

## Interactive Robot with Image Classification Techniques

**StiveHassan, Sridhar J, Mohan CR, M Supreeth, Sathyaprakash**

*School of Electronics and Communication Engineering, REVA University, India*

### ABSTRACT

*We are introduced to the world of automation, where everything is automatic and very accurate. The technology has replaced even the human presence one such example was the project Sophia. So, our team had thought to work out on a project that could be actually interact like a human. This is how the project Buggy came in to light. Buggy can interact with the human normally as other human do. Buggy takes and collect data at background read a detail study about the data and self-learn with the data collected. The webcam mounted on the top of the rover acts as an eye of the raspberry that can see around the environment and sense it's happening. This is a complete domestic assistance system that can be used as a surveillance system. This is an interactive bot which we can have a good time to chat with and the movement of the bot is also voice activated. There is a lot of future scope of such robot that reduce the workload on human beings. Buggy can right now classify 7 types of dance, 7 types of mobile company and few animals. The classification takes place on the amount of data fed to the system. Buggy uses two modules namely dialog flow and tensorflow which is a strong platform provided by google for machine learning developers, we are using dialog flow for the interactive bot and tensorflow for classification purpose.*

**Keywords -** Dialogflow, Image Classifier, Tensorflow, Voice Controlled Robot

### INTRODUCTION

The concept of the conversation is central to a chatbot. A chatbot can and does converse with human but as mentioned previously its capabilities are limited. That is not to say however that in very narrow ways the text or voice-based conversation can be as good or better than conversing with an actual human. Chatbots can have advantages over human agents. They are available 24/7 and they have access to a very broad array of information and functionality. They can also outperform humans in terms of the speed and accuracy in a narrow domain. The problem however is making sure the end users are aware of these limitations.

While chatbots have the capability to replace humans for certain tasks, they also can be used to augment what human agents can offer their clients. The chatbot can for example provide suggested responses for the human agent or bring up relevant information in a timely manner which the human agent can then act on. The fact that chatbots are used directly in a communication channel means that the collaboration between the bot and a human agent is far easier to achieve.

## II.METHODOLOGY

### 2.1 Introduction

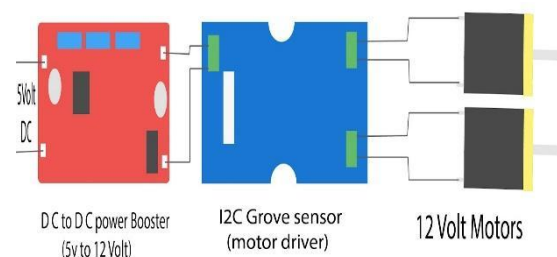
The working of the robot is divided into mainly three categories

- 1) Circuit Connection
- 2) Dialog Flow integration
- 3) TensorFlow Integration

### 2.2 Block Diagram

we are using CentIoT - CC and CV - Boost and Buck Voltage Converter to boost up the applied 5V from LiPo batteries to 12V because the motor that are used are of input voltage 12V and 500 Rpm. Any voltage less than 12V cause the motor to not function.

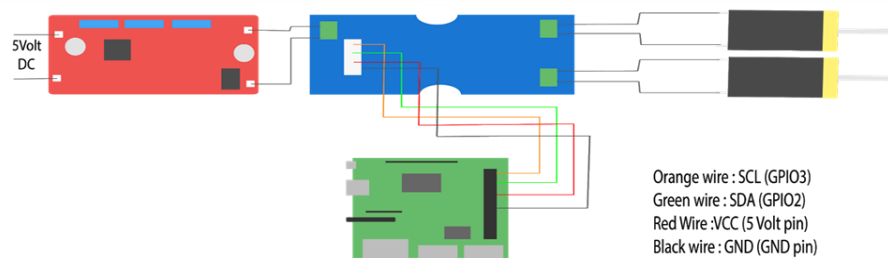
The I2C grove Sensor are used for interfacing the motor. The main reason for interfacing the I2C sensor are to reduce the number of wires to connect the motor and low speed micro-controllers. It mainly consists of 2 wires namely SDA (Serial Data) and SCL (Serial Clock). The connection of I2C circuit with the PI will be explained in the coming connection block.



**Fig 2.2.1 Connecting the motors to I2C grove sensor**

The Raspberry Pi has three types of serial interface on the GPIO header. You're probably already familiar with the UART serial port, which allows you to open a login session from a serial terminal application, such as PuTTY. The other two serial interfaces are the Serial Peripheral Interface ( **SPI** ) and Inter-Integrated-Circuit bus ( **I2C** ). SPI on the Pi allows for up to two attached devices, while I2C potentially allows for many devices, as long as their addresses don't conflict. Before the connection enable the I2C configuration in Raspberry PI by following procedure:

1. Run `sudo raspi-config`.
2. Use the down arrow to select 9 Advanced options.
3. Arrow down to A7 I2C.
4. Select yes when it asks you to enable SPI,
5. Also select yes when it asks about automatically loading the kernel module.
6. Use the right arrow to select the <Finish> button.
7. Select yes when it asks to reboot.

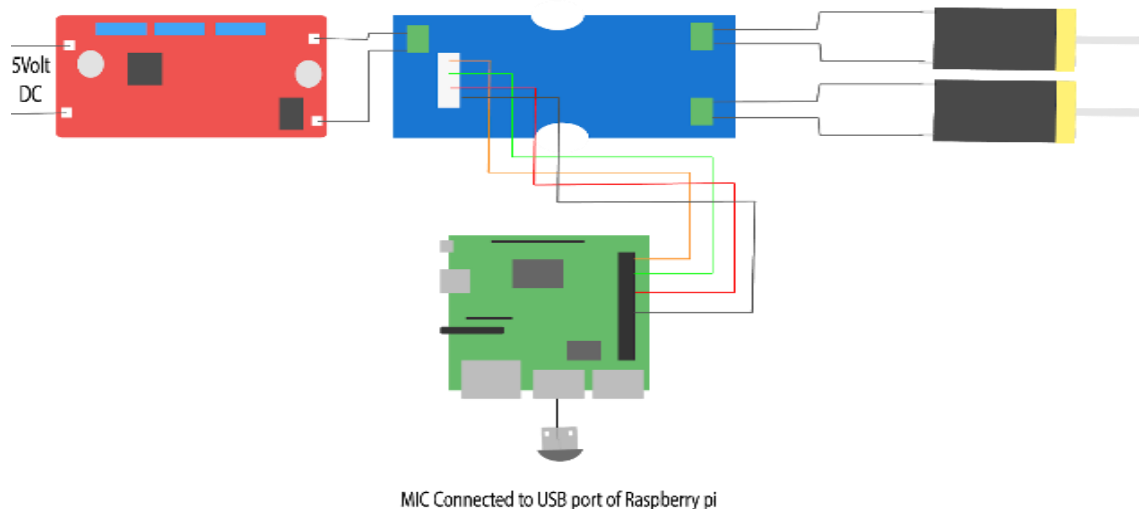


**Fig 2.2.2 Connecting I2C Grove to Raspberry Pi**

- 1.SCL pin to GPIO3 of Raspberry Pi.
- 2.SDA pin to GPIO2 of Raspberry Pi.
- 3.VCC pin to 5V pin of Raspberry Pi.
- 4.GND pin to GND pin of Raspberry Pi.

### 2.2.3 Mic Interfacing to Raspberry Pi 3

You need to check if the recording takes place properly. Use the command “arecord -l” to check if your mic/webcam is listed. Then, use the command “arecord -D plughw:1,0 test.wav” to record sound.

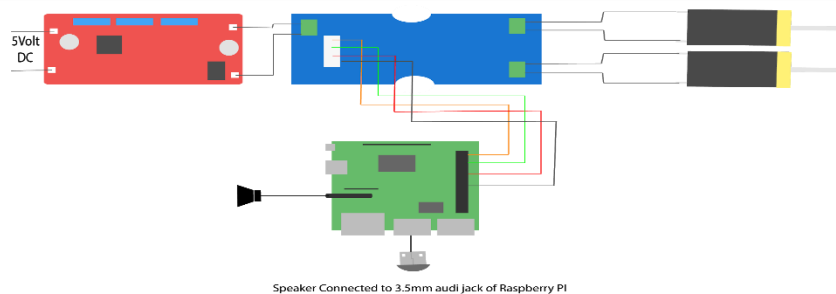


**fig 2.2.3.1 Connecting Microphone to Raspberry Pi**

The sound will be recorded in the file “test.wav”. To listen to it, plug in your headphones to the pi and enter the command “aplay test.wav” in the terminal. If you’re able to hear the sound, your microphones work perfectly.

### 2.2.4 Speaker Interfacing to Raspberry Pi

A Speaker is a device that converts analog audio signals into the equivalent air vibrations in order to make audible sound. When CRT monitors were the norm, speakers designed for computers were shielded to avoid magnetic interference with the CRT's magnetic coil.

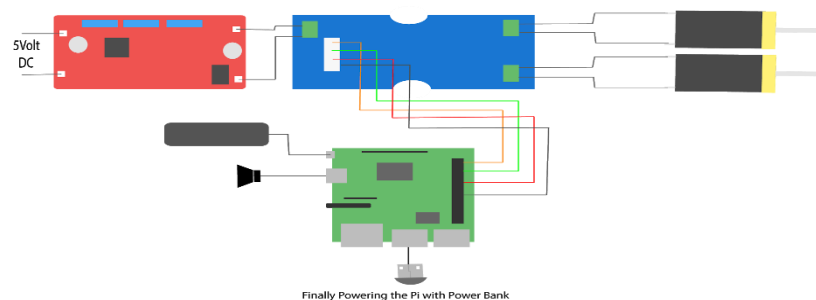


**fig 2.2.4.1 Connecting Speaker to Raspberry Pi**

Here Bluetooth speakers are used to hear the audio response from the PI. A normal AUX cable is used to interface the PI and Speaker. The AUX cable from the speaker is connected to the 3.5 mm female jack of the Raspberry PI.

### 2.2.5 Powering up the circuit

For this project we needed a power supply of 5V. We could use commercial power supply of similar ratings like a portable power supply. Portable Power Banks are comprised of a special battery in a special case with a special circuit to control power flow.



**fig 2.2.5.1 Connecting Portable Power Bank**

They allow you to store electrical energy (deposit it in the bank) and then later use it to charge up a mobile device (withdraw it from the bank). Power Banks have become increasingly popular as the battery life of our beloved phones, tablets and portable media players is outstripped by the amount of time we spend using them each day. By keeping a battery backup close by, you can top-up your device(s) while far from a wall outlet.

### 2.2.6 TensorFlow Integration

TensorFlow is an open source platform powered for machine learning initiated by Google developers.

To install TensorFlow in Raspberry, pi the following steps are to be followed:

1. Open PuTTY and get the Raspberry Pi connected using the SSH and enter the following command

“Pip install TensorFlow”

2. As TensorFlow is installed, a set of files and documents are needed to clone for further purpose. Enter the below command to the terminal -- “git clone <https://github.com/googlecode/tenorflow-for-poets-2>”



fig 2.2.6.1 TensorFlow with Raspberry Pi

3. Redirect into that particular library using CD command.
4. Download the Datasets from the Google images. Fatkun Batch downloader is the best software to download images at a larger chunk.
5. Paste these files into the tf\_files/flower\_photos
6. Train the Pictures to the model by entering the following command

```
python -m scripts.retrain \  
  --bottleneck_dir=tf_files/bottlenecks \  
  --how_many_training_steps=500 \  
  --model_dir=tf_files/models/ \  
  --summaries_dir=tf_files/training_summaries/"${ARCHITECTURE}" \  
  --output_graph=tf_files/retrained_graph.pb \  
  --output_labels=tf_files/retrained_labels.txt \  
  --architecture="${ARCHITECTURE}" \  
  --image_dir=tf_files/flower_photos
```

fig 2.2.6.2 Script for downloading pre-trained model

This script downloads the pre-trained model, adds a new final layer, and trains that layer on the flower photos you've downloaded the first retraining command iterates only 500 times. You can very likely get improved results (i.e. higher accuracy) by training for longer. To get this improvement, remove the parameter --how\_many\_training\_steps to use the default 4,000 iterations.

### 2.2.7 Dialog Flow Integration

Dialogflow is a Google-owned product. It was developed for human-computer interaction technologies based on natural language conversations. Dialogflow is a complete closed source product with API and web interface.



fig 2.2.7.1 DialogFlow with Raspberry Pi3

The client access token is a unique ID provided to each intent created by the user in dialog flow. This token acts as a key for the communication of the pi and google API.

#### 2.2.8 Code for Speech to Text

```
import speech_recognition as sr

def get_text():
    r = sr.Recognizer()
    with sr.Microphone() as source:
        print("Say something!")
        audio = r.listen(source)
    try:
        val = r.recognize_google(audio)
        return val
        #print("You said: " + r.recognize_google(audio))
    except sr.UnknownValueError:
        return 0
    except sr.RequestError as e:
        return 0
```

fig 2.2.8.1 Code for Voice Recognition

gTTS (*Google Text-to-Speech*) is a Python library and CLI tool that is used to interface with Google Translates text-to-speech API. which writes spoken mp3 data to a file, a file-like object (byte string) for further audio manipulation, or stdout.

Features include,

- Customizable speech-specific sentence tokenizer that allows for unlimited lengths of text to be read, all while keeping proper intonation, abbreviations, decimals and more;
- Customizable text pre-processors which can, for example, provide pronunciation corrections;

- Automatic retrieval of supported languages.

```
import cv2
from subprocess import call
def recog():
    vidcap = cv2.VideoCapture(0)
    success,image = vidcap.read()
    count = 0
    success = True
    cv2.imwrite("img.jpg", image)    # save frame as JPEG file
    success,image = vidcap.read()
    print('Read a new frame: ', success)
    call(["python -m scripts.label_image --graph=tf_files/retrained_graph.pb --image=tf_files/img.jpg"])
```

fig 2.2.9 Code for Image Recognition

OpenCV library are installed for computer vision to help in achieving to capture frame from the webcam and analyses the picture and pass on to the tensor flow for the image detection to finally pass out the results.

### III. HARDWARE IMPLEMENTATION

#### 3.1 Introduction

There is a large collection of hardware used in implementing the project, the functional brain of the circuits is Raspberry pi that controls the motors which is connected through I2C which is in turn powered by DC to DC power booster, Connects with the tensorflow and dialog flow api.

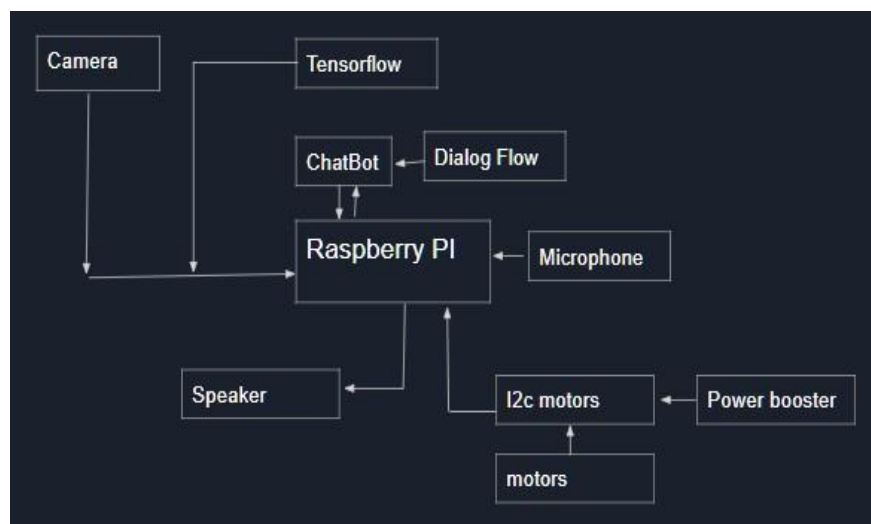


fig 3.1. Block Diagram of Data Flow and Architecture.



### 3.1.1 Raspberry Pi 3 :

CPU	ARM11 @700 MHz
RAM	512 MB SDRAM
Storage	micro-SD
Ethernet	10/100
Wireless	none
GPIO	40
Video Output	HDMI/Composite
WIFI	Onboard Wi-Fi

**Table 1. Raspberry Pi 3 Specification**

### 3.1.2 I2C Grove Sensor

The I2C bus is a simple and flexible way to transfer digital data between two electronic devices which may be physically separated or contained on the same printed circuit board (PCB). The main processor chip on the Quadrino board communicates with the accelerometer, gyro, mag and baro sensors using this I2C bus.



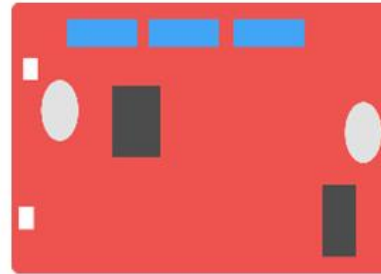
**Fig 3.1.2 Grove Sensor**

Only needs 2 wires to connect up to 255 devices. The two wires are called CLOCK and DATA. These are purely signal wires and provide no appreciable level of power. Each time a pulse occurs on the clock pin, the digital state (HI or LOW) of the DATA pin is taken. This would represent a single bit of the data transfer, a 1 or a 0. Data is transferred 8 bits at a time, so there will be 8 pulses on the clock line to transfer each byte.

### 3.1.3 Power Booster

A boost converter (step-up converter) is a DC-to-DC power converter that steps up voltage (while stepping down current) from its input (supply) to its output (load). It is a class of switched-mode power supply (SMPS) containing at least two semiconductors (a diode and a transistor) and at least one energy storage element:





**Fig 3.1.3 Power Booster and Computerized Picture**

a capacitor, inductor, or the two in combination. To reduce voltage ripple, filters made of capacitors (sometimes in combination with inductors) are normally added to such a converter's output (load-side filter) and input (supply-side filter).

### 3.1.4 Webcam

webcams work in broadly the same way, they use an image sensor chip to catch moving images and convert them into streams of digits that are uploaded over the Internet. The image sensor chip is the heart of a webcam.



**fig 3.1.4 Webcam**

## IV. CONCLUSION

A paragraph to summarize the concept of project and the implications (results),

This technology can be used as a home Surveillance system that can be used as a monitoring system. It can also be used in defense purpose after the required datasets are trained to the model. This model almost finds its application in all domain when the datasets pertaining to that domain is trained to the model.

challenges that need to overcome on for the model's smooth operation: Requires a very high internet to function, Constant maintenance, Datasets should be uploaded manually to recognize new objects.

## REFERENCES

- [1] Alexander Reben, Blabdroid, A team of miniature robots developed at MIT makes a documentary about humans, disarming their subjects with their engineered "cuteness.", MIT Open Documentary lab, \_docubase, US, 2013.

- [2]Andrey Victor Justo; Julio Cesar dos Reis; Ivo Calado; Rodrigo Bonacin; Felipe Rodrigues Jensen, Exploring Ontologies to Improve the Empathy of Interactive Bots, 2018 IEEE 27th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE).
- [3] M. Hans; B. Graf; R.D. Schraft Proceedings,Robotic home assistant Care-O-bot past-present-future, 11th IEEE International Workshop on Robot and Human interactive Communication,2002.
- [4]Peter Goldsborough, FakultätfürInformatik, A Tour of TensorFlow, Proseminar Data Mining, Technische Universität München, Germany, Oct 2016.
- [5]A Tool of Conversation: Chatbot, M. Dahiya,International Journal of Computer Sciences and Engineering, Volume-5, Issue-5, May 2017.
- [6] Miss. Renuka P. Kondekar,Prof. A. O. Mulani, Raspberrypi-based voice operated Robot, International Journal of Recent Engineering Research and Development (IJRERD) ISSN: 2455-8761Volume 02 – Issue 12 || December 2017
- [7]K. He, X. Zhang, S. Ren, and J. Sun. Delving deep into rectifiers: Surpassinghuman-level performance on imagenet classification. In ICCV, 2015.
- [8]Dan Jurafsky, Conversational Agents, Stanford University 124/LINGUIST 180, From Languages to Information ,2019.