iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** V **Month of publication:** May 2021

DOI: https://doi.org/10.22214/ijraset.2021.33940

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

Artificial Intelligence Based Mobile Robot

Bavithran. S[1], Jeswin Paul. P[2], Kandha Raj. M[3]
[1, 2, 3]B. E., Electronics and Communication Engineering Jeppiaar Engineering College, Chennai Anna University: Chennai 60025

Abstract: In this project we have developed an AI pathfinding robot with multiple functions in it. Our robot gets data from distance sensors and decides its path and navigates through a method called SLAM which is an algorithm developed in a robot operating system. We have used a robot operating system along with the SLAM algorithm for navigation and pathfinding of robots.

The SLAM provides a map of the environment at realtime. It can recognize human voice with the google voice recognition API and python voice recognition api. We have used this voice recognition to control the movement of the robot. It also can detect objects on the path using openCV. We used key points algorithm for image and object detection. The robot can be operated in two modes i. Personal Assistant mode and ii) Security mode. In personal assistant mode it operates according to user instruction through voice commands and in security mode, it uses SLAM navigation and path finding algorithm for operating on its own.

It also has an IOT(Internet of things) application which is operated through voice recognition. All the features added in this robot can be monitored in laptop, PC or through web.

I. INTRODUCTION

A. Introduction

The advancement of artificial intelligence made huge changes in the field of robotics every day. This results in development of robots which can plan their paths and decide what to do without the help of human beings. It also has the capability to assist humans in their work and monitor their health.

B. Mapping

For an AI mobile robot, mapping is an important step to explore the unknown environment, and the robot should be able to move and find the pathways to avoid obstacles and achieve the goal. Mobile robot navigation technology is now increasingly sophisticated, with the technology of 2D/3D mapping. The 3D/2D mapping of environment and path deciding is handled by ROS. ROS is an open source framework which has the capability to handle all software layers from low-level up to high-level layers. It has been used for many projects like SLAM, robotic arm, humanoid robots etc.

C. Path Finding

The localization of the robot turns out to be quite similar to classical Markov localization because the poses are modeled probabilistically. The pose distribution is computed recursively over time by performing a prediction step to model state changes and by performing a correction step to incorporate new measurements. In the prediction step, the pose distribution at time t is computed from the pose distribution at time using the probabilistic motion model . The prior can be updated by applying the product rule and by exploiting the fact that the motion model only depends on the previous pose and the latest control .

D. Object Detection

Object detection has become the basic feature in robots. TensorFlow's Object Detection API is an open-source framework built on top of TensorFlow that provides a collection of detection models, pre-trained on the COCO dataset, the Kitti dataset, the Open Images dataset and the iNaturalist Species Detection Dataset. Image recognition assigns a label to an image. A picture of a dog receives the label “dog”.

A picture of two dogs, still receives the label “dog”. Object detection, on the other hand, draws a box around each dog and labels the box “dog”. The model predicts where each object is and what label should be applied. In that way, object detection provides more information about an image than recognition.

E. Speech Recognition

AI Mobile Robots also act as a personal assistant, where Speech Recognition plays an important role. Google Speech Recognition API is used to recognize vocabulary words through a microphone. The speech input is processed and changes the current state of the robot. More advanced speech recognition software will use AI and machine learning. These systems will use grammar, structure, syntax as well as composition of audio and voice signals in order to process speech. Software using machine learning will learn more the more it is used, so it may be easier to learn concepts like accents.

II. LITERATURE SURVEY

A. You only look once: Unified, real-time object detection

Author: Redmon, Joseph

Year: 2018

In this paper, YOLO, a new approach to object detection. Prior work on object detection repurposes classifiers to perform detection. Instead, they frame object detection as a regression problem to spatially separated bounding boxes and associated class probabilities. It also outperforms all other detection methods, including DPM and R-CNN, by a wide margin when generalizing from natural images to artwork on both the Picasso Dataset and the People-Art Dataset.

B. SSD: Single shot multibox detector

Author: Liu, Wei

Year: October 2016

In this paper, a method for detecting objects in images using a single deep neural network. Our approach, named SSD, discretizes the output space of bounding boxes into a set of default boxes over different aspect ratios and scales per feature map location. Experimental results on the PASCAL VOC, MS COCO, and ILSVRC datasets confirm that SSD has comparable accuracy to methods that utilize an additional object proposal step and is much faster, while providing a unified framework for both training and inference.

C. Rich feature hierarchies for accurate object detection and semantic segmentation

Author: Ross, Girshick

Year : 2014

In this paper, An approach to knowledge-assisted semantic video object detection based on a multimedia ontology infrastructure is presented. Semantic concepts in the context of the examined domain are defined in an ontology, enriched with qualitative attributes (e.g., color homogeneity), low-level features (e.g., colormodel components distribution), object spatial relations, and multimedia processing methods (e.g., color clustering).

D. Fast Region-based Convolutional Network

Author: Girschick, Ross

Year: 2015

This paper proposes a Fast Region-based Convolutional Network method (Fast R-CNN) for object detection. Fast R-CNN builds on previous work to efficiently classify object proposals using deep convolutional networks. Compared to previous work, Fast R-CNN employs several innovations to improve training and testing speed while also increasing detection accuracy. Fast R-CNN trains the very deep VGG16 network 9× faster than R-CNN, is 213× faster at test-time

E. Context-Dependent Pre-Trained Deep Neural Networks for Large Vocabulary Speech Recognition

Author : Dahl, George E.; Yu, Dong; Deng, Li; Acero, Alex

Year : 2012

In this work, a novel context-dependent model for large-vocabulary speech recognition (LVSR) that leverages recent advances in using deep belief networks for phone recognition. We describe a pre-trained deep neural network hidden Markov model hybrid architecture that trains the DNN to produce a distribution over senones (tied triphone states) as its output. The deep belief network pre-training algorithm is a robust and often helpful way to initialize deep neural networks' generativity that can aid in optimization and reduce generalization error.

F. Automatic Speech Recognition – A Brief History of the Technology Development

Author : Juang, B. H.; Rabiner, Lawrence R Year : 2015

In this work, Designing a machine that mimics human behavior, particularly the capability of speaking naturally and responding properly to spoken language, has intrigued engineers and scientists for centuries. Since the 1930s, when Homer Dudley of Bell Laboratories proposed a system model for speech analysis and synthesis [1, 2], the problem of automatic speech recognition has been approached progressively, from a simple machine that responds to a small set of sounds to a sophisticated system that responds to fluently spoken natural language and takes into account the varying statistics of the language in which the speech is produced.

G. Simultaneous map building and localization for autonomous mobile robot

Author : Leonard, J.J.; Durrant-whyte, H.F Year : 1991

This work discusses a significant open problem in mobile robotics: simultaneous map building and localization, which the authors define as long-term globally referenced position estimation without a priori information. This problem is difficult because of the following paradox: to move precisely, a mobile robot must have an accurate environment map; however, to build an accurate map, the mobile robot's sensing locations must be known precisely.

H. Towards Real Time Handheld MonoSLAM in Dynamic Environments

Author : Perera, Samunda; Pasqual, Ajith

Year : 2011

This work discusses Traditional monoSLAM assumes stationary landmarks making it unable to cope with dynamic environments where moving objects are present in the scene.

This paper presents the parallel implementation of monoSLAM with a set of independent EKF trackers where stationary features and moving features are tracked separately.

I. Summary Of Literature Review

The chapters of these books show how SLAM, Object detection and Speech recognition works and the algorithm used in it. In SLAM they have explained about methods of mapping and methods of path finding and localisation. In object detection they explained about several methods used to recognise any object and machine and deep learning algorithms used. Various methods used in speech recognition are also explained in the above books.

III. PROPOSED SYSTEM

A. Existing System

In the existing system, the robot navigation system used sensors like UltraSonic sensor and Infrared sensor, to calculate the distance between the sensors and the obstacles.

These values are calculated and the current direction of the robot is modified. The object detection in the existing system detects surrounding objects and makes the robot follow a particular object or image. The voice detection in the existing system can detect 200 to 300 words and provide output regarding the words.

1) Drawbacks

a) It cannot generate and map the environment.

b) It doesn't have AI path finding ability.

c) The object detection has only the ability to detect a limited number of objects.

d) Building these types of robots is difficult and costly.

e) The voice recognition has Lack of Accuracy and Misinterpretation.

B. Proposed System

To overcome these drawbacks we have implemented SLAM for mapping the environment using the data collected from the distance sensor like ultrasonic sensor and localization by using map and vehicle odometry in place of ultrasonic obstacle avoidance. Also we have implemented unlimited real time object detection. We have used Google API for speech recognition and response.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

1) Advantages
a) This method can generate and map the environment.
b) It can generate paths and navigate the robot to a particular point.
c) Object detection system has the ability to detect more objects in real time.
d) Building these types of robots is cheap.
e) It can be used for IOT purposes.

C. Block Diagram

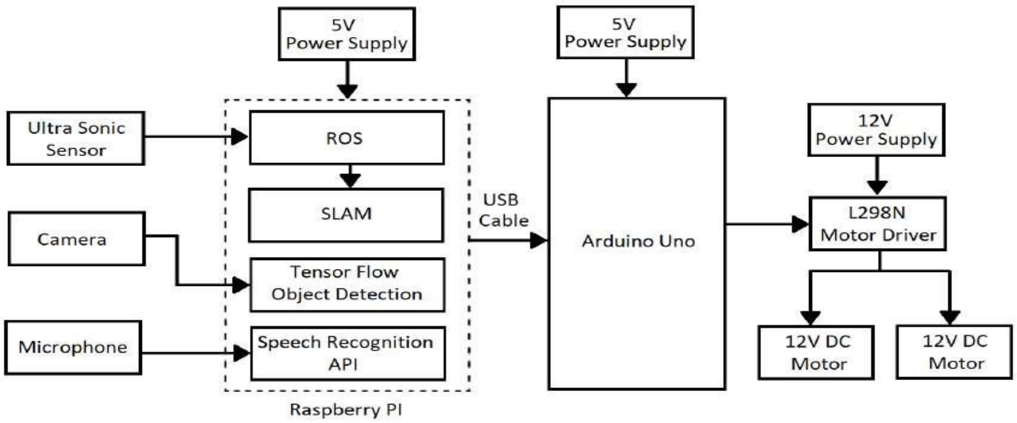


Figure 3.1 Block diagram

The above figure 3.1 represents the block diagram of the proposed system. Here, the dotted line parts are done in the Raspberry pi OS. The output is fed to the Arduino through USB cable. The USB cable acts as a bridge between the Arduino and Raspberry pi. Arduino gets input from Raspberry pi and control motors through motor driver L298N.
The distance sensor gives data as input to the raspberry pi and ROS map the environment using SLAM g-mapping. The localization data of SLAM is given to Arduino to control motor and robot odometry.
The image from the camera is given as input to raspberry pi. The image input is processed and recognized by openCV for object detection.
The voice is taken from the microphone on the camera and processed through speech recognition API and provides data to the arduino to control the motors.

D. Architecture Diagram

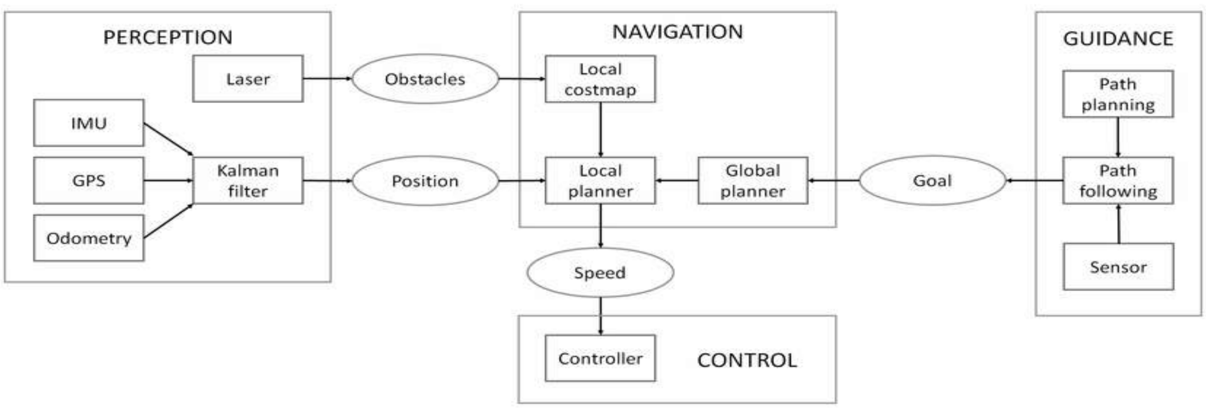


Figure 3.2 ROS Architecture

The figure 3.2 represents the architecture diagram and above procedures are done in Raspberry pi with ROS.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

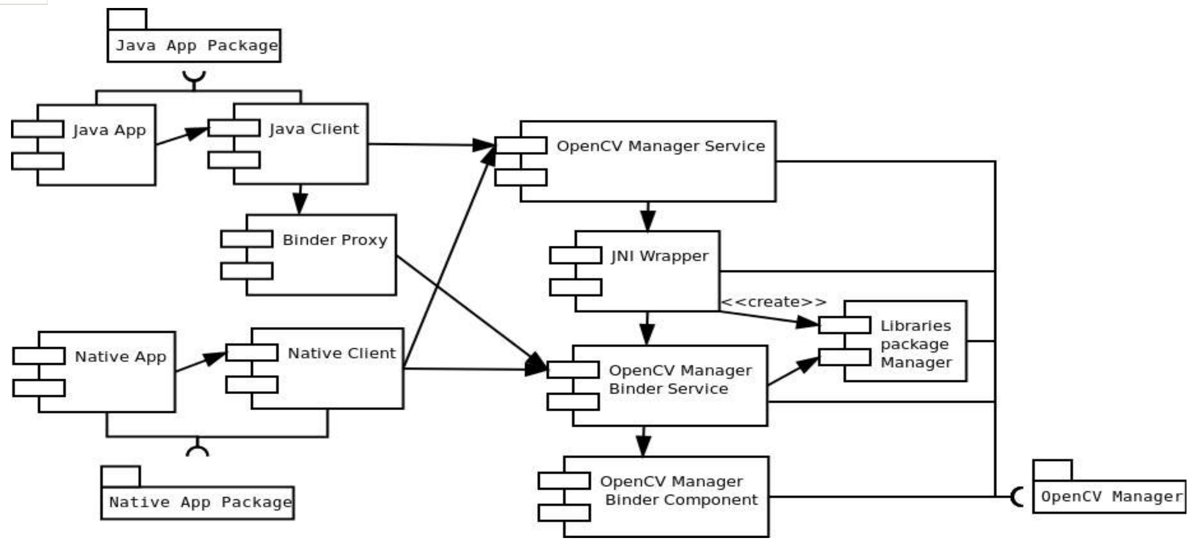


Figure 3.3 Object Detection Architecture

The figure 3.3 represents the architecture diagram of object detection and above procedures are done in Raspberry pi with openCV.

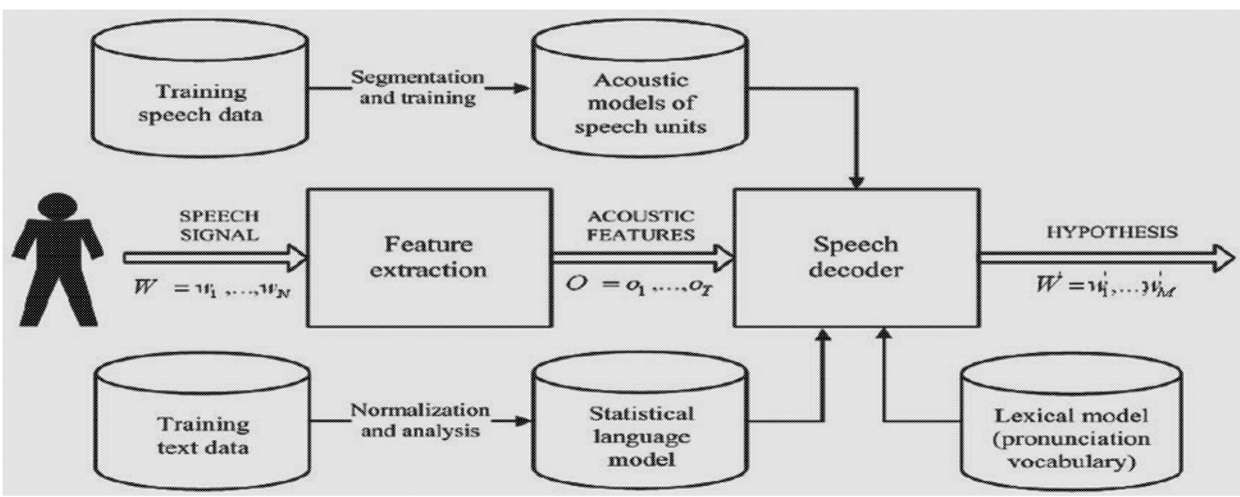


Figure 3.4 Speech Recognition Architecture

The figure 3.4 represents the architecture diagram of speech recognition and above procedures are done in Raspberry pi with speech recognition API.

E. ROS (Robot Operating System)

ROS is an open source robotics middleware suite. Although ROS is not an operating system but a collection of software frameworks for robot software development, it provides services designed for a heterogeneous computer cluster such as hardware abstraction, low-level device control, implementation of commonly used functionality, message-passing between processes, and package management. Running sets of ROS-based processes are represented in a graph architecture where processing takes place in nodes that may receive, post and multiplex sensor data, control, state, planning, actuator, and other messages.

The primary goal of ROS is to support code reuse in robotics research and development so you can find a built-in package system. Again, keep in mind that ROS is not an OS, a library, or an RTOS. It's a framework using the concept of an OS.

1) *SLAM:* The term SLAM is as stated an acronym for Simultaneous Localization And Mapping. It was originally developed by Hugh Durrant-Whyte and John J. Leonard based on earlier work by Smith, Self and Cheeseman [6]. Durrant-

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

Whyte and Leonard originally termed it SMAL but it was later changed to give a better impact. SLAM is concerned with the problem of building a map of an unknown environment by a mobile robot while at the same time navigating the environment using the map. SLAM consists of multiple parts; Landmark extraction, data association, state estimation, state update and landmark update. There are many ways to solve each of the smaller parts. We will be showing examples for each part. This also means that some of the parts can be replaced by a new way of doing this. As an example we will solve the landmark extraction problem in two different ways and comment on the methods. The idea is that you can use our implementation and extend it by using your own novel approach to these algorithms. We have decided to focus on a mobile robot in an indoor environment. You may choose to change some of these algorithms so that it can be for example used in a different environment. SLAM is applicable for both 2D and 3D motion. We will only be considering 2D motion.

2) *Mapping:* This package contains GMapping, from OpenSlam, and a ROS wrapper. The gmapping package provides laser-based SLAM (Simultaneous Localization and Mapping), as a ROS node called slam_gmapping. Using slam_gmapping, you can create a 2-D occupancy grid map (like a building floor plan) from laser and TF transforms from odom->base link(pose_data). Without the map, data association becomes much more difficult, namely the unknown mapping between landmarks and observations when the robot pose estimation is prone to uncertainty. We're currently stepping into a new robust-perception age, where performance with a low failure rate, a high-level understanding of the dynamic environments, the flexibility to adjust the computation load depending on the sensing and computational resources, and task-driven perception are required.
3) *Localization:* Localisation is the process of estimation of the position of the robot in a given environment. In this case, localisation specifically refers to estimation of the position and orientation of a robot within the previously generated map. A relatively straightforward way of doing this is using 'Dead Reckoning' , which constitutes estimating the robot's pose by calculating the distance travelled from an initial set-point with the help of odometry data. Robot localization requires sensory information regarding the position and orientation of the robot within the built map. Each method involves some major limitations, so proper sensor fusion techniques have been deployed to overcome the constraints of each sensor alone.
4) *Odometry:* The general definition of odometry is the use of data from distance sensors to estimate change in position over time. Odometry is used by our robot to estimate its position and orientation relative to a starting location given in terms of an x and y position and an orientation around the z (upward) axis. The topic is/odom and the command to view the form of the /odom message. The current position of the robot is taken from the map which is created using distance sensors.

F. OpenCV

OpenCV is a cross-platform library using which we can develop real-time computer vision applications. It mainly focuses on image processing, video capture and analysis including features like face detection and object detection. In this tutorial, we explain how you can use OpenCV in your applications. Using OpenCV, a BSD licensed library, developers can access many advanced computer vision algorithms used for image and video processing in 2D and 3D as part of their programs. The algorithms are otherwise only found in high-end image and video processing software.

1) *Machine Learning:* Machine learning is the science of getting computers to act without being explicitly programmed. In the past decade, machine learning has given us self-driving cars, practical speech recognition, effective web search, and a vastly improved understanding of the human genome. Machine learning is so pervasive today that you probably use it dozens of times a day without knowing it. Many researchers also think it is the best way to make progress towards human-level AI. Machine learning enables analysis of massive quantities of data. While it generally delivers faster, more accurate results in order to identify profitable opportunities or dangerous risks, it may also require additional time and resources to train it properly. Combining machine learning with AI and cognitive technologies can make it even more effective in processing large volumes of information.
2) *Deep Learning:* Deep learning is a particularly good example in this regard: It's related to – but not interchangeable with – the broader category of machine learning. This exacerbates the possibility for misnomers and misunderstandings. In fact, there's a Russian doll analogy here: Deep learning sits inside of machine learning, which sits inside of artificial intelligence. Deep learning is almost similar to that, but it makes more accurate decisions on its own by collecting various information about an object. It has many layers of analysis and takes a decision according to it. To fasten the process, it uses Neural Network and provides us with more exact results that we needed (means better prediction than ML). Something like how a human brain thinks and makes decisions.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

G. Hardware Component

1) *Raspberry Pi:* The Raspberry Pi is a low cost, credit-card sized computer that plugs into a computer monitor or TV, and uses a standard keyboard and mouse. It is a capable little device that enables people of all ages to explore computing, and to learn how to program in languages like Scratch and Python. It's capable of doing everything you'd expect a desktop computer to do, from browsing the internet and playing high-definition video, to making spreadsheets, word-processing, but it also provides a set of GPIO (general purpose input/output) pins that allow you to control electronic components for physical computing and explore the Internet of Things(IoT).

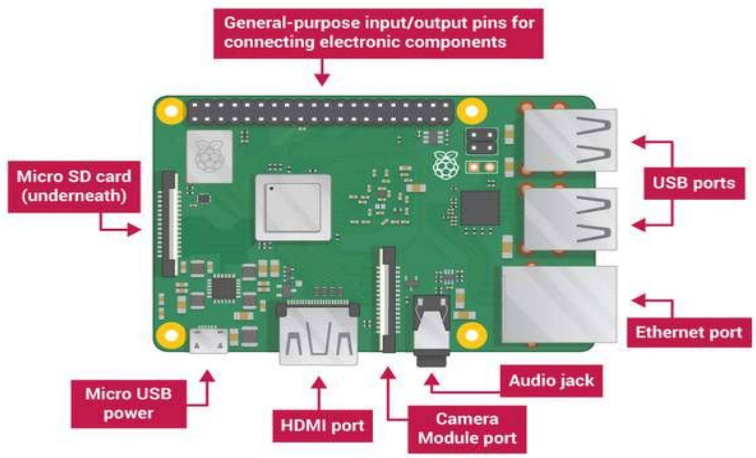


Figure 3.5 Raspberry pi

It is a low-cost module running Quad-core ARM (BCM2837) processor with 1GB RAM, 8/16/32GB eMMC (on CM3+ variant), bundled with a lot of IOs on a 200-pin SODIMM edge connector. Block diagram is shown in figure 3.5.

2) *Arduino UNO:* Arduino is a platform for microcontroller devices that makes embedded programming much easier than traditional methods. Thanks to Arduino's simplicity and ease-of-use, embedded systems and programming now have a much lower barrier of entry than before. For only about $25 you can get started in electronics, as opposed to a few hundred dollars to buy evaluation boards and hardware programmers. The Arduino Uno is a microcontroller board based on the ATmega328. It has 14 digital input/output pins (of which 6 can be used as PWM outputs), 6 analog inputs, a 16 MHz crystal oscillator, a USB connection, a power jack, an ICSP header, and a reset button.

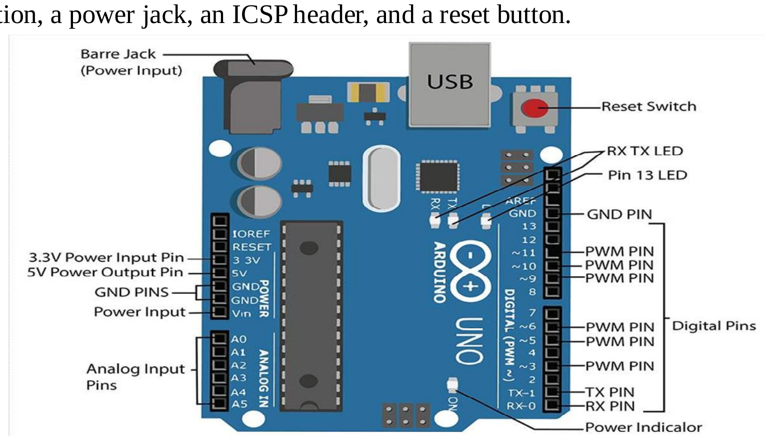


Figure 3.6 Arduino Uno

The most important advantage with Arduino is the programs can be directly loaded to the device without requiring any hardware programmer to burn the program. This is done because of the presence of the 0.5KB of Bootloader which allows the program to be burned into the circuit. All we have to do is download the Arduino software and write the code. Model and parts of Arduino Uno and its parts are shown in figure 3.6.

3) *Motor Driver:* This L298N Motor Driver Module is a high power motor driver module for driving DC and Stepper Motors. This module consists of an L298 motor driver IC and a 78M05 5V regulator. L298N Module can control up to 4 DC motors, or 2 DC motors with directional and speed control.

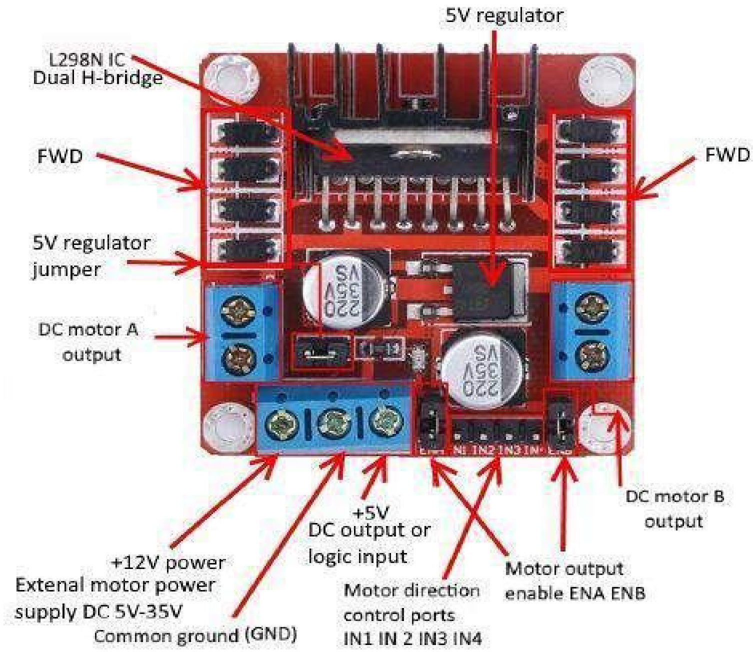


Figure 3.7 Motor Driver

The module can be applied to drive DC motors. Since the module uses a dual H-bridge drive, it can drive two motors at the same time. Parts of the motor driver are shown in figure 3.7.

4) *DC Motor:* A DC motor is any motor within a class of electrical machines whereby direct current electrical power is converted into mechanical power.

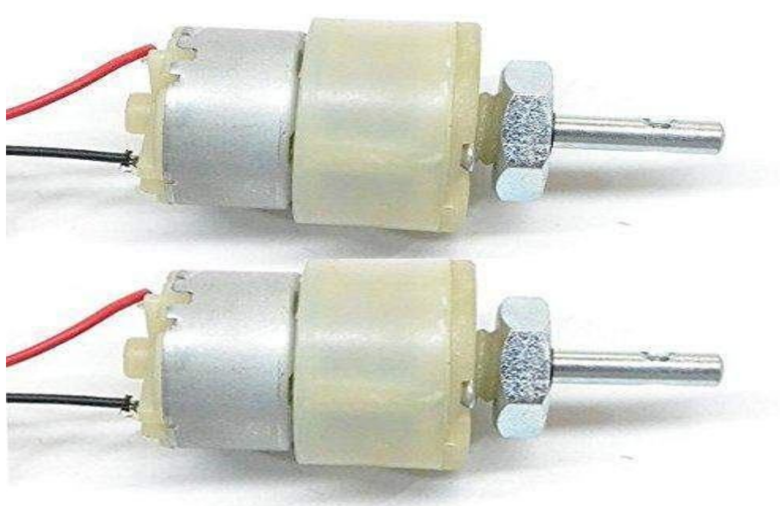


Figure 3.8 DC Motors

A 12v DC motor is small and inexpensive, yet powerful enough to be used for many applications. Image of DC motor is given in figure 3.8.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

5) Web Camera

Figure 3.9 Web Camera

A camera is an optical instrument used to capture an image. At their most basic, cameras are sealed boxes with a small hole that allows light in to capture an image on a light-sensitive surface (usually photographic film or a digital sensor). Cameras have various mechanisms to control how the light falls onto the light-sensitive surface. The image of camera is shown in figure 3.9

6) ESP32

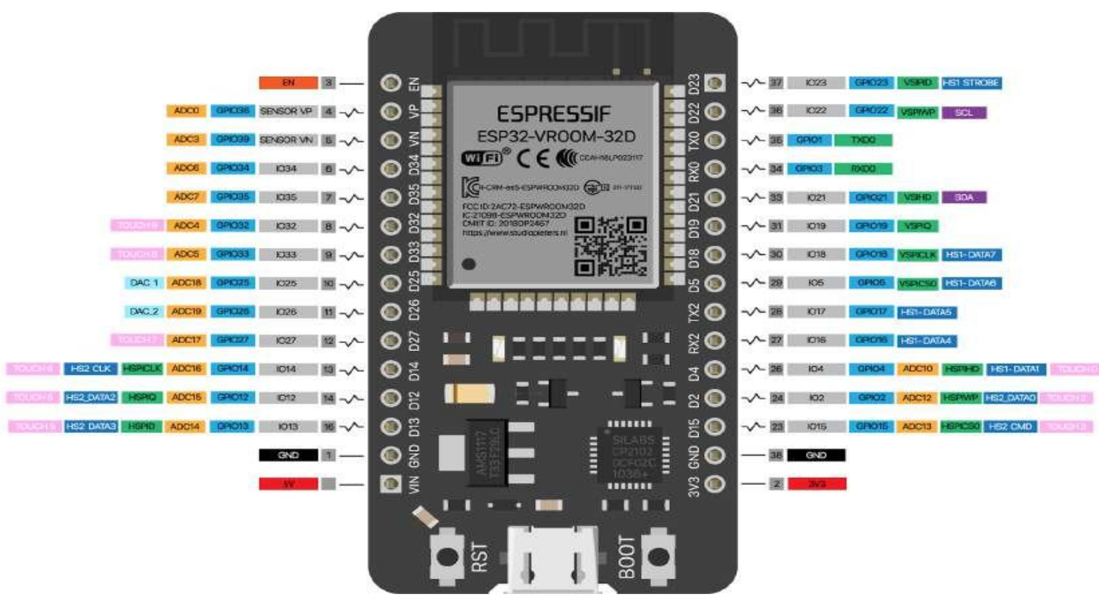


Figure 3.10 ESP32

ESP32 is a single 2.4 GHz Wi-Fi-and-Bluetooth combo chip designed with the ultra-low-power 40 nm technology. It is designed to achieve the best power and RF performance, showing robustness, versatility and reliability in a wide variety of applications and power scenarios. It is mainly used for IOT purposes. ESP8266 has a builtin processor. However, due to multitasking involved in updating the WiFi stack, most of the applications use a separate microcontroller for data processing, interfacing sensors and digital Input Output. The image and parts of ESP32 is shown in figure 3.10.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

7) Relay

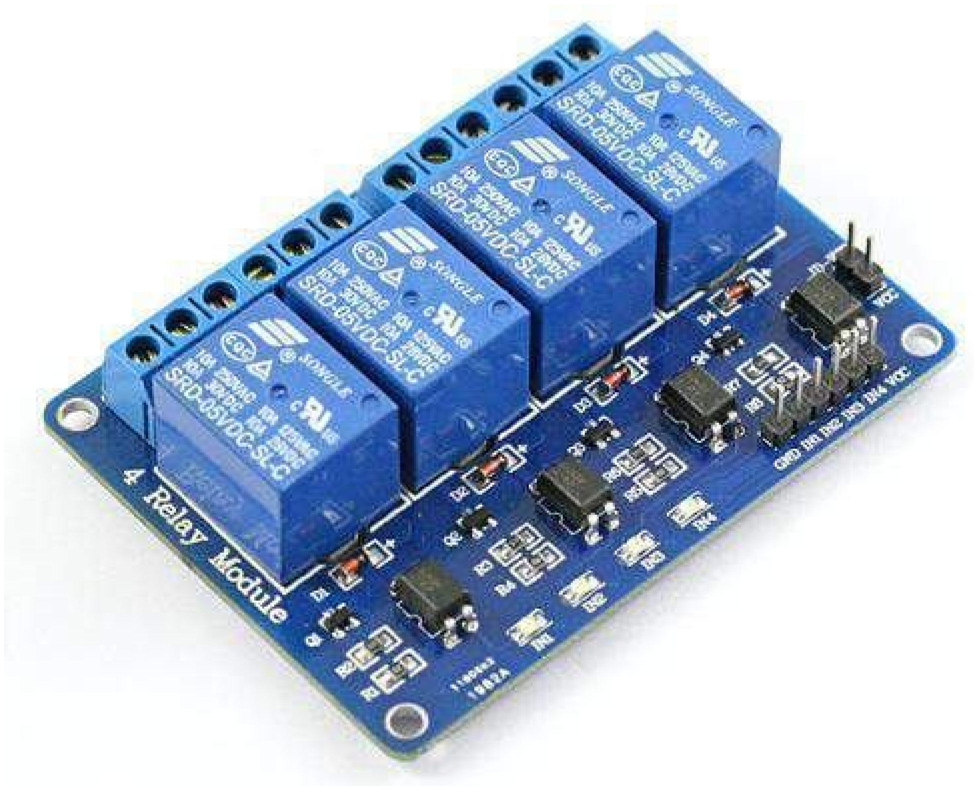


Figure 3.11 Relay

A relay is an electrically operated switch. It consists of a set of input terminals for a single or multiple control signals, and a set of operating contact terminals. The switch may have any number of contacts in multiple contact forms, such as make contacts, break contacts, or combinations thereof. The image of relay is shown in figure 3.11.

8) *12V Battery:* A 12v Battery is used. Which can power the robot upto 20 hours. The image of the battery is shown in figure 3.12.

Figure 3.12 12v Battery

9) *Wheels:* Wheels of 66mm diameter are used in the robot. Both the wheels are used on the rear side of the robot. The image of wheels are shown in figure 3.13.

Figure 3.13 Wheels

F. *Hardware Design*

The below figure 3.13 represents the hardware design of the proposed system. The Raspberry pi acts as the brain of this system, since it controls all the remaining hardware. Arduino is used to control motors using a motor driver. The Raspberry pi gets input from ROS and gives output to Arduino through USB cable.

Figure 3.13 Hardware design

In our hardware design, raspberry pi is connected with arduino through USB cable. Three megapixel web camera and ultrasonic sensor of 4 meter range is connected as input devices. The arduino is connected with the L298N motor driver which is used to drive two 12V dc motors. We have used a 12V battery to power up our robot. The motor driver is connected with 8, 9, 10, 11 digital pins of arduino. The ultrasonic sensor is connected with GPIO 27, GPIO 31 pins of raspberry pi.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

IV. RESULTS AND DISCUSSIONS

A. ROS (Robot Operating System)

We have used a Robot Operating System for mapping, localisation and navigation of the robot in an unknown environment.

1) *Mapping:* In ROS mapping the data taken from the distance sensor is given as input. The ROS mapping system runs an algorithm to plot the data on the graph related to odometry of the robot.

Figure 4.1 Environment

The environment setup is shown in figure 4.1. The robot is placed in this environment. The datas of distance between the ultrasonic sensor and the obstacles are taken as input data by ROS. The input datas are processed by the SLAM algorithm. The processed data are converted into key points and plotted in a graph. this. This algorithm also tracks the robot odometry values. The distance sensor value and odometry value together form a map of the environment.

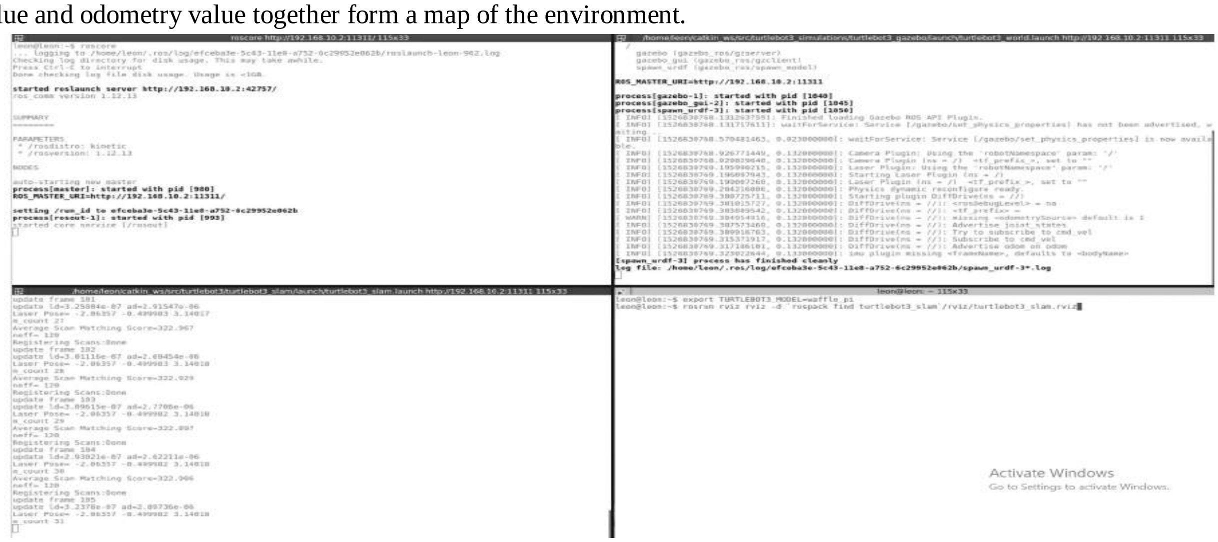


Figure 4.2 Terminal Output

The figure 4.2 shows the command prompt output for the given values from distance sensors and the change in position of the robot with odometry values. There are four terminals run at a time. In which one terminal shows the value of distance sensor input, the other terminal shows the current position, rotation and movement robot wheels, The other two terminals are used to initiate and give real time commands to the robot.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

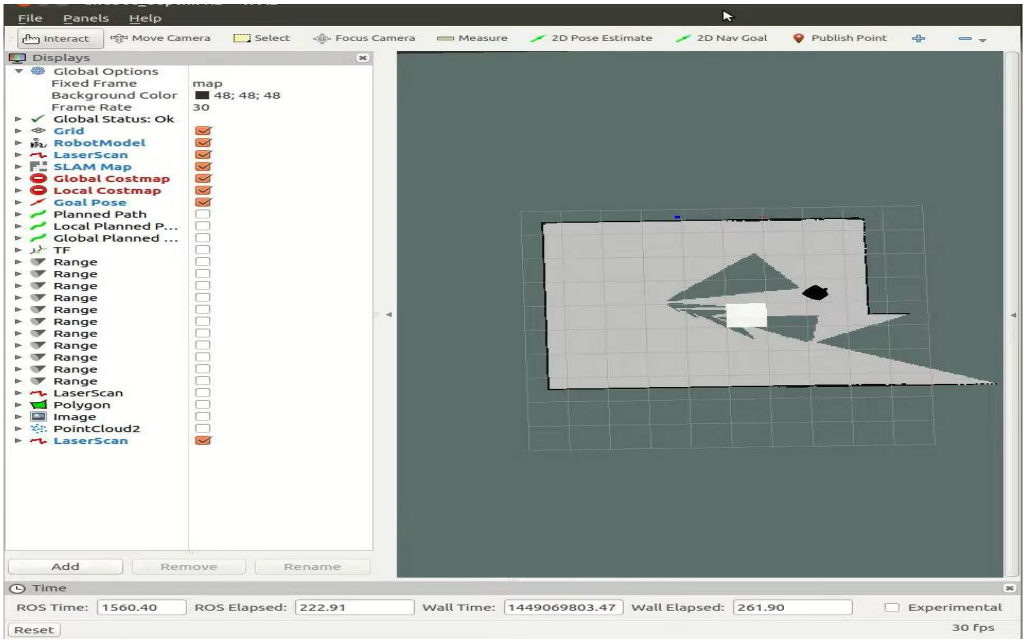


Figure 4.3 Mapping Output

The figure 4.3 shows the mapping of the environment. In figure 4.3 black lines represent obstacles, grey surfaces represent free space and dark grey represent unknown environments, dark black dots represent the current position of the robot. The left side window shows the components and instruments used in the robot. The bottom window shows the ROS time and wall time in milliseconds.

2) *Localization:* After mapping, the robot uses the map as the environment data and localize through the environment. ROS can help with keeping track of coordinate frames over time. Package for it is tf2 - the transform library, it comes with a specific message type: tf/Transform and it is always bound to one topic: /tf. Message tf/Transform consists of transformation (translation and rotation) between two coordinate frames, names of both frames and timestamp.

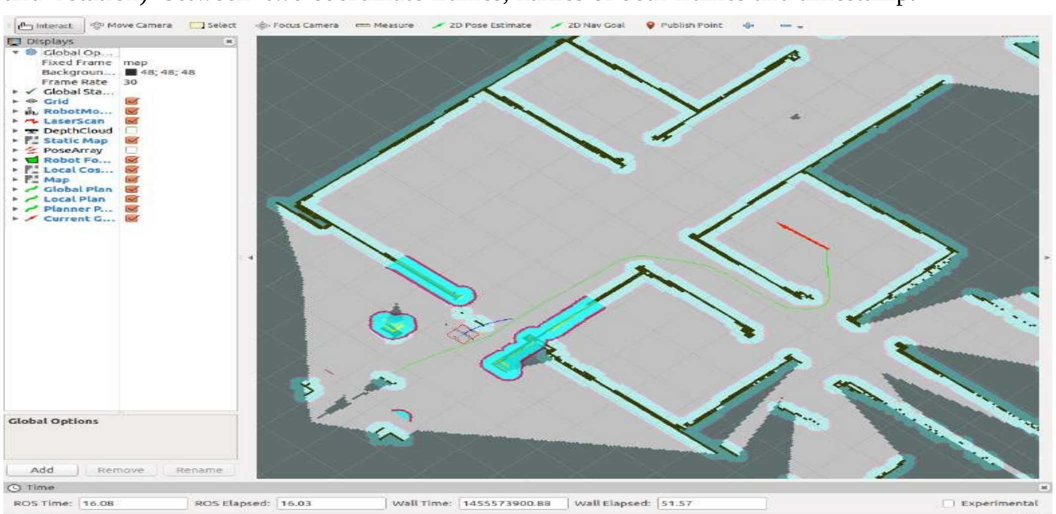


Figure 4.4 Path Planning

The figure 4.4 shows the path planned by the robot to reach the destination. In figure 4.4 red arrow represents the goal or destination of the robot. The green line represents the path planned by the ROS algorithm. This value coordinates with the wheel odometry to reach the destination.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

B. Object Detection

The camera collects images and transfers them to the raspberry pi. The object detection algorithm analyzes each frame and gives the name of the object with the help of object detection API.

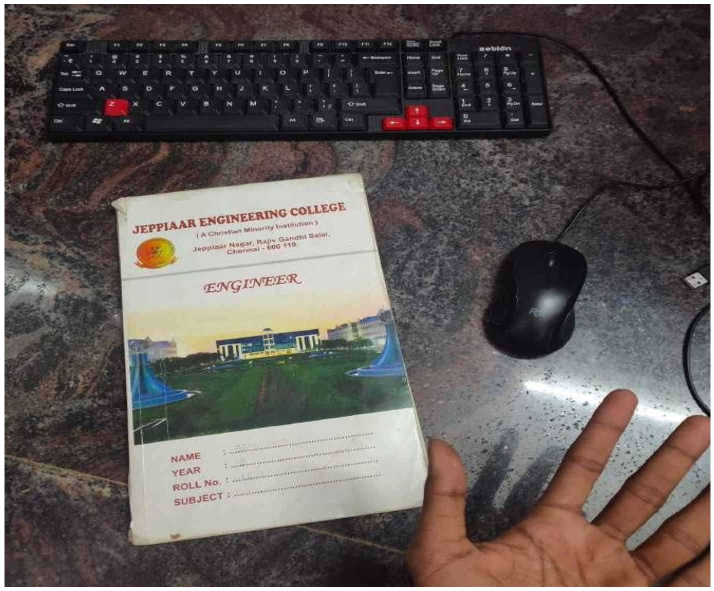


Figure 4.5 Input Image(1)

The image figure 4.5 with some general things has been taken and given as input to the object detection. The openCV image processing algorithm creates keypoints on the image and compares the key points with already existing images and gives the name of the object. The more number of key points gives more accurate and efficient output. The deep learning results in accurate output at very less time.

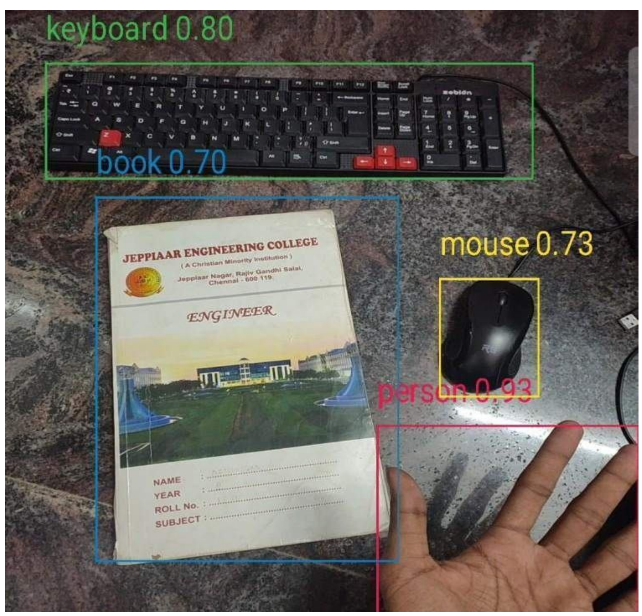


Figure 4.6 Output Image(1)

After processing it gives image figure 4.6 with the object's name marked with accuracy rate as output. In figure 4.6 the person is found by comparing skin color and shape of hand. Book, keyboard and mouse are found by comparing its shape, colors and pattern with per defined datas. The accuracy number given shows the percentage of key points matched in each object. The image with high quality and more colors gives high accurate outputs. Also we have tried with different image figure 4.7 whose output as shown in figure 4.8

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

Figure 4.7 Input Image(2)

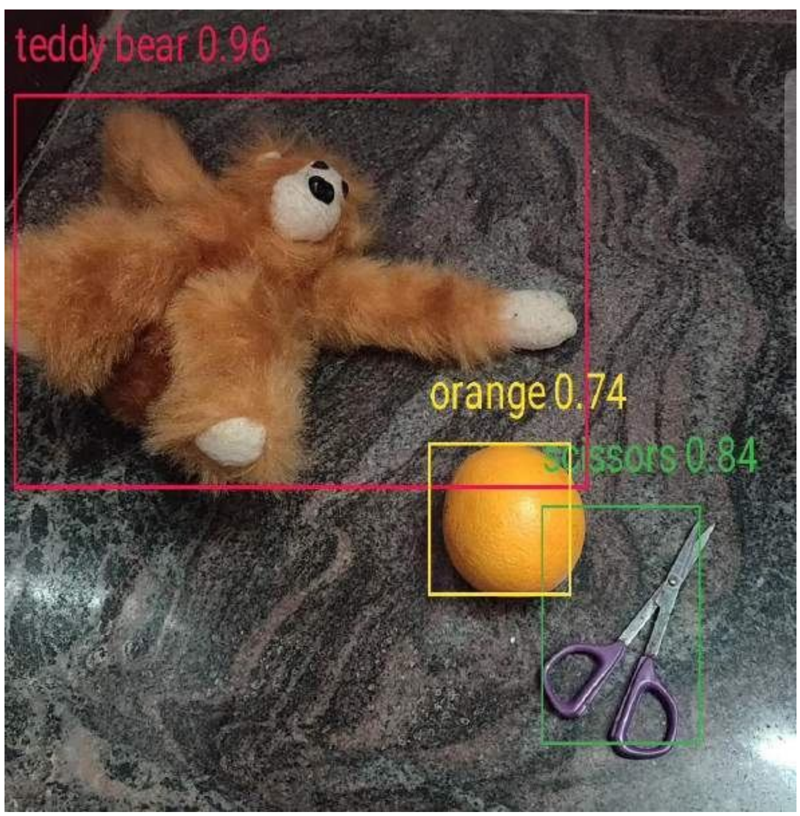


Figure 4.8 Output Image(2)

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue V May 2021- Available at www.ijraset.com

C. Speech Recognition

The human speech is given as input through a microphone to the raspberry pi which runs a python speech recognition library.

Figure 4.8 Before Speech Recognition

The figure 4.8 shows off the state of fan before speech recognition and figure 4.8 shows on state of fan before speech recognition. Here we used ESP32 wifi / bluetooth module which is connected with raspberry pi via bluetooth. The ESP32 module is loaded with a program to turn on and off the relay with particular data comments received through bluetooth. The relay is connected to an electric fan. 220v is given as input in relay.

Figure 4.9 After Speech Recognition

The phrase "Turn on the fan" is given as input to raspberry pi using a microphone. The python speech recognition API processes the speech and converts it into text. The text is compared with predefined commands. If the command matches, it will be sent to the ESP32 module through bluetooth. ESP32 module compares the received data with existing commands and turns on the particular relay which will turn on the fan.

V. CONCLUSION AND FUTURE WORK

The purpose of implementing this system is to help humans using AI. In this system we have implemented a personal assistant mode which gets commands from humans and gives responses. This will help people in their education, work and many other activities in daily life. Our robot also has an IOT application through which home appliances can be controlled remotely. This will reduce power consumption and prevent people from electric accidents. The object detection we used can detect many objects in real time. In personal assistant mode our robot can search and find objects using object detection. The robot movement also can be controlled through voice recognition which will be more helpful for aged and diabled people. The SLAM algorithm maps the house environment and saves the map which will help for future navigation. The saved map is updated every time the robot moves.

In order to improve mapping and localization, more powerful distance sensors like lidar and 3d cameras can be used. The use of cameras will improve the environment data and it will give more detail about surrounding objects. To improve image detection high definition stereo cameras can be used. Also many advanced APIs like tensorflow and YOLO can be used. Also usage of computers with powerful processors can detect more images and objects at real time without any error. Voice recognition can be improved by more advanced speech recognition APIs. High range microphones can be used for recognising voice from distance with less noise.

VI. ACKNOWLEDGEMENT

We are very truly indebted to Late Col. Dr. Jeppiaar, M.A., B.L., Ph.D., Founder, Jeppiaar Educational Trust, Dr. M. Regeena Jeppiaar, B. Tech., M.B.A., Ph.D., our Chairman and Managing Director and our Principal Dr. V. Natarajan M.E., Ph.D., for their blessings and the facilities to carry out the project.

We would like to express our gratitude to Dr. Anita Titus M.E., Ph.D, Professor and Head of the Department. We thank our assistant project coordinator Mr. E. Sakthivel, M.E. We are thankful to our supervisor Mr. S.K.Binu Siva Singh, M.Tech., for giving valuable suggestions and guiding us in all stages of the project. We finally thank all the faculty members of the Department of Electronics and Communication Engineering for their constant support.

We would also like to thank our parents and friends for all their support and we thank the Almighty God for helping us complete this project, successfully.

REFERENCES

[1] Z. Tufekci, J.N. Gowdy, S. Gurbuz, E. Patterson, Applied mel-frequency discrete wavelet coefficients and parallel model compensation for noise-robust speech recognition. Speech Commun. 48(10), 1294–1307 (2006)

[2] Z. Xueying, J. Zhiping, Speech recognition based on auditory wavelet packet filter, in Proceedings of 7th International Conference on Signal Processing, 2004 (ICSP '04), vol. 1, 2004, pp. 695–698

[3] O. Farooq, S. Datta, Wavelet-based denoising for robust feature extraction for speech recognition. Electron. Lett. 39(1), 163–165 (2003)

[4] O. Farooq, S. Datta, Mel filter-like admissible wavelet packet structure for speech recognition. IEEE Sig. Process. Lett. 8(7), 196–198 (2001)

[5] Ryan, M.R.: Exploiting subgraph structure in multi-robot path planning. Journal of Artificial Intelligence Research 31, 497–542 (2008)

[6] Dresner, K., Stone, P.: Multiagent traffic management: a reservation-based intersection control mechanism. In: Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems, pp.530–537 (2004)

[7] Wang, K.H.C., Botea, A.: Fast and memory-efficient multi-agent pathfinding. In: International Conference on Automated Planning and Scheduling (ICAPS), pp. 380–387 (2008)

[8] Parker, L.E.: Current state of the art in distributed autonomous mobile robotics. In: Distributed Autonomous Robotic Systems 4, pp. 3–12. Springer (2000)

[9] Van Den Berg, J.P., Overmars, M.H.: Prioritized motion planning for multiple robots. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 430–435 (2005)

[10] Desaraju, V.R., How, J.P.: Decentralized path planning for multi-agent teams with complex constraints. Autonomous Robots 32, 385–403 (2012)

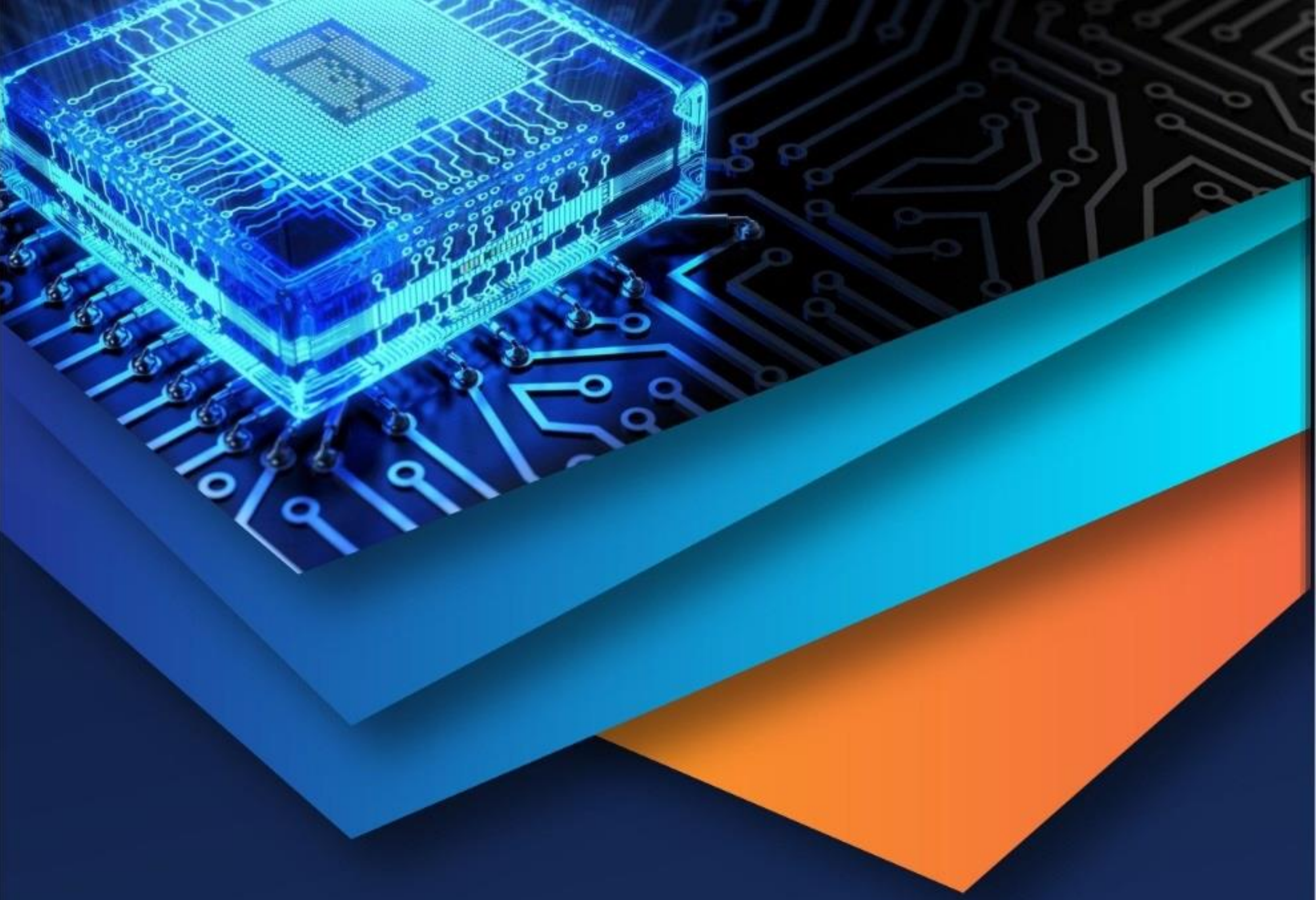
[11] Angeline, I. and E.Z., F. (2016). Image Contrast Enhancement Techniques: A Comparative Study of Performance. International Journal of Computer Applications, 137(13), pp.43-48.

[12] Dalal, N., and Triggs, B. (2005). Histograms of oriented gradients for human detection, in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEEComputer Society Conference on, Vol. 1 (San Diego, CA: IEEE).

[13] Herbert Bay, Andreas Ess, TinneTuytelaars, and LucVan Gool. Surf: Speeded up robust features. ComputerVision and Image Understanding (CVIU), 110(3):346–359.

[14] McGraw-Hill Inc.'s BYTE, the Magazine of Technology Integration, published a two-page review of IBM's Personal Dictation System software. May 1994; Volume 2 Pages 145-146

[15] Yifan Xia, Jie Li, Lin Qi, and Hao Fan. Loop closure detection for visual slam using pcanet features. (IJCNN), June 2016, Vol 1 Pages 56-75



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY