ISSN 2278-3091
Volume 10, No.3, May - June 2021
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse351032021.pdf
https://doi.org/10.30534/ijatcse/2021/351032021

THIRD EYE
360° Object Detection and Assistance for Visually Impaired People

Rohan Jain[1], Yukta Bhattad[2], Tejal Patil[3], Rohit Patil[4], Khalid Alfatmi[5]
[1]Department of Computer Engineering, Shri Vile Parle Kelawani Mandal's Institute of Technology, Dhule, India, rohanrj81099@gmail.com
[2]Department of Computer Engineering, Shri Vile Parle Kelawani Mandal's Institute of Technology, Dhule, India, yuktabhattad@gmail.com
[3]Department of Computer Engineering, Shri Vile Parle Kelawani Mandal's Institute of Technology, Dhule, India, tejalpatil1068@gmail.com
[4]Department of Computer Engineering, Shri Vile Parle Kelawani Mandal's Institute of Technology, Dhule, India, krohitpatil15@gmail.com
[5]Assistance Professor, Department of Computer Engineering, Shir Vile Parle Kelawani Mandal's Institute of Technology, Dhule, India, Khalid.Alfatmi@svkm.ac.in

ABSTRACT

Visual disability is a global issue. Visually impaired people confront several challenges every day. Many times, blindness affects a person's ability to self-navigate in known or unknown environments. The difficulties faced by them and how they deal with them are largely known and explored. The system we developed is an idea to overcome the challenges of detecting objects in a known environment or room environment with the help of Artificial Intelligence. The idea is based on the approach to aid visually impaired people with voice assistance to detect objects of the surrounding using 360° view cameras. The proposed system uses a 360° view camera of the mobile phone to assist the user for detecting desired objects in the room environment and provide localization. Using this system, users can search for the desired objects by giving voice commands and can be assisted to the object location. When the user wants to search any object, he/she simply gives a voice command using NLP to the system. The system then identifies commands and extracts the object name to be searched. With the help of image processing, first identifies and locates the object in surrounding and navigates the user to that object using a voice assistant.

Key words: Artificial Intelligence, Natural Language Processing, Voice Assistant, Convolutional Neural Network (CNN), Cloud Computing firebase, Raspberry pi, Support vector Machine(SVM), Robot(prototype).

1. INTRODUCTION

According to World Health Organization (WHO) in 2012, out of the 7 billion population, there were over 285 million people are visually impaired and 39 million were totally blind. Blind people face the problem of movability. Traditional tools like white cane and guide dogs do solve the problems completely. They cannot perceive enough knowledge to avoid all the obstacles. With research in science and technology, different systems were proposed to help visually impaired and blind people. To make them independent of aid, a different system using different technologies were proposed. To make them self-dependent and perceptive, we provide them a third eye i.e. 360° object detection and assistance. Object detection and extracting their feature is a very interesting topic and also one of the most challenging problems in the computer field. Object detection from 360-degree panoramic images is mostly used in recognizing assistance and has amazing potential in automatic driving and navigation of drones, robots, and other movable devices.

The International Classification of Diseases organization has declared the statistics that show, 285 million people are visually impaired worldwide of which 39 million are blind and 246 million have low vision. This is a widely adopted disease. Visual disability is a global issue. The visually impaired people in their daily lives face innumerable challenges. Amongst which, one is to independently navigate and perform daily activities are highly affected activities. Our idea is based on the approach to aid visually impaired people with voice assistance to detect objects of the surrounding using 360° view cameras. The proposed system uses a 360° view camera of the mobile phone to assist the user in detecting desired objects in the room environment and provide localization. An artificial intelligence-based system that provides voice-based commands with object tracking to search, locate, and navigates to objects.

This paper shows a system as well as an algorithm for detecting and locating objects in-room environment. After looking at the virtually impaired people's problems we decided to create a system that will be helpful for them to

make their day-to-day life activities easy. After talking to many people and after taking their views that what kind of problems they are facing we concluded that what kind of system should be designed for them. This system will be free and handy to use. It can be used by everyone having smartphones and internet connectivity with them. We have designed the android app which will help the user to perform his or her day to day activities by helping users to recognize and find the objects they are looking for. The mobile application will be operated through voice input using Google Assistance. The mobile camera will be used to capture the images of the surrounding, and this application will send these images to google firebase for further processing. After the images are received at the server-side, the image processing will be carried out, and the object detected in the images will be kept in the database in the form of table, with there meta-data also, like direction of the object, name of object and other information. The voice input which is received by the user through the mobile application will also be transferred to the severe side for processing. Then a simple program of natural language processing will take the sentence of the user and will extract the keywords from that sentence. This keyword will contain the name of the object which the user needs to find. Then the searching algorithm will run into the database system to find, is the object present in the room. If the object is present, then the direction command will be provided to the raspberry pi present on the prototype robot which is designed by us. This prototype robot will move to this object and will tell the user the location of the object with the help of voice assistance.

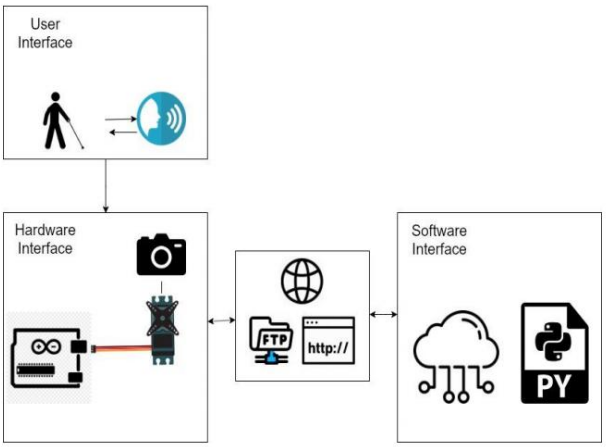

Figure 1: Communication Interface

The above figure shows the interface of the whole communication carried out in the system, from the user input to the hardware and software interfaces. So in this whole scenario, we use technologies like android, image processing, natural language processing, python, cloud firebase and integrate them with Raspberry Pi system present of the prototype robot.

2. SURVEY OF LITERATURE

Many mobility aids and object detection and recognition aids have been introduced for visually challenged people for the detection of objects in there surrounding so that they can find what they want in an easy way. Different aids are based on various technologies mainly ultrasonic sensors, infrared (IR) sensors, or image sensors for the detection of obstacles [1].

The first paper that we surveyed is, PARTHA: A visually Impaired Assistance System [1]. This system is a combination of smart glove and smartphone which will be used to detect the objects and also to avoid the obstacles that come their way [1]. The smartphone is used to detect the objects around them. This system also has the navigation through which the person using the device will get the instructions of the objects they are searching for, the navigation part is based on wifi [1]. The most important part of this device is the system shares the real-time location of the blind person which improves the chances of security of the person as the location will be shared. This point makes the system more feasible [1]. The limitation in this system is that the Current indoor navigation system needs to be trained for the location before even using it. The roots are set in the system, the person needs to set the roots before using the system to make these all facilities available [1].

Now another paper was Application for the Visually Impaired People With Voice Assistant [2]. This paper talks about an application that will help visually impaired to see the object using their Smartphones [2]. The smartphone will be used to click pictures around and with this, the user and perform the activities like reading or finding things or finding their roots [2]. They will have an app that will help the user to get the guidance throughout and the app will use the voice control feedback mechanism because of which the user can perform his or her daily life task easily without seeking anyone's help [2]. The limitations that we found is that the user needs to have a smartphone always to perform the task throughout and the instructions will have to be given always manually [2].

Another paper was, Multi-sensor – based Object Detection in Indoor Environment for Visually Impaired People[3]. This system is designed for people whose vision cannot be corrected to normal vision with standard high glasses or any type of contact lenses [3]. They have a multisensory based system which is used for object detection for the indoor environment [3]. Where their object detection will be performed on the images they capture using statistical parameters which has a support vector machine algorithm [3]. The multi-sensor here is employed by an ultrasonic sensor. Also, here infrared sensor is used to detect small objects [3]. Limitations observe in this system are limited to indoor-only, the person cannot use it when outdoor. The distance is limited to 7 feet. Their system finds obstacles in the range of only 300 cm using an ultrasonic sensor. The person always needs to have his headphones with him as their sensor is light sensitive [3].

We also came through a paper that was, Real-time Object Detection for 360-degree Panoramic Image using CNN [4]. This system is presenting real-time object detection for 360-degree panoramic images where they have used a convolutional neural network (CNN) [4]. They have also

adopted the CNN- based detection framework for object detection which is having a post-processing stage to find the result. they have also proposed a novel method where they will be reusing the existing database of ordinary images, they have given the examples of imageNet and PASCAL VOC[4]. They have also demonstrated with several examples where they have proven that their method yields higher accuracy and also recall rate [4]. Limitations that we came across are, as their system has CNN which is based on YOLO's CNN structure, their system also has the limitations that YOLO has [4]. Their system cannot detect small objects also the objects which will be near to pole of the 360-degree panoramic sphere might get twisted too much to detect. They might get duplicate detections too. If their IoU threshold does not suite [4].

This was another paper using, 360° View Camera-Based Visual Assistive Technology for Contextual Scene Information [5]. This system uses a 360° view camera with a mobile device to capture surrounding scene information and provide contextual information to the user in the form of audio. The scene information from the spherical camera feed is classified by identifying objects that contain contextual information of the scene [5]. That is achieved using convolutional neural networks (CNN) for classification. The result of this paper shows 92.8% of classification accuracy. The results are reported after 10 trails where training data was shuffled randomly [5]. Some of the limitations that were founded are, No lens distortion correction was included in the system pipeline. Also, the time is taken for image capture and then data transfer over network took 7 seconds on average. This is not because of the connection type rather the internal image processing [5].

Some more papers on the above topics that we got are the Adaptive 360 Degree Image Recognition Approach to Empower Visually Impaired [6]. They have explained the kind of problems the people who are the virtual impaired face, among which the ability to navigate and also to perform all the daily tasks is one of the most difficult jobs for these people [6]. This system is completely based on the abilities for using virtual feeds from cameras and image processing and machine learnings which will help the user the identify and detect the things objects around them which they need to find. They have audio assistance which will guide them throughout to get the objects they are finding for [6]. This system is limited to only detecting objects and might not even get a 360-degree angle. The user needs to have all the 360-degree angle to detect the image [6].

Some of the papers were based on the hardware devices that the user can wear and can use for finding the objects like, Third Eye: A Shopping Assistant for the Visually Impaired [7]. This system is a combination of cameras that can be worn also called wearable cameras also hardware accelerators and algorithms and a vision-based automatic shopping assistant these things will allow the user to detect the activities and get to know about them [7]. Here they have used the technology for human augmentation, where they will be working with the help of gloves which will guide the person to give the instructions, to click the pictures and other important stuff [7]. It is limited to a particular distance, the person will have to rotate the camera for finding the object in that case the person might miss few places which counts it to the biggest disadvantage, if the particular person who is using the device rotates everywhere properly then he or she has the chances of finding the object or else he or she might miss on few places [7].

Some papers that we came across uses the 3d technique and 3d cameras like, Quick 3D Object Detection and Localization by Dynamic Active Search with Multiple Active Cameras [8]. This system was designed to detect the known objects in the 3D environment and also to set their positions to pan-tili-zoom cameras. It was a time back then when 3D wasn't invented so basically this was kind of the procedure to start inventing it, multiple cameras were used to click the multiple pictures facing different angle, multiple cameras were accurate than the single-camera because the multiple cameras were faster than the single-camera [8]. Some of the problems that we can find out were single cameras can't be said as accurate as multiple cameras will be said. 3D room is needed for any of the activities, there is no voice assistant which will guide the person to perform multiple activities. the objects are not more accurate or clear [8]. There is no fast detection of the object we want. If the background is green and the object is also green this method cannot be applied to it. It means this method denies working where they have the same color of the object as the color of the background [8].

3. PROPOSED METHODOLOGY

We have designed this model, especially for blind people to make their day-to-day activities easier. This model can also be used by normal people as well. Now the question is how the app works and how it is going to be useful. The first step is to install the android app designed to take the pictures in 360 degrees to detect objects in the surrounding through the images taken. The android phone will be fitted on the robot. The UI of the android app will be as follows:

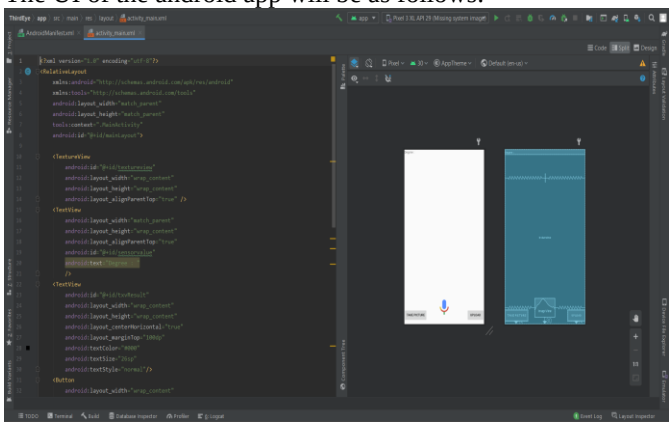

Figure 2: Mobile UI Coding

Figure 3: Mobile UI Real View

The application can be useful for normal people as well as for the visually impaired, so the UI of the app is designed in the way it can work automatically as well as if someone wants to use it manually it can be done as well. The main page of the application has Three buttons viz One will be Take Picture, another will be Upload Picture, and the last will be a Voice Input Button.

This application will be installed in an android mobile and then the application can be opened using google or Alexa voice assistance. As soon as the user tells through voice to open the application, the application will pop up and will start taking the photos of the surrounding and will make the Robot or any prototype to rotate in 360-degree fashion so that it can take 8 photos in different direction and angles so that it can sense the whole surrounding properly. As soon as the photos are taken, they are uploaded to the firebase system using internet.

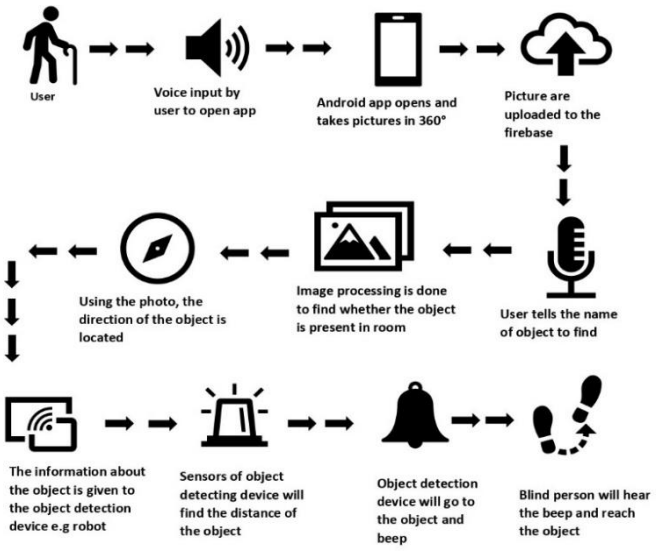

Figure 4: Working System Pictorial Flowchart

The whole proposed system is explained using this pictorial flowchart, here as soon as the system is openly charged in standby mode and after the user's instruction, as soon as the app opens, the robot rotates, and it takes pictures all over i.e in 360 degrees, the app takes around 1-2 sec to cover each angle, and slowly it will take images from all over. After taking pictures the app will uploads all the pictures to the firebase. Simultaneously the person has to give voice input to the app by saying the name of the object to be found. The voice message is uploaded to firebase and with the help of NLP algorithms like Stanfordnlp and other codes which includes nltk ie Natural Language Toolkits and pyttsx3, we extract the keyword and recognize the object to be found from that spoken sentence, Here in firebase, with the help image processing algorithm yolov3, imageai and tensorflow, all the 8 images are given as a input to object detection and with python script it is found whether the given object is present in room or not, then the code returns us the picture in which the given object is visible. If the object is available in the room the prototype robot as well as the person receives directions accordingly, like the 8 images will be named as L1, R1, etc and that will help with the directions for the person. When the Camera module takes the pictures, it will label the photos accordingly, like Front-1, Front2, and similarly Right, Left, and Back. Now the direction of the object is given to the robot and the DC motors fitted on the robot rotate in that direction. The prototype is fitted with the raspberry pi, which contains the python code, which helps to receive the direction of the object from the firebase and then it gives the command to the motor module connected to the robot to move in that direction and then reach till that object. After reaching the desired object, the mobile speaker will respond by loudly telling the direction of the object. For example, if the user tells to find the backpack in the room, and if the object is present in the room then the robot will move towards it and after reaching that object it will make a sound and it will speak, that "Backpack Object is in F1 (or maybe different) Direction". In the below given figure we can get an idea how the prototype rotates in 360° and how it names the images.

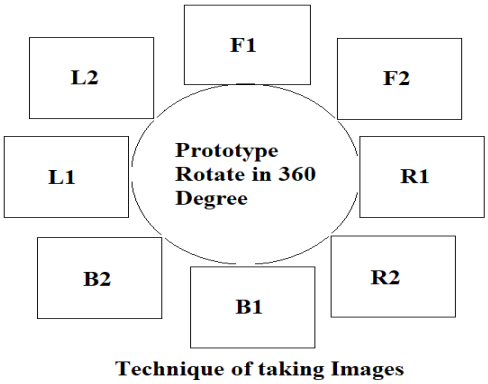

Figure 5: Technique Used to take Images

Now suppose if the object is present, the location of the object will be located and then the directions will be given to the user to reach towards the object. This device will give directions to

the blind person. The information of the object direction is given to the object detection device i.e. prototype in our case which is a robot. The following figure shows the working of the robot.

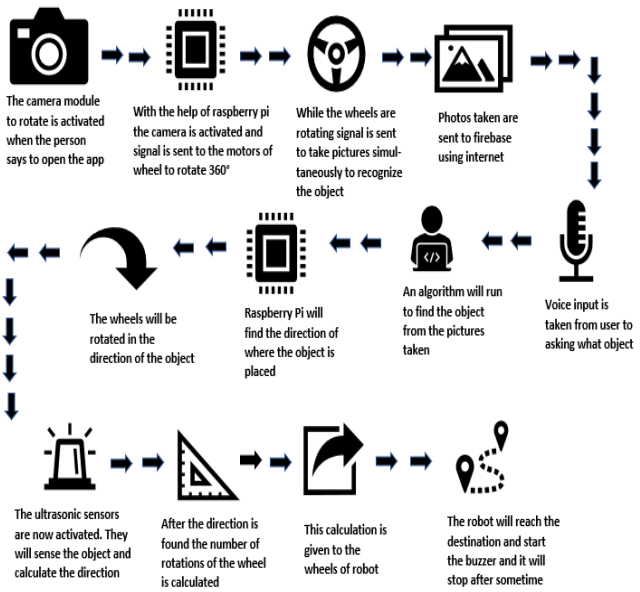

Figure 6: Pictorial flowchart of Hardware Working

We can see the total working of the hardware system in the above-given flowchart. As soon as the order is given to the model by the voice of the input of the user, the camera module gets activated. The Raspberry Pi which is mounted on the robot helps the robot to do the movement and it will make the necessary decision and will command the robot movement. The coding written inside the raspberry will now tell the robot to rotate so that the camera takes photos of the surroundings. The wheels rotate in circular format and will halt for 1-2 seconds after some time so that the camera can take a good blur-free image of the surrounding or indoor room. As soon as the photos are taken, with the help of the image compression inside the android application they are compressed and are uploaded to the firebase. Compression of the images is done by reducing the pixel size of the images when the image is taken, this is necessary as these images are too be uploaded over the firebase using the internet. As the camera of the mobile is of high megapixels, the image size is also high so we will need good internet connectivity, but with the help of the compression technique, we have made this process faster. As soon as the rotation of the Robot is finished and all the images are successfully uploaded to firebase, the application asks the user, " What object you have to Find? ". After hearing this the user has to give a voice input specifying the name of the object he is willing to find. The voice input which will be received by the user will be then uploaded to the firebase, and different algorithm of Natural Language Processing like Stanfordnlp and other codes which includes nltk ie Natural Language Toolkits and pyttsx3 will work to extract the key words as a token from the sentence and will identify the name of the object the user needs to find [9]. When this happens the images which were uploaded are given to Image Processing Algorithm ImageAi Resnet and the YoLoV3, which extracts the objects that can be identified in the images and will makes an entry into the database of that objects accordingly and gives it a proper naming as per the direction it was found in the image, Then it is checked if the object which the user needs to find is present in the Images taken or not, that is if it is present in that surrounding or not. If the object is found then the direction of that object is given to the raspberry pi module on the robot, which then orders the robot to rotate itself in that direction of that object.

After that, the ultrasonic sensors fitted on the robot help to sense where the object is placed. After sensing the object in that direction, raspberry module is used to calculate the distance of that object which is sensed. This calculation will result in calculating the rotations of the wheels that are to be made so that the robot can reach and can navigate to that object. After reaching the object, the mobile mounted on the robot will make a sound and will tell the direction of the object loudly using the speakers on the mobile, and hearing the sound of mobile the person will reach the object. After sometime this sound will get turnoff.

4. IMPLEMENTATION

The system is implemented using the Python 3.8 and the mobile application is developed in Android by using Java and all the other system is developed and integrated using the Python 3.8.

The follow of the task performed in the system can be illustrated by the below figure 7.

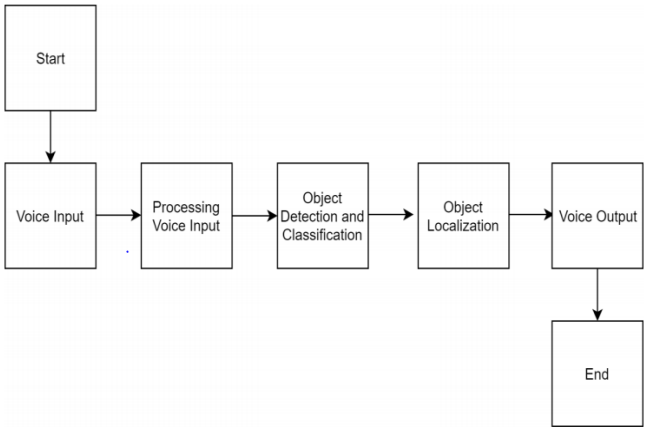

Figure 7: Task Network

The Requirement for this project can be classified into two categories wiz-

Hardware Requirement:

1. Camera Module i.e. Mobile
2. Raspberry Pi kit
3. Servo Motor and Motor Module
4. Batteries
5. Ultrasonic Sensor

Software Requirement:
1. Python
2. Cloud Environment
3. Android Studio
4. MySql Database

The Use Case diagram of our working system is as follows, it is showing us the connection between the User, System, Cloud, and all the Identification and Classification work it does.

Table 1: Information of Use case

Sr. No.	Use Case	Description	Actors	Assumptions
1	Send Command	Blind user gives voice command by specifying object name to be searched.	User	User is able to speak in English.
2	Capture Image	Image is get captured by camera of mobile	User	User is able to walk around the objects
3	Feature Extraction	Labelling different images from environment	Cloud server	User is able to sense vibration.
4	Comparing Extracted features	Matching labels with user requested object.	Cloud server	User got closer to object.
5	Object Detection and localization	Give direction to user.	Cloud server	User got closer to object.

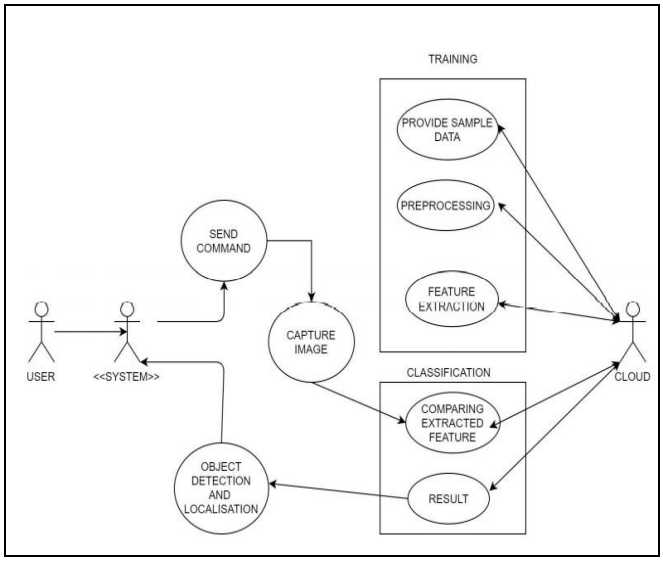

Figure 8: Use Case Diagram

The above table describes the Use case diagram which is present below it. The table explains the actors and work they do with some assumptions that we made that the user will perform while using the prototype robot in real world scenario. Data Description that is all the data objects that will be managed/manipulated by the software are described here. The database entities or files or data structures required to be described. For Data objects and their major attributes and relationships among data objects are described using an ER diagrams- like form.

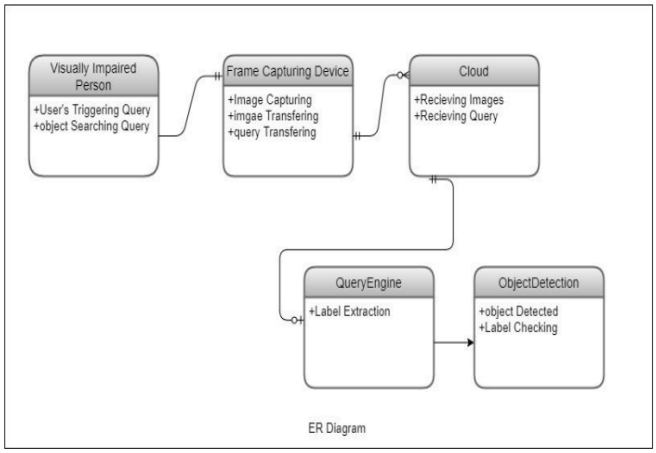

Figure 9: ER Diagram

The architecture of the project is explained in the below given figure. The figure completely describes how the model is going to work and how the processes will take place in simultaneously way. It gives the brief description from the user providing the voice input to the system till the response user get from the machine, in between all the processesing of the images as well as the voice input and extraction of the keywords and then localization of the object by the robot itself and then calculation of the distance and then response by the voice output.

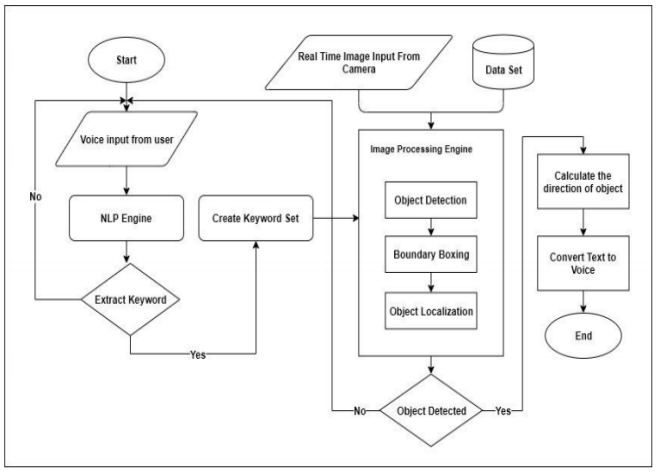

Figure 10: System Architecture

5. OUTPUT

The objects which are there in the surrounding are captured into the images clicked by the mobile mounted on the prototype robot are sent to the firebase and then they are given to the image processing and after that the output we get is in

the form of images which are shown below

Figure 11: Images of Object after Processing

The table of these objects are created into the local database into the table name “nlp”, the view of that database can be seen in the below figure 12

```
mysql_test.py > ...
1   import mysql.connector
2
3   mydb = mysql.connector.connect(host="127.0.0.1",user="root",passwd="password",auth_plugin='mysql_native_password',database='nlp')
4   print(mydb)
5   if(mydb):
6       print("Connection successful")
7   else:
8       print("Unsucessful")
9
10  mycursor = mydb.cursor()
11
12  mycursor.execute("select * from nlp1")
13  myresult = mycursor.fetchall()
14  for row in myresult:
15      print(row)

PROBLEMS  OUTPUT  DEBUG CONSOLE  TERMINAL  SQL CONSOLE MESSAGES        2: Python Debug Consc

Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Try the new cross-platform PowerShell https://aka.ms/pscore6

PS C:\Users\rohan\Desktop\Rohan Final Year Project\AAA New Project Update Till 5-2-21> & 'C:\Users\rohan\AppData\Local\Programs\Python\Python38\python
s\rohan\.vscode\extensions\ms-python.python-2021.4.765268190\pythonFiles\lib\python\debugpy\launcher' '53232' '--' 'c:\Users\rohan\Desktop\Rohan Final
AA New Project Update Till 5-2-21\mysql_test.py'
<mysql.connector.connection.MySQLConnection object at 0x0000021E581B8070>
Connection successful
('clock', '79.02817130088806', 'F1')
('book', '50.402283668518066', 'R2')
('toothbrush', '57.30130672454834', 'B1')
('suitcase', '62.4565064907074', 'B2')
PS C:\Users\rohan\Desktop\Rohan Final Year Project\AAA New Project Update Till 5-2-21>
```

Figure 12: Detected Objects Information Inside the Database

The prototype robot is as follow which has raspberry pi with motor module and motor to move and battery. There is detachable mobile mount which can be placed over the prototype. The view of robot can be seen in below figure 13.

Figure 13: Prototype Robot

6. CONCLUSION

In this paper, we provide an overview of the project that would assist visually impaired people. The project 360° object detection and assistance, an Artificial Intelligence based technology that helps in providing voice-based commands with object tracking to search, locate and navigate to objects. The aim of the system is enabling blind person to detect the desired object and proving localization of object. A voice command is given to the system when a person wants to search any object in his/her surroundings. The system, then recognizes the command and the object name to be searched is extracted. Using the technology of image processing, it identifies and locates the object in surrounding using voice assistant. The Voice assistant plays a major role in directing the blind person to the object. Hence the system aims at providing a third eye for the assistance of blind persons. The system helps in making their life easy and comfortable with the smart and advanced technology by enabling them to navigate them to the desired objects they want to get to. The system acts as a helping hand in the life of visually impaired ones, we have presented object detection system where we will get a 360-degree view of surrounding using 8 images and a robot as a prototype of an object detecting device which will have an android phone, raspberry pi, and ultrasonic sensors. This system is designed to detect the objects which blind people are not able to do and will help to make their day to day life easy. This system can also be used by a normal person to

find various things that will save the time of the person.

ACKNOWLEDGEMENT

This work was supported and guided by Prof. Khalid F. Alfatmi. He works at Shri Vile Parle Kelawani Mandal's Institute of Technology, Dhule. He is working as an Assitant Professor in the Computer Department. He helped us in going through this idea and guided us till the end.

REFERENCES

1. S Devashish Pradeep Khairnar, Rushikesh Balasaheb Karad, Apurva Kapse, Dr. Geetanjali Kale, and Prathamesh Jadhav, "PARTHA: A Visually Impaired Assistance System", 2020 3rd International Conference on Communication System, Computing and IT Applications (CSCITA), 2020.
2. Abhijeet Mohanta, Shah Yash Jitendra, Khandelwal Nikita Dinesh, Wable Saurabh Suhas, and Aruna K. Gupta," Application for the Visually Impaired People With Voice Assistant", International Journal of Innovative Technology and Exploring Engineering (IJITEE), Volume-9 Issue-6, April 2020.
3. Charmi T. Patel, Vaidehi J. Mistry, Laxmi S. Desai, and Yogesh Meghrajhani, "Multisensor – based Object Detection in Indoor Environment for Visually Impaired People," Proceedings of the Second International Conference on Intelligent Computing and Control Systems (ICICCS), 2018.
4. Yiming Zhang, Xiangyun Xiao, Xubo Yang, "Real-time Object Detection for 360-degree Panoramic Image using CNN", 2017 International Conference on Virtual Reality and Visualization (ICVRV), 2017.
5. Mazin Ali, Ferat Sahiny, Shitij Kumarz, and Celal Savurx, "360° View Camera-Based Visual Assistive Technology for Contextual Scene Information", 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC) Banff Center, Banff, Canada, October 5-8, 2017.
6. Ankur Jalan, Prafulla Bafna, Anagha Vaidya, Lalit Kathpalia, "Adaptive 360 Degree Image Recognition Approach to Empower Visually Impaired", IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), 2017.
7. Peter A. Zientara, Sooyeon Lee, and Gus H. Smith, Rorry Brenner and Laurent Itti, Mary B. Rosson and John M. Carroll, Kevin M. Irick, Vijaykrishnan Narayanan, "Third Eye: A Shopping Assistant for the Visually Impaired", THE IEEE International Conference on COMPUTER SOCIETY 0018-9162/17, 2017.
8. Takahito KAWANISHI, Hiroshi MURASE, Shigeru TAKAGI, "Quick 3D Object Detection and Localization by Dynamic Active Search with Multiple Active Cameras", IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI), 2002.
9. https://pythonspot.com/category/nltk/.
10. https://projects.raspberrypi.org/en.
11. https://imageai.readthedocs.io/en/latest/.
12. https://github.com/OlafenwaMoses/ImageAI/releases/tag/1.0/.