# Multilingual Voice Assistant

Sandip Chauhan[1], Ankur Singh[2], Tanya Kansal[3], and Priyanka Rai[4]

[1-4]Computer Science Department, Raj Kumar Goel Institute of Technology, Uttar Pradesh, India

*Abstract -* **Voice Assistants are used per day by many people worldwide to send messages, controlling many devices, provide data for input in small devices. The responsiveness of the system defines the user experience in this situation. For multilingual users, a further obstruction to natural interaction while communicating to the voice assistants is the monolingual character that is working on a single language of many Assistant systems, in which users are bound to a single preset language. In this process, we present the architecture of a multilingual voice assistant, that allows users to choose combinations of many regional spoken languages over the world. We leverage a new way in a specific method of real-time based language selection for the voice assistant to achieve the goal of real Natural interaction.**

*Index Terms -* **Assistant, Multilingual, Monolingual, Obstruction.**

## I.INTRODUCTION

Voice assistant is an android app that is commonly based on various technology. This multilingual voice assistant allows users to give the command to the voice assistant in multiple and foreign languages. It includes features of Artificial Intelligence mostly Natural Language Processing (NLP), Text analysis, voice recognition, speech synthesis. For this project Machine Learning (ML) model is used for well-served interaction. This type of application is very simple to use and useful in the real world. A multilingual voice assistant allows end users to interact in various common languages found in the globe but lacks when it comes to regional languages. This assistant looks like a Google interface in Google Assistant. Google API is very simple to operate by calling Hello Google and then interface ask how can I help you. For creating Voice assistant various modules related to Artificial Intelligence are used like Speech-To-Text, Text-To-Speech, Image-To-Text, OCR, Google API etc.

## II. INNOVATIVE FEATURES OF PROJECT

There is a need for such a voice assistant which can not only work and act infamous or commonly known languages but also work with the regional languages. The voice assistants should not be limited to voice-related work but also perform some action given below:

Translator:
In this multilingual voice assistant, we can translate one mode of input to another that is if anyone provides input in voice then it can give output in voice and text and vice versa.

Multilanguage support:
This application supports different type of languages which include many regional and foreign languages. Some of the languages included Hindi, English, German, French, Tamil, Sindhi Etc.

Documents read and write:
This application, we have provided documents read and write features, by this feature we can read the document in any language and make a document with speech to text feature.

Optical character recognition (OCR):
This feature allows the user to extract text from the images and that text can be further used as output and can be more processed. It is very useful to find numbers or small text from the images.

## III. LITRATURE REVIEW

As we know that Technology plays a very crucial role in making life simpler as much as.
In the whole world, there are many people who are deaf, dumb and blind. So, these people face many issues in accessing the phone or communicating with others.

Not only this, when people visited other countries so they facing issues in communicating and interacting with the people of that country.

So that we can design the app using the available technologies so that it helps the people up to some extent.

Using Machine Learning Technology, we build a Multilingual Voice Assistant for the people. The main purpose of the research paper is that there are many systems built-in limited language or in regional language.

The objective of this paper is to build a voice assistant system in multiple languages, not only in regional language but in a foreign language also with some extra features.

The action of the people is saved in the system, then the language of communication is asked and based on the user input communication is established and then it asked what action user want to perform such as want to call anyone, want to know about weather etc. in the same country or in another country whenever user want.

## IV. METHODOLOGIES

General Structure

Considering the overall research, any voice application can be done in 3 major steps: Firstly, giving the command to the computer whereas secondly to process the input on basis of algorithms. finally communicating with the desired people that is output. This project can be divided into three major parts: Each section is discussed below:

Input or Command to System

This allows us not to limit ourselves to the default mode of input in general voice-based input. Providing versatility in the assistant will provide ease to the end-user.

Mode of Command Input and Output received are discussed below:
1-Text to Speech
2-Speech to text
3-Speech to Speech

Text to Speech:

Used Different Artificial Intelligence Algorithm which converted text to speech in the same or different language. This, the algorithm processes the input text then identify the language in which text is written and it generates speech in many different languages. You can listen to output in any language.

Speech to Text:

Used Different Artificial Intelligence Algorithm which converted Speech to text in the same or different language. In this, the algorithm recognizes the speech using speech recognition techniques and generated the respective output in text and the user can see the text in any language and can listen to the text also, for this speech-to-speech technique is used.

Speech to Speech:

Used Different Artificial Intelligence Algorithm which converted Speech to Speech in the same or different language. In this, the algorithm recognizes the speech using speech recognition techniques and generated the respective output in speech using Natural Processing Language and the user can also see the output in text in any language.

Various Artificial Intelligence Algorithms and tools and techniques are used to generate this app are discussed below:

Natural Language Processing (NLP) algorithm:

NLP is a subfield of computer science and artificial intelligence for interaction between human and computer. NLP is one of the important features for an assistant. NLP is the most complicated topic of computer science because understanding the differences require more knowledge.

Indic NLP Library can be used to implement regional languages like Sindhi, Sinhala, Konkani in the voice assistant.

Artificial Intelligence algorithm:

An Artificial intelligence algorithm is used to processing the commands and also extracting the features from the images. These features later help to find text out of images. Artificial intelligence algorithms also can find basic details of image and only extract features those are useful to user that is text.

End Users:

Basically, end-users refer to this voice assistant user. They will be using this voice assistant for their daily chores and react according to that and they can use it on any device according to the compatibility

## V. TOOLS AND TECHNOLOGIES USED

There are various tools and technologies used in multilingual voice assistant. Some of the major tool and technologies are discussed below:

Python:
Basically, we are using python as the main development language in our project. Python is a high-level programming language. Python language can be integrated by various technologies like artificial intelligence and machine learning and data science.

Artificial Intelligence:
In this project, we are using Artificial intelligence and Natural language processing for processing the Input that is a command given by end-users. Also, OpenCV and pytesseract are used for Optical recognition purpose. Speech Recognition Modules are used to process the voice and test respectively. Artificial Intelligence algorithms are trained in such a manner that they can extract even the smallest useful feature and none of the user data which is useful is being altered or lost.
Inltk, Indic NLP libraries of NLP can be used to implement regional languages of the Indian Subcontinent.
Stanford NLP library can be used to implement Many Common and foreign languages.

Application Programming Interface (API):
API allows us to use real time data and help to execute pre-defined functions. Several API is integrated with the project. Google language API is used to use Common and popular languages already have in them. OpenWeatherAPI is used for weather forecasting which gives real-time Weather of any place.

## VI. SYSTEM ARCHITECTURE

The total design consists of these phases:
1) Input in the desired format that is the user with give input either in text or speech
Formatting or cleaning of the input should be done if necessary.
2) Analyze the voice or text and processing it and getting data from API if requires. This section is the heart of the whole system. Every Major work is done here and the output is going to be displayed to end-users, its accuracy and relevance depend on this section.
3) Output in the form of Speech or text from the input that is given by the user and processed.
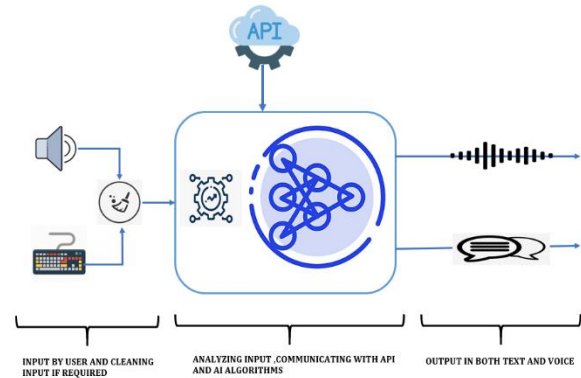


FIG 1: BASIC SYSTEM ARCHITECTURE

## VII. APPLICATIONS

This app used as an interface for setting communication in many languages in both regional and a foreign language.
Also, help the people who cannot read but can understand or the person who can't see but hear
This Project is providing a platform, not for the traditional way of using voice assistants but also proposing to integrate which make conversion of file easier or help you to take out text from images.
It can read documents for you or can make a document for you.
It can make your search easier and allow you to use to search content verbally and get output in different languages.

## VIII. CONCLUSIONS

The main goal of this paper was to establish communication between people in multiple languages. The app was succeeded to maintain communication between persons. In this app, we are providing many facilities and technologies also which helps people on daily basis. This not only implements basic voice assistant features but also integrated some features used in day-to-day life.
Hence, making it is cost-effective. The system can be used by both normal and dumb people in this digitalized electronic world.

## REFERENCES

[1] Dyah Rahma Kartika, Riyanto Sigit, S.T., M.Kom, Ph.D., Setiawardhana, S.T., "Sign Language Interpreter Hand Using Optical-Flow", 2016 International Seminar on Application for Technology of Information and Communication.

[2] Abhijith Bhaskaran K, Anoop G Nair, Deepak Ram K, Krishnan Ananthanarayanan, H R Nandi Vardhan, "Smart Gloves for Hand Gesture Recognition Sign Language to Speech Conversion System", 2016 International Conference on Robotics and Automation for Humanitarian Applications (RAHA).

[3] Sidra Syed, Seemeen Chagani, "Sign Recognition System for Differently Abled People", Proceedings of TENCON 2018 - 2018 IEEE Region 10 Conference (Jeju, Korea, 28-31 October 2018).

[4] Shivam Khare, "Finger Gesture and Pattern Recognition Based Device Security System ", Vol.2015, ISBN-15: 978-1-4799-6761-2'

[5] Alfat Jahan Rony, Khairul Hossain Saikat, Mahdia Tanzeem and F.M.Rahat Hasan Robi, "An Effective Approach to Communicate with the Deaf and Mute People by Recognizing Characters of One-hand Bangla Sign Language Using Convolutional NeuralNetwork", Vol.2018, ISBN-18:978-1-5386-8279-1, International Conference on Electrical Engineering and Information & Communication Technology.

[6] Suraksha Devi, Suman Deb, "Low cost tangible glove for translating sign gestures to speech and text in Hindi language", Vol. 2017 ISBN 17: 978-1-5090-6218-8, 3rd IEEE International Conference on "Computational Intelligence and Communication Technology" (IEEE-CICT 2017).

[7] Amiya Kumar Tripathy, Dipti Jadhav, Steffi A. Barreto, Daphne Rasquinha, Sonia S. Mathew, "Voice for The Mute ", Vol.2015 ISBN 15: 978-1-4799-8187-8, International Conference on Technologies for Sustainable Development (ICTSD-2015), Feb. 04– 06, 2015, Mumbai, India

[8] Vigneshwaran S, ShifaFathima M, VijaySagar V, "Hand Gesture Recognition and Voice Conversion System for Dumb People ", 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS).

[9] Aarthi M, Vijayalakshmi P, "Sign Language to speech Conversion", Vol. 2016, 5th International Conference on Recent Trends in Information Technology.

[10] DOUGLAS O'SHAUGHNESSY, SENIOR MEMBER, IEEE, "Interacting With Computers by Voice: Automatic Speech Recognition and Synthesis" proceedings of THE IEEE, VOL. 91, NO. 9, SEPTEMBER 2003