

Reconhecimento e tradução da linguagem de sinais brasileira (LIBRAS) utilizando visão computacional e Machine Learning

Eric Cerqueira Lima de Menezes¹; Lucas Amparo Barbosa²

¹ Bolsista; Iniciação Científica; ericmenezes2@gmail.com

² Centro Universitário SENAI CIMATEC; Salvador - BA; lucas.barbosa@fieb.org.br

RESUMO

O projeto em questão é um Identificador e Tradutor da Linguagem Brasileira de Sinais (LIBRAS), usando Visão Computacional e Aprendizado de Máquina (Machine Learning). Foi desenvolvido com Python, fazendo uso das bibliotecas OpenCV para Visão Computacional e PyTorch para Machine Learning. Onde um banco de dados de Libras foi feito em conjunto com a otimização do protótipo feito anteriormente..

PALAVRAS-CHAVE: Aprendizado de Máquina; Visão Computacional; LIBRAS; Banco de dados.

1. INTRODUÇÃO

Atualmente, o Brasil segundo a Pesquisa Nacional de Saúde (PNS) possui cerca de 2,3 milhões de pessoas com deficiência auditiva(1). No entanto, ainda enfrentamos uma carência de recursos e modelos dedicados à identificação da Linguagem Brasileira de Sinais (LIBRAS), especialmente quando se trata da identificação de sinais contínuos, ou seja, aqueles que envolvem movimento. Com embasamento nisso foi feito um identificador para identificar os sinais correspondentes aos números e as letras do alfabeto do Português brasileiro.

2. METODOLOGIA

De início para atingir os objetivos deste projeto,diversas imagens de múltiplos sinais foram capturadas,possibilitando a identificação e tradução da Linguagem Brasileira de Sinais (LIBRAS). A biblioteca MediaPipe, desenvolvida pela Google, foi utilizada para detecção das mãos e extrair pontos de referência, chamados de landmarks para cada imagem capturada. Esses landmarks forneceram conjuntos de coordenadas normalizadas, que serviram como base para a criação de uma base de dados(dataset inicial). Cada conjunto de landmarks representava um sinal estático em LIBRAS. Tendo em base este método, foi criado um base de dados contendo fotos com os estáticos que representam os números. Após isso, foi criada uma nova base de dados mais robusta, em que possuía os mesmos sinais, adicionando variações de posições e ângulos de captura. Além de que foi criado um banco de dados de cerca de 5000 imagens para os sinais que representam as letras do alfabeto português brasileiro e após a serialização dos landmarks foi possível detectar sinais dinâmicos simples. A partir disso foi feito um uma interface gráfica (GUI) para facilitar o uso deste protótipo.

3. RESULTADOS E DISCUSSÃO

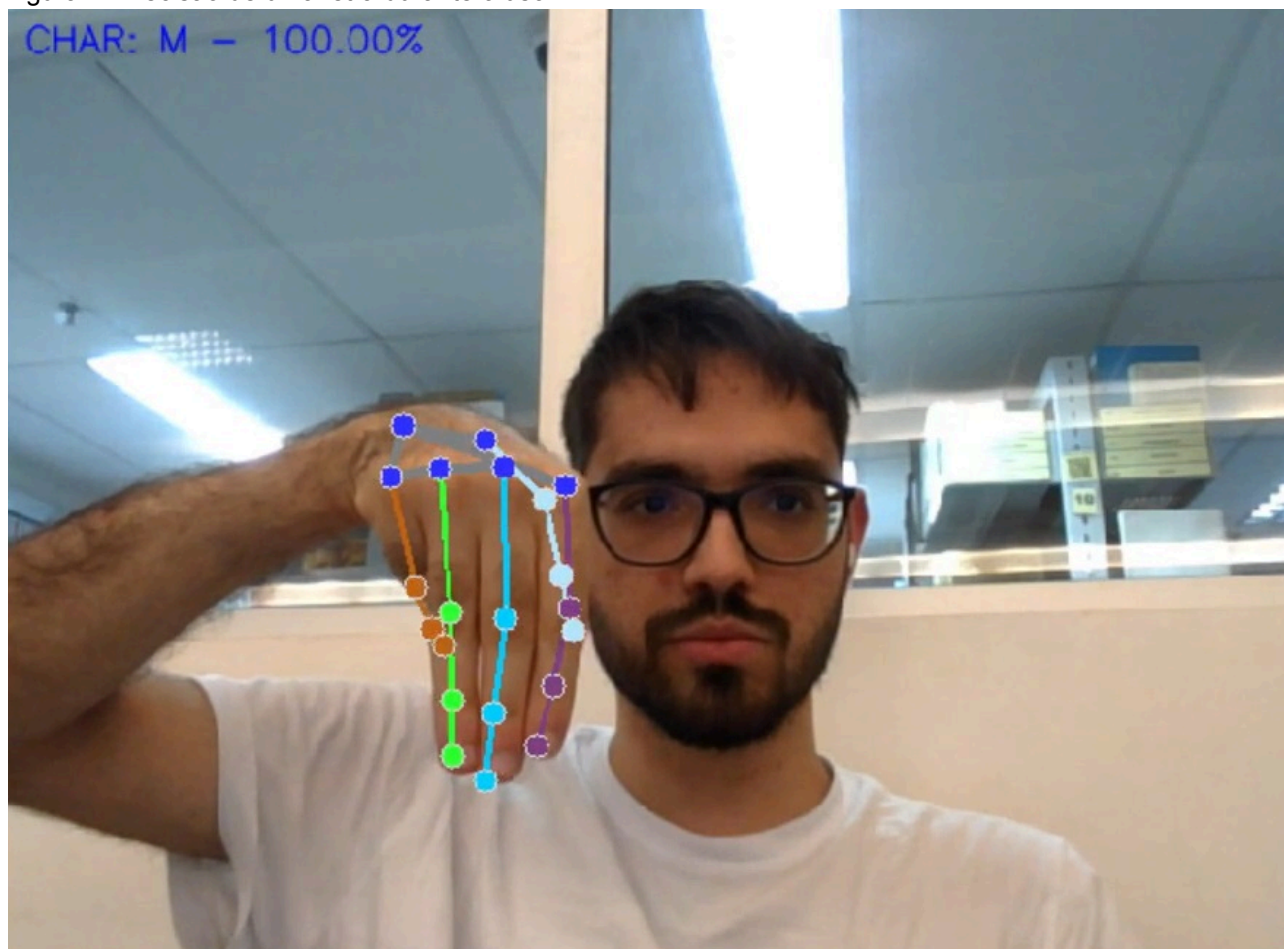
Estudar e criar uma base de dados para estudos futuros para o reconhecimento de sinais brasileiros é importante para a acessibilidade de pessoas portadoras de surdez e para sua inclusão na sociedade moderna e este estudo facilita a criação de futuros projetos nesta área.

Através de aprendizado de máquina como o MLP e o Random Forest em uma base de dados de imagens Fig. 1 é possível processar imagens e serializar seus landmarks e as detectar com uma acurácia boa de 90% dos sinais de LIBRAS estáticos e contínuos com baixo custo computacional Fig.2.

Figura 1: Exemplos de imagens da base de dados criada.

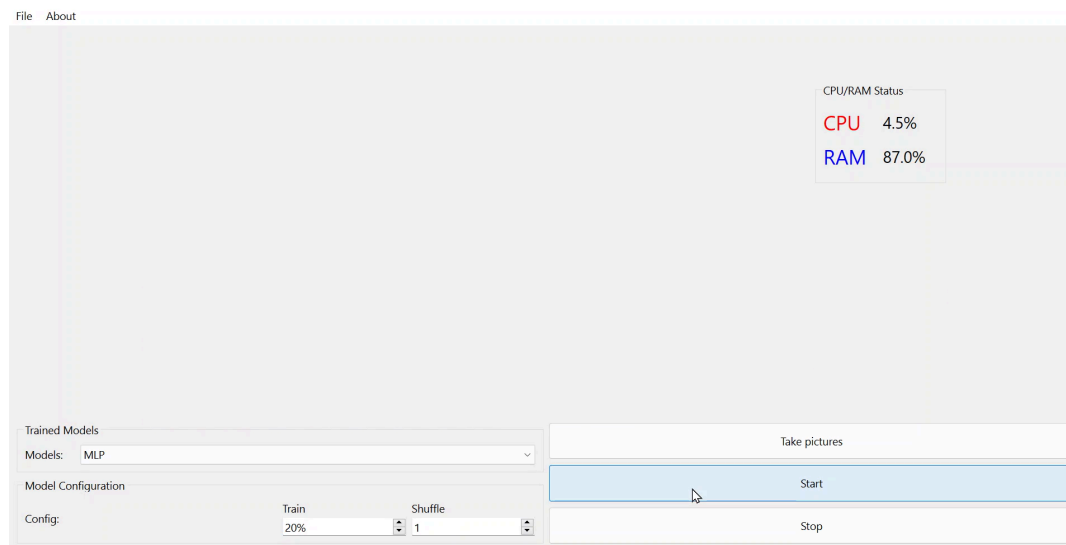


Figura 2: Precisão de uma letra durante o uso.



A interface feita em Qt4 utilizando a biblioteca PySide para o protótipo, apresentou uma interface simples, leve e amigável Fig. 3.

Figura 3: Interface do protótipo feita em Qt4.



4. CONSIDERAÇÕES FINAIS

Com os resultados apresentados, é possível notar que um sistema de tradução de LIBRAS pode facilitar a comunicação entre pessoas com deficiência e pessoas que não possuem conhecimento em LIBRAS. Neste contexto, os modelos implementados, incluindo a Random Forest e a MLP e a base de dados (dataset), demonstraram um desempenho satisfatório, superando os desafios de distinguir sinais semelhantes e fornecendo uma base sólida para futuros desenvolvimentos. A partir disso foi possível perceber que uma base de dados ainda mais robusta será necessária para detectar frases em libras e a futura implementação de um aprendizado de máquina como o (LSTM) para a detecção de frases e sinais que possuem contextos mais complexos.

Agradecimentos

Ao meu orientador Lucas Amparo Barbosa que me acompanhou durante este projeto e para meu colega de equipe João Pedro Almeida que idealizou o projeto e começou o desenvolvimento do mesmo.

5. REFERÊNCIAS

- ¹ IBGE. **PNS 2019: país tem 17,3 milhões de pessoas com algum tipo de deficiência**. Estatísticas sociais. 2021. Disponível em : <https://agenciadenoticias.ibge.gov.br/agencia-sala-de-imprensa/2013-agencia-de-noticias/releases/31445-pns-2019-pais-tem-17-3-milhoes-de-pessoas-com-algum-tipo-de-deficiencia/> Acesso em: 15 mar. 2024
- ² <https://developers.google.com/mediapipe>
- ³ D. Kelly, J. Mc Donald and C. Markham, "Continuous recognition of motion based gestures in sign language," *2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops*, Kyoto, Japan, 2009, pp. 1073-1080, doi: 10.1109/ICCVW.2009.5457585.
- ⁴ Sultan, Ahmed, Makram, Walied, Kayed, Mohammed and Ali, Abdelmaged Amin. "Sign language identification and recognition: A comparative study" *Open Computer Science*, vol. 12, no. 1, 2022, pp. 191-210. <https://doi.org/10.1515/comp-2022-0240>.