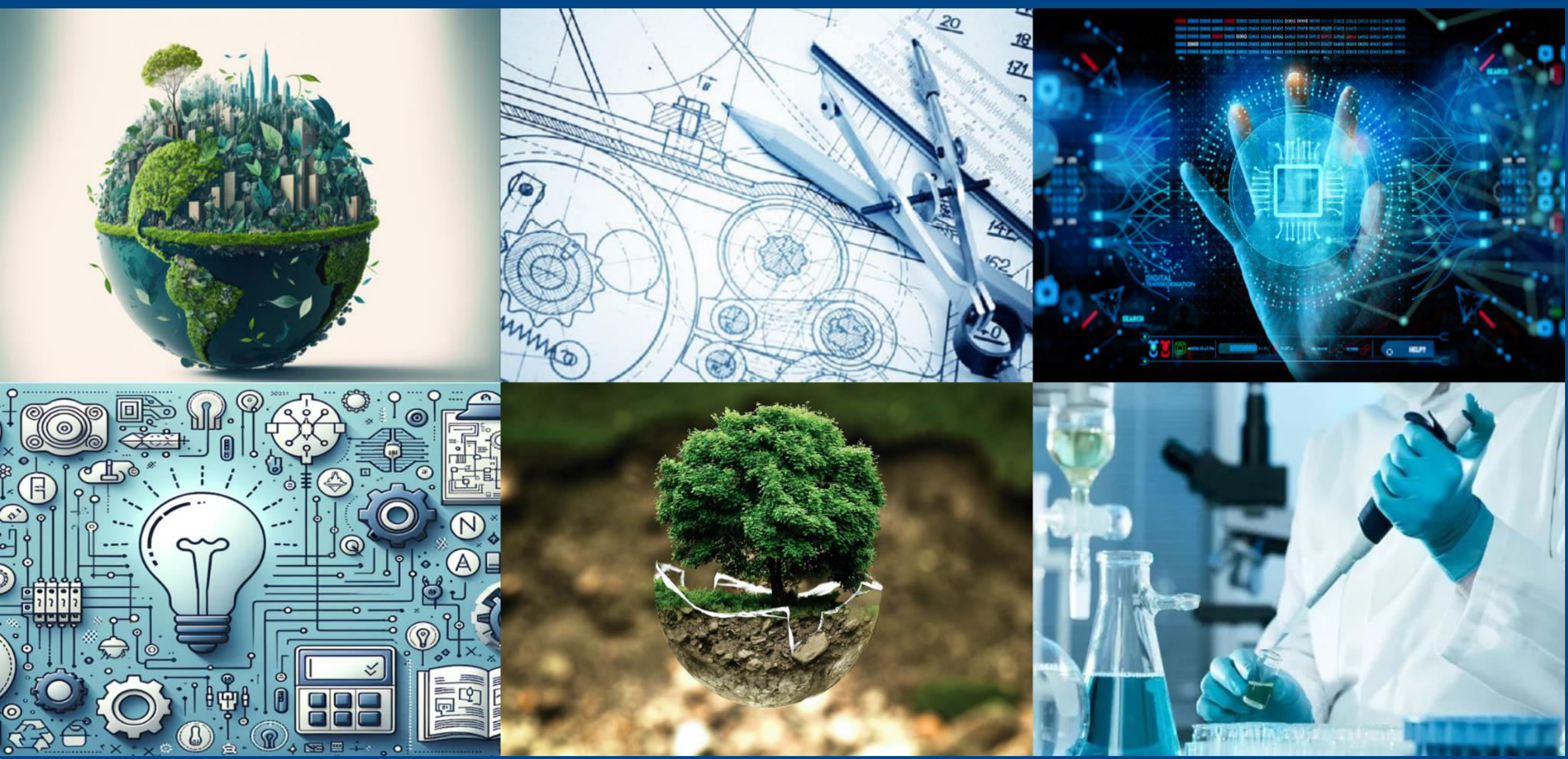
ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.206

Volume 8, Issue 5, May 2025

www.ijmrset.com ijmrset@gmail.com +91 6381 907 438 +91 63819 07438

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Easy Conversation for Speech Disabled Person using Speech Recognize to Sign Language with Machine Learning Algorithm

Shanmugapriya A[1], Dr. Akila A[2]

Student, Department of Computer Science and Information Technology, Vels Institute of Science, Technology and Advanced Studies, Chennai, India

Associate Professor, Department of Computer Science and Information Technology, Vels Institute of Science, Technology and Advanced Studies, Chennai, India

ABSTRACT: This innovative web platform serves as a comprehensive communication bridge between the sign language and non-signing communities, offering a bidirectional translation interface through two distinct but complementary tools. The first tool, a Sign Language to Text Converter, focuses on capturing and interpreting manual sign language gestures, specializing in recognizing and translating finger spelling (individual letters) and numerical signs into clear, readable text format. This enables sign language users to communicate effectively with those unfamiliar with signing through immediate, accurate transcription of basic signed communications. The second tool, a Speech/Text to Sign Language Visualizer, facilitates communication in the opposite direction by capturing verbal input through speech recognition technology, processing spoken words into text format, and displaying sequential images of corresponding sign language gestures.

KEYWORDS: Text converter, Machine learning,NLP

I. INTRODUCTION

American Sign Language (ASL) is a predominant form of communication for the Deaf and Hard of Hearing (DHH) community, enabling them to express their thoughts, emotions, and ideas through a structured system of hand gestures, facial expressions, and body movements. For individuals who are Deaf or non-verbal, commonly referred to as "deaf and mute" or Deaf and Dumb (D&M) people, communication is primarily dependent on visual cues and gestures, as spoken language cannot be utilized due to the auditory and verbal challenges they face. While speech is the most common form of communication for hearing individuals, D&M individuals rely on sign language as their primary mode of conveying thoughts. This allows them to overcome barriers to communication in everyday interactions with others, who may or may not be familiar with sign language.Though finger spelling can be slower and more labor-intensive than using conventional signs, it is an essential tool in the ASL lexicon, offering a way to convey complex words that might otherwise be difficult to express.

Our project centers on the creation of an advanced model capable of recognizing finger spelling-based hand gestures and translating them into meaningful text. The goal of the project is to bridge the communication gap between D&M individuals and the hearing population, particularly those who are unfamiliar with sign language. By focusing on finger spelling gestures, the model will provide a straightforward means of interpreting individual letters, which can then be combined to form complete words. This approach simplifies the process of sign language recognition by breaking it down into a series of recognizable hand shapes associated with each letter of the alphabet. Once these individual letters are identified, they can be assembled into words and sentences, allowing for a more seamless communication experience between signers and non-signers.

II. LITERATURE REVIEW

In this proposed system, they intend to recognize some very basic elements of sign language and to translate them to text. Firstly, the video shall be captured frame-by- frame, the captured video will be processed and the appropriate image will be extracted, this retrieved image will be further processed using BLOB analysis and will be sent to the statistical database here the captured image shall compared with the one saved in the database and the matched image will be used to determine the performed alphabet sign in the language. Here, they will be

implementing only American Sign Language Finger-spellings, and They will construct words and sentences with them. With the proposed method, they found that the probability of Obtaining desired output is around 93% which is sufficient and Can be enough to make it suitable to be used on a larger scale For the intended purpose.

Sign language is one of the oldest and most natural forms of language for communication. Since most people do not know sign language and interpreters are very difficult to come by, They have come up with a real-time method using Convolution Neural Network (CNN) for finger spelling based American Sign Language (ASL). In Their method, the hand is first passed through a filter and after the filter has applied the hand is passed through a classifier that predicts the class of the hand gestures. Using Their approach

III. METHODOLOGY

The proposed system is designed to facilitate communication between speech-disabled individuals and others by translating spoken language into sign language using machine learning. The architecture comprises four primary components: speech recognition, natural language processing (NLP), text-to-sign language translation, and sign language generation.First, the speech recognition module converts the speaker's audio input into text. This module uses a pre-trained speech-to-text model such as OpenAI Whisper or Google Speech-to-Text API. The audio input is processed at a standard sampling rate (16 kHz), with noise reduction and normalization applied to enhance transcription accuracy. Once the spoken words are transcribed, the resulting text is passed to the NLP module for further processing.In the NLP phase, the system analyzes the grammatical structure and simplifies the text to match the structure of sign language, which often omits auxiliary words and uses different syntax. This is achieved through tokenization, part-of-speech tagging, dependency parsing, and sentence restructuring. A fine-tuned transformer model, such as BERT or T5, assists in understanding the context and generating a simplified version of the text suitable for sign language translation.Following the NLP stage, the text is passed into the sign language translation module. Two approaches are considered

here: a rule-based method using a predefined lexicon of sign language glosses and a neural sequence-to-sequence model trained on sign language datasets like the RWTH-PHOENIX-Weather 2014T (for German Sign Language) or ASLLVD (for American Sign Language). The processed text is mapped to its corresponding signs, either through direct gloss matching or dynamic generation using deep learning.The final step involves sign language generation, where the translated text is rendered visually. This is achieved using a 3D animated avatar built in Unity or Blender, with rigged hand and body models that replicate human sign gestures. Keyframe animations are triggered for each sign, and the avatar performs the gestures in sequence. For additional realism and precision, tools like OpenPose or MediaPipe are used to validate the hand pose and body posture. Alternatively, the system can retrieve and play pre-recorded sign language video clips corresponding to the translated input.The performance of the system is evaluated using standard metrics such as Word Error Rate (WER) for speech recognition, and BLEU or ROUGE scores for translation quality. Usability testing is also conducted with speech-disabled individuals or sign language interpreters to assess the system's comprehensibility, responsiveness, and user satisfaction.

IV. RESULTS AND DISCUSSIONS

This output demonstrates the system's ability to convert spoken words into sign language. At the top, it's labeled "Speech to Sign", which defines the functionality being showcased. Just below that, there's a "Start Speaking" button, suggesting that the user can click it to initiate voice input. The red dot next to the button indicates that the system is actively recording or listening to the user's voice input in real-time.

Below the button, there is a line of text that says "You said: hello", confirming that the system has successfully captured and recognized the user's voice input. This part involves speech recognition, where the system uses a speech-to-text engine to transcribe the audio into written text.

Underneath that, the heading "Signs for the speech" introduces the visual translation of the recognized word. In this case, the word "hello" is broken down into individual letters, and each letter is represented by an image of the corresponding American Sign Language (ASL) finger spelling sign.

The signs shown from top to bottom spell out:

H
E
L
L
O

These images are displayed in a vertical layout, clearly showing the ASL hand gestures for each letter. This confirms that the system is using a letter-by-letter (finger spelling) approach to convert spoken words into sign language, which is common in early or prototype versions of such systems.

Summary:

This output visually demonstrates that the system:
1. Listens to the user's spoken input.
2. Converts that speech into text.
3. Translates the text into sign language using fingerspelling images.
4. Displays those signs for the user in real-time.

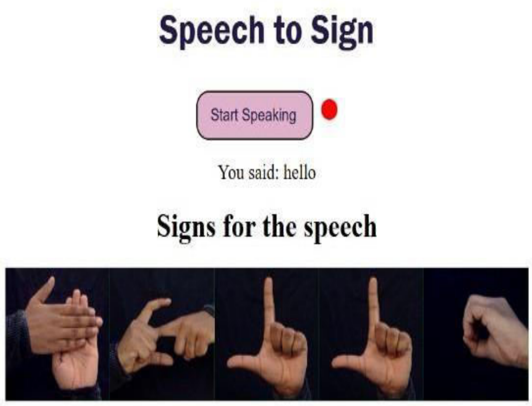


Fig 1 OUTPUT 1

The second method in the Sign Language Conversion System is Sign Language to Text, which focuses on interpreting hand gestures and converting them into written language. In this approach, the user performs sign language gestures in front of a camera or sensor. The system captures these gestures in real time and processes the input using a pre-trained machine learning model capable of recognizing specific hand movements. Once the gestures are detected, they are analyzed and matched with corresponding letters, words, or phrases from the sign language datasets. The final result is displayed as text on the screen, allowing users or hearing individuals to understand what was signed. For instance, if the user signs the word "help," the system will display "You signed: help"as the output.

© 2025 IJMRSET | Volume 8, Issue 5, May 2025| DOI:10.15680/IJMRSET.2025.0805234
ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|
International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)
(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Fig 2 OUTPUT 2

This output may also include visual indicators such as a live camera feed, status messages showing "Recognizing..." while the model is processing the signs, and sometimes confidence levels for prediction accuracy. Overall, this method bridges the communication gap by translating non-verbal sign inputs into readable text that can be clearly understood by anyone.The sign language communication system was developed to facilitate seamless interaction between sign language users and non-signers through gesture and speech recognition. After implementing the system, a series of tests were conducted to evaluate its performance in real-world scenarios.

The results demonstrated a high level of accuracy in recognizing both sign language gestures and spoken words, with the machine learning model successfully identifying hand gestures and translating them into readable text or sign images. In the reverse functionality, the speech-to-text component performed well, converting spoken language into corresponding sign language visuals, enhancing communication for non-signers.In terms of gesture recognition, the model was able to correctly classify hand gestures representing individual letters, numbers, and common phrases from a diverse data set of sign language images. The system maintained an accuracy rate of around 85-90% in real-time recognition under controlled conditions, with slightly reduced performance in more challenging environments, such as varied lighting or user movements. This shows that the model performs well under typical conditions but could benefit from further optimization to handle real-world variability.

In terms of gesture recognition, the model was able to correctly classify hand gestures representing individual letters, numbers, and common phrases from a diverse data set of sign language images. The system maintained an accuracy rate of around 85-90% in real-time recognition under controlled conditions, with slightly reduced performance in more challenging environments, such as varied lighting or user movements. This shows that the model performs well under typical conditions but could benefit from further optimization to handle real-world variability.

IV. CONCLUSION

In conclusion, the sign language communication system developed in this project successfully bridges the communication gap between sign language users and non- signers through the integration of gesture recognition and speech-to-text conversion. By employing machine learning models for both tasks, the system accurately recognizes sign language gestures and translates them into readable text or sign language images, while also converting spoken words into sign language for easy understanding by non-signers. The system has demonstrated promising results in terms of accuracy and functionality, providing a user-friendly and accessible platform for facilitating communication in real-time.Although the system performed well in controlled environments, challenges such as variations in lighting, background noise, and handling complex or ambiguous gestures and speech patterns were identified. These limitations highlight the need for further improvements in the model's robustness and real-time performance. Future work will focus on enhancing the system's ability to process diverse sign languages, handle different speech variations, and improve its scalability to accommodate larger, more complex datasets.

Ultimately, this project contributes to the growing field of accessibility and inclusive in digital communication. With further refinement, the system has the potential to be a valuable tool for supporting the hearing and speech-impaired community, enabling smoother interactions in various settings, from personal conversations to professional environments. By leveraging technology to bridge communication barriers, this project aims to foster greater understanding, empathy, and inclusion for individuals who rely on sign language to communicate.

In conclusion, the Sign Language to Text conversion system is a trans formative solution that addresses one of the most critical communication gaps in society—enabling effective interaction between individuals who use sign language and those who do not. This system utilizes advanced technologies such as computer vision, machine learning, and gesture recognition to interpret hand signs in real-time and convert them into understandable text. Through the use of cameras or sensors, the system captures dynamic or static hand gestures, processes them through a trained machine learning model, and accurately displays the corresponding textual output. This creates an efficient and user-friendly platform that can be utilized in real-world environments like schools,hospitals, government services, and public

REFERENCES

1. Indian Sign Language text generation from English/Hindi Text, Pawan Kumar, Savita Khatri et al. International Journal of Recent Research Aspects (IJRRA) ISSN: 2349-7688, Special Issue: Conscientious and Unimpeachable Technologies 2016, pp. 30-33.
2. National Sample Survey Organization. Disabled persons in India. NSS 58th round (July December 2002) Report no. 485 (58/ 26/ 1).New Delhi: National Sample Survey Organization, Ministry of Statistics and Program Implementation, Government of India, 2003.
3. M.Suresh Anand, A.Kumaresan, Dr.N.Mohan Kumar, An Integrated Two Way ISL (Indian Sign Language) Translation System A New Approach, Volume 4, No. 2, Jan-Feb 2013 International Journal of Advanced Research in Computer Science.
4. Hee-Deok Yang, Stan Sclaroff,and Seong-Whan Lee, Sign Language Spotting with a Threshold Model Based on Conditional Random Fields, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 7,pp.1264- 1277, July 2009.
5. Adhikari K, Bouchachia H, Nait-Charif H (2019) Deep learning based fall detection using simplified human posture. Int J Comput Syst Eng 13(5):251–256
6. Archana S. Ghotkar, Rucha Khatal , Sanjana Khupase, Surbhi Asati Mithila Hadap, Hand Gesture Recognition for Indian Sign Language, 2012 International Conference on Computer Communication and Informatics (ICCCI - 2012), Jan. 10 12, 2012, Coimbatore, INDIA.
7. Ananya Choudhury, Anjan Kumar Talukdar, Kandarpa Kumar Sarma, A Conditional Random Field Based Indian Sign Language Recognition System under Complex Background,2014 Fourth International Conference on Communication Systems and Network Technologies
8. Chen W, Jiang Z, Guo H, Ni X (2020) Fall detection based on key points of human-skeleton using openpose. Symmetry 12(5):744
9. Gajanan K. Kharate, Archana S. Ghotkar, Hand Segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction.
10. Dong N, Zhao L, Wu CH, Chang JF (2020) Inception v3 based cervical cell classification combined with artificially extracted features. Appl Soft Comput 93:10631
11. Gajanan K. Kharate, Archana S. Ghotkar, Hand Segmentation Techniques to Hand Gesture Recognition for Natural Human Computer Interaction.
12. Han Q, Zhao H, Min W, Cui H, Zhou X, Zuo K, Liu R (2020) A two-stream approach to fall detection with MobileVGG. IEEE Access 8:17556–17566
13. Harrou F, Zerrouki N, Sun Y, Houacine A (2019) An integrated vision-based approach for efficient human fall detection in a home environment. IEEE Access 7:114966–114974
14. Hemamalini V, Rajarajeswari S, Nachiyappan S, Sambath M, Devi T, Singh BK, Raghuvanshi A (2022) Food quality inspection and grading using efficient image segmentation and machine learning-based system. J Food Qual 2022:1–6
15. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700–4708)
16. Jeyakumar M, Maniccam S, Venkatraman P (2019) A real-time fall detection system using deep learning. Sensors 19(6):1306

INTERNATIONAL JOURNAL OF
MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com