**MULTIPULE COLOR DETECTION & IMAGE TO SPEECH Converter**

**International Institute Of Information Technology, pune**

Shaikh Sadiya Rafiq Jadhav Tejal Sharad Tehare Shreyash s

[Sksadiya2110@gmail.com](mailto:Sksadiya2110@gmail.com) [tejaljadhav798@gmail.com](mailto:tejaljadhav798@gmail.com) [sxreyash@gmail.com](mailto:sxreyash@gmail.com)

**Abstract:**

Human communication today is mainly via speech and text. To access information in a text, a person needs to have vision. However, those who are deprived of vision can gather information using their hearing capability. The proposed method is a camera based assistive text reading to help blind person in reading the text present on the text labels, printed notes and products. The proposed project involves Text Extraction from image and converting the Text to Speech converter, a process which makes blind persons to read the text. This is the first step in developing a prototype for blind people for recognizing the products in real world, where the text on product is extracted and converted into speech. This is carried out by using Python and gtts (google text to speech).

Keywords: image to speech, python and gtts.

**I. INTRODUCTION**

Machine replication of human functions like reading is an ancient dream from a few years . However, in past the last five decades, machine reading has grown from a imaginary to reality. Today, there are already some systems that have some promise for conversion use, like portable Universal Product Code readers designed to assist blind people identify different products in an intensive product database can enable users who are blind to access information about these products through speech and Braille. But an enormous limitation is that it's extremely hard for blind users to look out the what's written in image.Speech is probably the foremost efficient medium for communication between humans. To extract the text from image we use optical character recognition technique(OCR) tool for python sponsored by google(Pytesseract(Python-tesseract) ). Optical character recognition has become one in every of the foremost successful applications of technology within the sector of pattern recognition and AI. Character recognition or optical character recognition (OCR) is that the method of converting images of machine printed or handwritten text (numerals, letters, and symbols) into a computer format text. Speech synthesis is that the synthetic synthesis of human speech. PIL having capability to process image to your Python interpreter.gTTS (Google Text-to-Speech), a Python library and CLI tool to interface with Google Translate's text-to-speech API.and store in .mp3 file format.

**II. OBJECTIVE:**

Our aim during this project is to make a prototype system for recognition of text present within the image using OpenCV. As illustrated within the diagram (Figure 1) the system framework consists of 5 functional components:

• Image acquisition

• Text extraction

• Text to speech conversion and

• Speech output.

Existing systems for text recognition are typically limited either by explicitly looking forward to specific shapes or colour masks or by requiring user assistance or could even be of high cost. Therefore, we would like a low-cost system which will be ready to automatically locate and skim the text aloud to visually impaired persons. the foremost idea of this project is to acknowledge the text character and convert it into mp3 file.

**III. Existing SYSTEM:**

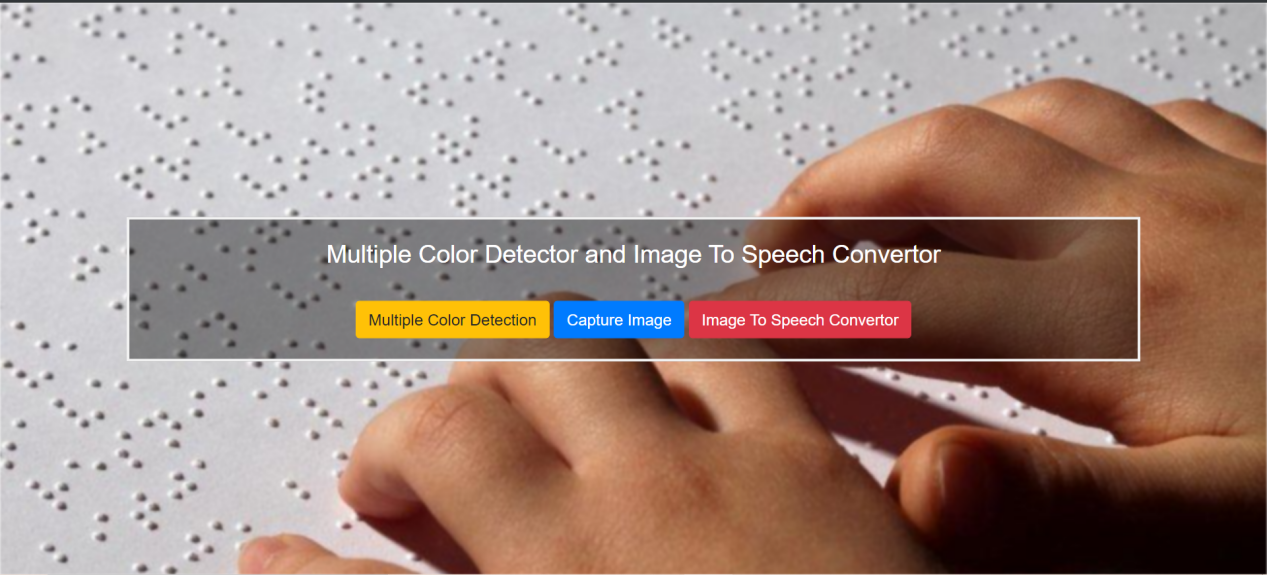
The Open CV could also be a python library which will be used for several tasks, like color detection, face recognisation, and mask detection . The Open CV board comes equipped with an SD card. This slot permits us to insert an SD card which may use it as their devices. Python is that the recommended programming language , but windows is its recommended OS. Nearly in every OS that works on Raspberry Pi.Optical character recognition, or OCR, could also be a way of converting a scanned image into text.When a page is scanned, it's typically stored as a bit-mapped enter TIF format. When the image is displayed on the screen, we'll read it. But to the pc, it's just a series of black and white dots. the pc doesn't recognize any "words" on the image. this is often often what OCR does. OCR looks at every line of the image and attempts to figure out if the black and white dots represents to a specific letter or number. OCR was actually developed for originally to assist sight-impaired individuals gain access to printed information.That same technology has been updated and improved and is now used to "read" computer files.

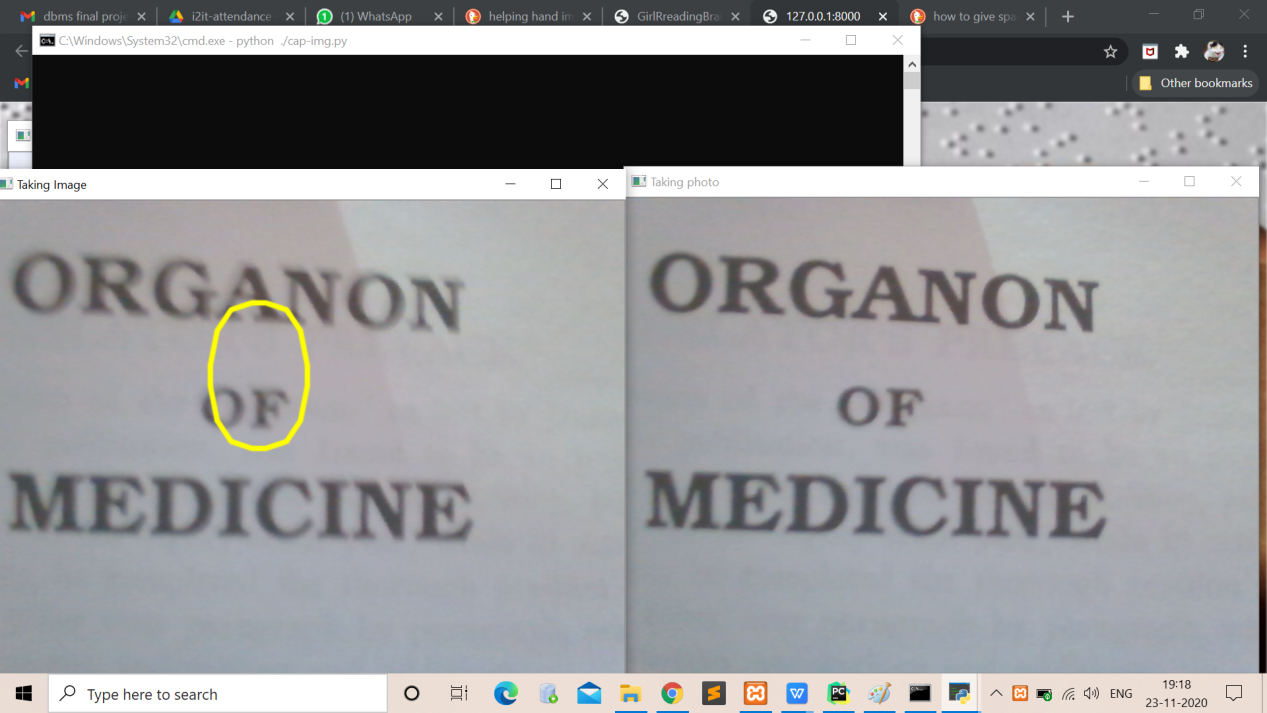
**IV. Proposed System:**

In this system we are using openCV for image capturing After capturing that image is stored then we fetching text from image by using Pytesseract(Python-tesseract) it's an optical character recognition (OCR) tool for python sponsored by google and storing in to file after this we are performing text to speech by using gtts(google text to image).In our proposed system no need of any hardware kit so it's low cost system.

**Image Acquisition**

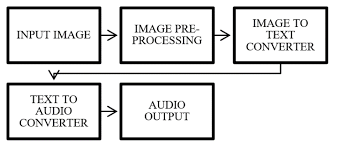
In this step the image of the text is captured using web camera. Time require to capture image is 5 sec timer for closing capturing window press ‘Q’ and for closing camera window press ‘esc’ key .The acquired image is then applied to the image text extraction step for extracting text from image.





**Text Extraction**

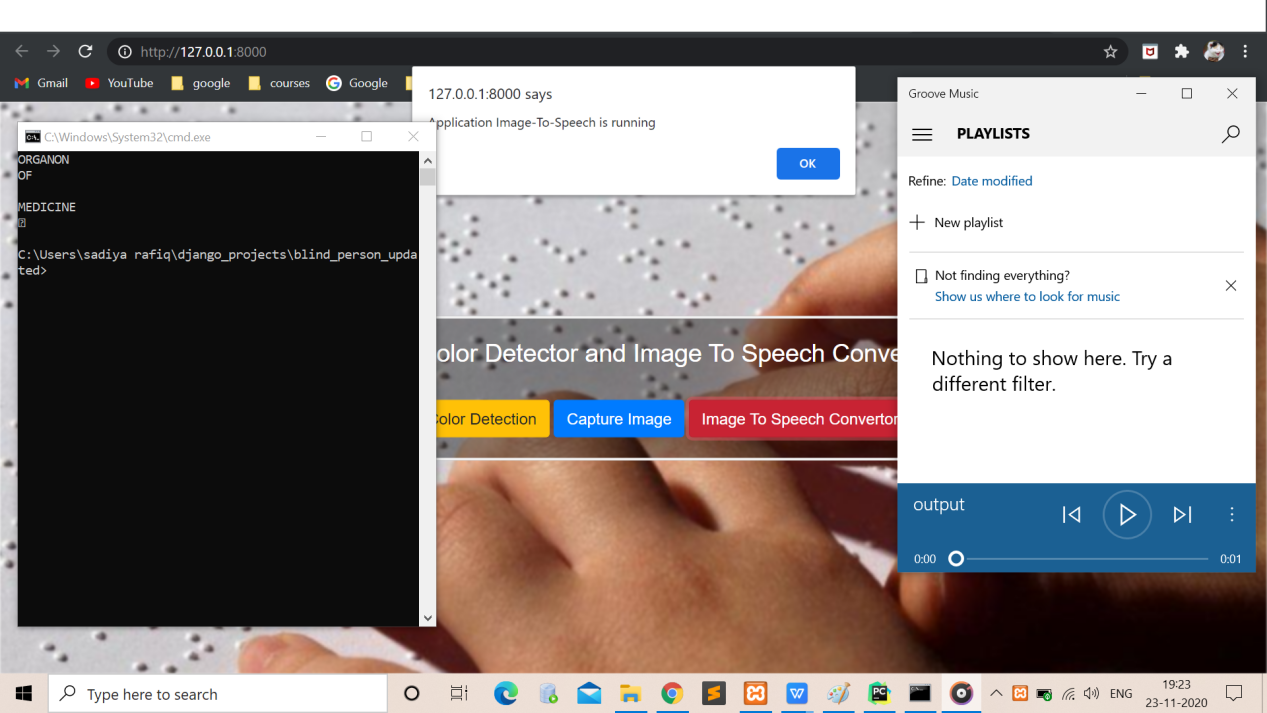
In this step the recognized text present within the image are extracted using OCR tool. during this project, we use Pytesseract library which helps to extract the recognized text.



**Text to Speech (TTS) System**

In this step the extracted text is first converted into speech using GTTS library

which is capable of converting text to speech. during this project the text to speech library is used for conversion of Text to speech.



**V. Software Requirement:**

Operating system: windows

Language: Python3.9

Library: OpenCV (python -library), Pytesseract(Python-tesseract), Python Imaging Library (PIL), gTTS (Google Text-to-Speech)

The software under which the proposed project is executed is windows 10. The algorithms are written in python language which may be a script language. OpenCV is an open source computer vision library, which is written under C and C++ and runs under multiple OS like Linux, Windows and Mac OS X. OpenCV was designed for robust consider real-time applications. OpenCV is written in optimized C and should make the foremost of multi-core processors.

**VI. CONCLUSION**

For the first stage project presentation, the required research work has been completed and also the validation of

project has been proved. The proposed system ensures to read text present within the image for assisting blind persons. Pre-processing part ensures efficient foreground extraction. But the system fails to extract the foreground after they possess a complicated background and tiny text size. An improved algorithm for background subtraction can reduce the results of complex backgrounds. The extracted text is then given to a spell corrector as OCR output isn't perfect. After getting the corrected output we send it to the TTS engine which provides a speech output. . the long run work are focused on developing an efficient portable product which will extract text from any image enabling the blind people to read text present on the products, banners, books etc.

**VIII. REFRENCES**

[1] Portable Camera Based Assistive Text and Product

Label Reading From Hand-Held Objects For Blind

Persons. Chucai Yi, Student Member IEEE, Yingli Tian,

Senior Member, IEEE, and Aries Arditi .

[2] Digital Image Processing, 3rd edition, Rafael C

Gonzalez, Richard E Woods.

[3] “Text Pre-processing and Text Segmentation for OCR.

International Journal of Computer Science Engineering

and Technology”, ARCHANA A. SHINDE, D. pp. 810-

812, 2012.

[4] ”Optical Character Recognition" International Journal

of Recent Technology and Engineering (IJRTE)”, R.

Mithe, S. Indalkar and N. Divekar. ISSN: 2277-

3878,Volume- 2, Issue-1, March 2013.

[5] "High quality text-to-speech synthesis: a comparison

of four candidate algorithms," Acoustics, Speech, and

[6]Ajantha Devi and S Santhosh Baboo, "Optical Character Recognition on Tamil Text Image Using Raspberry Pi", International Journal of Computer Science Trends and Technology (IJCST), vol. 2, no. 4, Jul–Aug 2014.

[7]Raja Venkatesan T, M. Karthigaa, P. Ranjith, C. Arunkumar and M. Gowtham, "Intelligent Transalate System for Visually Challenged People", International Journal for Scientific Research & Development (IJSRD), vol. 3, no. 12, 2016, ISSN 2321-0613.