RESEARCH ARTICLE OPEN ACCESS

Drishti for blind - a smart assistant for navigation and go-to text reading

Rupali Singh [1], Roopesh Shukla [2], Manish Thakur [3], Saylee Shinde [4], Abhijit Patil [5]

[1][2][3][4][5] Computer Engineering, MGM's College of Engineering and Technology, Mumbai University, Navi Mumbai, Maharashtra - India

ABSTRACT

Visual impairment can pose a challenge to accomplish everyday activities such as driving, reading, and walking. People with visual impairment experience their surroundings uniquely when compared to a sighted person. They have more active senses (touch, smell, sound, and taste) to learn about their surrounding but activity as easy as reading a book or detecting objects in front without touching seems impossible for them. Drishti (meaning "vision" in the Hindi language) is designed to describe this beautiful world to the visually impaired. Drishti helps to read the text in front of the user, and it also conveys the environment around the visually impaired, by describing the objects and relationship b/w the objects. Examples are a 'person sitting on a bench,' 'a dog is sleeping,' 'a stop sign on the road'. On user's command the image is captured with a camera connected to the raspberry pi. If the user wants text description of the image, the image is analyzed by Google's Cloud Vision API using Optical Character Recognition to detect the characters, letters in the picture, and if the user wants the scene description of the nearby environment, the image is sent to the Microsoft Cognitive Services to analyze the image using Computer Vision. Generated result is stored in the Dynamo DB (Cloud database). When the user asks Alexa to 'Read the text' or 'Explain the environment', the Alexa Skill Kit triggers AWS Lambda Function to fetch the results from the database (Dynamo Db). The Alexa app then recites the stored result on the user's mobile. This innovative device, designed to serve the visually impaired, is handy, easy to use with high accuracy.

Keywords— Drishti, visually impaired, Alexa, Raspberry Pi, Dynamo DB, Microsoft computer vision API, AWS lambda, OCR.

I. INTRODUCTION

Blind and visually impaired people even with complete or partial loss of vision face various problems in day-to-day life. But their ability to accomplish a day-to-day task is no different from any sighted person. They are just differently-abled who rely on different methods to achieve those tasks but with some struggles. They are more or less dependent, to do their daily life routines. Besides improved mobility, most people who have visual impairment need aids to perceive obstacles usually assisted by other persons. Effortless activity, such as reaching a place or reading a book, is challenging for them. Many solutions have been introduced before making their life a little more manageable, like the braille system, which uses a sense of touch to help read. Still, this system's reach and popularity are not on a significant scale as not everybody uses this system in their business. Other solutions like audio or digital systems have improved over the years but have still not reached a vast extent though it carries great potential. With the advancement in AI and with much higher availability of the internet, computational systems are now capable of automatically generating captions like describing objects, people, and scenery in images. While these systems vary in accuracy, they are prominent enough to use it in daily life. Modern mobile phones can capture photographs, making it possible for the visually impaired to take images of their surroundings Description of what is happening in their surroundings in real-time with the help of a Smart Assistant can empower the visually impaired and blind, but to accomplish the task with high accuracy is challenging. Image classification and object recognition tasks are prevalent in the domain of Computer Vision, but to recognize the relationship between two images is comparatively hard. Tech Giants like Google, Microsoft, and Amazon have developed various API's to do this task, each of them with varying performance levels. With Drishti, we aim to make visually impaired lives easier.

Drishti uses modern and highly useful tools such as - Raspberry pi, Computer Vision, Dynamo DB, OCR, and Alexa Skill kit. When the user requests Alexa to illustrate the area/location/surrounding, the camera captures the surrounding image and then using Computer vision APIs, objects, and the relationship between the objects are recognized. Furthermore, if the user wishes to read something, OCR detects the texts in the image. All the information is processed and converted into text, which is reported to the user by Alexa.

II. RELATED WORK

A. Smart Cap – Wearable System for blind

A smart portable device specially designed, to address the needs of visually impaired people. This device has proved to be valuable for blind and visually impaired people by helping them to experience their surroundings freely and independently without any human assistance. The device is build using a NoIR camera- to capture the scene in their surroundings, Raspberry Pi – the heart of the system, an earphone – a hearing aid which gives the voice output, and a power source to keep the whole lot running. The portable camera captures the surrounding in frames, using CV2 – an open-source library. This frame act as an input to the processor – Raspberry Pi, which processes the image using Tensor Flow API, built to perform tasks like object detection and classification, giving result in text format. This text is converted to speech using an open-source python library called gTTS (Text to Speech). The voice output generated can be easily heard using the attached earphone.[1]

B. Finger Reader

In this project, a device is designed to fit on the index finger, which helps the blind and visually impaired to read any printed text without braille. This device equipped with a camera is fitted on the index finger, assisting the user in reading printed text on the go.The device is a straightforward, practical solution. This device also ensures that the user follows the alignment and give real-time feedback with Vibrations if deviated from the alignment. Presently when technology has advanced so much, blind and visually impaired still depend on the ancient braille system to read, which is not practiced to a vast extent, limiting their interactions and access to the world. Finger Reader can benefit them with daily reading tasks like 'Reading newspaper,' book, or menu in a new way.[2]

C. Smart cap for visually impaired in a disastrous Situation

Performing routine activities for blind and visually impaired is already challenging, but when a disaster hits, like an Earthquake or Tsunami, that challenge can escalate to a new level. In these stressful situations, navigating to a secured location is tough without any assistance, and if the person lives alone, then navigation becomes even more challenging. To resolve this problem, a device called "Smart Cap," Working with the same technology as described in [1] and connected to the cloud to use it more efficiently, gives precise information about the situation in real-time. It informs any change in terrain, helping them decide which direction or path is safer and more
secure.[3]

D. Indriya - A Smart Guidance System for the Visually Impaired

An intelligent device, made using just two equipment, a smartphone, and a portable, handheld device. It is compact and cheap, making it available to every needy. Using resources like google maps and any handheld device, the user can foresee the obstacle present at two meters distance. A gadget so small can
fit in a pocket, purse, or can be attached to a belt. With added features like Voice calling and informing about current time and location, this device surely proves to be a bane for the visually impaired.[4]

III. PROPOSED SYSTEM

Intending to bring a very affordable and handy device, Drishti gives real-time information about the encompassing surrounding using ALEXA as its go-to assistant. A raspberry pi supported with a portable camera is fitted into a cap that captures the image after the user's request. The image is sent to Microsoft API – Computer Vision for recognizing and analyzing objects present in the image. The output generated is stored in the Dynamo Db (AWS database). When the user requests ALEXA for illustration, Alexa Skill triggers the AWS Lambda Function to retrieve the output stored in the database. The entire code is in python, which runs in Raspberry Pi – the backbone of the system

A. Additional OCR Support

A person strolling can also come across specific and relevant instructions like a "stop sign" or any warning sign, tags, and labels used in shops or reading newspapers, books. So, the device must be able to identify these texts and narrate it back to the user. Drishti delivers this feature with the help of Optical Character Recognition technology, which does the task of recognizing the texts from the image. Google Cloud Vision API provides this feature with high accuracy, trained with thousands of datasets. This simplistic feature of text extraction and recognition is applied to bring ease into the life of visually impaired.

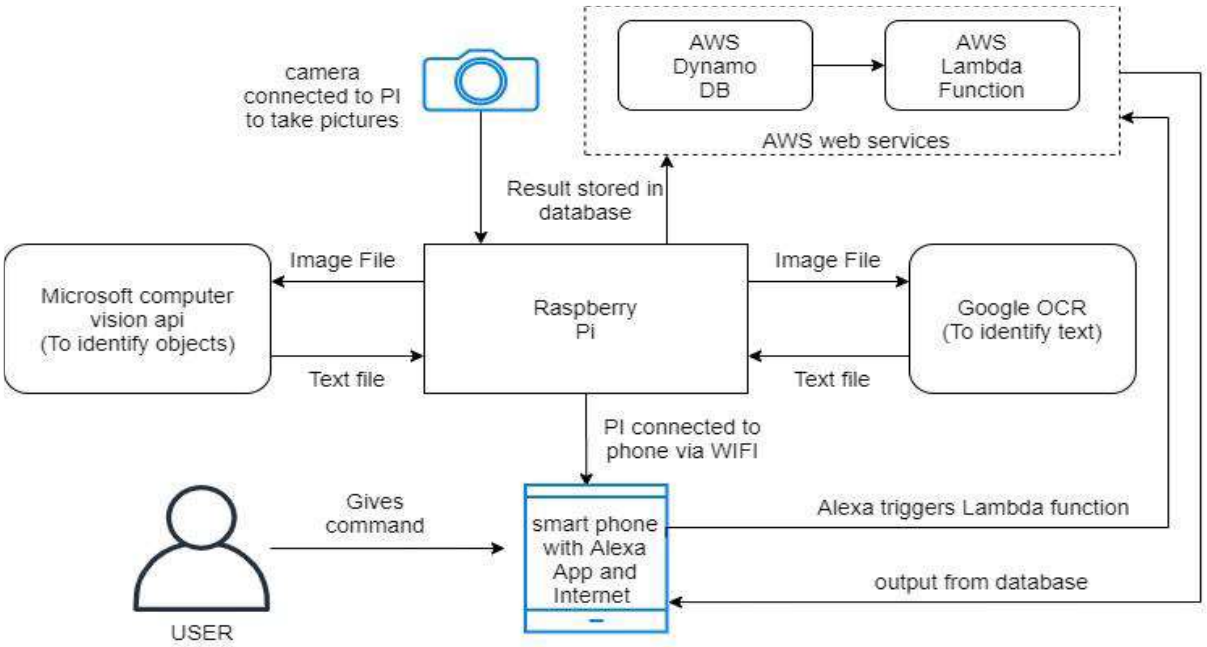

Fig. 1 Block Diagram

IV. DESIGN AND IMPLEMENTATION

In this project, the leading technologies, which form the heart of the system, are Raspberry pi, Computer Vision, OCR, AWS Dynamo Db and Alexa Skill kit. The camera captures the surrounding image and then using Computer vision APIs, objects, and the relationship between the objects are recognized. All the information is processed and converted into text, which is stored in the Dynamo DB database.
For communicating the message to the user, a custom skill (created using the Alexa skill kit) triggers the AWS Lambda function to extract the text from the Dynamo DB and forwards it to the Alexa app in an audio format recognizable to the user. The user needs to download the Alexa app on their phone, and they need to enable the custom skill in this app.

A. OCR Support

The task of Google's Cloud Vision API is to detect and extract text from any image. In our system, the captured image file is first converted into a base64 string. Then a layout analysis of the image is performed to segment the location of the text. Once the area is detected, the OCR module performs text extraction analysis on the identified site to generate text. In the post-processing stage, errors are minimized or corrected by feeding the texts through a dictionary or a language model. The final result is stored and recited by ALEXA to the user. recited by ALEXA to the user.

Given below is the design specification of Drishti:

- Scene Description- Describing its encompassing surrounding.
- Text Description- Identifying and Reading text present in the image.
- A portable Smart Assistant – A Custom Alexa Skill to give commands and receive voice output.
- Cloud Computing and IoT - To store data remotely and access via ALEXA.
- Portable, rechargeable, and affordable.

Table I
HARDWARE & SOFTWARE USED

S. No	Device/Software	Description
1.	Camera	Fitted to the cap. 2 to 3MP camera is used.
2.	Raspberry Pi 3 and Raspbian OS	Operating system and programming platform
3.	Microsoft Cognitive Service	Tool for image recognition.
4.	Amazon Alexa	Deployed to use AWS services.
5.	Google OCR	To extract text from the image.
6.	Cap	To hosts the camera and make the device wearable.
7.	Headset	Help the user to listen to Alexa.
8.	Anaconda	Software platform for Python programming language.
9.	Smartphone	Hosts the Alexa App, to provide the user interface for giving commands and receiving the output.

V. COMPARISON WITH EXISTING SOLUTION

The research works mentioned above have undoubtedly contributed to the betterment of blind and visually Impaired. However, when compared, Drishti compensates for the features which may lack in them. Similar work in paper [1] [3] give an excellent narration of the picture taken by the camera; however, it has limitations. The feature to recognize and extract text from the picture taken is not provided. Drishti overcomes this limitation by providing the feature to read the text from the image taken using OCR (optical character recognition). Research work [2] helps the blind and visually impaired to read any printed text without braille. However, this project is limited to reading, and it cannot provide a real-time narrative of the nearby environment. Research work [4] can inform the user of any obstacle ahead up to two meters with high accuracy. However, this project is incapable of illustrating objects, people, and scenery in the surrounding or read text from any signboard or books. Drishti overcomes this limitation by providing a real-time description of the nearby objects and text in the image taken

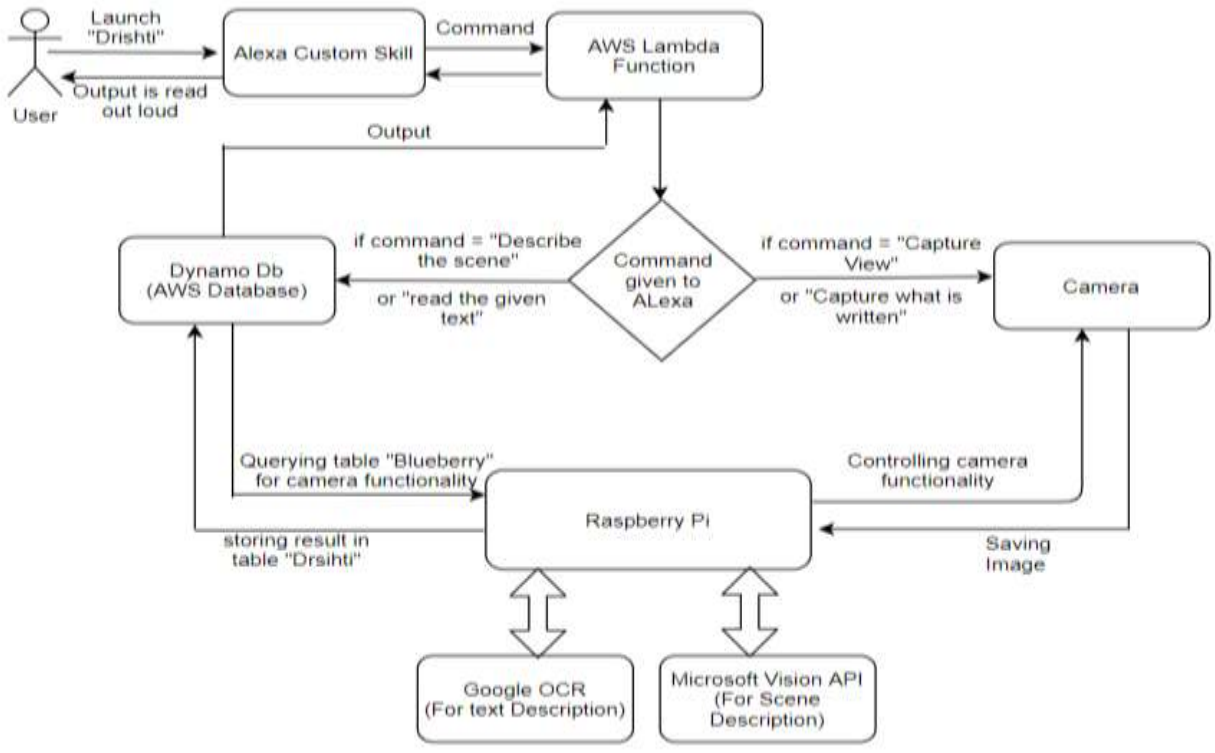

Fig. 2 Process Flow Diagram

Given below are some of the unique features of Drishti:

- It can read a text in the image as well as provide a text description of the image.
- Friendly user interface and easy to understand it's working. Portable, lightweight, and rechargeable.
- Use of Cloud Computing and IoT, giving real-time response with remote storage and access of data.
- Giving voice control input and getting voice control feedback making it manageable.
- Moderate cost

VI. TESTING AND EVALUATION

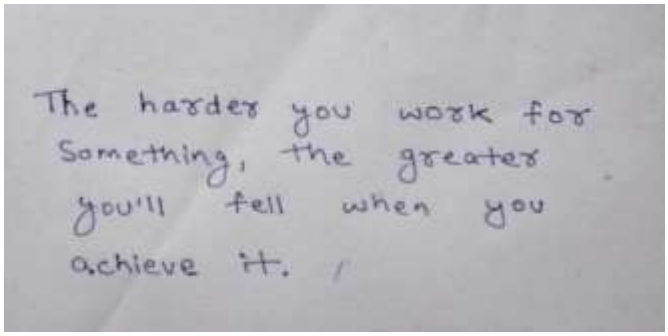

Fig. 3 Captured Image 1

Fig. 4 Alexa Output 1

Fig. 5 Captured Image 2

Fig. 6 Alexa Output 2

Fig. 7 Result from Microsoft vision API (Computer Vision)

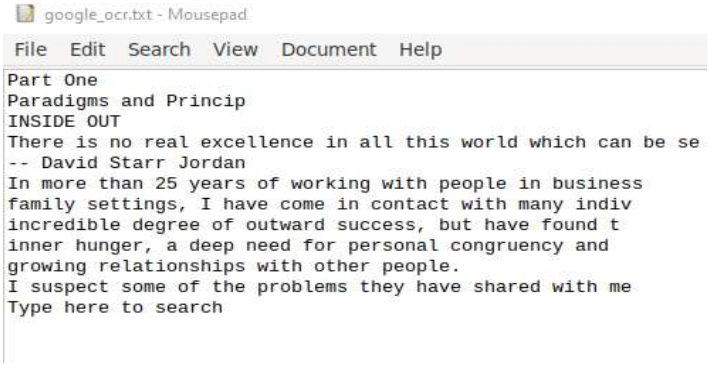

Fig. 9 Text file created after analyzing the "written *text" using google — Optical Character Recognition*

Fig. 10 "Help" section of Drishti – Alexa

Fig. 11 Raspberry pi kit with camera

VII. RESULT AND CONCLUSION

Technology, when applied right, can prove to bring excellent development in the quality of life. With Drishti, we aim to bring a significant change in the lives of Blind and Visually Impaired. It is an affordable device to assist them with their activities, giving them a sense of autonomy. Drishti has tried to address as many challenges as possible and has come up with an innovative solution. From being aware of their surroundings to be able to read any vital information, Drishti can surely be of great help to improve their mobility and communication with people. With an easy-to-use interface and with real-time results, it has the potential to reach a vast population of needy. We believe that Drishti can serve the visually-impaired people and provide them with the comfort and assistance they desperately need during numerous circumstances.

REFERENCES

[1] A. Nishajith, J. Nivedha, S. S. Nair and J. Mohammed Shaffi, "Smart Cap - Wearable Visual Guidance System for Blind," 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, 2018, pp. 275-278.

[2] Shilkrot, Roy & Huber, Jochen & Wong, Meng Ee & Maes, Pattie & Nanayakkara, Suranga. (2015). FingerReader: A Wearable Device to Explore Printed Text on the Go. 2363-2372. 10.1145/2702123.2702421.

[3] Madanian, Samaneh & Airehrour, David & Cherrington, Marianne & Patil, Nikhilkumar. (2018). Smart Cap for Visually Impaired in Disaster Situations.

[4] Kallara, Sujith & Raj, Mitu & Raju, Roshan & Mathew, Nidhin & V R, Padmaprabha & DS, Divya. (2017). Indriya Indriya - A Smart Guidance System for the Visually Impaired. 10.1109/ICICI.2017.8365359.