OBJECT IDENTIFICATION AND RECOGNITION FOR BLIND PEOPLE

A Main Project Report

Submitted to the Faculty of Engineering of

JAWAHARLAL NEHRU TECHNOLOGICAL UNIVERSITY KAKINADA, KAKINADA

In partial fulfillment of the requirements for the award of the Degree of

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

By

M.M.KAIF	**P.POOJITHA**
(19481A05E8)	**(19481A05H7)**
M.DHEERAJ	**M.SIVA NAGA SAI**
(19481A05F4)	**(19481A05E7)**

Under the Enviable and Esteemed Guidance of

Mrs.P.NagaMani, M.Tech(Ph.D)

Assistant Professor, Department of CSE

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SESHADRI RAO GUDLAVALLERU ENGINEERING COLLEGE

(An Autonomous Institute with Permanent Affiliation to JNTUK, Kakinada)

SESHADRIRAO KNOWLEDGE VILLAGE

GUDLAVALLERU – 521356

ANDHRA PRADESH

2022-23

SESHADRI RAO GUDLAVALLERU ENGINEERING COLLEGE

(An Autonomous Institute with Permanent Affiliation to JNTUK, Kakinada)

SESHADRI RAO KNOWLEDGE VILLAGE,

GUDLAVALLERU

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the project report entitled **"OBJECT IDENTIFICATION AND RECOGNITION FOR BLIND PEOPLE"** is a bonafide record of work carried out by **M.M.Kaif (19481A05E8)**, **P.Poojitha (19481A05H7), M.Dheeraj (19481A05F4), M.Siva Naga Sai (19481A05E7)** under the guidance and supervision of **Mrs.P.Naga Mani, M.Tech(Ph.D)** in the partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering of Jawaharlal Nehru Technological UniversityKakinada, Kakinada during the academic year 2022-23.

Project Guide
(Mrs.P.Naga Mani)

Head of the Department
(Dr. M. BABU RAO)

External Examiner

ACKNOWLEDGEMENT

The satisfaction that accompanies the successful completion of any task would be incomplete without the mention of people who made it possible and whose constant guidance and encouragements crown all the efforts with success.

We would like to express our deep sense of gratitude and sincere thanks to **Mrs.P.Naga Mani,** Assistant Professor, Department of Computer Scienceand Engineering for her constant guidance, supervision and motivation in completing the project work.

We feel elated to express our floral gratitude and sincere thanks to **Dr. M. Babu Rao**, Head of the Department, Computer Science and Engineering for his encouragements all the way during analysis of the project. His annotations, insinuations and criticisms are the key behind the successful completion of the project work.

We would like to take this opportunity to thank our beloved principal **Dr. G. V. S. N. R. V. Prasad** for providing a great support for us in completing our project and giving us the opportunity for doing project.

Our Special thanks to the faculty of our department and programmers of our computer lab. Finally, we thank our family members, non-teaching staff and our friends, who had directly or indirectly helped and supported us in completing our project in time.

Team members

M.M.Kaif (19481A05E8)
P.Poojitha (19481A05H7)
M.Dheeraj(19481A05F4)
M.Siva Naga Sai(19481A05E7)

CONTENTS

ABSTRACT

This project helps with the ability to inform blind individuals about things and their spatial placements, this initiative contributes to the transformation of the visual world into the audible world. The scene's identifiable objects are recognized by their labels and rendered as speech. Utilizing 3D binaural sound modeling, the 2-channel audio is used to encode their spatial positions. The system has many components. A portable camera device will be used to record video, which will then be sent to the server for real-time object identification and picture recognition (OPENCV). The position and size of the bounding boxes from the detection technique are used to estimate the 3D location of the items. The binaural sound with locations encoded is then rendered by a 3D sound creation program built on the Unity game engine. With wireless earbuds, the user receives the sound. When the detected item varies from the previous one, or every few seconds, whichever comes first, the sound is produced. The user will be able to effectively identify items that are three to five meters away with the aid of the gadget. For the current prototype, potential problems include detection failure when items are too close or too far away, and an information overload when the system tries to alert users of too many things.

LIST OF FIGURES

ABBREVIATIONS

Abbreviation	Explanation
SDIO	Secure Digital Input and Output
GPIO	General purpose input/output
SOC	System On Chip
UART	Universal Asynchronous Receiver/Transmitter
MIPI	Mobile Industry Processor Interface

CHAPTER 1

INTRODUCTION

1.1 Introduction

IoT- Internet Of Things
History Of IoT:
Kevin Ashton, the co-founder of the Auto-ID Center at the Massachusetts Institute of Technology (MIT), first mentioned the internet of things in a presentation he made to Procter &Gamble (P&G) in 1999. Wanting to bring radio frequency ID (RFID) to the attention of P&G's senior management, Ashton called his presentation "Internet of Things" to incorporate the cool new trend of 1999: the internet. MIT professor Neil Gershenfeld's book, *When Things Start to Think*, also appeared in 1999. It didn't use the exact term but provided a clear vision of where IoT was headed.

IoT has evolved from the convergence of wireless technologies, microelectromechanical systems (MEMSes), microservices and the internet. The convergence has helped tear down the silos between operational technology (OT) and information technology (IT), enabling unstructured machine-generated data to be analyzed for insights to drive improvements.

Although Ashton's was the first mention of the internet of things, the idea of connected devices has been around since the 1970s, under the monikers *embedded internet* and *pervasive computing*.

The first internet appliance, for example, was a Coke machine at Carnegie Mellon University in the early 1980s. Using the web, programmers could check the status of the machine and determine whether there would be a cold drink awaiting them, should they decide to make thetrip to the machine.

IoT evolved from M2M communication, i.e., machines connecting to each other via a network without human interaction. M2M refers to connecting a device to the cloud, managing it and collecting data.

Taking M2M to the next level, IoT is a sensor network of billions of smart devices that connect people, systems and other applications to collect and share data. As its foundation, M2M offers the connectivity that enables IoT.

The internet of things is also a natural extension of supervisory control and data acquisition (SCADA), a category of software application programs for process control, the gathering of data in real time from remote locations to control equipment and conditions. SCADA systemsinclude hardware and software components. The hardware gathers and feeds data into a computer that has SCADA software installed, where it is then processed and presented in a timely manner. The evolution of SCADA is such that late-generation SCADA systemsdeveloped into first-generation IoT systems.

The concept of the IoT ecosystem, however, didn't really come into its own until the middleof 2010 when, in part, the government of China said it would make IoT a strategic priority in its five-year plan.

The Internet of Things, or IoT, is a system of interrelated computing devices, mechanical and digital machines, objects, animals, or people that are provided with unique identifiers (UIDs) and the ability to transfer data over a network without requiring human-to-human or human- to-computer interaction.

A *thing* in the Internet of Things can be a person with a heart monitor implant, a farm animal with a biochip transponder, an automobile that has built-in sensors to alert the driver whentire pressure is low, or any other natural or man-made object that can be assigned an Internet Protocol (IP) address and is able to transfer data over a network.

Increasingly, organizations in a variety of industries are using IoT to operate more efficiently, better understand customers to deliver enhanced customer service, improve decision-making and increase the value of the business.

How IoT works?

An IoT ecosystem consists of web-enabled smart devices that use embedded systems, such as processors, sensors, and communication hardware, to collect, send and act on data they acquire from their environments. IoT devices share the sensor data they collect by connectingto an IoT gateway or other edge device where data is either sent to the cloud to be analyzed oranalyzed locally. Sometimes, these devices communicate with other related devices and act on the information they get from one another. The devices do most of the work without human intervention, although people can interact with the devices -- for instance, to set them up, give them instructions or access the data. The connectivity, networking and communication protocols used with these web-enabled devices largely depend on the specificIoT applications deployed.

IoT can also make use of artificial intelligence (AI) and machine learning to aid in making data collecting processes easier and more dynamic.

Why is IoT so important?

The internet of things helps people live and work smarter, as well as gain complete control over their lives. In addition to offering smart devices to automate homes, IoT is essential to business. IoT provides businesses with a real-time look into how their systems really work, delivering insights into everything from the performance of machines to supply chain and logistics operations.

IoT enables companies to automate processes and reduce labor costs. It also cuts down on waste and improves service delivery, making it less expensive to manufacture and deliver goods, as well as offering transparency into customer transactions.

As such, IoT is one of the most important technologies of everyday life, and it will continueto pick up steam as more businesses realize the potential of connected devices to keep them competitive.

IoT for organizations:

The Internet of Things offers several benefits to organizations. Some benefits are industry- specific, and some are applicable across multiple industries. Some of the common benefits of IoT enable businesses to:

- monitor their overall business processes;
- improve the customer experience (CX);
- save time and money;
- enhance employee productivity;
- integrate and adapt business models;
- make better business decisions; and
- generate more revenue.

IoT encourages companies to rethink the ways they approach their businesses and gives themthe tools to improve their business strategies.

Generally, IoT is most abundant in manufacturing, transportation and utility organizations, making use of sensors and other IoT devices; however, it has also found use cases for organizations within the agriculture, infrastructure and home automation industries, leadingsome organizations toward digital transformation.

IoT can benefit farmers in agriculture by making their job easier. Sensors can collect data onrainfall, humidity, temperature and soil content, as well as other factors, that would help automate farming techniques.

The ability to monitor operations surrounding infrastructure is also a factor that IoT can help with. Sensors, for example, could be used to monitor events or changes within structural buildings, bridges and other infrastructure. This brings benefits with it, such as cost saving, saved time, quality-of-life workflow changes and paperless workflow.

A home automation business can utilize IoT to monitor and manipulate mechanical and electrical systems in a building. On a broader scale, smart cities can help citizens reduce waste and energy consumption.

IoT touches every industry, including businesses within healthcare, finance, retail andmanufacturing.

Visual Impairment

Vision loss or visual impairment are other names for visual impairment. It is a visual issue that cannot be resolved with regular glasses. A person who is completely blind has lost allof their eyesight. It is difficult for persons of both types—whether they are blind or have mild to severe vision impairment—to carry out their daily tasks. They can only adjust to thisand that in a comfortable setting. However, things get more difficult for them when they are in a completely foreign setting. Millions of individuals throughout the world struggle to interpret their surroundings due to vision impairment. They may adopt different strategies todeal with everyday tasks, but they also have certain navigational issues and social difficulties. For instance, it is quite challenging for them to locate a certain room in a strange setting. Additionally, it might be challenging for those who are blind or visually impaired to determine who is being spoken to during a discussion. WHO research estimates that there are 217 million individuals with mild to severe vision impairment and 36 million persons who are blind. WHO study lists cataracts, trachoma and river blindness infections, glaucoma, diabetic retinopathy, uncorrected refractive errors, and certain occurrences of infant blindness among visual impairments. Many persons with severe visual impairment get advantages from vision therapy, environment adjustments, and new equipment. We are now able to solve almost any problem because of the advances in technology. In recent years, "deep neural network" technology, in particular, has advanced quickly. Blind users ofcommon computers can employ access technologies including screen readers, screen amplifiers, dynamic Braille displays applications, and cell phones. Tools are becoming more widely available, and there are coordinated efforts to ensure that information technology is accessible to all potential users, including the blind. An accessibility wizard, a magnification for those with low vision, and Microsoft Narrator, a straightforward screen reader are all included in later versions of Microsoft Windows. Vinux and Adriane Knoppix are Linux distributions (as live CDs) for the blind; Adriane Knopper, a person with vision impairment, helped build the latter. In addition to VoiceOver, which is included withmacOS and iOS, most Android devices also have Google TalkBack as a built-in screen reader.

"Blindsight," "TapTapSee," and other current solutions are available to assist these exceptional folks. "Blindsight" provides a smartphone app called Text Detective that uses OCR technology, or optical character recognition, to find and interpret a text from images taken with a camera [2]. "TapTapSee" is a smartphone application that defines an image taken by blind people in roughly 10 seconds using computer vision and crowdsourcing. Facebook is also working on image captioning technology to enable visually impaired people to interact with other users through images. Here, the image is translated into spoken language. However, neither of these solutions utilized spatial sound approaches to enhance the user experience nor were they concentrated on enhancing general visual sense for blind individuals. Some of the work created by Neil Harbisson, Daniel Kish, and VOICE technology developers falls within the general category of sensory replacement. Neil Harbisson, a colorblind artist, created a tool that converts colour information into sound frequencies. Daniel Kish, who is completely blind, has independently developed accurate echolocation skills utilising "mouth clicks" for solo riding and trekking [19]. The introduction of VOICE technology is an extreme attempt to translate visual perception to sound. The VOICE system examines each camera image from left to right, correlating brightness with loudness and height with pitch.

1.2 Objectives

The development of technology has made it possible for us to solve numerous issues. Even humans have advanced to the point where a team recently succeeded in clicking an image ofa black hole thanks to some amazing computational work. But as of right now, it's clear that blind individuals have a lot of difficulties carrying out their daily tasks. They rely on other people to perform all or many of their tasks. That inspired us to create a prototype. This is a modest effort to help blind individuals become independent. An autonomous navigation system for blind individuals is being developed as part of this research. It's an effort in whichI won't donate the eyeballs but will instead give them to these unique folks nearby. The project's primary goal is to help blind individual become familiar with their surroundings. This would help the individual manage his or her workload and inform them of what is coming next. They could recognize fundamental and frequent items in their daily lives with the aid of technology. The use of object detection algorithms can create new opportunitiesfor helping with both interior and outdoor navigation for differently abled people

1.3 Problem Statement

- People who are differently-abled have to depend upon the otherpeople for help while walking.

- Even though there are different approaches for solving this problemthey are have many limitations such as limited detection of things.

- Using such different approaches which might also cost more.

As we have seen each approach have it’s own limitations and keeping this in care,the proposed system has been proposed by overcoming some of the limitations.

CHAPTER 2

LITERATURE REVIEW

Several approaches have been proposed related to this issue in many papers. Ofthese, some specific papers have been analyzed in the following paragraphs.

Computer Vision Technologies

There exist multiple tools to use computer vision technologies to assist blind people. Some of these known technologies are:

TapTapSee: The mobile app TapTapSee uses computer vision and crowdsourcing to describe a picture captured by blind users in about 10 seconds.

Text Detective: The Blindsight offers a mobile app Text Detective featuring optical character recognition (OCR) technology to detect and read text from pictures capturedfrom the camera.

Facebook: Facebook is developing image captioning technology to help blind users engaging in conversations with other users about pictures.

However, these products were not focusing on enabling general visual sense for blind people and did not use the spatial sound techniques to further enhance the user experience.

Sensory Substitution Technologies

Some works exist in the general scope of sensory substitution. Some of these known technologies are:

- **Mouth Clicks:** Daniel Kish, who was totally blind, developed accurate echolocation ability using "mouth clicks" for navigation tasks including biking and hikingindependently.
- Colorblind artist Neil Harbisson developed a device to transform color information into sound frequencies. An extreme attempt of converting visual sense to sound is introduced by the vOICe technology. The vOICe system scans each camera snapshot from left to right, while associating height with pitch and brightness with loudness.[16]

However, all these attempts on sensory substitution arereported with very difficult learning process.

Electronic Travel Aids (ETA):

This technology offers a blind navigation system using RFID tags to set up a location alert infrastructure in buildings so that the blind can use an RFID-equipped ETA (such as a cell phone) to determine their location as well as software which can use this localization to generate vocal directions to reach a destination. Electronic travel aids (ETAs) are electronic devices designed to improve autonomous navigation of blind people. The ETA design differs from the sizes, the type of sensor used in the system, the method of transmitting information and also the usage method. Image of the scene before blind user is captured using the video camera and it is converted into sound pattern. The intensity of the pixel of the image is converted into hardness

LIMITATIONS OF EXISTING SYSTEMS

- These products were not focused on enabling general visualsense for blind people.
- They did not use spatial sound techniques to further enhancethe user experience.
- All these attempts at sensory substitution are reportedwith the very difficult learning processes.
- The ETA system uses expensive ultrasonic sensors thatcontain few electronic component

PROPOSED SYSTEM

Several processes are used in this work's object identification approach: object detection, thecreation of unique descriptors for each object, retrieval from a model database, and matching. The building of a model database is a necessary step before the matching process can be used. It has characteristics for all typical items seen in blind people's settings. An audio will be generated specifying the object before the user.

CHAPTER 3

PROPOSED METHOD

3.1 Methodology

Convolutional Neural Networks

CNNs have wide applications in image and video recognition, recommender systems and natural language processing. Here the example I will take is related to Computer Vision. However, the basic concept remains the same and can beapplied to any other use case!

A convolution neural network consists of one or more convolution layers (often with a subsampling step) and then followed by one or more fully connected layers as in a standard multilayer neural network. The architecture of a CNN is designed to take advantage of the 2D structure of an input image (or other 2D input such as a speech signal). It is achieved with local connections and tied weights, followed by a form of pooling that results in translation-invariant features. Another advantage of CNNs is that they are easier to train and have far fewer parameters than fully-connected networks with the same number of hidden units. Here I have discussed the architecture of a CNN and the reprocessing algorithm to calculate the gradient with respect to the parameters of the model using gradient-basedoptimization.

A CNN consists of a number of convolution and sub- layer layers, optionally followed by fully-linked layers. The input of a convolution layer is a m x m x r image where m is the height and width of the image and r is the number of channels, e.g. An RGB image has r = 3. The convolution layer will have k filters (or kernels) of size n x n x q, where n is smaller than the size of the image and q can be either the same as the number of channels r or smaller and can for each core varies. The size of the filters gives rise to the locally coupled structure that is each connected to the image to produce k-function cards of size m-n + 1. Each card is then subsampled typically with average or maximum pole over p x p. adjacent regions where p falls between 2 for small images (e.g., MNIST) and is usually not more than 5 for larger inputs. Before or after the sub-sampling layer, an additive bias and sigmoidal nonlinearity are applied to each function map. The figure below illustrates a complete layer in a CNN consisting of confluent and subsampling sublayers. Units of the same colour attached weights.

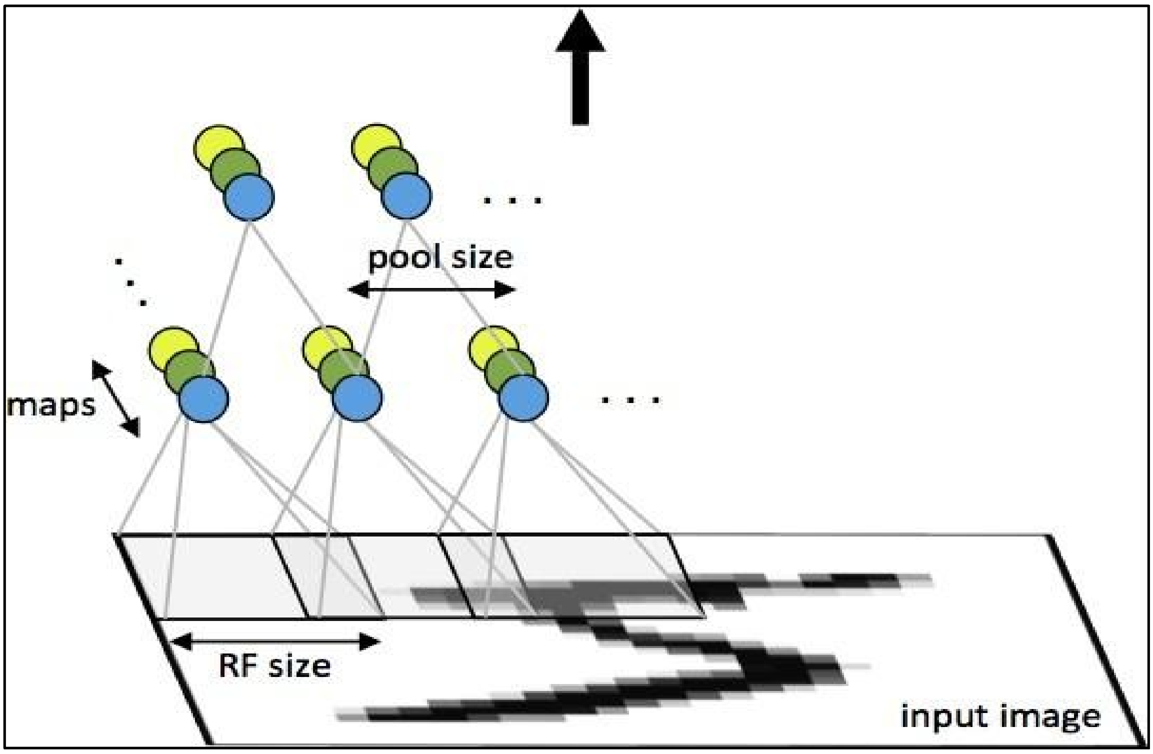


Fig: First layer of a convolutional neuralnetwork with pooling.

After the convolutional layers there may be any number of fully connected layers. The densely connected layers are identical to the layers in a standard multilayer neural network.

OpenCV

OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today's systems. By using it, one can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To Identify image pattern and its various features we use vector space and perform mathematical operations on these features.

The first OpenCV version was 1.0. OpenCV is released under a BSD license and hence it's free for both **academic** and **commercial** use. It has C++, C, Python and Java interfaces and supports Windows, Linux, Mac OS, iOS and Android. When OpenCV was designed the main focus was real-time applications for computational efficiency. All things are written in optimized C/C++ to take advantage of multi-core processing.

Fig: OpenCV

3.2 Implementation

In software engineering, a functional requirement defines
a function of a software system or its component. A function is described as a set of inputs, the behaviour, and outputs. Functional requirements may be calculations, technical details, data manipulation and processing and other specific functionality that define what a system is supposed to accomplish. Behavioural requirements describing all the cases where the system uses the functional requirements are captured in use cases.

Here, the system has to perform the following tasks:

- Capture images in the form of video of at least 4 fps from the user side.
- Live stream the video through the Raspberry Pi to the Server
- The server computes the result from the images in the form of video and detects object and tags each object.
- These tagged objects are converted to speech.
- These speeches are converted to audible form which are 3D sound.
- The 3D sound is transmitted back to the user.

RASPBERRY PI

If raspberry pi has to be represented in one sentence it will definitely be “It’s a low cost, credit card sized computer”. It’s a device which can be used plugged in with TV or computer and uses standard mouse or keyboard. It’s capable of performing all the computation tasks. The Raspberry Pi Foundation is a UK-based charity that works to put the power of computing and digital making into the hands of people all over the world. They do this so that more people are able to harness the power of computing and digital technologies for work, to solve problems that matter to them, and to express themselves creatively. With raspberry pi I can make use of programming languages like python and scratch. It’s a device which is easy to use and can be used by people of all the ages. It’s a device which is able to perform everything that a computer can do like from browsing internet to making spreadsheets and playing games to playing high definition video.

It’s a device capable of interacting with the outside world and is a choice of good stack of digital maker projects, from music machines and parent detectors to weather stations and tweeting bird houses with infra- red cameras.

In this project raspberry pi acts as a client. This along with the raspberry pi noir camera is used to live stream the video to the server end. For the better communication between the server and client the connection is wireless and the only condition is both server and client must be present in the same network.

Unity 3D

Unity is a cross-platform real-time engine developed by Unity Technologies, first announced and released in June 2005at Apple Inc.'s Worldwide Developers Conference as an OS X- exclusive game engine. As of 2018, the engine has been extended to support 27 platforms. The engine can be used to create both three- dimensional and two-dimensional games as well as simulations for its many platforms. Several major versions of Unity have been released since its launch, with the latest stable version being Unity 2019.1.0.

Unity gives users the ability to create games and interactive experiences in both 2D and 3D, and the engine offers a primary scripting API in C#, for both the Unity editor in the form of plugins, and games themselves, as well as drag and drop functionality. Prior to C# being the primary programming language used for the engine, it previously supported Boo, which was removed in the Unity 5 release, and a version of JavaScript called UnityScript, which was deprecated in August 2017 after the release of Unity 2017.1 in favor of C#.

The engine has support for the following graphics APIs: Direct3D on Windows and Xbox One; OpenGL on Linux, macOS, Windows; OpenGL ES on Android and iOS; WebGL on the web; and proprietary APIs on the video game consoles. Additionally, Unity supports the low-level APIs Metal on iOS and macOS and Vulkan on Android, Linux, and Windows, as well as Direct3D 12 on Windows and Xbox One.

Within 2D games, Unity allows importation of sprites and an advanced 2D world renderer. For 3D games and simulations, Unity allows specification of texture compression, mipmaps, and resolution settings for each platform that the game engine supports, and provides support for bump mapping, reflection mapping, parallax mapping, screen space ambient occlusion (SSAO), dynamic shadows using shadow maps, render-to-texture and full-screen post-processing effects.

Since about 2016 Unity also offers cloud-based services to developers, these are presently: Unity Ads, Unity Analytics, Unity Certification, Unity Cloud Build, Unity Everyplay, Unity IAP ("In app purchase" - for the Apple and Google app stores), Unity Multiplayer, Unity Performance Reporting, Unity Collaborate and Unity Hub.

Unity supports the creation of custom vertex, fragment (or pixel), tessellation and compute shaders. The shaders can be written using Cg, or Microsoft's HLSL.

CONNECTIVITY

 # SERVER SCRIPT

```
server_socket = socket.socket() server_socket.bind(('0.0.0.0', 777))server_socket.listen(0)
connection = server_socket.accept()[0].makefile('rb')try:
() image_stream.write(connection.read(image_len))image_stream.seek(0)
image = Image.open(image_stream)
data = np.fromstring(image_stream.getvalue(), dtype=np.uint8)imagedisp = cv2.imdecode(data, 1)
cv2.imshow("Frame", imagedisp)cv2.waitKey(1) cv2.destroyAllWindows()
```

 # CLIENT SCRIPT

```
client_socket = socket.socket() client_socket.connect(('my_server', 8000))connection =
client_socket.makefile('wb') try:
picamera.PiCamera() as camera: camera.resolution = (640, 480)camera.start_preview()
time.sleep(2) start=time.time() stream = io.BytesIO()
for foo in camera.capture_continuous(stream, 'jpeg'): connection.write(struct.pack('<L', stream.tell()))
connection.flush()
stream.seek(0) connection.write(stream.read()) if time.time()  - start > 30: break
finally:
seek(0)stream.trun cate()
connection.write(struct.pack('<L', 0))
connection.close() client_socket.close()
```

OBJECT DETECTION PSEUDO CODE

```
while True:

image_len = struct.unpack('<L', connection.read(struct.calcsize('<L')))[0]if image_len:
image_stream = io.BytesIO() image_stream

image_stream.seek(0)

image = Image.open(image_stream)

data = np.frombuffer(image_stream.getvalue(), dtype=np.uint8)frame = cv2.imdecode(data, 1)
img, orig_im, dim = prep_image(frame, inp_dim)im_dim = torch.FloatTensor(dim).repeat(1, 2)
image_stream.seek(0)
image = Image.open(image_stream) data = np.frombuffer(image_stream.getvalue(),
dtype=np.uint8) frame = cv2.imdecode(data,1)
img, orig_im, dim = prep_image(frame, inp_dim) im_dim = torch.FloatTensor(dim).repeat(1, 2)
if CUDA:
)
output = model(Variable(img), CUDA)
output = write_results(output, confidence, num_classes,nms=True, nms_conf=nms_thesh)
if
continue
output[:, 1:5] = torch.clamp(output[:, 1:5], 0.0, float(inp_dim)) / inp_dim
output[:, [1, 3]] *= frame.shape[1]classes = load_classes('Working/da ta/coco.names')
list(map(lambda x: write(x,orig_im), output)) cv2.imshow("frame", orig_im)
key = cv2.waitKey(1)
```

3.3 DATA PREPARATION

In the proposed system, the disadvantages of the existing system will be takencare. There would be 5 main modules the system will be working around.

OBJECT DETECTION ALGORITHM

To successful detection of surrounding objects, several existing detection systems could classify objects and evaluate it at various locations in an image. Deformable Parts Model (DPM) uses root filters that slide detection windows over the entire image [20].

RCNN uses region proposal methods to generate possible bounding boxes in an image. Then, it applies various ConvNets to classify each box. The results are then postprocessed and output finer boxes. The slow test-time, complex training pipeline and the large storage does not fit into the application. [21]

Fast R-CNN max-pools proposed regions, combines the computation of ConvNet for each proposal of an image, and outputs features of all regions at once. [22]
Based on Fast R-CNN, Faster R-CNN inserts a region proposal network after the last layer of ConvNet. [23]

Both methods speed up the computational time and improve the accuracy. The pipelines of these methods are still relatively complex and hard to optimize.

Considering the requirement of real-time objective detection, in this project, have used the OpenCV model. OpenCV could efficiently provide relatively good objective detection with extremely fast speed.[19]

DIRECTION ESTIMATION

After detecting the type of objects in a video frame, the next step is to obtain the direction and distance detected object from the user.
The approximated distance is 5-6m for detection. First ofall, human is good at inferring direction from binaural sound,and the relative distance, namely object A is closer than object Bor object is moving closer and closer between frames. However, absolute distance is difficult to deduce from binaural sound. This means my image processing algorithm needs to provide the accurate directional information and the relative distance, butnot the exact depth.

The direction is estimated by dividing the regioncaptured the camera into the coordinate system with (0,0) at the center. The spatial location of the object is conveyed through 3d sound. Ex: If the object is to the left of center then same is conveyed through the 3d sound played through the earphones. The intensity of sound decides how close or how far is an object. The approximated distance is for detection is 5-6m but finding the exact distance is bit problem. My system is able to find the direction correctly and an approximated distance.

Thus, I resort to estimate the direction and relative distance. I have used Raspberry pi Noir camera along with Raspberry pi 3b+. Giving the field of view of the camera, andthe bounding box of the object, the direction is estimated from the central pixel location of the bounding box.

DATA STREAMING

The architecture is shown in Figure. In this platform, a Raspberry pi Noir camera and transfers through video link directly to the OpenCV model running on a local server machine with high performance GPU capture the environment. The server detects objects, sends information directly to the unity sound generator and plays the binaural sound.

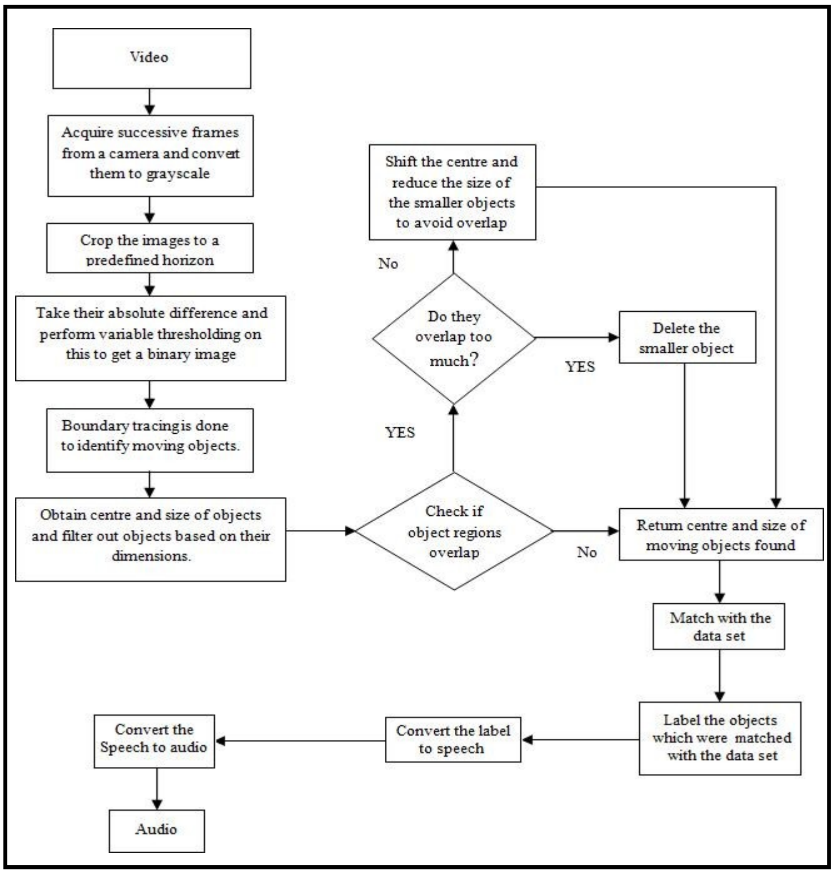


Fig: Data flow pipeline of the system.

In particular, a portable camera at 30 frames 1080p-720p resolution captures the environment picture. The video is live- streamed through the video link to the computer server. The object detection engine OpenCV then predicts objects in the stream. The OpenCV algorithm could process a single image frame at a speed of 4-60 frames/second depending on the image size I send to the engine. The outputs are sent to unity sound generator and the generated sounds are played through wireless earbuds. During the implementation, the platform is capable of processing all captured live stream at a minimum speed of 30 frames per second at 1080p-720p resolution.

RESULT FILTERING

OpenCV outputs the top classes and their probability for each frame. A probability above 20% will be taken as a confident detection result. To present the results to the user in a reasonable manner, my algorithm also has to decide whether to speak out a detected object and at what time. Obviously, it’s undesirable to keep speaking out the same object to the user even if the detection result is correct. It’s also undesirable if two object names are spoken overlapping or very closely that the user won’tbe able to distinguish.

To solve the first problem, I assume a cool-down-time of five seconds for each class. For example, if a person is detected inthe first frame and is spoken out, the program will not speak out “person” again until after five seconds. This is only a sub- optimal solution since it does not deal with multiple objects of the same class. Ideally, if there are two persons in the frame, the user should be informed about the two people, but he does not need to be informed about the same person continuously. One possible improvement, which I have to work upon, is to track the object using overlapping bounding box between frames. To solve the second problem, delay of half a second can be enforced between any spoken classes.

3D SOUND GENERATION

Plug-in for Unity 3D game engine will be used to simulate the 3D sound. Unity- based game program “3D Sound Generator” is developed using either a file watcher or TCP socket to receive the information about the correct sound clips tobe played as well as their spatial coordinates.

CHAPTER 4

RESULTS AND DISCUSSION

Since the output is in the form of audio the only output that can be documented and be used for verification is to display the name of the object detected and display it on the monitor in real-time. The object detected is labeled in at the top of the object in the video frame of each object. This module is coded separately just for the verification of the output generated and isnot the actual part of the project as it slows down the performance of the system. Following are the images and screenshots of the system and the output.

Fig: Raspberry pi top view

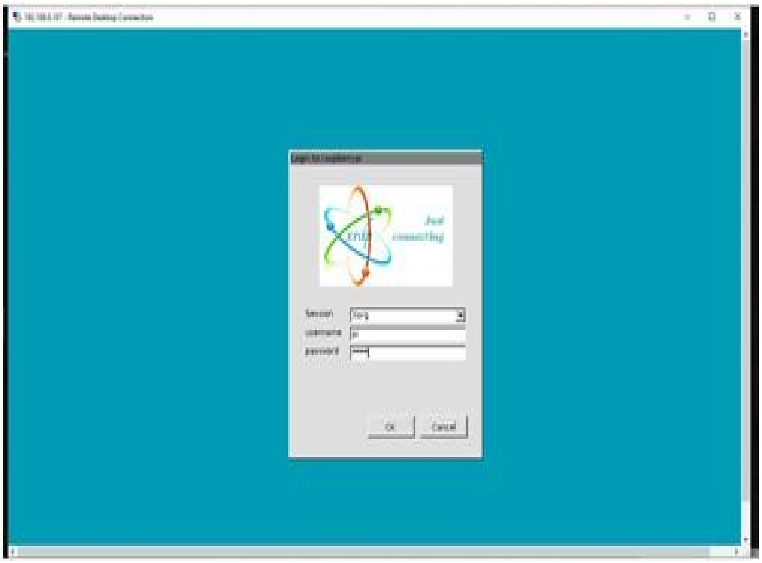


Fig: Login Panel

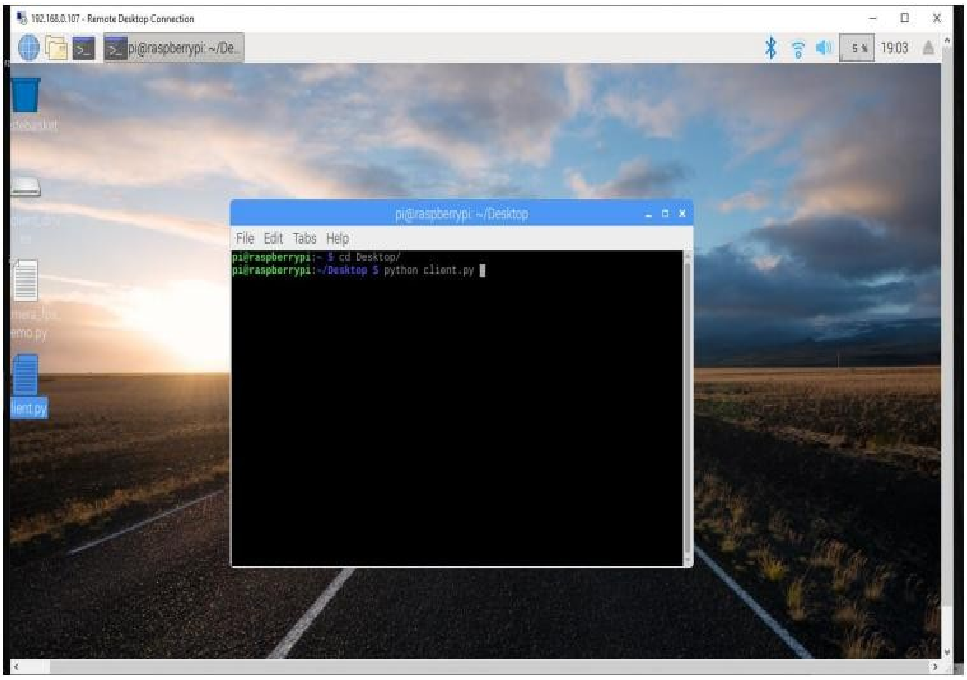


Fig: Raspberry pi terminal (client)

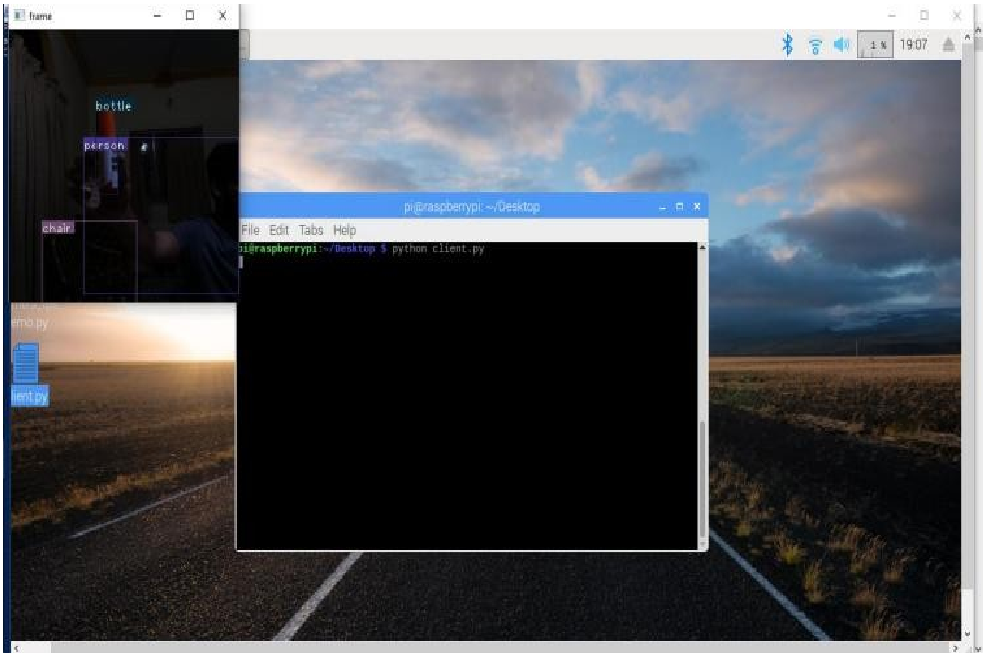


Fig: Object detected from a live stream video (Bottle)

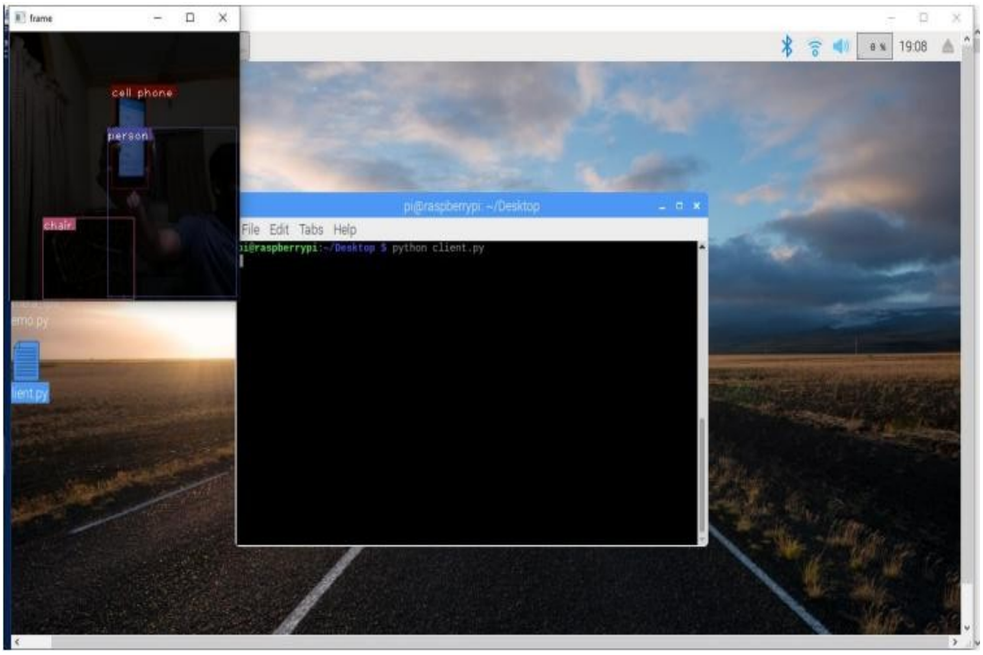


Fig: Object detected from a live stream video (Cell phone)

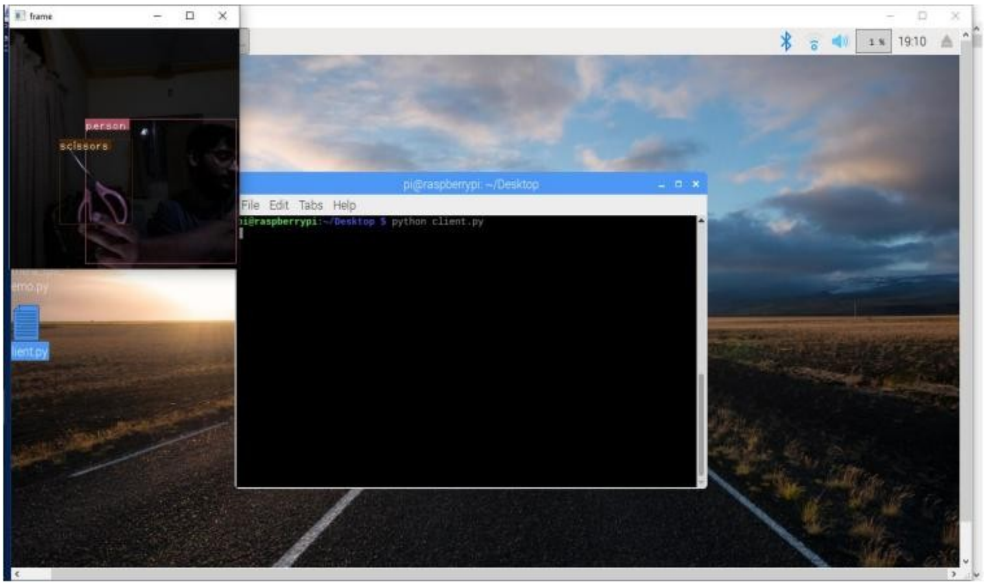


Fig: Object detected from a live stream video (Scissors)

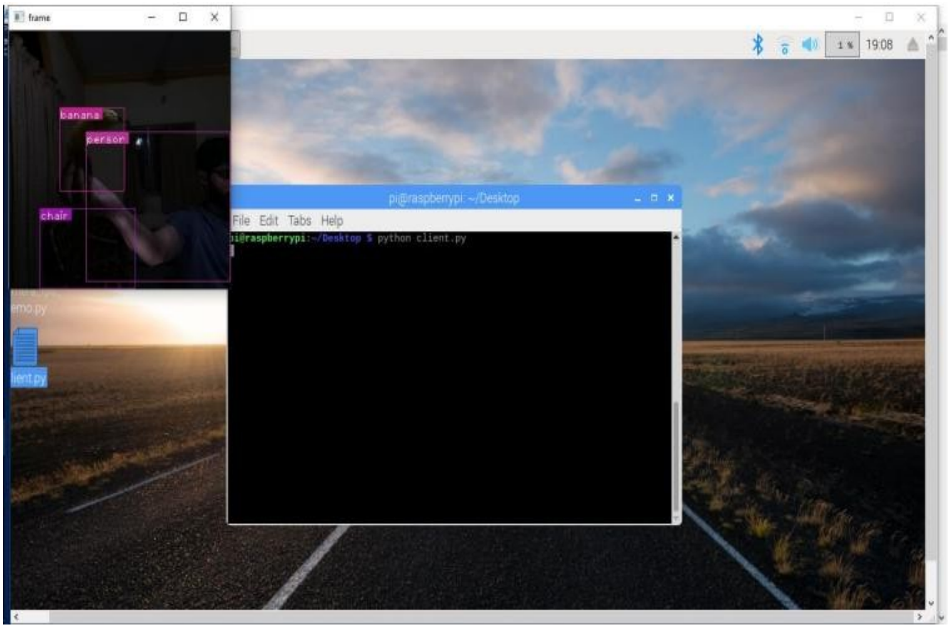


Fig: Object detected from a live stream video (Banana)

Fig: Depiction of the whole system

CHAPTER 5
CONCLUSION AND FUTURE SCOPE

Conclusion

To best assist the blind person, an object recognition code is built and combined with unity code for 3D sound engine. The system necessitates extensive network data transfers. In case of flawless data transport, the system offers the best outcomes. Many items that we come across in our daily lives are named and explained to the person.

In accordance with the concept described in the paper, we have developed a model for object detection. The solution is capable of performing precise real-time objective detection by utilizing the OpenCV algorithm and transmission of 3d audio. This project offers a real- time, portable solution.

Future Scope

As this project can be extended in future to detect images without concerning about the distance from the user. This may be done by the sensors such as the ultrasonic sensors.

BIBILOGRAPHY

[1] Let Blind People See: Real-Time Visual Recognition with Results Converted to 3D Audio Rui (Forest) Jiang Earth Science, Stanford forestj@stanford.edu Qian Lin Applied Physics, Stanford linqian@stanford.edu Shuhui Qu Civil and Environmental Engineering.

[2] Object detection and identification for a blind in video scene – Hanen Jabnoun, Hamid Amiri.

[3] Use of 4G waveform towards RADAR Nigidita pradhan1, Rabindranath Bera2, Debasish Bhaskar3 PG Student [DEAC], Dept. of ECE, Sikkim Manipal Institute of Technology, India1 Head of the Dept [HOD], Dept. of ECE, Sikkim Manipal Institute of Technology, India2 Assistant Professor, Dept. of ECE, Sikkim Manipal Institute of Technology, India3 Majitar, Rangpo, East Sikkim, India.

[4] Motion-Based Object Detection and Tracking in Color Image Sequences Bernd HeiseleyImage Understanding Group (FT3/AB) DaimlerChrysler Research Center Ulm, D-89013, Germany.

[5] Real-time multiple vehicle detection and tracking from a moving vehicle Margrit Betke1, Esin Haritaoglu2, Larry S. Davis2 1 Boston College, Computer Science Department, Chestnut Hill, MA 02167, USA; e-mail: betke@cs.bu.edu 2 University of Maryland, College Park, MD 20742, USA.

[6] Finding Objects for Blind People Based on SURF Features Ricardo Chincha and YingLi Tian Department of Electrical Engineering the City College of New York New York, NY, 1003.

[7] Portable Camera-based Assistive Text and Product Label Reading from Hand-held Objects for Blind Persons Chucai Yi, Student Member, IEEE, YingLi Tian, Senior Member,IEEE, Aries Arditi. Blind Leap Dept of CSE, NHCE 44

[8] An Approach for Object and Scene Detection for Blind Peoples Using Vocal Vision. Manisha Rajendra Nalawade, Vrushali Wagh, Shradha Kamble Shraddha A. Kamble Int. Journal of Engineering Research and Applications www.ijera.com ISSN: 2248- 9622, Vol. 4, Issue 12(Part 5), December 2014, pp.01-03.

[9] Research on Moving Vehicles Tracking Algorithm Based on Feature Points and AKF Jinhua Wang; Yu Li; Chongyu Ren; Jie Cao.

[10] A Step-by-Step Introduction to the Basic Object Detection Algorithms (Part 1) PULKIT SHARMA, OCTOBER 11, 2018

[11] https://www.mathworks.com/solutions/deep-learning/object-recognition.html Quattoni, and A. Torralba. Recognizing Indoor Scenes. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009.

[12] Saurabh Gupta, Ross Girshick, Pablo Arbelaez and Jitendra Malik, Learning Rich Features from RGBD Images for Object Detection and Segmentation (ECCV), 2014.

[13] Tadas Naltrusaitis, Peter Robison, and Louis-Phileppe Morency, 3D Constrained Local Model for Rigid and Non-Rigid Facial Tracking (CVPR), 2012.

[14] Andrej Karpathy and Fei-Fei Li, Deep Visual-Semantic Alignments for Generating Image Descriptions (CVPR), 2015.

[15] David Brown, Tom Macpherson, and Jamie Ward, seeing with sound? exploring different characteristics of a visual-to-auditory sensory substitution device. Perception, 40(9):1120–1135, 2011.

[16] Liam Betsworth, Nitendra Rajput, Saurabh Srivastava, and Matt Jones. Audvert: Using spatial audio to gain a sense of place. In Human-Computer Interaction–INTERACT 2013, pages 455–462. Springer, 2013.

[17] Jizhong Xiao, Kevin Ramdath, Manor Iosilevish, Dharmdeo Sigh, and Anastasis Tsakas. A low cost outdoor assistive navigation system for blind people. In Industrial Blind Leap Dept of CSE, NHCE 45 Electronics and Applications (ICIEA), 2013 8th IEEE Conference on, pages 828–833. IEEE, 2013.

[18] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. arXiv preprint arXiv:1506.02640, 2015.

[19] Pedro F Felzenszwalb, Ross B Girshick, David McAllester, and Deva Ramanan. Objectdetection with discriminatively trained part-based models. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 32(9):1627–1645, 2010.

SESHADRI RAO GUDLAVALLERU ENGINEERINGCOLLEGE

(An Autonomous Institute with Permanent Affiliation to JNTUK, Kakinada)
Seshadri RaoKnowledge Village, Gudlavalleru)

Department of Computer Science and Engineering

Program Outcomes (POs)

Engineering Graduates will be able to:

1. **Engineering knowledge**: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
2. **Problem analysis**: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.
3. **Design/development of solutions**: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
4. **Conduct investigations of complex problems**: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions, component, or software to meet the desired needs.
5. **Modern tool usage**: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.
6. **The engineer and society**: Apply reasoning informed by thecontextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

7. **Environment and sustainability**: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.
8. **Ethics**: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

9. **Individual and team work**: Function effectively as an individual,and as a member or leader

in diverse teams, and in multidisciplinary settings.

10. **Communication**: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
11. **Project management and finance**: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team,to manage projects and in multidisciplinary environments.
12. **Life-long learning**: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

Program Specific Outcomes (PSOs)

PSO1 : Design, develop, test and maintain reliable software systems andintelligent systems.

PSO2 : Design and develop web sites, web apps and mobile apps.

PROJECT PROFORMA

Classification Of Project	Application	Product	Research	Review
	√			

Note: Tick Appropriate category

Course Outcomes	
Course Outcome (CO1)	Identify and analyze the problem statement using prior technical knowledge in the domain of interest.
Course Outcome (CO2)	Design and develop engineering solutions to complex problems by employing systematic approach.
Course Outcome (CO3)	Examine ethical, environmental, legal and security issuesduring project implementation.

Course Outcome (CO4)	Prepare and present technical reports by utilizing differentvisualization tools and evaluation metrics.

CS1537 : MAIN PROJECT															
Course Outcomes	Program Outcomes and Program Specific Outcomes														
	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12		PSO 1	PSO 2
CO1	3	3	1					2	2	2				1	1
CO2	3	3	3	3	3			2	2	2		1		3	3
CO3	2	2	3	2	2	3	3	3	2	2	2			3	
CO4	2		1		3				3	3	2	2		2	2

Mapping Table

Note: Map each project outcomes with POs and PSOs with either 1or 2 or 3 based on level of mapping as follows:

1-Slightly (Low) mapped 2-Moderately (Medium) mapped 3-Substantially (High) mapped

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

Object Identification and Recognition System For Blind Persons

P.Naga Mani[1] , Department of CSE, Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru
nagamanichittibomma@gmail.com
Mohammed kaif[2], Department of CSE, Seshadri rao Gudlavalleru Engineering College, Gudlavalleru
mohammed.kaif0209@gmail.com
Poojitha[3], Department of CSE, Seshadri rao Gudlavalleru Engineering College, Gudlavalleru
poojithapakala@gmail.com
Merugu Deeraj[4] , Department of CSE, Seshadri rao Gudlavalleru Engineering College, Gudlavalleru
merugudheeraj@gmail.com
M.Siva Naga Sai[5] , Department of CSE, Seshadri rao Gudlavalleru Engineering College, Gudlavalleru
bannu5310@gmail.com

Abstract:
This paper helps with the ability to inform blind individuals about things and their spatial palcements, this initiative contibute to the transformation of the visual world into the audible world. The scene's identifiable objects are recognized by their lables and rendered as speech. Utilizing 3D binarual source modeling, the 2-channel audio is used to encode their spatial positions. The system has many components. A portable camera device will be used to record viedo, which will then sent to the server for real-time object identification and picture recognition (OpenCV).The position and size of the bounding boxes from the detection technique are used to estimate the 3D sound creation program bulit on the unity game engine. With wireless earbuds, the user receives the sound. When the detected item varies from the previous one, or every few seconds, whichever come first, the sound is produced. With the help of the device, the user will able to successfully detect objects that are three to five meters away. Potential issues for the current prototype include information overload when the system tries to tell users of too many things and detect failure when objects are too close or far away

Introduction:
Vision loss or visual impairment are other names for visual impairment. It is a visual issue that cannot be resolved with regular glasses. A person who is completely blind has lost all of their eyesight. It is difficult for persons of both types—whether they are blind or have mild to severe vision impairment—to carry out their daily tasks. They can only adjust to this and that in a comfortable setting. However, things get more difficult for them when they are in a completely foreign setting. Millions of individuals throughout the world struggle to interpret their surroundings due to vision impairment. They may adopt different strategies to deal with everyday tasks, but they also have certain navigational issues and social difficulties. For instance, it is quite challenging for them to locate a certain room in a strange setting. Additionally, it might be challenging for those who are blind or visually impaired to determine who is being spoken to during a discussion. WHO research estimates that there are 217 million individuals with mild to severe vision impairment and 36 million persons who are blind. WHO study lists cataracts, trachoma and river blindness infections, glaucoma, diabetic retinopathy, uncorrected refractive errors, and certain occurrences of infant blindness among visual impairments. Many persons with severe visual impairment get advantages from vision therapy, environment adjustments, and new equipment. We are now able to solve almost any problem because of the advances in technology. In recent

ISSN PRINT 2319 1775 Online 2320 7876

Research paper UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

years, "deep neural network" technology, in particular, has advanced quickly. Blind users of common computers can employ access technologies including screen readers, screen amplifiers, dynamic Braille displays applications, and cell phones. Tools are becoming more widely available, and there are coordinated efforts to ensure that information technology is accessible to all potential users, including the blind. An accessibility wizard, a magnification for those with low vision, and Microsoft Narrator, a straightforward screen reader are all included in later versions of Microsoft Windows. Vinux and Adriane Knoppix are Linux distributions (as live CDs) for the blind; Adriane Knopper, a person with vision impairment, helped build the latter. In addition to VoiceOver, which is included with macOS and iOS, most Android devices also have Google TalkBack as a built-in screen reader.

Literature Survey

A thorough investigation and effort have been done on this topic by numerous researchers. From among them, the following works have been listed:

Developers of [1] provided a portable, real-time solution. Using mobile cameras, a rapid HD video connection, and a powerful server, they showed off a platform that generates 3D sounds. The device used a powerful wireless transmitter and algorithm to do accurate real-time objective detection with live streaming at a rate of 30 frames per second and 1080P quality. The study creates a model for sensory replacement (from visual to hearing). Through the research, they showed how computer vision techniques may be used as a form of assistive technology.

The suggested visual replacement method in [2] is based on the recognition of things in the blind person's immediate environment. They suggested a method of identifying and locating 2D in the video. This approach identifies the properties of objects that are resistant to changes in viewpoint, offers identification, and lessens the difficulty of detection. The technique is focused on extracting and matching essential parts from videos. To identify objects in each frame, a comparison between the query frame and database objects is performed. An audio file with the object's details is activated for each object that is detected. As a result, both object detection and identification are handled. For object identification in this study, SIFTS key point extraction and feature matching were employed.

Researchers from the group [3] created an app with a buzzer and vibration mode that uses sensors to guide the blind. Using a key-matching method, the item is found by comparing it to an image stored in the database.

[4] presents a method for identifying items in a series of color photos captured by a moving camera. The estimate of motion in the picture plane is the initial step of the method. The motion of clusters is calculated and constructed by a grouping of pixels in a color/position feature space, as opposed to optical flow, monitoring single points, edges, or areas throughout a succession of pictures. A motion-based segmentation is the second phase when nearby clusters with comparable trajectory combinations are joined to create object hypotheses. Researchers forecast dynamic changes in cluster placements using Kalman filters. The primary use was for aid with vision.

In order to address the concerns of real-time factor and accuracy factors in video image tracking, a method for tracking moving vehicles in real-time based on feature points was put out in the paper [9]. They developed a simple and effective feature points matching method based on the Adaptive Kalman Filter (AKF). The results of simulation testing show that the algorithm does a good job of quickly monitoring moving targets, which not only allows for accurate target characterization but also reduces the amount of time required for matching calculations.

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

In an article [5], a different group of researchers concentrated on A forward-looking video camera in a car that is traveling down a highway that has been outfitted with a real-time vision system that can evaluate color films. To identify and track lane markers, road edges, and other moving objects, the system combines motion, edge, and color information. The internet input data is chopped into templates that match cars, and characteristics of the highway scene are detected and their relationships to one another are assessed. Additionally, automobiles may be found using motion tracking and temporal differentiation of car-specific motion characteristics. The system uses a recursive least-squares filter to identify and track lane markers and road borders. The findings of the experiments show reliable, real-time automobile tracking and detection across thousands of camera frames. The data contains footage that was captured in poor visibility circumstances.

However, Ricardo Chincha and Yingli Tian have developed an object identification technique in the article [6] named "Finding Objects for Blind People Based on SURF Features" to assist blind people in finding lost goods by utilizing Speeded-Up Robust Features (SURF).

The user-queried object's particular features are compared to a database of features from other personal things that have been stored in advance as part of the proposed recognition procedure. Out of 100 total items in the studies, 84 were spotted, demonstrating the need for higher performance; as a result, SIFT may be employed in place of SURF to improve object recognition.

"Portable Camera-Based Assistive Text and Product Label Reading From Hand-Held Items for Blind People" is the title of their research paper. A framework for camera-based assistive text reading was proposed by Chucai. They made a system that helps the blind to read the labels of the products, which are held by hand. They recommended a quick and easy motion-based method for determining the keyframes in a video. The effectiveness of the suggested text localization algorithm is assessed quantitatively. After that, they produced text data as audio using the Microsoft Speech Software Development Kit.

In [8], researchers proposed a technique known as "VOCAL VISION." Its operation is based on the video to audio conversion. The webcam detects the image, which is then processed with an item identification algorithm, enhanced, and compared to database photographs. If the comparison is successful, the user or blind person will receive a message with the name of the object and the name of the scene via headphones and an Android device. This system's various components included blurring, grey scalability, cropping, RGB to HSV, and histograms. Following a successful comparison of the object or scene with database images, the blind person is told about the object or scene.

Software and Hardware requirements:
Hardware requirements:

- Raspberry PI
- Power Supply
- LCD
- APR Voice Module
- Speaker
- Camera
- SD CARD
- WI-FI

Raspberry Pi Components:
Raspberry Pi can best be summed up in one sentence: "It is a small-sized computer at a low cost." It's a gadget that uses a regular mouse or keyboard and may be used with a TV or PC. It is able to carry out all the computation exercises. A UK-based non-profit organization

ISSN PRINT 2319 1775 Online 2320 7876

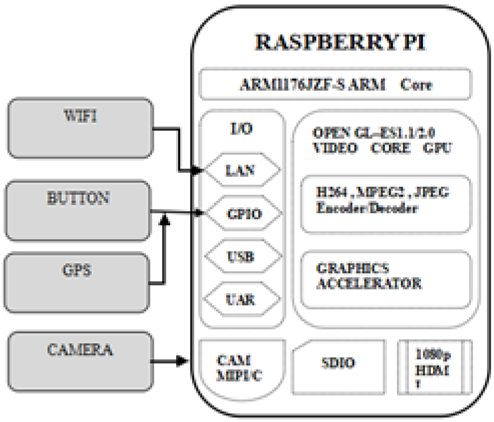
Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

called the Raspberry Pi Foundation seeks to empower individuals all over the world with access to computing and digital creating tools.

They do this to enable more people to use computing and digital technologies for work, find solutions to issues that are important to them, and engage in creative expression. I can utilize programming languages like Python and Scratch with Raspberry Pi.

Everyone of any age can use this tool, and it is easy to use. It is a device that is capable of carrying out all of a computer's functions, including watching high-definition videos, making spreadsheets, and using the internet. It is a device that can interact with the outside environment and offers a good variety of digital maker projects and many more.

It serves as the project's client. Here Footage is live-streamed to the server end using this and the Raspberry Pi Noir camera. The connection is wireless and the only prerequisite is that both the server and the client are present on the same network for improved communication between them.



Software Requirements:

- Raspberry PI IDE
- Python
- OpenCV

Proposed Work:

The development of technology has made it possible for us to solve numerous issues. Even humans have advanced to the point where a team recently succeeded in clicking an image of a black hole thanks to some amazing computational work. But as of right now, it's clear that blind individuals have a lot of difficulties carrying out their daily tasks. They rely on other people to perform all or many of their tasks. That inspired us to create a prototype. This is a modest effort to help blind individuals become independent. An autonomous navigation system for blind individuals is being developed as part of this research. It's an effort in which I won't donate the eyeballs but will instead give them to these unique folks nearby. The project's primary goal is to help blind individual become familiar with their surroundings. This would help the individual manage his or her workload and inform them of what is coming next. They could recognize fundamental and frequent items in their daily lives with the aid of technology. The use of object detection algorithms can create new opportunities for helping with both interior and outdoor navigation for differently abled people.

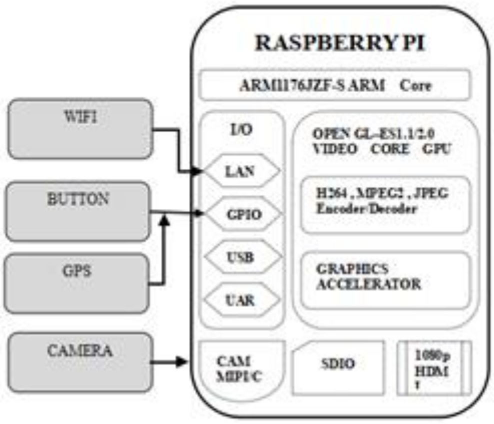
The system will be composed of several parts. The video is recorded on a portable raspberry pi and sent to the server via a raspberry pi camera device. The server features an object

ISSN PRINT 2319 1775 Online 2320 7876

Research paper © 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

detection model for real-time picture recognition using OpenCV models. To estimate the 3D location of the items, the position and size of the bounding boxes from the detection technique are employed. A 3D sound generation program built on the Unity game engine then renders the binaural sound with places encoded. The user hears the sound through wired or wireless earbuds. The sound is emitted when the detected object differs from the preceding one or every few seconds, whichever occurs first. With the use of the device, the user will be able to successfully detect objects from 3-5 meters away. Potential issues for the current prototype include information overload when the system tries to tell users of numerous objects and failure of detection when items are too close or far away.

Block Diagram:



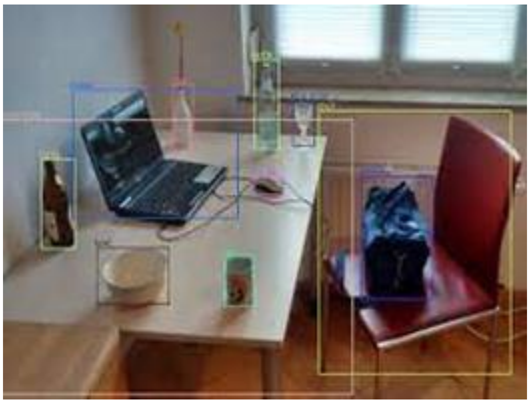
Working:
Open CV:
Computer vision is a method that enables us to comprehend how images and videos are stored, how to change them, and how to extract data from them. The foundation or primary tool utilized in artificial intelligence is computer vision. Robotics, photo-editing software, and self-driving cars all significantly rely on computer vision.
Importance of the real-time operation is, it is essential in systems which are contemporary, and is currently played by OpenCV, a big open-source library for computer vision and ML. In order to identify items like people, objects, and even human handwriting, it may be used to analyze images and movies. Python uses numerous libraries for handling OpenCV array structure for analysis. For recognizing a visual pattern, we employ vector space and perform mathematical operations on its features.

Applications of OpenCV:
There are many applications that OpenCV can be used to solve, some of them are described

- Recognizing faces
- Automated surveillance and inspection
- Count of individuals (foot traffic in a mall, etc)
- Retrieval of videos and images
- Robotic navigation and object identification in driverless vehicles
- 3D motion structure from motion TV channel ad recognition

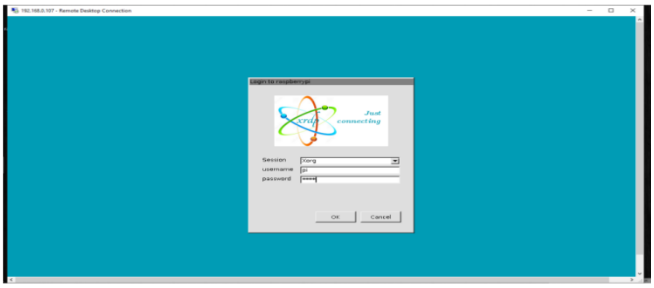
ISSN PRINT 2319 1775 Online 2320 7876

Research paper **© 2012 IJFANS. All Rights Reserved, UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022**

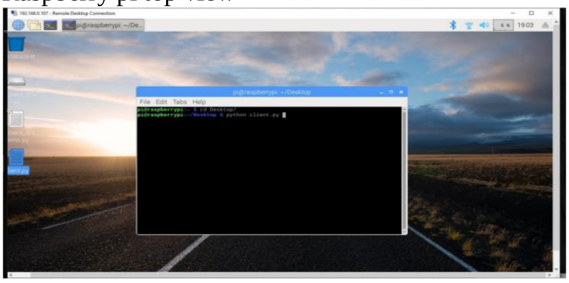


Results and Discussions

Raspberry pi top view



Remote desktop connection login panel



Raspberry pi terminal (client)

Depiction of the whole system

Conclusion:

To best assist, the blind person, an object recognition code is built and combined with unity code for a 3D sound engine. The system necessitates extensive network data transfers. In case of flawless data transport, the system offers the best outcomes. Many items that we come across in our daily lives are named and explained to the person.

In accordance with the concept described in the paper, we have developed a model for object detection. The solution is capable of performing precise real-time objective detection by utilizing the Open CV and transmission of 3d audio. This project offers a real-time, portable solution.

References:

1 Let Blind People See: Real-Time Visual Recognition with Results Converted to 3D Audio Rui (Forest) Jiang Earth Science, Stanford forestj@stanford.edu Qian Lin Applied Physics, Stanford linqian@stanford.edu Shuhui Qu Civil and Environmental Engineering.

ISSN PRINT 2319 1775 Online 2320 7876

 UGC CARE Listed (Group -I) Journal Volume 11, Iss 12, Dec 2022

2 Object detection and identification for a blind in video scene – Hanen Jabnoun, Hamid Amiri.

[3] Use of 4G waveform towards RADAR Nigidita pradhan1, Rabindranath Bera2, Debasish Bhaskar3 PG Student [DEAC], Dept. of ECE, Sikkim Manipal Institute of Technology, India1 Head of the Dept [HOD], Dept. of ECE, Sikkim Manipal Institute of Technology, India2 Assistant Professor, Dept. of ECE, Sikkim Manipal Institute of Technology, India3 Majitar, Rangpo, East Sikkim, India

[4] Motion-Based Object Detection and Tracking in Color Image Sequences Bernd Heiseley Image Understanding Group (FT3/AB) DaimlerChrysler Research Center Ulm, D-89013, Germany.

[5] Real-time multiple vehicle detection and tracking from a moving vehicle Margrit Betke1, Esin Haritaoglu2, Larry S. Davis2 1 Boston College, Computer Science Department, Chestnut Hill, MA 0216betke@cs.bu.edu 2 University of Maryland, College Park, MD 20742, USA.

[6] Finding Objects for Blind People Based on SURF Features Ricardo Chincha and Yingli Tian Department of Electrical Engineering the City College of New York, NY, 1003.

[7] T.Srinivasarao, G.Siva Nageswarao, Efficient scheduling Measures for improving real-time systems using prediction process, Journal of Computational and Theoretical Nanoscience,2019, Vol16, Issue 5, Pages 1876-1880

[8] Shaik Salma Begum & Dr. D. Rajya Lakshmi," GLCM of Fuzzy Clustering Means for Textural Future Extraction of Brain Tumor in Probabilistic Neural Networks", International Journal of Innovative Technology and Exploring Engineering, ISSN: 2278-3075, Volume-9, Issue-1, November 2019.