*Saii Maadhav Madderla, et. al. International Journal of Engineering Research and Applications*
*www.ijera.com*
*ISSN: 2248-9622, Vol. 12, Issue 6, (Series-II) June 2022, pp. 07-13*

RESEARCH ARTICLE OPEN ACCESS

# AI and ML based Product label and text reading System

Saii Maadhav Madderla
*Department of Electronics and Communication Engineering*
*NallaMalla Reddy Engineering College, JNTUH*
*Hyderabad, Telangana*

Kaushik Kumar K N
*Department of Electronics and Communication Engineering*
*NallaMalla Reddy Engineering College, JNTUH, Hyderabad, Telangana*

SatyaDurga P
*Department of Electronics and Communication Engineering*
*NallaMalla Reddy Engineering College, JNTUH, Hyderabad, Telangana*

Dr.B.KhaleeluRehman
*Associate Professor, Department of Electronics and Communication Engineering*
*, NallaMalla Reddy Engineering College, JNTUH, Hyderabad, Telangana*

**ABSTACT**
An idea that helps the physically challenged persons specifically the Blind in identifying the product or the text present on the product with the help of AI and ML based product label and text reading System.
The main aim of this paper is to design a system for the blind people to recognise the handheldproduct and its labels or any text that notifies it with voice output. So, as they can't see or read any text this designed system will enhance their ability to identify the products. The Machine Learning algorithms and Artificial intelligence is used for the predictionsand know the products for improvising the performance of the system.
This portable system which performs the tasks is RaspberryPi processor which is linked with the Pi-camera and the Headsets. These are the three main devices in the system.The Processor takesresponsibility to check the object name details and tells the blind person through voice using aheadset.The text characters are recognised by Optical Character Recognition (OCR).
Therefore, by above mentioned mechanism, developing this prototype system will help the Blind.
*Keywords: Raspberry Pi, Pi-Camera, OCR, Artificial intelligence, Machine Learning*

---

Date of Submission: 28-05-2022 Date of Acceptance: 10-06-2022

---

## I. INTRODUCTION

The main motive behind this work is to benefit the Blind people and giving them the equal importance in the society. As they are also much interested in knowing the worldly things and also to learns different things which encounters them in their daily life.Total number of visually effected people is 2.2 billion [1] among them, 36 million are completely blind, and the rest of the 1 billion have reasonable to fatiguing vision impairment. Regrettable fact is that the 80% of them are located in the developed cities where no good provisions are provided for their disability. We are walking ahead with greatest advancements in the technology but ignoring the fact of leaving behind the disabled society without any advancements in technological supports for their well being and making them feel same as us in the community.

The system has been designed to the visuallyimpaired people to identify the products and give them the voice announcement to recognize better through voice. The technologies used for making the system were Machine Learning and Artificial Intelligence and these two plays a major role in recognizing the product Label. The AI is the technology which makes the system to think, behave and react similar to the humans, where as this will give the user a feel good experience while operating.
**Chucai Yi, YingliTian and Aries Arditi** [2] proposed Portable Camera-Based Assistive Text and Product Label Reading from Hand-Held Objects for

Blind Persons that totally focus recent developments in computer vision, digital cameras, and portable computers make it feasible to assist these individuals by developing camera-based products that combine computer vision technology with other existing commercial products such optical character recognition (OCR) systems.

**Sunil Kumar, Rajat Gupta, NitinKhanna, SantanuChaudhury and Shiv Dutt Joshi** [3] proposed Text Extraction and Document Image Segmentation Using Matched Wavelets and MRF Model. This paper proposes the extraction of textual areas of an image using globally matched wavelet filters. A clustering-based technique has been devised for estimating globally matched wavelet filters using a collection of ground truth images and text extraction scheme.

**Patrick E. Lanigan, Aaron M. Paulos, Andrew W. Williams, Dan Rossi and PriyaNarasimhan**[4] proposed Trinetra: Assistive Technologies for Grocery Shopping for the Blind. This paper describes about the research and development of the Trinetra system, a barcode-based solution comprising COTS components, such as an Internet- and Bluetooth-enabled cell phone, text-to-speech software and a portable barcode reader.

**KANDAGATLA ROOPA, P. SURESH KUMAR, G. RAVI KUMAR** [5] proposed Portable Camera Based Assistance Label Reading for Blind Person. This paper proposes to isolate the object from chaotic backgrounds or objects in. They have firstly proposed an efficient and effective motion based method to define a region of interest (ROI). This method extracts moving object region by a mixture-of-Gaussians-based background subtraction method. In the extracted ROI, text localization and recognition are conducted to acquire text information.

**KwangIn Kim, Keechul Jung, and Jin Hyung Kim** [6] proposed Texture-Based Approach for Text Detection in Images Using Support Vector Machine (SVM) and Continuously Adaptive Mean Shift Algorithm. This paper show texture-based method for detecting texts in images

## II. PROPOSED SYSTEM

Theproposed system is based on Pi-camera which is interfaced with the Microprocessor (Raspberry-Pi3). An earphone which gives the voice announcement out and a SD card which is used as a hard drive to the processor. All the pre-captured images i.e. Datasets is stored in it. The camera acts as a main vision in detecting the product label and text which is placed in front of it. The OCR algorithm is used for text recognition.

2.1 Block Diagram:

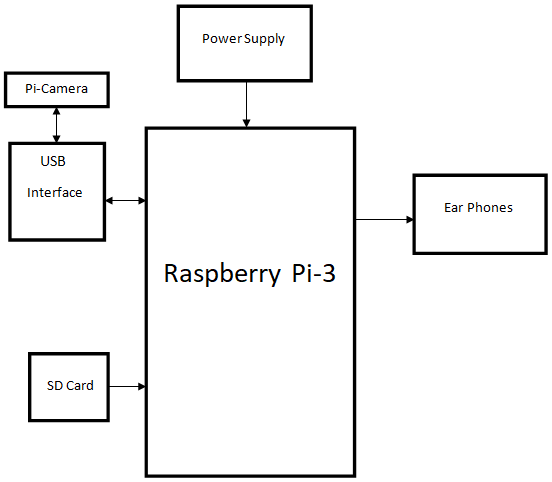


Fig.2.1 Block Diagram of the system

Figure 2.1 shows the Block diagram of the proposed system. The blocks involved in the proposed system are listed below.
2.1.1 Microprocessor (Raspberry-Pi 3)
2.1.2 Pi-Camera
2.1.3 SD card
2.1.4 Earphones/Bluetooth Headsets

2.1.1 Raspberry Pi-3:

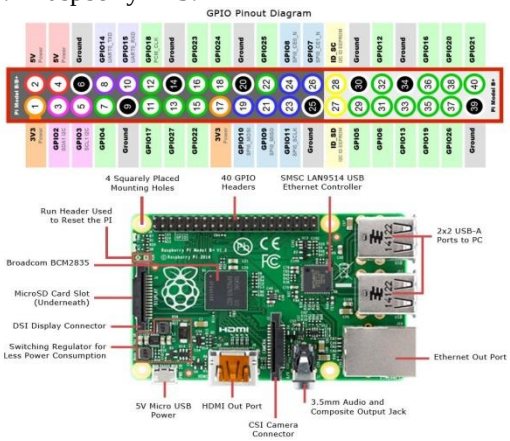


Fig.2.1 Raspberry pi 3 board

Fig. 2.1 shows Raspberry pi 3 boards. Raspberry Pi uses ARM1176JZF-S which has a frequency of 1.2 GHz 64-bit quad core ARM processor, RAM is 1 GB, 802.11n wireless LAN, Bluetooth 4.1 and Bluetooth Low energy.

This ARM-11 processor is as same as that of a PC, this can be easily connected to any monitor or TV or the keyboard and use it as a Mini PC. This Processor is in the size of ID card or credit card where we can easily carry and use it with multiple devices.

Specifications:

- 1GB RAM
- 4 USB ports
- 40 GPIO pins
- Full HDMI port
- Ethernet port
- Combined 3.5mm audio jack and composite video
- Camera interface (CSI)
- Display interface (DSI)
- Micro SD card slot (now push-pull rather than push-push)
- Video Core IV 3D graphics core

2.1.2 Pi-Camera:

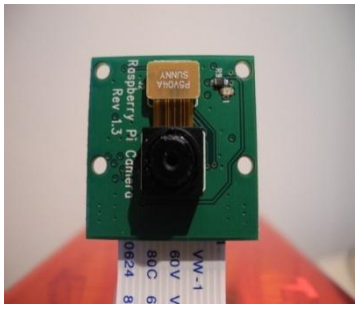


Fig. 2.2 Pi-camera

Fig.3.3 shows the Pi-camera. The Raspberry pi has declared a camera component which is compatible to its microprocessor in the year 2013. This Pi-camera is connected with a ribbon cable which is connected to the processor board. The size of the small circuit board is of 25mm/20mm/9mm.

Specifications:

- Fully Compatible with Both the Model A and Model B Raspberry Pi
- 5MP Omni vision 5647 Camera Module
- Still Picture Resolution: 2592 x 1944
- Video: Supports 1080p @ 30fps, 720p @ 60fps and 640x480p 60/90 Recording
- 15-pin MIPI Camera Serial Interface - Plugs Directly into the Raspberry Pi Board
- Size: 20 x 25 x 9mm
- Weight 3g
- Fully Compatible with many Raspberry Pi cases

2.1.3 SD card:

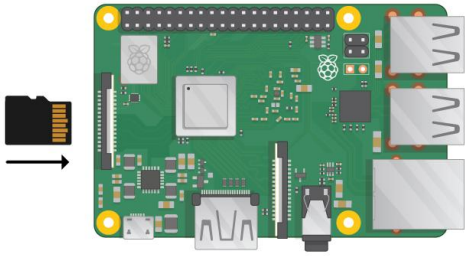

Fig.2.3 SD card insertion

Fig.2.3 shows the SD card insertion .The MMC or SD card is a flash memory card. A memory card or flash memory card is solid-state electronic flash memory data storage device capable of storing digital contents.

This SD card is a main part in the raspberry pi as it stores the datasets that is previously captured by the pi-camera and also stores other information. It acts as a hard disk to this processor.

2.1.4 Earphones:


Fig.2.4 Earphones plugging slot on the board

Fig 2.4 shows the Earphones plugging slot on the board. The Mp3 slot is present in the Raspberry processor which we use it to connect the headphones/Earphones for the voice output from the system.

## III. WORKING OF THE SYSTEM

3.1 Image Acquisition:

When any product or the text is kept in front of the Pi-camera the image is detected by the camera and captured. Thus, this would be the first step of working process of our proposed system. The Fig.4.1 shows the image captured by the Pi-camera.


Fig.3.1 Image captured by the Pi-camera

3.2 Gray scale conversion:

After the process of capturing the image through Pi-camera, the captured image is converted to gray using gray scale conversion technique where it is initialized by the python libraries. The main purpose of doing this is to get the enhanced and exact information of image for the later process. The Fig.4.2 shows the gray scale conversion of the image acquired by the pi-camera.


Fig 3.2 Gray scale converted image

3.3 Text Extraction:

Automatically text extraction algorithm is implemented to detect the text present on the Product label. The edge detection method or Method of blob detection in image is used for handling the complex backgrounds and identify it, as the system should understand what to read out and what should be considered as background or Co-variant region in a captured image. The Fig.4.3 shows the Text extraction from a complex image background.


Fig.3.3 Text extraction from a Complex image background

3.4 Optical Character Recognition:

The OCR is used to make the acquired image or text to convert it into machine-encoded text, to make the system understand and give the voice output. The output of the OCR is nothing but a text file containing the product label in text form. The audio output component (earphones) is to inform the blind users to recognize the text code in the form of speech or Audio. This Fig.4.4, 4.5 shows the pictures of the images detected and printed out by using OCR



Fig.3.4, 3.5 Pictures of the images detected and printed out by using OCR

## IV. FLOWCHART AND ALGORITHM

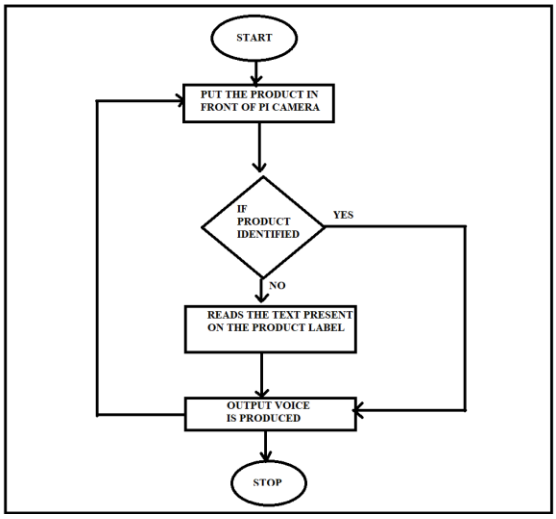


Fig 4.1 Flowchart and algorithm of the proposed code

Step-1: Startof the algorithm
Step-2: Put the product in front of pi camera
Step-3: If the product is identified. The go to step-5. Else go to step-4.
Step-4: The camera tries to read the text present on the product. Go to step-5.
Step-5: Vocal announcement is produced through audio jack. To know the name of another product go tostep-2 or directly to step-6.
Step-6: Stop of the algorithm

### 4.1.1Explanation of Algorithm

The input for the proposed system is captured using the pi camera. When the product is kept in front of the pi camera, it captures the image and sends it to the raspberry pi. The raspberry pi processor the image and the check whether the image is present in the dataset. If the image is not present in the dataset, then the system tries to read the text present on the product label.

## V. DATASETS

The datasets used for the system are the pictures taken previously by the Pi-camera of different products and its labels. The various pictures are stored in a folder in the microprocessor's hard disk (i.e. SD card) and when the code is made to run, it analyses whether the captured picture is present in the dataset or not, only then the decision is taken to give the voice output of either the name of the product by recognizing it from the dataset or to read out the text present on the product


Fig. 6.1, 6.2, 6.3some of the pictures from the Dataset

## VI. RESULTS AND CONCLUSIO

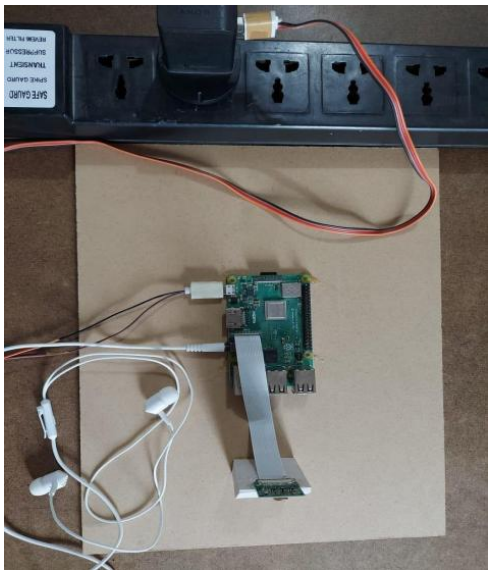

**Fig.7.1**Image of the system when it is switched OFF

Fig.7.1 shows Image of the system when it is switched OFF. This prototype model is shown when all the physical components are integrated and connected but not switched ON.

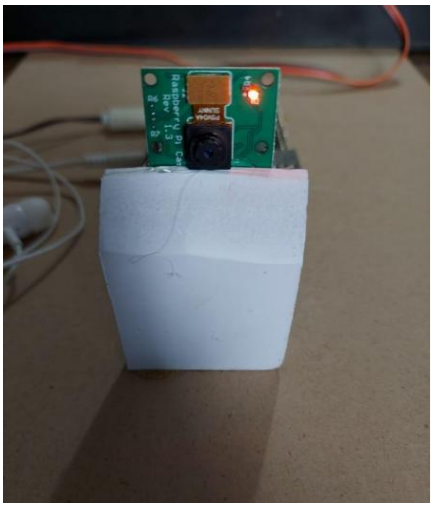

**Fig.7.2**Image of the system when switch is ON

Fig.7.2 shows the Image of the system when switch is ON. The prototype model is shown when all the physical components are integrated and connected when power is ON. The power ON and OFF is indicated by using a LED as shown in above figure. And also when the object is detected by the Pi-camera the LED light blinks for 1.4 seconds indicating the image is captured.

**Fig.7.3** Imageof the system when it is produces the output

Fig.7.3 shows the Image of the system when it produces the output. The output of this system is in the form of voiceannouncement through audio jack. The blind people can connect the wireless Bluetooth or wired Headsets.

## VII. CONCLUSION

This paper concludes to serve the blind with their disability to read or identify the products; this system will help them in achieving it. By this, the text recognizing is done and the text code converts it into the speech. Hence, Artificial Intelligence and Machine Learning models are used for detecting the different images.

## ACKNOWLEDGEMENT

The authors would like to thank the R&D Lab, Department of Electronics and Communication Engineering, NallaMalla Reddy Engineering College, Hyderabad, for providing cooperation and laboratory facility to carry out our research work.

## REFERENCES

[1]. Steinmetz, J. D., Bourne, R. R., Briant, P. S., Flaxman, S. R., Taylor, H. R., Jonas, J. B., ... & Morse, A. R. F. (2021). Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study. The Lancet Global Health, 9(2), e144-e160.

[2]. Yi, C., Tian, Y., &Arditi, A. (2013). Portable camera-based assistive text and product label reading from hand-held objects

for blind persons. IEEE/ASME Transactions On Mechatronics, 19(3), 808-817.

[3]. Kumar, S., Gupta, R., Khanna, N., Chaudhury, S., & Joshi, S. D. (2007). Text extraction and document image segmentation using matched wavelets and MRF model. IEEE Transactions on Image Processing, 16(8), 2117-2128.

[4]. Lanigan, P. E., Paulos, A. M., Williams, A. W., Rossi, D., &Narasimhan, P. (2006, October). Trinetra: Assistive Technologies for Grocery Shopping for the Blind. In ISWC (pp. 147-148).

[5]. ROOPA, K., KUMAR, P. S., & KUMAR, G. R. (2016). Portable Camera Based Assistance Label Reading for Blind Person.

[6]. Kim, K. I., Jung, K., & Kim, J. H. (2003). Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm. IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(12), 1631-1639.

Saii Maadhav Madderla, et. al. "AI and ML based Product label and text reading System." *International Journal of Engineering Research and Applications (IJERA)*, vol.12 (06), 2022, pp 07-13.