

# Conexão Linguística: Leitura e interpretação de sinais gramaticais de LIBRAS utilizando machine learning

Francisco Wallyson do Nascimento Silva<sup>1</sup>, Cornelia Janayna Pereira Passarinho<sup>2</sup>

<sup>1</sup>Universidade Estadual do Piauí (UESPI)

fwnsilva@aluno.uespi.br, janaynapassarinho@prp.uespi.br

**Resumo.** LIBRAS é uma língua de sinais utilizada pela comunidade brasileira como forma de expressão e comunicação, através de sinais e de expressões corporais e faciais, tendo como principal intuito a inclusão social e a acessibilidade linguística para pessoas surdas. Esta pesquisa visa ao desenvolvimento de um software leitor de sinais que representem expressões de LIBRAS, transcrevendo esses símbolos gramaticais para a língua portuguesa. Para tal, utilizou-se uma ferramenta desenvolvida para JavaScript, uma biblioteca chamada TensorFlow, utilizada para criação de modelos de aprendizado de máquina a nível de produção. O diferencial deste trabalho reside no desenvolvimento de uma ferramenta web para interpretação de sinais gramaticais de LIBRAS, utilizando técnicas de aprendizado de máquina e bibliotecas como TensorFlow e MediaPipe, sem a necessidade de servidores back-end robustos. A continuidade da pesquisa propõe direções futuras que visam ampliar a precisão e a abrangência dos resultados obtidos.

## 1. Introdução

A Língua Brasileira de Sinais (LIBRAS) é fundamental para a inclusão e comunicação da comunidade surda no Brasil. Reconhecida como uma língua completa e independente, possui sua própria estrutura linguística, incluindo sintaxe e gramática distintas [Silva 2024]. Ao contrário do que muitos pensam, a LIBRAS não é uma simples tradução do Português para sinais, mas uma língua natural com características únicas.

Cerca de 8,4% da população brasileira, ou aproximadamente 17,2 milhões de pessoas, possuem algum tipo de deficiência, incluindo 2,3 milhões com algum grau de deficiência auditiva, segundo dados da *Pesquisa Nacional de Saúde: 2019* [IBGE 2022]. A LIBRAS desempenha um papel crucial ao permitir a participação desses indivíduos no acesso à educação, cultura e informação.

Embora existam centenas de variações de línguas de sinais ao redor do mundo, poucas são amplamente conhecidas ou utilizadas pela maioria da população surda [Lobato 2021]. [Agência IBGE 2021] mostra que, no Brasil, a LIBRAS enfrenta desafios similares e que, segundo dados do IBGE de 2019, sua adoção não é generalizada entre os surdos.

Nesse contexto, este estudo propõe o desenvolvimento de uma ferramenta inovadora: um *software web* que traduz sinais gramaticais da LIBRAS para o Português utilizando *machine learning*. Essa ferramenta visa facilitar o aprendizado para pessoas surdas e ouvintes, promovendo uma experiência educacional inclusiva e interativa. Esta

pesquisa também abre caminhos para futuras investigações, focando em aprimorar a precisão em diferentes condições e na integração de novas tecnologias que permitam uma detecção mais robusta e abrangente dos sinais.

## 2. Trabalhos Relacionados

A Internet é considerada uma das Tecnologias de Informação e Comunicação (TIC). Essa ferramenta possibilita a elaboração de materiais didáticos interativos, como jogos, animações e simulações, que podem auxiliar no contexto do ensino de uma segunda língua [de Amorim Silva et al. 2016].

Em [da Silva 2022] apresentou-se uma ferramenta que utiliza o *framework MediaPipe* para capturar expressões faciais, posição e configuração das mãos em vídeos e traduzir palavras e letras em LIBRAS para o Português brasileiro. Esta ferramenta conseguiu encontrar o significado dos sinais referentes às palavras Deus, direito, cunhado e muro, além das letras A, B, C e D, identificando cada parâmetro formador do sinal individualmente, e depois combinando-os para traduzir a palavra. Em contraste, o uso do *TensorFlow.js* no presente trabalho permite a execução direta no navegador, reduzindo significativamente o tempo de resposta em comparação à solução anterior.

Utilizar *TensorFlow* e *MediaPipe* juntas é uma prática recomendada em muitos projetos de visão computacional e aprendizado de máquina devido à sinergia que essas duas bibliotecas oferecem. Em [Caiafa et al. 2023], o método proposto utiliza o *MediaPipe* para a extração de pontos-chave dos usuários, *K-Means* para agrupar esses pontos e classificadores KNN e *Random Forest* para classificar os agrupamentos. Para desenvolver as redes neurais, foi utilizada a biblioteca *TensorFlow*.

Comparado aos trabalhos apresentados anteriormente, este projeto se destaca por utilizar uma abordagem em tempo real para a detecção de sinais de LIBRAS, utilizando as bibliotecas *MediaPipe Handpose* e *TensorFlow.js*, fazendo com que a aplicação consiga executar perfeitamente no navegador, proporcionando uma experiência de aprendizado mais interativa e eficiente para os usuários.

## 3. Proposta do Trabalho

Este projeto propõe o desenvolvimento de um interpretador de sinais gramaticais da LIBRAS via *web*, com o objetivo de auxiliar no aprendizado, por meio da formação de palavras. O *pipeline* do sistema proposto, pode ser visto na Figura 1, e visa servir como base para criar ferramentas futuras e inclusivas.

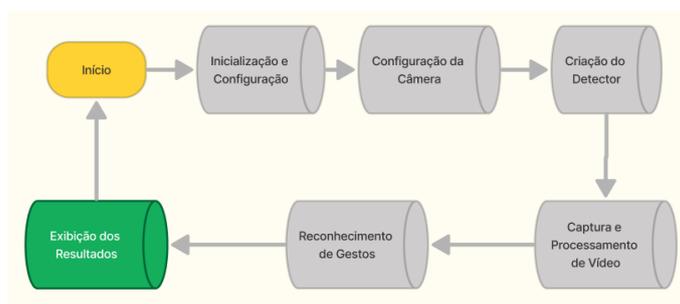


Figura 1. *Pipeline* do Sistema Proposto

### 3.1. Materiais e Métodos

Para o desenvolvimento do interpretador de sinais gramaticais, optou-se pelo ambiente de desenvolvimento *VS Code*, uma *IDE* que oferece recursos avançados de edição de código e depuração. Utilizou-se o *MediaPipe*, um *framework* de código aberto do *Google*, para processar as imagens capturadas pela câmera, permitindo o reconhecimento e a interpretação gramatical dos sinais da LIBRAS em tempo real. Especificamente, a parte *Handpose* do *MediaPipe* foi crucial, focando no rastreamento e detecção detalhada das mãos, identificando *landmarks* que são cruciais para a aplicação. A Figura 2 ilustra a divisão desses *landmarks*.



Figura 2. Pontos-chave *MediaPipe Handpose* [TensorFlow.js Models nd]

Para o estudo e análise de casos, utilizou-se a biblioteca *Fingerpose*, uma ramificação da *Handpose*, com ênfase na detecção de gestos e pontos chave dos dedos. A biblioteca *TensorFlow*, desenvolvida pelo *Google*, é crucial para o treinamento e monitoramento dos sinais de LIBRAS.

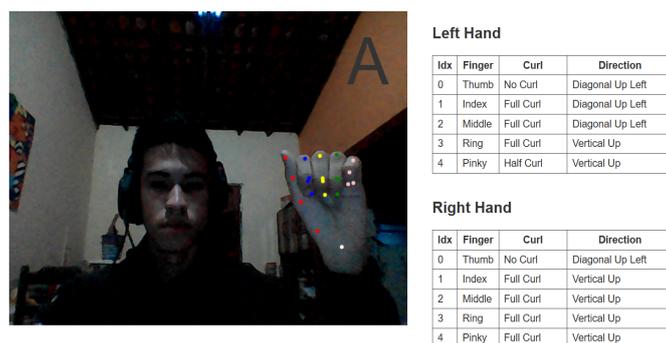
A aplicação desenvolvida se destaca por utilizar *TensorFlow.js* em vez do *TensorFlow* convencional, permitindo execução no navegador sem a necessidade de servidores *back-end* robustos. Para controle e versionamento do código, foi utilizado o *GitHub*, enquanto para transmissão de vídeo em tempo real, foi usado o *Open Broadcaster Software (OBS)*, uma ferramenta gratuita e de código aberto.

## 4. Resultados e Discussão

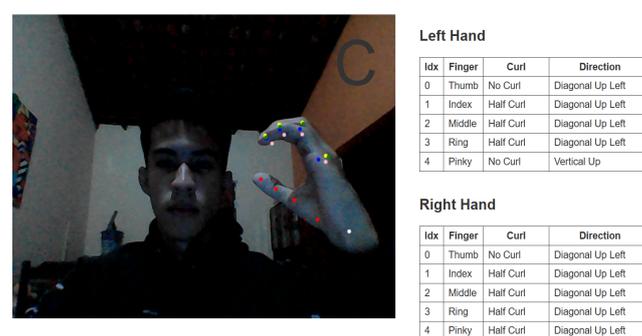
Com a biblioteca *FingerPose*, obteve-se um resultado positivo na detecção de sinais das mãos correspondentes ao alfabeto LIBRAS, utilizando as propriedades de curvatura dos dedos. A Figura 3 ilustra o sinal da letra A, onde a curvatura dos dedos indicador, anelar, médio e mindinho é necessária para a correta identificação.

Para uma identificação mais precisa, é essencial posicionar a mão corretamente em relação à câmera. Testes com diferentes fundos e iluminação, configurados via *OBS*, mostraram que variações no cenário não afetam a detecção, desde que a iluminação seja adequada. No entanto, a qualidade do *hardware* e limitações da biblioteca podem levar a detecções incorretas se a mão não estiver corretamente posicionada.

A letra A em LIBRAS requer que os dedos indicador, anelar, médio e mindinho estejam totalmente curvados, enquanto a letra B exige que esses dedos estejam esticados, representados como "Vertical up". A letra C, por sua vez, depende da posição correta da mão, verificada pela propriedade *Handpose*, como mostrado na Figura 4.



**Figura 3. Letra A do alfabeto em LIBRAS [Libras 2022]**



**Figura 4. Letra C do alfabeto em LIBRAS [Libras 2022]**

Para lidar com os sinais identificados para estudo tem-se as seguintes propriedades, onde cada numero informado representa um dedo, conforme mostrado na tabela 1.

**Tabela 1. Representação das Propriedades das Mãos na LIBRAS**

| Categoria                 | Nome                                  | Descrição                            |
|---------------------------|---------------------------------------|--------------------------------------|
| Dedos                     | <i>Finger.Thumb</i>                   | Polegar                              |
|                           | <i>Finger.Index</i>                   | Indicador                            |
|                           | <i>Finger.Middle</i>                  | Médio                                |
|                           | <i>Finger.Ring</i>                    | Anelar                               |
|                           | <i>Finger.Pinky</i>                   | Mínimo                               |
| Curvatura                 | <i>FingerCurl.NoCurl</i>              | Não Curvado                          |
|                           | <i>FingerCurl.HalfCurl</i>            | Meio-Curvado                         |
|                           | <i>FingerCurl.FullCurl</i>            | Curvado Completamente                |
| Direção                   | <i>Vertical Up</i>                    | Vertical para cima                   |
|                           | <i>Vertical Down</i>                  | Vertical para baixo                  |
|                           | <i>Horizontal Left</i>                | Horizontal para a esquerda           |
|                           | <i>Horizontal Right</i>               | Horizontal para a direita            |
|                           | <i>Diagonal Up Right</i>              | Diagonal para cima e para a direita  |
|                           | <i>Diagonal Up Left</i>               | Diagonal para cima e para a esquerda |
|                           | <i>Diagonal Down Right</i>            | Diagonal para baixo e para a direita |
| <i>Diagonal Down Left</i> | Diagonal para baixo e para a esquerda |                                      |

Além disso, identificaram-se algumas indicações linguísticas em LIBRAS com suas características distintas, como o gesto para "Beber", que simula o ato de beber de um copo com a palma voltada para baixo, e o sinal para "Nome", apontando o dedo indicador ao peito para indicar o nome de alguém. Para "Obrigado", forma-se a letra S com a mão dominante ao peito, realizando um movimento circular. Já os sinais para

”Ir”, ”Vou”e ”Fui”envolvem apontar o dedo indicador para si mesmo e indicar a direção desejada, variando conforme o contexto subsequente.

Utilizou-se um vídeo do *YouTube* de [Libras 2022], intitulado, espelhado via *OBS* para avaliar a precisão no reconhecimento de palavras-chave como 'IR', 'Beber', 'Nome' e 'Obrigado'.

No entanto, o estudo encontrou limitações na precisão da análise devido às características da biblioteca utilizada. Por exemplo, o sinal ”Beber”exibiu propriedades como *Diagonal Up* e *Full Curl* nos dedos indicador, médio, anelar e mínimo, enquanto o polegar assumiu à posição *Vertical Up*, como mostrado na Figura 6.

Em LIBRAS, elementos gramaticais semelhantes podem ter significados diferentes dependendo do contexto, como exemplificado pelo sinal ”IR”, que pode indicar tanto ”vou”quanto ”fui”, conforme visto na Figura 5 A aplicação desenvolvida alcança precisão comparável aos estudos anteriores na identificação de sinais, sendo executada diretamente no navegador para maior acessibilidade, facilidade de uso e menor latência na comunicação.



**Left Hand**

| Idx | Finger | Curl | Direction |
|-----|--------|------|-----------|
| 0   | Thumb  | -    | -         |
| 1   | Index  | -    | -         |
| 2   | Middle | -    | -         |
| 3   | Ring   | -    | -         |
| 4   | Pinky  | -    | -         |

**Right Hand**

| Idx | Finger | Curl      | Direction        |
|-----|--------|-----------|------------------|
| 0   | Thumb  | No Curl   | Vertical Up      |
| 1   | Index  | Half Curl | Diagonal Up Left |
| 2   | Middle | Full Curl | Diagonal Up Left |
| 3   | Ring   | Full Curl | Diagonal Up Left |
| 4   | Pinky  | Full Curl | Vertical Up      |

Figura 5. Sinal Ir em LIBRAS [lib 2020]



**Left Hand**

| Idx | Finger | Curl      | Direction         |
|-----|--------|-----------|-------------------|
| 0   | Thumb  | No Curl   | Horizontal Right  |
| 1   | Index  | Half Curl | Horizontal Right  |
| 2   | Middle | Full Curl | Diagonal Up Right |
| 3   | Ring   | Full Curl | Horizontal Right  |
| 4   | Pinky  | Full Curl | Horizontal Right  |

**Right Hand**

| Idx | Finger | Curl      | Direction        |
|-----|--------|-----------|------------------|
| 0   | Thumb  | No Curl   | Horizontal Left  |
| 1   | Index  | Full Curl | Diagonal Up Left |
| 2   | Middle | Full Curl | Diagonal Up Left |
| 3   | Ring   | Full Curl | Diagonal Up Left |
| 4   | Pinky  | Full Curl | Vertical Up      |

Figura 6. Sinal Beber em LIBRAS [lib 2020]

A biblioteca *Handpose FingerPose* é uma ferramenta bastante poderosa para a detecção de gestos em grande escala; no entanto, para o uso em uma aplicação *Web* simultânea, ela apresenta algumas limitações. Algumas limitações de *hardware* de baixo nível, como câmeras que não têm muita qualidade, podem comprometer a precisão na captação dos sinais, sendo assim um empecilho. O ângulo em que a câmera está posicionado é um crucial para o desempenho da ferramenta.

## 5. Conclusão

Os resultados indicam que, embora útil, a biblioteca *Handpose Fingerpose* tem limitações de precisão, afetada por fatores como a baixa luminosidade e qualidade da câmera. A expressividade dos sinais e a regionalização também apresentam desafios, visto que uma mesma expressão pode variar entre estados.

Apesar disso, a construção de uma ferramenta mais desenvolvida é viável usando diretamente a biblioteca *Handpose*, que se mostra mais completa e eficaz que sua ramificação *Fingerpose*. O uso de um modelo pré-treinado não foi ideal para lidar com a complexidade dos elementos gramaticais. Com avanços na biblioteca, será possível melhorar a precisão de detecção, minimizando as limitações de *hardware*.

Dentre as principais direções futuras, destacam-se aumentar a quantidade de sinais e expressões a serem identificados com a integração das bibliotecas. Este estudo espera contribuir para a tecnologia de reconhecimento de sinais e promover uma compreensão mais profunda das necessidades das pessoas surdas na interação com dispositivos tecnológicos.

## Referências

- (2020). Libras básico. <https://educarepedagogia.com.br/wp-content/uploads/2020/11/Libras-basico.pdf>. Acessado em: 2023-10-06.
- Agência IBGE (2021). Um em cada quatro idosos tinha algum tipo de deficiência em 2019.
- Caiafa, E. G. d. A., Basilio, A., de Lima, A. A., and Araujo, G. M. (2023). Interpretação de gestos de libras usando k-means e random forest. *XLI SIMPÓSIO BRASILEIRO DE TELECOMUNICAÇÕES E PROCESSAMENTO DE SINAIS*.
- da Silva, R. P. (2022). Visão computacional: um estudo de caso aplicado à língua brasileira de sinais (libras). *Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte*.
- de Amorim Silva, R., Lima, L. G., and Bastos, R. O. (2016). Aperfeiçoando o aprendizado de libras utilizando elementos de internet das coisas.
- IBGE (2022). Pesquisa nacional de saúde: 2019: ciclos de vida: Brasil e grandes regiões.
- Libras, N. (2022). Aprenda os 50 sinais mais fáceis mais usados da libras (são mais que 50). <https://www.youtube.com/watch?v=7yepxfXKF4M>. Acesso em: 14 set. 2023.
- Lobato, L. (2021). Ibge confirma: surdez não é sinônimo de libras. <https://desculpenaoouvi.com.br/ibge-confirma-surdez-nao-e-sinonimo-de-libras/>.
- Silva, D. N. (2024). Língua brasileira de sinais (libras); brasil escola. <https://brasilecola.uol.com.br/educacao/lingua-brasileira-sinais-libras.htm>. Acessado em: 2023-10-06.
- TensorFlow.js Models (n.d.). Mediapipe hands keypoints. <https://github.com/tensorflow/tfjs-models/tree/master/hand-pose-detection#mediapipe-hands-keypoints-used-in-mediapipe-hands>.