



iJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 Issue: IV Month of publication: April 2018

DOI: <http://doi.org/10.22214/ijraset.2018.4607>

www.ijraset.com

Call: ☎ 08813907089

E-mail ID: ijraset@gmail.com

ReeMiE – Intelligent Personal Assistant

Siddharth Gupta¹, Vipin Tiwari², Abhishek Rajak³, Dr. Ratish Agrawal⁴, Dr. Sachin Goyal⁵

^{1, 2, 3, 4, 5}Department of Information Technology

Abstract: *The purpose of this paper is to demonstrate the model of intelligent virtual assistant. It enables smart voice two way interactions for the purpose of browsing favorite recipes, playing news, listening music and controlling various IOT devices. It works in a secure way by using primary voice and face analysis techniques. Primary processing of user image and voice is done through Raspberry pi model b+ in which algorithms of text to speech and speech to text is implemented. This model can help people with blurred or no vision to interact with the cyber world through voice. It also provides a secure way for home automation as user can control the person who can interact with the assistant. It remembers your choices by linking with your Gmail account to provide global access to your favorite music and history.*

Keywords: *Reemie, Intelligent Personal Assistant, voice interaction, Voice command system, query processing engine, Raspberry pi*

I. INTRODUCTION

During day to day activities many people often forget to turn off the lights or water there plants thus it becomes challenging for them to keep manage day to day work. Also it is a challenge to get the answers for all you queries and questions at a single place. A personal assistant take care of all the devices around you, it could answers all your queries intelligently or by searching on the internet or to play your favorite song or video over the internet [1]. A personal assistant is a system that process voice response through natural language processing in the case English language and gives an appropriate voice output. Since human brain is more receptive to voice than to written text hence personal assistant are gaining wide popularity and is subject of major innovations all across the world [2]. As of 2017, the usage of personal assistants is expanding rapidly, with new products entering the market. The most popular in the market are Apple's Siri (34%), Google Assistant (19%), Amazon Alexa (6%), and Microsoft Cortana (4%). Apple and Google have huge installed bases of users on smartphones and Microsoft has wide numbers of Windows-based personal computers which uses the services of its personal assistant Cortana, Amazon Alexa was the first to get the ability to place e-commerce orders. The model uses natural language processing techniques to read the user input and query processing after matching it with certain available keywords [3] [4]. The rest of the paper is divided into 4 sections. Section 2 System Architecture describes the drawbacks in existing system and the proposed system to overcome them. Section 3 provides information regarding different modules implemented in the model. Section 4 provide the details of the final test and the results obtained after it. Section 5 provide conclusion and future applications of the proposed system.

II. SYSTEM ARCHITECTURE

A. Existing system

The existing system has major security drawbacks, it allows no unique identification mechanism which exposes your personal choices and browsing history to everyone in command to use the assistant just through a call of wake word. In current system no global access of your previous searches is present as it stores it in local storage .It is totally dependent on a single query processing engine which if may go down could put to failure the whole system [5] [6] [7].

B. Proposed system

The proposed system is such that it efficiently overcomes the drawback of the existing system by providing a biometric based unique identification system, separating the query processing in 2 stages and linking your searches with the Gmail account for global access There are two functional components in this project. They are Face recognition and response. Raspberry pie is programmed using OpenCv an open source Python based library for face recognition. To enable natural language processing and intelligent response we have used google assistant API available on [8]. Snowboy api available on [9]. Microprocessor is the brain of the system. Several devices connected to the controller's input pin. These devices can be controlled using commands set up externally on raspberry pi. The overall activity is reported in the form of microphone and speaker.

Main Parts of the model is

- 1) *Face Recognition*: This will be achieved by Opencv. Which is installed inside the raspbian. It is python based library which allows training and recognition of data sets of faces using built in library
- 2) *Intelligent Natural Language Processing*: On receiving queries, microphones will notify the pi by Searching for the catch phrase “hey ReeMie” or “Alexa”. The project uses an Logitech webcam with built in mic in combination with relay control switch to control the devices and overall functioning. Motor may be driven by external 9 volt battery with interfacing to microcontroller

C Hardware Implementation

1) *Raspberry pie 3*: The Raspberry Pi 3 Model B builds upon the features of its predecessors with a new, faster processor on board to increase its speed. It also features WiFi and Bluetooth Low Energy capabilities to enhance the functionality and the ability to power more powerful devices over the USB ports. It is based on Quad Core 1.2GHz Broadcom BCM2837 64bit CPU with 1 GB RAM , Wifi and Bluetooth onboard. It has 40 extended GPIO pins and 2 usb ports.



Fig.1 Raspberry Pi

2) *Logitech Web Camera and Inbuilt Microphone*: This Logitech webcam has a widescreen HD feature which enables you to record visuals in better quality. The 16:9 ratio enhances image detailing and offers effective wide screen technology. The 3 MP camera of this device helps in recording high quality visuals with better clarity.

With this webcam, you don't have to install an external microphone as it comes pre-included with one. This powerful microphone records clear audio and facilitates crisp sound delivery. The noise reducing capacity of this webcam allows you to speak effectively even when there is background noise. This webcam is equipped with effective light correction feature as well. This feature automatically adjusts the visual lighting and ensures brighter visuals even during dark conditions.



Fig. 2: Web camera

3) *The Relay Module*: Relay is an electrically operated switch. Many relays for switching solenoid mechanism mechanically operated, but can also be used for other principles of operation. Relays are widely used in early computers to telephones and perform logical operations.



Fig 3. Relay Module

4) *Speakers*: Speakers, once the response is collected by the assistant, using text to speech the output of that query is converted. Now this speech which is the audio output is sent to the user using the speakers which are running on audio output

D. Flow diagram

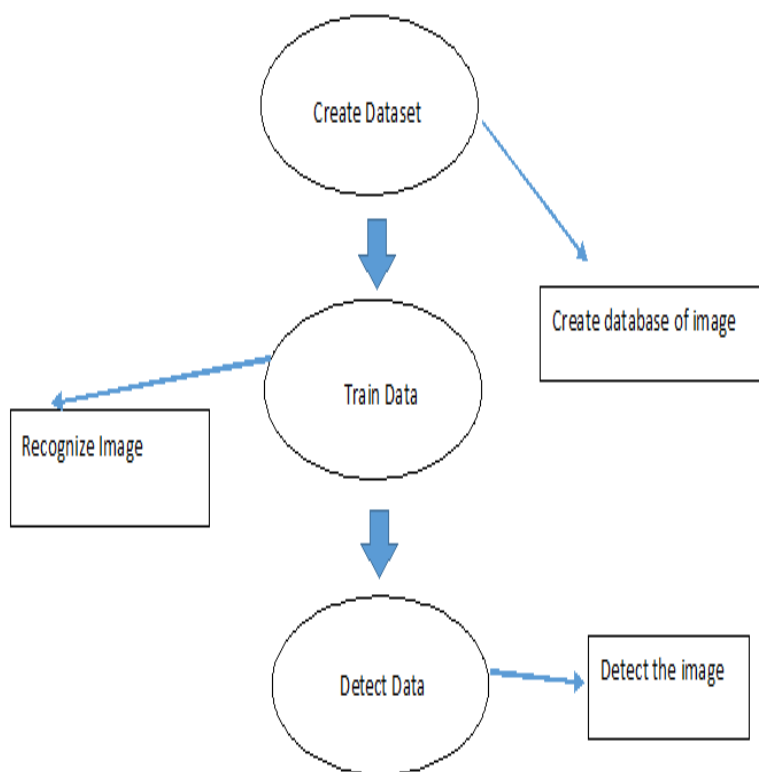


Fig 4: Flow diagram of face recognition

First create a dataset of the images of the individuals, convert the images to grayscale. Train the recognizer with these images to detect the user. As the user faces the web camera the recognizer predicts with a confidence value. If the confidence value is greater than 70%. Then the system is unlocked and assistant is started.

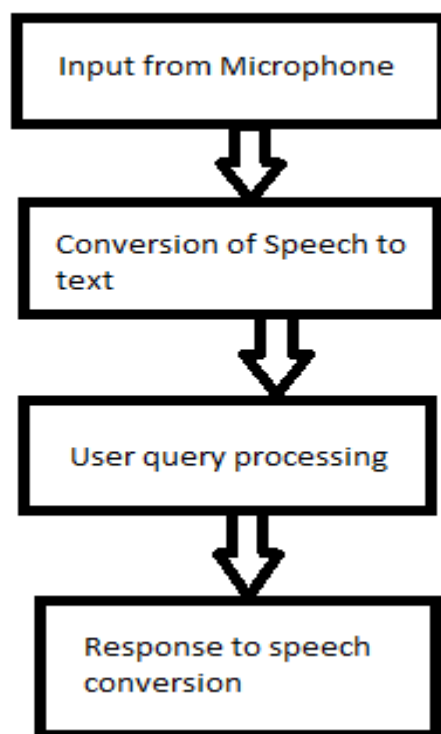


Fig 5: Flow diagram of voice command system

Voice Command system controls the working of assistant. User gives the input which is received through microphone in the raspberry pi. The api converts it to text and sends it to query analyzer. Query processor search for the keyword in dictionary and send it to the respective API.

III. MODULES IMPLEMENTED

A. Bio metric Security Using Face Recognition

Unique identification is provided using face recognition, it is implemented with open source computer vision library OpenCv. Datasets of the authorized users is created using 20 images processing in gray scale. Then database of all the images is stored in file which is used to predict the user with the confidence level of 80%. If the person is matched the local history is loaded in assistant and the assistant is triggered using the wake words. If the person is not found in the database then the assistant shuts down any further usage. The prediction uses LBPH face recognizer algorithm and its implementation in OpenCv [10].

B. Speech to Text Conversion

Google speech to text enables developers to convert audio to text using powerful neural networks implemented on high speed servers. To use google speech, application is created in the google developers console and the generated API key is used to access the speech engine. It requires OAuth authentication as data is sent over the Google servers.

C. Text to Speech Conversion

Google text to speech enables developers to convert text to speech using powerful machine learning techniques. To use google speech, application is created in the google developers console and the generated API key is used to access the speech engine. It requires OAuth authentication as data is sent over the Google servers.

D. Query Processing

The module Query Processing is used to fetch the relevant response for the user query. It is implemented using various open sources sdk and API like google assistant sdk, Snowboy sdk, Youtube api, Rss feed streaming, google play music.

The generated text from the user speech is searched for the dictionary words for the query like “play music from youtube”, “play from google”, “search”, “turn on lights” etc. the match query is send to the relevant api for the response

E. Voice Control of The Devices

This module controls the devices attach to the raspberry pi. It allows home automation and voice control of the pins. The default word for the devices is “trigger” followed by the name assigned to the gpio pin. If no device is attached to the gpio the response is not found device

F. Search

The module controls the searching of query on google and returns the response of the best result provided by search engine. It uses google assistant sdk which provides function to be implemented to use the services. Every transaction requires a basic authentication.

G. Youtube

The module controls playing the music through YouTube. YouTube personally maintain the list of most played songs by the users. The default sentence is “play songs form YouTube”. As user request for a particular song it adds it to my music and plays the response. It directly access the speakers attach to the raspberry pi.

H. News

This module is implemented using RSS feed streaming of the google news. The default word for the news “play some news”

I. Unclear

In case any query that does not match the dictionary keys is send to the google assistant which responds by through choice from an array of words “I don’t get that”, ”the query cannot be resolved”, “My apologies”

J. Wake Word Detection

This model takes care of Hotword detection which turns on the code for user input. The microphone is always listening and unlocks on the word “Reemie”. The personal model is created using the Snowboy API [10] and detection is done using the query analysis.

IV. RESULT

This testing stage was done with 10 different faces simultaneously because it give multiple capacity to the image recognizer and around 1000 queries were at an interval of the response to the assistant .. The results from this test on ReeMiE were very positive, as many of the modules are dependent on the api calls, the functioning requires high speed internet connection. For the queries that can resolved locally no internet is needed.

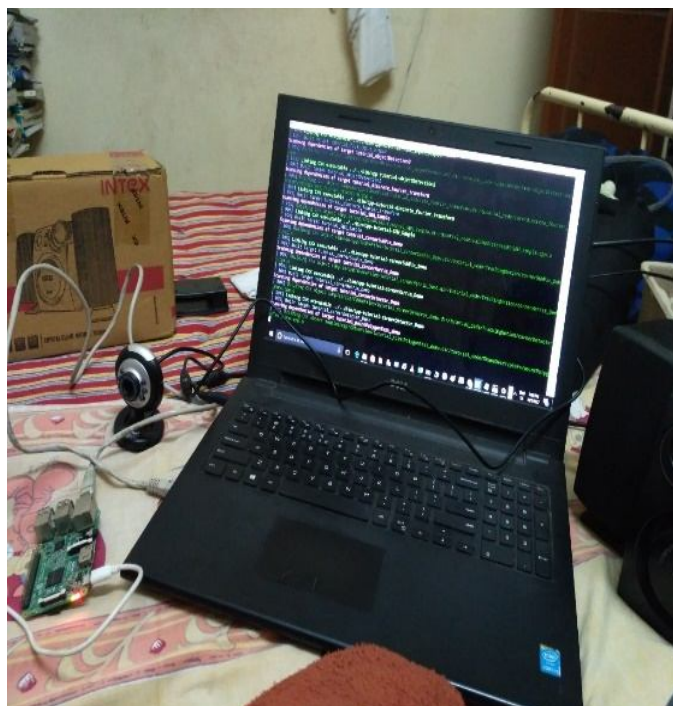


Fig. 6: Final Module

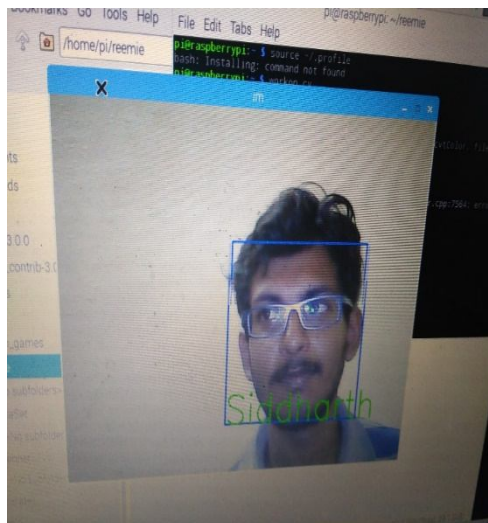


Fig. 7: face detected

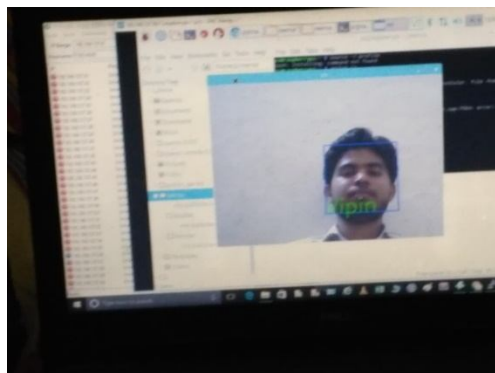


Fig 8: face detected

V. CONCLUSIONS AND FUTURE WORK

In this paper introduces the implementation of a secure biometric based personal assistant system. It indicates the flaws in the current system and ways to overcome the flaws. Many modules used are from open source libraries and are customized to suit the current hardware system. Personal assistant has huge purview of scope in the areas of centralized control over devices using internet of things. Through attachment of computer vision with the proposed work it can become complete home surveillance system.

REFERENCES

- [1] Wayne Wobcke, Anh Nguyen, Van Ho and Alfred Kqzywicki, School of computer science and engineering, University of new south walls, Sydney nsw 2052, Australia, "the smart personal assistant: an overview"
- [2] Emad S. Othman, Senior Member IEEE – Region 8, High Institute for computers and information Systems "Voice controlled personal assistant using Raspberry pi". International Journal of scientific and engineering Research volume (2017)
- [3] Black, Alan W., and Kevin A. Lenzo. "Flite: a small fast run-time synthesis engine." 4th ISCA Tutorial and Research Workshop (ITRW) on Speech Synthesis. 20
- [4] Digital Personal Assistant for the Enterprise 2013 Intel White Paper
- [5] Comeford, L, Frank, D, Gopalakrishnan P, Gopinath R and Sredivya J. 2001 The IBM Personal Speech Assistant Acoustics, Speech and Signal Processing Proc of International Conference on Speech and Signal processing
- [6] Steven Hickson, "Voice control and Face Recognition on the Raspberry Pi
- [7] Lee, Chin-Hui, Frank K. Soong, and Kuldip Paliwal, eds. Automatic speech and speaker recognition: advanced topics. Vol. 355. Springer Science & Business Media, 200
- [8] <https://developers.google.com/assistant/sdk/overview>
- [9] <https://docs.opencv.org/2.4/modules/contrib/doc/facerec/facerecutorial.htm>
- [10] <https://snowboy.kitt.ai/>



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)