JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A Survey on Random Forest for Smart Communication between Deaf and Mute Individuals

Vandana Bonu
Department of Information Technology
Assistant professor
GMR Institute of Technology, GMR nagar, Rajam, Vizianagaram, AP, India
vandanabonu4@gmail.com

Abstract : Communication is crucial for human interaction and independence, yet deaf and mute individuals face challenges with conventional methods. While sign language is an effective medium, its limited understanding in broader society poses barriers to inclusivity.By converting written or spoken sign language into sign language and vice versa, the Smart Communication System seeks to close this gap. This system uses AI, machine learning, and natural language processing to recognize, track, and classify hand movements in real time. It uses tools such as MediaPipe, OpenCV, and the Random Forest algorithm. OpenCV analyzes video frames for gesture identification, while MediaPipe guarantees reliable keypoint recognition. Random Forest offers precise categorization at little computational cost. By translating motions into speech or text, the device makes it easier to communicate with others who can hear. Developed to be portable and easy to use, this technology has potential uses in customer service, education, and healthcare, enabling deaf and mute people to participate in everyday activities.

IndexTerms - **American Sign Language, Random Forest, deep learning, deep neural networks, gesture recognition, sign language to text**

I. Introduction

In order to facilitate smooth communication between the deaf and mute and the general public, a "Smart Communication System for Deaf and Mute Individuals" is a cutting-edge technology solution.The growing relevance of such systems is evident from the extensive research into real-time sign language translation using various state-of-the-art techniques. Sign language, which relies on hand and facial expressions for communication, serves as a vital medium for people with DHL. Beyond its role in interpersonal communication, the analysis of hand kinematics also finds applications in human–machine interaction and broader human activity recognition (HAR) tasks, emphasizing its multidisciplinary potential.The amount of research being done on real-time sign language translation utilizing a variety of cutting-edge methods shows how important these systems are becoming. People with DHL use sign language, which is based on hand and facial motions, as a critical communication tool. The analysis of hand kinematics has transdisciplinary potential as it is used in human–machine interface and more general human activity recognition (HAR) activities in addition to interpersonal communication.[1]It can be difficult to find a reliable and professional sign language interpreter everywhere. However, a workable option that can be implemented in multiple locations is a human-computer interface system for sign language translation. Such an application's potential to greatly facilitate social integration and increase awareness of the difficulties experienced by deaf people is what inspired its development.The ability of human eyesight to recognize motions is particularly remarkable in the way that hearing and deaf people interact with one another through sign language. Through the proposal of a long-term solution to enable communication between deaf people and those who are not familiar with sign language, this study tackles one of the major social issues.

Indian, American, Chinese, and Arabic sign languages are only a few examples of the various regional variations of sign language. Numerous nations have carried out in-depth studies on image processing, pattern recognition, and hand gesture identification in an effort to improve applications and attain peak performance.[2]In order to simplify processing and reduce computing complexity, image processing is essential for converting identified sign language gestures into grayscale images. This technique includes taking gesture photos, preprocessing them to improve their quality and consistency, and converting them to grayscale to preserve important texture and shape features. After that, important gesture characteristics are highlighted using feature extraction techniques before the photos are fed into Convolutional Neural Networks (CNNs) for interpretation. For those who struggle with speech and hearing problems, this preliminary step ensures accurate recognition of sign language motions, permitting effective communication.[12]

Literature survey:

This research aimed to analyze ,the advancements in smart communication systems developed to assist deaf and mute individuals in overcoming communication barriers. The design and implementation of these systems leverage technologies such as machine learning, natural language processing, wearable sensors, and mobile applications to translate spoken or signed language into text or speech. The literature in this field is vast and covers a range of approaches that focus on making communication more seamless and accessible for users.

The paper highlights the increasing need for effective Sign Language Recognition (SLR) systems due to the projected rise in individuals with disabling hearing loss (DHL) according to the World Health Organization (WHO). This underscores the necessity for real-time communication tools to support effective interaction. Various methodologies have been explored in SLR, including the use of deep neural networks (DNNs) and multimodal approaches. For instance, a mixed CNN-LSTM model has achieved notable accuracy; however, many current systems struggle when encountering unseen data. The literature also indicates several significant challenges in SLR, such as variability in gestures, image noise, and the need for specialized algorithms for segmentation and classification—factors that complicate real-time processing and accuracy. To address these challenges, the paper proposes a novel approach that combines a DNN for handshape classification with a bidirectional LSTM for spelling correction. This dual-stage system aims to enhance both accuracy and robustness, particularly in real-world conditions with unseen data. Experimental validation conducted with a user community demonstrates that this proposed system outperforms existing methods in terms of accuracy and processing speed, offering valuable insights to advance the field of SLR.[1]

This research aimed to analyze the order to accurately recognize gestures, real-time hand gesture detection entails recording video, splitting it up into frames, and then extracting features like SIFT and Difference of Gaussian. Archana S. Ghotkar et al. created a technique for Indian Sign Language Recognition that combined a Genetic Algorithm with Camshift and HSV models; nevertheless, there were significant obstacles due to the Genetic Algorithm's time-consuming nature and incompatibility with MATLAB versions. Another labor-intensive method for detecting gestures in Indian Sign Language is a 7-bit orientation process that requires six modules, as suggested by P. Subha Rajan and Dr. G. Balakrishnan. T. Shanableh's approach for Arabic Sign Language Recognition uses K-Nearest Neighbors (K-NN) for classification after users wear gloves for color segmentation to separate hand motions polynomial networks as well. Furthermore, Byung-woo Min and colleagues investigated both static and dynamic gesture recognition without the use of external devices. They did this by utilizing Hidden Markov Models (HMMs) for dynamic gestures and image moments for static movements. This review of the literature emphasizes the variety of methods and technological difficulties in gesture identification, particularly for various sign languages.[2]

This research aims to provide the several different methods for recognizing American Sign Language (ASL), with a distinction between sensor-based and vision-based techniques. Although vision-based techniques are frequently more affordable and available, sensor-based techniques have historically achieved superior accuracy in sign language identification because of their accuracy in capturing hand motions. Hand tracking methods have been thoroughly investigated, and Kinect sensors have been a popular option. Previous studies utilizing Kinect revealed that using dynamic programming matching, they could recognize letters with accuracies of up to 98.9%. Significant progress has lately been made thanks to developments in computer vision and deep learning, including convolutional neural networks (CNNs). A study that used the InceptionV3 model, for instance, demonstrated the efficacy of deep learning by achieving 90% validation accuracy on a dataset of ASL characters.The present study estimates hand joints from webcam images using the Mediapipe hands algorithm, demonstrating the evolution of feature extraction techniques. By creating features based on joint angles and distances, this technique enhances recognition performance. On the Massey dataset, the suggested methodology in this study obtained 99.39% accuracy, demonstrating the promise of leveraging camera inputs for precise and easily accessible ASL detection. All things considered, this assessment highlights the development and efficacy of different approaches in ASL recognition, showing the shift from sensor- dependent methods to sophisticated vision-based systems that use deep learning for increased accuracy and practicality.[3]

In this work, there are several methods presented in the literature on Indian Sign Language (ISL) recognition that are intended to help deaf-dumb people communicate effectively. Aleem et al.'s and other gesture recognition systems use a speech synthesis module to turn motions into sound and a gesture-to-text dictionary to match input gestures with a predefined gesture database. Both static and dynamic motions have been attempted to be recognized; Byung-woo Min et al. concentrated on visual gesture detection on a 2D plane without the use of external devices. For static gesture recognition, they used picture moments, while for dynamic gesture recognition, they used Hidden Markov Models (HMMs). Also investigated are user-independent recognition methods, such the one suggested by T. Shanableh for isolated Arabic sign language motions.This method used K-Nearest Neighbors (K-NN) and polynomial networks for classification, and gloves for color segmentation to separate hand motions.The use of specialized equipment, which can be expensive and affect the naturalness of communication, is a major obstacle in ISL recognition. Furthermore, compared to the wealth of research on other international sign languages, such as American and Japanese Sign Language, there are notably few contributions that are unique to South Indian languages. Creating systems that enable human-computer interaction is another emerging field of interest that has the potential to significantly help deaf-dumb people and address the lack of certified sign language interpreters. This review of the literature emphasizes the variety of methods and difficulties in ISL recognition, highlighting the need for easily accessible, device-independent solutions as well as additional study.[4]

This study proposes a unique technique, despite the fact that Sign Language Recognition (SLR) has been researched for more than 20 years, SLT has only lately started to garner significant attention. This is indicative of a growing desire to bridge the gap between spoken language translation and sign language. Because sign language is multifaceted and depends on both manual and non-manual cues, such body language and facial expressions, it presents special difficulties for translation systems. Because SLT may not directly fit typical machine translation models, this intricacy makes it more difficult to build efficient translation techniques.Additionally, the lack of standardization in the transcription of sign language into textual glosses creates an additional layer of complexity because it is difficult to create consistent training datasets due to the heterogeneity in video frames reflecting a single gloss. The amount and vocabulary of existing datasets are frequently restricted, which is a major barrier to the development of SLT technologies. In order to overcome these obstacles, the authors develop Transformer networks and Spatial-Temporal Multi-Cue (STMC) networks, which constitute a methodological breakthrough in enhancing SLT performance. In order to create more reliable and efficient translation systems, it is critical to address the particular difficulties of sign language, as this review emphasizes.[5]

In this article, the communication gap between the general public and deaf and mute people is examined, with a focus on the societal issues experienced by the 1.7 million hearing-impaired and silent people in India. Although numerous communication aids have been studied in previous research, this work focuses on a full two-way communication system that makes use of mobile technology to make it efficient and accessible for users. The two main components of the suggested system are the first module, which uses the Google Speech-to-Text (STT) API to translate real-time audio speech into text and map the text to Indian Sign Language (ISL) gestures. The second module, which enhances interaction by translating natural language input into animated ISL gestures.The approach makes use of cutting-edge strategies like correlation-based mapping, which guarantees precise gesture conversion and recognition—both essential for efficient human-computer interaction. In order to further improve communication effectiveness, the report also makes recommendations for future enhancements, such as support for more sign languages and emotion recognition features. This review of the literature emphasizes how creatively the suggested method uses technology to break down obstacles to communication.[6]

The study emphasizes the value of efficient communication techniques by offering a thorough literature review on the developments in assistive technologies for people with speech impairments. In order to enhance gesture detection and recognition—which is crucial for precisely interpreting hand movements—a number of gesture recognition approaches are being investigated, including Region-based Convolutional Neural Networks (R-CNNs) and Hand Gesture Recognition based on Computer Vision (HGR-CV). Furthermore, the application of machine learning (ML) to pattern recognition models is covered, emphasizing the function of feature extraction and data normalization in the analysis of hand and wrist motions. The evaluation also highlights the importance of translating sign languages into regular English in order to improve communication between hearing and deaf people.Additionally, the Automatic Behavioral Analysis using a Gesture Detection Framework (ABA-GDF) is presented as a promising method that offers great precision and flexibility for translating the activities of deaf and mute people. The literature emphasizes the usefulness of these technologies in improving social integration and communication for the deaf and mute community by pointing to real-world applications.[7]

Sign languages are acknowledged as natural languages with distinct features that are similar to spoken languages. Beginning in the 1970s, linguistic research on Indian Sign Language (ISL) confirmed that it was native to the Indian subcontinent. Between 1977 and 1982, research conducted by Vasishta and others produced dictionaries for regional variants of ISL, which were then given to programs that assist the Deaf community in India. This endeavor demonstrated a methodical approach to ISL documentation and promotion.Nonetheless, the literature emphasizes the major obstacles that deaf people in rural locations must overcome. Since community connections are necessary for the development of a strong sign language, a lack of interaction with other deaf people hinders language development.The United Nations Convention on the Rights of Persons with Disabilities, which was adopted in 2008 with the goals of recognizing the Deaf community's cultural identity, promoting the use of sign language, and elevating their status, was a significant step in the advancement of their rights in India.Additionally, the research distinguishes between established sign languages and home sign systems. In general, home sign systems used by isolated deaf people are less complex linguistically than developed sign languages used by Deaf communities. This difference emphasizes how much more research is required to fully comprehend the special facets of communication among solitary deaf people.[8]

The purpose of the proposed SMART video-based sign language application is to help deaf and mute people communicate more easily by enabling more fluid hand gesture interactions. It highlights the necessity for efficient communication systems by addressing the everyday difficulties faced by disabled people, especially when attempting to communicate with others. This application incorporates features intended to help deaf-mute people engage with the wider community more successfully, in contrast to many other apps that only concentrate on learning or recognizing sign language.Its capacity to translate sign language into text and audio is a crucial component of the software, thus providing a "voice" to people who are mute. The literature also covers mobile devices and sensor-based systems that are utilized in the app to record and convert sign language gestures into formats that are accessible.To make learning more interesting, the app also has educational features like games and quizzes that teach deaf and mute kids vocabulary on a daily basis. The focus is on developing an intuitive user interface that facilitates smooth communication between deaf people and non-sign language users. The creative use of technology to improve communication for the deaf and mute communities is highlighted in this survey.[9]

The study discusses the substantial communication barriers that people with speech and hearing impairments encounter, stressing the difficulties they have expressing themselves verbally and the necessity of finding workable solutions to close these gaps. For deaf and mute people, sign language is the primary means of communication. To encourage inclusivity, efforts are being made to translate sign language into voice and text. The use of Convolutional Neural Networks (CNNs) to interpret sign language motions is a noteworthy technical integration covered in the research, illustrating how deep learning might improve communication accessibility.In order to create this technology, the authors describe a methodical approach to dataset building. This includes 300 photos representing 8 types of sign language actions. The CNN model cannot be successfully trained or tested without this organized dataset. Adding an emergency communication option is a creative way to improve user safety by allowing users to signal for assistance in dire circumstances. In general, the project seeks to promote a more inclusive society by dispelling myths and empowering people with disabilities via better communication techniques.[10]

The study discusses the difficulties Deaf and Dumb (D&D) people have communicating, highlighting the necessity of methods that successfully close the communication gap between them and the hearing community. It emphasizes how sign language serves as D&D people's main means of communication and how crucial it is to create a translator in order to promote comprehension. In order to facilitate communication between D&D users and others without the need for an interpreter, the suggested system seeks to create a bi-directional communication framework that can identify hand gestures and translate them into text and audio.With an emphasis on American Sign Language and the application of smart gloves for gesture detection, the literature review showcases developments in sign language recognition. In order to achieve high accuracy rates, the article outlines an approach that uses Leap Motion technology to recognize hand movements and an Android application to translate speech to text. In the end, the system is regarded as an important step in improving contact and lowering obstacles to communication between D&D people and the hearing community.[11]

The study examines the creation of MuteSpeak, a two-way electronic communication system principally controlled by an Arduino, to help people with speech problems. This system's key component is a glove with flex sensors that record hand movements and convert them into text and audio outputs so users may communicate efficiently. To improve the system's usefulness, an SD card is also included to save voice commands for playback.The research highlights the importance of hand movements in communication and suggests a wearable glove that can read sign language by using a variety of sensors that can identify words and letters.The paper examines current communication strategies for the deaf and hard of hearing and identifies shortcomings in these strategies, including the need for interpreters and text writing, which can be expensive and invasive at times. Additionally, the study discusses the difficulties that the mute population faces, especially when sign language is not universally understood.This research attempts to fill the gap in flexible, affordable sign language communication systems identified by the literature review by putting forward a more approachable option. The groundwork for the novel approach suggested in the study is established by this thorough analysis of related publications.[12]

The study emphasizes the value of sign language for those who are silent, noting that more than 80 million people in China alone use it to communicate. Even yet, a lot of silent people have trouble communicating with people who don't know sign language. Historically, vision-based methods for hand motion recognition have relied on cameras that are sensitive to external factors and need users to remain within camera range. A more dependable method is suggested in this work, which makes use of flex sensors that can recognize hand movements without the requirement for visual input.Flex sensors modify resistance according to the amount of bending, functioning as variable resistors. These analog inputs are ideal for gesture detection since they are transformed into digital signals by a microcontroller, like an Arduino.The technology has drawbacks, such as decreased accuracy in some directions, but it also has benefits, such as portability and simplicity of use. In addition to facilitating communication by translating hand gestures into speech signals, the suggested system has a safety function that allows users to send emergency messages to specified contacts. The project's ultimate goal is to improve silent people's communication skills so they can engage with the outside world more readily, which will increase their self-esteem and independence.[13]

The literature study follows the development of sign language recognition systems, starting with the VPL data glove, which was first released in the 1970s and had a 5Hz recognition rate of up to 45 American Sign Language (ASL) signals. Advances in technology led to the development of the SmartGlove, which was intended to record skeletal motion for gesture recognition. However, due to its weight and expense, this device was eventually replaced by vision-based systems.Although sensor- based systems witnessed tremendous advancements as well—a prominent translator from Cornell University used flex sensors and accelerometers to record 3D movement data—they were still cumbersome and costly in comparison to new vision-based options.A system that used a Dynamic Vision Sensor (DVS) camera was able to recognize bare hand motions in real-time, but its uses were restricted. Several categorization methods were investigated to increase accuracy, such as Support Vector Machines (SVM) and Local Contour Sequence.Using

Convolutional Neural Networks (CNNs), the suggested system in this paper is a portable sign language converter that can add new signs as needed and attain an accuracy of about 99%. The development from large sensor-based solutions to more effective vision-based methods—which led to the extremely accurate CNN-based converter—is highlighted in this literature review.[14]

The study discusses the significant communication barriers that deaf-mute people encounter, pointing out that only those who have received sign language training can interact with them successfully, which frequently results in social exclusion and emotional expression issues. It talks about how cloud computing and the Internet of Things (IoT) may improve healthcare and communication by enabling safe data management and flow, which can greatly enhance the lives of people with disabilities.Advances in natural language processing (NLP), particularly in speech-to-text and text-to- speech conversions, which may mimic human interaction and produce contextually relevant responses, are also highlighted in the study. The suggested solution makes use of sophisticated gesture detection technology to record delicate hand motions, which are subsequently translated into speech by Amazon Polly, a text-to-voice service that is based on deep learning.For those who struggle with communication, the system also has health monitoring features that enable caretakers to offer prompt support, enhancing their general quality of life. This literature review summarizes the key ideas and technological underpinnings that the study builds upon in order to meet the communication demands of those who are deaf-mute.[15]

VARIOUS METHODS ARE USED IN SMART COMMUNICATION

This project is broadly divided into four modules:

1. Speech to Text conversion
2. Sign language to Text conversion
3. Text to Speech conversion
4. Summarizing the video content

1. Speech to Text conversion

The goal of the Speech-to-Text (STT) conversion component is to help deaf people understand speech by translating spoken language into written text. With an emphasis on preserving accuracy and speed for real-time interaction, this conversion is carried out utilizing sophisticated speech recognition techniques.

We employ pre-trained models for speech recognition (such as OpenAI's Whisper and Google's Speech-to-Text API), which are renowned for their high accuracy in a variety of acoustic settings.

Data processing: The recognition model converts audio signals to text after they have been preprocessed to eliminate noise.

Techniques for Optimization: In noisy settings, language models and noise-canceling filters are used to increase transcribing accuracy.

To perform this task, we use some models and techniques like:

a. Speech recognizer
b. Pyttsx3

Fig 1:Google Speech Recognizer

A. oogle Speech Recognizer:

This recognizer model does the job of turning spoken audio into text. Using the speech_recognizer library

Use of the following method: Recognizer.recognize_google()

Using the Speech Recognition Library's Microphone class, this function listens to the audio input. Google's Speech to Text API is then contacted by the recognize_google() method to process this audio input.

he model makes use of sophisticated structures such as Transformers and Recurrent Neural Networks (RNNs), which are specifically used in Automatic Speech Recognition (ASR). These structures aid in the recognition and transcription of speech patterns into text.

Training of the model: Google pre-trains the model using large datasets of spoken words and phrases in a variety of languages and accents.

Accuracy and Reliability: Because of its extensive training and frequent upgrades, Google's Speech Recognition model is renowned for its high accuracy.

B. PYTTSX3 TTS ENGINE:

It is a Text to Speech model that converts text into audible speech. It is used to add voice feedback functionality, allowing the program to respond to the user with audible messages.

The library is used as: pyttsx3

This library abstracts underlying TTS engines like Microsoft SAPI5 on Windows or NSSpeechSynthesizer on macOS, allowing for platform independent TTS support in Python applications.

How it works:

Pyttsx3.init() initializes the TTs engine.

Engine.say() method is used to add text to the TTS engine’s queue.

The voice feedback is given to the user, particularly in situations where an error or timeout occurs in listen() function.

Accuracy: Although pyttsx3 is not as natural as neural network-based TTS systems, it can still produce clear and understandable speech. The rule-based approach in TTS engines provide a balancebetween clarity and efficiency. It also work offline.

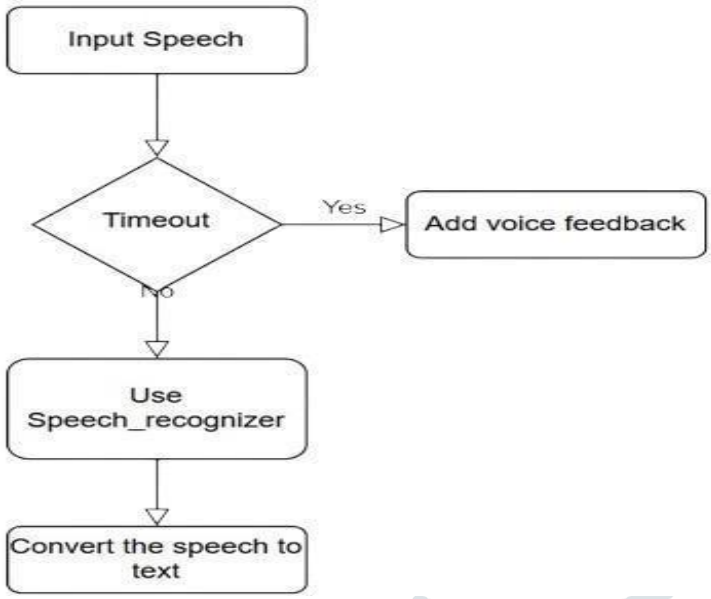


Fig 2: Speech to text

SIGN-LANGUAGE TO TEXT CONVERSION:

The purpose of the Sign Language to Text conversion component is to translate sign language into written text so that mute people can communicate using it. Techniques from computer vision and machine learning are used to do this:

Hand movements and facial expressions are recognized using a Convolutional Neural Network (CNN)-based model that has been refined using sign language datasets (such as the ASL Dataset and Sign Language MNIST).

Pose estimation and tracking: Methods like MediaPipe or OpenPose are used to identify important hand and body landmarks in order to precisely identify sign gestures.

Instantaneous Translation: Real-time camera input is processed by the model, which continuously translates gestures it detects into text. Methods such as Long Short-Term Memory (LSTM) networks are employed to record signing gesture sequences.

Fig 3:Hand gestures for sign language detection

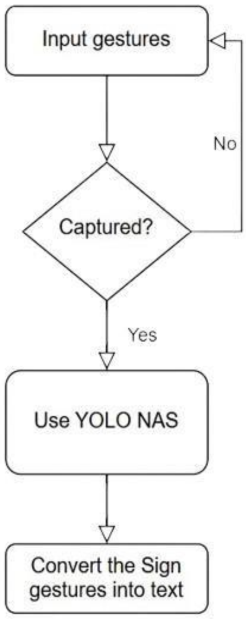


FIG 4: SIGN GESTURES INTO TEXT

TEXT TO SPEECH CONVERSION:

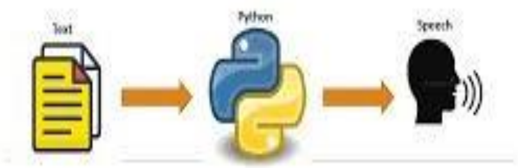
Text-to-Speech (TTS) improves accessibility and user interaction by transforming written text into spoken voice. Because it works offline and ensures privacy and usability without relying on the internet, PYTTSX3 is perfect for this. It is cross-platform and compatible with Linux, macOS, and Windows. For customized, natural-sounding speech, it also enables the adjustment of voice parameters like volume and pace.

PYTTSX3: A Python library called PYTTSX3 was created for offline text-to-speech (TTS) conversion, and it works with Linux, macOS, and Windows.

PYTTSX3 is ideal for applications where internet connectivity is scarce or user data security is an issue because it operates completely offline, protecting privacy.

PSTTSX3 is well-suited for accessibility tools like screen readers and educational applications, making it a valuable resource for enhancing user interaction and learning experiences.

A more customized and natural-sounding speech output is made possible by its ability to change the voice rate, volume, and voice selection.



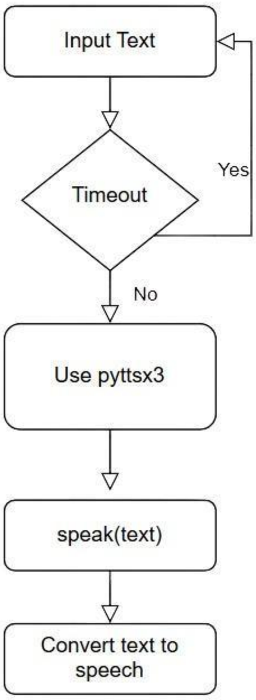


Fig 5: Text to speech

TEXT SUMMARISATION:

This module can condense longer conversations or texts into shorter, more manageable summaries. This is beneficial in group settings or educational contexts where quick understanding is essential, helping sign language users grasp the main points without overwhelming detail.

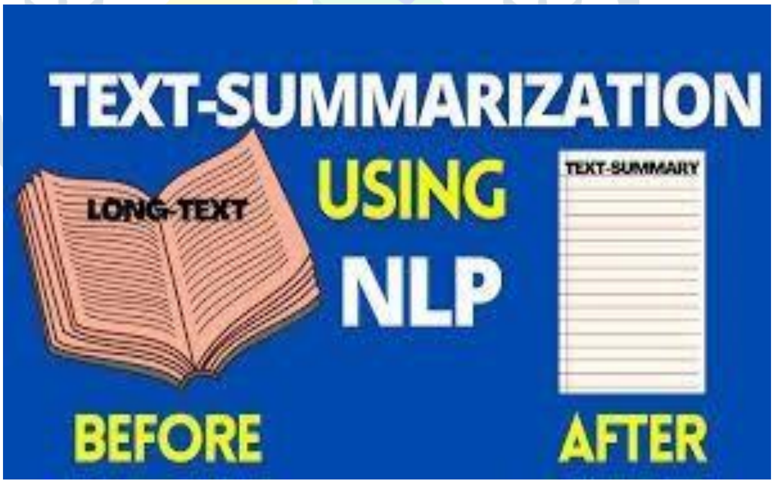


Fig 6:Text Summarization

Hugging Face stands out as a transformative toolkit in the NLP ecosystem. Hugging Face is an open- source machine learning tool that was initially developed to focus on NLP tasks.

Hugging Face hosts a Model Hub, which is a repository of thousands of pre-trained models contributed by the community and researchers. Users can browse, search, and download models for various tasks, including audio and vision tasks, in addition to text.

After obtaining the transcribed text, you can use Hugging Face models for summarization if needed. Models like BART or T5 are effective for this purpose.

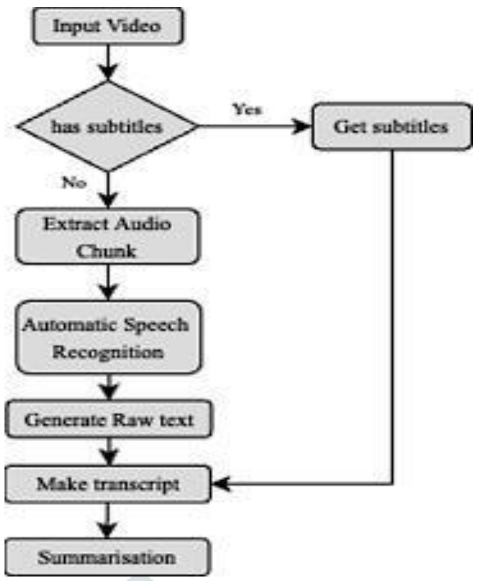


Fig 7: Video Summarization

RANDOM FOREST

This gesture recognition multi-class classification model is based on the Random Forest algorithm. The specific use of Random Forest for translating gestures into text is described in this section.

Training the Random Forest Model: Training and test sets are created from the preprocessed and labeled data (gesture images or frames with matching labels). To make sure the model can handle differences in gesture patterns, multiple decision trees are trained on subsets of the data.

Feature Importance: When understanding gestures, the model may prioritize important properties (such as hand form and motion trajectories) according to Random Forest's inherent feature importance scores. By concentrating on unique gesture features, this aids in making predictions that are more accurate.

Classification: The Random Forest algorithm classifies features from live gesture input using trained trees, producing the most likely gesture label during inference.

Real-Time Adjustment: In order to balance forecast speed and accuracy, the model parameters (such as the number of trees and maximum depth) are tuned in light of the requirement for real-time communication. This makes it possible to convert continuous movements into text quickly.

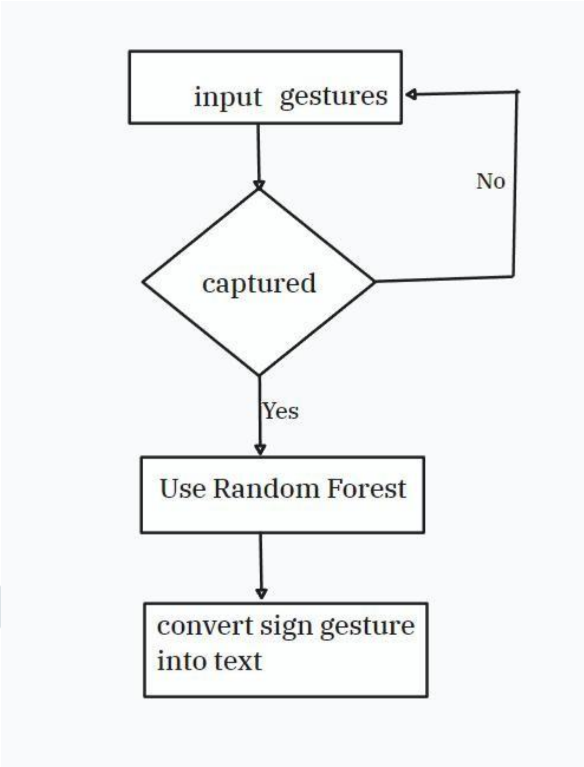


Fig 8:Sign gesture into text using Random Forest

CONCLUSION

THE PROPOSED SMART COMMUNICATION SYSTEM FOR DEAF AND MUTE INDIVIDUALS EFFECTIVELYADDRESSES THE COMMUNICATION BARRIERS FACED BY THOSE WITH HEARING OR SPEECH IMPAIRMENTS,ENABLING SEAMLESS INTERACTION WITH THE HEARING POPULATION. BY INTEGRATING SPEECH-TO-TEXTCONVERSION, GESTURE RECOGNITION FOR SIGN LANGUAGE, AND TEXT-TO-SPEECH TECHNOLOGIES, THESYSTEM OFFERS A REAL-TIME, BIDIRECTIONAL COMMUNICATION PLATFORM THAT IS ACCESSIBLE AND USER FRIENDLY. IT ENHANCES ACCESSIBILITY BY EMPOWERING USERS TO COMMUNICATE INDEPENDENTLYWITHOUT INTERPRETERS, PROMOTES INCLUSIVITY, AND ENSURES REAL-TIME PERFORMANCE WITH LOWLATENCY FOR DYNAMIC CONVERSATIONAL SCENARIOS. ADVANCED MACHINE LEARNING MODELS AND PRE TRAINED DATASETS DELIVER HIGH ACCURACY IN SPEECH RECOGNITION AND GESTURE INTERPRETATION, ENSURING RELIABILITY ACROSS DIVERSE INPUTS. USER FEEDBACK HAS BEEN OVERWHELMINGLY POSITIVE,EMPHASIZING ITS PRACTICAL UTILITY AND ABILITY TO FOSTER MEANINGFUL INTERACTIONS. THE MODULAR ARCHITECTURE SUPPORTS SCALABILITY, ALLOWING FUTURE ENHANCEMENTS AND ADAPTABILITY TO VARIOUSAPPLICATIONS AND ENVIRONMENTS. WHILE THE SYSTEM DEMONSTRATES SIGNIFICANT POTENTIAL INFOSTERING COMMUNICATION EQUITY, CERTAIN LIMITATIONS, SUCH AS CHALLENGES IN DYNAMIC GESTURERECOGNITION AND ENVIRONMENTAL VARIABILITY, SUGGEST OPPORTUNITIES FOR FURTHER REFINEMENT.FUTURE ITERATIONS WILL FOCUS ON IMPROVING ACCURACY, REDUCING LATENCY, AND EXPANDING THESCOPE OF SUPPORTED LANGUAGES AND GESTURES TO CREATE A GLOBALLY INCLUSIVE ASSISTIVE TOOL. THESEADVANCEMENTS AIM TO ENSURE THE SYSTEM BECOMES A TRANSFORMATIVE ENABLER OF COMMUNICATIONAND INCLUSIVITY.

REFERENCES

[1]. Rivera-Acosta, M., Ruiz-Varela, J. M., Ortega-Cisneros, S., Rivera, J., Parra-Michel, R., & Mejia-Alvarez, P. (2021). Spelling correction real-time american sign language alphabet translation system based on yolo network and LSTM. Electronics, 10(9), 1035.

[2]. Goyal, S., Sharma, I., & Sharma, S. (2013). Sign language recognition system for deaf and dumb people. International Journal of Engineering Research Technology, 2(4), 382-387.

[3]. Shin, J., Matsuoka, A., Hasan, M. A. M., & Srizon, A. Y. (2021). American sign language alphabet recognition by extracting feature from hand pose estimation. Sensors, 21(17), 5856.*[4]. Rajam, P. S., & Balakrishnan, G. (2011, September). Real time Indian sign language recognition system to aid deaf-dumb people. In 2011 IEEE 13th international conference on communication technology (pp. 737-742). IEEE.*

[5]. Yin, K. (2020). Sign language translation with transformers. arXiv preprint arrive:2004.00588, 2.

[6]. Ahire, P. G., Tilekar, K. B., Jawake, T. A., & Warale, P. B. (2015, February). Two way communicator between deaf and dumb people and normal people. In 2015 International Conference on Computing Communication Control and Automation (pp. 641-644). IEEE.

[7]. Swathiga, A., & Vinodha, N. (2024, June). Multimodal Communication for Deaf and Dumb. In 2024 International Conference on Smart Systems for Electrical, Electronics, Communication and Computer Engineering (ICSSEECC) (pp. 122-127). IEEE.

[8]. Saleem, M. I., Otero, P., Noor, S., & Aftab, R. (2020, October). Full duplex smart system for Deaf & Dumb and normal people. In 2020 Global Conference on Wireless and Optical Technologies (GCWOT) (pp. 1-7). IEEE.

[8]. Krishna, T. M., & Ponnusamy, M. (2024, May). MuteSpeak: Bridging Communication Gaps with Gestures. In 2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT) (pp. 1-5). IEEE.

[10]. Vigneshwaran, S., Fathima, M. S., Sagar, V. V., & Arshika, R. S. (2019, March). Hand gesture recognition and voice conversion system for dump people. In 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS) (pp. 762-765). IEEE.

[11]. Boppana, L., Ahamed, R., Rane, H., & Kodali, R. K. (2019, July). Assistive sign language converter for deaf and dumb. In 2019 International Conference on Internet of Things (iThings) Dept of IT, GMRIT and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData) (pp. 302-307). IEEE.

[12]. Duhan, M., & Gulati, U. (2020, February). Intelligent System to Make the World Hear Deaf-mute People. In 2020 3rd International Conference on Emerging Technologies in Computer Engineering: Machine Learning and Internet of Things (ICETCE) (pp. 46-51).

IEEE.[13].Jim, A. A. J., Rafi, I., Tiang, J. J., Biswas, U., & Nahid, A. A. (2024). KUNet-An Optimized AI Based Bengali Sign Language Translator for Hearing Impaired and Non Verbal People. IEEE Access, 12, 155052-155063.

[14]. Ajay, S., Potluri, A., George, S. M., Gaurav, R., & Anusri, S. (2021, July). Indian sign language recognition using random forest classifier. In 2021 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT) (pp. 1-6). IEEE.

[15]. Ahmed, S. A., Mahmood, B. N., Mahmood, D. J., & Namq, M. M. (2024). ENHANCING KURDISH SIGN LANGUAGE RECOGNITION THROUGH RANDOM FOREST CLASSIFIER AND NOISE REDUCTION VIA SINGULAR VALUE DECOMPOSITION (SVD). Science Journal of University of Zakho, 12(2),257-267