



Utilização de Avatar Virtual na Tradução de Termos Técnicos de Informática para Língua Brasileira de Sinais

Jamile Hassen Sá
Unioeste

Foz do Iguaçu, Brasil
jamil.sa@unioeste.br

Claudio Roberto Marquetto Mauricio
Unioeste

Foz do Iguaçu, Brasil
claudio.mauricio@unioeste.br

Fabiana Frata Furlan Peres
Unioeste

Foz do Iguaçu, Brasil
fabiana.peres@unioeste.br

Abstract—Brazilian Sign Language is the mother tongue of the deaf community in Brazil and the translation of technical computer terms into this language is a challenge due to the lack of direct equivalence between words and concepts. In this way, the areas of specific knowledge create a library of technical terms for better understanding. Given this scenario, this project under development aims to create a virtual avatar to help translate technical terms into Libras and promote the inclusion of deaf people in computer science teaching and the development of assistive technologies. Therefore, libraries for detecting human joints in videos will be used to identify the joints in the signage. The study will be experimental in nature and will have an interdisciplinary approach to knowledge in the areas of IT, computing, Portuguese and Libras. It is expected to contribute to the educational environment, mainly in the area of information technology, in the inclusion of deaf people in teaching information technology and the development of assistive technologies applied to processes alternative communication.

Keywords—Pounds; Detection algorithms; Human pose.

Resumo—A Língua brasileira de sinais é a língua materna da comunidade surda no Brasil e a tradução de termos técnicos de informática para essa língua é um desafio devido à falta de equivalência direta entre as palavras e conceitos. Dessa maneira, as áreas de conhecimentos específicos criam uma biblioteca de termos técnicos para a melhor compreensão. Diante deste cenário, este projeto em desenvolvimento tem por objetivo a criação de um avatar virtual para auxílio na tradução de termos técnicos para Libras e promover a inclusão dos surdos no ensino de informática e no desenvolvimento das tecnologias assistivas. Para tanto, serão utilizadas bibliotecas de detecção de articulações de humanos em vídeos para identificar as articulações na sinalização. O estudo será de natureza experimental e terá uma abordagem interdisciplinar de conhecimentos das áreas da informática, computação, português e Libras. Espera-se contribuir para o ambiente educacional, principalmente na área de informática, na inclusão dos surdos no ensino de informática e no desenvolvimento das tecnologias assistivas aplicadas aos processos de comunicação alternativa.

Palavras-chave—Libras; Algoritmos de detecção; Pose humana.

I. INTRODUÇÃO

A acessibilidade comunicacional nos processos educativos é uma necessidade cada vez mais presente para pessoas surdas, além da inclusão na sociedade, que tem sido uma preocupação crescente. O uso da Língua brasileira de sinais (Libras) é essencial para garantir a comunicação efetiva e a participação plena desses indivíduos. No entanto, a tradução entre línguas oral-auditivas e línguas de sinais tem sido um desafio constante, principalmente em relação a conceitos específicos em cada área do conhecimento. Logo, a criação de um avatar online com capacidade de tradução terminológica pode ser uma solução promissora para superar essa barreira de comunicação. Visto isso, existem atualmente várias tecnologias assistivas que servem de apoio para os surdos se comunicarem com ouvintes por meio da tradução simultânea, como os aplicativos VLibras [1] e Handtalk [2]; no entanto, tais tecnologias ainda se baseiam na dicionarização de palavras e expressões. O desenvolvimento de avatares para a tradução em Libras geralmente é baseado em um sistema inteligente capaz de interpretar movimentos e gestos de braços e mãos. Dessa forma, pode-se verificar que existem algumas bibliotecas comumente usadas nessa área, sendo elas: o OpenPose [3] e o *Open Source Computer Vision* (OpenCV) [4]. O OpenPose é Open Source, ou seja, de código aberto, e utiliza técnicas de aprendizado profundo para rastrear e detectar articulações humanas em imagens ou vídeos, sendo apto na identificação e localização das principais partes do corpo, como mãos, braços, tronco e cabeça, e estima suas posições em um espaço tridimensional. Nesse contexto, a biblioteca citada possibilita alcançar informações detalhadas sobre gestos e movimentações realizadas, formando uma base para a interpretação da língua de sinais.

Por sua vez, o OpenCV é mais conhecido no ambiente de visão computacional, pois fornece ferramentas para processar



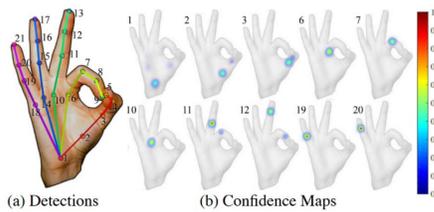


Fig. 1. Exemplificação da utilidade do OpenPose [5].

imagens e vídeos. Oferece funcionalidades mais específicas, como detecção de bordas, segmentação de objetos, reconhecimento de padrões e rastreamento de objetos em movimento. Dito isso, observa-se que o OpenPose e o OpenCV são frequentemente utilizados em conjunto, o que promove um melhor processamento de informações capturadas pelas articulações detectadas, permitindo a identificação e o reconhecimento dos gestos específicos da língua em tradução. Em suma, ressalta-se que a criação de um avatar nesse âmbito é um desafio complexo que envolve a coleta e o processamento de dados precisos, o desenvolvimento de algoritmos de reconhecimento de gestos, a criação de interfaces acessíveis e intuitivas, além da capacidade de assimilação do glossário de termos técnicos em Libras [6] que foi creditada pela Associação dos Surdos do Maranhão — ASMA [7]. Essas tecnologias têm o potencial de facilitar a tradução de termos técnicos de informática para Libras, proporcionando uma interface mais acessível e eficaz para a comunicação entre esses dois domínios.

II. MATERIAIS E MÉTODOS

Para desenvolver este projeto esta sendo utilizado um computador de alto desempenho, com processador I7 de 12ª geração, com 16 GB de RAM e placa de vídeo RTX 3070, disponibilizado pelo laboratório do grupo de pesquisa, onde o projeto esta em andamento.

A. Detecção e treinamento

Em primeira instância, foi feita uma avaliação das bibliotecas Open Source disponíveis na área de detecção de articulações humanas, como o OpenCV e o OpenPose, para determinar qual é a mais adequada entre os estudos criteriosos de precisão, facilidade de integração e eficiência. Visto isso, foi pesquisado o Framework MediaPipe, esse conjunto de bibliotecas foi criado pela Google, de código aberto, que identifica e detecta elementos em vídeos e imagens, como é visto na Figura 2, além de estar disponível na linguagem Python [8] para sua configuração. Adjunto a esse Framework, foi utilizada uma biblioteca de código aberto TensorFlow para aprendizado de máquina, sendo uma das principais ferramentas para *Deep Learning* [9] e *Machine Learning* [10], com a finalidade de

criar uma aplicação de acordo com o que será analisado nos modelos de aprendizagem de máquina.

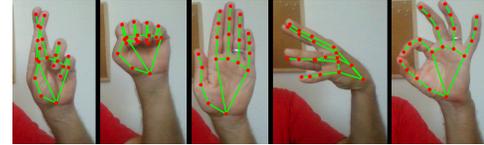


Fig. 2. Mediapipe Hands [11].

Dessa maneira, foi utilizado um site como suporte na identificação de imagens e vídeos para criar os modelos de aprendizagem de máquina, o Teachable Machine [12], onde é possível criar projetos com base em imagens, áudios ou poses. No projeto, está sendo analisado o projeto de imagem, que é capaz de sinalizar sinais da área de informática, como o símbolo do Windows classificado no glossário de termos técnicos em Libras. Dessa forma, a classe é renomeada como "Windows", são coletadas centenas de amostras de imagens gravadas, feitas de várias posições diferentes, e, por fim, é verificada a porcentagem de reconhecimento do sinal mostrado, como se vê na Figura 3.



Output



Fig. 3. Teachable Machine [12].

Em suma, foi abordada a utilização do Jupyter Notebook [13], como uma ferramenta capaz de compilar códigos de diferentes linguagens de programação, permitindo o acesso local ou online. Foi instalado a partir do Anaconda [14], que auxilia na simplificação do gerenciamento e implantação de pacotes com utilidade para a linguagem Python e R. Essas ferramentas estão sendo utilizadas para maior facilidade na

construção do projeto, pois ajudam na conexão do modelo com seu significado.

B. Modelagem 3D

À primeira vista, foi abordada a utilização do Blender [15], uma plataforma gratuita, de fácil entendimento e Open Source, capaz de fazer modelagem, animação, edição de vídeos, etc. Assim, está sendo utilizado na criação do avatar 3D, onde serão vistos as mãos e o tronco em estilo personificado, como exemplificado na Figura 4. Por fim, o avatar será animado para realizar os sinais prescritos pela modelagem de amostras.



Fig. 4. Exemplificação do avatar.

C. Coleta de dados e testes

No início, é necessário realizar o levantamento de termos para avaliar a modelagem e validação do desempenho do avatar. Isso inclui a gravação de vídeos realizando gestos específicos, a definição do escopo de termos técnicos que serão utilizados para a tradução em Libras, bem como a associação de cada termo com os gestos correspondentes. Posteriormente, é importante realizar o teste do protótipo com usuários proficientes em Libras, observar os resultados do protótipo, considerando a precisão e exatidão dos gestos e tradução, e identificar as correções, melhorias e pontos fracos do protótipo. É essencial registrar as etapas de pesquisa e testes, fazendo um detalhamento dos métodos utilizados, análises realizadas e conclusões.

III. OBJETIVOS

O objetivo principal do projeto é criar uma solução utilizando um avatar para sinalizar termos técnicos da informática utilizando a Língua Brasileira de Sinais, utilizando tecnologias assistivas. Por consequência, é necessário identificar objetivos específicos para se chegar à resolução principal do projeto,

como analisar e investigar o estado em que se encontram as técnicas de detecção de articulações humanas e a tradução de gestos do glossário citado anteriormente, comparar e pesquisar diferentes bibliotecas e ferramentas, como o OpenPose e o OpenCV, para examinar quais entre as estudadas melhor se adequam ao objetivo geral do projeto, realizar a identificação das articulações utilizadas nos sinais dos termos técnicos realizados por usuários proficientes em Libras para preparar o avatar virtual na tradução de termos sinalizados, e desenvolver um protótipo de avatar para sinalizar os termos técnicos da informática em Libras.

IV. CONCLUSÃO

A acessibilidade comunicacional para pessoas surdas é uma necessidade crescente, tanto para a inclusão na sociedade quanto para os processos educativos. A utilização da Língua Brasileira de Sinais (Libras) desempenha um papel fundamental na garantia da comunicação eficaz e na participação plena desses indivíduos. No entanto, a tradução entre línguas oral-auditivas e línguas de sinais tem sido um desafio persistente, especialmente quando se trata de conceitos específicos em áreas do conhecimento.

A proposta de criar um avatar online com capacidade de tradução terminológica emerge como uma solução promissora para superar essa barreira de comunicação. Atualmente, existem várias tecnologias assistivas, como os aplicativos VLibras e Handtalk, que oferecem tradução simultânea, mas ainda dependem da dicionarização de palavras e expressões.

O desenvolvimento de avatares para tradução em Libras geralmente se baseia em sistemas inteligentes capazes de interpretar movimentos e gestos de braços e mãos. Isso envolve o uso de bibliotecas como o OpenPose que utiliza técnicas de aprendizado profundo para rastrear e detectar articulações humanas, e o OpenCV, que é amplamente conhecido na visão computacional e fornece ferramentas para processar imagens e vídeos. A combinação dessas bibliotecas permite um processamento mais preciso das informações capturadas pelas articulações, facilitando a identificação e o reconhecimento dos gestos específicos da língua em tradução.

Em resumo, a criação de um avatar nesse contexto é um desafio complexo, envolvendo a coleta e o processamento de dados precisos, o desenvolvimento de algoritmos de reconhecimento de gestos, a criação de interfaces acessíveis e intuitivas, e a assimilação do glossário de termos técnicos em Libras. Essas tecnologias têm o potencial de facilitar a tradução de termos técnicos, proporcionando uma interface mais acessível e eficaz para a comunicação entre os domínios da informática e da Libras, contribuindo assim para uma educação mais inclusiva e uma participação plena na sociedade para as pessoas surdas.



AGRADECIMENTOS

Gostaria de agradecer ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ) pela concessão da bolsa de Iniciação Científica que permitiu o apoio financeiro a este projeto.

Agradeço também o Fundo Paraná pela concessão de equipamentos, os integrantes do Laboratório LABCG - Laboratório de Computação Gráfica pelo apoio e auxílio na realização deste projeto.

REFERÊNCIAS

- [1] G. F. Brasileiro, “Vlibras,” 2020. [Online]. Available: <https://www.gov.br/governodigital/pt-br/vlibras>
- [2] e. a. HandTalk, “Handtalk.me,” 2012. [Online]. Available: <https://www.handtalk.me/br/>
- [3] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, “Realtime multi-person 2d pose estimation using part affinity fields,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 7291–7299.
- [4] A. Kaehler and G. Bradski, *Learning OpenCV 3: computer vision in C++ with the OpenCV library*. ” O’Reilly Media, Inc.”, 2016.
- [5] Junxnone, “Hand keypoint detection,” 2019. [Online]. Available: <https://github.com/junxnone/examples/issues/164>
- [6] SENAI, “Curso técnico em informática. (2016). glossário de termos técnicos em libras,” 2016. [Online]. Available: <https://www.portaldaindustria.com.br/publicacoes/2016/8/glossario-de-termos-tecnicos-em-libras-informatica/>
- [7] A. dos Surdos do Maranhão, “Associação dos surdos do maranhão.” [Online]. Available: <https://asmaranhao.com.br/index/>
- [8] Alura, “Visão computacional: Hand tracking com opencv.” [Online]. Available: <https://www.alura.com.br/conteudo/visao-computacional-hand-tracking-opencv#:~:text=O%20MediaPipe%20C3%A9%20um%20framework,dispon%C3%ADvel%20para%20a%20linguagem%20Python.>
- [9] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [10] Z.-H. Zhou, *Machine learning*. Springer Nature, 2021.
- [11] LucianoSphere, “Exquisite hand and finger tracking in web browsers with mediapipe’s machine learning models,” 2021. [Online]. Available: <https://tinyurl.com/bdnakpkh>
- [12] MatthewStanciu, “Teachable machine.” [Online]. Available: <https://brasil.hackclub.com/workshops/teachable-machine/#:~:text=O%20Teachable%20Machine%20C3%A9%20um.aprendizado%20de%20m%C3%A1quina%20sem%20programa%C3%A7%C3%A3o.>
- [13] Alura, “Jupyter notebook: Exemplos de códigos e como usar.” [Online]. Available: <https://www.alura.com.br/artigos/conhecendo-o-jupyter-notebook>
- [14] T. Coutinho, “Como o anaconda ide pode ajudar na sua programação em python,” 2020. [Online]. Available: <https://www.voitto.com.br/blog/artigo/anaconda-python-ide>
- [15] —, “Conheça o blender, programa de sucesso na modelagem 3d,” 2021. [Online]. Available: <https://www.voitto.com.br/blog/artigo/conheca-o-blender>

