uma%20área,relevantes%20para%20uma%20aplicação%20futura.> Acesso em: 09 abr. 2023.

SIL International. Sign Languages [online]. 2023. Disponível em: <https://www.sil.org/sign-languages> Acesso em: 07 maio 2023.

SOFTTEK. Redução da dimensionalidade na Machine Learning. 2021. Disponível em:
<https://blog.softtek.com/pt/reducao-da-dimensionalidade-na-machine-learning>. Acesso em: 21 de dezembro de 2023.

SOUZA, J; SANTOS, M. Libras na educação: limites e possibilidades. Educação Pública, Rio de Janeiro, Jun. 2019. ISSN 1984-6290. Qualis B1 - avaliação CAPES 2020-2024. DOI: 10-18264/REP. Disponível em:
<https://educacaopublica.cecierj.edu.br/artigos/19/13/libras-na-educacao-limites-e-possibilidades>. Acesso em: 09 jun. 2023.

SZELISKI, R. (2011). Computer vision: algorithms and applications. Springer Science & Business Media.

TEACHABLEMACHINE. Teachable Machine, 2023. Disponível em:
<https://teachablemachine.withgoogle.com>. Acesso em: 30 de Out. 2023.

TKINTER. tkinter — Python interface to Tcl/Tk. Disponível em:
<https://docs.python.org/3/library/tkinter.html>. Acesso em: 20 nov. 2023.

WORLD HEALTH ORGANIZATION. Assistive Technology. Disponível em:
https://www.who.int/health-topics/assistive-technology#tab=tab_1. Acesso em: 08 mai. 2023.