

ICamSee – Sua vida mais independente

Alunos: Felipe Andrade (felipe_saff1@hotmail.com); Gabriel Pitol (newdracors@gmail.com); Gabriel Cervi (engineercervi@gmail.com); João Bosco (jcorreiadelimajunior@gmail.com) Rodrigo Wallace (rodrigo.wallace.ss@gmail.com)
Orientador: Isaac Jesus da Silva (isaacjesus@fei.edu.br)

RESUMO

Visando melhorar a vida de deficientes visuais, desenvolveu-se o conceito de um óculos inteligente. O modelo tem opções de detecção de textos, data de validade e de cores. Todos estes modos offline e contam com respostas emitidas por voz via fone Bluetooth. Para validar a ideia foram realizados muitos testes usando uma webcam e objetos do dia a dia, simulando situações reais de aplicação do mesmo.



PROPOSTA

Desejamos desenvolver um sistema de reconhecimento de imagem que possa trazer para o usuário uma maior facilidade no acesso à informações básicas, por exemplo: nome de produtos, validade, cores do objeto entre outros.

Diferentemente de qualquer produto parecido que já exista, este conceito visa o melhor custo-benefício para poder atender o maior número possível de pessoas, dentre elas indivíduos que apresentam deficiência visual em qualquer grau, pessoas com qualquer nível de analfabetismo e daltônicos.

Foi feito de forma a ser prático para transporte sendo estruturado com uma câmera acoplada em um óculos, no qual todo o processamento de imagem será feito em um microcomputador que pode ser transportado facilmente usando uma pequena pochete. Todos os dados captados serão traduzidos utilizando nosso sistema de reconhecimento de texto e serão convertidos para áudio sendo enviados para o usuário via fone bluetooth.



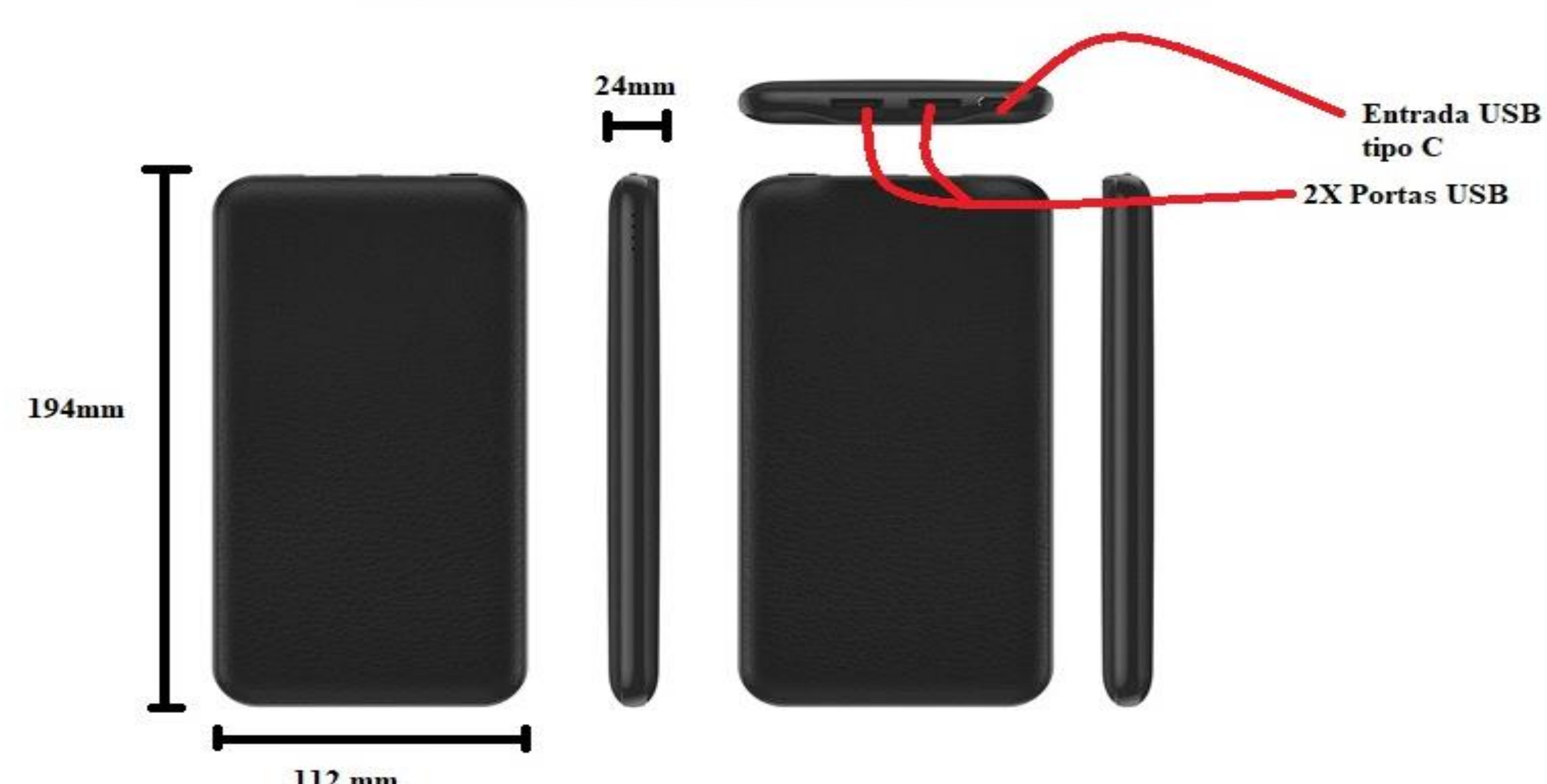
PROTÓTIPO

Visando o melhor desempenho e um custo acessível foram escolhidos como componentes os seguintes itens:

- Uma webcam com resolução igual ou superior a Full HD (1080p)
- Um microcomputador Nvidia Jetson Nano
- Uma Bateria Power Bank de 15000mAh para alimentação do sistema
- Fones bluetooth para captação do áudio gerado pelo programa
- Um sistema de armazenamento para o hardware, estilo pochete
- Um botão acoplado ao óculos para selecionar os modos de funcionamento.

Os modos do sistema são divididos em leitura de textos, como folhetos, jornais livros entre outros, detecção de data de validade, e reconhecimento de cores, onde ele lhe manda a cor predominante captada pela webcam, sendo todos estes com resposta em áudio.

Vale ressaltar que todos os objetos foram adaptados para melhor se ajustar no rosto como por exemplo a reestruturação do webcam para não atrapalhar o usuário.

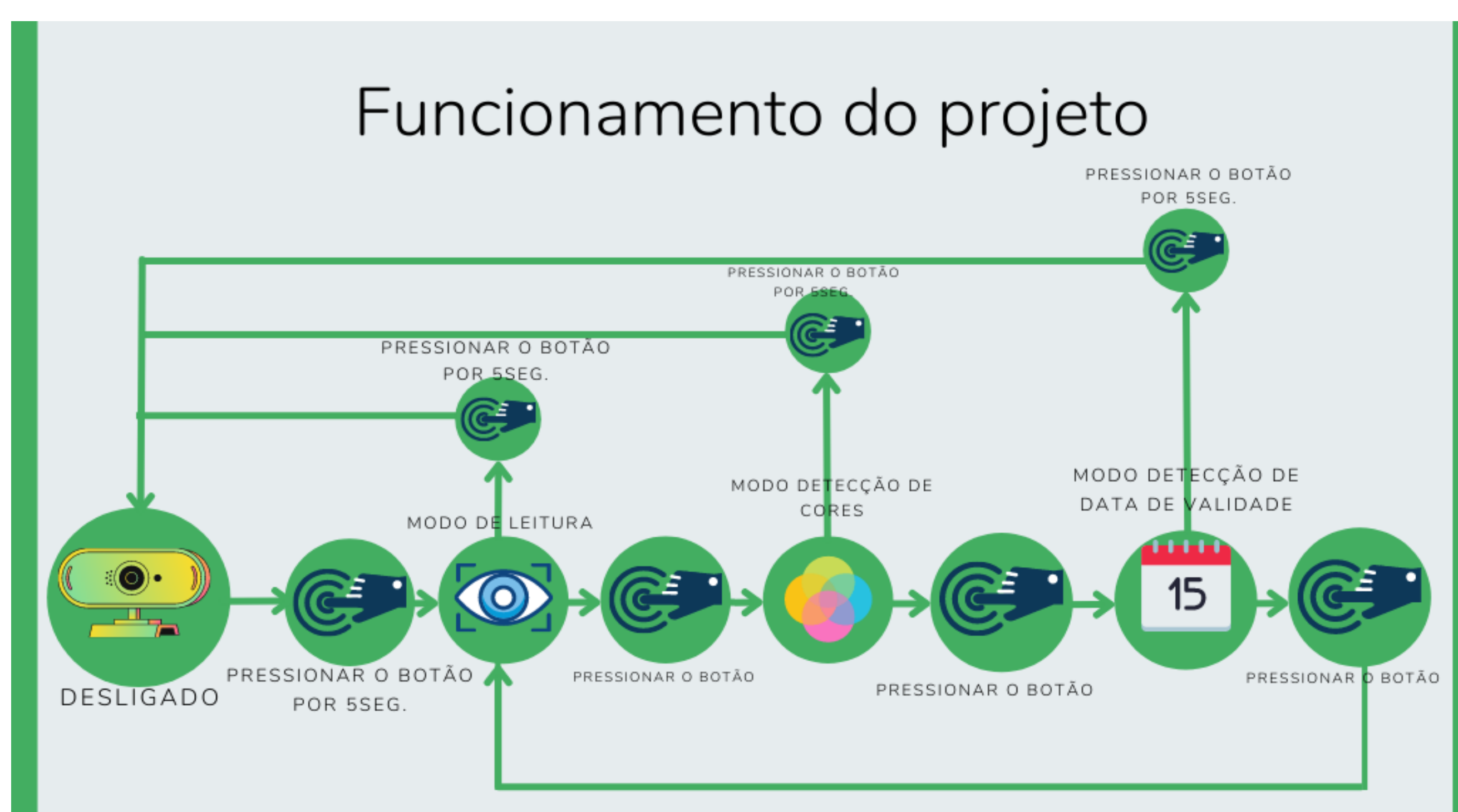


ICamSee – Sua vida mais independente

Aluno: Felipe Andrade (felipe_saff1@hotmail.com); Gabriel Pitol (newdracors@gmail.com); Gabriel Cervi (engineercervi@gmail.com); João Bosco(jcorreielimajunior@gmail.com) Rodrigo Wallace (rodrigo.wallace.ss@gmail.com)
Orientador: Isaac Jesus da Silva (isaacjesus@fei.edu.br)

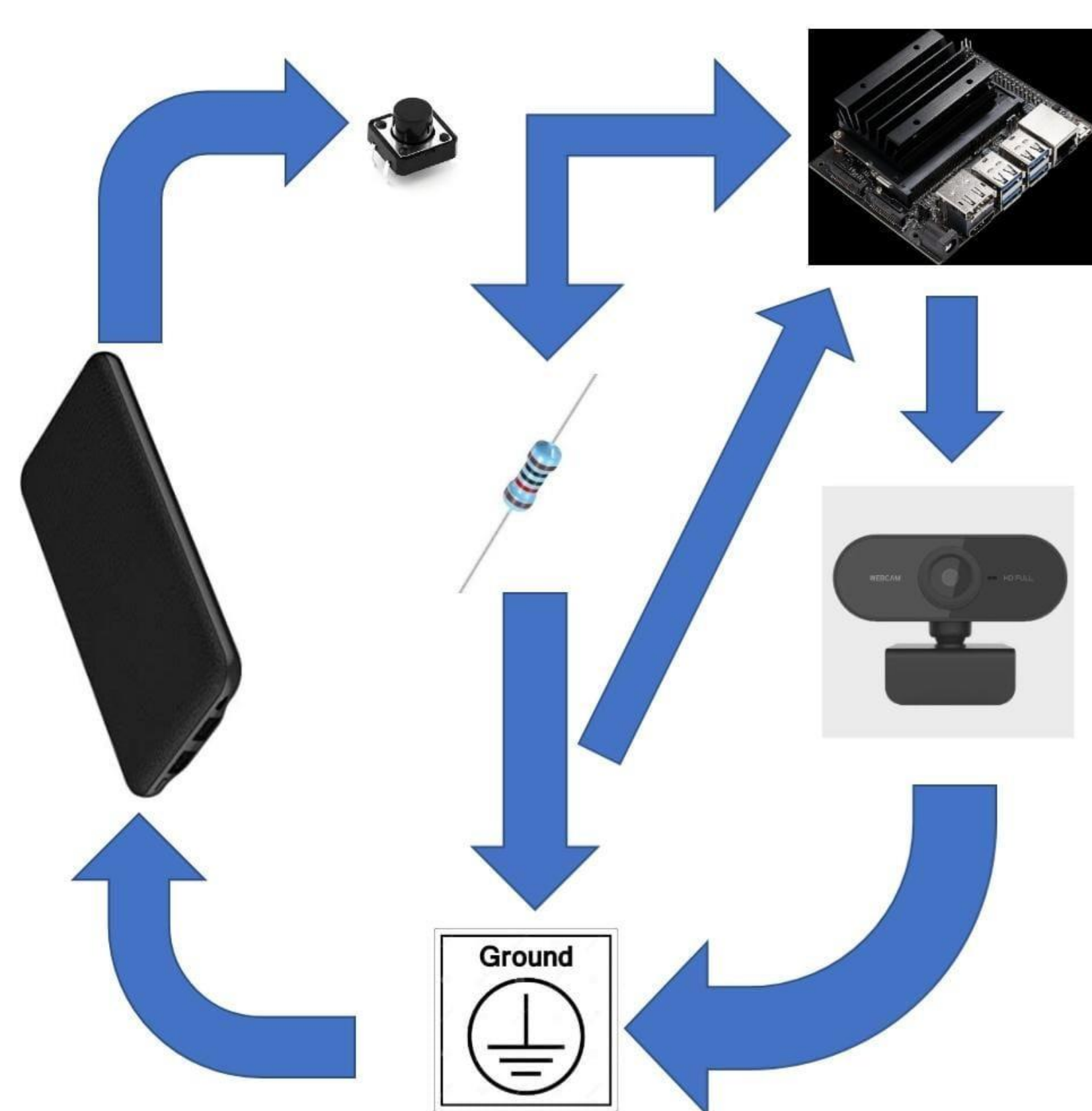
FUNCIONAMENTO DO SISTEMA

Com o sistema desligado, ao ser pressionado o botão por cinco segundos ele será ligado indo de forma padrão para o modo de leitura. Caso queira se alternar o modo de operação, será necessário apenas o toque do botão para selecionar os modos, sendo eles: detecção de cores e leitura de data de validade. Retorna-se para o modo leitura caso se aperte o botão novamente, quando se esta na leitura de data de validade. Vale lembrar que em qualquer momento com o sistema ligado, ao ser pressionado o botão por cinco segundos ele será desligado.



SISTEMA ELÉTRICO DO PROTÓTIPO

O circuito apresentado a seguir representa a ligação que deve ser feita entre os componentes já mencionados para que o protótipo funcione corretamente. Nele, pode ser visto que a modelagem apresenta os componentes como sendo blocos, dos quais há inputs e outputs, de acordo com seu respectivo funcionamento. Vale ressaltar que foi respeitada a tensão de alimentação de cada componente e da proveniente do Power Bank, ou seja, 5V. O botão fora ligado com pull-down para que possa funcionar corretamente como chave que liga ou desliga o circuito e como selecionador dos modos de operação.



RECURSOS UTILIZADOS

O primeiro recurso utilizado foi o Open Source Computer Vision Library conhecido como OpenCV. Esta biblioteca conta com um módulo para processamento de álgebra linear em grande velocidade. Por conta do seu desempenho possui diversos usos em muitos projetos na atualidade.

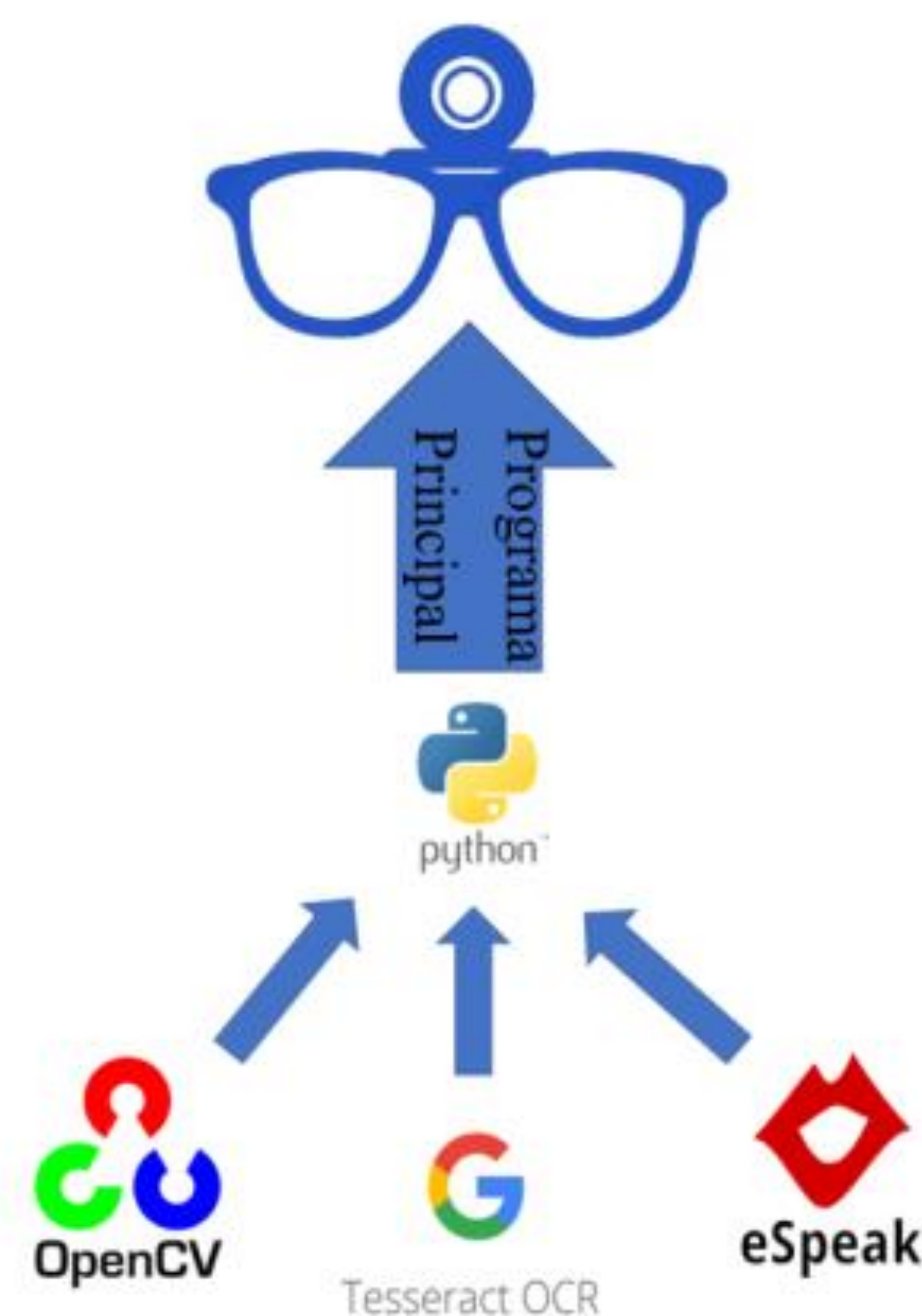
Este recurso permite a aquisição de imagens de câmeras digitais, tratamento e processamento delas, podendo ser feito sob dados previamente armazenados ou obtidos em tempo real.

O segundo recurso utilizado foi o PyTesseract que basicamente é uma biblioteca de reconhecimento de caracteres digitais, ou seja, é uma livraria utilizada para "ler" textos encontrados em imagens e convertê-los tanto para variáveis que podem ser utilizadas diretamente pelo programa, quanto para arquivos de texto.

Essa biblioteca se vale de inteligência artificial para, primeiramente, procurar e reconhecer caracteres em uma imagem e então completar qualquer falha no decorrer do texto que não tenha sido detectada, completando com os caracteres que mais façam sentido no contexto da frase.

Ao ser utilizado em conjunto com o OpenCV, o PyTesseract se vale das imagens adquiridas pelo mesmo, transformando todas as palavras encontradas em frases.

O último recurso utilizado foi o eSpeak que é a biblioteca usada na conversão de texto para voz e que conta com um software de código aberto. Ela utiliza um método de síntese mais compacto e rápido do que o dos softwares mais conhecidos de mesma função, não fazendo uso de internet e permitindo a conversão dos textos criados pelo Pytesseract, convertendo-os numa voz enviada diretamente para o fone bluetooth.



ICamSee – Sua vida mais independente

Aluno: Felipe Andrade (felipe_saff1@hotmail.com); Gabriel Pitol (newdracors@gmail.com); Gabriel Cervi (engineercervi@gmail.com); João Bosco(jcorreielimajunior@gmail.com) Rodrigo Wallace (rodrigo.wallace.ss@gmail.com)
Orientador: Isaac Jesus da Silva (isaacjesus@fei.edu.br)

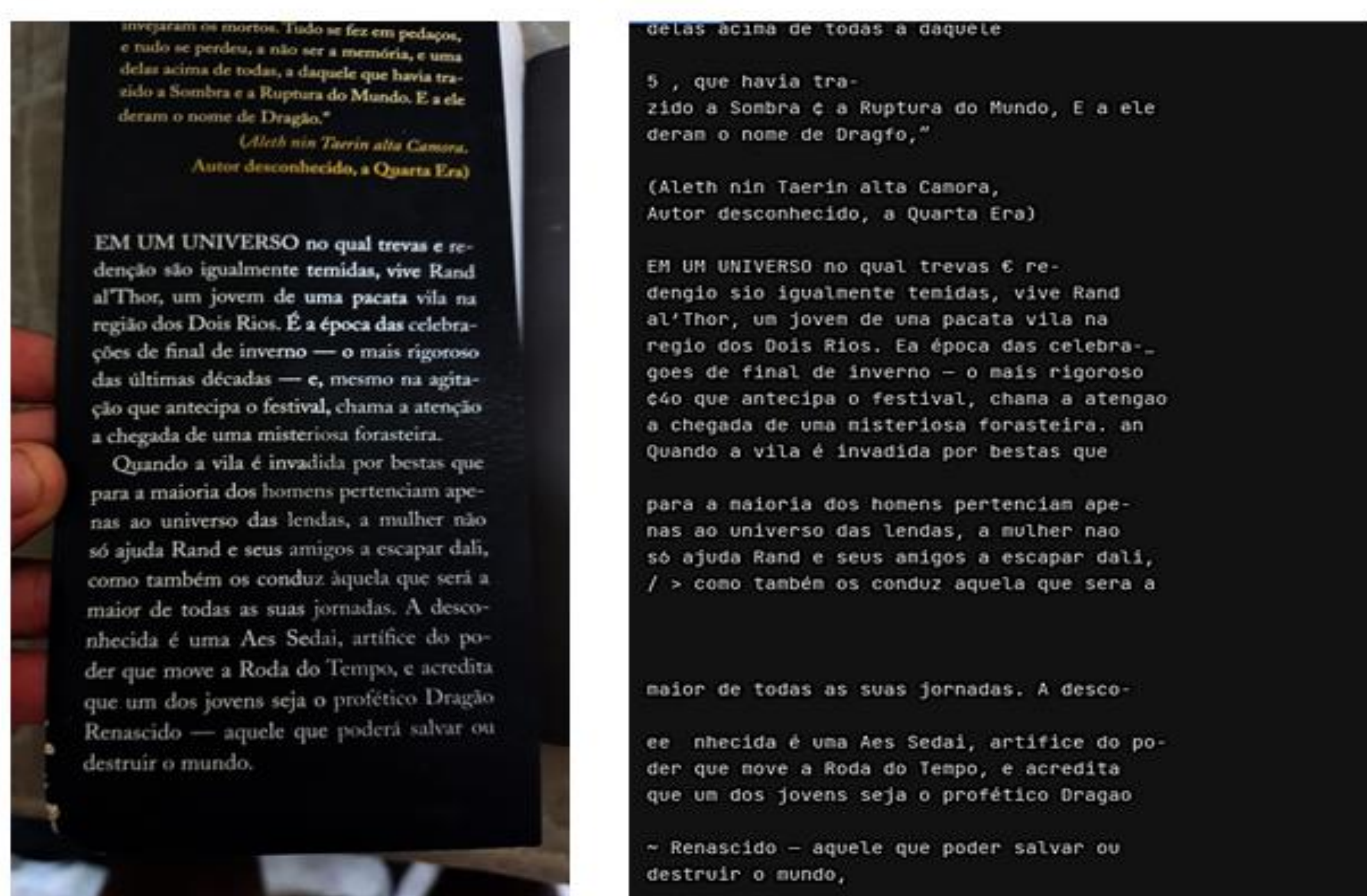
RESULTADOS OBTIDOS E DISCUSSÕES

Com base no código confeccionado com as bibliotecas mencionadas, foi obtido um programa capaz de atender todas as demandas requeridas neste projeto. A partir de uma webcam e de alguns objetos cotidianos conseguimos validar cada um dos modos de operação do protótipo.

Para o modo de detecção de textos é feito, inicialmente, uma análise do número de caracteres de tudo que foi identificado como texto. Esta análise é feita com a coloração normal, depois com o filtro monocromático e por último com mudança de contraste e dilatação. Os caracteres isolados, como “ç”, sem contexto, são removidos e, por fim, baseando-se nas coordenadas das palavras validadas da imagem, reestrutura-se uma frase que as contenham na respectiva ordem que se apresentam.



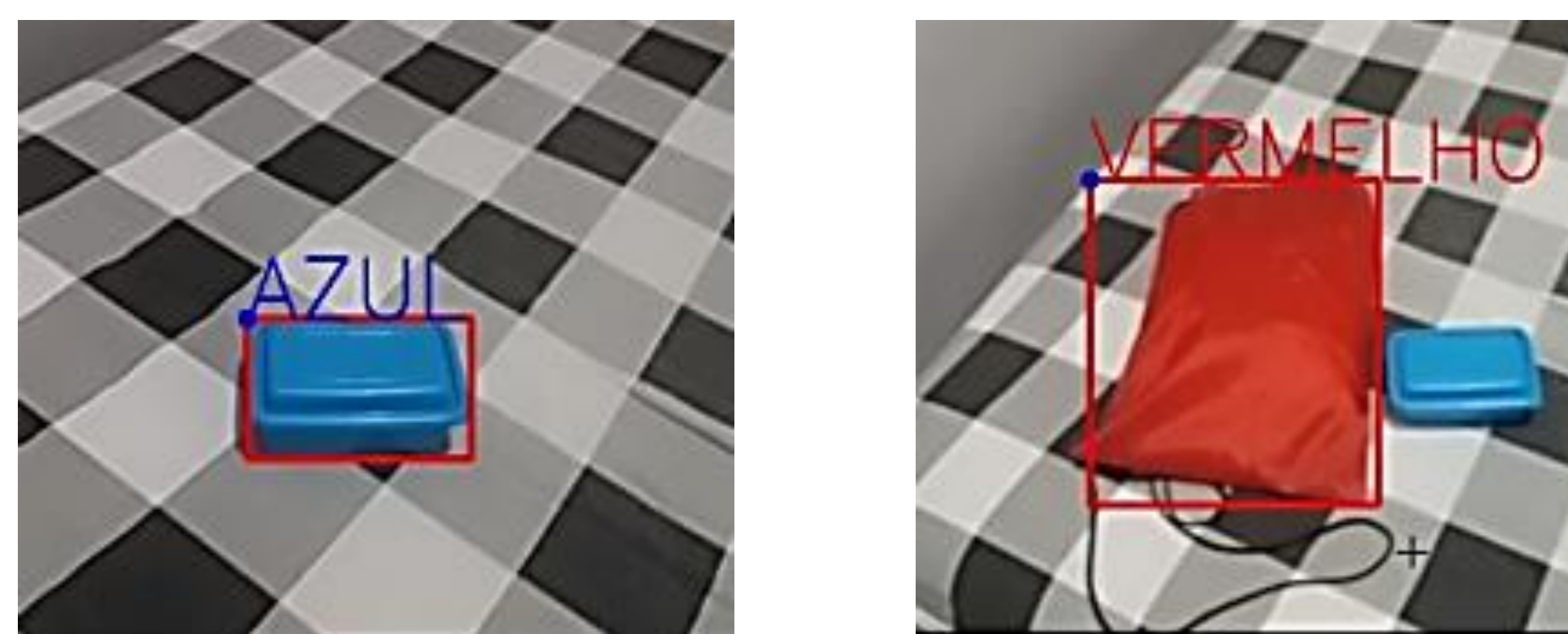
Já o segundo texto utilizado para teste foi de um livro em português, sendo ele capturado pela webcam supracitada.



Baseando-se neste teste em português pôde-se concluir que a implementação feita para detecção de textos apresenta dificuldade nos casos de leitura em imagens tremidas ou desfocadas e nos casos de interpretação de caracteres isolados.

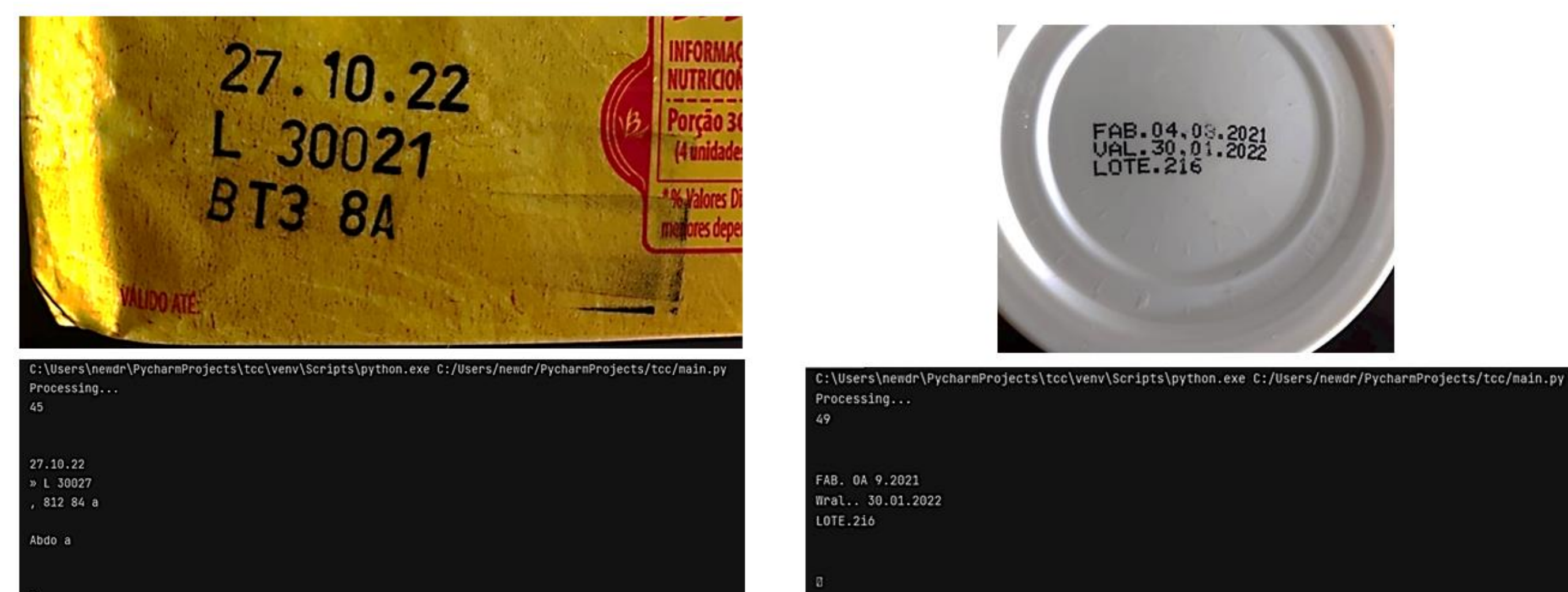
Quanto à estes dois testes que foram feitos, vale ressaltar que toda essa filtragem é feita apenas quando o texto capturado apresenta menos de 150 caracteres. No caso contrário, realiza-se uma leitura mais simples, aumentando apenas o contraste da imagem e capturando seu texto. Neste caso ganha-se mais velocidade de processamento e perde-se precisão.

Quanto à detecção de cores, faz-se um filtro da imagem capturada pela webcam analisando-se nela qual a cor dos pixels predominantes, de modo que se informa ao usuário esta cor. Os testes a seguir evidenciam os resultados obtidos para a detecção de cores em objetos.



As únicas dificuldades encontradas na detecção de cores são referentes às cores que sejam muito próximas no espectro, uma vez que o filtro confeccionado detecta apenas nove cores.

Por fim, quanto à detecção da data de validade, valendo-se de alguns produtos comuns em nossas casas, obteve-se:



Na detecção da data de validade no rótulo de produtos, faz-se a análise num filtro monocromático e de tudo que é capturado apenas é transmitido ao usuário a informação quanto a validade. Para localizar esta informação o código busca por palavras similares a “validade” e confirma que existam números a frente desta, caso positivo, confirma-se que é a informação buscada.

As principais dificuldades são nos casos em que a impressão não é contínua (letras pontilhadas), embalagens com irregularidades na superfície (amassados) e quando a cor da fonte é muito próxima a cor do fundo.

CONCLUSÕES

Por conseguinte, concluiu-se que o projeto atendeu as expectativas propostas, de forma que os modos de operação testados funcionam satisfatoriamente, validando assim a prova de conceito.

Este trabalho pode ser empregado como base para a confecção de um protótipo, o qual pode ser testado com o público-alvo e aperfeiçoado.