A REAL-TIME SMART LANGUAGE BARRIER BREAKING SYSTEM

Dr. K. Sai Venu Pratap[1], G. Mahesh[2], D. Harish Kumar Reddy[3], A. Charan Kumar[4], J. Nikhil[5]

[1] Associate Professor, Department of Electronics and Communication Engineering, Aditya College of Engineering, Madanapalle, India
[2,3,4,5] UG Students, Department of Electronics and Communication Engineering, Aditya College of Engineering, Madanapalle, India

Abstract: By facilitating contact through both spoken language and hand gestures, the Real-Time Smart Language Barrier Braking System aims to remove communication barriers. The central processing unit of the system is a Raspberry Pi. While a USB microphone records user speech, a USB camera records hand motions. To accurately translate these inputs into intelligible text, picture processing and speech recognition techniques are applied. Two modes of operation are supported by the system: Users can select the most practical mode of communication by using Sign Mode, which recognizes hand signals, and Voice Mode, which processes spoken language. A 16x2 LCD panel quickly displays the created text, offering concise and instantaneous feedback.

The system's multilingual support, which includes English, Telugu, and Hindi, makes it easier for users from various linguistic backgrounds to communicate with each other. The suggested system, which was created with simplicity, affordability, and user-friendliness in mind, provides a useful and economical option for real-time communication in public service, healthcare, education, and everyday settings where speech or language impairments might be present.

Keywords: Language Translation, Raspberry Pi, Speech Recognition, Image Processing, Assistive Technology

I. INTRODUCTION

In real-time human interaction, communication constraints resulting from linguistic variety and speech or hearing impairments pose serious challenges. Conventional communication systems frequently lack the potential to handle multi-modal input processing and unified output representation, hence restricting their efficacy in assistive and multilingual situations. A Real-time sign language to speech conversion using Bluetooth for communication and flex sensors [11] was a base for ours research. This research suggests a Real-Time Smart Language Barrier Breaking System that combines automatic speech recognition (ASR) and computer vision-based sign recognition into a single embedded platform in order to overcome these constraints. A Raspberry Pi is used to create the system, allowing for effective edge-level computing that is portable and uses little power.

The suggested system uses a camera to record hand motions, which are then analysed using image processing methods to extract useful characteristics for sign interpretation. Concurrently, speech recognition algorithms are used to process voice input obtained through a microphone in order to translate spoken words into text. To ensure consistent processing and interpretation, the input modalities are combined into a single textual representation.

Arduino microcontroller is used as a voice conversion system that converts predefined user input voice into audible speech for communication assistance [15]. An LCD module shows the processed data, giving consumers instantaneous, unambiguous visual feedback. The system supports several languages, including English, Telugu, and Hindi, to enhance accessibility and communication across various locations. The system provides more flexibility and enables users to interact in the way that best meets their needs by integrating sign and voice input modes with a single output interface. Because the suggested system is built to function in real time, there is very little time lag between recording the input and showing the output. It is also simple to expand and improve in the future because to its modular architecture. More features, such support for more languages, sophisticated machine learning techniques, or different output devices, can be added without meaningfully modifying the present architecture. Arduino based sign language translator that converts hand gestures into speech, improving communication accessibility [12].

All things considered, the Real-Time Smart Language Barrier Breaking System offers an easy, effective, and affordable way to get around communication obstacles. The system enhances communication for individuals with speech or hearing impairments as well as those with diverse linguistic origins by combining sign language identification, speech recognition, multilingual text conversion, and real-time LCD display into a single embedded platform. It is ideal for usage in assistive technology, educational institutions, healthcare services, public information systems, and other smart communication environments due to its portability, usability, and real-time performance. This research is useful for speech-to-speech people communication, that is very important for the people who can speak but do not have standardised language script for writing do have [13]

II. LITERATURE SURVEY

The development of real-time communication and language translation systems has greatly benefited from recent developments in AI and ML. Guillaume Lample and Alexis Conneau's [1] cross-lingual language models improved multilingual comprehension by facilitating efficient knowledge transfer between several languages. Singh et al. **[2]** reviewed major ML-machine learning algorithms for intelligent applications. In a similar vein, Minh-Thang Luong and Christopher D. Manning [3] showed how neural machine translation methods might enhance speech-based translation systems. Khan et al. **[4]** demonstrated image detection Algorithms using computer vision techniques. Kamaraj et al. **[5]** improved speech recognition using Artificial Intelligence (AI) translation techniques. Over time, speech recognition technology has also advanced significantly. DL techniques, which offer greater recognition accuracy, have rapidly supplanted conventional methods based on Hidden Markov Models (HMMs). While Song et al. [8] established the SpecSwap data augmentation method to enhance the performance and resilience of automatic speech recognition systems, Saksamudre et al. [10] provided a thorough overview of voice recognition strategies.

In 2018 Radwan et al. showed the efficacy of deep learning algorithms for optical character recognition (OCR), Tappert et al. (1990) set the groundwork for handwriting and gesture detection. These advancements have greatly improved the precision and dependability of visual recognition systems. Singh et al. **[6]** discovered artificial intelligence (AI) for intelligent human communication structures.

The increasing demand for real-time multilingual communication systems, especially in linguistically varied situations, has been highlighted by several recent research. In 2015, Kalyani and Sajja[9], and Pamudhurthy et al.[7] emphasized the significance of intelligent communication systems that can assist users from various linguistic backgrounds. Despite these developments, the majority of current systems lack an integrated platform that can handle several communication modalities at once and are only intended to process a single kind of input, such as speech or hand gestures.

Current communication devices developed to circumvent language and sensory obstacles mainly work independently. While sign language recognition systems use cameras and image processing methods to decipher hand gestures, speech-to-text systems use microphones and speech recognition algorithms to translate spoken language into text. These systems are usually restricted to a single input modality and lack seamless integration, despite the fact that they successfully carry out their intended functions.

The limited production capacity of current systems is another drawback. Most don't offer a consistent interface for various communication needs and simply deliver results in one format, such text or voice. Furthermore, a lot of systems include processing delays, which makes them less suitable for real-time communication. These solutions are less effective in settings where users speak multiple languages since multilingual support is frequently restricted.

Additionally, a lot of current solutions increase system complexity and implementation costs by requiring separate devices or several programs to carry out various communication activities. Furthermore, the lack of a straightforward and cohesive engagement method frequently limits its usage for people with speech or hearing difficulties. Research aims for the robustness of speech translation methods against speech recognition errors [14].

This article suggests a Real-Time Smart Language Barrier Breaking System that combines speech and sign language detection into a single embedded platform in order to get beyond these restrictions. Both voice and sign inputs are supported by the system, which processes them instantly and shows the identified data on an LCD panel. The suggested system, which has multilingual support for English, Telugu, and Hindi, provides a straightforward, affordable, and user-friendly way to enhance communication between individuals with speech or hearing impairments and those from diverse linguistic backgrounds.

III. MATERIALS USED

In order to provide a dependable and effective real-time communication platform, the Real-Time Smart Language Barrier Breaking System was created using a combination of hardware and software components. Together, these elements collect user input, process it, and quickly display the results.

All hardware and software functions are coordinated by the Raspberry Pi, which acts as the system's central processing unit. While a USB microphone records spoken input for speech recognition, a USB camera records hand movement that are processed for sign language recognition. A 16x2 LCD panel shows the identified data, giving users instantaneous, unambiguous visual feedback.

Push buttons for choosing the preferred language and operating mode, such as Sign Mode or Voice Mode, are included to enhance user interaction. The system operates steadily thanks to a regulated power supply unit, and the hardware components are interfaced with appropriate connecting cables.

The Raspberry Pi OS, the foundation of the software platform, offers the environment required for hardware interface and real-time execution. Because of its ease of use, adaptability, and broad support for image processing and speech recognition libraries, Python is the most popular programming language. The system incorporates a number of open-source libraries to carry out various functions.

PyAudio enables real-time audio collection from the microphone, MediaPipe is used for precise hand motion detection and tracking, Speech Recognition is used to translate spoken language into text, and OpenCV is used for image acquisition and processing.

The system can recognize hand movements and spoken language, translate them into understandable text, and show the results on the LCD in real time thanks to the integration of these hardware and software components. This comprehensive strategy ensures ease of use, portability, and dependable performance while offering a straightforward, effective, and economical way to overcome communication hurdles.

IV. SYSTEM ARCHITECTURE

The Raspberry Pi, which acts as the central processing unit and manages all input, processing, and output operations, is the foundation of the suggested Real-Time Smart Language Barrier Breaking System. The overall architecture is designed to provide real-time communication by integrating sign language recognition and speech recognition into a single embedded platform. The system is powered by a regulated power supply that delivers stable electrical power to the Raspberry Pi and all connected peripherals. The Raspberry Pi boots from a memory card containing the operating system, application software, and the necessary program files required for system execution.

The overall architecture of the system is shown in Figure 1. Several devices are used by the system to receive inputs. While a USB microphone records spoken language for speech recognition, a USB camera records hand gestures for sign language recognition. Additionally, users can choose their desired output language and operation mode (Sign Mode or Voice Mode) using push buttons. The Raspberry Pi receives all input data and processes it using algorithms for speech recognition and picture processing.

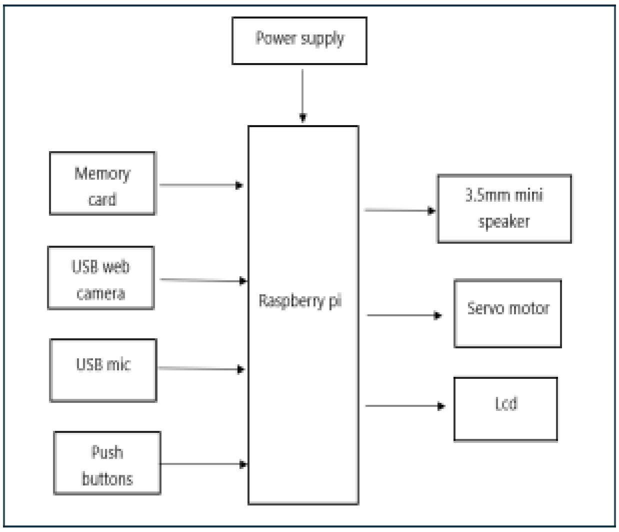

Figure 1: Block Diagram of Smart Language Convertor

Following processing, the identified data is translated into text and shown on a 16x2 LCD module, giving consumers instantaneous, clear visual feedback. The identified text can be shown in Hindi, Telugu, or English based on the user's preference thanks to the system's multilingual output capabilities. This function improves accessibility overall and facilitates communication amongst users who have diverse language preferences.

The system operates using a straightforward input-process-output methodology. Through the appropriate input devices, hand movements and vocal commands are recorded, processed by the Raspberry Pi, and converted into understandable text. The LCD then shows the processed result in real time, guaranteeing quick and dependable communication with less latency.

The suggested system's modular design makes future improvements and extension simple. It is possible to incorporate more input devices, sophisticated machine learning models, or different output interfaces without significantly altering the current design. The system is ideal for applications including assistive technology, smart communication systems, educational institutions, healthcare services, and public information centers because of its real-time performance, portability, and user-friendly operation.

V. METHODOLOGY

The Real-Time Smart Language Barrier Breaking System's methodology employs a methodical workflow to facilitate real-time communication using both spoken language and hand gestures. The system is built to handle two distinct input modes, Voice Mode and Sign Mode, and translate the identified data into text that is shown on an LCD panel.

Initializing the Raspberry Pi and all attached hardware, such as the USB camera, USB microphone, LCD module, and push buttons, is the first step in the process. Following initialization, the user uses the push buttons to choose the preferred mode of operation. The system switches between Sign Mode and Voice Mode based on the selection. The USB camera continuously takes live pictures of the user's hand gestures while in Sign Mode. MediaPipe is used to identify and track hand landmarks, while OpenCV is utilized for image capture and pre-processing. The technology recognizes the matching motion and translates it into understandable text based on the retrieved attributes.

The USB microphone records the user's speech in real time when in Voice Mode. PyAudio is used to record the audio, and speech recognition algorithms are used to translate spoken words into text. After that, the identified speech is turned into a comprehensible textual representation. The identified data is transformed into a single text format regardless of the input option that is chosen. The 16x2 LCD module instantly displays the created text, giving users instantaneous, unambiguous visual feedback. Additionally, the system allows for multilingual output, thus depending on the user's preferred language, the identified text can be shown in Hindi, Telugu, or English. The entire procedure is organized as follows: input acquisition, data processing, text conversion, and LCD display. The suggested approach guarantees quick, precise, and dependable communication by reducing the time between input capture and output synthesis. This comprehensive approach offers a straightforward, effective, and user-friendly way to get around language and communication obstacles in real-time settings.

VI. RESULTS

By combining speech recognition, multilingual text conversion, sign language detection, and LCD-based output, the suggested Real-Time Smart Language Barrier Breaking System was effectively realized on a Raspberry Pi platform. To confirm the system's performance in real-time communication, it was assessed under various operational scenarios. The results of the experiments show that the suggested system efficiently recognizes spoken language and hand gestures, transforms the recognized input into intelligible text, and displays the output with little processing latency. Push buttons are used to switch between the system's two separate modes, Sign Mode and Voice Mode. Below is a discussion of each mode's performance.

A. Sign Mode Gesture Recognition

The USB camera continuously takes live pictures of the user's hand gestures while in Sign Mode. OpenCV and MediaPipe are used to process the recorded frames in order to identify hand landmarks and extract pertinent characteristics like palm orientation and finger locations. The matching sign is then

determined by comparing these characteristics with predetermined gesture patterns.
The testing results demonstrate that, in normal illumination settings, when the user's hand is clearly visible to the camera, the system can accurately recognize common hand movements. When a legitimate gesture is identified, it is translated into the appropriate language and shown on the LCD panel.

Fig 2. Sign Mode Gesture Recognition Output

Additionally, the system confirms successful recognition by displaying the gesture detection status. These outcomes show how well the computer vision module works to provide accurate and timely sign language recognition.

B. Voice Mode Speech Recognition

The USB microphone records the user's speech in Voice Mode, which is then processed using speech recognition methods. The identified message is then translated into the user-selected language, such as Hindi, Telugu, or English, once the recorded audio has been converted to text.

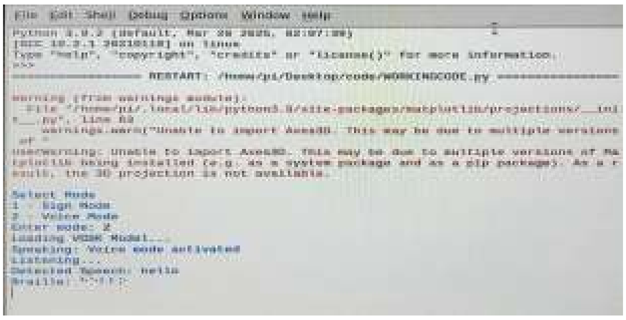

Fig 3: Voice Mode Speech Recognition Output

According to experimental findings, the speech recognition module works best when there is less background noise and clear input speech. The system facilitates natural and seamless communication by producing accurate text with little processing latency. The effective display of the identified sentence shows how the speech recognition module may facilitate multilingual conversation without the need for human text insertion.

C. LCD Output

A 16x2 LCD panel shows the processed data from both Sign Mode and Voice Mode, giving users quick and easy visual feedback. The first row of the LCD shows the chosen operating mode or system status, and the second row shows the message that has been recognized or translated. Users can immediately comprehend the output thanks to its straightforward presentation format, which also enhances readability.

The LCD module runs without an internet connection, uses very little power, and is simple to interface with the Raspberry Pi. Because of these features, the system can be used in embedded and portable applications where dependable real-time communication is crucial.

Fig 4: LCD Output

Overall, the experimental results verify that the suggested system, which combines speech recognition, sign language recognition, multilingual text conversion, and LCD-based output into a single embedded platform, offers precise and effective real-time communication. The system is a workable and affordable way to get around communication obstacles because of its low latency, dependable performance, and user-friendliness. Applications in assistive technology, healthcare, education, public service centres, and other smart communication contexts can benefit from its mobility and real-time operation.

VII. CONCLUSION

The Real-Time Smart Language Barrier Breaking System offers a practical way to get over communication obstacles brought on by speech or hearing impairments and language hurdles. The system allows users to communicate using both hand gestures and spoken language by combining multilingual text conversion, speech recognition, and sign language recognition on a Raspberry Pi platform. To provide prompt and dependable communication, the identified data is processed in real time and shown on an LCD. The findings of the experiment show that there is little processing time and adequate recognition accuracy. The system is portable, inexpensive, and appropriate for use in assistive technology, healthcare, education, and public services because to the utilization of inexpensive hardware and open-source software. Hindi, Telugu, and English are currently supported by the system.

More regional and international languages will be added in the future; recognition accuracy will be increased through sophisticated deep learning techniques, the vocabulary of supported gestures and speech commands will be expanded, and more output options, such as mobile applications and cloud-based services, will be integrated.

REFERENCES

1. G. Lample and A. Conneau, “Cross-lingual language model pretraining,” *arXiv preprint arXiv:1901.07291*, 2019.
2. P. Singh, S. Verma, I. Khan, and S. Sharma, “Machine Learning: A Comprehensive Survey on Existing Algorithms,” *J. Comput. Sci. Eng. Softw. Test.*, vol. 7, no. 3, pp. 1–9, 2021.
3. M.-T. Luong and C. D. Manning, “Stanford neural machine translation systems for spoken language domains,” in *Proc. 12th Int. Workshop Spoken Language Translation (IWSLT)*, 2015, pp. 76–79.
4. P. A. Khan, S. Sharma, I. A. Khan, and P. Singh, “Image Detection based technique for shutting down the Applications opened in Windows Operating Environment,” *Contemp. Issues Soc. Sci.*, p. 290.
5. S. A. Kamaraj, M. Gautham, S. Karthikeyan, and R. Parkavi, “Enhancing Automatic Speech Recognition and Speech Translation Using Google Translate,” in *Handbook of Research on Data Science and Cybersecurity Innovations in Industry 4.0 Technologies*, IGI Global, 2023, pp. 220–241.
6. S. Singh, C. Sharma, S. Sharma, and N. K. Verma, “Re-Learning Emotional Intelligence Through Artificial Intelligence,” in *Proc. 9th Int. Conf. Reliability, Infocom Technologies and Optimization (ICRITO)*, 2021, pp. 1–5.
7. P. S. V. Pamudhurthy, R. Komali, A. B. Panyam, A. Katarukonda, and H. Raithatha, “Real Time Language Translation Using AI and ML,” 2024.
8. X. Song, Z. Wu, Y. Huang, D. Su, and H. Meng, “SpecSwap: A Simple Data Augmentation Method for End-to-End Speech Recognition,” in *Proc. Interspeech*, 2020, pp. 581–585.
9. Kalyani and P. S. Sajja, “A review of machine translation systems in India and different translation evaluation methodologies,” *Int. J. Comput. Appl.*, vol. 121, no. 23, 2015.
10. S. K. Saksamudre, P. P. Shrishrimal, and R. R. Deshmukh, “A review on different approaches for speech recognition system,” *Int. J. Comput. Appl.*, vol. 115, no. 22, 2015.
11. F. Khan and R. S. Chahar, "Real-Time Sign Language to Speech Conversion Glove Using Arduino and Wearable Sensors," *Research & Reviews: Journal of Space Science & Technology*, vol. 14, no. 3, pp. 10–18, 2025.
12. N. Manoj, V. Surendran, V. Rathan, V. T., and S. Emmanuel, "Sign Language Translator: A Hand Gesture Recognition Device Using Arduino," *International Journal of Engineering Research & Technology (IJERT)*, vol. 11, no. 1, 2023.
13. C. Zhang *et al.*, "UWSpeech: Speech to Speech Translation for Unwritten Languages," arXiv:2006.07926, 2020.
14. Q. Cheng, M. Fang, Y. Han, J. Huang, and Y. Duan, "Breaking the Data Barrier: Towards Robust Speech Translation via Adversarial Stability Training," arXiv:1909.11430, 2019.
15. M. A. A. Rakib *et al.*, "Arduino Uno Based Voice Conversion System for Dumb People," *European Journal of Engineering and Technology Research*, vol. 7, no. 2, pp. 118–123, 2022.