Group no.37

A

PROJECT REPORT ON

RTVB FOR BLIND PEOPLE

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE B.E. PROJECT

BACHELOR OF ENGINEERING

IN

ELECTRONICS AND TELECOMMUNICATION ENGINEERING

By

Sourabh N. Mohite Exam.No: B150433079

Omkar D. Ibitkar Exam. No:B150433047

Mahesh M. Pharande Exam.No:B150433099

Shubham K. Pande Exam.No:B150433088

Under the Guidance of

Mr.S.B.Pawar

Sinhgad Institutes

Submitted to

DEPARTMENT OF ELECTRONICS AND TELECOMMUNICATION ENGINEERING

STES'S SINHGAD ACADEMY OF ENGINEERING, PUNE-411048

2019-2020

CERTIFICATE

This is to certify that the project report entitled

REAL TIME VISUALIZATION FOR BLIND PEOPLE

Submitted By

Students Name	ExamNo.
1.Sourabh N.Mohite	B150433079
2.Omkar D. Ibitkar	B150433047
3.Mahesh M.Pharande	B150433099
4.Shubham K. Pande	B150433088

Is a bonafide work carried out by the under the supervision by Prof. S.B.Pawar and it is approved or the partial fulfillment of the requirement of Savitribai Phule Pune University for the BE of Electronics and Telecommunication Engineering.

Prof.S.B.Pawar	Dr.K.M.Gaikwad	Dr.K.P.Patil
Project Guide	H.O.D & Vice-Principal	Principal,SAOE,Pune

Prof.S.S.Shah

Project Co-ordinator

Place : Pune

Date:

External Examiner

Table of Contents

List of Figures

ACKNOWLEDGEMENT:

It gives me an immense pleasure to submit this report on **"REAL TIME VISUALIZATION BOX(RTVB) FOR BLIND PEOPLE".**I have tried my level best to represent this topic into compact and to the point framework.

I truly express my deepest thanks to all the people who have given their precious time and encouragement and my guide **Prof.S.B.Pawar** & co-guide **Prof.S.V.Thakur** for his patience he showed me during the process of preparation of the report from the initial conception to final presentation and all staff of E&TC department. I would like to thanks from the core of my heart to our H.O.D of E&TC department **Dr.K.M.Gaikwad** and Principal **Dr.K.P.Patil** for being constant pillar of support, during the preparatory stages.

Finally yet importantly, I would like to thank my family and friends for their never-ending support and love.

Project Group Members:

Sourabh N. Mohite
Omkar D. Ibitkar
Mahesh M. Pharande
Shubham K . Pande

ABSTRACT:

Object detection is the main objective of this project besides surveillance. The aim of the project is to detection of object to guide blind people. This project tries to transform the visual world into the audio world with the potential to inform blind people objects as well as their spatial locations. Objects detected from the scene are represented by their names and converted to speech.

Our system composes of several modules. Video is captured with a portable camera device on the client side, and is streamed to the server for real-time image recognition with existing object detection models (YOLO). The 3D location of the objects is estimated from the location and the size of the bounding boxes from the detection algorithm. Then, a 3D sound generation application based on Unity game engineer enders the binaural sound with locations encoded. The sound is transmitted to the user with wireless earphones. Sound is play at an interval of few seconds, or when the recognized object differs from previous one, whichever earliest.

The majority of visualizations on the web are still stored as raster images, making them inaccessible to visually impaired users. We propose a deep-neural-network-based approach that automatically recognizes key elements in a visualization, including a visualization type, graphical elements, labels, legends, and most importantly, the original data conveyed in the visualization. We leverage such extracted information to provide the reading of the extracted information to visually impaired people. Based on interviews with visually impaired users, We compared the performance of the back-end algorithm with existing methods and evaluated the utility using qualitative feedback from visually impaired users.

CHAPTER NO:-1
INTRODUCTION

1.1. Motivation:

In searching our project topic we struck an idea about problem faced by blind people. So we are interested in searching how blind people surviving in day to day life. Next we visited to the Organization of Blind People ,while interacting with blind people we notice problem faced by them. So we decided to reduce their efforts in daily life to survive them as a normal people. For that purpose we decided to make a system which help them for their survive.

There is always a motivating factor behind every work we do. Just as there is a guiding line behind every step taken, We had certain motivating factors behind the selection of this project. Since we will be studying subjects like incorporate our knowledge gained by these subjects into our project. As we have already studied microcontroller and its programming in the previous semester, we were more comfortable using the same. The main factor behind motivation is concept of Digital & Smart India. So, we wanted to contribute in our idea.

The motivation for a object detection using image based approach is due to rapid growth of the electronics. The previous object detection system for blind people is based on sensors to make decision. Most of the available sensors used such as ultrasonic sensor and vibration sensor ,take time to response as well as they cannot analyze or visualize the object. So, we can replace sensors with the help of Camera in video mode, due to this mode we can visualize object and name of object is heard for blind people.

1.2. Problem Statement:

To prepare a system that can be detect object for blind people using Raspberry pi.

1.3. Objective :

1) Learning of Python and basic algorithm to detect object such as pen,bottle etc.
2) Integration of camera module with Raspberry Pi to detect object in real time.
3) To write algorithm to convert image detected by Raspberry Pi to audio format.
4) Real time testing and integration of RTVB.

CHAPTER NO:-2
LITERATURE REVIEW

LITERATURE REVIEW:

India is the home of the largest number of blind people in the world . Out of the 37 million blind people all over the globe more than 15 million people are in India. For supporting to avoid obstruction visually impaired people commonly use sticks.

At present numerous devices are available for offering guidance to a distant location but these are either costly or make use of braille interface. Blind people can get data from the unwilling interaction with entities, persons or animals, by exploring the environment and using their hands to recognize the shape of an object, moreover, blind people can sense other features of the objects as temperature, texture, weigh and though the tact has certain limits in confront of sight, it has a very significant function to unveil to blind persons the world around them.

There are many systems which are designed to help navigate the visually impair literature review was carried out throughout whole project to gain the knowledge and improve skill needed to complete the project. The main sources for this project are pervious related project ,reach ,books, journals and online tutorial the chapter focus on basic concept and all fundamental theories which is related to project.

Based on the impetus of the CNN, they develop a blind visualization system that helps blind people better explore the surrounding environment. By using YOLO algorithm and advanced wireless transmitter, the solution could perform accurate real time objective detection. A prototype for sensory substitution (vision to hearing) is established in the project.[1]

A portable and real time solution is provided in the work. They present a platform that utilizes portable cameras, fast HD video link and powerful server to generate 3D sounds[2].

This paper [3] proposes a system in which two cameras are put on the

glasses of a blind person. The proposed work has a wearable device and consists of a blind stick and sensor based detection circuit. It uses an infrared sensor which uses infrared waves to scan the surroundings of a person. It uses object detection and gives them audio information about it. [3] The system must be trained about object information. Feature extraction is also a part of the process. [3]

Another system [4] proposed in this paper focuses on giving the information about what are the different types of obstacles in front of the user, their size and their distance from the user. [4] MATLAB Software is used for signal processing. The camcorder is used for recording videos. Video processing methods are used after that. [4] The output of this system not only gives output in audio format but also vibration. A vibrating motor has been connected with an ultrasonic sensor. The ultrasonic sensor detects objects coming in its range and this makes the vibrating motor vibrate. [4]

This system [5] tries to detect multiple objects in an image. That is the core specialty of the system. It is a system where N object detectors are trained for N different objects. [5] When an image is sent to the system, all object detectors do their work. If an object is found by a detector, it will mark its boundary and label the object name. After the process completes for all N detectors, the image is displayed with all the tags. Moving a cursor over an object in the image shows the complete boundary of the object with its label beside. [5] This system is a little slower than other systems because a lot of object detectors are working on a single image. The performance can increase by allowing more than one object detectors to run in parallel. [5]

Yet there is another system [6] which first subtracts the current frame from the previous one and obtains a maximum value of the difference between two pixel values. Maximum value > given pixel = Foreground. Maximum value < given pixel = Background. [6] The brightness distortion and Chromaticity distortion are also taken care of in this project by using shadow detection technique theory. [6]

By using morphological algorithm, the solution could perform accurate real time objective detection with 1080Presolution. A prototype for sensory substitution (vision to hearing) is established in the work. Through this work, they hope to demonstrate the possibility of using computer vision techniques as a type of assistive. This will definitely navigate the blind people to detect the obstacles in front of them. Even there are plenty of techniques are available to guide them but this work efficiently help the blind people to guide and navigate them.

We consider above reference and decided to make a system, which transfer visual world into audio world for blind people.

CHAPTER NO:-3
DESIGN AND IMPLEMENTATION

3.1 Introduction:

Millions of people live in this world with incapacities of understanding the environment due to visual impairment. Although they can develop alternative approaches to deal with daily routines, they also suffers from certain navigation difficulties as well as social awkwardness. For example, it is very difficult for them to find a particular room in an unfamiliar environment , blind and visually impaired people find it difficult to know whether a person is talking to them or someone else during a conversation. Blind people face several problems in their life, one of these problems that is the most important one is detection the obstacles when they are walking.

In this research, we suggested a system with cameras placed on blind person's glasses that their duty is taking images from different sides. By comparing these images, we will be able to find the obstacles. In this method, first we investigate the probability of existence an object by use of special points that then we will call them "Equivalent points", then we utilize binary method, standardize and normalized cross-correlation for verifying this probability. Blindness is a state of lacking the visual perception due to neurological or physiological factors. The partial blindness represents the lack of integration in the growth of the optic visual or nerve centre of the eye, and total blindness is the full absence of the visual light perception.

In this work, cheap, a simple friendly user, smart blind guidance system is designed and implemented to improve the mobility of both blind and visually impaired people in a specific area. The proposed work includes a wearable equipment consists of light weight blind stick and sensor based obstacle detection circuit is developed to help the blind person to navigate alone safely and to avoid any obstacles that may been countered, whether fixed or mobile, to prevent any

possible accident. The main component of this system is the infrared sensor which is used to scan a predetermined area around blind by emitting-reflecting waves.

Computer vision technologies, especially the deep convolution neural network, have been rapidly developed in recent years. It is promising to use the state-of-art computer vision techniques to help people with vision loss. In this project, we want to explore the possibility of using the hearing sense to understand visual objects. The sense of sight and hearing sense share a striking similarity both visual object and audio sound can be spatially localized. It is not often realized by many people that we are capable at identifying the spatial location of a sound source just by hearing it with two ears.

In our project, we build a real-time object detection and position estimation pipeline, with the goal of informing the user about surrounding object and their spatial position using binaural sound. Discuss the relate works on sensory substitution, assistive products using computer vision for blind people, and the exploration of 3D sound introduces different components of our prototype.

3.2 Block Diagram :

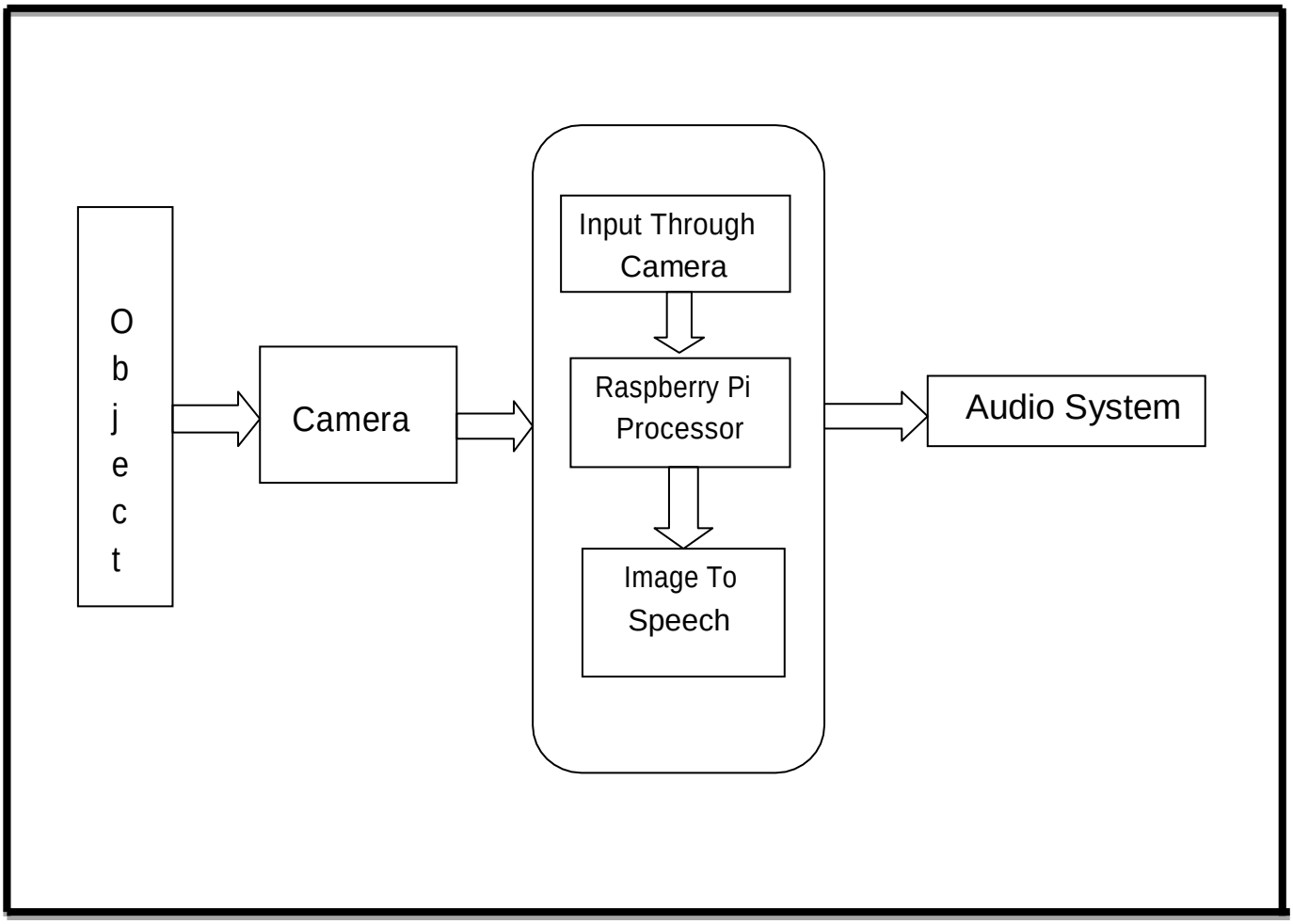


Fig.3.2.1

3.3 Working:

Here we will develop a system RTVB for blind people ,that can help them to guide.Our system works on principle convolutional neural network.

In this we use a processor i.e. Raspberry pi and a camera module.First we capture an images using camera module.Then we extract the features of an image and accordingly we develop a module. We used microprocessor for extracting the feature of an image.Here, we have used an YOLO V3 like algorithm for object detection. Here, we are not interested in whole object features, We require only boundary details of an object.

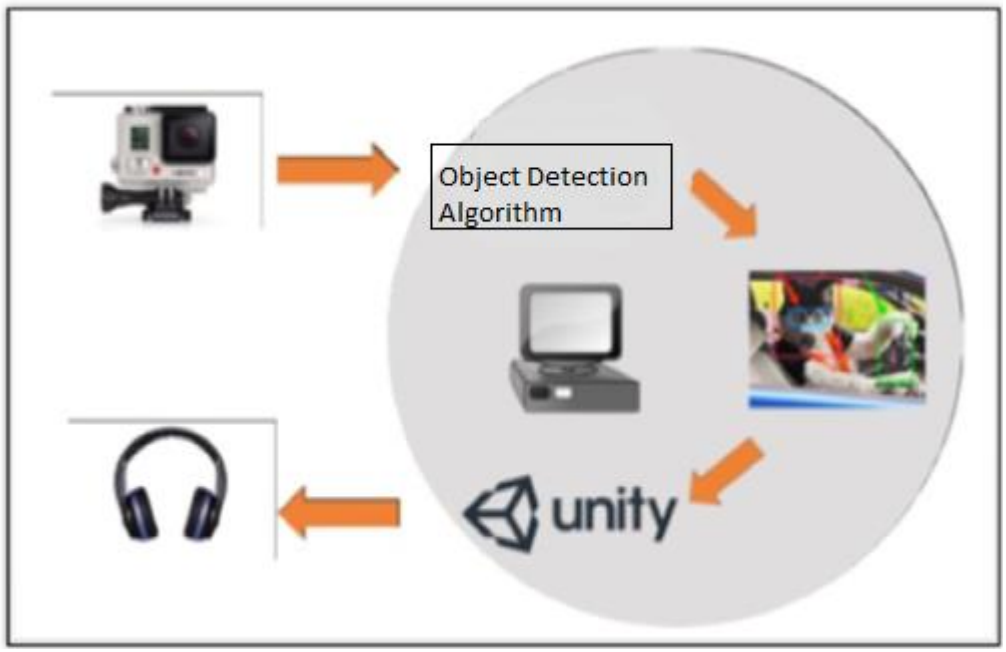


Fig.3.3.1 Object Detection

For the detection of object in real time, we use Raspberry Pi Camera Module V2 which has 8 megapixel native resolution which is connected to the Raspberry Pi board via a short ribbon cable.

By scanning various images we treat module for object detection. For processing the image we use processor name Raspberry Pi 3 Model B which has memory of 1GB RAM. We use external SD card for booting purpose of Raspberry Pi Operating System name Raspbian .We here detect an image with the help of

camera and then process the image boundaries and detect object accordingly.

For processing the image as per our requirement we use YOLO V3 algorithm,You Only Look Onceas it says, it work on captured image one time. Captured image is processed using this algorithm used to processes the boundaries of image as it use Neural Network it precises the image and convert into required form.

Then we have to transfer 3-D world into audio world. The Audio System program is used to convert detected image into audio,then object detected by system speech out by it's name. The 3.5mm audio jack (headphone jack) present in Raspberry Pi module is used for sound output of image.

With the help of audio it make a sense for blind people to get appropriate object detection and make decision according to it.

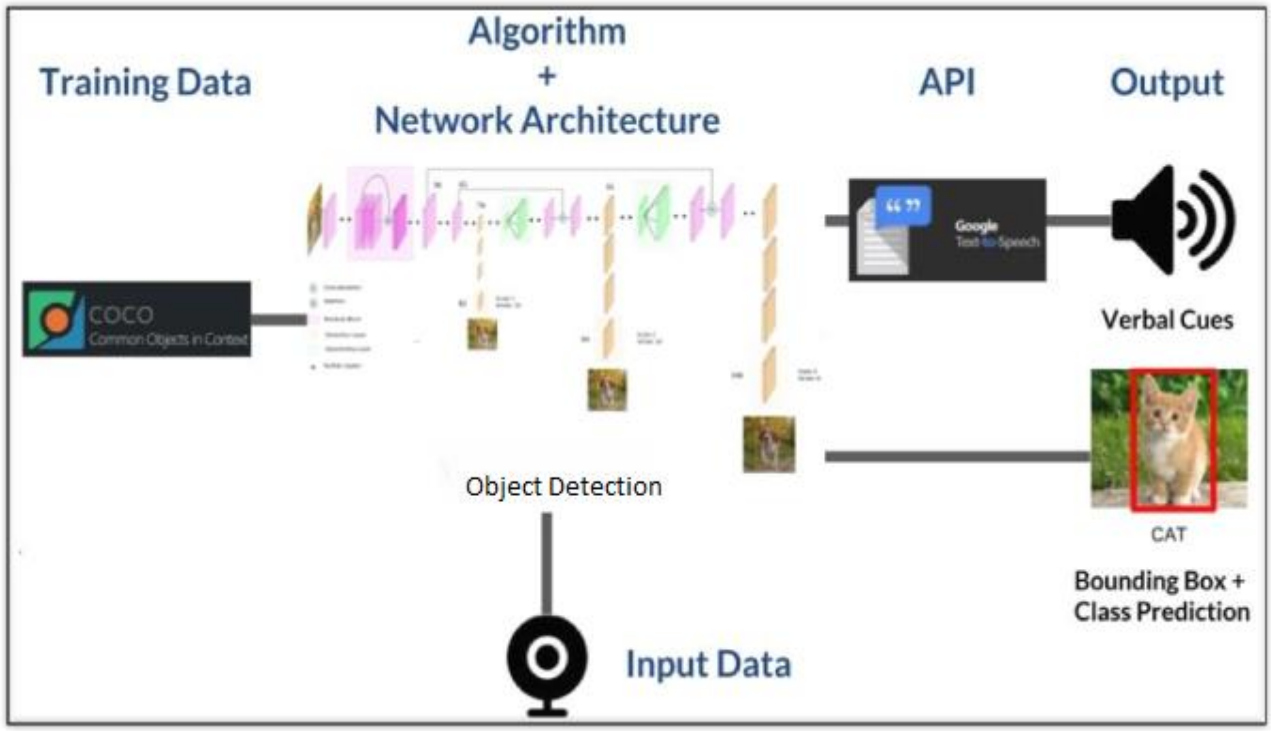


Fig.3.3.2 Object Detection With Audio

3.4 Raspberry Pi:

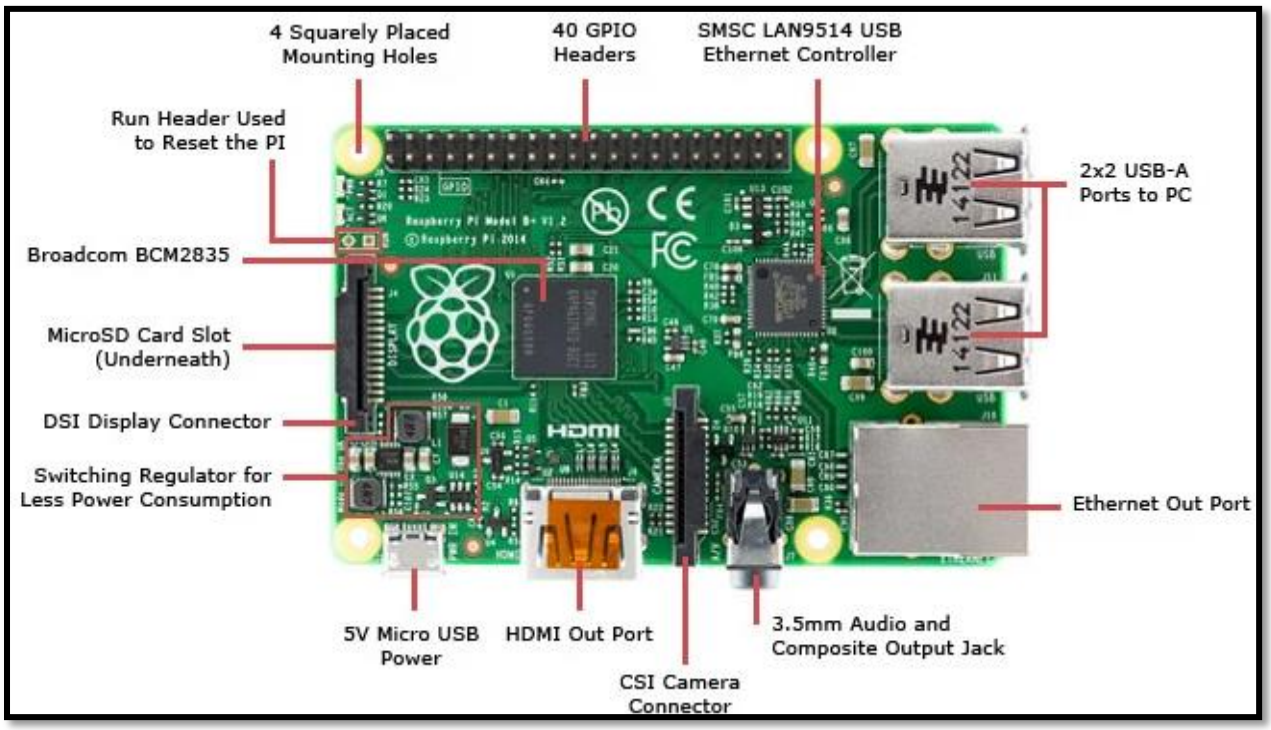


Fig.3.4.1 Raspberry Pi Module

Features:

- CPU: Quad-core 64-bit ARM Cortex A53 clocked at 1.2GHz
- GPU: 400MHz Video Core IV multimedia
- Memory: 1GB LPDDR2-900 SDRAM (i.e.900MHz)
- USB ports:4
- Video outputs: HDMI, composite video (PAL and NTSC) via 3.5 mmjack
- Network: 10/100Mbps Ethernet and 802.11n WirelessLAN
- Peripherals: 17 GPIO plus specific functions, and HAT IDbus
- Bluetooth:4.1
- Power source: 5 V via MicroUSB or GPIOheader
- Size: 85.60mm ×56.5mm

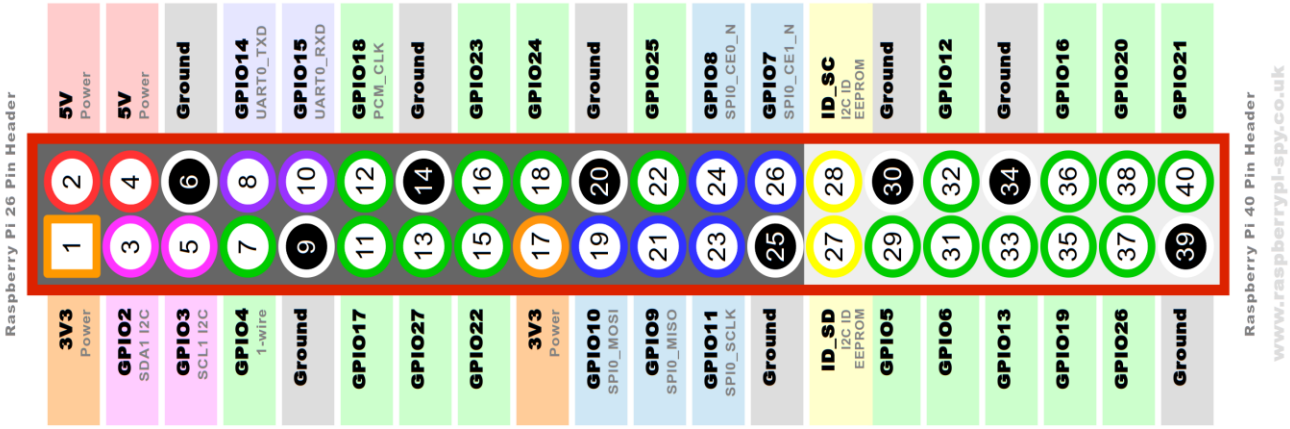


Fig: 3.4.2 Pin Diagram

Memory:-

The raspberry pi model A board is designed with 256MB of SDRAM and model B is designed with 51MB.Raspberry pi is a small size PC compare with other PCs. The Normal PCs RAM memory is available in gigabytes. But in raspberry pi board, the RAM memory is available more than 256MB or 512MB

CPU (Central Processing Unit):-

The Central processing unit is the brain of the raspberry pi board and that is responsible for carrying out the instructions of the computer through logical and mathematical operations. The raspberry pi uses ARM11 series processor, which has joined the ranks of the Samsung galaxy phone.

GPU (Graphics Processing Unit):-

The GPU is a specialized chip in the raspberry pi board and that is designed to speed up the operation of image calculations. This board designed with a Broadcom video core IV and it supports OpenGL

Ethernet Port:-

The Ethernet port of the raspberry pi is the main gateway for communicating with additional devices. The raspberry pi Ethernet port is used to plug your home router to access the internet.

GPIO Pins:-

The general purpose input & output pins are used in the raspberry pi to associate with the other electronic boards. These pins can accept input & output commands based on programming raspberry pi. The raspberry pi affords digital GPIO pins.These pins are used to connect other electronic components. For example, you can connect it to the temperature sensor to transmit digital data.

XBee Socket:-

The XBee socket is used in raspberry pi board for the wireless communication purpose.

Power Source Connector:-

The power source cable is a small switch, which is placed on side of the shield. The main purpose of the power source connector is to enable an external power source

UART:-

The Universal Asynchronous Receiver/ Transmitter is a serial input & output port. That can be used to transfer the serial data in the form of text and it is useful for converting the debugging code.

Display:-

The connection options of the raspberry pi board are two types such as HDMI and Composite. Many LCD and HD TV monitors can be attached using an HDMI male cable and with a low-cost adaptor. The versions of HDMI are 1.3 and 1.4 are supported and 1.4 version cable is recommended. The O/Ps of the Raspberry Pi audio and video through HMDI, but does not support HDMI I/p. Older TVs can be connected using composite video. When using a composite video connection, audio is available from the 3.5mm jack socket and can be sent to your TV. To send audio to your TV, you need a cable which adjusts from 3.5mm to double RCA connectors.

3.5 Camera:

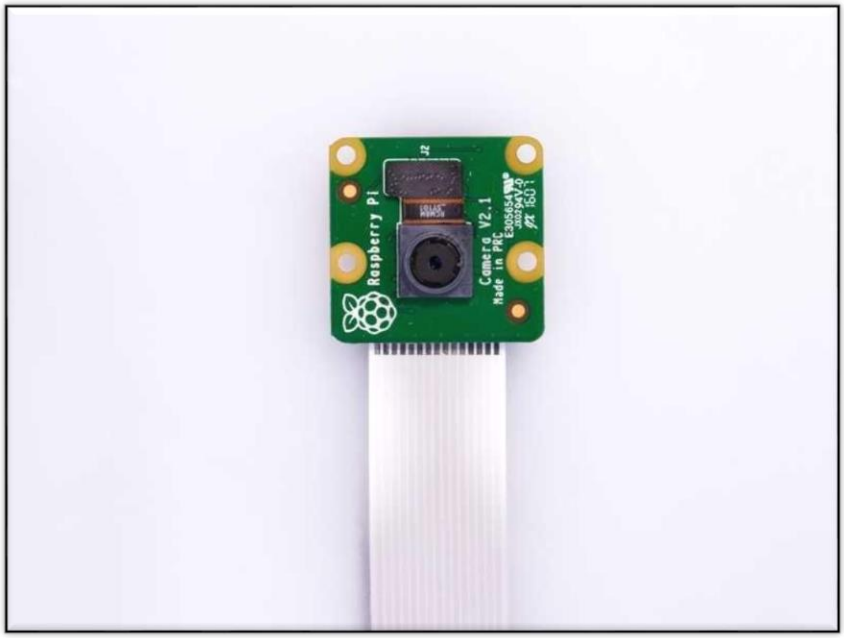


Fig.3.5.1 Camera Module

Features :

- Fixed focus lenson-board
- Improved resolution - 8 megapixel native resolution sensor-capable of 3280 x 2464 pixel staticimages
- Supports 1080p30, 720p60 and 640x480p90video
- Size 25mm x 23mm x9mm
- Weight just over3g
- Connects to the Raspberry Pi board via a short ribbon cable(supplied)
- Camera v2 is supported in the latest version of Raspbian , Raspberry Pi's preferred operating system
- Uses the Sony IMX219PQ image sensor - high-speed video imaging and high sensitivity
- 1.4m X 1.4m pixel with Omni BSI technology for high performance (high sensitivity, low crosstalk)

3.6 Object Detection Algorithm:

You only look once, or YOLO, is one of the faster object detection algorithms. Though it is no longer the most accurate object detection algorithm, it is a very good choice when you need real-time detection, without loss of too much accuracy.

"YouOnlyLookOnce" is an algorithm that uses convolutional neural networks for object detection. You only look once, is one of the faster object detection algorithms out there. It is a very good choice when we need real-time detection, without loss of too much accuracy.In comparison to recognition algorithms, a detection algorithm does not only predict class labels but detects locations of objects as well. So, It not only classifies the image into a category, but it can also detect multiple Objects within an Image. This Algorithm applies a single Neural network to the Full Image. It means that this network divides the image into regions and predicts bounding boxes and probabilities for each region. These bounding boxes are weighted by the predicted probabilities.

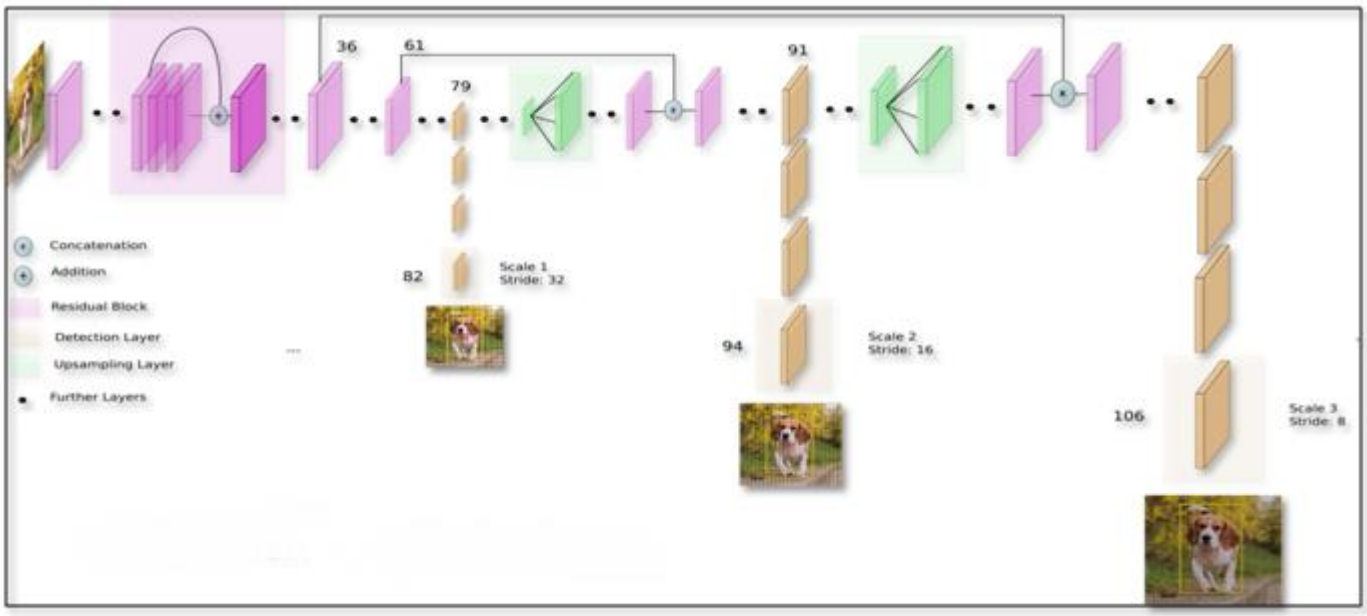


Fig.3.6 Object Detection Algorithm

Algorithm makes use of only convolutional layers, making it a fully convolutional network (FCN). In paper, the authors present new, deeper architecture of feature extractor called Darknet-53. As it's name suggests, it contains of 53 convolutional layers, each followed by batch normalization layer and Leaky ReLU activation. No form of pooling is used, and a convolutional layer with stride 2 is used to down sample the feature maps. This helps in preventing loss of low-level features often attributed to pooling.Algorithm is invariant to the size of the input image.

However, in practice, we might want to stick to a constant input size due to various problems that only show their heads when we are implementing the algorithm.

A big one among these problems is that if we want to process our images in batches (images in batches can be processed in parallel by the GPU, leading to speed boosts), we need to have all images of fixed height and width. This is needed to concatenate multiple images into a large batch.The network down samples the image by a factor called the stride of the network. For example, if the stride of the network is 32, then an input image of size 416 x 416 will yield an output of size 13 x13.Generally, stride of any layer in the network is equal to the factor by which the output of the layer is smaller than the input image to the network.

Algorithm makes prediction at three scales, which are precisely given by down sampling the dimensions of the input image by 32, 16 and 8 respectively. Object Detection performs multilevel classification for objects detected in images.

Feature of Object Detection Algorithm:

- The biggest advantage of using YOLO is its superb speed it's incredibly fast and can process 45 frames per second.
- This also understands generalized object representation.
- This is one of the best algorithms for object detection and has shown a comparatively similar performance to the R-CNN algorithms.
- Fast, Good for real-time processing.
- Predictions (object locations and classes) are made from one single image,
- Can be trained end-to-end to improve accuracy.
- Algorithm is more generalized.
- Most silent feature is it detect at three different scales.

Chapter No:-4
RESULT

1. Capture Image using pi camera:

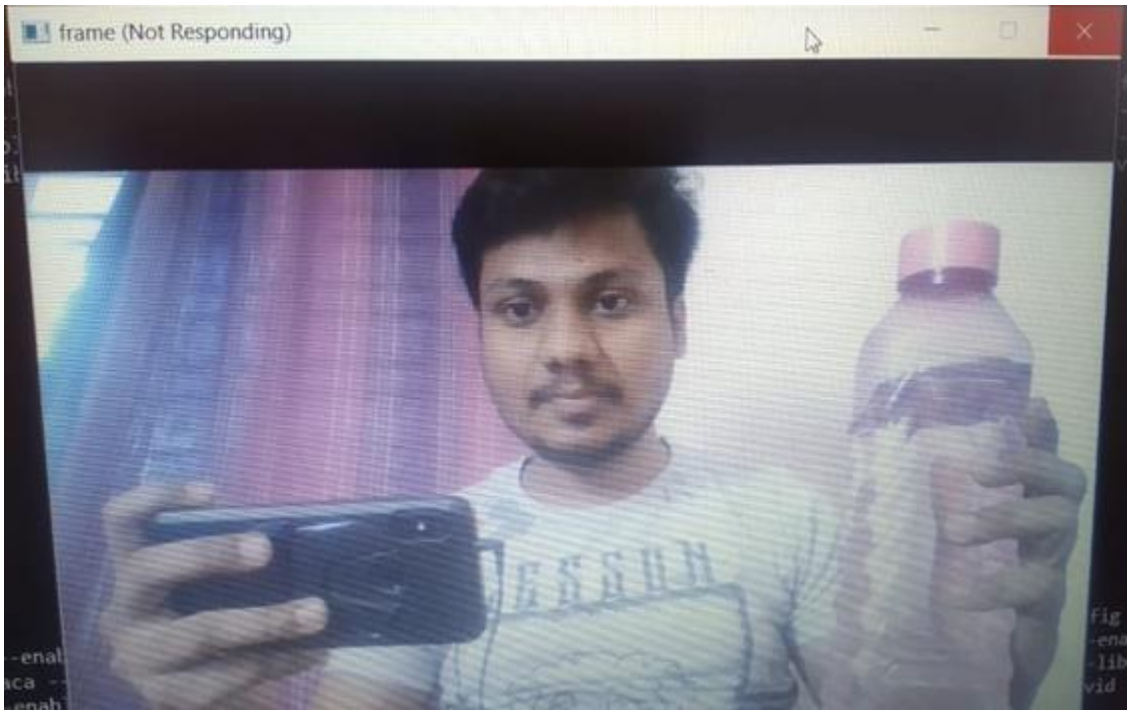


Fig 4.1 Captured Image

2. Feature Extraction and Boundary box:

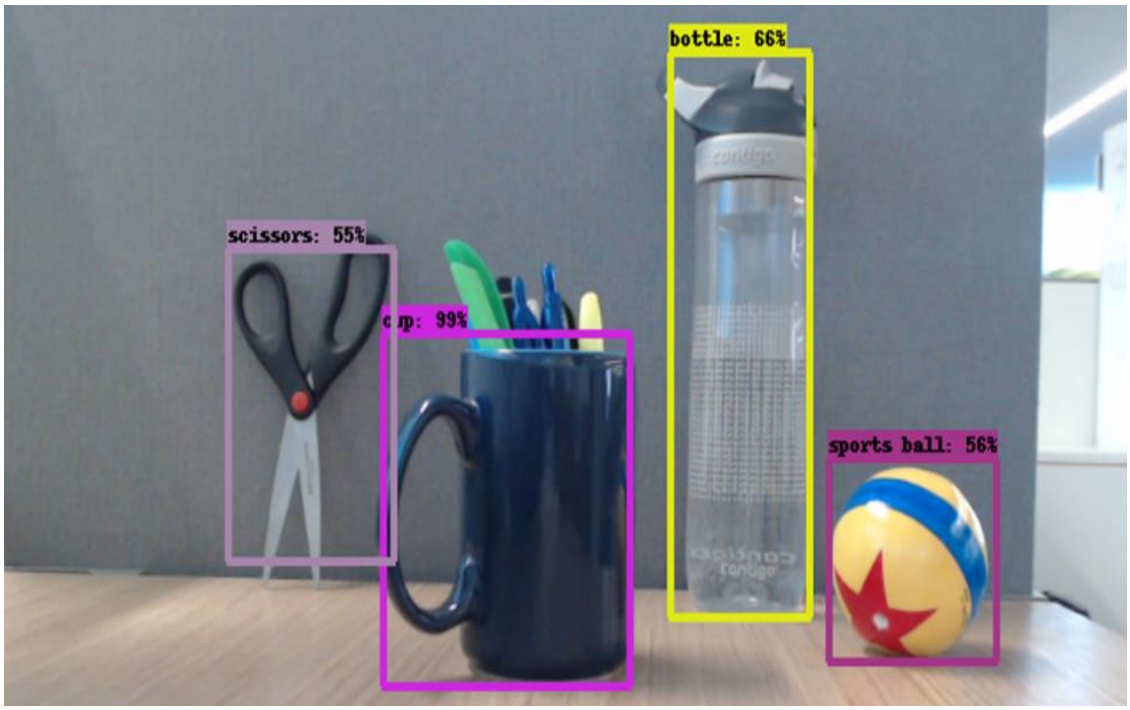
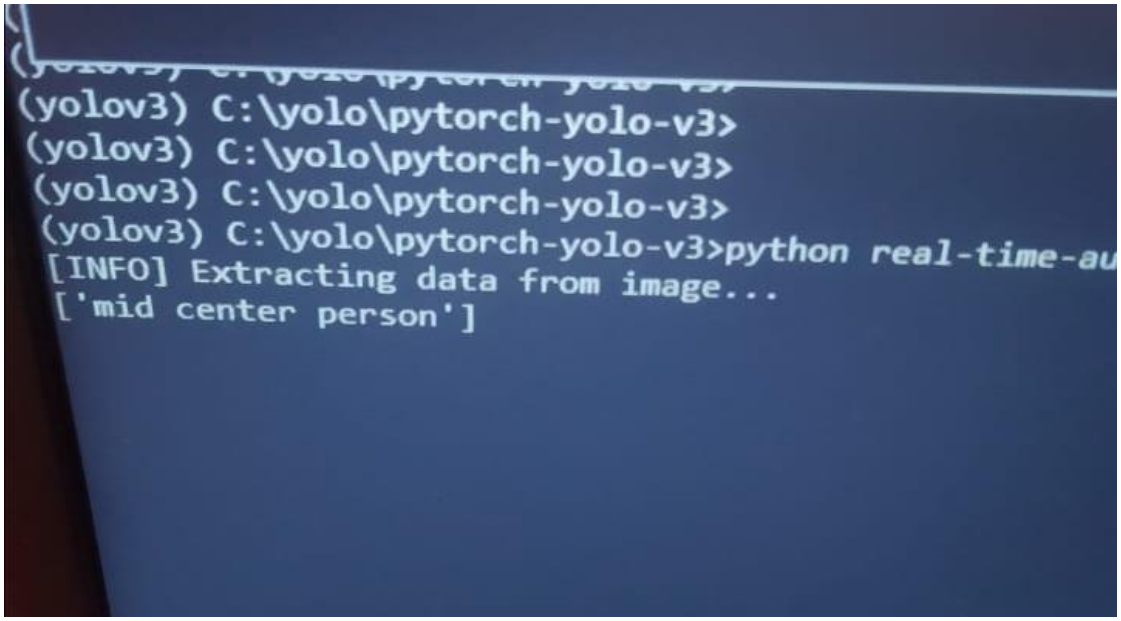


Fig 4.2 Feature Extraction



3. Conversion from Text To Speech:

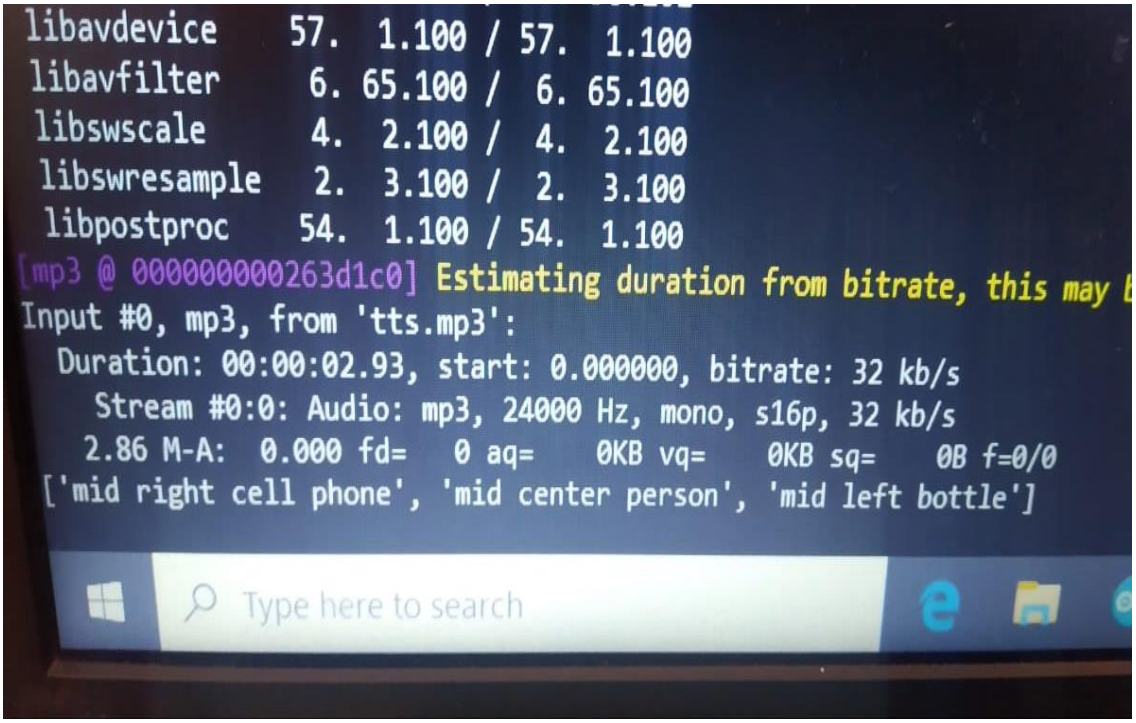


Fig 4.3 Text To Speech

CHAPTER NO:-5
SPECIFICATION

5.1 Advantages of system:

- It helps to realizes real time environment for the blind people.
- To avoid accidents of blind peoples.
- It reduces dependency of the blind people.
- Predictions (object locations and classes) are made from one single networks can
- be trained end to end to improve accuracy.
- It can find Objects in image grid at parallel their fore object detection is much faster than other systems.
- The network understand generalized object representation and convert into audio
- at real time.

5.2 Disadvantage of System:

- Only front view detection.
- As number of frames per second increases detection is difficult.
- Detection may fail if number of objects increase.

5.3 Conclusion:

In this project, we investigate the need from blind and visually impaired people. We develop a blind visualization system that helps blind people better explore the surrounding environment. A portable and real time solution is provided in the project. We present a platform that utilizes portable cameras, fast HD video link and powerful server to generate 3D sounds. By using Object Detection algorithm the solution could perform accurate real time object detection with live stream at a speed of 30 frames, 1080Presolution. A prototype for sensory substitution (vision to hearing) is established in the project.

Future Scope:

Future real time object detection based on GPS system that guide a blind people in day to day life.GPS system guide blind using google map technique. Other online services are added to it to guide blind.

CHAPTER NO:-6
REFERENCES

References:

[1] Rui (Forest) JiangEarth Science, Stanford

[2] N.Saranya, M.Nandinipriya, U.Priya Assistant Professor, Department of Electronic sand Communication Engineering, Bannari Amman Institute of Technology, Sathyamangalam Erode.(India)

[3] Prof. Seema Udgirkar, Shivaji Sarokar, Sujit Gore, Dinesh Kakuste, Suraj Chaskar, 2006, "Object Detection System for Blind People," International Journal of Innovative Research in Computer and Communication Engineering

[4] Daniyal Rajput, Faheem Ahmed, Habib Ahmed, Engr Zakir Ahmed Shaikh, Aamir Shamshad, 2014, "Smart Obstacle Detector for Blind Person," Journal of Biomedical Engineering and Medical Imaging.

[5] Khushboo Khurana, Reetu Awasthi, 2013, "Techniques for Object Recognition in Images and Multi-Object Detection," International Journal of Advanced Research in Computer Engineering & Technology.

[6] Payal Panchal, Gaurav Prajapati, Savan Patel, Hinal Shah and Jitendra Nasriwala, 2008, "A Review on Object Detection and Tracking Methods," International Journal For Research In Emerging Science And Technology

[7] ShaoqingRen, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn:Towards real-time object detection with region proposal networks. In Advances in Neural Information Processing Systems, pages 91–99, 2015.

[8] Ross Girshick. Fast.In Proceedings of the IEEE International Conference on Computer Vision, pages 1440–1448, 2015.

[9] B.Yu, L.Xu, and Y.Li, " Bluetooth Low Energy (BLE) Base Mobile Electrocardiogram MonitoringSystem," in Proc. of IEEE Int. Conf. on information and automation .pp,763767,june,2014.

[10] Jizhong Xiao, Kevin Ramdath, Manor Iosilevish, Dharmdeo Sigh,and

Anastasis T sakas. low cost outdoor assistive navigation system for blind people. In Industrial Electronics and Applications (ICIEA),2013 8th IEEE Conference on, pages 828–833.IEEE, 2013.

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

[12] David Brown, Tom Macpherson, and Jamie Ward,Seeing with sound exploring different characteristics of a visual-to-auditory sensory substitution device.Perception, 40(9):11201135, 2011.

[13] W.J.Li,Y. L.Luo, Y. S. Chang, and Y. H. Lin, "A wireless blood pressure monitoring system for personal health management," in Proc.32nd Annu. Int. Conf. IEEE EMBS

[14] Bonato, P., Wearable Sensors and Systems. IEEE Eng.Med. Biol.,volume 29(3)

[15] May year 2010. Chen X, Ho CT, Lim ET,KyawTZ ,Cellular Phone based online ECG Processing forAmbulatory and Continuous Detection.

[16] S. M. Lucas, A. Panaretos, L. Sosa, A. Tang, S.Wong and R.Young. "ICDAR 2003 Robust Reading Competitions",In 7th International Conference on Document Analysis and Recognition- 2003.

[17] P. Viola and M. Jones, "Fast and Robust Classification using Asymmetric AdaBoost and aDetector Cascade", In Proc. of NIPS01, 2001

[18] Aliverti,R Dellaca, P. P elosi, D.Chiumello, A. P edotti, and Gattinoni, "Optoelectronic plethysmography in intensive care patients," Am. J. Respir.Crit.

[19] Y. Wu J. Lim, and M.-H. Yang, "Online object tracking: A benchmark," IEEE 2013.

[20] Hersh M., Johnson M., " Assistive technology for visually impaired and blind people", Springer, 2012.

[21] Rodriguez, "Assisting the Visually Impaired: Obstacle Detection and Warning System Acoustic Feedback 17476 – 17496, Sensors",IEEE.

[22] Nanayakkara S. C., Shilkrot R. and Maes P,“EyeRing:An Eye on a Finger,”IEEE, 2012.

[23] John Canny,“A computational approach to edge detection.Pattern Analysis and Machine

Intelligence,” IEEE ,,1986.

[24] Joseph Schlecht, Björn Ommer, “Contour-based Object Detection” IEEE, 2011.

[25] Ostu N.A.,“Threshold Selection Method from Gray-Level Histograms,” IEEE,2014.

CHAPTER NO:-7
DATASHEETS