International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

Sign-Voice Bidirectional Communication System for Normal, Deaf/Dumb and Blind People based on Machine Learning

Jyothsna M [1], Srinika Kontham [2], Sharanya Balachandran [3], Meghana Danta [4], Sindu Naine [5]
Electronics and Telematics Engineering , G.Narayanamma Institute of Technology and Science , Hyderabad , India

Abstract- — The SignVoice system is based on artificial intel ligence technology that offers a bidirectional communication system for deaf, mute, and visually impaired persons to communicate smoothly with normal persons. The SignVoice system is based on machine learning, deep learning, and computer vision technologies to offer different types of communication such as sign language, speech, text, and image-based communication. The hand gestures are recorded through the webcam and processed through MediaPipe to identify the landmark and classify the image through machine learning to produce text output. The input is converted into text through Whisper for speech input, and the text output is generated through an artificial intelligence- based chatbot and then converted into audio through text- to-speech technology. The SignVoice system is based on the hybrid approach to process the gestures through client-side processing and computationally intensive operations such as speech recognition through cloud-based services. In addition to this, the chatbot can perform image input, text output, and speech output that can be helpful for visually impaired persons. The proposed SignVoice system can communicate efficiently and accurately for impaired persons through gesture, speech, and intelligent text-based responses.

Keywords—Multimodal Communication, Artificial Intelligence, Machine Learning, Deep Learning, Computer Vision, MediaPipe, Whisper Model, AI Chatbot, Assistive Technology, Accessibility, Human-Computer Interaction.

I. INTRODUCTION

Communication is one of the most important aspects of the interaction of humans, as it gives the people involved an opportunity to exchange the required information, ideas, and emotions. However, deaf, mute, and blind people face a great challenge of communicating with the general public [3] The deaf and mute people primarily use the sign language as the medium of communication, whereas the blind people rely on voice-based communication. However, the general public does not know the sign language, making it difficult for the differently-abled people and the normal public to interact. In the context of the above problem, the Sign-Voice Bidirectional Communication System would be provided as the solution, which would help narrow the gap of communication with the help of the latest technologies. In the initial stages of the project, the Sign-to-Text, Speech-to-Text, and Text-to- Speech systems were introduced, which would help convert the gestures and speeches into understandable forms. In the second phase of the project, the system would be improved by the addition of the multimodal chatbot, which would have the ability to take inputs from the user in any of the forms: text, speech, or image, and provide intelligent output[7],[6]. Moreover, a dedicated AI chatbot module is created to help visually impaired users through voice-based interaction and image description. A separate sign language recognition module is also created to help deaf and mute users convert their gestures into text.

II. BACKGROUND

The importance of assistive communication technology has increased significantly in the past few years due to the need to address the requirements of people with hearing, speech, and visual disabilities. In general, communication poses a challenge for differently abled people, as deaf and mute people use sign language, and visually impaired people rely on auditory communication. These types of communication cannot be understood by anyone, so there is a need to use intelligent systems that can communicate through technology. The development of speech-to-text systems has progressed from conventional methods such as Hidden Markov Model (HMM) to advanced methods such as deep learning tech- niques, including Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Transformer. Mel Frequency Cepstral Coefficient (MFCC) and Bidirectional Kalman Filter-based noise filtering techniques have improved the performance of speech recognition systems. These tech- niques have improved the performance of speech recognition systems

in noisy environments. However, there are some challenges in speech recognition systems[12],[13].

Similarly, sign language recognition systems have also progressed by integrating computer vision and deep learning techniques. Techniques involving Convolutional Neural Networks (CNN) and MediaPipe have achieved accurate results in real-time hand gesture detection and classification. Moreover, techniques involving RNN-LSTM have further enhanced sign language recognition by incorporating temporal dependency in dynamic hand gestures. Although high accuracy has been achieved in sign language recognition systems, they remain vulnerable to environmental factors such as lighting conditions and hand positions[11],[15],[10].

With regard to text-to-speech (TTS) systems, Natural Language Processing (NLP) and Digital Signal Processing (DSP) techniques are commonly used for text-to-speech synthesis. Although traditional techniques involving concatenative and formant synthesis have achieved accurate results in text-to-speech synthesis, they lack human-like quality in generated speech. Although recent techniques involving deep learning have achieved high-quality voice, TTS systems remain effective for visually impaired users[14].

Multimodal interaction is a new concept that has been identified as a potential solution for the current limitations of single input systems. It is based on multiple input modes, including speech and gesture. Multimodal systems have shown that they can improve communication accuracy and minimize errors[8]. Various studies on assistive robotics and artificial intelligence-based systems have confirmed that multiple input modes for interacting with users is effective in improving interaction and making systems more robust, though it adds more complexity and computational requirements to the systems[2],[7],[8].

Another important component of modern communication systems is artificial intelligence-based chatbot systems. These systems make use of Natural Language Processing and deep learning algorithms for effective analysis of user queries and generation of responses based on context. These systems, along with speech recognition and text-to-speech systems, make it possible for users to communicate hands-free, making it highly beneficial for differently abled users[4],[5].

The latest advancements in assistive technology for visually impaired people include object detection, OCR, and voice assistance. These advancements include the use of object detection techniques such as YOLO and machine learning algorithms, which have greatly improved the accuracy of these systems. However, the main issue with these systems is their efficiency in terms of real-time performance[6].

The above research indicates that there is a need for an integrated multimodal communication system, which can efficiently integrate the various individual systems such as speech recognition, sign language recognition, and chatbot systems into a single system to provide efficient, real-time, and inclusive communication, which is the main idea behind the proposed SignVoice system.

III. METHODOLOGY

A. Input Acquisition

The input acquisition stage represents the first and most vital stage of the Sign-Voice Bidirectional Communication System, where data acquisition takes place from the user. The Sign-Voice Bidirectional Communication System aims at supporting the acquisition of various forms, thus allowing the user to interact with the system conveniently depending on the nature of the user. These forms include gestures, speech, text, and images.

1) Gesture Input Acquisition:

- The gesture input is obtained through a webcam in real time.
- Video frames of hand movement are recorded through the webcam.
- These video frames are processed to identify the presence of hands and movement of hands.
- Lighting and background conditions are taken into consideration for proper hand movement capture.
- These video frames are sent to the gesture recognition module for processing.

This method enables deaf/mute individuals to communicate naturally through sign language.

2) Speech Input Acquisition:

- The speech input is recorded through a microphone connected to the system.
- The system receives audio signals, which are then converted to digital form.
- The recorded audio is temporarily stored in digital form, for example, in mp3 or wav formats, for further processing.
- The level of noise and speech clarity are significant factors in speech input quality.

- The recorded audio is then sent to the backend for speech recognition.

This allows users, particularly blind people, to interact with the system through voice commands.

3) Text Input Acquisition:

- The user can directly write the text using the keyboard or interface.
- The system has a space where the user can write the text or query.
- This method can be used if the user wants to communicate through text or if other forms of input cannot be used.
- The written text can be sent directly to the chatbot or processing system.

This text input is a simple way of communicating.

4) Image Input Acquisition:

- The input for the images is taken either through a camera capture or an upload.

- The user can take images in real time or can upload existing images.
- The images will be temporarily stored and then used for processing.
- The quality and resolution of the images will affect the accuracy of the object detection and text extraction.

This type of input will be helpful for visually impaired users.

5) Input Selection and Routing:

- After the acquisition of the input, the system determines the type of the input.
- Based on the type of the input, the data processing takes place[6].

The types of the inputs are as follows:

1) Gesture → Sign Language Module
2) Speech → Speech Processing Module
3) Text → Chatbot Module
4) Image → Image Processing Module

B. Gesture Recognition Process

- Hand gesture is detected through the webcam.
- MediaPipe detects hand landmarks such as finger joints and positions.
- Landmarks x and y coordinates are extracted as feature points[11].
- A trained machine learning model is used to classify the gesture[10].
- The detected gesture is converted into text.

C. Speech Processing Workflow

- The speech input is recorded through the microphone.
- The audio file is transmitted to the cloud backend.
- The Whisper model is used to process the audio and convert it into text.
- The text is cleaned and prepared for processing[12].

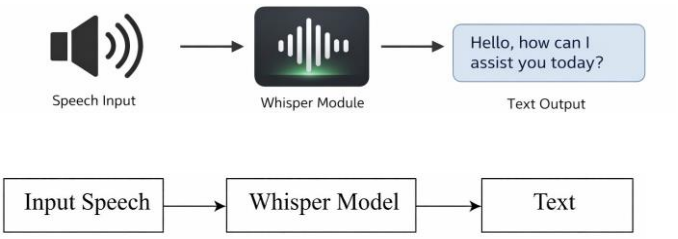

Figure. 1. Speech-to-Text Conversion flowchart

The detailed architectural workflow for capturing user audio and processing it through the Whisper model into text is depicted in Fig. 1.

D. Text-to-Speech Conversion

- The processed text is converted into speech using a TTS engine.
- The output audio is generated in a clear and natural voice. The sequence of operations required to synthesize natural-sounding speech from the chatbot's generated text is shown in Fig. 2.

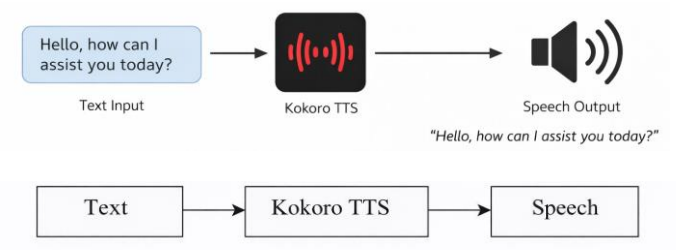

Figure. 2. Text-to-Speech Conversion flowchart

E. Chatbot Processing

- The input from gesture or speech modules is fed into the chatbot.
- NLP techniques are employed to comprehend the user's queries.
- The chatbot formulates suitable responses.
- The responses are sent in text and/or speech form[4].

F. Image Processing Workflow

- The system takes image data as input from the user.
- Object detection and OCR techniques are used.
- Information relevant to the task at hand is extracted from the image.
- A meaningful description of the image is created as text.
- The description can be converted to speech[6].

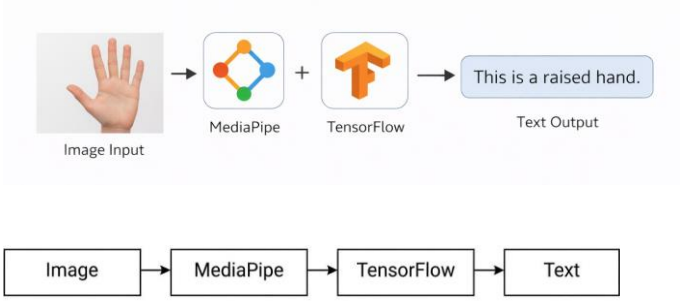

Figure. 3. Image-to-Text Conversion flowchart

The image processing pipeline, which utilizes MediaPipe for landmark identification and TensorFlow for classification, is illustrated in Fig. 3.

G. Output Generation

- Final output is provided in:

1) Text format for readability
2) Audio format for accessibility

- Enables effective communication among all user types.

IV. SYSTEM DESIGN

The proposed Sign-Voice Bidirectional Communication System is planned as a multimodal AI-based assistive system[7], which will allow normal users, deaf/mute users, and visually impaired users to communicate with each other. The system will incorporate advanced technologies, including machine learning, deep learning, natural language processing, and computer vision, to process different kinds of inputs, including gestures, speech, text, and images.

The system will be based on a hybrid approach, in which some tasks, including gestures, will be performed locally, and computationally complex tasks, including speech recognition, computer vision, and chatbot interaction, will be performed using cloud services, including Gemini Flash 2.5[1].

A. System Architecture Design

The system architecture is organized into multiple layers that work together to process inputs and generate outputs effectively.

1) Input Layer: The input layer is responsible for collecting user inputs in different forms:

- Gesture Input: Captured using a webcam for sign language recognition
- Speech Input: Captured through a microphone for voice interaction
- Text Input: Entered through a user interface
- Image Input: Captured or uploaded for visual analysis This layer ensures flexibility, allowing users to interact using their preferred mode of communication.

2) Processing Layer: The processing layer is the core of the system where all computations take place.

- Sign Language Processing:

MediaPipe detects hand landmarks, and a machine learning model classifies gestures into text.

- Speech Processing:

Audio input is processed using the Whisper model to convert speech into text.

- Image Processing:

Images are analyzed using object detection and OCR to extract meaningful information.

- AI Chatbot (Gemini Flash 2.5):

Acts as the central processing unit. It receives text input from all modules, understands user queries using NLP, and generates intelligent responses.

3) Integration Layer:

- Converts all inputs into a common text format
- Ensures smooth communication between modules
- Manages data flow across the system

4) Output Layer: The output layer provides results in accessible formats:

- Text Output: For deaf/mute users
- Audio Output: For blind users using Text-to-Speech
- Combined Output: For better accessibility

V. MACHINE LEARNING AND DEEP LEARNING TECHNIQUES IN ASSISTIVE COMMUNICATION SYSTEMS

A. Machine Learning Techniques

Machine learning is essential in processing structured and semi-structured data in the system. The main methods of ML in the system are:

- Feature Extraction: In speech processing, Mel Frequency Cepstral Coefficients (MFCC) is applied to extract significant audio features such as sound frequency and volume. In gesture recognition, hand landmark features such as finger joints and finger positions are extracted using MediaPipe, making it easier and less complex compared to image processing[12].
- Supervised Learning: The system applies a supervised approach in training models. This includes speech samples for speech-to-text conversion and gesture samples for sign language recognition.
- Classification Models: Machine learning models classify features and map them to a specific output. In hand gesture recognition, hand features are classified and mapped to text output such as alphabets or words[11].
- Real-Time Inference: After training the models, they make predictions in real-time. This allows for real-time user interaction and communication.
- Evaluation Metrics: The performance of the models is evaluated based on accuracy and prediction confidence.

B. Deep Learning Techniques

Deep learning algorithms are integrated into the system to deal with unstructured and complex information such as speech, images, and text. Some of the deep learning algorithms integrated into the system include:

- Speech Recognition (Whisper Model): A deep learning algorithm based on the transformer model has been integrated into the system to recognize speech and convert it into text. This algorithm offers high accuracy even in noisy conditions and can be applied in real-time scenarios[12].
- Convolutional Neural Networks (CNN): Convolutional neural network algorithms are integrated into the system to deal with gestures and images by recognizing features in images[9].
- Recurrent Neural Networks (RNN) and LSTM: These algorithms are integrated into the system to deal with continuous gestures and speech patterns[15].
- Transformer Models: Advanced algorithms based on the transformer model are integrated into the system to deal with speech recognition and chatbots.
- Natural Language Processing (NLP): The NLP algo- rithm has been integrated into the system to deal with queries from the end-users and generate responses accordingly[4].

C. Multimodal Integration of ML and DL

The SignVoice system integrates ML and DL for the processing of multimodal input in the following ways:

- Speech Input – Deep Learning(Whisper) – Text
- Gesture Input – Machine Learning – Text
- Image Input – CNN/OCR – Text Description
- Text Input – NLP chatbot – Response

Each of these is converted into a standard text form and is fed into the AI chatbot for processing.

International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 2, Mar-Apr-2026, ISSN (Online): 2395-566X

VI. RESULTS AND DISCUSSIONS

The proposed Sign Voice Bidirectional Communication System has been implemented, and the system has been evaluated using various real-time scenarios. The system has various modules, including the AI chatbot for the blind, sign language recognition for the deaf/mute, and the multimodal interface. The results show the system's ability to offer accurate, real-time, and accessible communication for differently abled people.

A. AI Chatbot for Blind Individuals

The module of the chatbot is designed in a way that it can be helpful for visually impaired individuals. The chatbot provides the user the facility of giving input to the chatbot in the form of speech and images taken through the camera of the device. The chatbot processes the user's queries and provides a suitable answer based on the understanding of the user's queries through natural language processing techniques. From the experimental results, it is clear that the chatbot provides accurate and contextual responses to the user's queries. This helps the chatbot to be a useful tool for the user as a virtual assistant. The image understanding module of the chatbot provides the user the facility of capturing images of the surroundings, which are then analyzed and provided to the user in the form of audio. This is quite helpful for the user to increase the level of environmental awareness as a blind user. The chatbot is efficient in performing the functions as expected. Minimal latency is experienced while generating responses for the user from the chatbot[4].

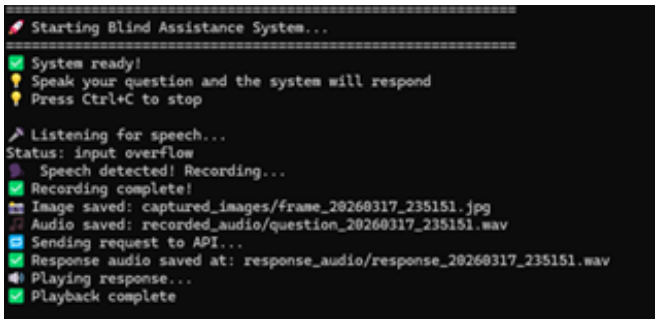

Figure. 4. AI Chatbot Interface for Blind Assistance Showing Audio Interaction and System Responses.

The primary user interface designed for blind assistance, showing the layout for audio interaction and system feedback, is presented in Fig. 4. The real-time execution of the AI chatbot, including the processing of user queries and the generation of context-aware responses, is captured in Fig. 5.

B. Sign Language to Text for Deaf/Mute Individuals

The sign language recognition module was created to recognize hand movements and translate them into text. This would help communicate with deaf and mute people. The webcam would take images of hand movements in real- time and process them using MediaPipe for hand landmark detection. The features are then used to create text using a

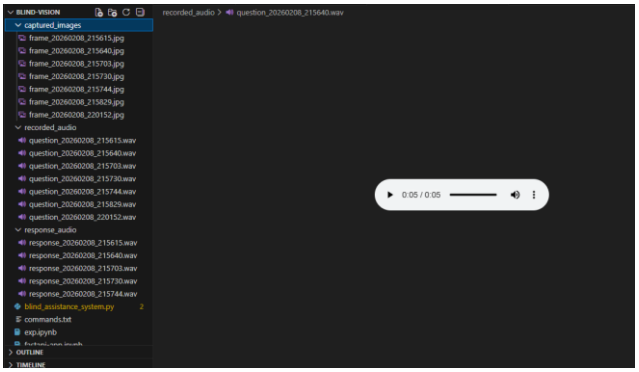

Figure. 5. AI Chatbot Execution Output Showing Captured Frames, User Queries, and Generated Responses

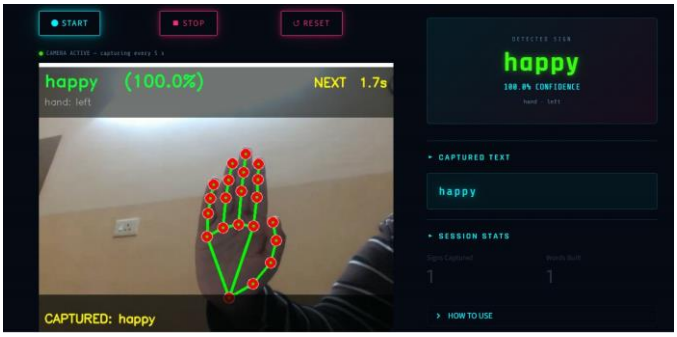

Figure. 6. Sign Language Recognition Output Showing ASL Gesture Interpreted as "Happy"

machine learning model[11]. The classification outputs for different emotional and functional signs are shown in Figs. 6 and 7.

The results show that the system is accurate in recognizing hand movements and translating them into text, provided that the background and lighting are correct. The features are extracted using landmark detection, which makes it easier and faster to

process. The translated text is displayed in real-time, and this would help communicate with deaf/mute people. The translated text would also be converted into speech to communicate with normal users. Some limitations were experienced in recognizing complex hand movements and in situations where the background and lighting are poor.

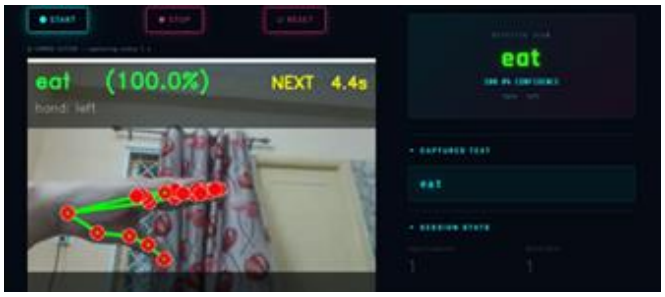

Figure. 7. Sign Language Recognition Output Showing ASL Gesture Interpreted as “Eat”

However, this module would be a good solution for sign language translation in real-time.

C. Multimodal Communication System

The multimodal communication module includes the text- to-speech, speech-to-text, and image-to-text functionality. These three functions are combined into a unified system. The multimodal communication module allows the user to interact with the system using multiple modes of input, such as speech, text, or image. It is highly flexible and user-friendly. The speech-to-text module allows the system to convert the spoken word into text. The text-to-speech module allows the system to convert text into speech. The image processing module processes the image, allowing the system to convert the image into text, which can be spoken[7]. The complete integration

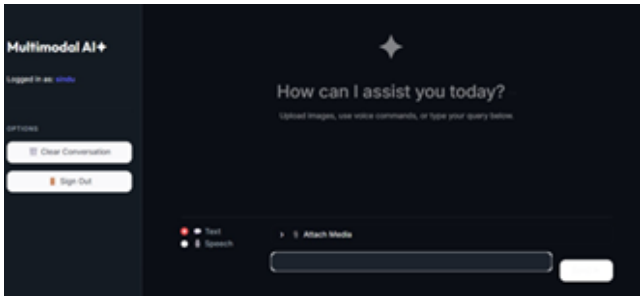

Figure. 8. Multimodal AI interface

of the various communication modules into a single, user-friendly dashboard is presented in Fig. 8, which shows the unified multimodal AI interface designed for seamless user interaction.

The results show that the multimodal system improves the efficiency of the entire communication process. It allows the user to interact with the system using multiple modes of input. It ensures the user can interact with the system despite the presence of any disability. The combination of all the modules removes the barriers of communication. It works well with the real-time system, as the response time of the system is acceptable. However, the system might be affected if the noise level of the environment, the image, or the internet speed are high. The multimodal system improves the efficiency of the system.

The performance of the audio synthesis component is demonstrated in Fig. 9, illustrating a successful Text-to-Speech conversion where the system generates clear audible output from processed text.

Fig. 10 depicts the execution of the Speech-to-Text module, showcasing the system’s ability to accurately capture vocal input and transcribe it into digital text using the Whisper model. The visual recognition capabilities of the framework are further evidenced in Fig. 11, which displays the Image-to- Text conversion process identifying objects within a breakfast spread and providing a descriptive textual summary.

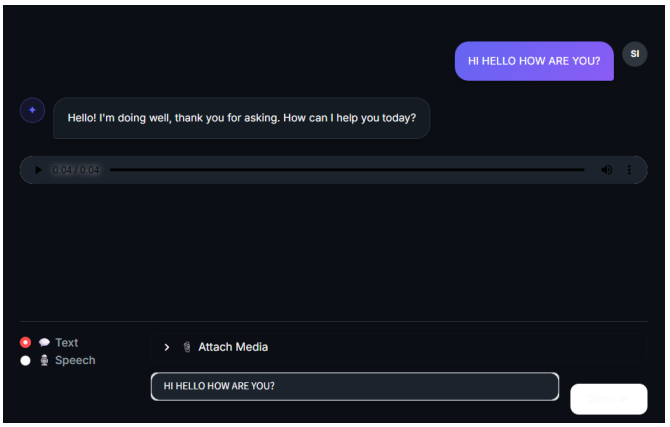

Figure. 9. Text-to-Speech Conversion

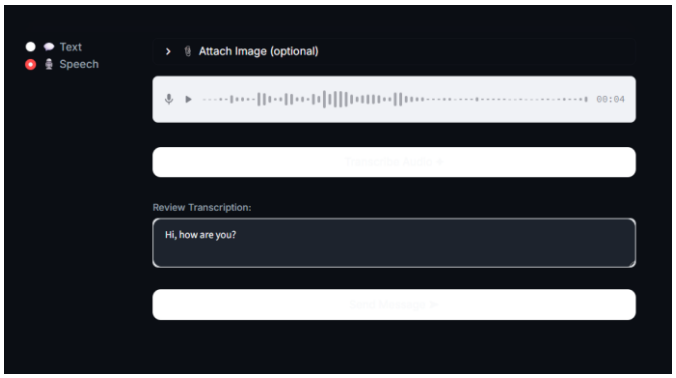

Figure. 10. Speech-to-Text Conversion

Figure. 11. Image-to-Text Conversion

VII. CONCLUSION AND SCOPE

A. Conclusion

The proposed Sign-Voice Bidirectional Communication System effectively addresses the communication problems faced by deaf, mute, and visually impaired people through an integrated and intelligent multimodal system. The system uses advanced technologies such as machine learning, deep learning, natural language processing, and computer vision to facilitate communication through gestures, speech, text, and images.

The proposed system's implementation of the sign language recognition module through hand landmark detection and machine learning effectively translates gestures into text, thus facilitating communication for deaf and mute people. The speech processing module, which uses the Whisper model, effectively translates speech into text, while the text-to-speech module generates speech from text, thus facilitating communication for visually impaired people. The integration of an AI chatbot into the system enhances its functionality through intelligent responses to user queries.

The multimodal chatbot further enhances the system's functionality through text, speech, and image-based interaction, thus making it highly flexible and user-friendly. The system's efficiency in performing tasks is ensured through the use of a hybrid approach, which combines local and cloud-based processing for efficient computation. The system's efficiency in performing tasks in real-time scenarios with good accuracy is evident from the experimental results.

In conclusion, the proposed system offers a unified, inclu- sive, and barrier-free communication system, which can help reduce the barriers of communication, foster independence, and enhance accessibility for differently-abled individuals. It represents a notable move towards the development of assistive technologies, which can offer inclusive, barrier-free, and effective communication[6].

B. Future Scope

Though the proposed system is successful in achieving its objectives, there are certain areas where enhancements and improvements are possible for better efficiency and effectiveness in real-world applications.

Firstly, the proposed system could be extended to incorpo- rate other sign languages in addition to ASL, which would en- able global usage and access for users across different regions. The gesture recognition system could also be made more efficient using more powerful deep learning techniques for better efficiency in different conditions, including illumination and background conditions.

Secondly, the proposed AI chatbot could also be made more efficient using more powerful and efficient AI techniques for better effectiveness in providing responses to users. The proposed system could also be integrated with translation technology for better efficiency in handling users who communicate in different languages.

Thirdly, the proposed image processing system could also be made more efficient using more powerful image processing techniques for better efficiency in handling visually impaired users. The proposed system could also be integrated with wearable technology for better efficiency and ease of access for users.

Furthermore, the system can also be optimized to function even without internet connectivity. Performance optimization techniques can also be employed to make the system respond faster and handle real-time processing. The system can also be enhanced to add security and privacy features to ensure that users' information is well taken care of, especially for voice and image input. The system can also be made accessible to many users through its deployment as a cloud-based system or even as a mobile application. Therefore, it is evident that the proposed system has the potential to grow and make significant contributions to the development of assistive technology systems.

REFERENCES

1. W. Y. Leong, "Personalized AI Solutions for Supporting Communi- cation Needs of Disabled Students," in Proc. 14th Int. Conf. Ed- ucational and Information Technology (ICEIT), 2025, pp. 1-6, doi: 10.1109/ICEIT64364.2025.10976060.
2. R. Jarvis, "Multimodal Robot/Human Interaction In An Assistive Tech- nology Context," in Proc. 2nd Int. Conf. Adv. Computer-Human Inter- actions (ACHI), 2009, pp. 1-6.
3. M. S. Shakib, Kamaruzzaman, M. E. Ali, R. Khisa, and S. Bhowmik, "Machine Learning Based Smart System for Blind and Visually Im- paired People," in Proc. 2nd Int. Conf. Information and Commu- nication Technology (ICICT), Dhaka, Bangladesh, Oct. 2024, doi: 10.1109/ICICT64387.2024.10839742.
4. S. Jancy, A. V. A. Mary, M. P. Selvan, and L. K. J. Grace, "An Intelligent Chatbot for the Disabled People," in Proc. 7th Int. Conf. Intelligent Sustainable Systems (ICISS), 2025, pp. 926-931, doi: 10.1109/ICISS63372.2025.11076528.
5. A. Balamurugan and D. Thiruppathi, "Artificial Intelligence-Based Chat- bot with Voice Assistance," in Proc. Int. Conf. Trends in Quantum Computing and Emerging Business Technologies (TQCEBT), Pune, India, Mar. 2024, doi: 10.1109/TQCEBT59414.2024.10545197.
6. P. Patel, S. Pampaniya, A. Ghosh, R. Raj, D. Karuppaih, and S. Kandasamy, "Enhancing Accessibility Through Machine Learning: A Review on Visual and Hearing Impairment Technologies," IEEE Access, vol. 13, pp. 33285-33307, 2025, doi: 10.1109/ACCESS.2025.3539081.
7. S. Bhat, P. Bhat, and S. V. Kolekar, "From Vision to Voice: A Multi-Modal Assistive Framework for the Physically Impaired," IEEE Access, vol. 13, pp. 128105-128121, 2025, doi: 10.1109/AC- CESS.2025.3590237.
8. S. S. Wazalwar, "Multimodal Interface for Deaf and Dumb Communi- cation," Journal of Information and Optimization Sciences, vol. 38, no. 6, pp. 895-910, Oct. 2017.
9. R. J, S. H. Nivas S, and C. Rajasekaran, "Deep Learning and Computer Vision based Smart Intercom Device for the Deaf and Mute People," in Proc. Int. Conf. Visual Analytics and Data Visualization (ICVADV), 2025, pp. 1-6, doi: 10.1109/ICVADV63329.2025.10961867.
10. V. P. A M and R. S. G. Prasad, "Domain Based Sign Language Recognition and Translation System with Multimodal Outputs," in Proc. 9th Int. Conf. Computational System and Information Technol- ogy for Sustainable Solutions (CSITSS), Bengaluru, India, 2025, doi: 10.1109/CSITSS67709.2025.11295544.
11. M. Chaudhary, S. Gaur, and R. Sharma, "Sign Language to text Con- version Using Machine Learning (For Deaf and Mute People)," in Proc. 1st Int. Conf. Advances in Computing, Communication and Networking (ICAC2N), 2024, doi: 10.1109/ICAC2N63387.2024.10895283.
12. V. M. Reddy and T. Vaishnavi, "Speech-to-Text and Text-to-Speech Recognition using Deep Learning," in Proc. 2nd Int. Conf. Edge Computing and Applications (ICECAA), 2023, pp. 1152-1157, doi: 10.1109/ICECAA58104.2023.10212222.
13. N. Sharma, "A Real Time Speech to Text Conversion System Using Bidirectional Kalman Filter in MATLAB," in Proc. Int. Conf. Advances in Computing, Communications and Informatics (ICACCI), Jaipur, India, 2016, pp. 1297-1301, doi: 10.1109/ICACCI.2016.7732224.
14. P. Mukherjee, A. Paul, S. Santra, and P. Chatterjee, "Development of GUI for Text-to-Speech Recognition using Natural Language Process- ing," in Proc. Int. Conf. Computational Intelligence and Communication Networks (CICN), 2017, pp. 1-5.
15. A. Seviappan, A. S. Reddy, K. Ganesan, B. V. Krishna, A. Anbumozhi, and D. S. Reddy, "Sign Language to Text Conversion using RNN- LSTM," in Proc. Int. Conf. Data Science, Agents and Artificial Intelli- gence (ICDSAAI), 2023, doi: 10.1109/ICDSAAI59519.2023.10444321.