

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA  
DE SÃO PAULO  
CÂMPUS SALTO

ANGELO GABRIEL SOARES DE QUEIROZ  
BRUNA LUIZA ALVES REAL

**Reconhecedor de gestos de Libras**

Salto  
2022

ANGELO GABRIEL SOARES DE QUEIROZ

BRUNA LUIZA ALVES REAL

**Reconhecedor de gestos de Libras**

Trabalho de conclusão de curso apresentado ao Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – Câmpus Salto, como parte dos requisitos para obtenção do título de Bacharel em Ciência da Computação.

Orientador:  
Prof. Me Fabiano Gonzaga Fumes

[Ficha catalográfica]

A ficha deve ser solicitada à Biblioteca pelo e-mail [biblioteca.salto@ifsp.edu.br](mailto:biblioteca.salto@ifsp.edu.br), informando os seguintes dados:

- Nome completo
- Orientador(a)
- Título do trabalho e subtítulo (se houver)
- Curso
- Número de páginas
- Há ilustração (tabela, gráfico, foto, quadro, etc.) sim ou não?
- Palavras-chave.

[Folha de aprovação]

A folha deve ser solicitada ao orientador.

Gostaríamos de dedicar este trabalho a nossos pais, Rogério Graziano Soares de Queiroz e Angela Cristina Soares de Queiroz, por parte do Angelo, e Sueli Cristina da Silva, por parte da Bruna. Por serem ótimas pessoas na vida a quem nos espelhamos ao longo dos anos e por todo o apoio nos momentos de dificuldades. Nos faltam palavras para descrever o quão grato e orgulhoso somos por termos vocês. Essa é somente uma das conquistas de nossas vidas, mas que sem o apoio, incentivo e confiança, não seria possível de ser realizado.

## **AGRADECIMENTOS**

Primeiramente agradecemos a todos os docentes e funcionários do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo – Campus Salto, que estavam dispostos a fornecer o suporte necessário para uma vivência agradável e rica em conhecimento.

Em especial, deixamos nossos agradecimentos ao professor Me Fabiano Gonzaga Fumes, por toda sua orientação e dedicação que foi fundamental para o desenvolvimento deste trabalho.

Por último, mas não menos importante, deixamos nossa gratidão a familiares e amigos, que sempre estiveram ao nosso lado, e pelo apoio demonstrado ao longo de todo o período dedicado a este trabalho.

"Nós não devemos deixar que as incapacidades das pessoas nos impossibilitem de reconhecer as suas habilidades."  
(Hallahan e Kauffman, 1994)

## **RESUMO**

O presente trabalho situa-se nas áreas de computação e acessibilidade, visto que o objetivo principal é a utilização da tecnologia, especificamente a visão computacional, para reduzir as barreiras de comunicação entre deficientes auditivos falantes de Libras e ouvintes não falantes da Língua. O trabalho dividiu-se em duas frentes: O estudo teórico das tecnologias e da Língua de Sinais Brasileira, e o desenvolvimento de um sistema capaz de reconhecer letras do alfabeto manual e traduzir para a Língua Portuguesa.

Palavras-chave: Visão Computacional; Libras; Tradução Libras para Língua Portuguesa.



## **ABSTRACT**

The presented work is situated in the computing and accessibility areas, and the main objective is the use of technology, specifically computer vision, to reduce communication barriers between hearing impaired speakers of Libras and listeners who do not speak the language. The work was divided into two fronts: The theoretical study of technologies and Brazilian Sign Language, and the development of a system capable of recognizing letters of the manual alphabet and translating them into Portuguese.

Keywords: Computer Vision; Libras; Translation from Libras to Portuguese Language.

## LISTA DE ILUSTRAÇÕES

Figura 1 - Sinal usado para dizer banco (financeiro)	15
Figura 2 - Sinal usado para dizer meia (peça de vestuário)	15
Figura 3 - Sinal usado para dizer telefone	16
Figura 4 - Sinal usado para dizer casa	16
Figura 5 - Comparação entre os gestos usados para falar “verão” na Língua de Sinais Americana (ASL) e na LIBRAS, respectivamente	16
Figura 6 - Alfabeto Manual ou Datilologia	17
Figura 7 - Configuração da mão	18
Figura 8 - Ponto de articulação do sinal de camelo	19
Figura 9 - Movimento espiral para realizar sinal de barco	19
Figura 10 - Funcionamento de uma CNN	21
Figura 11 - Exemplo de uma operação de convolução	21
Figura 12 - Exemplo de funcionamento de um <i>pooling</i>	22
Figura 13 - Tabela comparativa de trabalhos relacionados usando kinect	24
Figura 14 - Representação da letra A	27
Figura 15 - Representação da letra B	27
Figura 16 - Representação do sistema após a formação de uma palavra com todas as telas sendo exibidas	28
Figura 17 - QR Code do vídeo de demonstração do sistema	28

## **LISTA DE TABELAS**

Tabela 1 - Resultado das buscas de trabalhos relacionados

23

## LISTA DE ABREVIATURAS E SIGLAS

ASL	<i>American Sign Language</i>
CM	Configuração da mão
CNN	<i>Convolutional Neural Network</i>
HMM	<i>Hidden Markov Model</i>
IBGE	Instituto Brasileiro de Geografia e Estatística
INES	Instituto Nacional de Educação de Surdos
LIBRAS	Língua Brasileira de Sinais
M	Movimento
MLP	<i>Multilayer Perceptron</i>
PA	Ponto de articulação
SGD	<i>Stochastic Gradient Descent</i>

# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO .....</b>	<b>13</b>
1.1.	JUSTIFICATIVA .....	13
1.2.	OBJETIVOS .....	14
1.2.1.	<i>Objetivos Específicos .....</i>	<i>14</i>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA .....</b>	<b>15</b>
2.1.	LÍNGUAS DE SINAIS E A LIBRAS .....	15
2.2.	VISÃO COMPUTACIONAL .....	19
2.3.	REDES NEURAIS CONVOLUCIONAIS .....	20
2.4.	TRABALHOS RELACIONADOS.....	22
<b>3</b>	<b>ARQUITETURA DA SOLUÇÃO .....</b>	<b>25</b>
3.1.	MODELO DA SOLUÇÃO .....	25
3.1.1.	<i>Treinamento.....</i>	<i>25</i>
3.1.2.	<i>Captura de imagem e feedback ao usuário.....</i>	<i>26</i>
3.1.3.	<i>Processamento de imagem .....</i>	<i>26</i>
<b>4</b>	<b>RESULTADOS .....</b>	<b>27</b>
<b>5</b>	<b>CONCLUSÃO .....</b>	<b>29</b>

# 1 INTRODUÇÃO

Um fator indispensável para o desenvolvimento do ser humano é a comunicação, e é por meio da linguagem que o pensamento se estrutura e se torna possível expor ideias e expressar sentimentos (DOMINGOS, 2010).

O principal sentido humano que auxilia na comunicação é a audição, e quando essa está de alguma forma prejudicada é possível que a fala também esteja comprometida (BARROS, 2010). De acordo com dados do IBGE (2010), no Brasil aproximadamente 10 milhões de pessoas possuem algum nível de deficiência auditiva.

No decorrer dos anos, os deficientes auditivos obtiveram várias conquistas relacionadas a legislação e políticas de inclusão, dentre elas o reconhecimento da LIBRAS como meio legal de comunicação e expressão (Brasil, 2002).

Usar a língua de sinais é falar com as mãos e ouvir com olhos, é interagir com uma cultura que percebe o mundo através dos sinais, que experimenta deslumbramento quando entende e se faz entender, que convida, que troca, que aceita, que oferece. (DOMINGOS, 2010)

Ainda que o reconhecimento da Libras configure uma conquista para a comunidade de pessoas surdas, a falta de conhecimento e domínio pela maioria das pessoas que não possuem deficiência auditiva gera um isolamento desses indivíduos dentro da sociedade, uma vez que, tarefas corriqueiras como assistir uma aula ou abrir uma conta no banco, tornam-se atividades árduas quando não há um intérprete no local, e em grande parte não há (GALVÃO, 2018).

## 1.1. JUSTIFICATIVA

Embora seja uma língua independente, a Libras não se configura como uma mera gestualização da língua portuguesa, e sua aquisição cognitiva e o uso trazem consigo dificuldades inerentes a uma língua não-fonética (VIEIRA, 2016).

Ainda nas palavras de Vieira (2016) algumas das principais dificuldades enfrentadas no uso da língua de sinais são:

- Resistência no hábito dos pais ouvintes de uma criança surda em utilizar a língua oral, o que dificulta o aprendizado da língua de sinais.

- Seu aprendizado não ocorre de maneira espontânea ou natural, embora se organize neurologicamente da mesma maneira que uma língua oral.
- Choques culturais com a tradição em determinados países, que resistem à língua de sinais e defendem a comunicação exclusivamente oral.
- Dificuldade de formação bilíngue de pais e professores.
- Necessidade de adaptações curriculares, mudança de hábito de todo o corpo docente e cooperação do corpo discente para um bom aproveitamento do aluno surdo no contexto escolar e acadêmico.
- Aprimoramento técnico-pedagógico à disposição da comunicação para a inclusão do aluno surdo.
- Aumento da comunicação escrita em sala de aula, para facilitar a complementação oral, em caso de leitura labial.
- Necessidade de ensino particularizado de acordo com as limitações do deficiente auditivo.

À vista disso, <sup>então</sup> percebe-se a carência de acessibilidade linguística. <sup>Sendo</sup> assim, este trabalho visa ampliar a comunicação e interação entre surdos e ouvintes por meio da tecnologia. <sup>DE RECONHECIMENTO DE IMAGEM.</sup>

## 1.2. OBJETIVOS

O intuito deste trabalho é utilizar a tecnologia para facilitar a comunicação de indivíduos com deficiência que se comunicam por meio da Libras com indivíduos que não utilizam a língua, reconhecendo sinais, símbolos e gestos e transcrevendo seu significado na Língua Portuguesa.

### 1.2.1. Objetivos Específicos

- Identificar os gestos referentes às letras estáticas do alfabeto
- Traduzir os gestos identificados para a respectiva letra na Língua Portuguesa
- Construir palavras e/ou frases com as letras registradas

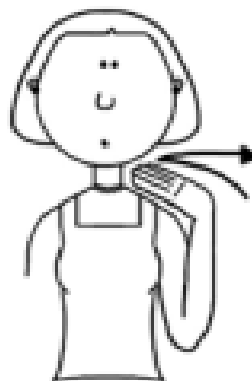
## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1. LÍNGUAS DE SINAIS E A LIBRAS

A falta de informação sobre as línguas de sinais ocasiona diversos pensamentos equivocados, dentre esses, a concepção de que a língua de sinais é universal. Esse equívoco se dá por conta de mitos que permeiam o senso comum, como por exemplo, achar que a língua de sinais é apenas mímica, ou que a mesma língua é utilizada em todos os lugares (FORSTER, 2004).

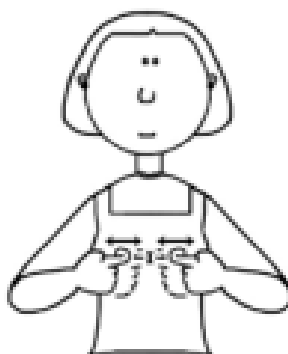
As figuras a seguir ajudam a comprovar, que apesar de existirem gestos imitativos, isso não é um padrão para todos os sinais, pois há sinais que dificilmente são compreendidos por não falantes da língua, e que assim como nas línguas orais, existem diferentes gestos que expressam o mesmo significado (FORSTER, 2004).

Figura 1 - Sinal usado para dizer banco (financeiro)



Fonte: FORSTER (2004)

Figura 2 - Sinal usado para dizer meia (peça de vestuário)



Fonte: FORSTER (2004)

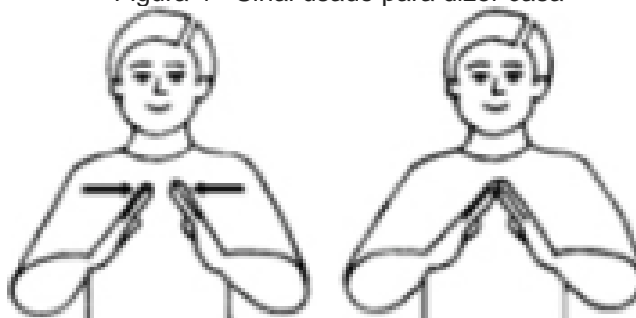


Figura 3 - Sinal usado para dizer telefone



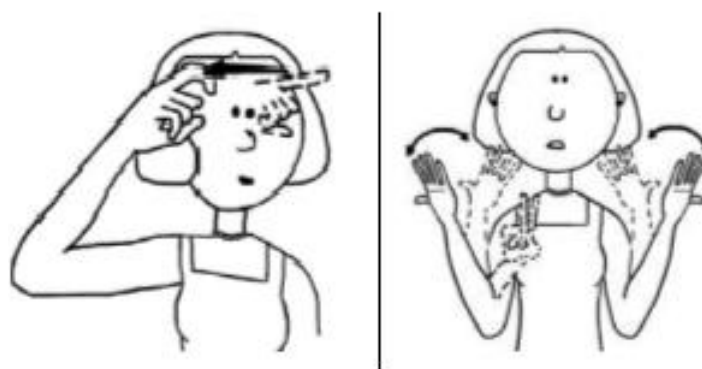
Fonte: Capovilla e Raphael (2001)

Figura 4 - Sinal usado para dizer casa



Fonte: Capovilla e Raphael (2001)

Figura 5 - Comparação entre os gestos usados para falar “verão” na Língua de Sinais Americana (ASL) e na LIBRAS, respectivamente



Fonte: Adaptada pelos autores

Esses mitos além de prejudicar o desenvolvimento linguístico de alguns surdos, ainda tiram o prestígio da língua falada por estes indivíduos.

As línguas de sinais possuem aspectos linguísticos equivalentes às línguas orais em todos os níveis, ou seja, nos níveis fonológicos (quirológico), morfológico, sintático, semântico e pragmático, porém isso ocorre na modalidade visual espacial.

Estudos focados em investigações linguísticas evidenciam que as mesmas restrições das línguas faladas se aplicam às línguas de sinais. (QUADROS; CRUZ, 2011).

No Brasil, a LIBRAS foi reconhecida a partir da Lei nº 10.436/2002, que de acordo com o parágrafo único do art.º 1:

Entende-se como Língua Brasileira de Sinais - Libras a forma de comunicação e expressão, em que o sistema linguístico de natureza visual-motora, com estrutura gramatical própria, constitui um sistema linguístico de transmissão de ideias e fatos, oriundos de comunidades de pessoas surdas do Brasil. (BRASIL, 2002)

Esse reconhecimento se tornou possível graças a aspectos estruturais da língua, que a faz ter incorporações lexicais, sintáticas e morfológicas. Segundo Almeida e Almeida (2012) os três principais parâmetros da Libras são: Configuração da mão (CM), Ponto de articulação (PA) e Movimento (M).

A CM refere-se à configuração que a mão terá ao executar um sinal. Na figura 7 observa-se que as configurações podem assumir características do alfabeto manual (figura 6), e formas diferentes da datilologia.

Figura 6 - Alfabeto Manual ou Datilologia



Fonte: Barros (2010)

Figura 7 - Configuração da mão



Fonte: Almeida e Almeida (2012)

O PA é a posição espacial onde a configuração de mão se realiza, podendo tocar alguma parte do corpo.

Figura 8 - Ponto de articulação do sinal de camelo



Fonte: Almeida e Almeida (2012)

E o M que indica que os sinais podem ter um movimento ou não. De acordo com Almeida e Almeida (2012, p. 12) “O movimento é definido como um parâmetro complexo que pode envolver uma vasta rede de formas e direções, desde os movimentos internos da mão, os movimentos do pulso e os movimentos direcionais no espaço.”

Figura 9 - Movimento espiral para realizar sinal de barco



Fonte: Almeida e Almeida (2012)

## 2.2. VISÃO COMPUTACIONAL

Um dos sentidos mais marcantes do ser humano é, sem dúvidas, a sua capacidade de processamento da informação, que possibilita a interpretação de inúmeros aspectos do mundo.

A visão computacional é o ramo da Inteligência Artificial que investiga maneiras de dar às máquinas a capacidade de interpretar visualmente as informações, em outras palavras, tenta reproduzir a visão humana (KOT ENGENHARIA, 2021).

As aplicações que utilizam visão computacional, geralmente, pretendem resolver problemas particulares de forma específica. Dessa forma, os sistemas de visão computacional envolvem reconhecimento de objetos em imagens e transformações destes objetos em informações úteis (MILANO; HONORATO, 2014).

Ainda nas palavras de Milano e Honorato (2014) existem algumas funcionalidades comuns na maioria dos sistemas de visão computacional, sendo elas:

- Aquisição de Imagem: Processo de obtenção de uma imagem ou de um conjunto de imagens.
- Pré-processamento: Procedimento realizado antes de obter informações de uma imagem, por exemplo destaque de contornos, bordas, a fim de facilitar a identificação de um objeto. *Justificado*
- Extração de características: Extração de aspectos da imagem, como textura, bordas, formatos, movimento.
- Detecção e segmentação: Recurso necessário para destacar regiões relevantes da imagem, fracionando-a para processamento posterior.
- Processamento de alto nível: Validação dos dados, estimativa de parâmetros e classificação dos objetos obtidos em diferentes categorias.

De fato, a visão computacional não se trata de uma atividade trivial e de acordo com Brandizzi (2020), os pesquisadores de aplicações de Redes Neurais Convolucionais viabilizaram sua utilização.

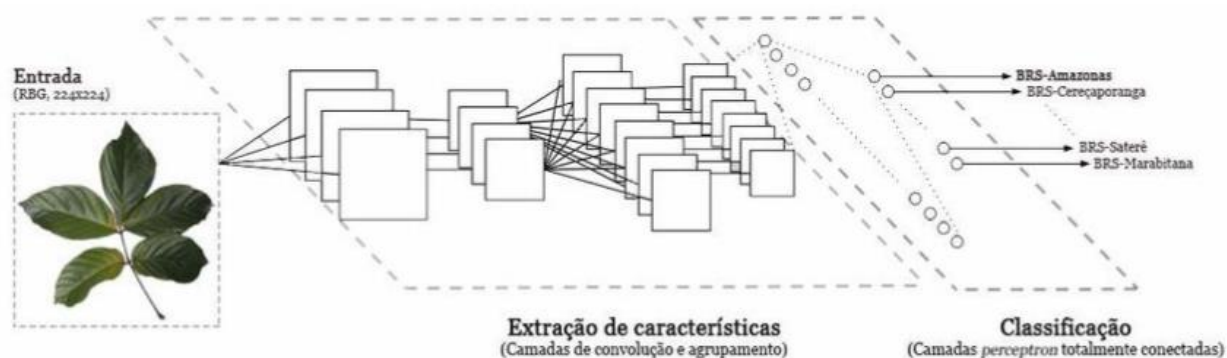
Nas palavras de Nishio et al. (2018), a *Convolutional Neural Network* (CNN) é uma classe de métodos de aprendizado profundo, e se tornou dominante em várias tarefas de visão computacional além de atrair interesse em vários domínios.

### 2.3. REDES NEURAIAS CONVOLUCIONAIS

Uma CNN é dividida em dois blocos, no qual o primeiro é responsável pela extração de características e é onde se encontra as camadas de convolução e

camadas de *pooling*, já o segundo bloco é encarregado da classificação e possui as camadas totalmente conectadas que se comportam igual a uma *Multilayer Perceptron* (MLP) (LACERDA, 2019).

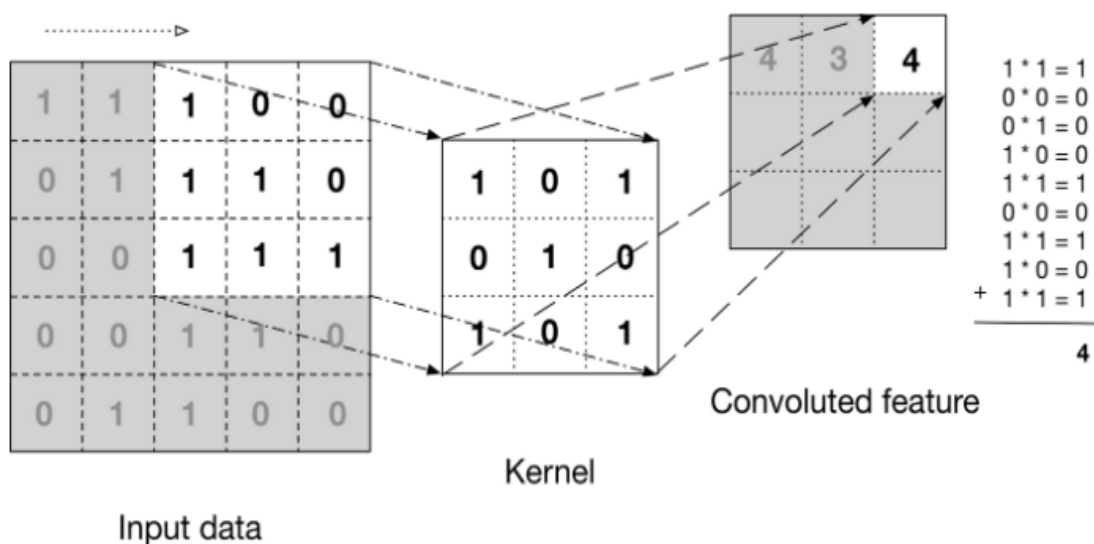
Figura 10 - Funcionamento de uma CNN



Fonte: Lacerda (2019)

No bloco de extração a camada de convolução é constituída por diversos neurônios, os quais são responsáveis por aplicar filtros, que é a matriz de convolução designada para a detecção de características importantes na imagem (LACERDA, 2019).

Figura 11 - Exemplo de uma operação de convolução



Fonte: Lacerda (2019)

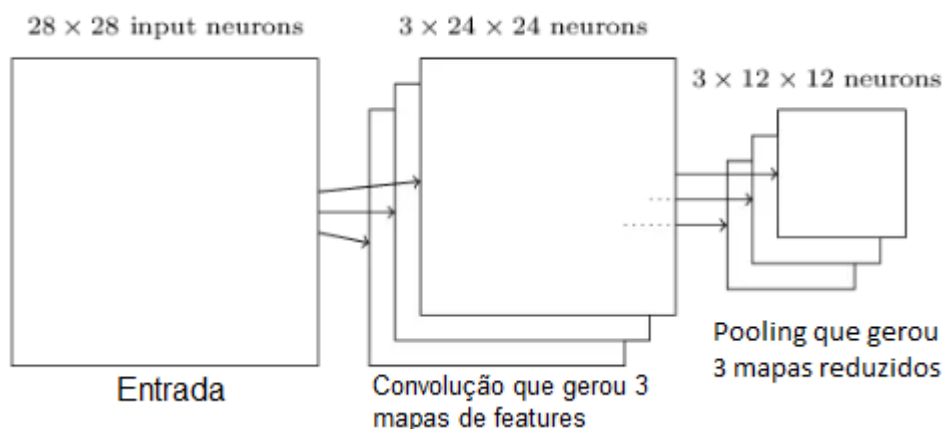
A camada de *pooling* é incumbida de reduzir a quantidade de parâmetros para a próxima camada, seja ela de convolução ou as camadas totalmente conectadas.



Isto significa que a camada de *pooling* simplifica a informação e reduz a dimensionalidade da convolução (LACERDA, 2019).

Esta camada é importante pois a redução gerada auxilia na agilidade do treinamento (LACERDA, 2019).

Figura 12 - Exemplo de funcionamento de um *pooling*



Fonte: Elaborada pelos autores

Por fim, no bloco de classificação, as camadas totalmente conectadas têm os neurônios com conexões completas com todos os neurônios ativados da camada anterior (LACERDA, 2019).

A entrada desta camada é a saída da camada anterior que contém informações das características da imagem. A resposta final será N neurônios, no qual N representa a quantidade de classes para a classificação (LACERDA, 2019).

## 2.4. TRABALHOS RELACIONADOS

A busca por trabalhos relacionados ocorreu na plataforma *Google Scholar* com as seguintes palavras chaves: reconhecedor de Libras, identificador de Libras, tradutor de Libras e classificador de Libras com visão computacional. Os critérios de inclusão foram:

- Tradução de Libras para o Português, excluindo o inverso.
- Identificação de sinais estáticos

- Utilização de visão computacional e processamento de imagens

Os resultados obtidos estão apresentados na tabela abaixo.

Tabela 1 - Resultado das buscas de trabalhos relacionados

BUSCA REALIZADA	QUANTIDADE ENCONTRADA
Reconhecedor de Libras	3
Identificador de Libras	1
Tradutor de Libras	6
Classificador de Libras com visão computacional	2
<b>TOTAL:</b>	12

Fonte: Elaborado pelos autores

Analisando os trabalhos selecionados, observou-se que este tema ainda é recente e não há um consenso sobre qual a melhor forma de resolvê-lo, visto que encontramos trabalhos utilizando Kinect (figura 13 apresenta uma tabela de comparação entre trabalhos relacionados que utilizaram o kinect), usando rede convolucional, rede convolucional 3D, rede neural recorrente, entre outros.

Outro ponto constatado foi a falta de um amplo *dataset* para treinamento e testes, isso se dá por conta da quantidade de gestos, as variações dessas articulações e a possibilidade de existir movimento. Sendo assim nos deparamos com projetos que além de reconhecer gestos ainda se propunham a criar um *dataset*, e outros que apenas faziam uso de algum *dataset* pré-existente.

Foram encontrados trabalhos semelhantes que tinham por objetivo identificar gestos em Libras com movimento, e no geral, estes faziam uso de Modelo oculto de Markov (HMMs). No entanto, estes trabalhos não fazem parte dos selecionados, já que foi delimita<sup>do</sup> para o presente trabalho apenas gestos estáticos. Contudo faz-se necessário ressaltar a existência destes trabalhos para pesquisas futuras.



Figura 13 - Tabela comparativa de trabalhos relacionados usando kinect

Trabalho	Objetivo do Trabalho	Dispositivos e Técnicas Utilizadas	Questões a Destacar
<b>Recognizing hand gestures with Microsoft Kinect (TANG, 2011)</b>	Estudar a viabilidade do reconhecimento de gestos em uma escala pequena como, por exemplo, agarrar e apontar.	São utilizados os recursos da câmera RGB e o sensor de profundidade do Kinect.	Como resultado foi obtida uma taxa de sucesso de 90% para reconhecimento de gestos de agarrar e soltar.
<b>Construção de ambiente para desenvolvimento de jogos educacionais baseados em interface de gestos (CEFET-RJ, 2013)</b>	Promover o estudo do Kinect e suas aplicações em contextos educacionais.	Kinect SDK e GreenFoot ( <i>framework</i> didático para desenvolvimento java)	Um dos trabalhos desenvolvidos consiste em uma espécie de quadro virtual, o qual possibilita ao usuário escrever em uma tela imaginária com sua mão.
<b>PRODEAF Aplicativo de Libras (PRODEAF, 2013)</b>	Traduzir diversas palavras para gestos em Libras, utilizando computadores e smartphones.	-	O aplicativo já é utilizado por mais de 130.000 pessoas de acordo com informações do site.
<b>Kinect SDK Dynamic Time Warping (DTW) Gesture Recognition (KINECTDTW, 2011)</b>	Reconhecer gestos rápidos, confiáveis e altamente personalizáveis, utilizando rastreamento do esqueleto em 2D.	Kinect SDK, Algoritmo de vizinho mais próximo baseado em vetor para monitorar e classificar seus gestos.	O elemento <i>Dynamic Time Warping</i> significa que pode reconhecer gestos executados em velocidades diferentes, sem importar a velocidade em que foram gravados, porém, o fato de utilizar apenas 2D limita seu uso para gestos simples.
<b>O uso do Kinect na acessibilidade de pessoas com deficiência visual (GASSEN, 2013)</b>	Promover autonomia de deficientes visuais por meio de um software desenvolvido com o Kinect, o qual calcula e informa a distância entre o usuário e os objetos.	Utiliza Kinect e sua SDK, juntamente com um algoritmo de processamento de imagem denominado "Crescimento de Regiões"	O <i>software</i> é capaz de detectar a distância de objetos com pequenas variações, portanto acredita-se que seja possível aplicar a tecnologia num ambiente real.

Fonte: Brandt (2015)

### 3 ARQUITETURA DA SOLUÇÃO

A partir do estudo realizado, a conclusão obtida foi de que a melhor opção para o treinamento de uma rede neural convolucional seria com a utilização da biblioteca Keras, que viabiliza uma rápida e fácil configuração de todos os parâmetros para a quantidade necessária de camadas a serem utilizadas.

Além disso, selecionamos a biblioteca OpenCV para o processamento de imagens e vídeo (entrada e saída), e manipulação da Interface Gráfica do Usuário (GUI), pois possui diversos algoritmos de visão computacional como: filtros de imagem, calibração de câmera, reconhecimento de objetos, entre outros.

Seguindo as definições dos objetivos específicos, apenas as letras estáticas do alfabeto manual serão reconhecidas pela aplicação, sendo assim as letras H, J, K, X, Z não serão identificadas pelo sistema.

Durante as pesquisas, foi encontrado o trabalho do Lacerda (2019) disponível no GitHub com um *dataset*, amplo e de fácil adaptação para este trabalho, e uma rede treinada. Após uma breve análise do treinamento da rede já treinada, foi decidido aproveitar algumas configurações como base para o desenvolvimento, devido a otimização empregada.

#### 3.1. MODELO DA SOLUÇÃO

##### 3.1.1. Treinamento

Para realizar o treinamento, foi empregado o Método do Gradiente Estocástico (SGD) para a otimização, após isso fez se necessário o uso de 11 camadas, sendo elas 9 compostas por camadas de convolução 2D (Conv2D), *Leaky Rectified Linear Unit* (LeakyReLU) e *Max pooling operation for 2D spatial data* (MaxPooling2D) modificando alguns parâmetros de algumas delas, e 2 camadas do tipo *Dense*, com ativações relu e softmax respectivamente.

O treinamento, fez uso de um *dataset* de aproximadamente 30 mil imagens, no qual para chegar a um resultado equilibrado entre qualidade e velocidade de processamento, fez se necessário que o tamanho das imagens fossem de 64x64. O

pixels .

tempo de execução ficou em torno de 45 minutos a 1 hora e a rede obteve a uma precisão de 98%.

### **3.1.2. Captura de imagem e feedback ao usuário**

Tratando-se da parte de interação do usuário e feedback de qual letra está sendo reconhecida, a biblioteca OpenCV forneceu as ferramentas ideais, sendo possível inicializar câmera para a captura de imagem, demarcar as áreas de entrada de imagem para validar a letra que está sendo expressa e de retorno do processamento.

Além disso, foi possível abrir novas janelas, no qual uma delas apresenta a frase que está sendo transcrita e outra tela opcional exibe o alfabeto manual com as letras em libras.

### **3.1.3. Processamento de imagem**

O fluxo após a captura da ação do usuário, até a resposta em tela do que foi reconhecido, trata-se do processamento de imagem e comparação com o modelo gerado.

Antes de efetivamente comparar o que foi capturado da tela de entrada, o sistema precisa efetuar alguns tratamentos nessa imagem, sendo eles: O redimensionamento dessa entrada e a conversão para tons de cinza. Somente após esse tratamento de imagem, que se torna possível a comparação com o modelo gerado em busca da tradução do gesto praticado.

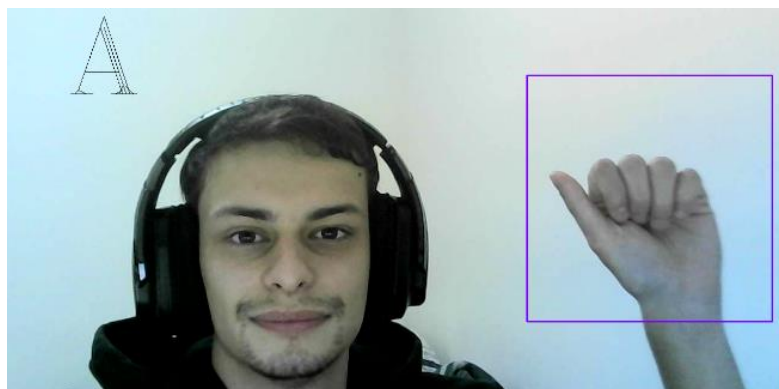
## 4 RESULTADOS

O software desenvolvido é composto por 3 telas, a principal em que apresenta o local de posicionamento da mão para captura dos gestos (uma área quadrada em roxo), a letra identificada e a palavra sendo formada, a segunda tela exibe apenas a transcrição das letras reconhecidas, e a terceira, que pode ser opcional, divulga a imagem do alfabeto manual.

O programa é capaz de reconhecer as letras estáticas apresentadas pelo usuário, além de armazená-las com o intuito de formar palavras e frases, contudo, para algumas letras em que a configuração da mão é mais agrupada, pode haver uma oscilação no reconhecimento (ver figura 17). Independentemente disso, o sistema apresenta uma taxa 98% de acerto no reconhecimento do gesto praticado pelo usuário.

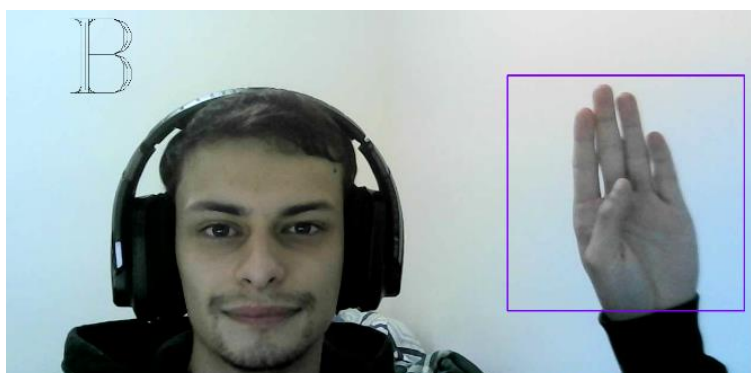
Nas figuras abaixo, é exibido apenas a tela principal com o reconhecimento de alguma letra.

Figura 14 - Representação da letra A



Fonte: Elaborada pelos autores

Figura 15 - Representação da letra B



Fonte: Elaborada pelos autores

A seguir, é apresentado uma imagem do sistema com todas as telas em exibição, após o reconhecimento de algumas letras e a formação da palavra obtida. E na figura 17, é possível escanear o QR Code exposto e assistir a uma breve demonstração do sistema em execução.

Figura 16 - Representação do sistema após a formação de uma palavra com todas as telas sendo exibidas



Fonte: Elaborada pelos autores

Figura 17 – QR Code do vídeo de demonstração do sistema.



Fonte: Elaborada pelos autores

## 5 CONCLUSÃO

Este trabalho foi desenvolvido para facilitar a comunicação entre falantes e não falantes de Libras através da tecnologia e seus recursos.

Em atenção aos objetivos propostos, concluiu-se que todos dentro do escopo foram alcançados, visto que o sistema elaborado é capaz de reconhecer e transcrever para a Língua Portuguesa os gestos referentes a letras estáticas do alfabeto manual, e ainda formar palavras e frases.

Apesar dos resultados já obtidos com o presente trabalho, ainda há espaço para melhorias futuras, que podem ser focadas por exemplo no reconhecimento das letras dinâmicas e de outros gestos com movimento, ou até mesmo a criação de um *dataset* que possa auxiliar em futuras aplicações.

Diante do que foi estudado e desenvolvido nesse trabalho acredita-se que o uso da visão computacional e o processamento de imagens podem ser grandes aliados na tentativa de superar as barreiras de comunicação existentes atualmente.

## REFERÊNCIAS

ALMEIDA, Magno Pinheiro de; ALMEIDA, Miguel Eugênio. **História de LIBRAS: característica e sua estrutura**. Revista Philologus, Ano, v. 18, 2012.

BARROS, Jussara. **Libras – Forma de comunicação**. Disponível em: <http://www.escolakids.com/libras-forma-de-comunicacao.htm>. Acesso em 15 maio 2022

BRANDIZZI, Loreane Evelyn Nazareth. **Visão computacional: O que é? Como funciona?**. [S. l.], 7 abr. 2020. Disponível em: <https://www.serpro.gov.br/menu/noticias/noticias-2020/o-que-eh-visao-computacional>. Acesso em: 27 maio 2022.

BRANDT, André Henrique. **LiRANN : sistema de reconhecimento de LIBRAS baseado em redes neurais artificiais com kinect**. 2015. Trabalho de Conclusão de Curso (Bacharelado em Ciência da Computação) - Universidade de Santa Cruz do Sul, [S. l.], 2015. Disponível em: <http://hdl.handle.net/11624/1031>. Acesso em: 26 maio 2022.

BRASIL. **Lei nº 10.436, de 24 de abril de 2002**. LEI Nº 10.436, DE 24 DE ABRIL DE 2002. Dispõe sobre a Língua Brasileira de Sinais - Libras e dá outras providências. Brasília, 24 abr. 2002. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/leis/2002/l10436.htm](http://www.planalto.gov.br/ccivil_03/leis/2002/l10436.htm). Acesso em: 2 maio 2022.

CALCAGNO, Fabio; BAIS, Ricardo Ferraz de Arruda; FERREIRA, Samuel Macedo. **Desenvolvimento de uma aplicação web para educação de surdos. 2019. Trabalho de Conclusão de Curso** (Tecnólogo em Análise e Desenvolvimento de Sistemas) - Faculdade de Tecnologia de Americana, Americana, [S. l.], 2019. Disponível em: [http://ric.cps.sp.gov.br/bitstream/123456789/3923/1/20192S\\_CALCAGNOFabio\\_OD0706.pdf](http://ric.cps.sp.gov.br/bitstream/123456789/3923/1/20192S_CALCAGNOFabio_OD0706.pdf). Acesso em: 1 maio 2022.

GALVÃO, Jardiana Silva. DESAFIOS E DIFICULDADES DA COMUNIDADE SURDA POR MEIO DO APRENDIZADO DA LÍNGUA DE SINAIS. **DESAFIOS E DIFICULDADES DA COMUNIDADE SURDA POR MEIO DO APRENDIZADO DA LÍNGUA DE SINAIS**. 2018. Relato de Experiência (Graduanda em Secretariado Executivo) - Universidade Federal do Ceará, [S. l.], 2018. Disponível em: <http://brazil.enactusglobal.org/wp-content/uploads/sites/2/2018/11/DESAFIOS-E-DIFICULDADES-DA-COMUNIDADE-SURDA-POR-MEIO-DO-APRENDIZADO-DA-L%C3%8DNGUA-DE-SINAIS-94408.pdf>. Acesso em: 30 maio 2022.

IBGE – INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Censo Brasileiro de 2010**. Rio de Janeiro: IBGE, 2011. Disponível em: <https://www.ibge.gov.br/apps/snig/v1/?loc=0&cat=-1,-2,-3,128&ind=4643>. Acesso em: 30 abr. 2022

CAPOVILLA, F. C.; RAPHAEL, W. D. **Dicionário Enciclopédico Ilustrado Trilíngue – Língua de Sinais Brasileira**. São Paulo: EDUSP, 2001.

DOMINGOS, Maria Cristina da Silva. **Cartilha de Libras**, UNIFENAS, Alfenas, 2010. Disponível em: <http://www.unifenas.br/extensao/cartilha/CartilhaLibras.pdf>. Acesso em: 20 maio 2022.

FORSTER, Renê. **Desfazendo Mitos e Mentiras Sobre Línguas de Sinais**. 2004. Disponível em: [http://www.pglettras.uerj.br/linguistica/textos/livro02/LTAA02\\_a19.pdf](http://www.pglettras.uerj.br/linguistica/textos/livro02/LTAA02_a19.pdf). Acesso em: 1 jul. 2022.

KOT ENGENHARIA. **Visão Computacional: O que é?**. [S. l.], 1 jun. 2021. Disponível em: <https://kotengenharia.com.br/visao-computacional-o-que-e/>. Acesso em: 25 maio 2022.

LACERDA, Lucas. **Deep Learning & Visão Computacional — REDES NEURAIS CONVOLUCIONAIS**. [S. l.], 10 ago. 2019. Disponível em: <https://medium.com/@lucaaslb/deep-learning-vis%C3%A3o-computacional-redes-neurais-convolucionais-c21f19f5ec34>. Acesso em: 27 maio 2022.



MACHADO, Marcelo Chamy. **Classificação Automática de Sinais Visuais da Língua Brasileira de Sinais Representados por Caracterização Espaço-Temporal**. 2018. Dissertação (Mestrado em Informática) - Instituto de Computação da Universidade Federal do Amazonas, [S. l.], 2018. Disponível em: [https://tede.ufam.edu.br/bitstream/tede/6645/5/Disserta%C3%A7%C3%A3o\\_Marcelo%20Machado%20PPGI](https://tede.ufam.edu.br/bitstream/tede/6645/5/Disserta%C3%A7%C3%A3o_Marcelo%20Machado%20PPGI). Acesso em: 30 abr. 2022.

MILANO, Danilo de; HONORATO, Luciano Barrozo. **VISÃO COMPUTACIONAL**. Academia: Accelerating the world's research., [s. l.], 2014. Disponível em: [https://d1wqtxts1xzle7.cloudfront.net/35825905/2010\\_IA\\_FT\\_UNICAMP\\_visaoComputacional-with-cover-page-v2.pdf?Expires=1651640440&Signature=CR267BhgGndR631jN4yYTjYH1rFZHXB~nfTU~fA9SH~gj42n2N9WQUpFY4KoQrxwJO9MZldKYSQlqrcumKbghURt2ANentq5D BCqYYF~hPeB5SUXEW86zdH7Lhk8hK8FnGRDGIKySJGu6jP7nJ7kRIUih2k6-RTsh9qfyEUbQ--u4wJewXp8RxbjOMyl~zHoIUvqj6RIldNoLNSJ8TPHyTlqi-iMujdfybVeefxc7wO1d4vW2-KQh5aQ6Yyx8Qqbqcridr0TzYByHnIIrDRsFc9-uZ7gt~gMDGYem-Errc~D94T8JnALT7Y-psF1jjzFT90ULtPZWiyuNXOoPMRz7w\\_\\_&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA](https://d1wqtxts1xzle7.cloudfront.net/35825905/2010_IA_FT_UNICAMP_visaoComputacional-with-cover-page-v2.pdf?Expires=1651640440&Signature=CR267BhgGndR631jN4yYTjYH1rFZHXB~nfTU~fA9SH~gj42n2N9WQUpFY4KoQrxwJO9MZldKYSQlqrcumKbghURt2ANentq5D BCqYYF~hPeB5SUXEW86zdH7Lhk8hK8FnGRDGIKySJGu6jP7nJ7kRIUih2k6-RTsh9qfyEUbQ--u4wJewXp8RxbjOMyl~zHoIUvqj6RIldNoLNSJ8TPHyTlqi-iMujdfybVeefxc7wO1d4vW2-KQh5aQ6Yyx8Qqbqcridr0TzYByHnIIrDRsFc9-uZ7gt~gMDGYem-Errc~D94T8JnALT7Y-psF1jjzFT90ULtPZWiyuNXOoPMRz7w__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA). Acesso em: 2 maio 2022.

NISHIO, Mizuho et al. **Convolutional neural networks: an overview and application in radiology**. Insights into Imaging, [s. l.], 22 jun. 2018. DOI <https://doi.org/10.1007/s13244-018-0639-9>. Disponível em: <https://link.springer.com/article/10.1007/s13244-018-0639-9#citeas>. Acesso em: 30 maio 2022.

QUADROS, Ronice Muller de; CRUZ, Carina Rebello. **Língua de Sinais: Instrumentos de avaliação**. [S. l.], 2011. Disponível em: [https://www.google.com.br/books/edition/L%C3%ADngua\\_de\\_Sinais/JVYWP6btj4C?hl=pt-BR&gbpv=1&dq=inauthor:%22Ronice+M.+QUADROS%22&printsec=frontcover](https://www.google.com.br/books/edition/L%C3%ADngua_de_Sinais/JVYWP6btj4C?hl=pt-BR&gbpv=1&dq=inauthor:%22Ronice+M.+QUADROS%22&printsec=frontcover). Acesso em: 30 maio 2022.

VIEIRA, Thiago. **LIBRAS: principais dificuldades no uso da língua de sinais**. [S. l.], 2016. Disponível em: <https://psicopedagogiacritica.wordpress.com/2016/07/25/libras-as-principais-dificuldades-no-uso-da-lingua-de-sinais/>. Acesso em: 21 ago. 2022.