



VI Encontro de Iniciação Científica e
Tecnológica
VI EnICT
ISSN: 2526-6772
IFSP – Câmpus Araraquara
21 e 22 de outubro de 2021



Desenvolvimento de aplicativo para promover comunicação e interação social para surdos

Miguel Alves¹, Gabriel Justino², Franciele Neves³, Eliane Comoli⁴

^{1 2 3} Graduando em Tecnologia de Informática para Internet, IFSP, miguel.alves@aluno.ifsp.edu.br, g.justino@aluno.ifsp.edu.br, franciele.neves@aluno.ifsp.edu.br

⁴ Doutora em Ciências Fisiológicas e Mestre em Neurociência e Comportamento. FMRP, USP, ecomoli@fmrp.usp.br

Área de conhecimento (Tabela CNPq): 1.03.03.05-7 Processamento Gráfico (Graphics)

RESUMO: Interagir e comunicar-se com as outras pessoas é uma necessidade humana básica. A surdez é uma especificidade na qual o indivíduo percebe o mundo de forma diferente, porém capaz de exercer funções linguísticas e cognitivas desde que interaja adequadamente com o meio para se desenvolver e exercer a sua função social. Utilizar uma língua adequada como Libras (Língua Brasileira de Sinais) é essencial, porém pouco difundida. Muitas vezes excluídos da sociedade e sem amparo suficiente, os surdos necessitam de uma ferramenta que auxilie na comunicação. Visando esta problemática entre surdos e falantes da língua portuguesa, o projeto atual objetivou desenvolver um software que promova uma melhor comunicação e interação social. Com o emprego da linguagem de programação Python, surgiu o Librare, que através da captação de voz reconhece padrões textuais da fala e utilizando Linguagem Natural converte essa informação em sinais de Libras. O mapeamento dos pontos chave do corpo humano é feito pela biblioteca Mediapipe, que junto ao aprendizado de máquina classifica os sinais gerando um padrão que permite reconhecer a mensagem comunicada e transmutá-la em som e texto. O Librare poderá contribuir para a autonomia e facilitar a inclusão e socialização do surdo.

PALAVRAS-CHAVE: Libras; Python; Neurociência; Tecnologia; Mediapipe;

INTRODUÇÃO

Segundo a Organização Mundial da Saúde (LAYTON ET AL., 2019), cerca de 360 milhões de pessoas apresentam algum grau de surdez, sendo no Brasil em torno de 10 milhões ou 5% da população. Isso reflete a necessidade de desenvolvimento de ações inclusivas e tecnologias assistivas que aumentem a habilidade e independência dessas pessoas para as atividades sociais, educacionais e de trabalho.

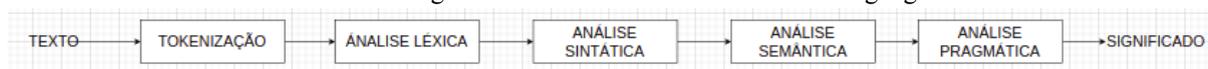
No Brasil os surdos se comunicam através de Libras, como outras línguas de sinais, esta foi construída sob os seguintes parâmetros: a configuração das mãos para formar o sinal, o movimento que o sinal possui, a localização onde é feito, orientação e posicionamento da palma da mão, além da expressão facial/corporal (ROCHA E MELGAÇO, 2018). A linguagem dos sinais é acessível pela visão, e a neurociência aponta que habilidades visuais específicas como a discriminação de objetos e expressões faciais, funções visuais periféricas como localização, e sincronização visuo-motora estão aumentadas em surdos. Ocorre uma reorganização funcional de modo que áreas corticais auditivas são responsáveis pelas habilidades visuais em surdos (ALENCAR ET AL., 2019). Assim, a tecnologia assistiva tem um papel importante, buscando suprir as necessidades sensoriais do indivíduo com deficiência e facilitar a sua comunicação e interação social. Como os softwares disponíveis são capazes de fazer o “português sinalizado”, um sistema de tradução de sons em sinais de Libras em tempo real e fiel à interpretação da comunidade surda é desejável. O software deve captar a voz e reconhecer padrões textuais, e convertê-los em sinais de Libras.

FUNDAMENTAÇÃO TEÓRICA

A Programação Neurolinguística, nasceu da observação de comportamentos e padrões linguísticos de pessoas de destaque e aos poucos foi se construindo um código de relacionamento, o qual não levava em consideração a estrutura cerebral dos envolvidos e tampouco seus sistemas neurais de interação com o ambiente. Tem como objetivo estudar aspectos neurais, de linguagem, cognitivos e psicológicos.

Criado originalmente em 2001, o NLTK é uma plataforma usada para construir programas que trabalham com dados de linguagem humana para aplicação em processamento de linguagem natural (PLN). A sua biblioteca em Python com o mesmo nome fornece classes básicas para representar dados relevantes para o PLN; utilizando a abordagem clássica do PLN conforme os estágios ilustrados na Figura 1.

FIGURA 1. Estágios da análise no Processamento de Linguagem Natural.



Fonte: Adaptado de Nascimento (2017).

A implementação do software exigiu um pré-conhecimento em relação a linguagem Kivy. O Kivy é uma biblioteca Open Source escrita em Python para o desenvolvimento de aplicações multiplataforma, seja para desktop ou mobile (ANDRADE, 2021). Para o processamento do vídeo o framework Mediapipe ofereceu suporte para as várias plataformas envolvidas no desenvolvimento do software, visto que é uma biblioteca Open Source multiplataforma, que utiliza soluções de aprendizado de máquina para resolução em tempo real de streaming (HAYS & MULLEN, 2021).

METODOLOGIA

A identificação de termos passíveis de tradução se dá após o processamento do áudio em texto. A análise dos termos é encaminhada ao banco de dados que retorna a sequência de execução e a exibição é feita dentro do próprio aplicativo. A captação do vídeo é realizada através da biblioteca OpenCv, que conecta de forma independente à webcam e dá acesso aos frames que serão processados. Entretanto, essa abordagem não seria diferente da encontrada no aplicativo Hand Talk (TENÓRIO, 2021), que é utilizado somente para mostrar para os ouvintes a forma de se comunicar com o deficiente auditivo, além de ser extremamente impreciso e incorreto do ponto de vista léxico, apontado pelo apoiador intérprete.

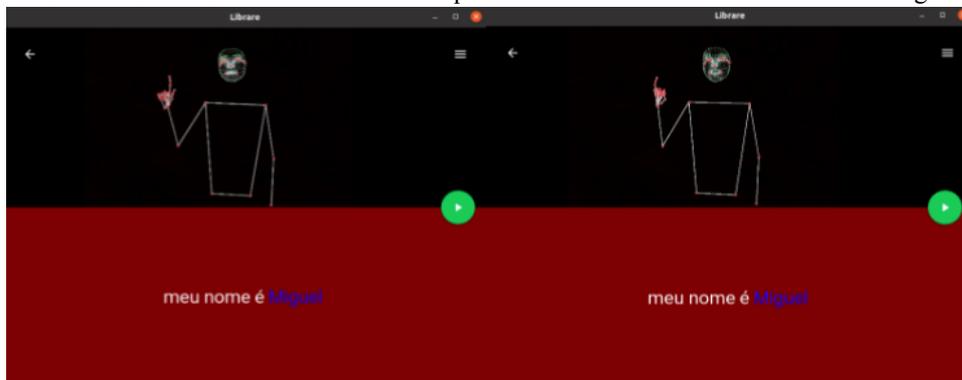
Assim a implementação do software encaminhou-se para o processamento de imagem por meio da utilização de redes neurais, entretanto a inexistência de dados impossibilitou sua continuação. Graças à quantidade massiva de vídeos disponibilizados no youtube por intérpretes surgiu a iniciativa de criar um banco de dados de Libras gratuito e que facilitasse a continuação do desenvolvimento do software.

A captação dos dados foi realizada pela biblioteca Mediapipe, que possui compatibilidade tanto com Python quanto com Android. Como resultados parciais foi desenvolvido uma espécie de glossário aderente a novas traduções e variações de uma mesma tradução. A criação de um glossário se dá pela identificação da frase ou palavra a ser traduzida e o isolamento dessa frase em um único vídeo. Ou seja, a partir de um vídeo de glossário do youtube, onde há a citação de mais de uma frase, o processamento retornaria vídeos menores com as transcrições de apenas um sinal em cada vídeo. A realização técnica desse processo se dá pela identificação de frases na imagem utilizando a biblioteca EasyOcr, o vídeo menor é então criado e os frames são inseridos enquanto a frase ou palavra identificada não se altera. Quando há a modificação da frase o vídeo é fechado e a construção do próximo é iniciado, de forma totalmente independente, o que permite a execução em larga escala e sem um intenso supervisionamento.

RESULTADOS E DISCUSSÃO

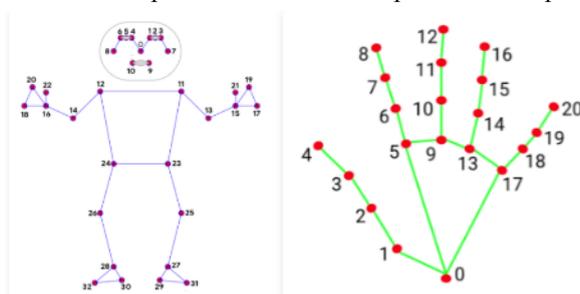
O Librare, converte nomes e substantivos próprios com facilidade, pois, em sua maior parte, seus significados correspondem às letras que integram a escrita dessas palavras. Assim sendo, um pequeno banco de dados com as 26 letras do alfabeto foi criado e a tradução para esses nomes ocorre através de sinais individuais para cada uma das letras conforme ilustrado na Figura 2.

FIGURA 2. Conversão das letras ‘G’ à esquerda e ‘I’ à direita inseridas no nome “Miguel”.



Com o Librare é possível traduzir frases curtas como “meu nome é”, mas não frases extensas, uma vez que há a ausência do banco de dados sobre sinais de libras com livre acesso. O software também foi programado para ler os movimentos, expressões e gestos realizados em libras para a sua tradução falada. Com a biblioteca Mediapipe foi possível mapear o corpo humano em mais de 542 pontos específicos em diversas partes do corpo. Sendo 32 pontos corporais (Figura 3, esquerda), 21 pontos em cada mão (Figura 3, direita) além de 468 pontos faciais.

FIGURA 3. Mapeamento do corpo e da mão direita disponibilizados pelo site oficial do Mediapipe.



Fonte: Mediapipe.

Esse mapeamento é coordenado, ou seja, o software retorna as coordenadas espaciais em terceira dimensão, transpondo-as ao ponto numérico correspondente (1,2,3...) no mapa corporal. Conforme a pessoa se move, o software é capaz de rastrear os pontos no espaço enquanto retorna as coordenadas modificadas a uma velocidade de 30fps, ou seja 30 frames ou imagens por segundo. As linhas conectam alguns pontos específicos do corpo gerando um esqueleto. Por essa razão optou-se por utilizar a ilustração do próprio esqueleto a fim de evitar qualquer cópia de imagem de terceiros. Numa etapa futura pretende-se materializar o boneco dando a ele uma aparência mais simpática.

O glossário gerado com os vídeos do youtube possui pequenas frases comuns no dia-a-dia, como “Desculpa”, “Bom dia”, “Boa tarde”, “Como é o seu nome?”, “Eu me chamo...” e algumas outras. A Figura 4 ilustra a captação dos sinais para o termo “Desculpa”.

A implementação do Mediapipe no Librare ocorre quando o aplicativo acessa a câmera, ilustrado na Figura 5, e é desenhado em tempo real o mapeamento dos pontos detectados.

FIGURA 4. Captação do áudio na imagem à esquerda e a tradução do termo ‘Desculpa’ a direita.

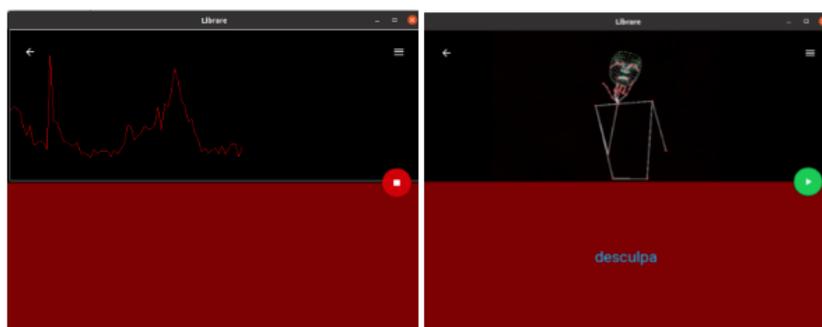
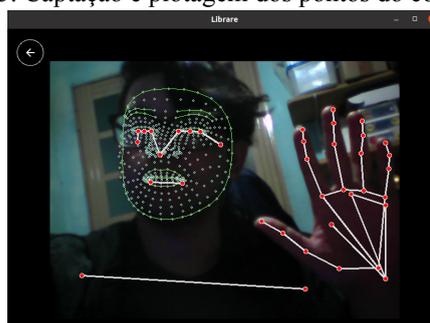


FIGURA 5. Captação e plotagem dos pontos do corpo humano.



CONCLUSÕES

O emprego dos conhecimentos em neurociência tem muito a contribuir para o desenvolvimento de tecnologias assistivas que sejam úteis à sociedade. Nesse sentido foi elaborado um software que atendeu as expectativas quanto à interpretação e tradução do som para sinais básicos de Libras. Um dos grandes desafios em fornecer acessibilidade para que os surdos consigam interagir com pessoas não falantes de Libras é encontrar bancos de dados com informações suficientes para manter no mínimo uma comunicação simples. Por isso, a implementação desses bancos é algo de extrema importância para que essa acessibilidade de fato ocorra. Assim, o presente trabalho trouxe resultados expressivos quanto a tradução da voz para alguns sinais em Libras. A próxima etapa visa avançar no desenvolvimento da conversão de sinais para a fala para a criação de um banco de dados e possivelmente a ampliação massiva dos sinais passivos de tradução.

AGRADECIMENTOS

Agradecemos ao CNPq pela bolsa de iniciação científica do Programa PIBIC-EM e também o apoio da coordenadora Célia Kawabata.

REFERÊNCIAS

- ALENCAR, C. D. C. **What and How the Deaf Brain Sees.** *Journal of Cognitive Neuroscience*. 2019.
- ANDRADE, Ana Paula. **O que é Kivy?**. Disponível em: <<https://www.treinaweb.com.br/blog/o-que-e-kivy>> Acesso em: 13 out. 2021
- HAYS, M & MULLEN, Tyler. **MediaPipe on the Web.** Disponível em: <<https://developers.googleblog.com/2020/01/mediapipe-on-web.html>> Acesso em: 12 ago. 2021
- LAYTON, Natasha; BORG, Johan. **Global perspectives on assistive technology.** WHO, ago, 2019.
- NASCIMENTO, J. L. **Introdução ao Processamento de Linguagem Natural usando Python.** III Escola Regional de Informática do Piauí. Livro Anais - Artigos e Minicursos, v. 1, n. 1, p. 336-360, jun, 2017.
- ROCHA, C. & MELGAÇO, S.C. **O uso de aplicativos para tradução de Libras.** V Simpósio Internacional de Inovação em Mídias Interativas, mai, 2018.
- TENÓRIO, Ronaldo. **Hand Talk.** Disponível em :<<https://www.handtalk.me/br/Aplicativo>> Acesso em: 18 set. 2021