Kannada Speech to Sign Language Converter: A Novel Platform Integrates Indian Sign Language and Speech-to-Text Technology

Pushpalatha M. N
1st Dept. of Information Science and Engineering(ISE), Ramaiah Institute of Technology (RIT), Bengaluru, Karnataka, India
(pushpalathamn@msrit.edu)

Parkavi A
2nd Dept. of Computer Science and Engineering(CSE), Ramaiah Institute of Technology(RIT),Bengaluru, Karnataka, India
(parkavi.a@msrit.edu)

Mushtaq Ahmed D M
3rd Dept. of Information Science and Engineering(ISE), Ramaiah Institute of Technology(RIT), Bengaluru, Karnataka, India
(mushtaq.ahmed@msrit.edu)

ARKODEEP KOLEY
4th Dept. of Information Science and Engineering(ISE), Ramaiah Institute of Technology(RIT),Bengaluru, Karnataka, India
(1ms20is022@msrit.edu)

ANKESH M NAIK
5th Dept. of Information Science and Engineering(ISE), Ramaiah Institute of Technology(RIT), Bengaluru, Karnataka, India
(1ms20is015@msrit.edu)

ASHISH KUMAR
6th Dept. of Information Science and Engineering(ISE), Ramaiah Institute of Technology(RIT), Bengaluru, Karnataka, India
(1ms20is027@msrit.edu)

2025 International Conference on Microwave, Optical, and Communication Engineering (ICMOCE) | 979-8-3315-6642-5/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICMOCE64100.2025.11077098

***Abstract*— India faced an enormous challenge in providing accessible educational resources for its deaf community, particularly in rural areas where specialized schools and sign language interpreters were scarce. An estimated 63 million people suffered from hearing impairments, many of whom were children with severe to profound hearing loss, highlighting the critical need for innovative solutions. This paper introduces an online platform designed specifically for this purpose, leveraging technology to facilitate kannada speech-to-text conversion and matching it with visual 3D sign language Avatar. This helps normal population who are not familiar with sign language to learn sign language independently and communicate with deaf community.**

Keywords—(ISL)Indian Sign Language, Speech-to-Text,3D Avatar, Visually Impaired.

Introduction

Sign Language aids as a vibrant mode of communication for individuals who are deaf or mute, utilizing visual gestures and signs to express meaning. In India, around 63 million people face with hearing impairments, including a significant number of children experiencing severe to intense hearing loss, the need for accessible educational resources is pronounced. However, such resources are notably lacking, especially in rural areas where specialized schools and sign language interpreters are scarce.

Knowing this persistent need, there is a call for innovative solutions that can bridge the gap in access to communication between normal population and the deaf community. In this paper, we addressed this pressing issue by introducing a novel online platform designed to close the accessibility gap for the deaf community in terms of communication and education. Kannada speech is converted to Kannada text using Google web speech API, Google cloud translate API is used for translating kannada text to English text, next parsing of English text is performed and 3D Avatar is displayed for the corresponding Indian Sign Language gestures. Solution presented in this paper helps for the normal people who are not familiar with sign language to learn independently and directly communicate with deaf community through kannada speech. It promotes the social inclusion by providing autonomous sign language learning and communication.

In response to this imperative, an online platform has been developed specifically tailored to address these challenges within the Indian context. This platform harnesses the power of technology to facilitate Kannada speech-to-text conversion and seamlessly match it with visual sign language cues, thereby normal person can learn and communicate with deaf community. [4] kannada, boasts a significant population of 44 million native speakers. The language, included in India's constitution, features a unique script with 34 consonants and 13 vowels. The Unicode system incorporates special characters and diacritic symbols to form otthakshara and conjugate letters, enhancing the language's phonetic representation. An alphabetic table in Siri-Bhuvalaya lists 64 characters representing different phonetic sounds, categorized as vowels, consonants, and semi-consonants. This work extracts 64 sounds from spoken words to represent them in the dataset. [13] Translator module that allows regular users to access sign animations that correlate to Malayalam text, and a Sign Editor module that allows professionals to contribute new words to the database using HamNoSys notation. When compared to motion capture, synthetic animation using Signing Gesture Markup Language (SiGML) is more flexible and less expensive. [2] English speech is converted to sign language 3D Avatars. The system architecture comprises three main stages: First stage involves converting audio input into text using Google's cloud speech API, tokenizing the text into words, and creating a phrase tree structure using Stanford Parser. This structure is then modified to adhere to ISL grammar rules. In second stage, data preprocessing steps are applied, including word elimination, stemming, and lemmatization, followed by ISL conversion using a transfer- based approach. The HamNoSys generation tool is utilized to obtain its notations, which are then converted into SIGML representations. Stage III involves retrieving its files from the database and generating avatar animations.

Literature Survey

[1] This research introduces a groundbreaking system that translates speech from six major Indian languages—Telugu, Hindi, Malayalam, Marathi, Kannada, and Tamil—into Indian Sign Language (ISL), aiming to improve communication for the deaf community in India. By integrating advanced speech recognition with natural language processing and techniques like MFCC, GMM, and LSTM, the system ensures precise translation. The study highlights the system's effectiveness, acknowledges current

limitations and suggests future improvements to enhance its adaptability and real-world application. Which potentially helps empowering individuals with hearing and speech impairments. [2] The system is designed to bridge the communication gap between hearing-impaired individuals in India by converting English speech into Indian Sign Language (ISL) animations. This innovative system integrates Natural Language Processing (NLP) techniques, Google's cloud speech recognizer API, and a predefined sign language database to achieve its core functionality. It represents an advancement over existing models like Text to Sign Language Translator and Audio to Sign Language Translator. Experimental results from a dataset of 101 natural speech inputs showed that It outperformed these models in avatar animation accuracy, achieving an average accuracy of 77%. It also demonstrated significant improvements in processing time, with actions performed in approximately 1 second after voice recognition and 0.7 seconds between words/characters. The system's memory- efficient approach, storing vocabularies and SIGML files in a single JSON file, sets it apart from existing systems. Its superior performance in accuracy, processing time, and memory utilization positions it as a promising solution for facilitating communication between hearing-impaired individuals and others through the conversion of English speech into ISL animations. In this work [3] "Deaf Talk" is a system that enhances communication between people with and without hearing impairments by translating between spoken language and sign language. It has two modules: one converts gestures into spoken words using Kinect sensors, and the other translates spoken words into animated sign language. The system utilizes machine learning techniques for gesture recognition and integrates Unity for text-to-speech and animation display. The methods of AdaBoost and Random Forest Regression (RFRProgress) are combined to achieve gesture recognition, guaranteeing both discrete trigger detection and ongoing progress tracking. The system is being implemented with many components: gesture recording, tagging, testing, and recognition.

[4] Automatic speech recognition is particularly challenging for languages with limited research resources, such as Kannada. This paper introduces a novel, crowd-sourced Kannada speech dataset crucial for developing an Automatic Kannada Speech Recognition system. Kannada, spoken by 44 million people in Karnataka, India, has a unique script with 34 consonants and 13 vowels. The dataset captures 64 distinct phonetic sounds based on these characters. While previous research on Kannada speech recognition focused on continuous speech and dialect identification with around 70% accuracy, this work aims to improve upon those methods by expanding the dataset. [5] This study develops educational tools for students with special needs, particularly those with hearing impairments, by introducing a signing avatar for Turkish Sign Language (TID). This avatar aids communication through visual gestures, making it a more practical alternative to text-based materials. The study found that the avatar was effective in helping students learn specific sign language words. This approach is crucial for deaf children's cognitive development, as they often face language and social challenges that can impact their literacy learning in primary school. [6] This research addresses communication barriers faced by the deaf-mute community in India by developing a real-time solution using Indian Sign Language (ISL) gestures. With ISL recently standardized, the study focuses on conveying essential Covid-19 information through these gestures. The research aims to improve communication accessibility for individuals with hearing and speech impairments, highlighting ISL's importance not only for the deaf-mute community but also for those with other disabilities like Autism and Down syndrome. [7] This study introduces a human interface framework that translates spoken language into Indian Sign Language (ISL) gestures, aiming to facilitate real-time communication between the deaf-mute community and the general population. Using the Microsoft Xbox Kinect 360 for motion capture and Unity3D for animation, the system is integrated into an Android app. With around 63 million people in India experiencing significant hearing disabilities, and many lacking ISL proficiency due to a shortage of resources, this system addresses communication barriers, promoting inclusivity and effective interactions for those with hearing and speech impairments. [8] This paper presents a real-time Sign Language Identification System using an American Sign Language (ASL) dataset and Convolutional Neural Networks (CNN). The system translates static ASL hand signals into written text and integrates an Android app for text-to-speech conversion. By leveraging CNN, K-Nearest Neighbors (KNN), and You Only Look Once (YOLO) algorithms, it enhances gesture recognition across various applications, improving communication for individuals with speech impairments. This technology offers a comprehensive solution to bridge communication gaps and facilitate effective interaction with the general public.

[9] The paper explores a translation system architecture that converts English sentences into sign language using computational methods. It outlines the process of transforming linguistic structures into Discourse Representation Structures (DRS) and Attribute Logic Engine (ALE) Head-driven Phrase Structure Grammar (HPSG) to produce sign language expressions. The system incorporates signing space, syntactic considerations for nominal and verbs, and generates HamNoSys phonetic forms for signs. It also covers creating realistic virtual human animations using avatars and morphs, highlighting the importance of three-dimensional signing space. The paper discusses the challenges of automatic sign language translation, including the need for large sign language corpora, and showcases applications of synthetic animation technology for sign language interpretation.

[10] The paper presents an advanced system designed to enhance communication between people with hearing impairments and those who can hear. The system uses 3D avatar technology to convert spoken language or text into Indian Sign Language (ISL) gestures. It consists of three main components: the speech/text processing module, which translates input into English phrases using Natural Language Processing (NLP); the ISL sentence translation module, which converts these English phrases into ISL sentences; and

the 3D avatar module, which animates the ISL gestures. This comprehensive approach aims to bridge communication gaps and improve interaction across different language modalities.
[11] The proposed system aims to automate the translation of spoken English into Indian Sign Language (ISL) animations to channel communication gaps between hearing impairments people and the normal population. It integrates natural language processing (NLP) techniques, Google's cloud speech recognition API, and a sign language gesture database. The system operates in four stages: recording speech with Python PyAudio, converting it to text using Google's Speech-to-Text API, tokenizing and processing the text with NLP tools, and matching the text to sign language videos from the ASL Video Dictionary. The final output is a continuous video of sign language gestures created using Python's moviepy package. The system was tested with 45 audio samples, and its accuracy was measured based on the ratio of correct to total operations.
[12] The proposed system develops an automated software solution to convert sign language gestures into audible speech and text, aiming to bridge communication gaps between hearing or speech-impaired individuals and the general population. It integrates image processing, machine learning classifiers, and speech synthesis.

Workflow

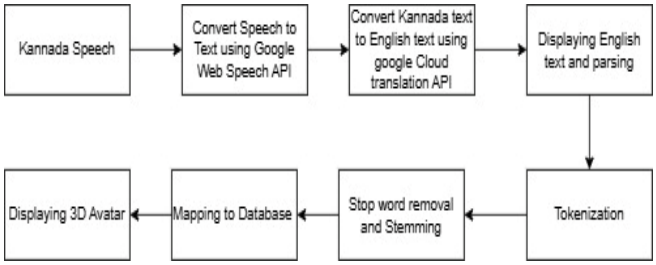

Fig. 1.1: The System Workflow

The steps of Fig.1.1 are explained as below
Input Kannada Speech: Kannada speech input is received through microphone. User input can be a kannada word or sentence.
Convert Kannada Speech to Text using Google Web Speech API: kannada audio input is converted into kannada text using Google Web Speech API. We Used python speech_recognition library recognizes kannada speech and convert to English text.
Convert Kannada text to English text using Google cloud translation API: Google Cloud Translate_v2 API is used for translating Kannada text to English text.
Displaying English text and Parsing: Converted English text is displayed on the front end and parsing is performed to understand the text's structure and meaning of sentence, which involves syntactic and semantic analysis. Using Natural Language Processing(NLP) parsing quality can be enhanced.
Tokenization and Stopword Deduction: Tokenization segments the parsed text into individual words or tokens, facilitating linguistic analysis. Common stop words like an, at, in, the etc. are removed to emphasis on content-carrying words.

Lemmatization: In this step, words are condensed into root or base forms for identification of accurate keywords. Different forms of words are standardized to a single base form for consistent analysis.

Mapping to Database: The system uses an Indian Sign Language (ISL) dataset, the keywords identified in the earlier steps are matched with entries in the dataset, mapping keywords to their corresponding signs or gestures which is displayed through avatar.

Visual Output: Each word sign is displayed through 3D avatar,if 3D avatar is not available for the word then it breaks the word into character and 3D avatar for each character will be displayed.

Algorithm: Conversion of Speech to Kannada Sign Language

Input: Speech input in Kannada

Output: Indian Sign Language representation

1. kannada speech input (audio) taken using a microphone.
2. Kannada speech is recognized and corresponding kannada text is generated using Google Web Speech API.
3. Translate kannada text to English text using Google Cloud Translate_v2 API.
4. Display converted English Text on the Front end.
5. Parsing English Text.
6. Tokenization and Stopword Removal.
7. Keyword Identification through Lemmatization.
8. Generating Visual Output Based on the matched entries from the ISL database.

Results and Discussions

3D Avatar is created using Blender, where intricate hand movements and gestures are meticulously designed. These gestures are then animated to represent various words, letters, and commonly used phrases in sign language. Once the animations are generated, they are collected into a dataset. English alphabet and frequently used words are included in the dataset. The user-friendly interface is created which helps for easy start. By a simple click on the microphone icon Kannada audio input is taken. Google Web Speech API, which is free API is used for speech recognition and kannada text generation. Kannada text is translated into English text using Google Cloud Translate_v2. Linguistic analysis will be performed on the English text after translation. Tokenization is performed on the English text, stop words are eliminated and needed keywords are identified. For knowing the context and

extracting the meaningful information this step is very important. Identified keywords undergo lemmatization, reducing words to their base forms. This linguistic refinement ensures precision in understanding and processing the content.

The initial stage website is shown in Fig. 4.1.

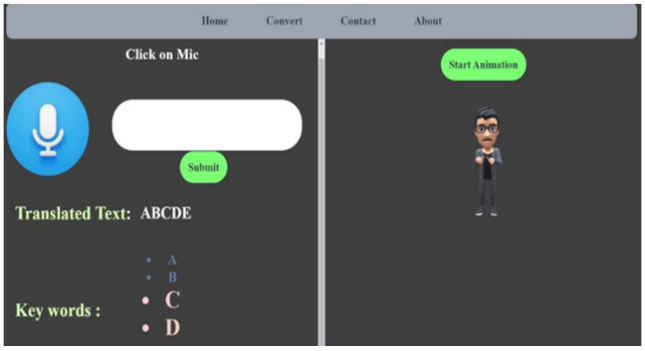

Fig. 4.1 : Website in initial stages

ISL library, the system intelligently supplements the communication with an image representation of the word.

In the example presented in Fig. 4.2, the word 'India' is associated with an image of the India map. This strategic inclusion of visual representations enhances the comprehensibility of the communication, ensuring that users receive meaningful feedback even for words without direct sign language counterparts.

Fig. 4.2: Images for some common words

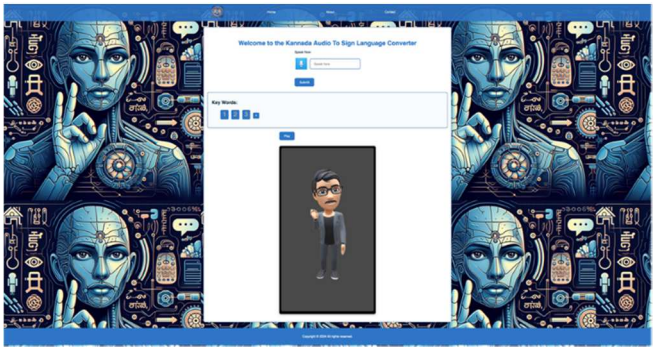

Fig. 4.3: Evolved Website

We have improved the feel and functionality of the website as shown in Fig. 4.3 and also working on taking feedback from specially-abled schools, organization and people concerned about deaf and dump community.

Avatar animation accuracy are represented Z_a in equation 3.1. The time it takes to complete an action following speech recognition to sign display through avatar for single word is around 1 second which is denoted with T_a [13].

$$Z_a = (\xi \times 100)/ N \quad (3.1)$$

ξ, and N stand for the action correctly performed by the avatar, and the total number of examples respectively. The speaker delivers a variety of brief and lengthy remarks in kannada and datasets available is used to make it work.

The animation accuracy is Za=85% we got for proposed work i.e. Kannada Speech to Indian Sign Language (KS2ISL). In earlier approaches [14] [2] T2ISL (text to Indian Sign language) achieved accuracy is 48% and ES2ISL (English Speech to Indian Sign Language) achieved accuracy is 85%, In the current work also we achieved same accuracy for kannada speech to sign language animation (KS2ISL).

Conclusion and future work

Conclusion

The research paper introduces an innovative interactive system developed to facilitate communication between normal people and hearing impairments people. This system effectively translates kannada speech into dynamic 3D avatar animations that demonstrate Indian sign language gestures. Unlike traditional methods, this novel approach utilizes realistic 3D avatars, significantly enhancing user engagement and improving memory retention.

Future Work

Future works include the development of a customized speech recognition system specifically tailored to the unique needs of the target user group. Reverse of speech to sign language conversion, i.e., sign language to speech will be implemented for two-way communication.

Acknowledgement

This project is supported by Faculty RIT Seed Funding (029/23-24), Ramaiah Institute of Technology(RIT). Bengaluru-54.

References

[1] J. Peguda, V. S. S. Santosh, Y. Vijayalata, A. D. R. N, and V. Mounish, "Speech to sign language translation for Indian languages," in 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS), vol. 1, 2022.

[2] B. D. Patel, H. B. Patel, M. A. Khanvilkar, N. R. Patel, and T. Akilan, "Es2isl: An advancement in speech to sign language translation using 3d avatar animator", in 2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), 2020, pp. 1–5.

[3] M. Ahmed, M. Idrees, Z. ul Abideen, R. Mumtaz, and S. Khalique, “Deaf talk using 3d animated sign language: A sign language interpreter using microsoft’s kinect v2,” in 2016 SAI Computing Conference (SAI), 2016, pp. 330–335.

[4] P. N and S. Kini, “Phoneme based kannada speech corpus for automatic speech recog nition system,” in 2022 IEEE International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), 2022, pp. 1–5.

[5] R. Yorganci, A. Kindiroglu, and H. Kose, “Avatar-based sign language training interface for primary school education,” 09 2016.

[6] G. Gunasagari, N. Poornima, A. Yaji, M. Achuth, A. Dsilva, and R. Chethana, “Keller: Indian sign language based gesture to text and speech assistant for covid-19 protective quotes,” 06 2022, pp. 241–247.

[7] H. Monga, J. Bhutani, M. Ahuja, N. Maid, and H. Pande, “Speech to indian sign language translator,” 12 2021.

[8] V. Mane, T. Adsare, T. Dharmik, N. Agrawal, D. Patil, and P. Atram, “Sign language translator for speech impaired people,” in 2023 5th Biennial International Conference on Nascent Technologies in Engineering (ICNTE), 2023.

[9] Elliott, R., Glauert, J.R.W., Kennaway, J.R. et al. Linguistic modelling and language-processing technologies for Avatar-based sign language presentation. Univ Access Inf Soc 6, 375–391 (2008). https://doi.org/10.1007/s10209-007-0102-z.

[10] Das Chakladar D, Kumar P, Mandal S, Roy PP, Iwamura M, Kim B-G. 3D Avatar Approach for Continuous Sign Movement Using Speech/Text. Applied Sciences. 2021; 11(8):3439. https://doi.org/10.3390/app11083439.

[11] R. Bharti, S. Yadav, S. Gupta, and R. B, “Automated speech to sign language conversion using google api and nlp,” in Proceedings of the International Conference on Advances in Electronics, Electrical & Computational Intelligence (ICAEEC) 2019, June 2019.

[12] P. C. Narvekar, S. Mungekar, A. Pandey, and M. Mahadik, “Sign language to speech conversion using image processing and machine learning,” 2020. [Online]. Available: https://api.semanticscholar.org/CorpusID:245984165.

[13] M. S. Nair, A. P. Nimitha, and S. M. Idicula, “Conversion of malayalam text to indiansign language using synthetic animation,” in Conversion of Malayalam Text to Indian Sign Language Using Synthetic Animation, Kottayam, India, 2016, pp. 1–4.

[14] D. Singh, “Text to sign language,” Thapar University, 2016