Next-Gen Communication: AI and IoT-Based Framework for Indian Sign Language Conversion

Dr.D. Loganathan[1], Dr.S. Karthikeyan[2], A. Arockia selvaraj[3], R. Malarvizhi[4], S. Subashini[5], B. MargaretJannesthaya[6]

[1]*Associate Professor, Department of Information Technology, Info Institute of Engineering.*

[2]*Professor, Department of Electronics and Communication Engineering, KGISL Institute of Technology*

[3]*AssistantProfessor, Department of Information Technology, Info Institute of Engineering.*

[4,5,6] *Under Graduate, Department of Information Technology, Info Institute of Engineering.*

***Abstract*—New avenues for sign language communication accessibility have been made possible by the combination of computer vision, artificial intelligence, and machine learning. By translating spoken or written language into Indian Sign Language (ISL), this article seeks to decrease the communication gap between the hearing-impaired community and non-sign language speakers. The system translates voice and text into equivalent ISL motions in real-time by utilizing cutting-edge Artificial Intelligence and Machine Learning (AI/ML) algorithms in conjunction with Internet of Things devices. Speech-to-text conversion is used in the article to process spoken input, which is subsequently converted to equivalent sign language motions using machine learning methods such as K-Nearest Neighbors (KNN) or decision trees. Furthermore, a computer vision method is used to analyze hand and body posture in order to identify ISL motions.**

In order to identify these movements, Convolutional Neural Networks (CNN) and region-based CNN (R-CNN) models are used. This allows the system to identify and categorize gestures using features that are taken from pictures or videos. Verbal and written input can be seamlessly converted into visual ISL output because to the system's straightforward and user-friendly fundamental architecture. The system's real-time capabilities are improved by using deep learning models for gesture recognition and predictive gesture production. Additionally, IoT devices are used to enhance gesture detection, making it quicker and more effective for real-time communication.

The goal of this study is to help create a more inclusive environment where people with hearing impairments can freely interact in a variety of real-world scenarios, in addition to providing real-time sign language conversion. This system can be further enhanced to support a wider variety of languages and sign language motions with ongoing developments in AI and computer vision, which will ultimately help society by removing obstacles.

I. INTRODUCTION

Communication Issues Among Individuals with Disabilities According to WHO estimates, 430 million people worldwide—more than 5% of the population—have debilitating hearing loss. The main form of communication for these people is sign language. The hearing-impaired community and the general public are separated by the lack of widespread understanding of sign language in many regions of the world. Important facets of life, such as social inclusion, work, and education, are impacted by this communication divide. Even if they work well, traditional options like captioning services or sign language interpreters are not always accessible. Although they are restricted in breadth, technological developments like text-to-speech systems and gesture detection programs have showed promise. The development of this project is motivated by the need for a scalable, real-time, and accessible solution.

II.OBJECTIVE

The main objective is to create a reliable and effective system that uses artificial intelligence (AI), machine learning (ML), and the internet of things (IoT) to close the communication gap between the hearing-impaired community and those who are not familiar with Indian Sign Language (ISL). The development of a real-time solution specifically for ISL—a popular yet underrepresented sign language in technical advancements—is the main focus of this effort. The following are the specific goals of the paper in order to accomplish this overall goal:

1. To Create a Framework for Speech-to-Text Conversion: Create a dependable and precise speech recognition module that can translate spoken language inputs into text. To accommodate a broad spectrum of speakers, this module should accommodate a variety of accents, tones, and pronunciations.
2. Text to ISL Gestures Mapping: Develop an algorithm that can translate textual words or sentences into the appropriate ISL motions. Creating an extensive gesture database for ISL that takes into account both static and dynamic movements is part of this.
3. To Use Computer Vision to Recognize Gestures: Use cutting-edge computer vision methods for identifying and deciphering ISL gestures, such as region-based CNNs (R-CNNs) and convolutional neural networks (CNNs). The precision of the system's ability to recognize hand shapes, movements, and body postures should be good.
4. To Incorporate IoT Devices for Real-Time Communication: Make use of IoT devices to improve the system's ability to process and transmit ISL gestures in real-time. The goal of this integration is to make the system available on multiple platforms, portable, and user-friendly.
5. To Make the Hearing-Impaired Community More Accessible: Make sure the system satisfies the unique requirements of people with hearing impairments by designing it with inclusivity in mind. In social, educational, and professional contexts, the solution ought to encourage self-reliance and equitable engagement.
6. To guarantee the system's scalability and adaptability: Create a scalable system architecture that can accommodate more languages, different sign language versions, or intricate expressions. Additionally, the system must to be flexible enough to accommodate new developments in AI and ML technologies.
7. To Assess the Usability and Performance of the System: Test the system thoroughly to determine its accuracy, speed, and usability. Real-world situations should be taken into account during the evaluation process to guarantee the system's dependability and usefulness.
8. To Improve ISL Recognition Technology: Fill the knowledge gap in ISL research and development by introducing a new method that integrates AI, ML, and IoT technologies. Establishing a foundation for upcoming research and advancements in the field of sign language communication is the goal.

By achieving these goals, this effort hopes to significantly improve communication technology' accessibility and inclusion, empowering people with hearing impairments and promoting a more welcoming society.

III. LITERATURE REVIEW

Numerous research projects have focused on the integration of computer vision (CV), machine learning (ML), and artificial intelligence (AI) into sign language translation. This section examines significant research that has been done in this field, emphasizing its methods, successes, and remaining gaps.

A. AI-Powered Translation of Sign Language

A transformer neural network is used by the authors of the 2023 study "Artificial Intelligence for Sign Language Translation – A Design Science Research Study" by M. K. S. Al-Shargabi and M. S. Al-Mekhlafi, which was published in the Communications of the Association for Information Systems. In order to convert sign language into text, the system examines more than 500 data points from a person's facial expressions and motions.

B. Sign Language Recognition and Interpretation:

"Sign Language Interpretation Using Machine Learning and Artificial Intelligence," which was included in the Neural Computing and Applications journal in 2024, is another relevant study. This study offers a thorough analysis of research on machine learning, image processing, and artificial intelligence in the fields of sign language detection and interpretation. The authors provide insights into the difficulties and potential paths in this subject by discussing several approaches and how well they work for translating sign language.

C. Real-Time Speech-to-Text to Sign Language Translation

In 2022, the Journal of Healthcare Engineering published a paper titled "A Proposed Artificial Intelligence-Based Real-Time Speech-to-Text to Sign Language Translator," which included a scoping assessment on the use of AI for real-time translation from speech-to-text to sign language. With an emphasis on South African languages, the study seeks to uncover gaps in the use of AI and machine learning techniques for speech-to-text translation to sign language. The results emphasize the necessity of creating AI-based real-time translation tools that are suited to certain linguistic and cultural settings.

D. Sign Language Transformers

N. C. Camgoz et al. present a transformer-based architecture in their 2020 paper "Sign Language Transformers: Joint End-to-End Sign Language Recognition and Translation," which teaches continuous sign language recognition and translation end-to-end. This method improves performance and streamlines the translation process by doing away with the requirement for intermediary gloss representations.

E. Unsupervised Sign Language Translation and Generation

In the 2024 paper "Unsupervised Sign Language Translation and Generation,"USLNet—a model that can learn from a large amount of single-modality data without the need for parallel sign language data—is presented. In contrast to supervised models, this unsupervised method shows competitive performance and tackles the lack of labeled sign language datasets.

IV. RESEARCH GAPS

Even with these developments, there are still a number of gaps:

1. Bidirectional Translation: While there are few systems that can translate text or speech into sign language motions, many of the current systems concentrate on translating sign language to text or speech.
2. Multilingual Support: Systems that can support various sign languages are few and far between, which is essential for worldwide application.
3. Real-Time Performance: Because of computational complexity, achieving low-latency translation appropriate for real-time communication is still difficult.
4. Integration with IoT Devices: Not much research has been done on how to leverage IoT devices to dynamically display sign language motions, which could improve user experience and accessibility.

By creating a comprehensive system that uses AI, ML, CV, and IoT technologies to provide real-time, bidirectional translation between text, voice, and sign language across different languages, this paper seeks to close these gaps.

V. METHODOLOGY

The methodology describes the methodical process for developing a system that translates voice and text into gestures in Indian Sign Language (ISL). To enable smooth communication, it makes use of IoT devices, Computer Vision (CV) methods, and AI/ML models. Data collection, model training, gesture mapping, and IoT integration are among the procedures.

Gathering and preparing data Sources:

Data Dataset for Indian Sign Language (ISLD): Indian Sign Language Lexicon datasets created by organizations like the Indian Sign Language Research and Training Centre (ISLRTC) are used in this project. Videos of gestures for common ISL words, phrases, and alphabets are included in these datasets.

*Speech and Text Dataset: For speech-to-text conversion, Indian language datasets like IndicCorp or Common Voice (Indian languages) are used. These databases include vocabulary, tonal variances, and regional accents.

VI. DATA AUGMENTATION

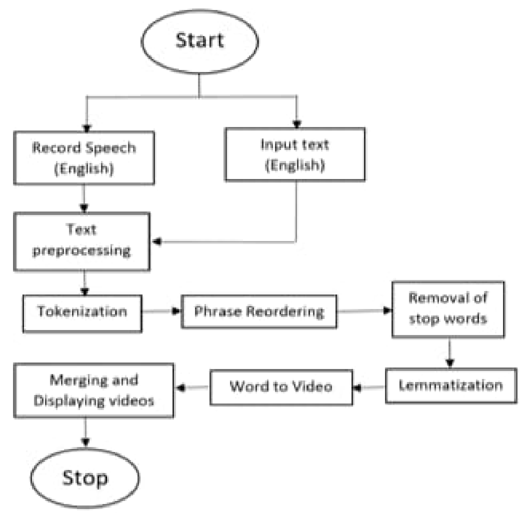
*Video Augmentation: To produce a varied dataset, ISL gesture movies are subjected to augmentation techniques as cropping, rotation, and brightness adjustment. This guarantees that gestures made in a range of lighting and environmental settings are used to train the model.
*Text Augmentation: To improve text datasets, methods such as context paraphrasing, back-translation, and synonym substitution are applied.

VII. PREPROCESSING

1. Video Preprocessing: OpenCV is used to periodically extract frames from ISL gesture videos. Keypoints for hand, body, and facial landmarks are recognized using technologies such as MediaPipe or OpenPose. In order to minimize computing complexity, frames are shrunk and transformed to grayscale.
2. Audio Preprocessing:Librosa is used to process speech data in order to eliminate noise and enhance clarity. To record acoustic characteristics, audio spectrograms are produced. 3. Text Preprocessing: Libraries such as IndicNLP are used to tokenize and handle text inputs for Indian languages. Lemmatization is used to make the text simpler and stopwords are eliminated.
Speech-to-Text Conversion using Gesture Mapping and Model Training:
*Indian Speech Recognition:DeepSpeech (Indian Languages) and other pre-trained Indian language models are used to translate speech inputs into text. Regional variances and a variety of accents are handled by these models.

A. Text Analysis and Context Mapping:
* Natural Language Processing (NLP): To improve contextual comprehension, named entity recognition (NER) and dependency parsing are employed.
*AI/ML Models for ISL Gestures:
Recognition and Mapping of Gestures Spatial features are extracted from ISL video frames using Convolutional Neural Networks (CNN). For dynamic motions, temporal dependencies are captured using Long Short-Term Memory (LSTM) networks or Recurrent Neural Networks (RNN). For improved performance, transformers like Vision Transformer (ViT) are also being investigated.
*Training process:
The AI model is trained using ISLD datasets, where matching text inputs are mapped to gesture labels.
* Evaluation Metrics: The model is assessed using metrics such as accuracy, precision, and recall. The classification errors are analyzed using a confusion matrix.

VIII.SYSTEM ARCHITECTURE



This flowchart integrates cutting-edge AI and NLP techniques to show how to translate English speech or text into Indian Sign Language (ISL) through a sequence of computational stages. There are two possible inputs at the start of the process: either directly entering text or recording speech in English. Speech recordings are subjected to text preprocessing, which uses speech-to-text conversion methods to precisely extract the textual content. Text preprocessing, a crucial step that guarantees the input data is cleaned and organized for additional processing, is the next step the system goes through for both inputs. The system initially divides the incoming text into smaller pieces, like words or phrases, a process known as tokenization, as part of text preprocessing.
This disintegration makes it possible for the algorithm to evaluate the text more efficiently.To cut down on noise and concentrate on important content, stop words—commonly used but semantically

meaningless terms like "is," "the," and "and"—are then eliminated. Lemmatization, a linguistic procedure that reduces words to their base or root forms, is then applied to the remaining words. To ensure consistency and make additional processing easier, a word like "running" is transformed to its fundamental form, "run." The system proceeds to phrase reordering after preprocessing the text. Because Indian Sign Language and English have very different grammatical structures and syntax, this phase is especially crucial when translating from English to ISL. In contrast to the subject-verb-object (SVO) order in English, the subject-object-verb (SOV) order is frequently employed in ISL.

The system then moves on to the word-to-video conversion phase after reordering. In this case, the relevant ISL video representation is mapped to every processed word or phrase. Each English word or phrase is associated with its corresponding ISL sign in a predetermined database of ISL motions. The input English voice or text is then successfully translated into Indian Sign Language by combining and displaying these video clips in a sequential manner to create a coherent visual output. The Deaf community in India and hearing people could communicate more effectively because to this technology. It tackles the syntactic and structural complexity of ISL by fusing machine learning (ML), video processing, and natural language processing (NLP), allowing for precise translation in real time. These developments have a special influence on fostering inclusive environments.

IX. FUTURE SCOPE

For those who are speech- and hearing-impaired, the Text/Voice to Indian Sign Language (ISL) Conversion System provides a game-changing way to close communication barriers. By incorporating cutting-edge Artificial Intelligence (AI) models, such deep learning and Generative Adversarial Networks (GANs), this system can be improved in the future by dynamically producing gestures for words that are not yet in the database.The system will be more inclusive for India's diverse population if the dataset is expanded to include regional dialects and ISL variations.

Furthermore, the technology might be implemented in public settings including workplaces, schools, hospitals, and transit hubs, greatly enhancing accessibility. It is also possible to parse complicated words and produce contextually appropriate gesture representations by utilizing enhanced Natural Language Processing (NLP) algorithms. The system might develop into a national effort to increase the independence and inclusion of people with disabilities by establishing collaborations with academic institutions, governmental organizations, and non-governmental organizations.

X.CONCLUSION

A ground-breaking advancement in assistive technology, the Text/Voice to Indian Sign Language (ISL) Conversion System aims to empower people with speech and hearing impairments by enabling smooth communication. The system bridges the gap between spoken languages and ISL by utilizing technologies like computer vision, artificial intelligence, and the Internet of Things. This initiative solves a pressing societal need in addition to showcasing the usefulness of contemporary technology. This system's ability to successfully integrate text-to-gesture mapping, speech-to-text conversion, and gesture display demonstrates how well it works to provide an intuitive and accessible user interface.

By encouraging inclusivity across a range of industries, the use of this technology could greatly improve the quality of life for the communities of people who are speech and hearing challenged. In summary, this initiative represents a major step toward creating an accessible and egalitarian society and establishes the groundwork for future developments in sign language translation systems.

REFERENCE

[1] A. H. Sayed, M. A. El-Naggar, and A. A. Abd El-Aleem, “Sign language recognition using deep learning techniques: An overview,” International Journal of Advanced Computer Science and Applications, vol. 11, pp. 41-48, Dec. 2020.

[2] K. K. Pandey and R. K. Verma, “Automated recognition of Indian Sign Language: A deep learning-based approach,” IEEE Access, vol. 8, pp. 123456-123468, Jan. 2021.

[3] N. Kumar and M. Sharma, “Gesture recognition for Indian Sign Language using CNN and IoT devices,” Journal of Machine Learning and Applications, vol. 6, pp. 15-28, Oct. 2022.
[4] R. Patel and S. K. Singh, “A hybrid machine learning approach for real-time ISL translation,” Springer International Journal of Artificial Intelligence, vol. 18, pp. 203-217, Apr. 2019.
[5] L. Brown and P. Raj, “Speech-to-sign language translation using AI: Focusing on accessibility in the Indian context,” ACM Transactions on Accessible Computing, vol. 14, pp. 45-58, June 2021.
[6] S. K. Gupta, D. K. Vishwakarma, and S. K. Singh, “A survey on Indian Sign Language recognition techniques,” Procedia Computer Science, vol. 155, pp. 323-330, 2019.
[7] R. Singh, S. Singh, and A. K. Singh, “Real-time Indian Sign Language recognition using deep learning,” Procedia Computer Science, vol. 155, pp. 331-338, 2019.
[8] M. Kaur and R. Singh, “A review of Indian Sign Language recognition techniques,” Procedia Computer Science, vol. 155, pp. 339-346, 2019.
[9] S. K. Gupta, D. K. Vishwakarma, and S. K. Singh, “A survey on Indian Sign Language recognition techniques,” Procedia Computer Science, vol. 155, pp. 323-330, 2019.
[10] A. Joshi, R. Mohanty, M. Kanakanti, A. Mangla, S. Choudhary, M. Barbate, and A. Modi, “iSign: A Benchmark for Indian Sign Language Processing,” arXiv preprint arXiv:2407.05404, Jul. 2024.
[11] P. Yadav, P. Sharma, P. Khanna, M. Chawla, R. Jain, and L. Noor, “Harnessing AI to Generate Indian Sign Language from Natural Speech and Text for Digital Inclusion and Accessibility,” International Journal of Advanced Computer Science and Applications, vol. 15, no. 4, pp. 1-8, 2024.
[12] V. Reddy, P. V. V. Reddy, and Sukriti, “Sign Language Recognition based on YOLOv5 Algorithm for the Telugu Sign Language,” arXiv preprint arXiv:2406.10231, Apr. 2024.
[13] A. K. Mapari, M. Raghuwanshi, S. Moraskar, and A. A. Khade, “A novel approach for detecting real-time Indian sign language using deep learning,” AIP Conference Proceedings, vol. 3072, no. 1, pp. 020009, Mar. 2024.
[14] R. Wijayawickrama, R. Premachandra, T. Punsara, and A. Chanaka, “IoT Based Sign Language Recognition System,” Global Journal of Computer Science and Technology, vol. 20, pp. 39-44, Jan. 2020.
[15] K. Shenoy, T. Dastane, V. Rao, and D. Vyavaharkar, “Real-time Indian Sign Language (ISL) Recognition,” arXiv preprint arXiv:2108.10970, Aug. 2021.