RASPBERRY PI BASED OPTICAL CHARACTER RECOGNITION USING COMPUTER VERSION AND CNN ALGORITHM

A PROJECT REPORT

Submitted by

V.THANUSH - 510918106014

J. THIYAGARAJAN - 510918106015

C.MEIYAZHAGAN - 510918106009

K. KAMALESH - 510918106302

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

in

ELECTRONICS AND COMMUNICATION ENGINEERING

GLOBAL INSTITUTE OF ENGINEERING AND TECHNOLOGY MELVISHARAM

ANNA UNIVERSITY: CHENNAI 600 025

June 2022

ANNA UNIVERSITY : CHENNAI 600025

BONAFIDE CERTIFICATE

Certified that this project report **"RASPBERRY PI BASED OPTICAL CHARACTER RECOGNITION USING COMPUTER VISION BASED CNN ALGORITHM"** is the bonafide work of **"V.THANUSH** (5109178106014),**J.THIYAGARAJAN**(510910106015),**C.MEIYAZHAGAN** (510918106009),**K.KAMALESH**(510918106302)**"** who carried out the project work under my supervision.

SIGNATURE	**SIGNATURE**
Mr. P. SRIDHAR M.E.,	Mr. P. SRIDHAR M.E.,
HEAD OF THE DEPARTMENT	**SUPERVISOR**
	ASSISTANT PROFESSOR
DEPARTMENT OF ECE	DEPARTMENT OF ECE
Global Institute of Engineering	Global Institute of Engineering
and Technology,	and Technology,
Melvisharam-632509	Melvisharam-632509

Submission for the project viva voce held on ___________

INTERNAL EXAMINER EXTERNAL EXAMINER

ACKNOWLEDGEMENT

We express our sincere thanks to our Honorable managing Trustee Chairperson **Shri. HEMA PRASAD**. We wish to express us since thanks to our director of administration **Shri. RAM PRASAD VEDULA**. Our sincere thanks to Managing Director of Academics, **Shri. Dr. SUHAS VEDULA** or their constant support and the facility provided for successful completion of the project.

Our sincere thanks to our **Principal Dr. S. SENGOTTAIYAN M.Tech., Ph.D.,** and Our **Head of the Department P. SRIDHAR M.E.,** Electronics and Communication Engineering, for her inspiring support to our project.

We are also grateful to our internal guide **Mr.P.SRIDHAR M.E.,** Assistant Professor Department of Electronic and Communication and Engineering for her valuable suggestion towards our project.

We take this opportunity to thank all our **Staff member** for their help to make this project a success. Their astute guidance, invaluable suggestion, enlightening comments and constructive criticism always kept our spirits up during this proclaim work.

Our experience in working on this project has been wonderful. We hope the knowledge, practical and theoretical, that we have gained through this project will helpme in future endeavor.

ABSTRACT

Reading is apparently essential in today’s society. Printed text everywhere in the form of reports, receipts, bank statements, classroom handouts, instructions on medicine bottles, etc. And while optical aids, video magnifiers, and screen readers can help blind users and those with low vision to access, there documents. There are few devices that can provide good common hand-held objects printed with text such as invention packages. Objects printed with text such as prescription medication bottles. The most well-known answers for tag restriction in computerized pictures are through the execution of edge extraction, morphological administrators, and Sobel administrator. An edge methodology is ordinarily straightforward and quick. Sobel administrator for edge discovery gives constructive outcomes on the picture.The confinement of tags through morphologically based methodologies is not defenseless to clamor but rather is moderate in execution. After the limitation of the tag comes the character division process. Normal character division procedures depend on histogram investigation andthresholding. Other late methodologies proposed are the utilization of counterfeit neural systems.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVATION

OCR	OPTICAL CHARACTER RECOGNITION
TTS	TEXT TO SPEECH
RTOS	REAL TIME OPERATION SYSTEM
SOC	SYSTEM ON CHIP
PIR	PASSIVE INFRARED SYSTEM
CMS	CONTENT MANAGEMENT SYSTEM
GUI	GRAPHICAL USERS INTERFACES
OOP	OBJECT ORIENTED PROGRAMMING
HW	HARDWARE
SW	SOFTWARE
USB	UNIVERSAL SERIAL BUS
GPIO	GENERAL PURPOSE INPUT/OUTPUT
HDMI	HIGH DEFINITION MULTIMEDIA INTERFACE

CHAPTER 1

INTRODUCTION

1.1 OVERVIEW

Blind people are unable to perform visual tasks. For instance, text reading requires the use of a braille reading system or a digital speech synthesizer (if the text is available in digital format). The majority of published printed works does not include braille or audio versions, and digital versions are still a minority. Thus, the development of a mobile application that can perform the image to 1.or in another support,has a great potential and utility. The technology of optical character recognition (OCR) enables the recognition of texts from image data. This technology has been widely used in scanned or photographed documents, converting them into electronic copies, which one can edit, search, play its content and easily carry. The technology of speech synthesis (TTS) enables a text in digital format to be synthesized into human voice and played through an audio system. The objective of the TTS is the automatic conversion of sentences, without restrictions, into spoken discourse in a natural language, resembling the spoken form of the same text, by a native speaker of the language. This technology has had significant progress over the last decade, with many systems being able to generate asynthetic speech very close to the natural voice. Research in the area of speech synthesis has grown as a result of its increasing importance in many new applications.

This part presents an introduction to OCR and TTS technologies used in this project as well as studies and work within the theme of this project.

1.1.1 Optical Character Recognition (OCR) overview

Optical character recognition, usually designated by the acronym OCR, is the process of recognition and automatic conversion of existing characters in the written-support image into the text format, which can then be used in various applications. The OCR has been widely studied and has displayed considerable advances regarding the performance and accuracy of the obtained results.

The process of optical character recognition can be summarized as a process that follows a set of steps:

- Optical image acquisition
- Location and segmentation
- Preprocessing
- Feature extraction
- Classification
- Postprocessing

The final text is then converted into the desired document format (rtf, txt, pdf, etc.).

1.1.2 Text-to-Speech (TTS) overview

Voice synthesis, defined as TTS (acronym for Text-To-Speech), is a computer system that should be able to read aloud any text, regardless of its origin.

The use of TTS aims to produce human voice artificially. Voice synthesis is a complex process and complex algorithms are needed to produce an intelligible and natural result. TTS synthesis makes use of techniques of Natural Language Processing. Since the text to be synthesized is the first entry of the system, it must be the first to beprocessed.

There are several techniques to create a synthesized voice :

- Articulatory synthesis
- Formant synthesis
- Concatenation synthesis
- Hidden Markov models synthesis

The main synthesis techniques, presented above, are the methods used in the study and development of speech synthesis systems. However, a way to profit from the inherent advantages of each technique is to use a hybrid of the various techniques in the development of future systems speech synthesis.

The quality of a speech synthesis can be determined by its naturalness and intelligibility. Naturalness is the characteristic that describes how close to the human voice is the sound obtained by TTS; intelligibility refers to the ease with which the sound is understood in complex situation.

1.2 RASPBERRY PI-3

The Raspberry Pi 3 Model B+ is the latest product in the Raspberry Pi 3 range, boasting a 64-bit quad core processor running at 1.4 GHz, dual-band 2.4 GHz and 5 GHz wireless LAN, Bluetooth 4.2/BLE, faster Ethernet, and PoE capability via a separate PoE HAT

The dual-band wireless LAN comes with modular compliance certification, allowing the board to be designed into end products with significantly reduced wireless LAN compliance testing, improving both cost and time to market.

The Raspberry Pi 3 Model B+ maintains the same mechanical footprint as both the Raspberry Pi 2 Model B and the Raspberry Pi 3 Model B.

Fig 1.1: Raspberry Pi 3 B+

1.3 BASIC PRINCIPLE

In this project Raspberry Pi is the main processor, Camera is used to detect the characters. In this concept we integrated smart sensors such as PIR and Ultrasonic sensor, here PIR sensor isused to detect the motion of human presence once disabled person entered in to the room then it will trigger the fan. Ultrasonic is used to detect any obstacles in front of disable people if any obstacle is there vibrator motor will activate and give a alert to the disable person. In this proposed system optical character recognition integrated in raspberry pi so that it will be detectthe characters then voice will be played in the headphone.

CHAPTER 2

LITERATURE SURVEY

2.1 LITERATURE SURVAY

1. Monica Gupta, Alka Choudhary, and Jyotsna Parmar "**Analysis of Text Identification Techniques Using Scene Text and Optical Character Recognition**",2021

This paper presents, reviews, and analyses recent research expansion in the area of optical character recognition and scene text recognition based on various existing models such as convolution neural network, long short-term memory, cognitive reading for image processing, maximally stable extreme regions, stroke width transformation, and achieved remarkable results up to 90.34% of F-score with benchmark datasets such as ICDAR 2013, ICDAR 2019, IIIT5k. The researchers have done outstanding work in the text recognition field. Yet, improvement in text detection in low-quality image performance is required, as text identification should not be limited to the input quality of the image.

Drawbacks:

- We need more man power and monitoring the blind peoples.

2. Rodney Pino, Renier Mendoza, Rachelle Sambayan "**Optical character recognition system for Baybayin scripts using support vector machine**",2021

The proposed system considers the normalization of an individual character to identify if it belongs to Baybayin or Latin script and further classify them as to what unit they represent. This gives us four classification problems, namely: (1) Baybayin and Latin script recognition, (2) Baybayin character classification, (3) Latin character classification, and (4) Baybayin diacritical marks classification. To the best of our knowledge, this is the first study that makes use of Support Vector Machine (SVM) for Baybayin script recognition. This study has largely contributed to the script and character recognition community by providing a new set of data for Baybayin characters, its diacritics, and Latin characters.

Drawbacks:

- In existing system having more disadvantages, here we need to prepare Braille is a system of reading and writing by touch used by the blind. It consists of arrangements of dots which make up letters of the alphabet, numbers, and punctuation marks. So its take more time to read by blind peoples.

3. Shalini Sonth and Jagadish S Kallimani "**OCR Based Facilitator for the Visually Challenged**" 2017

A novel implementation of a smart OCR based reader for the visually impaired. This project has been implemented on an embedded platform, and uses various technologies such as optical character recognition, image processing and text to speech engines. The major goal of this project was to provide an affordable hand held device to the under-represented sections of the society, i.e., the blind and the visually impaired.

Drawbacks:

- Low Accuracy
- High cost

4. R.Anand,T.shanthi R.S.Sabeenian and S.Veni."Real **time noisy dataset implementation of optical character identification using CNN**"2020

Optical character recognition (OCR) is one of the major research problem in realtime applications and it is used to recognise all the characters in an image. As English is a universallanguage, character recognition in English is a challenging task. Deep learning approach is one of the solution for the recognition of optical characters. Aim of this research work is to perform character recognition using convolution neural network with LeNET architecture. Dataset used in this work is scanned passport dataset for generating all the characters and digits using tesseract.

Drawbacks:

- In existing system having more disadvantages, here we need to prepare Braille is a system of reading and writing by touch used by the blind. It consists of arrangements of dots which make up letters of the alphabet, numbers, and punctuation marks. So, its take more time to read by blind peoples.

5. Abhishek Mathur, Akshada Pathare, Prerna Sharma and Sujata Oak. "**AI based Reading System for Blind using OCR**"2019

Then system is an OCR reading system which uses camera application present in your smart phones combined with OCR (Optical Character Recognition). OCR is a mechanism which coverts images of typed, handwritten, or printed text into machine encoded text. This system will help you to take a picture or scan the document present with user using the phone's camera, the image will be scanned and the application will read the text written in English language and convert the output in speech format. The speech output is generated using Text To Speech Module. The purpose of delivering the output in form of voice/speech is to serve the information that is present on the document to the visually impaired.

Drawbacks:

- We need more man power and monitoring the blind peoples.

2.2 DRAWBACKS IN EXISTING SYSTEM

- ❖ In existing system having more disadvantages, here we need to prepare Braille is a system of reading and writing by touch used by the blind.

It consists of arrangements of dots which make up letters of the alphabet, numbers, and punctuation marks. So its take more time to read by blind peoples.

- We need more man power and monitoring the blind peoples.
- Low Accuracy
- High cost

2.3 PROPOSED SYSTEM

- To overcome from the problems in the existing system we have developed a project for Blind People using neural OCR in Open CV.
- The proposed system is to assist blind persons to read text from challenging pattern and background for the purpose of reading document.
- The main objective of our system is to identify the text in the documents.
- Firstly the object image is captured by using a webcam which is embedded within Raspberry Pi and is followed by the image processing.
- To execute an automated system, which check a document and read out its substance to the individual on click of a button.
- The vocal is delivered with the assistance of speaker which would help the individual to readout the content in the scan document.
- Our system helps the blind people to for the purpose of reading without consuming much space.
- The proposed model has a raspberry pi 3B+ surrounded by Ultrasonic sensor, PIR sensor, and a camera. In a pair of glasses we fixed the camera in middle of the glass. This camera can capture the images or live video, according to the command given by the user. PIR sensor will detect the motion around the person and ultrasonic sensor will give the distance of the object from the user. Headphones were used

here to intimate to the user through the speaker.

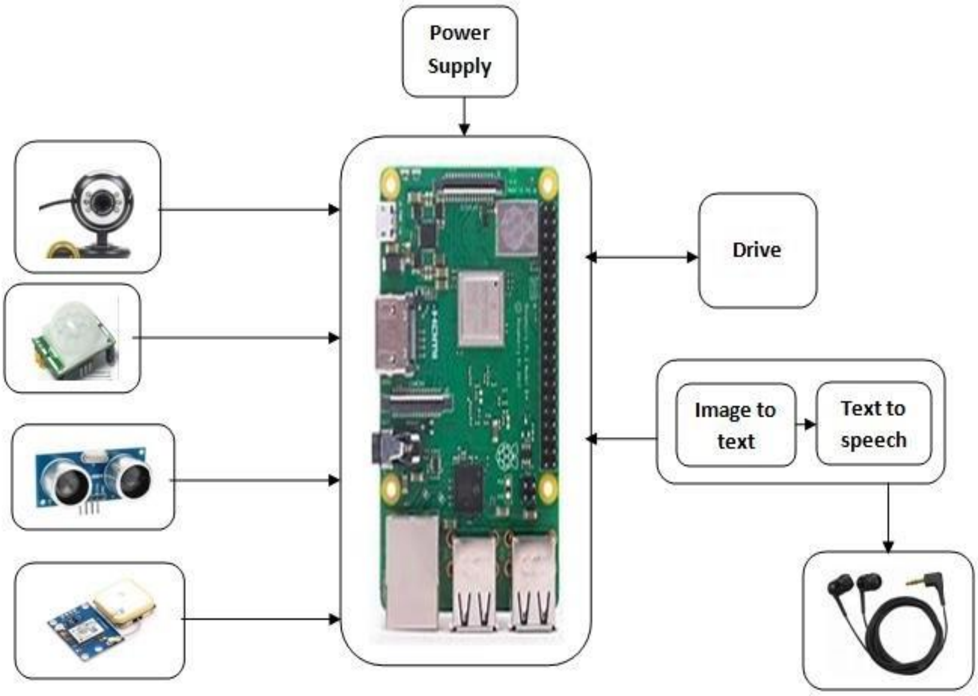

Fig 2.1: Hardware Model of the Proposed System

In this section, the methodology to pursue the project is presented, from image capture with the camera, through implementation of OCR and TTS tools.

The overall system architecture is represented by the following:

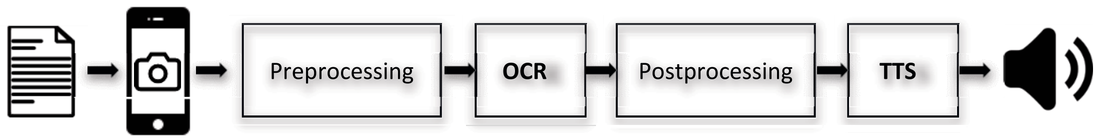

Fig 2.2: System Architecture

2.4 APPLICATIONS

2.4.1 Image capture with the camera

The camera of the mobile device is critical to use the application, since it will be essential for the user to take the picture of the brackets containing text that will be recognized and synthesized.

The image capture is done inside the application itself, thus avoiding the use of additional applications such as access to the photo gallery.

The device used in the project was an Apple iPhone 5. The display screen size is 4.0' with a resolution of 640x 1136 pixels. The camera of the device has a resolution of 8 megapixels with 3264x2448 pixels, autofocus and LED Flash.

The application will run in full screen mode. The captured image will be the one displayed on the screen. The autofocus and auto flash camera options will be set as switched on.

Fig 2.3: Prototype of Proposed Model

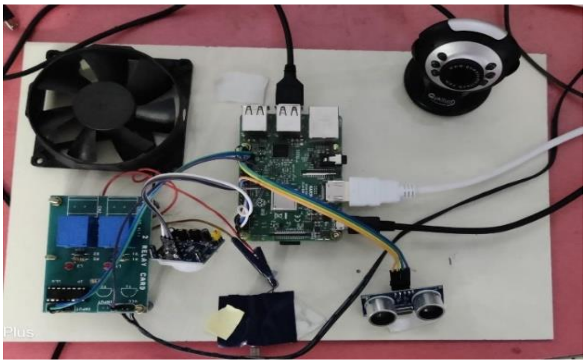

Fig 2.4: Top View

2.4.2 Finding the distance using Ultrasonic

The Ultrasonic sensor is used for finding the distance of an obstacle which is in front of the blind people. The ultrasonic sensor mounted to the glasses beneath the PIR sensor. The sensor basically has Vcc, GND, TRIG and ECHO. Echo will generate the signal at 40 Khz samples, Trig will receive those samples continuously. Whenever we give 5V supply to the ultrasonic sensor the Txr will send the samples in the form of ultra sound which can’t hear by human being. If any obstacle is present in front of the sound signal it will be reflected back, with the help of time lapsed and speed of the signal the distance will be measured.

The distance = (Time lapsed * Speed) / 2

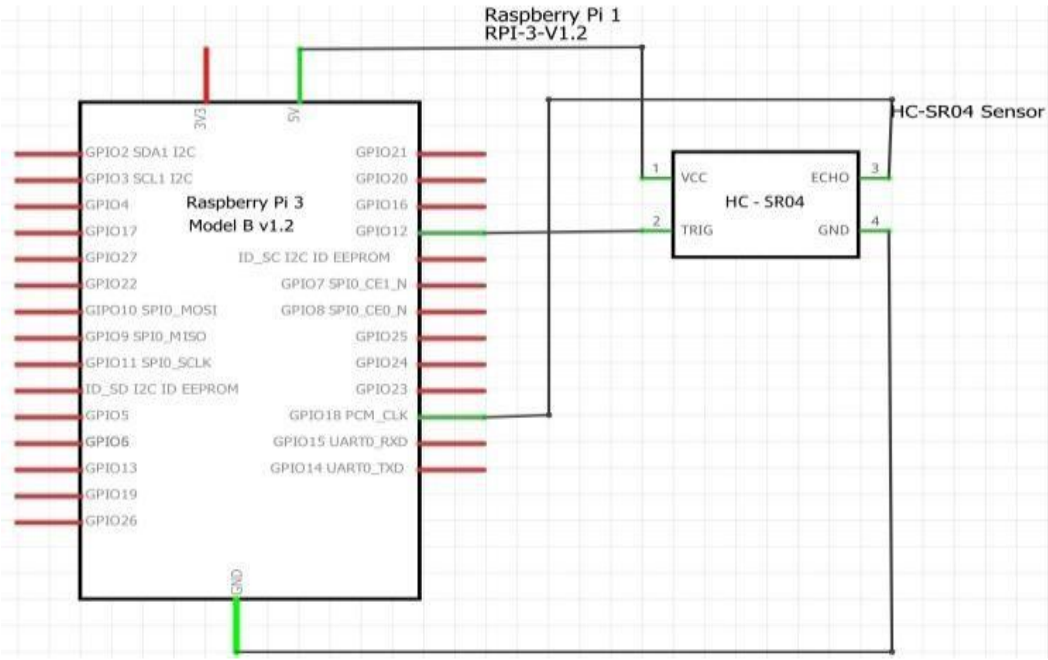

Fig 2.5: Ultrasonic connection with Raspberry Pi

2.4.3 PIR sensing

The PIR sensor will give the logic high when it detects a warm body in its surroundings. PIR sensor has a data pin, Vcc and GND. The data pin will give the data to the raspberry pi, raspberry pi will process the related action. In our project on detection of movement by PIR sensor that will be intimated to the user through the speaker.

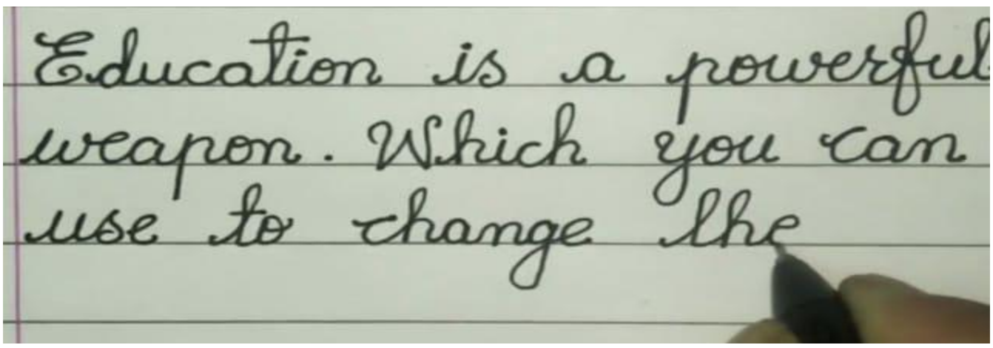

Fig 2.6: Handwriting Image

CHAPTER 3

PROJECT DESCRIPTION

3.1 BLOCK DIAGRAM

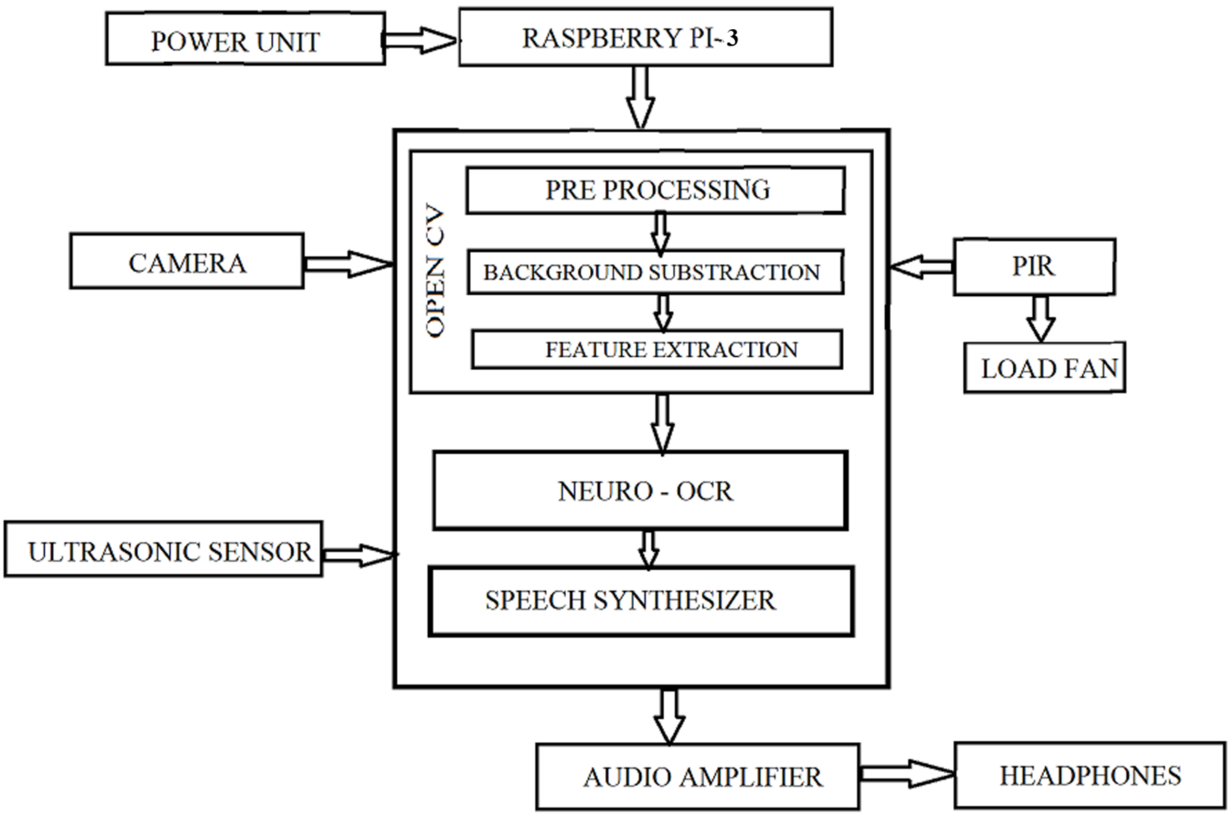

Fig 3.1: Block Diagram

- ***Power unit:***

  Power supply applied is about 5.25V or 2.5A.

- ***Raspberry pi:***

  It is the operating system of the device. It is capable of doing the entire job of an average desktop computer.

- ***Camera:***

  Camera is used to capture the page or the screen of the monitor.

- ***Open CV:***

  OpenCV is a library of programming functions for real time computer vision. It is built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products.

- ***Pre-processing:***

  A program that processes its input data to produce output that is used as input to another program like a compiler.

- ***Background subtraction:***

  Background subtraction, also known as foreground detection, is a technique in the fields of image processing and computer vision wherein an image's foreground is extracted for further processing(text recognition etc.).

- ***Ultrasonic senor:***

  It is used to measure the non-contact distance between the page and the sensor

- ***Neuro-OCR:***

  It is mechanical or electronic conversion of image of typed, handwritten or printed into machine encoded text.

- ***Speech synthesizer:***

  It gives an artificial production of human voice.

- ***Audio Amplifier:***

  It reproduces a low power electronic audio signal.

- ***Headset:***

  It is used to get the audio output.

3.2 SCHEMATIC DIAGRAM OF OUR PROJECT

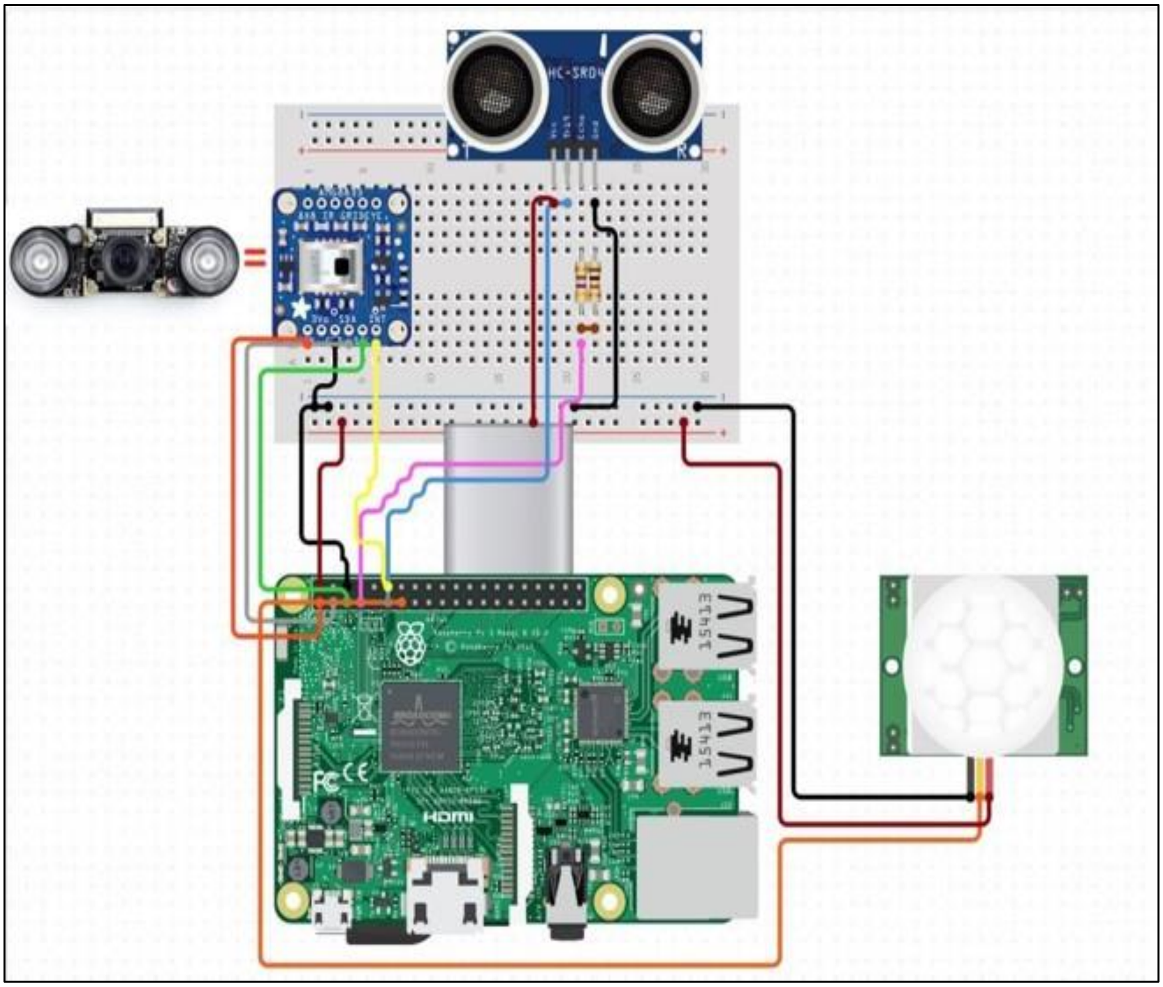

Fig 3.2: Schematic Diagram

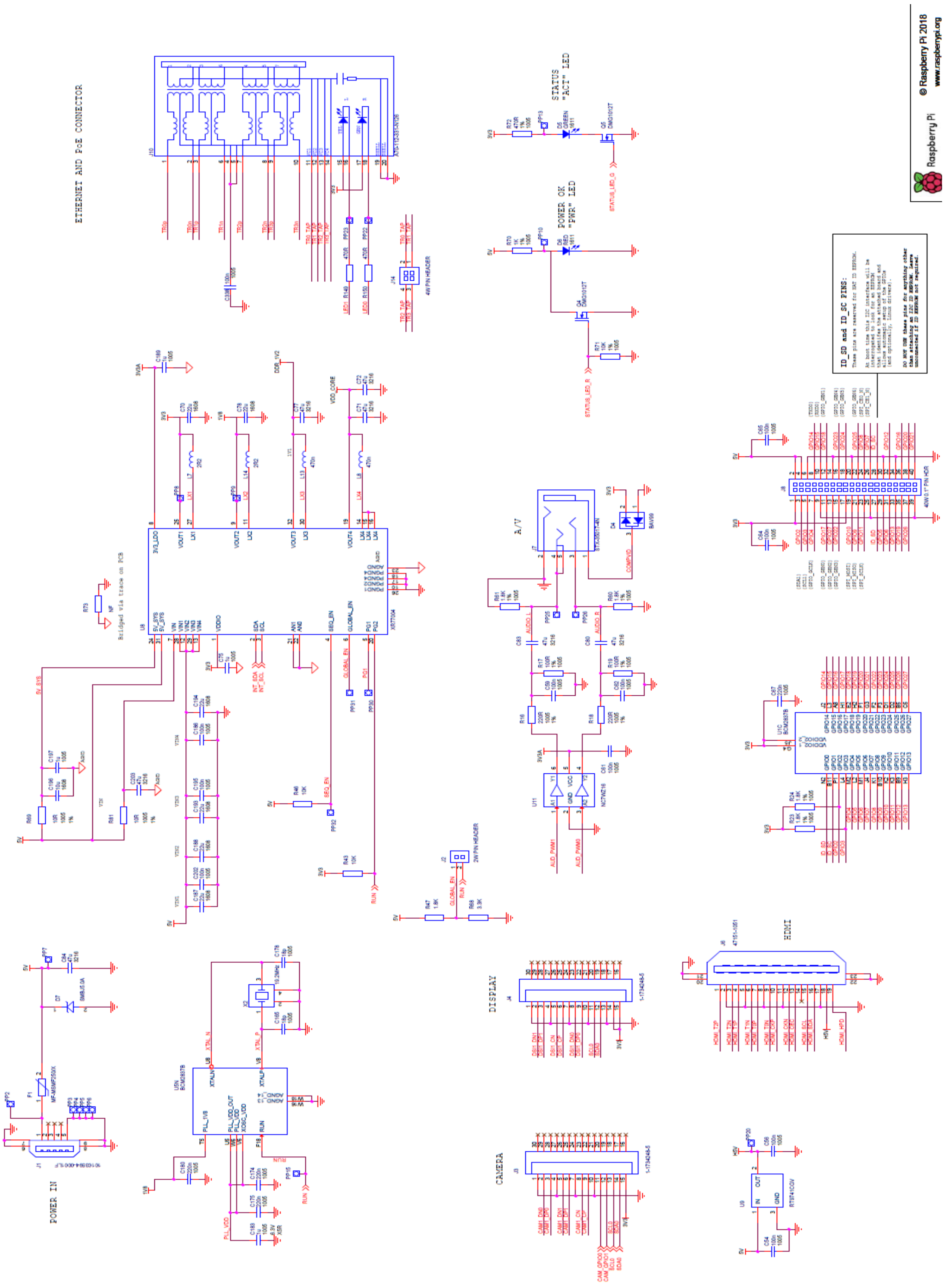

Fig 3.3: Layout Architecture

3.3 LAYOUT OF RASPBERRY Pi

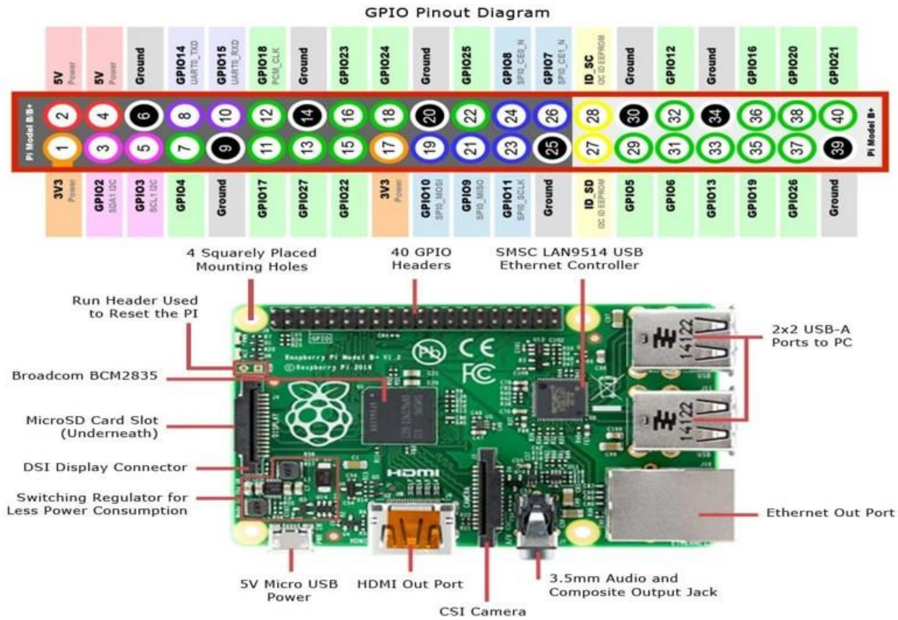

Fig 3.4: Layout of Raspberry Pi-3

3.4 RASPBERRY MECHANICAL DIAGRAM

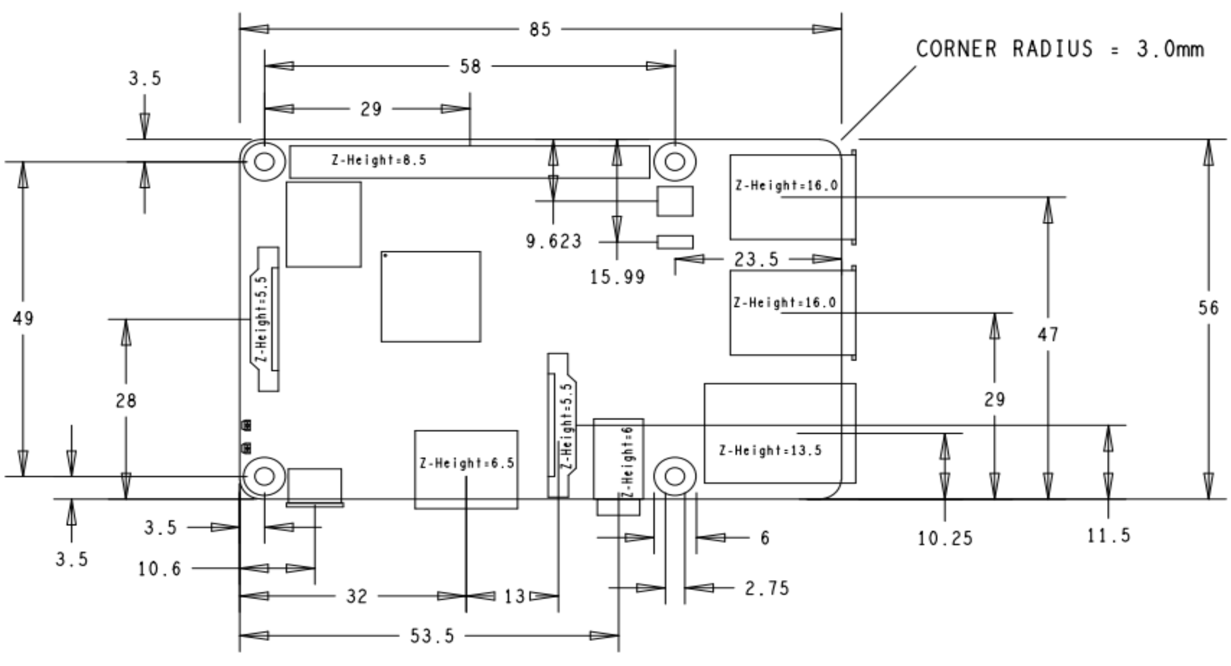

Fig 3.5 Raspberry Mechanical Diagram

CHAPTER 4

HARDWARE DESCRIPTION

4.1 EMBEDDED SYSTEM

As its name suggests, Embedded means something that is attached to anotherthing. An embedded system can be thought of as a computer hardware system havingsoftware embedded in it. An embedded system can be an independent system or itcan be a part of a large system. An embedded system is a microcontroller or microprocessor-based system which is designed to perform a specific task. For example, a fire alarm is an embedded smoke.

An embedded system has three components:

- It has hardware.
- It has application software.
- It has Real Time Operating system (RTOS) that supervises the application software and provide mechanism to let the processor run a process as per scheduling by following a plan to control the latencies. RTOS defines the way the system works. It sets the rules during the execution of application program.A small-scale embedded system may not have RTOS.

So we can define an embedded system as a Microcontroller based, software driven and reliable, real-time control system.

4.1.1 Characteristics of an Embedded System

- ***Single-functioned:***

  An embedded system usually performs a specialized operation and does the same repeatedly. For example: A pager always functions as a pager.

- ***Tightly constrained:***

  All computing systems have constraints on design metrics, but those on an embedded system can be especially tight. Design metrics is ameasure of an implementation's features such as its cost, size, power, and performance. It must be of a size to fit on a single chip, must perform fast enough to process data in real time and consume minimum power to extend battery life

- ***Reactive and Real time:***

  Many embedded systems must continually react to changes in the system's environment and must compute certain results in real time. without any delay. Consider an example of a car cruise controller; it continually monitors and reacts to speed and brake sensors. It must compute acceleration or de-accelerations repeatedly within a limited time; a delayed computation can result in failure to control of the car.

- ***Microprocessors based:***

  It must be microprocessor or microcontroller based.

- ***Memory:***

  It must have a memory, as its software usually embeds in ROM. It does not need any secondary memories in the computer.

- ***Connected:***

  It must have connected peripherals to connect input and output devices.

- ***HW-SW systems:***

  Software is used for more features and flexibility. Hardware is used for performance and security.

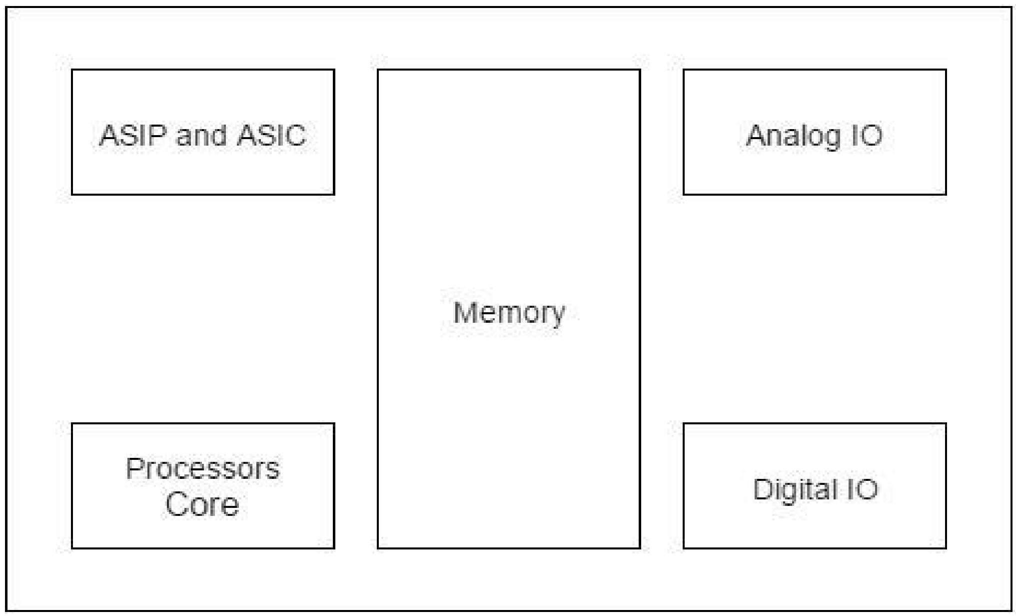

Fig 4.1: Memory Block Diagram

4.1.2 Basic Structure of an Embedded System

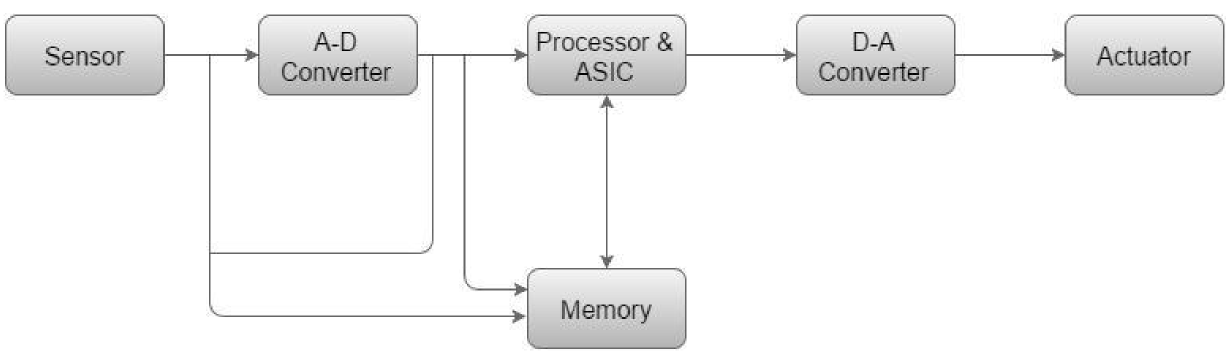

Fig 4.2: Structure of Embedded System

- ***Sensor:***

  It measures the physical quantity and converts it to an electrical signal which can be read by an observer or by any electronic instrument like an A2D converter. A sensor stores the measured quantity to the memory.

- ***A-D Converter:***

  An analog-to-digital converter converts the analog signal sent by the sensor into a digital signal.

- ***Processor & ASICs:***

  Processors process the data to measure the output and store it to the memory.

- ***D-A Converter:***

  A digital-to-analog converter converts the digital data fed by the processor to analog data

- ***Actuator:***

  An actuator compares the output given by the D-A Converter to the actual (expected) output stored in it and stores the approved output.

4.2 MECHANICAL SPECIFICATION

The Compute Modules conform to JEDEC MO-224 mechanical specification for 200 pin DDR2 (1.8V) SODIMM modules (with the exception that the CM3, CM3L modules are 31mm in height rather than 30mmof CM1) and therefore should work with the many DDR2 SODIMM sockets available on the market. (Please note that the pinout of the Compute Module is not the same as a DDR2 SODIMM module; they are not electrically compatible.)

The SODIMM form factor was chosen as a way to provide the 200 pin connections using a standard, readily available and low-cost connector compatible with low-cost PCB manufacture.

The maximum component height on the underside of the Compute Module is 1.2mm. The maximum component height on the top side of the Compute Module is 1.5mm. The Compute Module PCB thickness is 1.0mm +/- 0.1mm.

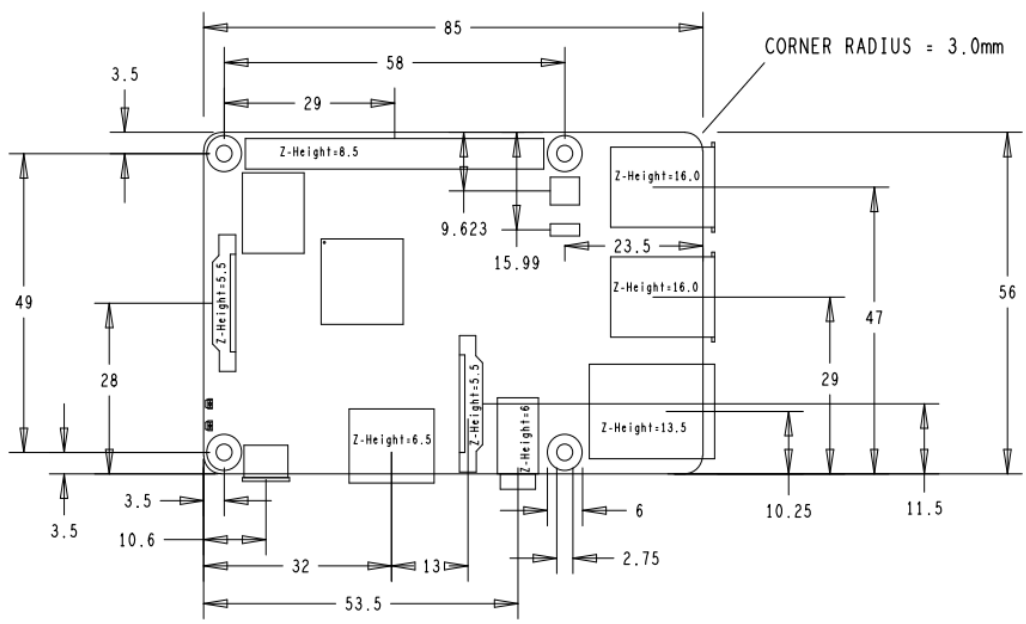

Fig 4.3: Mechanical Specification

4.3 ELECTRICAL SPECIFICATION

Stresses above those listed in Table 4 may cause permanent damage to the device. This is a stress rating only; functional operation of the device under these or any other conditions above those listed in the operational sections of this specification is not implied. Exposure to absolute maximum rating conditions for extended periods may affect device reliability.

Table 4.1: Electrical Specification

Symbol	Parameter	Minimum	Maximum	Unit
VBAT	Core SMPS Supply	-0.5	6.0	V
3V3	3V3 Supply Voltage	-0.5	4.10	V
1V8	1V8 Supply Voltage	-0.5	2.10	V
VDAC	TV DAC Supply	-0.5	4.10	V
GPIO0-27_VDD	GPIO0-27 I/O Supply Voltage	-0.5	4.10	V
GPIO28-45_VDD	GPIO28-27 I/O Supply Voltage	-0.5	4.10	V
SDX_VDD	Primary SD/eMMC Supply Voltage	-0.5	4.10	V

Table 4.2: DC Characteristics

Symbol	Parameter	Conditions	Minimum	Typical	Maximum	Unit
V_{IL}	Input low voltage[a]	VDD_IO = 1.8V	-	-	0.6	V
		VDD_IO = 2.7V	-	-	0.8	V
V_{IH}	Input high voltage[a]	VDD_IO = 1.8V	1.0	-	-	V
		VDD_IO = 2.7V	1.3	-	-	V

I_{IL}	Input leakage current	TA = +85ºC	-	-	5	µA
C_{IN}	Input capacitance	-	-	5	-	pF
V_{OL}	Output low voltage[b]	VDD_IO = 1.8V, IOL = -2mA	-	-	0.2	V
		VDD_IO = 2.7V, IOL = -2mA	-	-	0.15	V
V_{OH}	Output high voltage[b]	VDD_IO = 1.8V, IOH = 2mA	1.6	-	-	V
		VDD_IO = 2.7V, IOH = 2mA	2.5	-	-	V
I_{OL}	Output low current[c]	VDD_IO = 1.8V, VO = 0.4V	12	-	-	mA
		VDD_IO = 2.7V, VO = 0.4V	17	-	-	mA
I_{OH}	Output high current[c]	VDD_IO = 1.8V, VO = 1.4V	10	-	-	mA
		VDD_IO = 2.7V, VO = 2.3V	16	-	-	mA
R_{PU}	Pullup resistor	-	50	-	65	kΩ
R_{PD}	Pulldown resistor	-	50	-	65	kΩ

Table 4.3: AC Characteristics

Pin Name	Symbol	Parameter	Minimum	Typical	Maximum	Unit
Digital outputs	T_{rise}	10-90% rise time[a]	-	1.6	-	Ns
Digital outputs	T_{fall}	90-10% fall time[a]	-	1.7	-	Ns
GPCLK	t_{JOSC}	Oscillator-derived GPCLK cycle-cycle jitter (RMS)	-	-	48	Ps
GPCLK	t_{JPLL}	PLL-derived GPCLK cycle-cycle jitter (RMS)	-	-	20	Ps

4.3.1 Power Requirements

The Compute Module has six separate supplies that must be present and powered at all times; you cannotleave any of them unpowered, even if a specific interface or GPIO bank is unused. The six supplies are as follows:

- VBAT is used to power the BCM283x processor core. It feeds the SMPS that generates the chip core voltage.
- 3V3 powers various BCM283x PHYs, IO and the eMMC Flash.
- 1V8 powers various BCM283x PHYs, IO and SDRAM.
- VDAC powers the composite (TV-out) DAC.
- GPIO0-27 VREF powers the GPIO 0-27 IO bank.
- GPIO28-45 VREF powers the GPIO 28-45 IO bank.

Table 4.4: Power Requirements

Supply	Descripion	Minimum	Typical	Maximum	Unit
VBAT	Core SMPS Supply	2.5	-	5.0+5%	V
3V3	3V3 Supply Voltage	3.3-5%	3.3	3.3+5%	V
1V8	1V8 Supply Voltage	1.8-5%	1.8	1.8+5%	V
VDAC	TV DAC Supply[a]	2.5-5%	2.8	3.3+5%	V
GPIO0-27_VDD	GPIO0-27 I/O Supply Voltage	1.8-5%	-	3.3+5%	V
GPIO28-45_VDD	GPIO28-27 I/O Supply Voltage	1.8-5%	-	3.3+5%	V
SDX_VDD	Primary SD/eMMC Supply Voltage	1.8-5%	-	3.3+5%	V

Exact power requirements will be heavily dependent upon the individual use case. If an on-chip subsys- temis unused, it is usually in a low power state or completely turned off. For instance, if your application does not use 3D graphics then a large part of the core digital logic will never turn on and need power. This is also the case for camera and display interfaces, HDMI, USB interfaces, video encoders and decoders, and so on.

Power chain design is critical for stable and reliable operation of the compute module. We strongly recom- mend that designers spend time measuring and verifying power requirements for their particular use caseand application, as well as paying careful attention to power supply sequencing and maximum supply voltage tolerance.

Table 4.5: Power Units

Supply	Minimum Requirement	Unit
VBAT (CM1)	2000*a*	mW
VBAT (CM3,3L)	3500*a*	mW
3V3	250	mA
1V8	250	mA
VDAC	25	mA
GPIO0-27_VDD	50*b*	mA
GPIO28-45_VDD	50*b*	mA
SDX VDD	50*b*	mA

a. Recommended minimum. Actual power drawn is very dependent on use-case
b. Each GPIO can supply up to 16mA, aggregate current per bank must not exceed 50Ma.

4.3.2 Booting

The 4GB eMMC Flash device on CM3 is directly connected to the primary BCM2837 SD/eMMC interface. These connections are not accessible on the module pins. On CM3L this SD interface is available on the SDX_pins.

When initially powered on, or after the RUN pin has been held low and then released, the BCM2837 will try to access the primary SD/eMMC interface. It will then look for a file called boot code .bin on the primary partition (which must be FAT) to start booting the system. If it cannot access the SD/eMMC device or the boot code cannot be found, it will fall back to waiting for boot code to be written to it over USB; in other words, its USB port is in slave mode waiting to accept boot code from a suitable host.

A USB boot tool is available on Github which allows a host PC running Linux to write the BCM2837 boot code over USB to the module. That boot code then runs and provides access to the SD/eMMC as a USB mass storage device, which can then be read and written using the host PC. Note that a Raspberry Pi can be used as the host machine.

4.4 PERIPHE RAL

BCM283x has in total 54 GPIO lines in 3 separate voltage banks. All GPIO pins have at least two alternative functions within the SoC. When not used for the alternate peripheral function, each GPIO pin may be set as an input (optionally as an interrupt) or an output. The alternate functions are usually peripheral I/Os, and most peripherals appear twice to allow flexibility on the choice of I/O voltage.

On CM1, CM3 and CM3L bank2 is used on the module to connect to the eMMC device and, on CM3 and CM3L, for an on-board I2C bus (to talk to the core SMPS and control the special function pins). On CM3L most of bank 2 is exposed to allow a user to connect their choice of SD card or eMMC device (if required).Bank0 and 1 GPIOs are available for general use. GPIO0 to GPIO27 are bank 0 and GPIO28-45 make up bank1. GPIO0-27_VDD is the power supply for bank0 and GPIO28-45_VDD is the power supply for bank1. SDX_VDD is the supply for bank2 on CM3L. These supplies can be in the range 1.8V-3.3V (see Table 7) and are not optional; each bank must be powered, even when none of the GPIOs for that bank are used.

Note that the HDMI_HPD_N 1V8 and EMM_ EN_N 1V8 pins (on CM1 these were called GPIO46_1V8 and GPIO47_1V8 respectively) are 1.8V IO and are used for special functions (HDMI hot plug detect and boot control respectively). Please do not use these pins for any other purpose, as the software for the Compute Module will always expect these pins to have these special functions. If they are unused please leave them unconnected.

All GPIOs except GPIO28, 29, 44 and 45 have weak in-pad pull-ups or pull-downs enabled when the device is powered on. It is recommended to add off-chip pulls to GPIO28, 29, 44 and 45 to make sure they never float during power on and initial boot.

4.5 GPIO PIN ASSIGNMENT

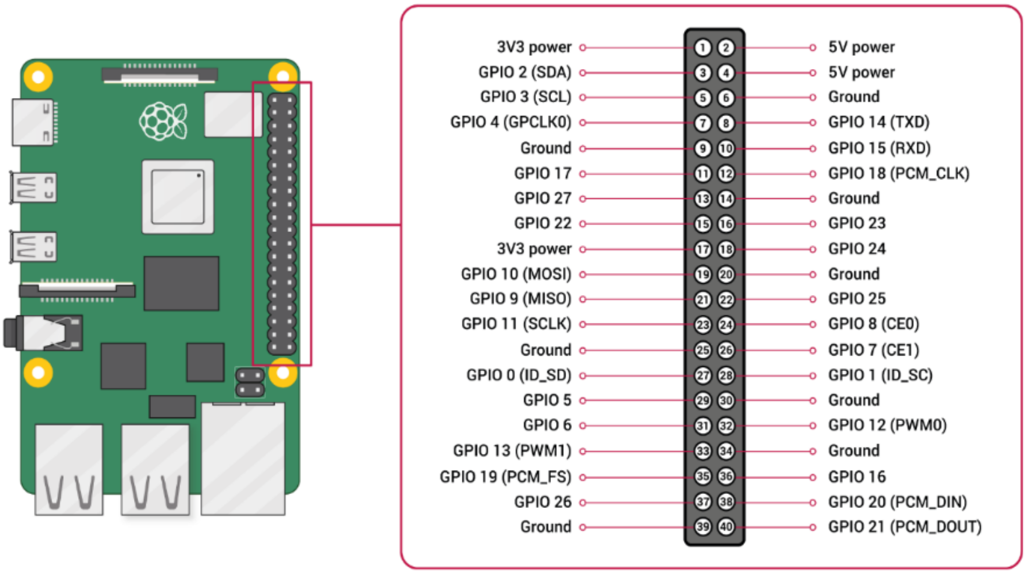

Fig 4.4: GPIO Pin Assignment

4.6 GPIO ALTERNATE FUNCTION

Table 4.6: GPIO Alternate Function

GPIO	Default Pull	ALT0	ALT1	ALT2	ALT3	ALT4	ALT5
0	High	SDA0	SA5	PCLK	-	-	-
1	High	SCL0	SA4	DE	-	-	-
2	High	SDA1	SA3	LCD_VSYNC	-	-	-
3	High	SCL1	SA2	LCD_HSYNC	-	-	-
4	High	GPCLK0	SA1	DPI_D0	-	-	ARM_TDI
5	High	GPCLK1	SA0	DPI_D1	-	-	ARM_TDO
6	High	GPCLK2	SOE_N	DPI_D2	-	-	ARM_RTCK
7	High	SPI0_CE1_N	SWE_N	DPI_D3	-	-	-
8	High	SPI0_CE0_N	SD0	DPI_D4	-	-	-
9	Low	SPI0_MISO	SD1	DPI_D5	-	-	-
10	Low	SPI0_MOSI	SD2	DPI_D6	-	-	-
11	Low	SPI0_SCLK	SD3	DPI_D7	-	-	-
12	Low	PWM0	SD4	DPI_D8	-	-	ARM_TMS
13	Low	PWM1	SD5	DPI_D9	-	-	ARM_TCK
14	Low	TXD0	SD6	DPI_D10	-	-	TXD1

15	Low	RXD0	SD7	DPI_D11	-	-	RXD1
16	Low	FL0	SD8	DPI_D12	CTS0	SPI1_CE2_N	CTS1
17	Low	FL1	SD9	DPI_D13	RTS0	SPI1_CE1_N	RTS1
18	Low	PCM_CLK	SD10	DPI_D14	-	SPI1_CE0_N	PWM0
19	Low	PCM_FS	SD11	DPI_D15	-	SPI1_MISO	PWM1
20	Low	PCM_DIN	SD12	DPI_D16	-	SPI1_MOSI	GPCLK0
21	Low	PCM_DOUT	SD13	DPI_D17	-	SPI1_SCLK	GPCLK1
22	Low	SD0_CLK	SD14	DPI_D18	SD1_CLK	ARM_TRST	-
23	Low	SD0_CMD	SD15	DPI_D19	SD1_CMD	ARM_RTCK	-
24	Low	SD0_DAT0	SD16	DPI_D20	SD1_DAT0	ARM_TDO	-
25	Low	SD0_DAT1	SD17	DPI_D21	SD1_DAT1	ARM_TCK	-
26	Low	SD0_DAT2	TE0	DPI_D22	SD1_DAT2	ARM_TDI	-
27	Low	SD0_DAT3	TE1	DPI_D23	SD1_DAT3	ARM_TMS	-
28	None	SDA0	SA5	PCM_CLK	FL0	-	-
29	None	SCL0	SA4	PCM_FS	FL1	-	-
30	Low	TE0	SA3	PCM_DIN	CTS0	-	CTS1
31	Low	FL0	SA2	PCM_DOUT	RTS0	-	RTS1
32	Low	GPCLK0	SA1	RING_OCLK	TXD0	-	TXD1
33	Low	FL1	SA0	TE1	RXD0	-	RXD1
34	High	GPCLK0	SOE_N	TE2	SD1_CLK	-	-
35	High	SPI0_CE1_N	SWE_N	-	SD1_CMD	-	-
36	High	SPI0_CE0_N	SD0	TXD0	SD1_DAT0	-	-
37	Low	SPI0_MISO	SD1	RXD0	SD1_DAT1	-	-
38	Low	SPI0_MOSI	SD2	RTS0	SD1_DAT2	-	-
39	Low	SPI0_SCLK	SD3	CTS0	SD1_DAT3	-	-
40	Low	PWM0	SD4	-	SD1_DAT4	SPI2_MISO	TXD1
41	Low	PWM1	SD5	TE0	SD1_DAT5	SPI2_MOSI	RXD1
42	Low	GPCLK1	SD6	TE1	SD1_DAT6	SPI2_SCLK	RTS1
43	Low	GPCLK2	SD7	TE2	SD1_DAT7	SPI2_CE0_N	CTS1
44	None	GPCLK1	SDA0	SDA1	TE0	SPI2_CE1_N	-
45	None	PWM1	SCL0	SLC1	TE1	SPI2_CE2_N	-

4.7 DISPLAY PARALLEL INTERFACE

A standard parallel RGB (DPI) interface is available on bank 0 GPIOs. This up-to 24-bit parallel inter- face can support a secondary display. Again this interface is not documented in the Broadcom Peripherals Specification but documentation can be found here.

4.8 SD/SDIO INTERFACE

- The BCM283x supports two SD card interfaces, SD0 and SD1.
- The first (SD0) is a proprietary Broadcom controller that does not support SDIO and is the primary interfaceused to boot and talk to the eMMC or SDX_x signals.
- The second interface (SD1) is standards compliant and can interface to SD, SDIO and eMMC devices; for example on a Raspberry Pi 3 it is used to talk to the on-board BCM43438 WiFi device in SDIO mode.
- Both interfaces can support speeds up to 50MHz single ended (SD High Speed Mode).

4.9 USB

- The BCM283x USB port is On-The-Go (OTG) capable. If using either as a fixed slave or fixed master, please tiethe USB_OTGID pin to ground.
- The USB port (Pins USB_DP and USB_DM) must be routed as 90 ohm differential PCB traces. Note that the port is capable of being used as a true OTG port however there is no official documentation.Some users have had success making this work.
- The Raspberry Pi 3 shares the same SMSC LAN9514 chip as its predecessor, the Raspberry Pi 2, adding 10/100 Ethernet connectivity and four USB channels to the board. As before, the SMSC chip connects to the SoC via a single USB channel, acting as a USB-to-Ethernet adaptor and USB hub.

4.10 HDMI

- BCM283x supports HDMI V1.3a.
- It is recommended that users follow a similar arrangement to the Compute Module IO Board circuitry forHDMI output.

- The HDMI CK_P/N (clock) and D0-D2_P/N (data) pins must each be routed as matched length 100 ohm differential PCB traces. It is also important to make sure that each differential pair is closely phase matched. Finally, keep HDMI traces well away from other noise sources and as short as possible.
- Failure to observe these design rules is likely to result in EMC failure.

4.11 AUDIO AND COMPOSITE (TV OUT)

- The TVDAC pin can be used to output composite video (PAL or NTSC). Please route this signal away from noise sources and use a 75 ohm PCB trace.
- Note that the TV DAC is powered from the VDAC supply which must be a clean supply of 2.5-2.8V. It is recommended users generate this supply from 3V3 using a low noise LDO.
- If the TVDAC output is not used VDAC can be connected to 3V3, but it must be powered even if the TV-out functionality is unused

4.12 TEMPERATURE RANGE AND TERMS

- The operating temperature range of the module is set by the lowest maximum and highest minimum of any of the components used.
- The eMMC and LPDDR2 have the narrowest range, these are rated for -25 to +80 degrees Celsius. Therefore the nominal range for the CM3 and CM3L is -25C to +80C.
- However, this range is the maximum for the silicon die; therefore, users would have to take into account the heat generated when in use and make sure this does not cause the temperature to exceed 80 degrees Celsius.

4.13 SYSTEM-ON-CHIP (SOC)

Built specifically for the new Pi 3, the Broadcom BCM2837 system-on-chip (SoC) includes four high-performance ARM Cortex-A53 processing cores running at 1.2GHz with 32kB Level 1 and 512kB Level 2 cache memory, a Video Core IV graphics processor, and is linked to a 1GB LPDDR2 memory module on the rear of the board.

4.14 ANTENNA

There’s no need to connect an external antenna to the Raspberry Pi 3. Its radios are connected to this chip antenna soldered directly to the board, in order to keep the size of the device to a minimum. Despite its diminutive stature, this antenna should be more than capable of picking up wireless LAN and Bluetooth signals – even through walls.

4.15 ULTRASONIC SENSOR

4.15.1 General Description

Ultrasonic sensor emit ultrasonic pulses, and by measuring the time of ultrasonic pulse reaches the object and back to the transducer. The sonic waves emitted by the transducer are reflected by an object and received back in the transducer. After having emitted the sound waves, the ultrasonic sensor will switch to receive mode. The time elapsed between emitting and receiving is proportional to the distance of the object from the sensor.

4.15.2 Product Description

Ultrasonic transmitter emitted an ultrasonic wave in one direction and started timing when it launched. Ultrasonic spread in the air and would return immediately when it encountered obstacles on the way. At last the ultrasonic receiver would stop timing when it receives the reflected wave. The distance of sensor from the target object is calculated.

It offers excellent non-contact range detection with high accuracy and stable readings in an easy-to-use package. It operation is not affected by sunlight or black material. The supply voltage to the sensor is 5VDC. The sensor has two pins namely trig and echo which is connected to the controller to give digital input.

Fig 4.5: Ultrasonic Sensor

4.15.3 Ultrasonic Features

- **Working Voltage:** 5VDC
- **Quiescent Current :** <2mA Working
- **Current:** 15Ma
- **Detecting Range:** 2cm - 4.5m
- **Trigger Input Pulse width:**10Us

4.15.4 Ultrasonic Application

- Robot navigation
- Obstacle avoidance
- Engineering measurement tools
- Industrial control system

4.16 PASSIVE INFRARED SENSOR (PIR)

A passive infrared sensor (PIR sensor) is an electronic sensor that measures infrared (IR) light radiating from objects in its field of view. They are most often used in PIR-based motion detectors. PIR sensors are commonly used in security alarms and automatic lighting applications.

A PIR detector is a motion detector that senses the heat emitted by a living body. These are often fitted to security lights so that they will switch on automatically if approached. They are very effective in enhancing home security systems.

4.16.1 Operation Of PIR Sensor

A few mechanisms have been used to focus the distant infrared energy onto the sensor surface. The window may have Fresnel lenses moulded into it. Alternatively, sometimes PIR sensors are used with plastic segmented parabolic mirrors to focus the infrared energy; when mirrors are used, the plastic window cover has no Fresnel lenses melded into it. A filtering window (or lens) may be used to limit the wavelengths to 8-14 micrometers, which is most sensitive to human infrared radiation (9.4 micrometers being the strongest).

PIRs can have more than one internal sensing element so that, with the appropriate electronics and Fresnel lens, it can detect direction. Left to right, right to left, up or down and provide an appropriate output signal.

Fig 4.6: PIR Sensor

4.17 WEB CAMERA

4.17.1 General Description

Active WebCam captures images up to 30 frames per second from any video device including USB cameras, Analog cameras connected to capture card, TV- boards, camcorders with FireWire (IEEE 1394) interface and from Network cameras.

When the program detects motion in the monitored area, it can sound an alarm, e-mail you the captured images, and start broadcasting or record a video. The program has features to add text captions and image logos to the images, to place a date/time stamp on each video frame, and to adjust the frame rate, picture size, and quality.

4.17.2 Product Description

A webcam is a video camera that feeds or streams its image in real time to or through a computer to computer network. When "captured" by the computer, the video stream may be saved, viewed or sent on to other networks via systems such as the internet, and email as an attachment.

When sent to a remote location, the video stream may be saved, viewed or on sent there. Unlike an IP camera (which connects using Ethernet or Wi-Fi), a webcam is generally connected by a USB cable, or similar cable, or built into computer hardware, such as laptops.

Fig 4.7: web camera

4.17.3 Applications

- ❖ Video calling and video conferencing.
- ❖ Video monitoring.
- ❖ Health care (for capturing arterial pulse rate).
- ❖ Input control devices.
- ❖ Commerce (e.g Webcam social shopper).
- ❖ Home security.

4.18 VIBRATOR MOTOR

This vibration motor is a very powerful and user-friendly motor, which can be used with a variety of products. This includes pagers, GPS devices, mobile phones or even toys.

4.18.1 Features

1. Small Compact Size
2. Metal Body
3. Power: 3.7V@0.195A
4. Speed 19000RPM
5. Powerful Motor

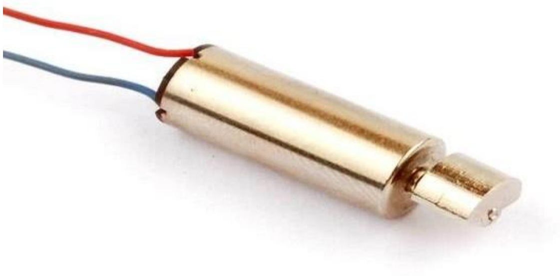

Fig 4.7: Vibrator Motor

CHAPTER 5

SOFTWARE DESCRIPTION

5.1 PYTHON

- Python is a powerful multi-purpose programming language created by Guido vanRossum.
- It has simple easy-to-use syntax, making it the perfect language for someone trying tolearn computer programming for the first time.

5.2 EMBEDDED SYSTEM SOFTWARE

A typical industrial microcontroller is quite unsophisticated compared to a typical enterprise desktop computer and generally depends on a simpler, less-memory- intensive program environment. The simplest devices run on bare metal and areprogrammed directly using the chip CPU's machine code language. Often, however, embedded systems use operating systems or language platforms tailored to embedded use, particularly where real-time operating environments must be served. At higher levels of chip capability, such as those found in SoCs, designers have increasingly decided that the systems are generally fast enough and tasks tolerant of slight variations in reaction time that "near-real-time" approaches are suitable. In these instances, stripped-down versions of the Linux operating system are commonly deployed, though there are also other operating systems that have been pared down to run on embedded systems, including Embedded Java and Windows IoT (formerly Windows Embedded). Generally, storage of programs and operating systems on embedded devices make use either of flash or rewritable flash memory.

5.3 RASPBERRY PI SOFTWARE

- ARMv6 (CM1) or ARMv7 (CM3, CM3L) Instruction Set
- Mature and stable Linux software stack
- Latest Linux Kernel support

- ❖ Many drivers upstreamed
- ❖ Stable and well supported userland
- ❖ Full availability of GPU functions using standard APIs

5.4 FEATURES OF PYTHON

- ***Easy to code:***

  Python is high level programming language. Python is very easy to learn language as comparedto other language like c, c#, java script, java etc. It is very easy to code in python language and anybody can learn python basic in few hours or days. It is also developer-friendly language.

- ***Free and Open Source:***

  Python language is freely available at official website and you can download it from the givendownload link below click on the Download Python keyword. Since, it is open-source, this means that source code is also available to the public. So you can download it as, use it as wellas share it.

- ***Object-Oriented Language:***

  One of the key features of python is Object-Oriented programming. Python supports object oriented language and concepts of classes, objects encapsulation etc.

- ***GUI Programming Support:***

  Graphical Users interfaces can be made using a module such as PyQt5, PyQt4, wxPython or Tk in python. PyQt5 is the most popular option for creating graphical apps with Python.

- ***High-Level Language:***

  Python is a high-level language. When we write programs in python, we do not need to remember the system architecture, nor do we need to manage the memory.

- ***Extensible feature:***

Python is a Extensible language. we can write our some python code into c or c++ language and also we can compile that code in c/c++ language.

- ***Python is Portable language:***

Python language is also a portable language. for example, if we have python code for windowsand if we want to run this code on other platform such as Linux, Unix and Mac then we do not need to change it, we can run this code on any platform.

- ***Python is Integrated language:***

Python is also an Integrated language because we can easily integrated python with other language like c, c++ etc.

- ***Interpreted Language:***

Python is an Interpreted Language. because python code is executed line by line at a time. likeother language c, c++, java etc there is no need to compile python code this makes it easier to debug our code. The source code of python is converted into an immediate form called bytecode.

Large Standard Library Python has a large standard library which provides rich set of module and functions so you do not have to write your own code for every single thing. There are many libraries present in python for such as regular expressions, unit-testing, web browsersetc.

- ***Dynamically Typed Language:***

Python is dynamically-typed language. That means the type (for example- int, double, long etc)for a variable is decided at run time not in advance. Because of this feature we don’t need to specify the type of variable.

5.5 APPLICATIONS OF PYTHON

5.5.1 Web Applications

- You can create scalable Web Apps using frameworks and CMS (Content ManagementSystem) that are built on Python. Some of the popular platforms for creating Web Appsare:Django, Flask, Pyramid, Plone, Django CMS.
- Sites like Mozilla, Reddit, Instagram and PBS are written in Python.

5.5.2 Scientific and Numeric Computing

- There are numerous libraries available in Python for scientific and numeric computing.There are libraries like:SciPy and NumPy that are used in general purpose computing. And, there are specific libraries like: EarthPy for earth science, AstroPy for Astronomyand so on.
- Also, the language is heavily used in machine learning, data mining and deep learning.

5.5.3 Creating Software Prototypes

- Python is slow compared to compiled languages like C++ and Java. It might not be a good choice if resources are limited and efficiency is a must.
- However, Python is a great language for creating prototypes. For example: You can usePygame (library for creating games) to create your game's prototype first. If you like the prototype, you can use language like C++ to create the actual game.

5.5.4 GOOD LANGUAGE TO TEACH PROGRAMMING

- Python is used by many companies to teach programming to kids It is a good language with a lot of features and capabilities. Yet, it's one of the easiest language to learn because of its simple easy-to-use system.

CHAPTER 6

RESULTS AND DICUSSION

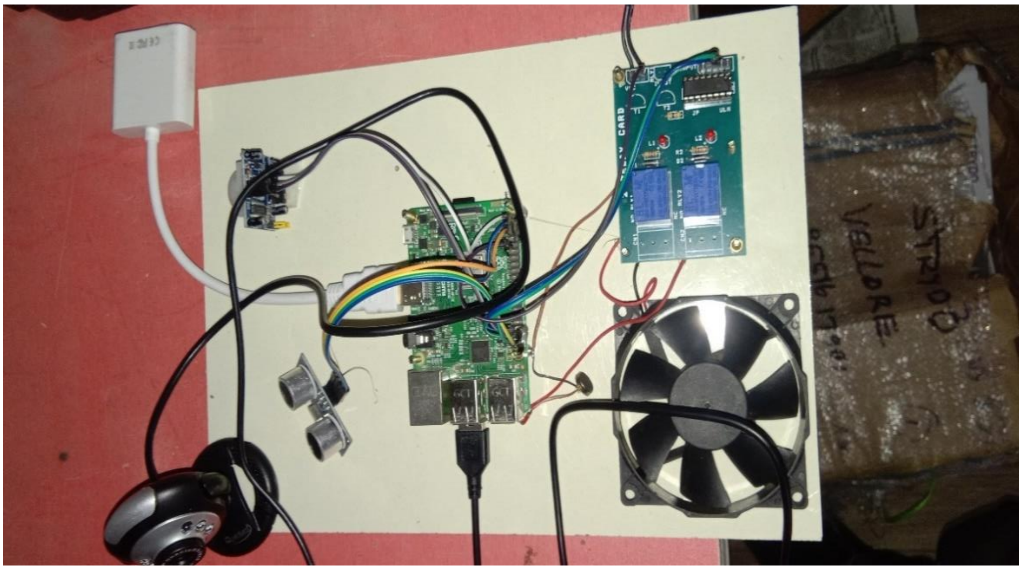

Fig 6.1: Final Result

Image Processing Simulation Environment for English Language.

6.1 IMAGE ACQUISITION

The image is acquired through the Raspberry Pi Camera Module.

6.2 BINARY CONVERSION

The acquired image is converted to a binary format. Bit 0 indicates a black pixel and Bit 1 indicates a white pixel. There are three threshold algorithms namely: Simple threshold,Adaptive Threshold and Otsu threshold. All the 3 algorithms work on a simple principle: if a pixel value is greater than the threshold, it is assigned one value (say white), else it is assignedanother value (say black) Algorithm: a) Convert the image to grayscale and flip the foreground and background to ensure foreground is now white and the background is black. b) Thresholdthe image setting all foreground pixels to 255 and background pixels to 0.

6.3 SKEW CORRECTION

The image acquired may have orientation issues, an thus de-skewing is an important step in image processing. Algorithm: a) Detect the block of text in the image b) Compute the angle ofrotated text c) Rotate the image to de-skew it.

6.4 DE-NOISING

The acquired image may contain noise. Noise is any random variable with zero mean. Algorithm: a) Apply global threshold. b) Apply Otsu's threshold c) Apply Gaussian Blur.

6.5 SEGMENTATION IMAGE

Segmentation partitions an image into multiple segments. The isolated blocks of characters are segmented and are labelled for identity. Connected component labelling is used in computer vision to detect connected regions in binary digital images, although colour images and data with higher dimensionality can also be processed. Algorithm: a) Apply Adaptive Threshold b) Apply dilation and erosion to join the gaps c) Find contours in the image. d) For each contour, draw a bounding rectangle around it.

6.6 TEXT RECOGNITION

AI based reading system for blind converts the image file into text and display o the screen andgive a voice output. For conversion and recognition of text it uses OCR.OCR uses in built library for recognition of text. It successfully recognizes English alphabets from A-Z or a-z.

6.7 NUMBER RECOGNITION

This system not just recognizes the alphabetic words and letters but it also identifies the integeror numerical value from 0-9 and it is also delivered in form of speech.

6.8 EFFECTS OF DIFFERENT LOSS FUNCTIONS

To show the effects of 3 loss functions, we perform an ablation study by trying different loss combinations. Specifically, we try using Lcontent alone, and combining Lcontent with Lperceptual. Then we generate synthetic images with the same style-skeleton pair setting usingthese 2 newly-trained models. Visualizations of some synthetic images are listed in Table IV. We compare their FIDs in Table V. When only Lconten is used, the backgrounds of synthetic images are smooth and clean, which leads to poor FID scores indicating that there is a high mismatch between synthetic and real images. It is observed that

Lperceptual and Ladv can helpmitigate this problem to generate more realistic images that better match the styles in the corresponding style images.

We also compare the WERs/CERs of trained handwritten OCR systems. Experimental resultsare summarized in Table VI. The results show that using L content alone achieves the worst recognition accuracy, which is consistent with the visualization results and FID scores.

6.9 RESULTS ON LARGE SCALE HANDWRITTEN OCR TASK NEXT

We verify the effectiveness of our proposed approach on a large scale handwritten OCR task. We first build a synthesis model with 145k handwriting text line images extracted from whiteboard and handwritten notes. The minibatch formation and hyper-parameters in training are the same as in IAM task except that W is set as 2H. With this model, we render 506.8k online handwriting lines sampled from an in-house online handwriting dataset to text line images. Table VII gives some examples of synthetic images, whose styles look quite similar with the corresponding style images. It is hard to distinguish synthetic images from real images.

Then we combine the synthetic dataset with a large scale real dataset to build handwritten OCR systems. The real dataset contains about 248K handwriting text line images extracted from whiteboard and handwritten note images as training set.

We also use a validation set which contains 28k text line images to guide model training. During evaluation, 2 test sets are used, one is for end-to-end testing called “E2E” with 3,953 text lines, the other is from IAM testingset called “IAM” with 1,861 text lines. As for the handwritten OCR system, a compact DarkNet-DBLSTM character model as in [38] is used. Results are summarized in Table VIII. We observe that adding synthetic data improves the recognition result on both testing sets, especially on the “IAM” testing set. This is because the writing styles of online handwriting samples we use are similar with that of the “IAM” testing set. These results show that with theaugmented synthetic data, the performance of the handwritten OCR system can be improved.

The experimental results of the prototype tabulated according to the performance of the device. The obstacle detection of the device as shown below.

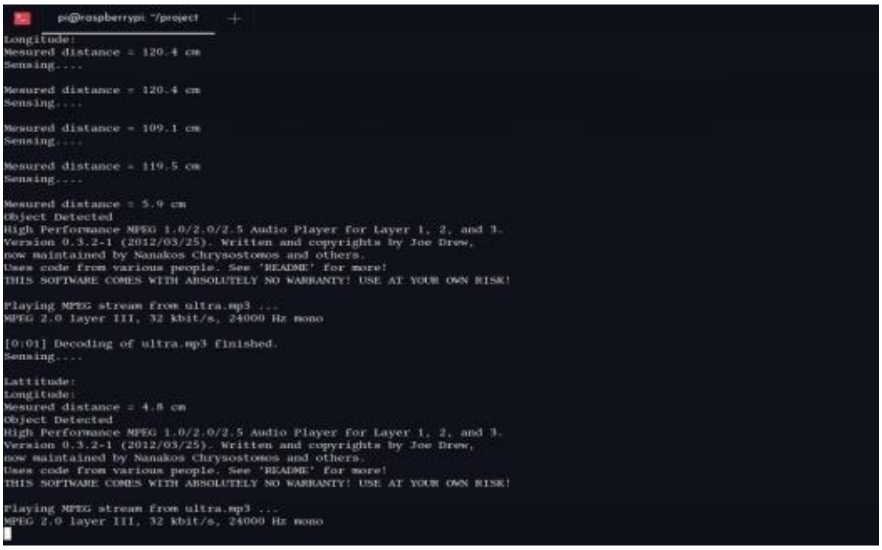

Fig. 6.2: Screenshot of Obstacle Detection Result

Table 6.1: Testing the Ultrasonic Sensor Accuracy

Condition	Instruction	Accuracy
Obstacle detected	Attention, obstacle detected at 45cm	9
Obstacle detected	Attention, obstacle detected at 25cm	10
Obstacle detected	Attention, obstacle detected at 15cm	10
No obstacle	-	10

As well as the PIR sensor is also detecting the movement, a screenshot of result is provided below during the test. The live OCR streaming is the big task in this prototype. Even though it is live streaming we need to translate the image into text and to the speech. So for that we set the loop of live OCR to one time only, when the person pressed the button then only the OCR live will takes place.

At the same time the converted text will be read through the speaker. The converted text will be saved to memory. This can be retrieved thereafter.

Fig 6.3: Image Captured by Camera

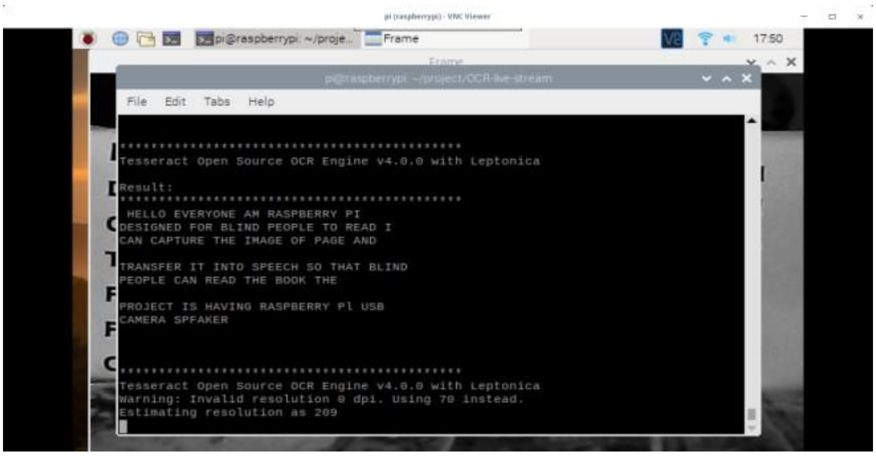

Fig 6.4: Output of the Above Image (image to text)

The image detection is also a part of our prototype. The detected image will be intimated on speaker. The process is shown in the below images.

The image capturing, converting it into text and object detection are experimented many times to get the accurate results. We tabulated the no. of tests and accuracy of the device for image detection process as below.

Table 6.2: Testing the Object Detection and Accuracy

Test case	No. of times	Result
Cell phone	10	85-95%
Book	7	60-75%
Cat	4	70-80%
Tooth brush	3	80-95%
Laptop	5	76-87%
Person	10	75-91%
Pen	8	73-85%

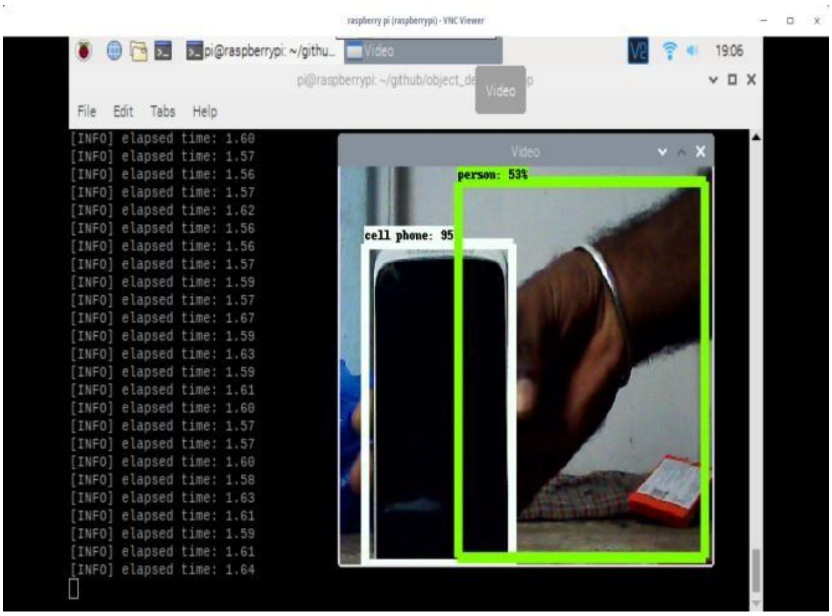

Fig 6.5: Image Showing live Objet Detection

According to the designer's perspective we designed a very well device. Based on our testing the device performance is very well it should be giving 90-93% accurate output. While user's perspective it is little hard to place the reading text book in front of the camera. To overcome this problem we can exchange the more high resolution camera instead.

CHAPTER 7

CONCLUSIONS AND FUTURE SCOPE

7.1 CONCLUSIONS

This paper proposes a novel implementation of a smart OCR based reader for the visually impaired. This project has been implemented on an embedded platform, and uses various technologies such as optical character recognition, image processing and text to speech engines. The major goal of this project was to provide an affordable hand-held device to the under-represented sections of the society, i.e., the blind and the visually impaired.

AI based reading system using OCR is an artificial intelligence reading system developed usinga smart phones camera combined with OCR (Optical Character Recognition). This applicationdetects the text using the camera and scans the text and then converts it into digital text whichis recognized by the system and displays the translated text and gives speech output. To understand the dynamics of the project, a basic idea about what is AI and OCR is required. This report explains the entire working of Language Translator, along with minimum requirements needed to implement it. Hence, visually impaired person can easily use this AI based Reading system as a friendly simple application in all around the globe.

A comprehensive evaluation was carried out under the categories of complementary Sinhala books, old Sinhala books and old Sinhala newspapers. Considering the Sinhala newspaper category, most of the training images are not in a human-readable format. According to the results of our system, the model trained with font iskolapotha gave an accuracy of 87.63% in contemporary Sinhala books.

In the Sinhala old book category, models developed using fonts Malithi Web, LKLug and combined font models using Noto Sans, LKLug and Malithi Web gave accuracies of 87.07%, 87.15% and 87.52%. Meanwhile, in the old

Sinhala newspaper, category 67.02% of accuracy was obtained from the model developed by font iskolapotha. Analyzing linguistics rules and mapping them with computer science is quite challenging for low resource languages like Sinhala and Tamil. As future improvements, we will work on identifying touching and conjoining letters which are frequently occurred in Sinhala and Pali writing systems. Moreover, we plan to integrate n-gram or word embedding based post- processing techniques to enhance the accuracy of the proposed system.

A comprehensive evaluation was carried out under the categories of complementary Sinhala books, old Sinhala books and old Sinhala newspapers. Considering the Sinhala newspaper category, most of the training images are not in a human-readable format. According to the results of our system, the model trained with font iskolapotha gave an accuracy of 87.63% in contemporary Sinhala books. In the Sinhala old book category, models developed using fonts Malithi Web, LKLug and combined font models using Noto Sans, LKLug and Malithi Web gave accuracies of 87.07%, 87.15% and 87.52%. Meanwhile, in the old Sinhala newspaper, category 67.02% of accuracy was obtained from the model developed by font iskolapotha. Analyzing linguistics rules and mapping them with computer science is quite challenging for low resource languages like Sinhala and Tamil. Moreover, we plan to integrate n-gram or word embedding based post- processing techniques to enhance the accuracy of the proposed system.

In this project, we propose a style-conditioned GAN (SCGAN) that can transfer styles of real handwriting images to skeleton images extracted from online handwriting samples. With a novel training set generation strategy, we can successfully build such a model and synthesizephoto-realistic handwriting images. Experimental results show that when combined with real data, these synthetic images can improve the performance of handwritten OCR systems significantly. The proposed approach provides a potential solution to constructing high performance handwritten OCR systems without collecting and labeling massive real handwritten images. As future work, we will investigate better skeletonization method and improve synthetic image quality by better training stratergies and more advanced model architectures.

7.2 FUTURE SCOPE

The scope of this project is promising. Optical Character Recognition finds its applications inMedicine, Online Retails, Education, and more. The proposed system works well for both English and Hindi scripts. Our aim for future work is to extend the same functionality for other Indian regional languages such as Tamil, Kannada, Telugu, etc. Optical character recognition finds its use in the following applications.

- Captcha
- Optical Music Recognition
- Handwriting Detection
- Invoice Imaging
- Automatic Number Detection

REFERENCES

[1]. Faiz Alotaibi, Muhamad Taufik Abdullah, "Optical Character Recognition for Quranic image similarity matching,"IEEE,2017.

[2]. Shalini Sonth, Jagadish S kallimai, "OCR based facilitator for the visuallychallenged,"IEEE,2017.

[3]. Michael D.kim and Jun Ueda, "Dynamics-based motion de-blurring improves the performance of Optical Character Recognition during fast scanning of a robotic eye,"IEEE,2017.

[4]. Sanjana, Monisha, "OCR based automatic book reader for the visually impaired using raspberry pi,"IEEE,2016.

[5]. Velumurugan D. et al, "Hardware Implementation of Smart Reader for Visually Impaired using Raspberry Pi", IJAREEIE, March 2016.

[6]. N. Venkata Rao et al, " Optical Character Technique Recognition Algorithms, Journal of Theoretical and Applied Information Technology",20th January 2016.

[7]. K. Nirmala Kumari et al, "Image Text to Speech Conversion using OCR Technique in Raspberry Pi", IJAREEIE, May 2016.

[8]. Seiichi Uchida, Koichi Kise, "OCR fonts revisited for camera-based characterrecognition,"IEEE,2016.

[9]. D.Velmurugan et al, " A Smart Reader for Visually Impaired People Using RaspberryPI",IJSEC, Vol 6 ,pp 2997-3001

[10]. Nagaraja L. et al, "Vision Based Text Recognition using Raspberry Pi, International Journal of Computer Applications", NCPSIA 2015.

[11]. Kwang in Kim, Keechul Jung,"Texture-based approach for text detection inimages,"IEEE,2013.

[12]. Monica Gupta, Alka Choudhary, and Jyotsna Parmar, "Analysis of Text Identification Techniques Using Scene Text and Optical Character Recognition",IEEE,2021.

[13]. Rodney Pino, Renier Mendoza, Rachelle Sambayan, "Optical character recognition system for Baybayin scripts using support vector machine",IEEE,2021.

[14]. R. Anand, T. Shanthi, R.S.Sabeenian and S.Ven, “Real time noisy dataset implementation of optical character identification using CNN”,IEEE,2020.

[15]. Abhishek Mathur, Akshada Pathare, “Prerna Sharma and Sujata Oak,”AI based Reading System for Blind using OCR”,IEEE,2019

APPENDIX

PYTHON OCR CODE

```
import cv2
import numpy as np
from PIL import Image
from Tkinter import Tk, Label, Button
from pytesser import *
import re
from time import sleep
#from espeak import espeak
import os
#from PIL import Image
os.system("sudo pulseaudio --kill")
os.system("sudo jack_control start")
os.system("sudo jack_control exit")
os.system("sudo pulseaudio --start")
IMAGE_FILE = 'fonts_test.png'
cam = cv2.VideoCapture(0)
s, img = cam.read()
winName = "Movement Indicator"
#cv2.namedWindow(winName, cv2.CV_WINDOW_AUTOSIZE)
while True:
  #cam = cv2.VideoCapture(0)
  #s, img = cam.read()
  s, img = cam.read()
  cv2.imshow( winName,img )
  key = cv2.waitKey(10)
  if key == 27:
    cv2.imwrite(IMAGE_FILE, img)
    #sleep(4)
    img = Image.open(IMAGE_FILE)
    words = image_to_string(img).strip()
    words = re.sub(r"\s+", " ", words, flags=re.UNICODE)
    #words = words.replace(' ', '')
    print words
    text_file = open("Output.txt", "w")
    text_file.write(words)
    text_file.close()
```

```
    os.system("espeak -s110 -f Output.txt --stdout | aplay")
    #ww="hai"
    #os.system("sudo espeak text detected")
    #sleep(2)
    #os.system("sudo espeak -g10 -s80 "+ words)
    #espeak(words)
    sleep(2)
print "ok"
```

ULTRASONIC AND PIR CODE

```
import numpy as np
from time import sleep
import os
from gpiozero import LED
from time import sleep
from gpiozero import Button
import RPi.GPIO as GPIO
import time
motor = LED(26)
buzzer=LED(19)
GPIO_TRIGGER = (20)
GPIO_ECHO = (21)
sw=Button(16)
GPIO.setup(GPIO_TRIGGER, GPIO.OUT)
GPIO.setup(GPIO_ECHO, GPIO.IN)
amt=0
nt=0
def distance():
    # set Trigger to HIGH
    GPIO.output(GPIO_TRIGGER, True)
    # set Trigger after 0.01ms to LOW
```

```
    time.sleep(0.00001)
    GPIO.output(GPIO_TRIGGER, False)
    StartTime = time.time()
    StopTime = time.time()
    # save StartTime
    while GPIO.input(GPIO_ECHO) == 0:
        StartTime = time.time()
    # save time of arrival
    while GPIO.input(GPIO_ECHO) == 1:
        StopTime = time.time()
    # time difference between start and arrival
    TimeElapsed = StopTime - StartTime
    # multiply with the sonic speed (34300 cm/s)
    # and divide by 2, because there and back
    distance = (TimeElapsed * 34300) / 2
    return distance
i=0
j=0
nt=0
while True:
        dist = distance()
        print ("Measured Distance = %.1f cm" % dist)
        print(nt)
        nt=nt+1
        if dist < 20:
            motor.on()
        else:
            motor.off()
```

```
if sw.is_pressed:
    buzzer.off()
else:
    buzzer.on()
sleep(1)
```