

Abstract

Table of Contents

List of Figures

Chapter

1. Preamble

a. Introduction

b. Problem Statement

2. System Design

3. SRS

4. Implementation / Methodology

5. Software Testing

6. Code

7. Conclusion

8. References

Abstract

In today's advanced hi-tech world, the need of Independent living is recognized in cases of visually impaired people who are facing the main problem of social restrictiveness. They suffer in strange surroundings without any manual aid. Visual information is the basis for most tasks, so visually impaired people are at Disadvantages because necessary information about the surrounding environment is not available. With the recent advances in inclusive technology, it is possible to extend the support given to people with visual impairment.

This project is proposed to help those people who are blind or visually impaired using Artificial Intelligence, Machine Learning, Speech and Text recognition. We can implement different modules to which the assistance is provided such as Wikipedia, opening of YouTube, Google, Gmail, Checks the weather, time, news, searching the information and some basic modules as well. It also provides an interface to log out of the system using this speech recognition technique. This is used to help the people who are visually impaired to make their life easier. They do not have to depend on anyone for help. It creates an interactable environment which is very helpful.

Chapter

1. Preamble

a. Introduction

Assistive technology for visually impaired persons has been studied for many years. In this information society, visually impaired persons need assistive tools to help them in operating digital devices so that they can get and apply digital information while learning, living and working. According to the report of the ministry of Interior, Taiwan, 80% of visually impaired persons are not originally blind and most of them have the ability of hearing. It arises that the visually impaired persons operate digital devices by the help of voice guidance.

A screen reader supports voice features to help users operating computers and dramatically reduces the difficulty of operating computers for visually impaired persons. For example, Jaws and NVDA are both famous screen readers on MS-Windows platforms. The Smartphone is one of the commonly used digital devices and the smart phone with Android system is inexpensive and so popular in the Market. Nearly 50% of visually impaired persons are unemployed. This model proposes a method for voice assistance and from this it is easier for visually impaired people. They can just speak and they receive responses instantly and the required action is performed. Due to this it benefits people and employment is also provided for visually impaired people as well.

Keeping all the above factors in mind we came up with the solution of virtual assistant. The primary objectives to bridge the accessibility gap between the average user and the visually impaired individuals with regards to the internet. The internet is blind to the visually impaired, but to not make the converse the truth, in this paper we present an end-to-end voice-based software for the visually impaired to enable them to access the internet with minimal to no keystrokes required. The user will provide the commands he wants to execute as a voice input instead of using a keyboard. The software then uses a speech to text module to convert the input speech to text which will be the command to be executed. The command is executed using selenium web driver. Once executed the user will have three options: - either to read the entire content of the website, read a summary or ask a question. The second and third options are implemented using machine learning. Once the voice input is taken and the command is executed the output is said to the user using the text to speech module. Thus, the software manages to make the internet more accessible easily, quickly and more effectively for the visually impaired. Figure 1 gives a gist of the overall solution and how the software works.

input

Speech is recognized using speech and convert to text, the commands are then recognized and executed using a selenium web driver used to automate systems resulting output is played back to the user using text to speech

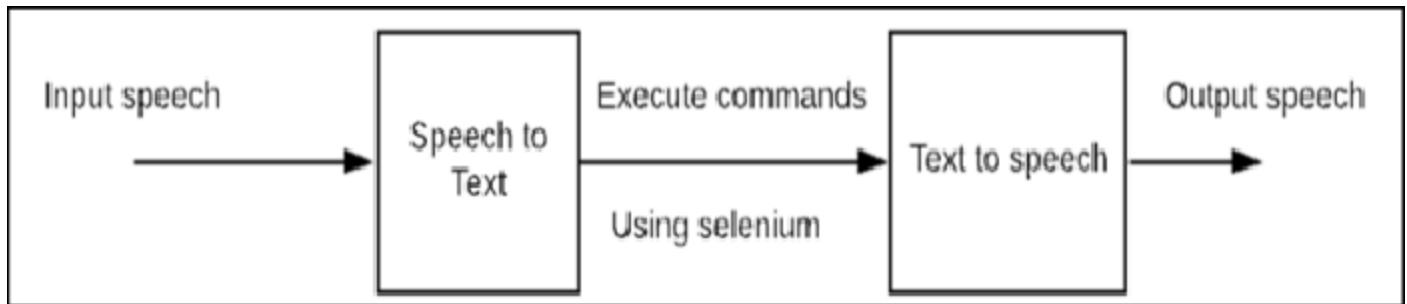


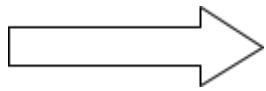
Figure :Flow diagram of the proposed solution

This paper describes the implementation of the software modules automating the three most frequently used websites that are Google, Wikipedia and Gmail by users, so as to fulfill the needs of the visually impaired as much as possible.

b. Problem Statement

To recognize a word spoken by the user.

To perform the command interpreted by the user. It ensures whether the command given by the User in the clear and precise.



2. System Design

3. SRS

3.1 Functional Requirements:

Hardware Requirements

- Pentium Processor IV or Higher
- Min 10 GB HDD
- RAM 512 MB or Higher
- 2.4 GHz or faster Processor

Software Requirements

- Windows Vista onwards, Linux, Mac OS
- In the case of building the Project from the source
 - Python Compiler
 - Tensor flow Machine learning library
 - KERAS
 - SCIKIT Learn
 - Pandas
 - NUMPY
 - Flask

3.2 Non Functional Requirements:

Performance Requirements:

- The formats of the scanned copies should be in the standard format
- Should have a training error of as low as possible

Software Quality Attributes

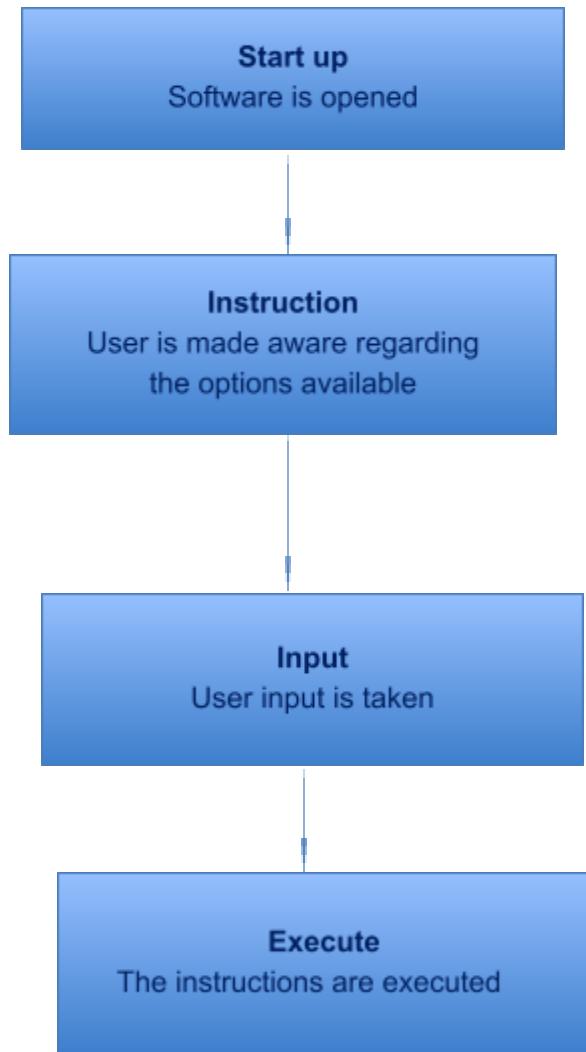
- Robustness
- Reliability
- Better learning methods
- Acquiring good accuracy results

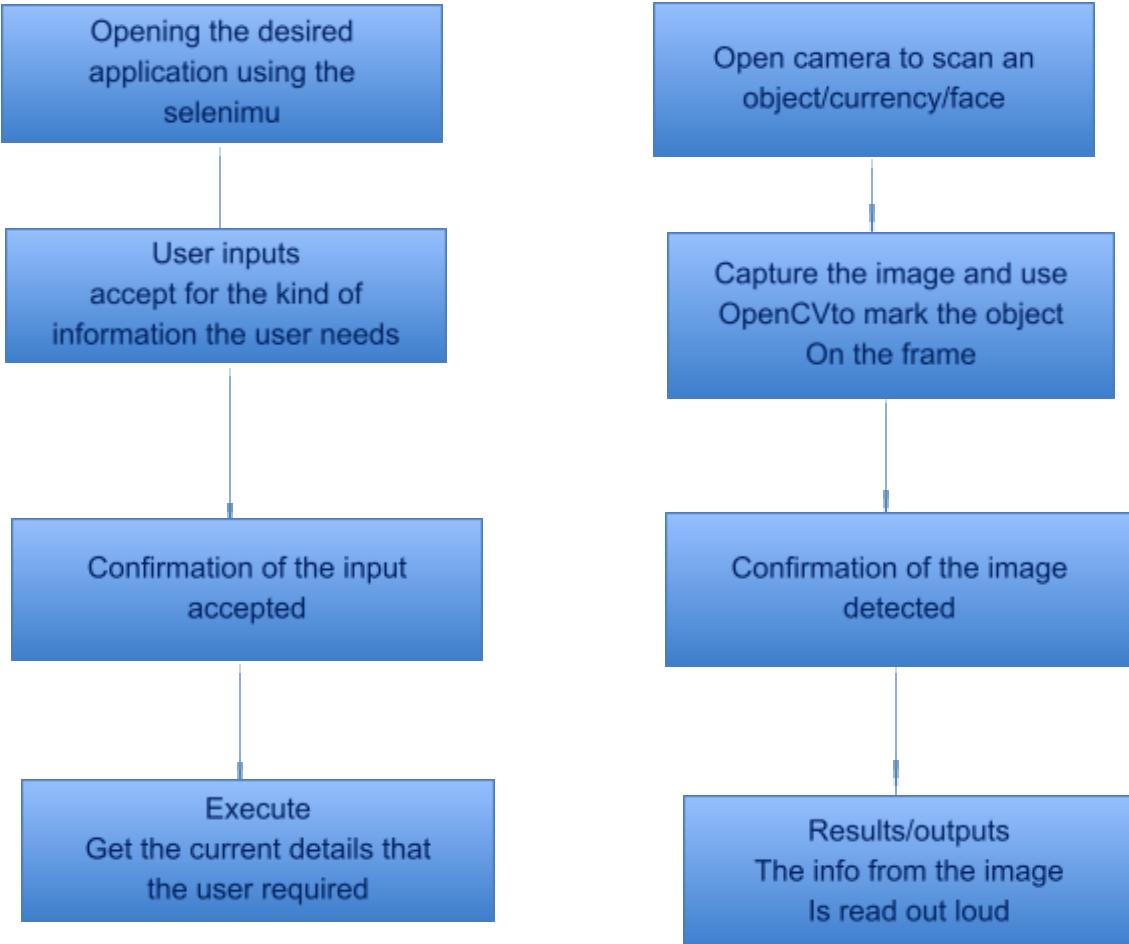
4.Implementation / Methodology

Methodology and survey

Methodology & Survey The system comprises a modular client-server distributed architecture. The system consists of the main menu which first runs on the startup of the software and the website modules. The client communicates with the server and back with the use of REST APIs, thus the website modules are not local to the client. Throughout the system, the user communicates with the software via a speech-to-text interface. The Google library of speech-to-text (Speech Recognition) for Python is used for this purpose. For communicating the system's output to the user as well as for confirming the user input, the recognized input is played back to the user using the Python text-to-speech library (pyttsx3). The modules are written in Python and make use of Selenium for automation of the respective module and Beautiful Soup for scraping the contents of the web page. The "Script" component of each module consists of customized code that entails the features of the website contained in the module. For instance, the Wikipedia module consists of a Question and Answer and A summary feature along with the traditional feature of reading out the entire article. The former is implemented by training a BERT model on the Stanford Question Answering Dataset of the APIs that hold the system together are written in Flask. The software is operating system independent to support hassle-free application and usage of the system. Many researchers have contributed to this field. Various combinations of existing technologies have been used. Braille systems, screen magnifiers, etc. went through some developments but later faced technical issues. Pilling et al. conducted a study to determine whether the internet provides opportunities for disabled people to carry out activities which they were previously unable to do or whether it leads to greater social exclusion. Sinks and kings et al. state that there is no known research to determine the reasons people with disabilities can't access the internet. Muller et al on the other hand state that the primary barrier to inaccessibility is that of economic and technical capabilities. This thought is seconded by Kirsty et al. who state that bad HTML code and use of pdf causes a hindrance in accessing the internet for the visually impaired. Although the W3C mentions a list of guidelines for maintaining a high level of accessibility for the visually impaired, Power et al. [5] state that only 50.4% of the problems encountered by users were covered by Success Criteria in the Web Content Accessibility Guidelines 2.0 (WCAG 2.0) and 16.7% of websites implemented techniques recommended in WCAG 2.0 but the techniques did not solve the problems. Android phone-controlled voice gesture and touchscreen operated wheelchair where voice and gesture are recognized through android. Developers also created a universal voice control on android which is used to launch android applications via voice commands

a. Main Menu





Time/Data/Location/Weather/News/Dictionary

5. Software Testing

The research in this thesis focuses on predicting the general sentiment polarity of the reactions to the news on Twitter/Reddit before a news article is published. To answer our research questions regarding the influence of category of product acceptance

4. Code

<https://drive.google.com/file/d/1w--U3KUq79KQRIONFBPdVi6DuLNt1HqF/view?usp=dri vesdk>

OUTPUT



THE TIMES OF INDIA

Briefs TOI+ Videos City India World Business Tech Cricket Sports Entertainment TV Web Series Lifestyle Blogs Photos Education Science All Sections

NEWS HEADLINES

Lawrence Bishnoi 'admits' to his gang killing Moose Wala How, in her death, this six-year-old gave life to four

Live: After Sonia, Priyanka Gandhi tests Covid +ve SC junks pleas seeking stay on Jagannath Temple Corridor project

TOP TECH SALE
A PERFECT TIME TO REFRESH

UPI/NET-BANKING DISCOUNT! Inspiron M 2-in-1

AMD DELL Buy Now



```
subprocess.call(['shutdown','-/1'])
time.sleep(3)
LOADING YOUR A.I PERSONAL ASSISTANT Ranbir
Hello, Good afternoon Ms.Aalia
I AM LISTENING
I AM LISTENING
I AM LISTENING
user said:are you
I AM LISTENING
user said:are you
I AM LISTENING
user said:what is the weather now
I AM LISTENING
    temperature in kelvin units is 306.29
    humidity in percentage is 41
    weather descriptionclear sky
I AM LISTENING
user said:are you
I AM LISTENING
user said:who are you
I am Ranbir version 1 point 0 your personal assistant.I am programmed to minor task like'
    'opening youtube,google chrome,gmail and stack overflow, predict time,take a photo,predict weather'
    'in different cities, get top headline news from times of India
I AM LISTENING
user said:ok bye
Your personal A.I assistant Ranbir is shutting down, Good bye
```

5.Conclusion

Various techniques to implement the aforementioned system are analyzed and summarized. Different systems have different ways of implementation along with some limitations and restrictions. These types of systems are very critical for multiple reasons and the occurrence of an error in such a system/device may cause catastrophic damage and loss. The system we are achieving overcomes the limitations of the already implemented systems. Our system consists of a basic UI on a web-based application and comprises several Deep learning models; some of them are object detection, face recognition using TF, TTS, speech recognition and so on. These modules will work together and assist in vital activities like

References

- [1] Yu tiecheng. The current development of speech recognition [J]. Communication World, 2005.
- [2] Ren tianping. Application of speech recognition technology [J]. Henan Science and Technology, 2005.
- [3] LA Liporace. Maximum Likelihood for Multivariate Observation of Markov Sources. IEEE Trans. IT, 1982, 28(5): 729-734
- [4] Zhang Ping, Zhang Qiong. Based on HMM and BP neural network for speech recognition [J]. Cross-century, 2008.
- [5] Yin Peng, Li Tao, Wang Haibing. Intelligent neural network system composed of the principle in speech recognition. Mini-Micro Systems, 2000, 21(8): 836-839.
- [6] Jiang Ming Hu, in the Yuan Baozong, Lin Biqin. Neural networks for speech recognition research and progress. Telecommunications Science, 1997, 13(7): 1-6.
- [7] Huang Shan. Voice recognition systems in the telecom prepaid business applications [J]. Information Science, 2010.
- [8] Yangshang Guo, Yang Jinlong. The speech recognition technology overview [J]. Computer, 2006.