SMART READER FOR BLIND PEOPLE

Prof.Heena Kachhela
Dept. of Computer Science & Engineering, Tulsiramji Gaikwad-Patil College of Engineering & Technology Nagpur,India, heenakachhela.@gmail.com

Ms.Veena D. Harinkhede
Dept. of Computer Science & Engineering, Tulsiramji Gaikwad-Patil College of Engineering & Technology Nagpur,India, veenaharinkhede38@gmail.com

Prof.Sneha Bhange
Dept. of Computer Science & Engineering, Tulsiramji Gaikwad-Patil College of Engineering & Technology Nagpur,India, snehab.poly@tgpcet.com

Mr.Abhishek G. Chinchole
Dept. of Computer Science & Engineering, Tulsiramji Gaikwad-Patil College of Engineering & Technology Nagpur,India, abhishekchinchole42@gmail.com

Mr.Swayam S. Satpute
Dept. of Computer Science & Engineering, Tulsiramji Gaikwad-Patil College of Engineering & Technology Nagpur,India, swayamsatpute2021@gmail.com

Ms.Sayali S. Khangar
Dept. of Computer Science & Engineering, Tulsiramji Gaikwad-Patil College of Engineering & Technology Nagpur,India, khangarsonu359@gmail.com

Abstract
The venture centers on making a report peruser for outwardly impeded people, actualized on a Raspberry Pi. The framework transforms pictures of written, transcribed, or printed text into machine-encoded text, changing over them into sound . The transformation of printed records into textual files through conversion. is accomplished through the utilization of both the Tesseract library and Python programming. This combination of tools enables the device to effectively transform physical text into digital formats, enhancing accessibility and allowing for further processing or utilization of the extracted content .The ensuing preparing of content records and era of sound yield are encouraged utilizing OpenCV and Python. The operation of this device relies on a Raspberry Pi embedded board, equipped with a Pi camera. The actual or real-time capturing of banknotes is performed and taken care of by means of particular picture managing with methodologies like edge region, division and Integration of feature extraction and classification is a key aspect of this project. In this envisioned endeavor, the classification algorithms. These calculations are actualized in Python lingo with OpenCV on raspberry pi equipment stage. Moreover, another perspective of this venture includes a comprehensive machine learning approach, joining OpenCV and profound learning strategies to address the challenge of address location in an end-to-end mold. The demonstrate is prepared on a challenging freely accessible dataset, chosen for its trouble and significance, serving as the yearly testing ground for an protest localization competition. The coming about system illustrates exceptional speed and accuracy, rendering it profoundly profitable for applications requiring vigorous address location capabilities.

I. Introduction

The Smart Reader for Blind People encompasses a spectrum of features designed to enhance the overall user experience. Beyond mere accessibility, this innovation emphasizes the importance of inclusivity, ensuring that individuals with visual impairments not only access information but do so with 'efficiency' and 'comfort'. The device incorporates advanced natural language processing algorithms, contributing to a more nuanced and human-like rendition of text, fostering a more immersive and enjoyable reading experience.

In the realm of accessibility technology, the Smart Reader for Blind People emerges as a revolutionary innovation. Leveraging the computational power of Raspberry Pi, this revolutionary device addresses a significant challenge faced by visually impaired individuals – the ability to access and comprehend printed text. By seamlessly converting written content into spoken words, this reader empowers individuals with visual impairments to engage with literature and information in a manner that was once elusive.

The portability of the Smart Reader for Blind People further amplifies its impact, allowing users to seamlessly integrate this technology into their daily lives. Whether at home, in the classroom, or on the go, individuals with visual impairments can carry this compact yet powerful device, breaking down geographical barriers to information and education. This mobility underscores the commitment to fostering independence, ensuring that users have unhindered access to a world of knowledge at their fingertips.

Automated recognition of objects in pictures may be essential step for numerous real-world computer vision applications. It is the assignment of finding a given object in an picture or video arrangement without or slightest human intercession/ help. As we know, exceptionally small exertion is required at our portion to distinguish and recognize gigantic number of classes of objects in pictures in spite of the fact that picture of the question may be distinctive with respect to measure / scale, perspective position or orientation.

Optical character acknowledgment came into presence of computer vision, written by hand information on its essential exchanging instrument. The outlined information basically utilized for the daze as a development. Outwardly impeded individuals discover it hard to identify the littlest detail with sound eyes.

The reason behind this may root to the reality that "Object Recognition" requires understanding of human visual discernment and so gets to be multidisciplinary inquire

about zone including information inability of areas li ke

optics, brain, design acknowledgment, manufactured insight,machine learning and most vitally cognitive science which in itself needs modern concepts and apparatuses from arithmetic as well as computer science.

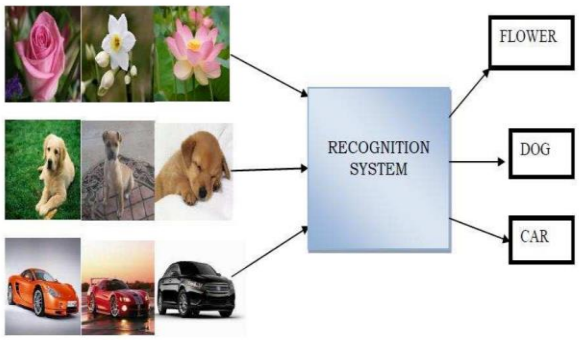


Fig 1. Generic Object Recognition

A. *Methodology*

1) ***Datasets****:*

In the realm of developing algorithms for Natural Language Processing (NLP), the dataset assumes a pivotal role, serving as the foundational material for the Raspberry Pi-based reader to learn and comprehend written text. The chosen dataset is akin to a diverse library, enabling the Raspberry Pi to adapt and handle various language types, expressions, and contexts.

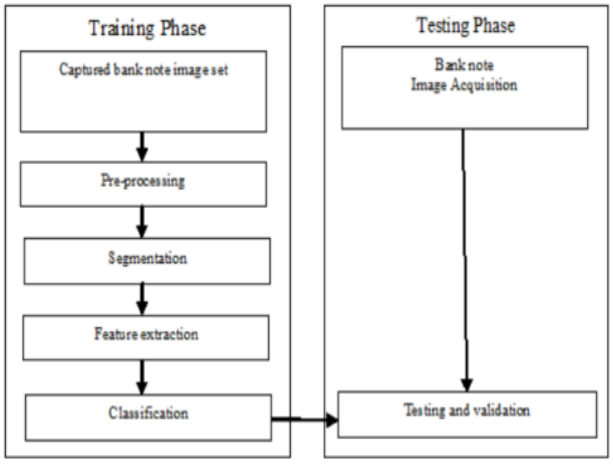


Fig 4. Schematic Overview of the Planned Approach

During the instructional phase, a variety of pictures is taken, featuring diverse notations and angles to represent a comprehensive dataset for banknote training. This varied collection of images helps in training the system to recognize and process banknotes accurately under different conditions and perspectives.

2) ***Procedural Design***

In procedural design, the initial step involves a visually impaired person capturing an image through the camera. Once captured, the image is stored in the system, proceeding to the pre-processing stage to facilitate object identification. Following the identification process, it becomes essential to calculate the distance between the identified object and the blind person. This step is crucial to ensure that only one object is present in front of the blind person, considering the possibility of multiple objects. The calculation of distance becomes imperative in scenarios where more than one object may be present.

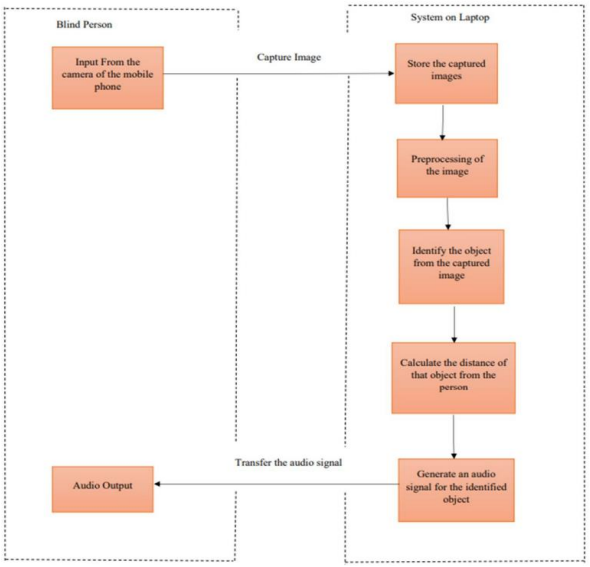


Fig 5. Procedural Design

B. System Design And Development

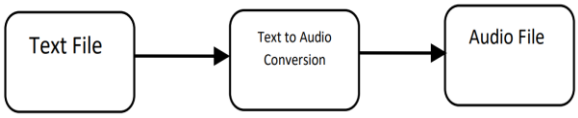
The undertaken work involved the utilization of diverse methods for processing images, encompassing capturing images, identifying edges, segmenting, categorizing, and extracting features. This comprehensive set of methods contributes to the through processing and analysis of captured images, ensuring the efficient recognition and extraction of pertinent information.

a) ***Image Capturing:***

The crucial step involves placing the document under the camera, which then takes a picture of the document. The image's high quality is ensured for swift and accurate identification.

b) ***Text to Speech:***

This module's scope commences with the completion of the Character Recognition module. It is responsible for converting the recognized content into an audible form. The Raspberry Pi incorporates an on-board sound jack, generating on-board sound through a PWM output that is minimally filtered. However, the use of the utilization of a USB sound card can greatly improve both the quality and volume of the sound output.



c) ***Image to Text Converter***

The ASCII values assigned to the characters that have been identified undergo handled by the Raspberry Pi board. Every character is paired with its corresponding arrangement and subsequently saved as normalized text translation. This translated text is subsequently conveyed to the Sound output, enabling the conversion of recognized characters into audible information.

d) ***Software Implementation****:*

Operating system: “Raspbian”.

Language: “Python”.

The proposed project operates under the Raspbian operating system. The algorithms for the project are scripted using the Python language, a versatile scripting language known for its readability and ease of use.

Fig 6. Raspberry pi

e) ***Image acquisition:***

The image acquisition process involves capturing pictures using an advanced 8MP Raspberry Pi camera. The images are initially obtained in the RGB format and subsequently transformed into grayscale. Figure 2 showcases diverse samples of input banknote images, featuring denominations including all Indian rupee notes.

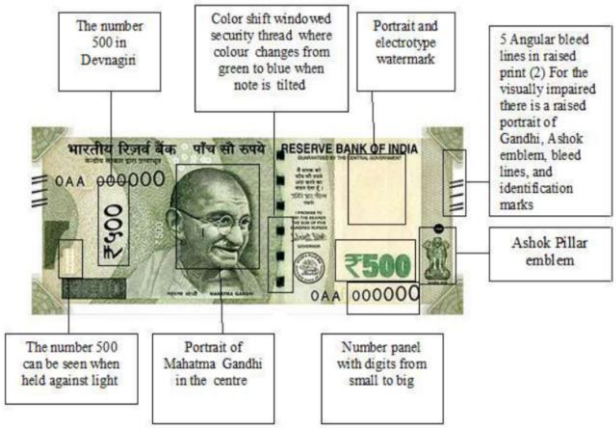


Fig 2. Front side of 500 rupees note

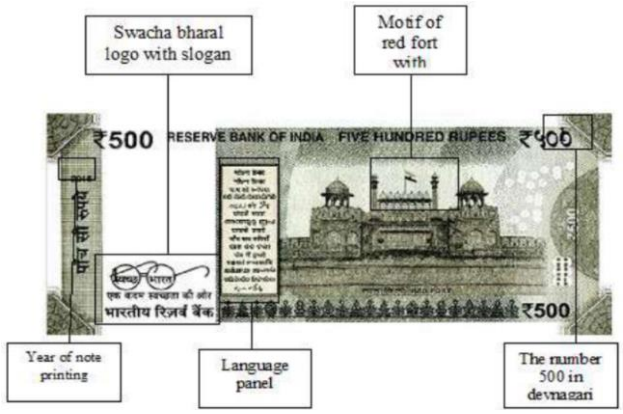


Fig 3. Back side of 500 rupees note

f) ***Pre-processing:***

To improve image quality and eliminate noise, Gaussian blurring is applied. The Gaussian function in two dimensions is expressed by equation 1. This process contributes to enhancing the clarity and overall quality of the images.

$$G(x,y)=\frac{1}{2\pi\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}$$

g) ***Segmentation:***

In order to separate the currency document from the backdrop, the image is transformed into a binary format. Canny edge and Sobel detection techniques are then implemented to extract distinct regions, as outlined in the mathematical equation. This process entails calculating the intensity gradient of the image. Subsequently, the smoothed image undergoes separation using a Sobel filter in both the horizontal and vertical directions, yielding the first

derivative along the horizontal and vertical axes .These derivative images provide information about edge gradient and direction for each pixel, facilitating further analysis and extraction.:

$$Edge\,Gradient(G) = \sqrt{G_x^2 + G_y^2}$$

$$Angle(\theta) = \tan^{-1}\left(\frac{G_y}{G_x}\right)$$

h) ***Feature extraction:***

Feature extraction is a specialized method of reducing dimensionally that is utilizes to encapsulate the crucial visual information inherent in an image.This process is employed for various purposes, including image retrieval and indexing. When dealing with excessively large input data that contains abundant but not necessarily informative details, the algorithm may become impractical to handle.

In such cases, the input data is transformed into a condensed set of features, simplifying the representation and reducing the computational resources needed to describe the extensive dataset. Feature extraction serves the purpose of streamlining the data representation, making it more manageable and efficient.

S.No.	Feature Specifications
1.	Vertical dimension of the pictures
2.	Horizontal dimension of the pictures
3.	Euler number of pictures
4.	Comparing the image to a template image using correlation analysis

Table 1. Feature specifications

II. Working Principle

Upon pressing the capture button, our system initiates the process of capturing a visual representation of the document positioned in front of the camrea . This camera is connected to an ARM microcontroller through a USB interface, facilitating seamless communication and data transfer between the imaging device and the processing unit. After capturing the image, activating the process button triggers the implementation of Optical Character Recognition (OCR) technology.OCR enables the conversion of scanned picture containing printed text or symbols to editable and understandable text using a computer program.

After OCR, the extracted text undergoes conversion to audio using a Text-to-Speech library. The camera serves as the primary visual input for detecting the document image. Internally, the system processes the image using the OpenCV library, separating the label from the image, and then identifies the text , which is subsequently pronounced through voice.

Fig 5.3 Project module

Step 1: Connect the camera , speaker, Adapter and any display with raspberry pi.

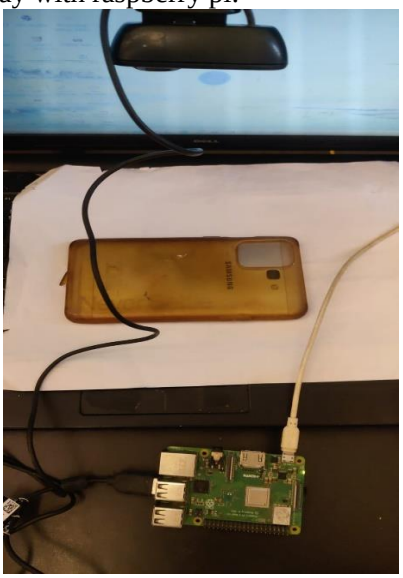


Fig 5.4 Object Recognition

Step 2:After connecting turn on the power supply and start the pi.

Fig 5.5 Raspberry pi Home Screen

Step 3: After starting the pi open the Thonny IDE and install all the required libraries.

1. sudo apt-get update
2. sudo apt-get upgrade
3. sudo apt-get install build-essential cmake pkg-config libjpeg-dev libtiff5-dev libjasper-dev libpng-

dev libavcodec-dev libavformat-dev libswscale-dev libv4l-dev libxvidcore-dev libx264-dev libfontconfig1-dev libcairo2-dev libgdk-pixbuf2.0-dev libpango1.0-dev libgtk2.0-dev libgtk-3-dev libatlas-base-dev gfortran libhdf5-dev libhdf5-serial-dev libhdf5-103 python3-pyqt5 python3-dev -y

4. pip3 install opencv-python==4.5.3.56
5. sudo apt-get install python-opencv
6. pip install -U numpy
5. sudo apt-get install tesseract-ocr
6. pip3 install pytesseract
7. pip3 install pyttsx3

Step 4: After installing all the required libraries, run the code.

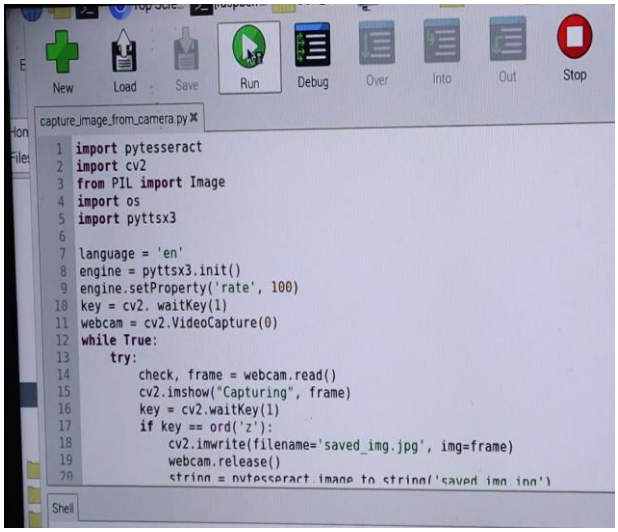


Fig 5.6 Execution

Step 4: After running the code, the camera starts capturing the image.

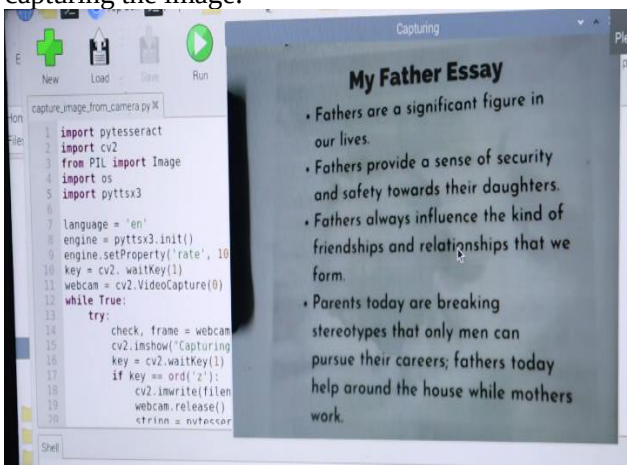


Fig 5.7 Image Capturing

Step 5: After the capturing is complete the system extracts the text from the image and starts reading it.

Step 6: Similarly , also by executing the object recognition code, the system opens the camera and start recognizing the different objects.

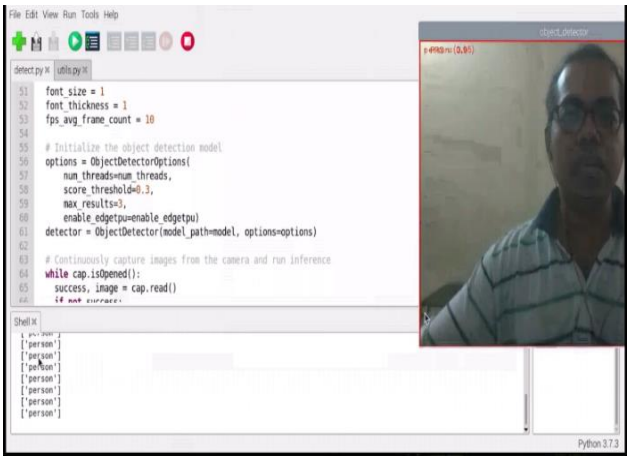


Fig 5.8 Object recognition

Step 7: The output is given through speaker.

III. Conclusion

We've created an exceptionally precise object detection system that rivals the performance of top-tier solutions, employing state-of-the-art methodologies in computer vision and advanced deep learning techniques. Our project involves the creation of a specialized dataset through image labeling, ensuring a reliable evaluation process. Notably, achieving these impressive results necessitated additional computing power during the training phase.

Additionally, we have designed a Text-to-Speech device that efficiently converts text images into sound, boasting a high-performance level and maintaining a readability tolerance of less than 2%. This portable device operates independently, without requiring an internet connection, and can process A4-sized documents in under three minutes. It serves as a valuable tool for individuals, simplifying the editing process for books or web pages and facilitating easy enhancement and modification of content and additionally , we have designed a device that identitfies the different numbers of currencies which is useful to visually challenged people.

Acknowledgments

The development and realization of the Smart Reader for Blind People have been a collaborative effort fueled by the commitment and dedication of numerous individuals and organizations. We extend our deepest gratitude to the brilliant minds behind the creation of the Raspberry Pi, whose innovative technology serves as the cornerstone of our groundbreaking reader. Additionally, we acknowledge the invaluable contributions of researchers, engineers, and accessibility advocates who have tirelessly worked to enhance the functionality and user experience of the device.

Our heartfelt appreciation goes out to the visually impaired individuals who actively participated in

user testing and provided invaluable insights, shaping the reader into a truly user-centric innovation. Their feedback has been instrumental in refining the design and ensuring that the device meets the diverse needs of the visually impaired community.

We express gratitude to the academic and research institutions that have supported this project, providing resources and fostering an environment conducive to innovation. Furthermore, we acknowledge the ongoing collaboration with organizations dedicated to promoting accessibility and inclusivity, as their shared commitment has propelled the Raspberry Pi Based Reader for the Blind to the forefront of assistive technology.

References

1. William A. Ainsworth A system for converting English text into speech IEEE Transactions on Audio and Electroacoustics, Vol. 21, Issue 3, Jun 1973.
2. Dimitrios Dakopoulos and Nikolaos G.Bourbakis Wearable Obstacle Avoidance Electronic Travel Aids for Blind IEEE Transactions on systems, man and cybernetics,Part C (Applications and Reviews). Vol. 40, issue 1, Jan 2010.
3. Larisa Dunai Dunai, Mónica Chillarón Pérez, Guillermo Peris-Fajarnés, Ismael Lengua, Euro Banknote Recognition System for Blind People Sensors, ISSN 1424-8220,2017.
4. Faiz M. Hasanuzzaman, Xiaodong Yang, and YingLi Tian, Senior Member, IEEE, Robust and Effective Component-Based Banknote [3] Recognition for the Blind, IEEE transactions on systems, man, and cybernetics—part c: applications and reviews, vol. 42, no. 6, november 2012.
5. Ms.AthiraPanicker Smart Shopping assistant label reading system with voice output for blind using raspberry pi, Ms.Anupama Pandey, Ms.Vrunal Patil YTIET, University of Mumbai ISSN: 2278 – 1323 International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Vol. 5, Issue 10, Oct 2016 2553