B.A.R.F

Blind Abled Reading Frame

An Internal Funded Project Report

Submitted By

Kshitij Sharma	185001080-III Year
Kiruthika J	185001078-III Year
Raghav R	185001119-III Year

COMPUTER SCIENCE AND ENGINEERING

SSN COLLEGE OF ENGINEERING

KALAVAKKAM 6030110

SSN COLLEGE OF ENGINEERING

CHENNAI 603110

BONAFIDE CERTIFICATE

This is to certify that the project report titled **BLIND ABLED READING FRAME** is the bonafide work of Kshitij Sharma (185001080 -III Year), Kiruthika J (185001078 -III Year) and Raghav R (185001119 -III Year), who carried out the project work under the supervision of our supervisor as an internal funded project in the department of CSE .

Dr. D.Thenmozhi
Supervisor,
Associate Professor,
Department of CSE,
SSNCE

Dr. Chitra Babu
Head of the Department,
Professor,
Department of CSE,
SSNCE

Place:
Date:

ACKNOWLEDGEMENTS

I would like to thank and deep sense of gratitude to my guide **Dr. D.Thenmozhi**, Associate Professor, Department of Computer Science and Engineering, for her valuable advice and suggestions as well as his continued guidance, patience and support that helped me to shape and refine my work.

My sincere thanks to **Dr. CHITRA BABU**, Professor and Head of the Department of Computer Science and Engineering, for her words of advice and encouragement and I would like to thank our project Coordinator **Dr. T. T. MIRNALINEE**, Professor, Department of Computer Science and Engineering for her valuable suggestions throughout this first phase of project.

I express my deep respect to the founder **Dr. SHIV NADAR**, Chairman, SSN Institutions. I also express my appreciation to our **Dr. V.E. ANNAMALAI**, Principal, for all the help he has rendered during this course of study.

I would like to extend my sincere thanks to all the teaching and non-teaching staffs of our department who have contributed directly and indirectly during the course of my project work. Finally, I would like to thank my parents and friends for their patience, cooperation and moral support throughout my life.

Kshitij Sharma **Kiruthika J** **Raghav R**

ABSTRACT

B.A.R.F is a cognitive frame that helps the visually impaired, with these glasses they can expand their horizon and do things that they have always wanted. B.A.R.F is easy to use and has assistant called Jarvis which runs the command the users ask for along with the ease-of-use B.A.R.F is very comfortable to wear.

B.A.R.F performs three functions as of now which are reading any sort of text, this varies from and printed text (books, magazines etc.) to handwritten text. The next feature is to detect objects in front of the user thereby helping them to perceive objects around them. The last feature is of detecting currency this helps them to pay people and not rely on others. All these features are available in any language the user wants.

B.A.R.F uses a raspberry pi for processing and a raspberry pi camera for taking pictures. It also consists of a go-pro head strap which is used by vloggers for comfortability.

Overall, B.A.R.F is user friendly and will change the lives of many people and let them live a comfortable life which they deserve.

Table Of Content

INTRODUCTION

Globally, around 2.2 billion people suffer from vision impairment or blindness. Despite this, it is estimated that around half of sight loss is actually avoidable.

Vision impairment is classified into two groups; distance and near presenting vision impairment. Each person's experience of sight loss may vary on a number of factors, including availability of prevention and access to rehabilitation. There are particular groups in society that are at higher risk than others when it comes to vision problems. For instance, those later in life or children who were born prematurely.

The two main causes of sight loss are uncorrected refractive errors and cataracts. This is closely followed be age-related macular degeneration (AMD), glaucoma and diabetic retinopathy.

Approximately 230 million people have severe visual impairment and this number is predicted to increase in the coming years. This leads them to be dependent on another individual but not everyone has someone to help them out every time, so as to help them and not make them dependent on other people we have developed B.A.R.F.

B.A.R.F is cognitive frame that can help the visually disabled to read and live a better life. It aims at removing the use of fingers to read a book and expand the options available for the blind to read books and help them to perceive objects around them. B.A.R.F is portable, light weight and has a long battery life. It has an inbuilt assistant called Jarvis which runs the respective program asked by the user. B.A.R.F recognizes multiple languages and can perform text to speech conversion to any desired language.

HARDWARE

Raspberry Pi 4: We are using a 2gb ram raspberry pi for the processing and ease of use. Due to the compact design of the pi its easily portable and can be powered by any 5v power source such as power banks. We have place it in a pi case for support and protection.

Raspberry pi with case.

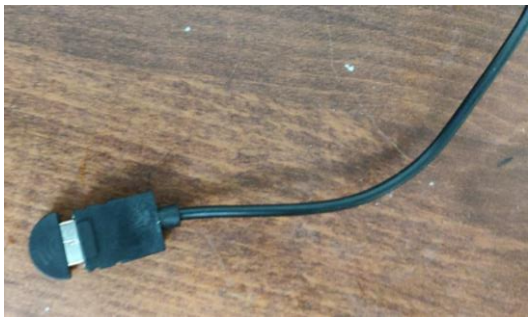
Raspberry pi camera: The raspberry pi camera has a resolution of 1080p and the focus can be adjusted manually the camera also has an infrared sensor so as to detect objects at low light. It is mounted on a base frame for support. The camera is connected to the pi via the flex cable.

Raspberry pi camera mounted on the frame.

Go-Pro Head Strap: We have used this head strap as it is professionally used by vloggers for shooting videos and supports long term use with comfortability.

Go-Pro head strap

Voulmio Mic: This USB mic is chosen due its small size and compatibility with the raspberry pi. It is attached via a USB pin extender and has a great sound quality.



USB Mic with extender

Power Supply: The main advantage of this glasses is that you can use any sort of power supply available at home such as power banks or battery packs.

Headphones: The user can use any available headphone either via Bluetooth or with an audio Jack which can be directly connected to the raspberry pi.

SOFTWARE

Modules:

gTTS: gtts(*Google Text-to-Speech*), a Python library and Command line Interface tool to interface with Google Translate's text-to-speech API(Application programming interface). It is used to write spoken mp3 data to a file. It can also be used to write to a file-like object (bytestring) for further audio manipulation, or stdout. Or simply pre-generate Google Translate TTS request URLs to feed to an external program.

google.cloud-vision: The google cloud vision API enables us to understand the content of an image by encapsulating powerful machine learning models. It quickly classifies images into thousands of categories (e.g., “sailboat”, “lion”, “Eiffel Tower”), detects individual objects, and finds and reads printed words contained within images.

pygame: It is a free and open-source cross-platform library for the development of multimedia applications like video games using Python. It is used here as an audio library to play the audio file generated .

Speech Recognition: This is a library for performing speech recognition, with support for several engines and APIs, online and offline.
The Speech Recognition library acts as a wrapper for several popular speech APIs and is thus extremely flexible.

This can be used to avoid writing and building scripts separately for accessing microphones and processing audio file from scratch.

Mutagen: It is a Python module to handle audio metadata. It supports ASF, FLAC, MP4, Monkey's Audio, MP3, Musepack, Ogg Opus, Ogg FLAC, Ogg Speex, Ogg Theora, Ogg Vorbis, True Audio, WavPack, OptimFROG, and AIFF audio files.
Mutagen works with Python 3.5+ (CPython and PyPy) on Linux, Windows and macOS, and has no dependencies outside the Python standard library.
This library allows to extract data about the audio file.

Tensorflow.keras: TensorFlow is an open source software library for high performance numerical computation. Its flexible architecture allows easy deployment of computation across a variety of platforms. It comes with strong support for machine learning and deep learning and the flexible numerical computation core is used across many other scientific domains.The model to detect currency was trained using this module.

Pillow(PIL): This Python Imaging Library adds image processing capabilities to your Python interpreter. It also provides extensive file format support, an efficient internal representation, and fairly powerful image processing capabilities.
The core image library is designed for fast access to data stored in a few basic pixel formats. It should provide a solid foundation for a general image processing tool.

Numpy: It is a fundamental package for array computing with python. It provides N-dimensional array object, useful linear algebra, random number capabilities and related functions.

OS: The OS module in Python provides functions for interacting with the operating system. OS comes under Python's standard utility modules. This module provides a portable way of using operating system-dependent functionality. This module was used to run command terminal system commands.

Time: This is a basic python module which provides various time related functions. In the program the module is used for time code delay for the execution of previous called module.

Functions:

Google_text_to_speech: This function takes the text to be as input and converts it to an audio file and then plays the file.This function also takes an optional language parameter for output audio in the specified language.

Image_Text_Speech: This function takes the path of the image as the input and detects the text and language of text in the image file and then converts to audio.

Object_Detection: This function takes the path of the image as the input and detects the object in the image file and then converts it to audio.

Currency: This function takes the path of the image as the input and detects the currency in the image file and then converts to audio.

Voice_Input: This function accesses the microphone connected to the system and removes the ambient noise in the surrounding before taking the audio input from the user and converts to string.

Dataset:

For training the model to detect currency, a dataset of 2500 images was selected. Various perspective of the currency were taken and added to the dataset.
Our model had 11 classes comprising of

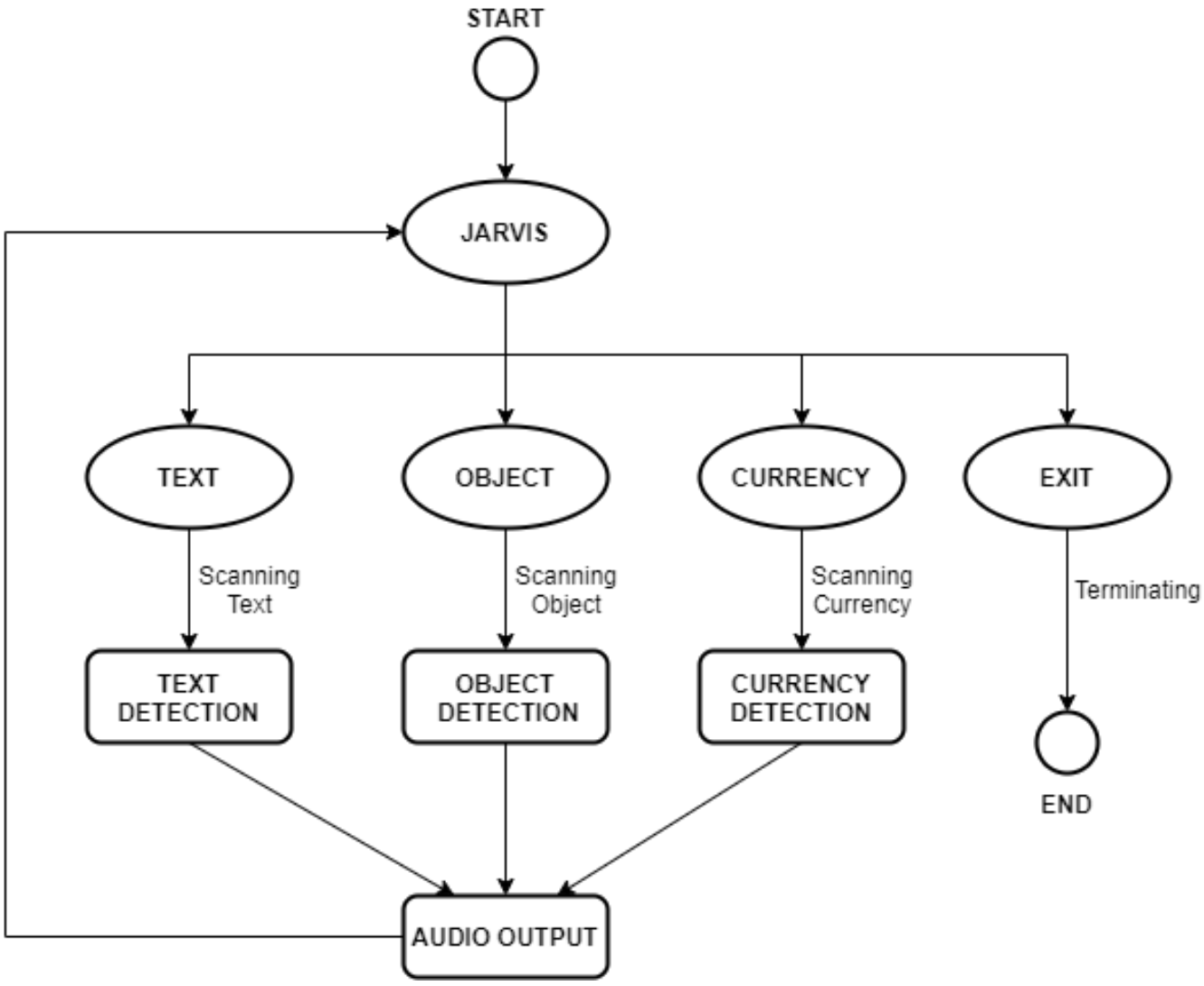
- Rs 10 old
- Rs 10 new
- Rs 20 old
- Rs 20 new
- Rs 50 old
- Rs 50 new
- Rs 100 old
- Rs 100 new
- Rs 200
- Rs 500
- Rs 2000

Few classes were separated into two different classes(old and new) as their features were at extreme ends and could not be trained as single class as it yielded very low accuracy.

WORKING

The model is activated using a voice command. The voice recognition looks for the keyword “Jarvis” to get activated. There are four possible commands to do four different tasks:

- **Jarvis Text** for text detection
- **Jarvis Objec**t for object detection
- **Jarvis Currency** for currency detection
- **Jarvis Exit** to terminate and close

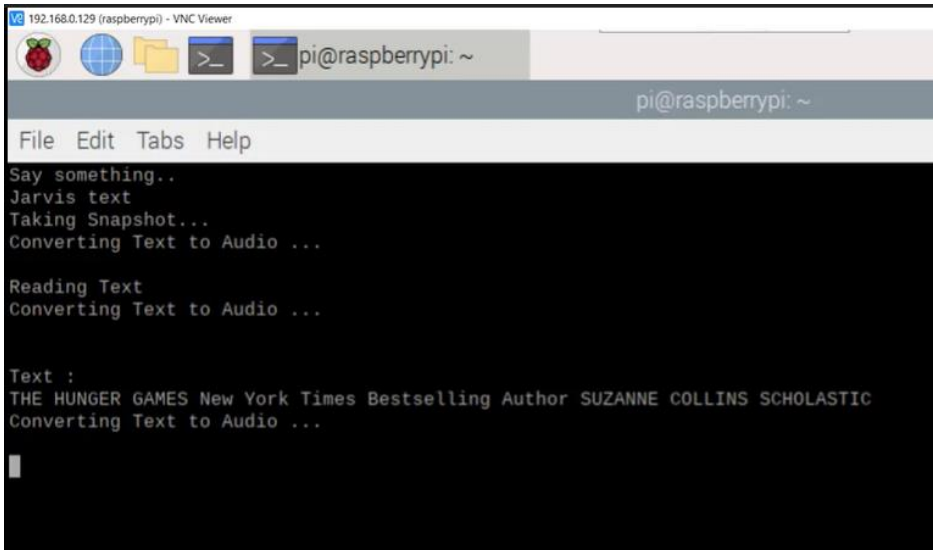


Text Detection:

Upon receiving the voice command “Jarvis Text”, the text detection module is activated. A picture is taken using the Raspberry Pi camera. The text in the image is detected using Google Cloud Vision API. The detected text is then converted to audio using the Pygame module. This audio is played to the user through earphones connected to the Raspberry Pi’s audio output jack.

Example:

a. Image taken



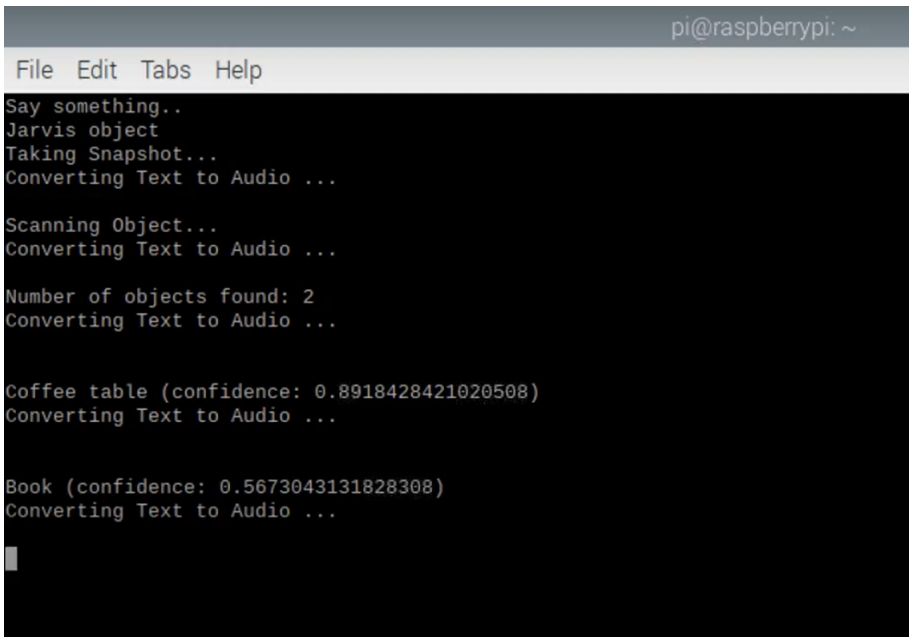
b. Text detected

Object Detection:

Upon receiving the voice command “Jarvis Object”, the object detection module is activated. The Raspberry Pi camera takes a snapshot. Google Cloud Vision API is used to detect the objects present in the image fed to it. It gives the number of objects detected and their names. The output is again converted to audio using Pygame and played through the user’s earphones.

Example:

a. Image taken



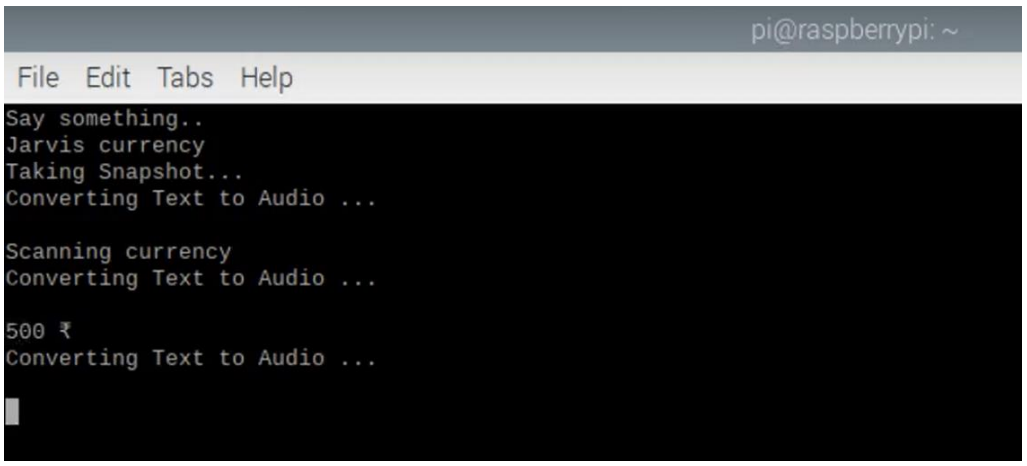
b. Objects Detected

Currency Detection:

Upon receiving the voice command “Jarvis Currency”, the currency detection module is activated. Raspberry Pi Camera is used to take a picture of the currency placed in front of the glasses. This image is fed to the trained machine learning model to detect the currency present in the image. The denomination of the currency along with the type of note (old or new) is given as the output. This output is read out to the user in the form of audio through his earphones.

Example:

a. Image taken



b. Currency detected

FUTURE IMPLEMENTATION

- **Face Recognition:**

  It is difficult for the blind to recognize people in a variety of social interactions. Our model can be improvised to identify and verify the identity of an individual using their face. Upon a voice command, a picture of the person in front is taken. This module will perform one of the 2 operations. One, if a person's face is new to the blind person, the model can be made to learn the face and a name can be assigned to the face. Two, if the person's face exists in the database, the model will give the name of the person. More advanced deployment methodologies will access existing face image databases. This will improve the access, integration, and independence of the blind in workplace or educational settings.

- **Barcode Scanning:**

  We can also implement barcode scanning to identify products using their barcodes as such information are inaccessible for the visually impaired. Barcodes uniquely identifies almost all commercial products. OpenCV can be used to facilitate the process of reading barcodes and QR codes. Once we have the image, we can then pass it to a dedicated Python barcode decoding library such as a Zbar.

- **Reading Newspapers:**

  Knowledge of current affairs is very essential in today's life. The blind cannot read newspapers in a timely manner. For this purpose, we can also integrate newspaper reading into our model. This is an application of text detection but the text in newspapers is printed column-wise which makes the detection a bit complicated. This can be implemented by making the model detect and isolate columns with margins around them and convert the news to audio.

- **Time Detection:**

  Time detection can be implemented to read the time from both digital and analog clocks or watches. Digital time detection is again an application of text detection. Time in analog clocks can be detected by finding the positions of the hour and minute hands of the clock and calculate the time.

LIMITATIONS

A few drawbacks in our model are:

- The model detects objects like bottles (water, beverages, cosmetics, etc.) as packaged goods. Since there are infinitely possible packaged goods, it is hard to classify them. Only objects belonging to major categories are identified. But this problem can be solved in the future by improving the dataset and adding specific categories for such products.

- The object detection takes some time to detect objects. This may be solved by using a faster processing chip. Another possible solution could be improving the object detection algorithm while keeping the accuracy factor in mind.

- The currency detection module does not recognize very old notes that are not in use now. These notes are not in circulation currently and are not accepted also. Hence the detection of these is not very important.

SUMMARY

We conclude that our product proves to be very useful for the blind and visually impaired. It helps them lead a much more independent life without depending on others to read things for them. It is very easy to understand and use. Today almost everyone owns a power bank. Hence the cost of buying a power supply for the product is avoided. It's portability and compactness are some important feature as it can easily be worn and carried around. It is also comfortable for long time use.
The product can perform 3 operations – text detection, object detection and currency detection. The modes are activated using voice commands which makes the use very hassle-free. The results are as accurate as possible.
It is also very cost efficient. Our budget for the project was Rs. 9000 and we have used it perfectly. But this cost can definitely be cut down to about Rs. 4000 – 5000 by removing the middlemen. Many parts of the Raspberry Pi like the extra USB ports, GPIO pins for LCD display, etc. are not used which can be removed, further reducing the cost.

TIMELINE

S.no	Activity	Months										
		0-1	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	10-11	11-12
1	Discussion on project design												
2	Ordering Components												
3	Object detection												
4	Text detection												
5	Text to speech conversion												
6	Voice recognition												
7	Currency detection												
8	Debugging												
9	Testing												

ACCOUNT STATEMENT

S.No.	Item	Price in Rs
1	Raspberry Pi 4	4099
2	Raspberry Pi Camera	1799
3	Raspberry Pi Case	399
4	Micro SD Card	799
5	Flex Cable	445
6	USB Microphone	381
7	USB Extension Cable	199
8	Go Pro Head Strap	249
9	Pole Mount for Go Pro	229
10	Miscellaneous	401
	Total	**9000**

The total cost of the project was Rs 9000 we estimate that we can reduce the cost to less than Rs 4000 by removing the middle man, manufacturing our own products and by bulk production.

PATENT BY OTHER COMPANIES

- WO1997017043A1 (Patent number): https://patents.google.com/patent/WO1997017043A1/en (Abstract).
- WO2015172418A1 (Patent number): https://patents.google.com/patent/WO2015172418A1/en (Abstract).
- US9805619B2 (Patent number): https://patents.google.com/patent/US9805619B2/en (Abstract).
- EP2490155A1 (Patent number): https://patents.google.com/patent/EP2490155A1 (Abstract).
- US20150002676A1 (Patent number): https://patents.google.com/patent/US20150002676A1/en (Abstract).

We are yet to file a patent but have started the process.

REFERENCES:

- https://www.pyimagesearch.com/2018/05/21/an-opencv-barcode-and-qr-code-scanner-with-zbar/
- https://towardsdatascience.com/
- https://www.sbir.gov/sbirsearch/detail/1137#:~:text=This%20project%20proposes%20to%20develop,wireless%20earpiece%20or%20Braille%20display.
- https://www.who.int/blindness/publications/globaldata/en/#:~:text=Globally%20the%20number%20of%20people,blindness%20is%20cataract%20(51%25).
- https://pypi.org/
- https://realpython.com/python-speech-recognition/
- https://cloud.google.com/vision/docs
- https://www.tensorflow.org/api_docs/python/tf/keras