Software Requirements Specification

For

Saksham – AI Powered Communication Solution for Disabled

Jan 2022 - May 2022

Prepared by

Specialization	SAP ID	Name
B.tech CSE - AI & ML	500075374	Utkarsh Gupta
B.tech CSE - AI & ML	500075358	Aradhya Singh
B.tech CSE - AI & ML	500076110	Priyal Gupta

Department of Informatics
School Of Computer Science
UNIVERSITY OF PETROLEUM & ENERGY STUDIES,
DEHRADUN- 248007. Uttarakhand

Table of Contents

Revision History

Date	Change	Reason for Changes	Mentor Signature

1. INTRODUCTION

1. 1 Purpose of the Project

The project aims to develop a single point solution for the communication gap between blind, deaf, dumb and normal population. Application is able to convert sign language, voice, Braille & text to each other which enables almost everyone to be able to communicate with anyone. The sign language recognition will be achieved through computer vision, voice recognition will be achieved by DNN, and Braille will be achieved by a sophisticated wireless hardware.

1.2 Target Beneficiary

The beneficiaries of the project consist of people with disability like blind, dumb, deaf and normal people. Also, this application could be clubbed with other applications that requires user's interaction or chat bots.

1.3 Project Scope

This project work can be deployed anywhere ranging from a government hospital to a school. It is capable of translating a conversation for blind, deaf, mute or any of the combination of these three disabilities. The system will enable the disabled user to very easily communicate with anyone who doesn't have any knowledge of sign language or Braille script. The system will be light weight enough that it will be able to run properly on even a pc with minimal features.

1.4 References

Paper [1] demonstrates the use of Convolution Neural Network to process the images. First the images are preprocessed for identifying the hand gestures. This process reduces the chances of error by a huge margin. These images are fed directly to Convolutional Neural Network. The model then identifies the gesture of the hand and predicts the label for the gesture. Further those labels are transferred to text to speech engine.

Paper [2] uses a sensor-based approach. It uses a series of motion sensors deployed in a hand glove. The motion sensed by the sensors will be transmitted to a nearby computer. The computer will then preprocess the data received for any redundant movement. Then this data will be sent to RNN for prediction of the label. Further the output label will be handed over to the text engine.

Paper [3] uses a vibrator-based approach for Braille communication. The system consists of six vibration motors in a 3x2 matrix. The system receives a text message from a computer through serial communication. The received text will then be converted to the Braille script. Now the Braille script having a 3x2 matrix in the form of true falls will be used for output. In the matrix if a value is true, the vibration motor will be turned on.

The hardware discussed in paper [4] is using a combination of servos for creating a 3x2 matrix for Braille scripts. The internal processing of the system is quite similar to that of paper [3]. The only difference is that this system is using a servo for providing a sense of touch instead of a vibration motor.

Paper [5] demonstrates a use of Gradient decent for Text to Speech engine. Stochastic differential equation is used along with forward differential neural network is used for conversion of text to speech.

Paper [6] discusses about software that allows users to use their voices to control computer functions and dictate text so this system is made up of two parts: the first part is for processing acoustic signals acquired by a microphone, and the second part is for interpreting the processed signals and then mapping them to words. They have used Hidden Markov Models to create models for each letter (HMM) and Mel Frequency Cepstral Coefficients will be used to extract features (MFCC). Their dataset's features will be trained using vector quantization, and the dataset's features will be tested using the Viterbi algorithm. A speech recognition technology will be used only for home automation.

2. PROJECT DESCRIPTION

2.1 Development Environment

Type	Name
Operating System	Windows 10 V-21H1 B-19042.1110 64bit
Memory	8 GB 2400 MHz DDR4
Storage	250 GB SATA3 2.4 SSD
Python	Version 3.8
IDE	PyCharm 2022 V16.11

2.2 Characteristic of Data

The dataset that is going to be used contains multiple images. Multiple images correspond to a single sign that is defined in the sign language. The signs range from alphabets, numbers to punctuations. The pictures are of size 256x256 in RBG format.

2.3 SWOT Analysis

Strength	Weakness
• Efficient to translate a conversation for blind, deaf, mute or any of the combination of these three disabilities • User Friendly • Light weighted system that it will be able to run properly on even a pc with minimal features	• Requirement of a smart system to run the application

Opportunities	Threat
• The system will enable the disabled user to very easily communicate with anyone who doesn't have any knowledge of sign language or Braille script • Can be clubbed with other applications as a chat bot	

2.4 Project Features

This project aims to develop remote communication software accessible to normal, blind, dumb and deaf people for interaction. Thereafter, construct hardware i.e., an interactive, wireless refreshable "Braille" device that would be connected to the software to facilitate communication for blind people.

2.5 User Classes and Characteristics

The project will facilitate people with disability like blind, dumb, deaf and normal people. This application could be clubbed with other applications that require user's interaction or chat bots.

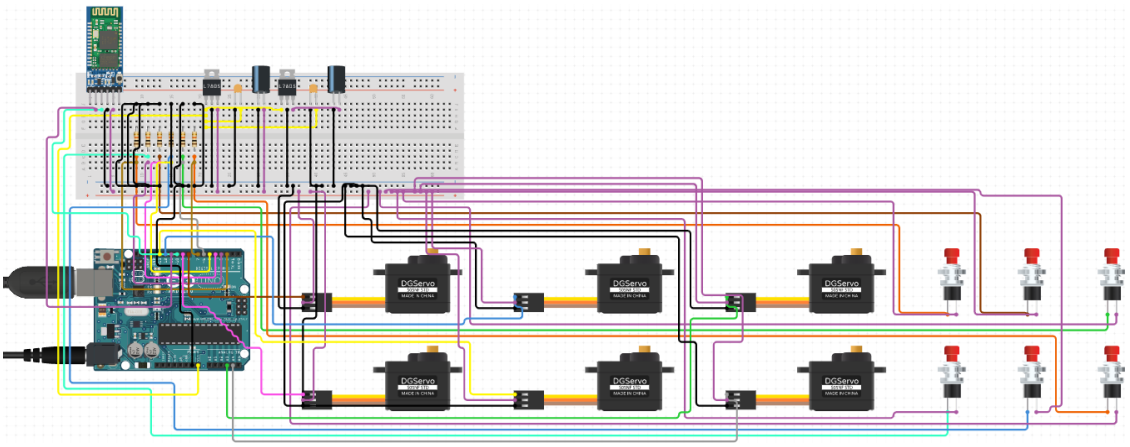
2.6 Design and Implementation Constraints

The system should be distributed in fragments such as, user application, sever application & database. This will help in the deployment phase so that the system can be easily implemented on a central server and user can access the system remotely from his own PC.

2.7 Design Diagrams

User Interface: -

Hardware: -



2.8 Assumption and Dependencies

During this project development we are assuming that out target audience that is blind, deaf & mute people know at least on of the communication method i.e., Text, Voice, Braille & Sign language.

There will be presence of an operator for operating the software for the people who are not fond with computers.

3 SYSTEM REQUIREMENTS

3.1 User Interface

The UI (user interface): point at which users interact with a computer, website, or application. The goal is to make the user's experience easy and require minimum effort on the user's part to receive the maximum desired outcome.

3.2 Software Interface

Different components in the system will communicate through a central server. Most of the processing will be done by the Django server like sign to text conversion and message sending.

3.3 Database Interface

Oracle’s MySQL will be used as the database management system.

3.4 Protocols

POST protocol will be used most of the time for sending data from User application to the Django server.

4 NON-FUNCTIONAL REQUIREMENTS

4.1 Performance Requirements

- The application should be fast enough to facilitate a seamless conversation between two users.
- The sign language recognition should be extremely accurate.
- The hardware should be easy enough to be used by any blind person.

4.2 Security Requirements

The data transfer from the user application to the Django server should be extremely secure to prevent any hacking attempt.

4.3 Software Quality Attributes

The software will be tested for all possible combination of communication mode for ensuring the smooth working after deployment. The server will also undergo penetration test to ensure the privacy of the users.