

Estudo para implementação de software de tradução automática para libras

Study for implementing an automatic translation software for libras

Vanina Carrara Sigrist

vanina.sigrist01@fatec.sp.gov.br
Fatec Rubens Lara

Anna Haydée de Lima Siqueira 

anna.siqueira@fatec.sp.gov.br
Fatec Rubens Lara

Danubia Pinheiro Cunha 

danubia.cunha@fatec.sp.gov.br
Fatec Rubens Lara

RESUMO

O objetivo deste artigo é identificar como a inteligência artificial junto com o reconhecimento óptico de caracteres (OCR) podem ajudar na tradução de documentos e livros escritos na língua portuguesa para LIBRAS. Por meio de artigos científicos e teses publicados sobre as duas tecnologias analisou-se a possibilidade de aplicação de ambas as tecnologias no processo de tradução, automatizando-a e tornando-a mais acessível para pessoas com deficiência auditiva, independentemente da língua de alfabetização. Para este artigo, realizou-se pesquisa teórica exploratória, com revisão bibliográfica, para entender a aplicação das tecnologias no cenário proposto, elaborando um protótipo tecnológico. Por fim, mediante o que foi evidenciado, chegou-se ao resultado da viabilidade de um aplicativo que utilize módulos separados de OCR e de uma rede neural interligados para realizar a tradução automática.

PALAVRAS-CHAVE: Acessibilidade. Inteligência Artificial. Reconhecimento Óptico de Caracteres. Rede neural artificial. Língua brasileira de sinais (LIBRAS).

ABSTRACT

This article aims to identify how artificial intelligence together with optical character recognition (OCR) can help in the translation of documents and books written in Portuguese into the Brazilian Sign Language. Through scientific articles and theses published on the two technologies, we analyzed the possibility of applying both in the translation process, automating it, and making it more accessible to the people who are Deaf or Hard of Hearing (DHH), whether they are literate in Portuguese or not. For this article, we performed exploratory theoretical research and a bibliographic review to present the application of these technologies in the proposed scenario, elaborating a technological prototype. Finally, based on what has been evidenced, we concluded the feasibility of an application that uses separate modules of OCR and a neural network interconnected to perform automatic translation.

KEY-WORDS: *Accessibility. Artificial intelligence. Optical character recognition (OCR). Machine translation. Artificial neural network. Brazilian Sign Language.*

INTRODUÇÃO

Segundo um estudo feito, em setembro de 2019 no Brasil, pelo Instituto Locomotiva em conjunto com a Semana da Acessibilidade Surda (GANDRA, 2019, s/p), existem 10,7 milhões de pessoas com deficiência auditiva no Brasil. Dentre eles, 2,3 milhões têm deficiência severa e, na sua maioria, são pessoas na faixa etária de 60 anos de idade para cima.

Nesse cenário, é importante ressaltar que, mesmo com a Lei Brasileira de Inclusão (nº 13.146, de 6 de julho de 2015), muitos ainda encontram diversas dificuldades no dia a dia, tendo um acesso limitado a oportunidades básicas como trabalho e educação. Ainda de acordo com esse estudo (GANDRA, 2019, s/p), do total de brasileiros com deficiência auditiva, somente 7% deles concluíram o ensino superior; 15% cursaram até o médio completo; 46%, têm fundamental completo; e 32%, sem instrução alguma (Quadro 1).

Quadro 1 – Nível de escolaridade entre pessoas com deficiência auditiva no Brasil

Nível de Escolaridade	Quantidade (%)
Ensino superior completo	7
Médio completo	15
Fundamental completo	46
Não possuem nenhuma instrução	32

Fonte: Gandra (2019)

No que concerne a legislação vigente, o governo brasileiro, no Decreto Federal Nº 5.296, de 2 de dezembro de 2004, definiu deficiência auditiva como a “perda bilateral, parcial ou total, de quarenta e um decibéis (dB) ou mais, aferida por audiograma nas frequências de 500Hz, 1.000Hz, 2.000Hz e 3.000Hz” (Brasil, 2004). E através do Decreto Federal Nº 5.626, de 22 de dezembro de 2005, que:

para os fins deste Decreto, considera-se pessoa surda aquela que, por ter perda auditiva, compreende e interage com o mundo por meio de experiências visuais, manifestando sua cultura principalmente pelo uso da Língua Brasileira de Sinais – Libras (Brasil, 2005).

O Art. 3º, deste mesmo decreto, tornou também obrigatória a disciplina de LIBRAS como componente curricular para professores em formação (Brasil, 2005). É importante ressaltar que a Língua Brasileira de Sinais foi oficializada no país pela Lei nº 10.436 de 2002 (Brasil, 2002). “Entretanto, a LIBRAS é pouco difundida nas escolas e na sociedade entre as

peças que não possuem deficiência, existindo assim, dificuldade por parte dos surdos de se comunicarem com a maioria da população e vice-versa.” (FURLAN, 2016, p. 33).

O Art. 8º do decreto federal de 2004 descreve o conceito de acessibilidade como sendo a

[...] condição para utilização, com segurança e autonomia, total ou assistida, dos espaços, mobiliários e equipamentos urbanos, das edificações, dos serviços de transporte e dos dispositivos, sistemas e meios de comunicação e informação, por pessoa portadora de deficiência ou com mobilidade reduzida (BRASIL, 2004).

Montardo e Passerino (2007) apud Furlan (2016, p. 34) afirmam que o objetivo da acessibilidade é a “inclusão social, a fim [...] de melhorar qualidade de vida, autonomia financeira, direitos iguais e oportunidades para pessoas ou grupos de pessoas que possuem [...] desvantagem em relação a outros indivíduos”. Furlan também declara que uma das formas de promover a acessibilidade é a tecnologia.

Mencionando o Comitê de Ajudas Técnicas da Coordenadoria Nacional para Integração da Pessoa Portadora de Deficiência (Corde), Ribeiro, Matos e Pimentel (2017) definem Tecnologia Assistiva (TA) como

[...] uma área do conhecimento, de característica interdisciplinar, que engloba produtos, recursos, metodologias, estratégias, práticas e serviços que objetivam promover a funcionalidade, relacionada à atividade e participação, de pessoas com deficiência, incapacidades ou mobilidade reduzida, visando sua autonomia, independência, qualidade de vida e inclusão social (BRASIL, 2007).

Já Santos (2018, p. 52) estabelece o significado de TA como “um termo novo, utilizado para identificar todo o tipo de recursos que, de alguma maneira, contribui para proporcionar vida autônoma ao portador de deficiência”. Além disso, ele cita Sartoretto e Bersch (2017) para explicar que “em seu sentido geral, pode-se entender que todos os artefatos usados por uma pessoa, em seu dia a dia, desde talheres, ferramentas, etc., são objetos de Tecnologia Assistiva”.

Tendo em vista esse panorama, para garantir a acessibilidade desses indivíduos, é necessária a criação de uma ou mais tecnologia assistivas. Nesse âmbito, verifica-se a possibilidade de aplicações tecnológicas que consigam traduzir as línguas maternas para a língua brasileira de sinais (LIBRAS), levando em conta as diferenças gramaticais e de semântica existentes entre esta e a língua portuguesa.

Na atualidade, já existem aplicativos que funcionam como tradutores, entretanto eles apenas traduzem um texto digitado na hora ou aquele que é ditado por áudio. Esses programas não permitem que um usuário escaneie uma imagem, tirada ou não na hora, e, a partir dela, possa ser feita a tradução do que está escrito nela, opção que é encontrada hoje em dia em

alguns apps de tradução de línguas naturais para dispositivos móveis, como o Google Tradutor e Naver Papago (tradutor da empresa coreana Naver).

Visando essa lacuna, o objetivo deste artigo é identificar como a inteligência artificial junto com o reconhecimento óptico de caracteres (OCR) podem ajudar na tradução de documentos e livros escritos na língua portuguesa para LIBRAS. Apresentar-se-á, então, uma base teórica para analisar a possibilidade da implementação de uma aplicação voltada para a tradução de textos escaneados, com o propósito de aumentar a inclusão de pessoas portadoras de deficiência auditiva, em especial no mercado de trabalho, e, assim, facilitar o manuseio de suas documentações por parte deles.

Utilizando o reconhecimento óptico de caracteres, que é uma tecnologia desenvolvida para converter diversos tipos de documentos em dados pesquisáveis e editáveis (VELAZ, 2018), serão reconhecidas as letras e palavras escritas nos papéis escaneados ou imagens. Tendo identificado as informações presentes no texto, através de uma rede neural artificial, ocorrerá a tradução da língua portuguesa para a língua brasileira de sinais, levando em conta as diferenças estruturais entre as duas.

Para este artigo, realizou-se uma pesquisa de natureza teórica com revisão bibliográfica com o propósito de analisar as tecnologias citadas acima, buscando demonstrar as aplicações delas por intermédio de artigos científicos e teses com fundamentos similares. Por fim, mediante os resultados encontrados, apresentaram-se maneiras e recursos para viabilizar a geração de um *software* com a finalidade de tornar mais acessíveis textos impressos para os brasileiros com deficiência auditiva.

1. EMBASAMENTO TEÓRICO

Atualmente, existem diversas tecnologias assistivas (TA) no mercado voltadas para pessoas com deficiência auditiva. Em uma pesquisa feita na *Google Play Store*, com dados até dia 20 de novembro de 2020, encontraram-se diversos aplicativos para a plataforma *Android*. Dentre eles, foram selecionados apenas três deles com a funcionalidade principal de serem tradutores de texto e áudio para LIBRAS, representando os sinais através de um avatar. São eles:

- **Hand Talk:** é uma plataforma que traduz simultaneamente conteúdos em português para a língua brasileira de sinais. Ele possui duas modalidades: o aplicativo para dispositivos móveis e um *plugin* para que empresas instalem em seus sites tornando-os acessíveis. Possui dois avatares: o Hugo e a Maya. Foi criada por Ronaldo Tenório, Carlos Wanderlan e Thadeu Luz e lançada em julho de 2013;
- **VLIBRAS:** conforme seu site oficial, é o:

resultado de uma parceria entre o Ministério da Economia (ME), por meio da Secretaria de Governo Digital (SGD) e a Universidade Federal da Paraíba (UFPB), [...] é um conjunto de ferramentas computacionais de código aberto, que traduz conteúdos digitais (texto, áudio e vídeo) para Língua Brasileira de Sinais - LIBRAS, tornando computadores, celulares e plataformas Web acessíveis para pessoas surdas (VLIBRAS, s/p).

- O aplicativo possui dois avatares: o Ícaro e a Hozana. Está disponível também para computadores com sistemas operacionais *Windows* e *Linux* e para dispositivos *IOS*, além de extensões para os navegadores *Chrome*, *Firefox* e *Safari* e como *widget* para sites;
- **Rybená:** uma TA brasileira desenvolvida por surdos que surgiu de uma parceria entre o Grupo de Usuários Java do Distrito Federal (DFJUG) e o Instituto CTS em 2003. Em seu site oficial é descrito como:

uma tecnologia assistiva que traduz textos do português para Libras e Voz. Assim, surdos e pessoas com deficiências intelectuais, disléxicos e outros com dificuldades de leitura podem consumir conteúdos e interagir com diversos sites e plataformas online (RYBENÁ, s/p).

Na tela principal dos aplicativos, aparece, em foco, seus respectivos avatares, com ícones que quando clicados é possível digitar em uma caixa ou ditar pelo microfone do celular o texto a ser representado em sinais. Com exceção do Rybená, eles também têm a possibilidade de trocar entre um avatar feminino e masculino, além de uma aba de dicionário, onde são encontrados diversos sinais e seus respectivos significados.

No *Hand Talk*, em especial, existe além disso, uma aba com vídeos ensinando sobre a língua de sinais e uma outra chamada “Loja” onde se tem a possibilidade de trocar as roupas dos avatares Hugo e Maya, sendo algumas grátis e outras pagas através de moedas, que podem ser compradas no app com custo de R\$9,90 a R\$ 139,90.

De acordo com Barbosa (2020, p.92) *apud* Nilsson (2009), inteligência artificial (IA, ou em inglês *AI – artificial intelligence*) é conjunto de técnicas para a construção de máquinas inteligentes, capazes de resolver problemas que requerem inteligência humana.

A autora configura que IA é um campo do conhecimento atualmente bastante explorado pelo cinema e pela literatura, mas ainda pouco conhecido no que se refere à sua gênese (2020, p.92). Para Barbosa (2020), o objetivo da inteligência artificial:

é entender e construir sistemas inteligentes, o que representa um elevado impacto em nossa cultura ocidental, uma vez que nela há crenças humanistas e especistas que nos levam a pensar que somos seres superiores e que inteligência e pensamento são dádivas exclusivas à nossa espécie - o que nos diferenciaria e nos torna superiores às demais criaturas (BARBOSA, 2020, p. 93).

Já Lobo (2018) define inteligência artificial como

um ramo da ciência da computação que se propõe a desenvolver sistemas que simulem a capacidade humana na percepção de um problema, identificando seus componentes e, com isso, resolver problemas e propor/tomar decisões. [...] IA envolve várias etapas ou competências como reconhecer padrões e imagens, entender linguagem aberta escrita e falada, perceber relações e nexos, seguir algoritmos de decisão propostos por especialistas, ser capaz de entender conceitos e não apenas processar dados, adquirir “raciocínios” pela capacidade de integrar novas experiências e, pois, se auto aperfeiçoar (“self learning”), resolvendo problemas, ou realizando tarefas (LOBO, 2018, p. 3).

Segundo Barbosa (2020), o marco-zero da inteligência artificial foi o ano de 1956. Ela, então, mencionando Russel e Norvig (2009), explica que nesse ano ocorreu a Conferência do Dartmouth College, em New Hampshire (USA), onde o termo “inteligência artificial” foi registrado pela primeira vez, referindo-se a um novo campo do conhecimento.

Barbosa (2020), entretanto, salienta que o desenvolvimento de ideias concernentes a essa área é anterior a 1956, remontando à Segunda Guerra Mundial. Ela relata que:

no que se refere à primeira produção bibliográfica correlata ao tema da IA, em 1943 Warren McCulloch e Walter Pitts escreveram um artigo sobre estruturas de raciocínio artificiais em forma de modelo matemático que imitam o sistema nervoso humano. Esse modelo matemático deu base para diversas outras formulações acadêmicas sobre o tema. (BARBOSA, 2020, p. 93).

As técnicas de IA necessitam de uma grande quantidade de conhecimentos e de mecanismos de manipulação de símbolos para que os computadores se tornem capazes de pensar e tomar decisões. Eles devem ter a possibilidade de representação, modificação e ampliação de suas capacidades de interpretação e de análise de dados (SANTOS, 2018, p. 52).

Voigt (2018) configura que Redes Neurais Artificiais (RNAs, ou em inglês ANNs – *Artificial neural networks*):

são sistemas computacionais vagamente inspirados na estrutura das Redes Neurais Biológicas de animais. Na maioria dos casos é um sistema de aprendizagem supervisionada, capaz de aprender determinada tarefa mesmo sem que sejam programados detalhes específicos referentes à ela (VOIGT, 2018, p. 20).

Ele continua explicando que:

é formada por um conjunto de unidades, representando neurônios, que possuem conexões uns com os outros, representando sinapses. Cada neurônio então recebe um determinado sinal como entrada, efetua um determinado processamento no mesmo, e o repassa para os neurônios adiante (VOIGT, 2018, p. 20).

Para Voigt (2018, p. 21), um grande atrativo das Redes Neurais é a capacidade de reduzir problemas não-lineares para outros problemas menores e mais simplificados, e então resolvê-los de forma mais eficiente.

O Reconhecimento Óptico de Caracteres ou *Optical Character Recognition* (OCR) é uma metodologia computacional capaz de extrair de forma completamente automática o texto (impresso ou manuscrito) contido em uma imagem digital (OLIVEIRA; PITERI; ARTERO; SILVA; 2014, p. 1). O OCR transforma através da conversão os tipos diferentes de documentos digitalizados em dados pesquisáveis e editáveis (IFRS, 2018, s/p).

De acordo com Mithe, Indalkar e Divekar (2013), o reconhecimento óptico de caracteres é uma tecnologia que funciona como a capacidade humana de leitura, mas que ainda não consegue competir com ela. Eles ressaltam também que, o OCR:

[...] consegue reconhecer tanto texto escrito a mão quanto impresso. Mas a performance do OCR está diretamente dependente na qualidade de entrada do documento. OCR é designado para processar imagens que consistem quase inteiramente de texto, com pouquíssimo lixo não-textual obtido de uma imagem capturada por uma câmera móvel. (MITHE; INDALKAR; DIVEKAR, 2013, p. 1, tradução nossa).¹

Conforme relatado por Wanzeller (2018), o OCR:

[...] segundo relatos, já existia e era praticado há vários anos; contudo, só se tornou objeto de pesquisas e registros em 1950, quando David Shepard e Louis Tordella pesquisavam uma forma de automatizar dados da Agência de Segurança das Forças Armadas que, em alguns anos, se tornaria a Agência de Segurança Nacional Americana (NSA) (HEITLINGER, 2007, p. 43). Auxiliados por Harvey Cook, Shepard e Tordella criaram o primeiro software de OCR, conhecido como Gismo e, após esse passo inicial, Shepard fundou a IMR, empresa que se tornou pioneira no desenvolvimento de *software* comercial de OCR (Wanzeller, 2018, p. 15).

O termo *Optical Character Recognition* e a sigla OCR surgiram após a compra da licença da empresa pela IBM em 1953, 3 anos após a criação do primeiro software.

¹ No original: *can recognize both handwritten and printed text. But the performance of OCR is directly dependant on quality of input documents. OCR is designed to process images that consist almost entirely of text, with very little non-text clutter obtain from picture captured by mobile camera.*

2. RESULTADOS E DISCUSSÕES

Embora existam diversos estudos científicos realizados em reconhecimento óptico de caracteres ou inteligência artificial e seus ramos, em especial redes neurais artificiais, pouquíssimos do que se encontra teve um foco próximo de uma tradução automática de língua de sinais para uma língua natural por meio de um texto.

Romero (2017) buscou tratar da extração de texto em imagens complexas de documentos, especificamente os de identificação, com o objetivo de ler e reconhecer o conteúdo deles.

O autor descreve que, atualmente, os sistemas de OCR possuem altas taxas de reconhecimento de caracteres (95 a 99%) para as imagens simples caracterizadas como documentos, como um trecho de um artigo ou uma notícia de jornal; porém esses tais sistemas são incapazes de reconhecer em imagens complexas, que em sua maioria, são coloridas e de baixa resolução (Romero, 2017, p. 30).

Para o processamento de imagem, Romero (2017, p. 43-44) abordou diversas técnicas, convertendo inicialmente uma imagem colorida para tons de cinza para reduzir a quantidade de informações presentes. Segundo o mesmo estudo, o pré-processamento procura corrigir os defeitos na imagem através de algoritmos eficientes.

O autor configura que:

as técnicas de melhoria da qualidade de imagens podem ser divididas em duas famílias: as de realce e as de restauração de imagens. Quando a melhoria é usada para combater um processo de degradação conhecido ou avaliado por métodos da teoria da filtragem, a palavra usada é restauração de imagens. A restauração difere de realce pelo fato de que a primeira procura obter a imagem “real” tendo, se possível, um conhecimento a priori da degradação (ROMERO, 2017, p. 44).

Ele destaca a técnica MSER (*Maximally Stable Extremal Regions* – regiões extremas maximamente estáveis) que é usada como um método de detecção de componentes em imagens. O algoritmo é capaz de localizar com precisão texto em diferentes tamanhos, estilos e cores, independente de perspectiva e rotação, portanto, é amplamente usado na localização de texto (Romero, 2017, p. 62-65).

Feito o processamento de imagem, entra a etapa de reconhecimento, onde:

[...] a ideia principal do OCR é o reconhecimento de um caractere em uma determinada imagem digitalizada. Para isso, as características desse caractere são comparadas com características de padrões de um determinado alfabeto. Após o reconhecimento de cada um dos caracteres, deve-se fazer a associação entre eles com a finalidade de se ter uma sequência de caracteres que tenha algum significado, ou seja, eles devem ser agrupados de modo a formar palavras (ROMERO, 2017, p. 66).

O autor, então, ressalta a importância de considerar a ocorrência de erros nessa etapa de classificação, sugerindo uma atividade de detecção e correção dos erros.

Utilizando vários tipos de imagens e banco de imagens, Romero (2017) testou o sistema proposto. Ele efetuou testes da etapa de seleção, reconhecimento e o sistema geral usando as métricas de revocação e precisão.

Os resultados obtidos apontaram alguns erros pelas imagens de texto apresentarem uma grande diversidade de fontes, estilos, tamanhos e orientações. Entretanto, os algoritmos propostos pelo autor tiveram resultados satisfatórios nas métricas empregadas, apresentando, no sistema geral, uma precisão de 89%.

Romero (2017, p. 135) conclui que o algoritmo é bastante robusto, tendo selecionado 84% do texto existente nas imagens na etapa de seleção, no qual 94% são corretamente classificados. E na etapa de reconhecimento, reconheceu 88% do texto existente nas imagens, onde foi classificado corretamente 96% deles.

Para a tradução de língua natural textual para língua de sinais brasileira, analisou-se o trabalho de conclusão de curso no qual Gaio (2020) realizou o processo inverso, utilizando uma luva sensorial para reconhecer os sinais efetuados por seu usuário. O autor estudou LIBRAS profundamente para poder buscar as alternativas que contemplassem as suas características morfológicas.

Conforme cita seu estudo:

[...] como LIBRAS não é composta apenas por sinais de mão, para aumentar a compatibilidade de um tradutor com a língua em questão, notou-se a necessidade de um sistema de reconhecimento facial em conjunto com um reconhecimento dos referidos gestos (GAIO, 2020, p. 45).

Para a programação do sistema, como o aprendizado por máquina é dependente de dados, o autor (2020, p. 45) buscou por *datasets*, que são conjuntos de dados mais específicos, que melhor representassem as expressões faciais e a composição de sinais com as mãos de LIBRAS.

Obtido os dados, Gaio (2020) explicou que:

[...] foi necessário criar um modelo com base nos dados obtidos, para então executar uma aplicação que, em tempo real, classificasse os movimentos executados pelo usuário que utiliza a luva. Foi feita uma sequência de testes, com diferentes algoritmos de aprendizado de máquina, e, em seguida, executado o “ajuste fino” de parâmetros para cada um deles. [...] Posteriormente, foi selecionado o modelo que tivesse o melhor desempenho dentre os testados (GAIO, 2020, p. 59).

Através de uma aplicação *Python*, efetuou-se a execução dos modelos de classificações. De todos os testados, Gaio (2020, p. 64-65) constatou que o classificador do *XGBoost* obteve o melhor rendimento, com base na acurácia em relação ao conjunto de teste.

O autor constata que o:

[..] modelo final encontrado foi capaz de ter uma taxa de acerto muito interessante, de 96,8%, que de fato se reflete na realidade dos testes da luva em modo de tradução. Quando os sinais são realizados, o sistema consegue obter uma alta precisão de acerto (GAIO, 2020, p.64).

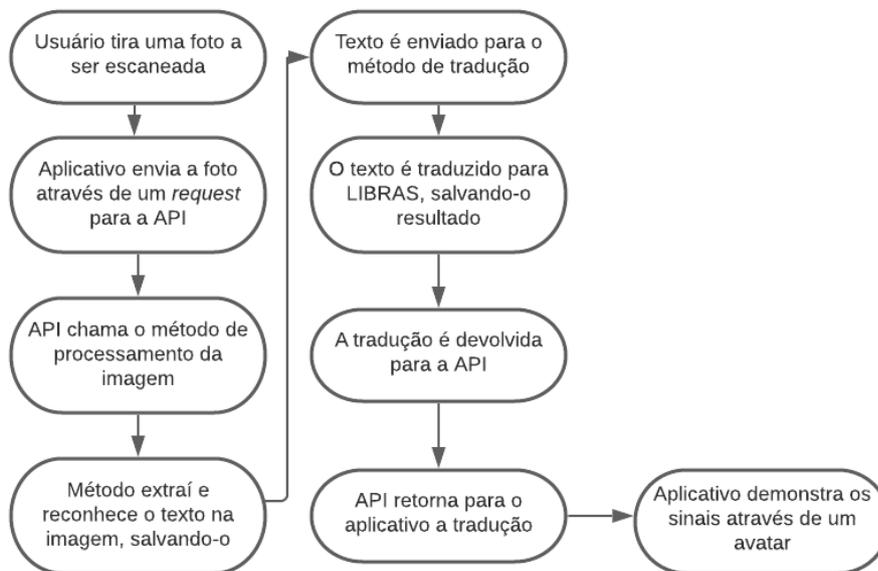
Desse modo, ele conclui que é importante ter os dados específicos deste nicho, com todos os tipos de intensidade e variação de expressões faciais.

Entretanto, é imprescindível reiterar que o trabalho de Gaio, analisado previamente, aborda o processo inverso do objetivo deste artigo. Para se chegar mais próximo, o necessário seria aplicar o reconhecimento dos sinais de forma contrária, possibilitando assim retornar o sinal equivalente a palavra que foi extraída.

Dessa maneira, leva-se em consideração a possibilidade da implementação de um software para a tradução automática para LIBRAS, que permita que o usuário escaneie um texto impresso e visualize a tradução para a língua de sinais através de um avatar, método esse parecido ao que ocorre no *Google Tradutor* e *Naver Papago*.

O processo de funcionamento do *software* de tradução envolveria o aplicativo e/ou site desenvolvido onde o usuário acessa e faz o upload da imagem a ser escaneada, que pode ser tirada na hora ou uma foto já existente. A partir disso, o aplicativo efetua *requests* para uma API desenvolvida, que chama o *back-end* do sistema, o chamado *server-side*, realizaria através de métodos o processamento da imagem, extraindo e reconhecendo o texto presente nela, traduzindo-o depois para LIBRAS.

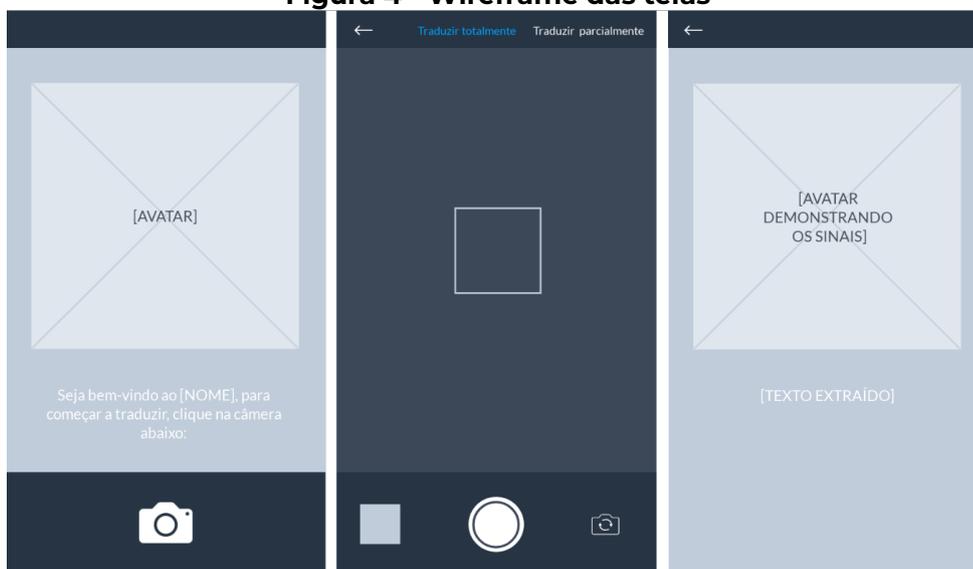
Com a tradução realizada, a API retorna para o aplicativo os sinais correspondentes que deverão ser demonstrados através de um avatar.

Figura 3 - Processo de funcionamento do software de tradução

Fonte: Elaborada pelos autores (2021)

Propôs-se, inicialmente, uma aplicação desenvolvida tanto para plataformas *mobile* quanto para uma plataforma *web*, podendo ser acessada por diversos sistemas operacionais, como *Windows*, *IOS*, *Android* e *Linux*. Escolheram-se ambas as plataformas por serem amplamente utilizadas e possuírem diversas bibliotecas e *frameworks* grátis.

De início, haveria 3 telas: uma inicial apresentando o avatar e com um botão para abrir a câmera; a tela da câmera com a opção de tirar a foto na hora ou escolher uma existente, podendo escolher se quer traduzir totalmente ou parcialmente o texto, selecionando o trecho desejado; e a terceira e última, onde apresentaria o avatar demonstrando os sinais e com o texto extraído em baixo na mesma tela.

Figura 4 - Wireframe das telas

Fonte: Elaborada pelos autores (2020)

No que se refere as tecnologias para o desenvolvimento, para o módulo de extração e reconhecimento de texto, a linguagem Python oferece o *wrapper pytesseract* que permite a compatibilidade com a API *libtesseract* do Tesseract OCR, um *software* de reconhecimento óptico que foi desenvolvido pela *Hewlett-Packard* (HP) e é, atualmente, hospedado pelo *Github*. O *Tesseract* tem suporte ao Unicode (UTF-8) e consegue reconhecer mais de 100 linguagens, além de retornar em diversos formatos os *outputs*, como texto simples, PDF, TSV e hOCR (HTML).

Já para a tradução automática, a linguagem *Python* também possibilita o manuseio de redes neurais, podendo ser aplicado em *machine learning* e outras áreas da inteligência artificial. Ressalta-se que o VLIBRAS, outro tradutor já existente no mercado, emprega a mesma linguagem, o que reforça a possibilidade de sucesso da implementação da tradução em Python.

Para a API que conectará o *front-end* do *app* com o módulo de extração e tradução, pensou-se na própria linguagem *Python*, mencionada anteriormente, ou então linguagens como *Java* e *C#*, ambas bastante utilizadas para a criação de APIs.

A fim de armazenar as informações, citam-se as plataformas de *database* MongoDB ou *Firebase*, se decidido criar um banco NoSQL, ou o *SQL Server*, para bancos que possuem relacionamento entre suas tabelas.

A página a ser acessada pelo usuário seria criada através de arquivos HTML (*HyperText Markup Language*), estilizadas por arquivos CSS (*Cascading Style Sheets*) e arquivos Javascript, sendo feita seguindo os padrões do W3C e a cartilha E-MAG do governo brasileiro para torná-la acessível. Além disso, o uso de bibliotecas torna-se possível através da plataforma *Node.js*, podendo-se usar *frameworks* como *Angular* e *React* para gerar a estruturação das páginas.

Por fim, para dispositivos móveis, o Apache Cordova permite criar aplicações híbridas para diferentes plataformas mobile, que utiliza HTML, CSS e *Javascript*, renderizando-os com base no *WebView*.

3. CONSIDERAÇÕES FINAIS

Mediante os trabalhos estudados, especialmente os dois citados neste artigo de Romero (2017) e Gaio (2020), pode-se estudar bases teóricas necessárias para o planejamento da

implementação de um protótipo que era o objetivo deste artigo, ou seja, desenvolver um sistema para uma tradução automática de texto reconhecido em imagens para LIBRAS.

Assim, sabendo que já existem aplicativos que funcionem como tradutores de LIBRAS no mercado, eles, no entanto, não possuem como funcionalidade a opção de o usuário tirar uma foto e/ou fazer *upload* de uma imagem, onde o conteúdo dela seria escaneado e traduzido para a língua de sinais.

Em virtude disso, concluiu-se que a possibilidade se encontra no desenvolvimento de um *software* de plataforma híbrida com módulos que integrem as tecnologias tratadas anteriormente onde um método receba a imagem a ser escaneada, processe-a e, assim, possa extrair o texto presente nela enviando-o para o outro módulo que realize a tradução do texto para a língua de sinais. Por fim, o sistema mostraria para o usuário, por meio de um avatar, o resultado obtido.

Porém, por se tratar de um artigo de estudo inicial, é indispensável enfatizar que ainda assim é preciso aprofundar conhecimentos teóricos e tecnológicos sobre os assuntos abordados. Com a falta de estudos científicos com o presente foco, acabou-se dando enfoque na extração e reconhecimento dos textos, o que fez com que alguns pontos não pudessem ser abordados em suas totalidades.

Assim, deixa-se como sugestão de pesquisa para trabalhos futuros a implementação do avatar para demonstrar os sinais de LIBRAS no aplicativo e uma possível melhoria no processo de tradução, levando em conta a semântica e complexidade das duas línguas, para que o resultado seja ainda mais preciso.

Por afetar milhões de pessoas todos os dias, é essencial também efetuar uma pesquisa de campo com o público-alvo para que seja possível verificar quais pontos devem ser abordados e realmente desenvolver uma tecnologia assistiva que possa ser empregada no dia a dia, assim ajudando as pessoas com deficiência auditiva a se comunicarem de forma mais clara seja no mercado de trabalho ou em qualquer outra situação de interação comunicativa.

A área de tradução automática ainda tem um campo amplo para ser explorado e desenvolvido até que se consiga construir um software capaz de realizar esse processo de forma mais dinâmica, que considere todas as nuances das línguas naturais, que continuamente se modificam ao longo dos anos.

REFERÊNCIAS

BARBOSA, Xênia de Castro. **Breve introdução à história da inteligência artificial.** *Jamaxi*, v. 4, n. 1, 2020.

BERSCH, Rita. **Introdução à tecnologia assistiva.** Porto Alegre, 2013.

BRASIL. Lei n. 13.146, de 6 de julho de 2015. **Institui a lei brasileira de inclusão da pessoa com deficiência (estatuto da pessoa com deficiência).** Diário oficial da união. Brasília, 7 jul. 2015. Seção 1, p. 2.

BRASIL. Decreto Federal n. 5.296, de 2 de dezembro de 2004. **Regulamenta as leis nºs 10.048, de 8 de novembro de 2000, que dá prioridade de atendimento às pessoas que especifica, e 10.098, de 19 de dezembro de 2000, que estabelece normas gerais e critérios básicos para a promoção da acessibilidade das pessoas portadoras de deficiência ou com mobilidade reduzida, e das outras providências.** Diário oficial da união. Brasília, 3 dez. 2004. Seção 1, p. 5.

BRASIL. Decreto Federal n. 5.626, de 22 de dezembro de 2005. Regulamenta a lei nº 10.436, de 24 de abril de 2002, **que dispõe sobre a língua brasileira de sinais - libras**, e o art. 18 da lei nº 10.098, de 19 de dezembro de 2000. Diário oficial da união. Brasília, 3 dez. 2005. Seção 1, p. 28.

BRASIL. Secretaria Especial dos Direitos Humanos. Coordenadoria Nacional para Integração da Pessoa Portadora de Deficiência. **Ata da 7ª Reunião do Comitê de Ajudas Técnicas.** Brasília, 2007. Disponível em: https://www.assistiva.com.br/Ata_VII_Reuni%C3%A3o_do_Comite_de_Ajudas_T%C3%A9cnicas.pdf. Acesso em: 20 nov. 2020.

DGE. DEPARTAMENTO DE GOVERNO ELETRÔNICO. 2014. **eMAG - Modelo de Acessibilidade em Governo Eletrônico.** Disponível em: <http://emag.governoeletronico.gov.br/>. Acesso em: 27 mar. 2021.

FURLAN, Andreson Luís. **Desenvolvimento de um protótipo de aplicativo móvel para conversão de voz em texto e texto em voz, orientado ao apoio à comunicação de deficientes auditivos.** 2016.

GAIO, Rennan de Lucena. **STILL: sistema tradutor inteligente de LIBRAS com luva.** 2020.

GANDRA, Alana. **País tem 10,7 milhões de pessoas com deficiência auditiva, diz estudo.** Agência Brasil, Rio de Janeiro, 13 de out. de 2019. Disponível em: <https://agenciabrasil.ebc.com.br/geral/noticia/2019-10/brasil-tem-107-milhoes-de-deficientes-auditivos-diz-estudo>. Acesso em: 5 jun. 2020.

GITHUB. **tesseract-ocr / tesseract: Tesseract Open Source OCR Engine.** Disponível em: <https://github.com/tesseract-ocr/tesseract>. Acesso em: 27 mar. 2021.

HAND TALK. **Sobre o Hand Talk – Acessibilidade em Libras.** Disponível em: <https://handtalk.me/>. Acesso em: 5 jun. 2020.

IFRS. **Ferramentas OCR – entenda o que são e sua relação com a acessibilidade.** Instituto Federal Rio Grande do Sul, Rio Grande do Sul, 18 de dez. de 2018. Disponível em: <https://cta.ifrs.edu.br/>. Acesso em: 20 nov. 2020.

LOBO, Luiz Carlos. **Inteligência artificial, o Futuro da Medicina e a Educação Médica.** Rev. bras. educ. med., Brasília, v. 42, n. 3, p. 3-8, set. 2018. Disponível em: <http://www.scielo.br/>. Acesso em: 20 nov. 2020.

MENEZES, João Victor; DE PAIVA GUIMARÃES, Marcelo; MARTINS, Valéria Farinazzo. **Inclusão de Pessoas com Deficiências Auditivas por meio do uso de um Sistema para Geração Automática de Legendas em Tempo Real em Apresentações Orais e de Material de Apoio.** Revista Ibérica de Sistemas e Tecnologias de Informação, n. E27, p. 515-527, 2020.

MITHE, Ravina; INDALKAR, Supriya; DIVEKAR, Nilam. **Optical character recognition.** International journal of recent technology and engineering (IJRTE), v. 2, n. 1, p. 72-75, 2013.

MONTARDO, S.P.; PASSERINO, L. **Inclusão social via acessibilidade digital: proposta de inclusão digital para Pessoas com Necessidades Especiais (PNE).** E-Compós, v.8, 2007.

OLIVEIRA, Helder; PITERI, Marco; ARTERO, Almir; SILVA, Francisco A. **Desenvolvimento de uma Plataforma de Software para o Reconhecimento Óptico de Caracteres.** 2014.

RIBEIRO, Sátilla Souza; MATOS, Aline Pereira Santos; PIMENTEL, Susana Couto. **Inclusão de estudantes surdos no ensino superior: contribuições dos recursos de tecnologia assistiva.** 2017.

ROMERO, Rodolfo Valiente. **Processo automático de reconhecimento de texto em imagens de documentos de identificação genéricos.** 2017. Dissertação (Mestrado em Sistemas Digitais) - Escola Politécnica, Universidade de São Paulo, São Paulo, 2017. Acesso em: 13 fev. 2021.

RYBENÁ. **Rybená – Tradutor de Português-Libras.** Disponível em: <https://portal.rybena.com.br/>. Acesso em: 20 nov. 2020.

SANTOS, Adenir Fonseca dos. **Uso das tecnologias da informação e comunicação - TIC – tecnologia assistiva – sistema inteligente para a capacitação e inclusão de pessoa com deficiência visual no mercado de trabalho – um estudo de caso.** 2018.

VELAZ, Bruno. OCR: **Reconhecimento Óptico de Caracteres.** Taugor Corporation. Disponível em: <https://www.taugor.com.br/>. Acesso em: 20 nov. 2020.

VLIBRAS. **O que é o VLIBRAS?** Disponível em: <<https://www.vlibras.gov.br/>>. Acesso em: 5 jun. 2020.

VOIGT, Johann Felipe et al. **Aprendizagem profunda para reconhecimento de gestos da mão usando imagens e esqueletos com aplicações em Libras.** 2018.

W3C. World Wide Web Consortium. 2021. Disponível em: <https://www.w3.org/>. Acesso em: 27 mar. 2021.

WANZELLER, Moisés Monteiro Pacheco. Tecnologias de apoio à tradução: tutorial de OCR. 2018.