

# TRADUAPP E A ATUAÇÃO DO RECONHECIMENTO ÓTICO DE CARACTERES NA TRADUÇÃO DE IDIOMAS

Caroline Eri Sato Ushirobira<sup>1</sup>, Rodrigo de Oliveira Plotze<sup>1</sup>

<sup>1</sup>Faculdade de Tecnologia de FATEC Ribeirão Preto (FATEC)

Ribeirão Preto, SP – Brasil

eriushirobira@hotmail.com, rodrigo.plotze@fatec.sp.gov.br

**Resumo.** *Com o avanço da globalização e conseqüentemente, da tecnologia presente em diversas áreas da sociedade atual, o OCR (Reconhecimento Ótico de Caracteres - Optical Character Recognition) desempenha um papel cada vez mais significativo no âmbito da tradução de idiomas. Além de auxiliar a compreensão de textos escritos em línguas estrangeiras, os métodos utilizados pelo OCR oferecem praticidade, agilidade e eficiência em sua execução. O TraduApp é um aplicativo desenvolvido em Flutter que busca contribuir com a área da Tradução de idiomas através dos princípios do OCR.*

**Abstract.** *With the advancement of the globalization process and, consequently, the technology present in several fields of our contemporary society, OCR (Significant Character Recognition - Optical Character Recognition) has an increasingly important role in the context of language translation. In addition to helping to understand texts written in foreign languages, the methods used by OCR are practical, agile and efficient in execution. TraduApp is an application developed in Flutter that seeks to contribute to an area of language translation through the principles of OCR.*

## 1. Introdução

A compreensão entre diferentes idiomas é uma demanda crescente no mundo globalizado, presente em diversas situações do cotidiano da sociedade atual. Desde uma simples interação em redes sociais internacionais até em situações formais de negociação internacional, a tradução é uma área que permite a compreensão mútua de idiomas estrangeiros de maneira rápida e prática. Para providenciar traduções instantâneas de textos escritos em idiomas estrangeiros, empresas e pessoas físicas podem utilizar softwares que realizam o OCR, efetuado a partir do processamento digital de imagens.

Segundo Anugrah e Bintoro (2017), a tecnologia tem facilitado a digitalização de documentos físicos e o OCR nos permite a conversão de textos impressos em digitais economizando tempo de digitação, por exemplo. Em linhas gerais, o processo de reconhecimento ótico de caracteres possui 6 etapas: aquisição das imagens, pré-processamento, segmentação, extração das *features* (atributos), identificação dos padrões e geração de relatório com as análises obtidas (GONZALEZ e WOODS, 2008).

Durante o processo de reconhecimento, os sistemas de OCR podem ser baseados em duas categorias: Análise Estrutural, que extrai de cada caractere um conjunto de *features* que compõe a estrutura do mesmo, e Comparação de Modelos, que utiliza uma espécie de matriz para representar a forma do caractere conforme o posicionamento de

seus pontos (TRIER, JAIN e TAXT apud Ivanski et al., 2008).

Este artigo tem como objetivo de analisar a atuação do reconhecimento ótico de caracteres no campo da tradução de idiomas, assim como detalhar a construção do aplicativo TraduApp que busca fornecer traduções instantâneas de textos extraídos de imagens de forma simples e prática para o usuário. O código utilizado para a construção do TraduApp está disponível em: <https://github.com/CarolEri/TraduApp>.

## 2. Referencial Teórico

A necessidade da tradução de idiomas em prol do entendimento humano existe desde meados de 300 a. C., quando surgiram os primeiros indícios de traduções escritas de textos sagrados do hebraico para o aramaico (SALES, 2021). Com o objetivo de transmitir a mensagem escrita para outros idiomas com comprometimento e fidelidade ao texto original, a tradução se faz presente na rotina da maioria das pessoas na atualidade, auxiliando a compreensão de músicas, filmes, notícias, livros, mensagens e diversos outros elementos da mídia e do cenário internacional.

Para facilitar o acesso à tradução de idiomas estrangeiros, diversos softwares foram desenvolvidos a exemplo dos aplicativos Google Tradutor, DeepL Tradutor, Cambridge Tradutor, entre outros, com o objetivo de proporcionar a tradução automática e instantânea para o usuário, de forma simples e intuitiva. Alguns softwares de tradução também permitem, além da inserção manual de palavras e frases, o envio de imagens que contenham textos para tradução. Esta opção facilita a tradução para usuários que, por exemplo, se deparam fisicamente com textos em idioma estrangeiro e não possuem acesso ao teclado de tal idioma em seu dispositivo móvel, precisam traduzir de maneira rápida um texto extenso que não se encontra no ambiente digital, ou até mesmo usuários portadores de deficiência visual.

Para realizar o processo de tradução de textos através de imagens digitais, é necessário utilizar um software OCR que extraia o conteúdo da imagem e transforme-o em texto. Segundo o Centro Tecnológico de Acessibilidade (CTA):

A ferramenta OCR funciona analisando o documento e comparando seus caracteres com fontes armazenadas em seu banco de dados e/ou reconhece características típicas de determinado caractere. De maneira bem simples, um OCR é um programa que reconhece caracteres, transformando imagens de texto em texto puro. (CTA, 2018).

Segundo Gonzalez e Woods (2008), uma imagem digital pode ser entendida como uma função que apresenta duas dimensões com valores e amplitude finitos:  $f(x, y)$ . Os elementos que compõe uma imagem digital são comumente chamados de pixels e cada pixel possui uma localização e valor específico dentro de uma imagem. Na Figura 1 elaborada por Queiroz e Gomes (2001) podemos observar uma representação matricial geral para a leitura dos pixels em uma imagem bidimensional. O primeiro índice apresenta a posição da linha denominada “m” e o segundo apresenta a posição da coluna, denominada “n”. Nesta ilustração, a imagem digital apresentaria M linhas e N colunas, com o índice m variando de 0 a M-1, e índice n variando de 0 a N-1:

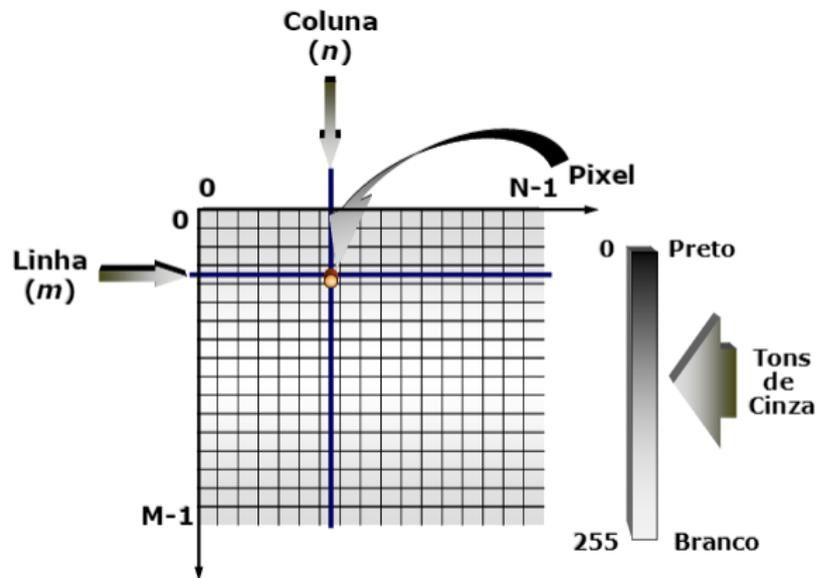


Figura 1 – Representação de uma imagem digital bidimensional.

Fonte: Queiroz e Gomes (2001)

Uma imagem digital colorida, por exemplo, pode ser entendida como a junção de pixels/vetores compostos por três imagens monocromáticas com intensidades luminosas de vermelho, verde e azul, pertencentes ao sistema de cores RGB (*Red, Green, Blue*) (QUEIROZ E GOMEZ, 2001).

### 3 – Materiais e Métodos

#### 3.1 - Modelos de contexto e de interação

Segundo Sommerville (2011) é essencial estabelecer os requisitos do sistema que satisfaçam as exigências dos *stakeholders*, assim como detalhar as funcionalidades exigidas para a organização e o desenvolvimento do sistema. O autor também argumenta que durante a fase de levantamento e análise de requisitos é necessário observar sistemas já existentes sobre o tema proposto, discutir ideias com potenciais usuários, prototipar e modelar as funcionalidades com o objetivo de proporcionar um entendimento mais claro sobre o produto a ser desenvolvido.

A UML (*Unified Modeling Language*) é uma ferramenta que auxilia a compreensão de sistemas de baixa até alta complexidade em um alto nível de abstração, permitindo o *design* dos processos do software que será desenvolvido (UML, 2005). No presente trabalho, foi desenvolvido o Diagrama de Caso de Uso através da ferramenta Astah, desenvolvida em prol da praticidade na modelagem, visualização e design de sistemas (ASTAH, 2022).

A Figura 2 foi elaborada com o objetivo de mapear as principais funções do aplicativo TraduApp, considerando que a principal é a tradução de textos através do reconhecimento óptico de caracteres.

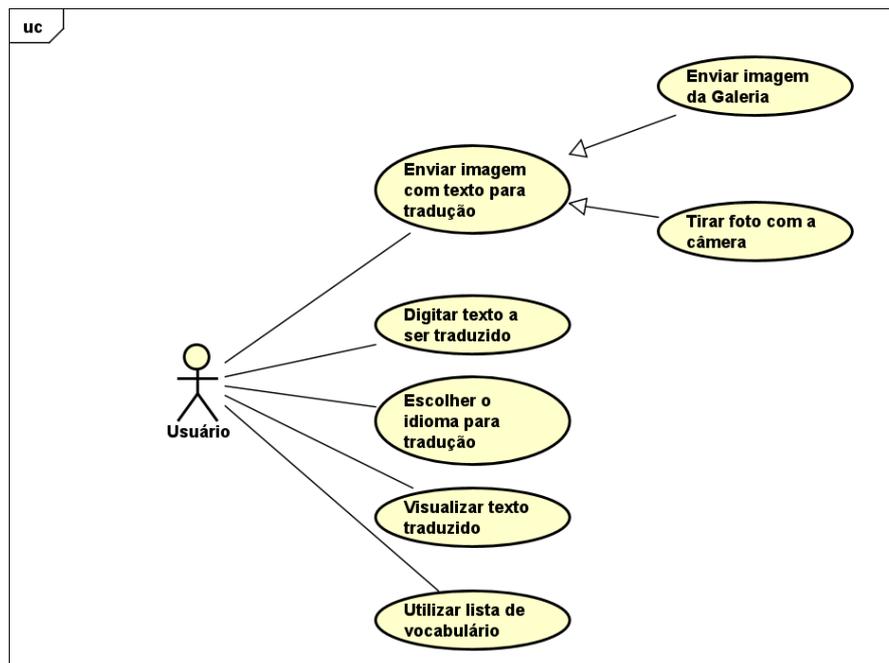


Figura 2 – Diagrama de Caso de Uso do Aplicativo TraduApp.

Fonte: Autoria própria

### 3.2 - Flutter integrado ao Firebase

O Flutter foi o *framework* multiplataforma escolhido para o desenvolvimento do TraduApp por possibilitar, segundo Wu (2018), a execução nos sistemas operacionais Android e iOS, um mecanismo de renderização próprio de alto desempenho, o recarregamento da aplicação de forma dinâmica e prática através do *hot-reload*, entre outros fatores que facilitam a codificação.

Para a construção do TraduApp foi utilizada a linguagem de programação Dart, desenvolvida pela Google e lançada em 2011. A linguagem apresenta algumas características marcantes do JavaScript (ES7) e sintaxe relativamente parecida com a da linguagem de programação Java (WU, 2018), além de ser otimizada para a codificação de interfaces de usuário que contenham codificação orientada a eventos (DART, 2022).

O Firebase foi a plataforma escolhida para criar e autenticar os usuários do aplicativo, assim como armazenar dados e palavras inseridas nas listas de vocabulário. O Firebase Authentication proporciona serviços de *back-end* através de kits práticos de desenvolvimento, bibliotecas prontas e praticidade durante o processo de autenticação de usuários através de e-mail e senha, logins do Google, Facebook, Apple, entre outros (FIREBASE, 2021). Durante a construção do TraduApp, foi escolhido o método de autenticação e cadastro de usuários através do e-mail e senha.

Outro serviço oferecido pelo Firebase e utilizado durante a construção do presente trabalho é o Cloud Firestore, que atua como um banco de dados não-relacional, “flexível e escalonável para desenvolvimento focado em dispositivos móveis, Web e servidores pelo Firebase e do Google Cloud” (FIREBASE, 2021). O Firestore foi utilizado no TraduApp para armazenar dados cadastrais dos usuários e, em uma pasta distinta, as palavras da lista de vocabulário.

### 3.3 – Tesseract OCR

Neste trabalho, o Tesseract OCR foi escolhido como ferramenta para realizar o reconhecimento óptico de caracteres. O Tesseract é um programa que utiliza linhas de comando em conjunto com um motor de OCR (a *libtesseract*) desenvolvidos majoritariamente por Ray Smith e atualmente mantido por Zdenko Podobny. A ferramenta apresenta suporte a *unicode* (UTF-8) e consegue reconhecer prontamente mais de 100 idiomas (TESSERACT, 2022). Os arquivos do programa e a documentação do Tesseract podem ser acessados através do repositório principal do projeto no GitHub, através do link: <https://github.com/tesseract-ocr/tesseract>.

### 3.4 – Google Cloud Translation API

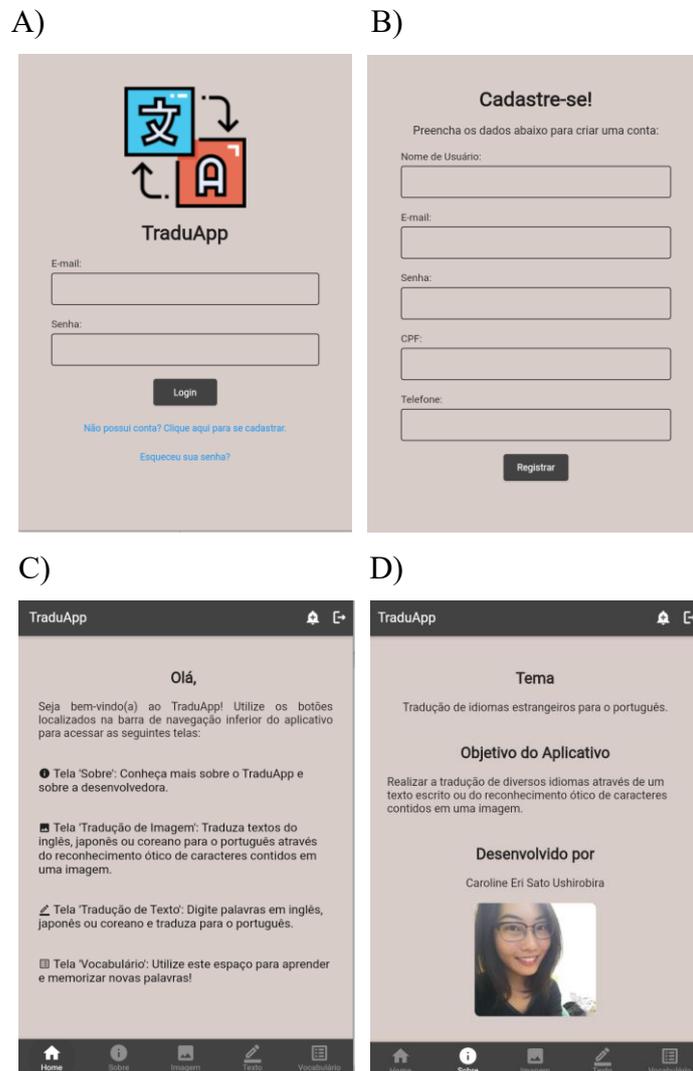
A Google Cloud Translation API (Application Programming Interface), desenvolvida pela Google, foi a ferramenta de tradução escolhida para a construção do TraduApp por utilizar a inteligência artificial para oferecer traduções rápidas, automáticas e dinâmicas para mais de cem idiomas (GOOGLE CLOUD, 2022). Neste projeto, foi utilizada a versão básica da API para realizar traduções do idioma inglês para o português.

## 4. Resultados

### 4.1 – Aplicativo TraduApp

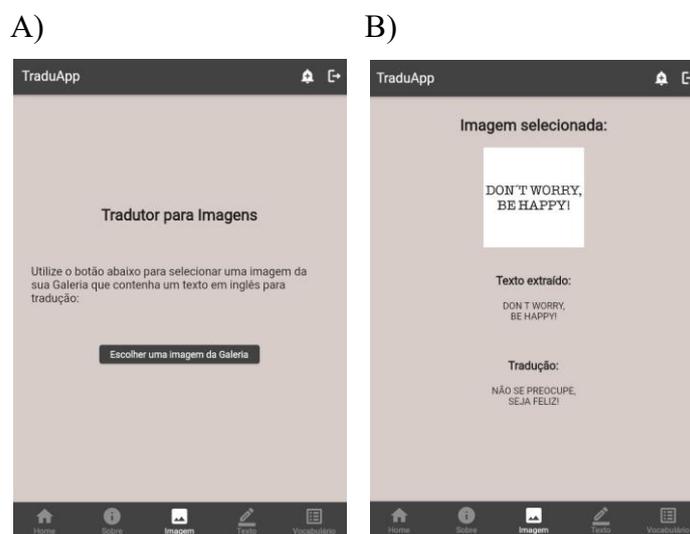
O aplicativo TraduApp foi desenvolvido para operar em Android e Web através do Flutter, com o objetivo principal de oferecer traduções instantâneas e confiáveis da língua inglesa para a língua portuguesa através do reconhecimento óptico de caracteres. Para realizar a tradução, o usuário deve efetuar o *login* no aplicativo (Figura 3A), utilizando e-mail e senha previamente cadastrados no sistema. Caso o usuário não possua cadastro, poderá clicar na frase “Não possui conta? Clique aqui para se cadastrar.”, que fará a alteração do botão de *login* para o de registrar uma nova conta no aplicativo. A tela de Cadastro (Figura 3B) solicita o nome do usuário, e-mail, senha, CPF e telefone para registro. Após o preenchimento dos dados, o usuário poderá efetuar *login* e acessar as funcionalidades do aplicativo.

A tela da *Home* (Figura 3C) é a primeira a ser visualizada dentro do aplicativo, que apresenta uma breve explicação sobre as demais telas e suas respectivas funcionalidades. Para navegar dentro do aplicativo, o usuário deve utilizar a barra de navegação inferior, que carrega cinco ícones. O segundo ícone da esquerda para a direita leva o usuário até a tela Sobre (Figura 3D), que apresenta o tema, o objetivo principal do aplicativo e informações sobre a desenvolvedora.



**Figura 3. A) Tela de login do aplicativo TraduApp. B) Tela de cadastro de novos usuários. C) Tela Home, primeira tela apresentada após o login. D) Tela “Sobre” com informações sobre o aplicativo.**

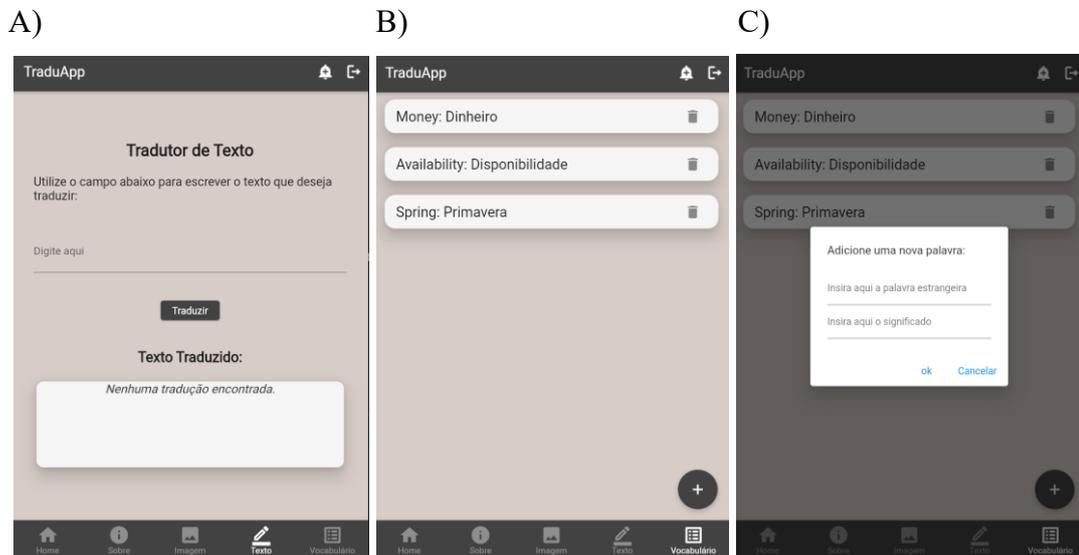
Para utilizar a principal funcionalidade do TraduApp, a tradução da língua inglesa para a língua portuguesa através do reconhecimento ótico de caracteres contidos em uma imagem, o usuário deve navegar até a tela do Tradutor para Imagens (Figura 4A) clicando no ícone central da barra de navegação inferior, e escolher uma imagem da Galeria de Fotos. Após selecionar a imagem desejada contendo o texto em inglês o aplicativo, após alguns segundos, mostrará na tela a imagem selecionada em tamanho reduzido, o texto original extraído e o resultado da tradução para o idioma português (Figura 4B). Neste exemplo, foi selecionada uma imagem com fundo branco e letras pretas que formam a frase “*Don't worry, be happy!*”, traduzida corretamente para “*Não se preocupe, seja feliz!*”:



**Figura 4. A) Tela do Tradutor para Imagens. B) Tela do resultado da tradução por imagem.**

Além da funcionalidade principal, o TraduApp também oferece ao usuário a possibilidade de realizar traduções de textos digitados através do teclado ou copiados de ambientes dentro do dispositivo. Para acessar o Tradutor de Textos (Figura 5A), o usuário deve navegar até o quarto ícone da barra de navegação, inserir o texto em idioma estrangeiro e acionar o botão “Traduzir”, que retornará a tradução para a língua portuguesa dentro da área branca.

O TraduApp também permite ao usuário armazenar uma lista de vocabulário, com palavras em idioma estrangeiro seguidas por suas respectivas traduções em português, para estudo e memorização de vocabulário. Para acessar, editar e inserir novas palavras, o usuário deve navegar até a tela da Lista de Vocabulário (Figura 5B), clicar nos cartões existentes para editar as palavras previamente armazenadas, clicar no ícone da lixeira para excluir palavras indesejadas, ou clicar no botão circular com sinal de adição (“+”) no canto inferior direito da tela. Ao clicar neste botão, o aplicativo abrirá um *pop-up* contendo dois campos para adição da palavra estrangeira e seu respectivo significado (Figura 5C):



**Figura 5. A) Tela do Tradutor de Textos digitados/copiados. B) Tela da lista de Vocabulário. C) Imagem da Tela da lista de Vocabulário apresentando o *pop-up* para inserção de novas palavras.**

## 5. Conclusão e Perspectivas Futuras

O framework Flutter atendeu todas as necessidades de desenvolvimento do TraduApp, assim como possibilitou uma vasta coleção de pacotes disponíveis para integração no aplicativo através do repositório oficial pub.dev (<https://pub.dev/>) e suas respectivas documentações. A utilização da API Firebase também foi uma ótima escolha para utilização no TraduApp por proporcionar as funcionalidades de cadastro, *login* e armazenamento de informações em um banco de dados não relacional de forma simples, prática e dinâmica.

A ferramenta Tesseract OCR também atendeu as necessidades do aplicativo de forma satisfatória, apesar de não proporcionar o reconhecimento perfeito de caracteres escritos com cores não contrastantes com o fundo da imagem. Para futuras iterações do aplicativo, serão adicionadas as funcionalidades de reconhecimento de outras línguas estrangeiras, assim como a possível integração da funcionalidade de reconhecimento inteligente de caracteres (ICR) manuscritos.

A API Google Cloud Translation apresentou resultados e desempenhos satisfatórios que atenderam os objetivos do TraduApp, proporcionando traduções corretas do idioma inglês para o português. Como perspectiva futura, a API será utilizada também para traduções de outros idiomas estrangeiros.

Por fim, o TraduApp atendeu todos os requisitos obrigatórios planejados no início do desenvolvimento, permitindo que o usuário envie uma imagem contendo um texto da língua inglesa para tradução automática para a língua portuguesa, além de proporcionar a tradução de textos digitados e armazenamento de vocabulário em lista para estudo. Após passar por fases de testes e consequentes ajustes no futuro, o

aplicativo poderá ser homologado e disponibilizado para o público geral.

## Referências

- ANUGRAH, Rio; BINTORO, Ketut B. Y. Latin Letters Recognition Using Optical Character Recognition to Convert Printed Media Into Digital Format. *Jurnal Elektronika dan Telekomunikasi (JET)*, Vol. 17, No. 2, December 2017, pp. 56-62. Disponível em: < <https://www.jurnalet.com/jet/article/view/163>>. Acesso em 19 mar. 2022.
- ASTAH. Powerful and Fast UML Diagramming Software, 2022. Disponível em: <<https://astah.net/products/>>. Acesso em 14 mar. 2022.
- CENTRO TECNOLÓGICO DE ACESSIBILIDADE. Ferramentas OCR – entenda o que são e sua relação com a acessibilidade. 17 dez 2018. Disponível em: < <https://cta.ifrs.edu.br/ferramentas-ocr-entenda-o-que-sao-como-funcionam-e-qual-sua-relacao-com-a-acessibilidade/>>. Acesso em 15 jan. 2022.
- DART. Dart Overview, 2022. Disponível em: <<https://dart.dev/>> Acesso em 15 mar. 2022.
- FIREBASE. Firebase, 2021. Disponível em: <<https://firebase.google.com/?hl=pt-br>>. Acesso em 15 mar. 2022.
- GONZALEZ, Rafael C.; WOODS, Richard E. *Digital Image Processing*, 2008.
- GOOGLE CLOUD. IA de tradução, 2022. Disponível em: < <https://cloud.google.com/translate>> . Aceso em 04 abr. 2022.
- IVANSKI, William; SILVA, Luciano; BELLON, Olga R. P. Método Híbrido de Reconhecimento Ótico de Caracteres. Grupo IMAGO de Pesquisa – UFPR. Disponível em: < <http://www.gpec.ucdb.br/sibgrapi2008/wuw/47964.pdf>> . Acesso em 19 mar. 2022.
- QUEIROZ, José Eustáquio Rangel de; GOMES, Herman Martins. Introdução ao Processamento Digital de Imagens, 2001. *Revista RITA*, Volume VIII, Número 1. Disponível em: <<http://www.dsc.ufcg.edu.br/~hmg/disciplinas/graduacao/vc-2016.2/Rita-Tutorial-PDI.pdf>> . Acesso em 15 jan. 2022.
- SALES, Antonia de Jesus. Estudos da Tradução: uma introdução, 03 Abril 2021. Disponível em: < [https://www.researchgate.net/publication/350610357\\_Estudos\\_da\\_Traducao\\_uma\\_introducao](https://www.researchgate.net/publication/350610357_Estudos_da_Traducao_uma_introducao)> . Acesso em 15 jan. 2022.
- SOMMERVILLE, Ian. *Software engineering 9th Edition*. ISBN-10, v. 137035152, p. 18, 2011. Disponível em: < <https://engineering.futureuniversity.com/BOOKS%20FOR%20IT/Software-Engineering-9th-Edition-by-Ian-Sommerville.pdf>> Acesso em 14 mar. 2022.
- TESSERACT, 2022. Tesseract OCR ReadMe.md, GitHub. Disponível em: <<https://github.com/tesseract-ocr/tesseract>> . Acesso em 02 abr. 2022.
- UNIFIED MODELING LANGUAGE - UML®. What is UML. Disponível em: <<https://www.uml.org/what-is-uml.htm>> Acesso em 14 mar. 2022.
- WU, Wenhao. React Native vs Flutter, cross-platform mobile application frameworks.

Metropolia University of Applied Sciences, Bachelor of Engineering, Thesis. 01 de Março de 2018. Disponível em: <<https://www.theseus.fi/handle/10024/146232>>. Acesso em 14 mar. 2022.