

Indian Sign Language to Text/Speech Translation

Miss. Nandini J. Kakde¹, Asst. Prof. Arsalan A. Shaikh²

Research Scholar, Department of Computer Applications, SSBT's COET, Jalgaon Maharashtra, India¹

Professor, Department of Computer Applications, SSBT's COET, Jalgaon Maharashtra, India²

Abstract: Indian Sign Language (ISL) is an important visual language used by deaf and hard-of-hearing people in India for communication. However, many people who do not understand sign language face difficulties in communicating with ISL users. In many situations, communication depends on another person who understands sign language, which may not always be available. Therefore, there is a need for a system that can automatically recognize Indian Sign Language gestures and convert them into understandable text and speech.

This research proposes an Indian Sign Language Text/Speech Translation System to make communication between ISL users and non-sign-language users easier and more accessible. The system provides a digital method to capture ISL gestures through a camera, process the input, identify the signs, and convert the recognized gestures into text. The generated text can then be converted into speech using a text-to-speech component. The proposed system can use computer vision, image processing, machine learning or deep learning techniques, and natural language processing for recognizing and translating ISL gestures. The system can support real-time communication and reduce the need for manual interpretation in basic communication situations. It can also provide a foundation for future development of multilingual translation and improved recognition of continuous signs.

Keywords: Indian Sign Language, ISL Translation, Sign Language Recognition, Text Translation, Speech Generation, Computer Vision, Deep Learning, Natural Language Processing.

I.INTRODUCTION

1.1 Background

Indian Sign Language (ISL) is a visual language used by deaf and hard-of-hearing people in India for communication. Sign language allows users to express information through hand gestures, movements, facial expressions, and body positions. However, people who do not understand ISL may face difficulties communicating with ISL users. This creates a communication barrier in situations such as education, healthcare, public services, workplaces, and daily activities. Research in ISL translation has therefore focused on developing computer-based systems that can recognize signs and convert them into understandable language.

Recent developments in computer vision, machine learning, deep learning, and natural language processing have made automatic sign language recognition possible. Existing research has explored systems that capture ISL gestures using cameras, process the captured images or video, identify the corresponding signs, and generate text or speech output. Some systems focus on recognizing individual signs or alphabets, while newer approaches also investigate continuous sign language translation. The ISLTranslate dataset, for example, provides more than 31,000 ISL-English sentence or phrase pairs for continuous ISL translation research.

The proposed Indian Sign Language Text/Speech Translation system is mainly focused on recognizing ISL gestures and converting them into text and speech. The system can use a camera as an input device, followed by preprocessing, feature extraction, gesture recognition, text generation, and text-to-speech conversion. In this way, the system can provide a digital method for reducing communication difficulties between ISL users and people who do not understand sign language.

1.2 Importance

An Indian Sign Language Text/Speech Translation system can play an important role in improving communication accessibility for deaf and hard-of-hearing people. It can provide a simple method for converting visual sign language into text that can be read and speech that can be heard by other people. This can be useful when a person who understands ISL is not available to interpret the communication.

The use of computer vision and deep learning can help the system recognize hand gestures from camera input. Natural language processing can further help in arranging recognized signs into meaningful textual output, while text-to-speech technology can convert the generated text into audio. Research on ISL translation has shown the importance of combining sign recognition and language translation, particularly for continuous sign language.

The system can therefore support communication in educational institutions, hospitals, offices, public places, and other situations where communication between ISL users and non-sign-language users is required. It can also provide a foundation for future systems supporting multiple Indian languages and more complex continuous signs.

1.3 Translation System

The proposed translation system is a digital platform designed to recognize Indian Sign Language gestures and convert them into understandable text and speech. The system can receive ISL gestures through a camera and process the captured images or video frames. The input is first preprocessed to improve the quality of the captured data and to identify the relevant hand or gesture information.

After preprocessing, important features are extracted from the input and provided to a machine learning or deep learning model for sign recognition. The recognized sign is then converted into text. The generated text can be displayed to the user and can also be provided to a text-to-speech component to produce audio output. Similar approaches have been explored in existing ISL translation research using computer vision, deep learning, NLP, and speech generation techniques.

The system can therefore provide a complete translation flow from visual sign input to textual and spoken output. It can also maintain a structured approach in which camera input, preprocessing, recognition, translation, and speech generation are handled as separate stages.

1.4 Types of Sign Language Translation

The proposed system can consider different types of sign language translation depending on the type of input and expected output. **Static sign recognition** deals with individual signs, alphabets, numbers, or predefined gestures represented by a single image or frame. This type of recognition is comparatively simpler because the system processes one sign at a time.

Dynamic sign recognition deals with gestures that involve hand movement over a sequence of frames. The system needs to analyze changes in position and movement over time to recognize the complete sign. **Continuous sign language translation** goes further by processing a sequence of signs and generating a sentence or phrase in a spoken language. The ISLTranslate research specifically addresses continuous ISL translation and provides a large ISL-English dataset for this purpose.

Another type is **text and speech translation**, in which the recognized ISL is converted into written text and then into spoken output. Recent research has investigated real-time ISL-to-text/speech systems using computer vision, machine learning, deep learning, NLP, and speech-processing techniques.

1.5 Research Problems

Communication between ISL users and people who do not understand sign language can be difficult because both users may not share a common communication method. Manual interpretation generally requires a person who understands ISL, but such assistance may not always be available. Therefore, an automatic and accessible translation system can help reduce this communication barrier.

Another important problem is the accurate recognition of ISL gestures from camera input. Changes in hand position, background, lighting conditions, gesture speed, viewing angle, and differences between users can affect recognition performance. A system must therefore process the input carefully before providing the final translation.

Existing systems may also focus on limited sets of signs, alphabets, or predefined gestures. Continuous sign language translation is more challenging because the system needs to understand a sequence of signs and generate meaningful language. The availability of suitable datasets is also an important research issue. The ISLTranslate study highlights the limited availability of large ISL translation resources and introduces a dataset of approximately 31,000 ISL-English sentence/phrase pairs.

Therefore, there is a need for a system that combines gesture recognition, language processing, and text-to-speech generation to provide a more complete ISL translation process.

1.6 Scope

The scope of the proposed Indian Sign Language Text/Speech Translation system covers the recognition and translation of ISL gestures captured through a camera. The system can process static signs as well as selected dynamic gestures depending on the available dataset and recognition model. It can identify the input gesture using image processing, computer vision, and machine learning or deep learning techniques.

The recognized signs can be converted into text and displayed on the system interface. The generated text can then be converted into speech using a text-to-speech component. The system can therefore provide both visual and audio output for communication.

The proposed system can be used for basic communication in educational institutions, healthcare environments, offices, public places, and other daily situations. The system can also maintain recognized text during a session so that a sequence

of signs can be converted into meaningful phrases. Future development can extend the system to continuous ISL translation, multilingual Indian language output, improved gesture recognition, and mobile or web-based applications.

1.7 Objectives

The main objective of the proposed system is to develop a computer-based system for translating Indian Sign Language gestures into text and speech. The system aims to capture ISL gestures through a camera and identify the corresponding signs using computer vision and machine learning or deep learning techniques.

Another objective is to convert the recognized signs into understandable textual output. The system also aims to generate speech from the translated text so that people who do not understand ISL can receive the information through audio output.

The proposed system further aims to reduce communication difficulties between ISL users and non-sign-language users. It is designed to provide a simple and accessible translation process while reducing dependence on manual interpretation for basic communication. The system can also provide a foundation for future improvements such as continuous sign recognition, multilingual translation, and improved real-time performance.

1.8 Need of the System

A dedicated Indian Sign Language Text/Speech Translation system is needed to reduce communication difficulties between people who use ISL and people who do not understand sign language. In daily situations, an ISL user may need to communicate with teachers, doctors, employees, service providers, or other people who may not know ISL. The absence of a common communication method can make such interactions difficult.

A computerized translation system can provide an alternative communication method by recognizing ISL gestures and converting them into text and speech. The use of a camera allows the system to receive visual input without requiring special physical equipment. Machine learning and deep learning techniques can then be used to identify the signs, while NLP and text-to-speech techniques can help produce understandable output.

The availability of ISL translation datasets is also important for developing and improving such systems. The ISLTranslate dataset provides more than 31,000 ISL-English sentence/phrase pairs and demonstrates the importance of data resources for continuous ISL translation research.

Therefore, the proposed system is needed to provide a simple, technology-based approach for supporting communication and to create a foundation for future development of real-time and multilingual ISL translation systems.

1.9 Selection of Life Cycle Model for Development

The **Prototype Model** is suitable for developing the proposed Indian Sign Language Text/Speech Translation system because the system contains several components that may require repeated testing and improvement. The recognition of ISL gestures depends on factors such as the quality of input images, dataset, preprocessing method, selected features, and performance of the machine learning or deep learning model.

In the Prototype Model, an initial version of the system can first be developed with basic features such as camera input, preprocessing, gesture recognition, and text output. The prototype can then be tested using sample ISL gestures. Based on the observed recognition results and user requirements, changes can be made to the preprocessing, recognition model, interface, or translation process.

After improving the recognition component, speech generation can be added to the system. Further testing can be performed to evaluate recognition accuracy, translation quality, response time, and speech output. User feedback can also be considered during different stages of development. Therefore, the Prototype Model provides a flexible approach for developing and improving the proposed ISL Text/Speech Translation system according to practical requirements.

II. LITERATURE REVIEW

Indian Sign Language (ISL) translation has received increasing attention because of the communication difficulties faced by deaf and hard-of-hearing people when interacting with people who do not understand sign language. Researchers have explored computer vision, machine learning, deep learning, natural language processing, and speech technologies to develop systems that can automatically recognize ISL gestures and convert them into text or speech.

Punekar et al. proposed an Indian Sign Language to Text and Speech Translator, demonstrating the possibility of converting recognized ISL signs into understandable text and audio output. The study provides an early foundation for automated ISL translation systems and shows the usefulness of combining sign recognition with speech generation. [4] More recent work has focused on deep learning techniques for improving recognition and real-time translation. Khetam et al. proposed an ISL to Text/Speech Translation system using deep learning and computer vision, showing the potential of automated gesture recognition for producing both text and speech output. [1]

Bontha et al. presented another ISL Text/Speech Translation approach using computer vision, machine learning, NLP, CNNs, and transformer-based techniques. Their work focuses on providing real-time translation and improving communication between ISL users and non-sign-language users. [2] Other recent studies have also explored lightweight deep learning models and multilingual output to make ISL translation systems more practical for different applications. For continuous sign language translation, Joshi et al. introduced ISLTranslate, a dataset containing 31,222 ISL-English sentence and phrase pairs. The dataset provides an important resource for studying continuous ISL-to-English translation rather than recognizing only individual signs. [3] This work shows that large and suitable datasets are important for developing reliable translation models.

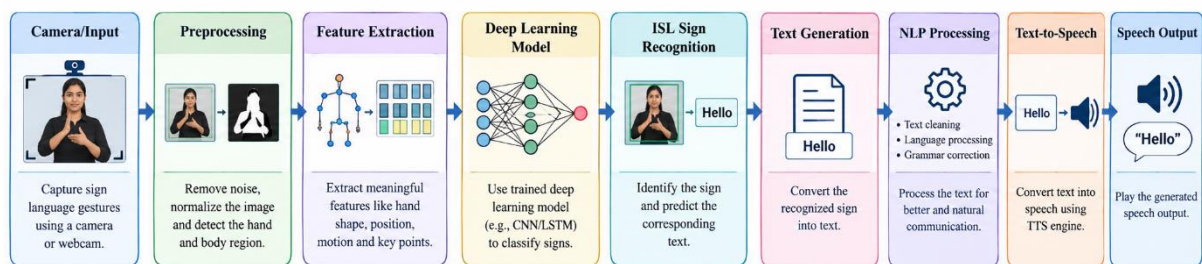
Overall, existing research demonstrates progress in ISL gesture recognition and text/speech generation. However, challenges remain in continuous sign recognition, variations in gestures, lighting and background conditions, dataset availability, translation accuracy, and real-time performance. Therefore, the proposed research focuses on combining camera-based gesture recognition, deep learning, NLP, and text-to-speech techniques to develop an accessible Indian Sign Language Text/Speech Translation System.

III. METHODOLOGY

The proposed research follows a system development methodology for designing and developing an Indian Sign Language Text/Speech Translation system. The first step is to study the communication problems faced by ISL users and understand the existing methods used for sign language recognition and translation. Based on this study, the functional and non-functional requirements of the proposed system are identified. Existing research shows that computer vision, machine learning, deep learning, natural language processing, and speech technologies can be combined to recognize ISL gestures and produce text and speech output. [1], [2]

After identifying the requirements, the structure and working of the proposed system are designed using suitable process-flow and UML diagrams. The system receives ISL gestures through a camera and performs preprocessing and feature extraction before applying a trained recognition model. The recognized signs are converted into text and can then be converted into speech using a text-to-speech component. The system can be tested using sample ISL images or videos to evaluate recognition accuracy, response time, and translation performance. The results obtained from testing can be used to identify errors and improve the proposed system. [1], [3]

Fig. 1. Basic workflow of the proposed Indian Sign Language Text/Speech Translation system.



3.1 Feasibility Study

The feasibility study is carried out to determine whether the proposed Indian Sign Language Text/Speech Translation system can be developed and used successfully in a practical environment. The study considers technical, operational, economic, and schedule feasibility. Existing research has demonstrated the use of computer vision and deep learning for ISL recognition and text/speech generation, which supports the technical feasibility of the proposed system. [1], [2]

Technical Feasibility: The proposed system can be developed using a camera, image-processing techniques, Python-based machine learning or deep learning libraries, and a suitable dataset of ISL gestures. A trained model can be used to recognize the signs and generate the required output. Existing ISL research has demonstrated the feasibility of camera-based recognition and deep-learning approaches. [1], [4]

Operational Feasibility: The system is designed to be simple and practical for users. A user can show an ISL gesture in front of a camera, after which the system processes the input and provides text and speech output. The interface can be designed with simple controls so that users can operate the system without requiring advanced technical knowledge.

Economic Feasibility: The proposed system can be developed using commonly available hardware such as a computer and webcam and software tools that are freely available or open source. Therefore, the basic prototype can be developed without requiring expensive specialized equipment.

Schedule Feasibility: The Prototype Model is suitable because the system can be developed in stages. Camera input and basic sign recognition can be developed first, followed by text generation, speech output, testing, and refinement. This makes it easier to manage the development process within the available time.

3.2 Risk Analysis

Risk analysis is an important part of developing the proposed ISL Text/Speech Translation system because several factors can affect recognition and translation performance. One major concern is recognition accuracy risk. The system may incorrectly recognize a sign because of differences in hand position, gesture speed, background, lighting, or camera angle. This risk can be reduced by using suitable preprocessing techniques and a sufficiently varied training dataset. [1], [3]

Another concern is data quality risk. If the training dataset contains insufficient examples or inconsistent sign representations, the trained model may not perform correctly on new inputs. The use of suitable datasets and proper data preparation can help reduce this risk. The ISLTranslate research demonstrates the importance of large datasets for continuous ISL translation. [3]

There is also a technical risk related to camera quality, processing speed, or hardware limitations. A low-quality camera or insufficient computing resources may affect real-time recognition. Testing the system on suitable hardware and optimizing the recognition model can help minimize this problem.

Speech-output risk may occur if the recognized text is incorrect or incomplete. Since speech is generated from the recognized text, errors in recognition can also affect the final audio output. Proper validation of the recognized text can help reduce this problem.

Finally, user and privacy risks should be considered because camera input may contain personal information. The system should process and store only the information required for translation and should avoid unnecessary storage of camera data.

3.3 Requirement Collection and Identification

The requirements for the proposed ISL Text/Speech Translation system can be identified by studying existing ISL communication and translation processes. Information can be collected by reviewing previous research papers, existing ISL recognition systems, available datasets, and the requirements of potential users. Research papers related to ISL text/speech translation provide useful information about recognition techniques, datasets, deep learning models, and output generation. [1], [2], [3]

The main stakeholders of the proposed system include ISL users, non-sign-language users, system administrators, and developers. ISL users require an easy method for communicating their signs, while non-sign-language users require understandable text or speech output. Developers require suitable datasets, software tools, and model requirements for implementing the recognition system.

The major requirements identified for the proposed system are:

- Camera-based ISL input
- Image or video preprocessing
- Hand or gesture detection
- Feature extraction
- Sign recognition
- Text generation
- NLP-based text processing
- Text-to-speech conversion
- User-friendly interface
- Performance evaluation and reporting

3.4 Software and Hardware Requirements Specification (SRS)

The proposed Indian Sign Language Text/Speech Translation system requires suitable hardware and software resources for development and operation. The hardware requirements include a computer or laptop with a webcam or camera for capturing ISL gestures. Depending on the complexity of the deep learning model, additional processing resources such as a suitable GPU may be useful during model training.

The software requirements include a suitable programming language such as Python, image-processing and computer-vision libraries, machine learning or deep learning frameworks, and text-to-speech tools. Libraries such as OpenCV can be used for image and video processing, while suitable deep learning frameworks can be used for training and testing the recognition model.

A dataset containing ISL images or videos is also required for model training and evaluation. For continuous ISL translation research, the ISLTranslate dataset provides 31,222 ISL-English sentence/phrase pairs and can be considered as a research resource. [3]

The proposed system can operate on commonly used operating systems such as Windows or Linux. A reliable camera and sufficient processing capability are required for real-time operation.

3.5 Functional Requirements

The functional requirements describe the main functions that the proposed Indian Sign Language Text/Speech Translation system should provide.

The system should allow the user to start the camera and capture ISL gestures. It should process the captured image or video and identify the relevant hand or gesture information. The system should then recognize the sign using the trained machine learning or deep learning model.

The system should provide the following functions:

1. **Camera Input:** Capture ISL gestures through a webcam or camera.
2. **Preprocessing:** Improve and prepare the captured image or video for recognition.
3. **Feature Extraction:** Extract useful hand or gesture features from the input.
4. **Sign Recognition:** Identify the corresponding ISL sign using the trained model.
5. **Text Generation:** Convert the recognized sign into understandable text.
6. **NLP Processing:** Arrange or process recognized signs to form meaningful output.
7. **Speech Generation:** Convert the generated text into speech.
8. **Output Display:** Display the translated text and provide audio output.
9. **Performance Evaluation:** Calculate measures such as accuracy, precision, recall, F1-score, recognition time, and error rate.

These functions provide a complete process from ISL camera input to text and speech output. Similar combinations of recognition and output generation have been explored in previous ISL translation studies. [1], [2]

3.6 Non-Functional Requirements

The non-functional requirements define the quality and performance standards that the proposed ISL Text/Speech Translation system should maintain.

Accuracy: The system should recognize ISL gestures with suitable accuracy and minimize incorrect predictions.

Usability: The interface should be simple and easy to understand so that users can operate the camera and view the translated output without unnecessary difficulty.

Performance: The system should provide the translated result within a reasonable time, particularly when real-time translation is required.

Reliability: The system should provide consistent results when similar gestures are presented under similar conditions.

Security and Privacy: Camera input and any stored user information should be handled carefully. Unnecessary personal or video data should not be stored.

Scalability: The system should allow additional ISL signs, phrases, languages, and improved recognition models to be added in future versions.

Maintainability: The software should be organized into separate components such as preprocessing, recognition, text generation, and speech generation so that individual components can be modified or improved.

3.7 Assumptions

The proposed Indian Sign Language Text/Speech Translation system is developed based on certain assumptions about the user and operating environment. It is assumed that the user will perform the ISL gesture within the camera's visible area and that sufficient lighting will be available for capturing the gesture.

It is also assumed that the camera is functioning properly and that the computer has sufficient processing capability to run the recognition model. The system assumes that the signs presented by the user belong to the set of signs represented in the training dataset. If an unknown or unsupported sign is presented, the system may not be able to provide the correct translation.

The system also assumes that the training dataset contains sufficiently representative examples of the selected ISL signs. Properly labelled data is important because the quality of training data directly affects the recognition model. Continuous translation systems additionally require suitable sentence-level or phrase-level data, such as the resources provided by ISLTranslate. [3]

Finally, it is assumed that the user understands the basic operation of the system and that the camera, software, and required libraries are correctly installed and configured.

IV. DIAGRAM DESIGN

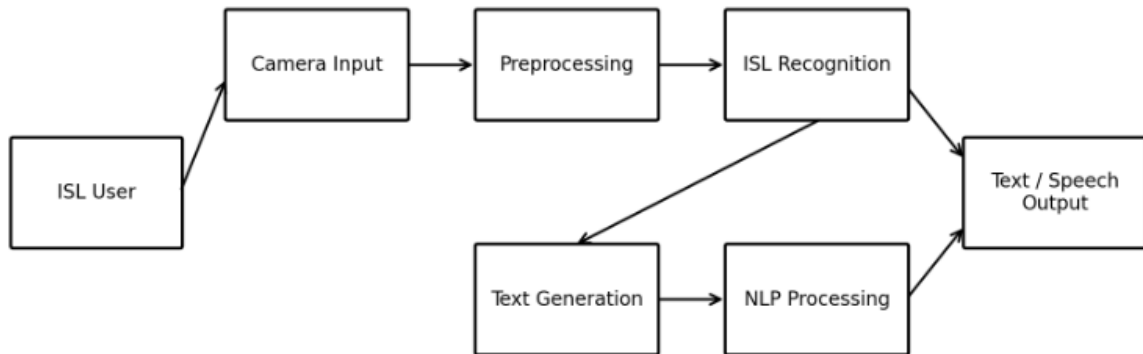
The diagram design of the proposed Indian Sign Language Text/Speech Translation system is used to present its structure, working process, and interaction between the user and system components. Data Flow Diagrams (DFD) show how ISL input moves through different processing stages, while UML diagrams represent the functions and structure of the system.

These diagrams help to understand the process of capturing ISL gestures, recognizing signs, generating text, and producing speech output.

4.1 Data Flow Diagram (DFD)

The DFD represents the flow of ISL data from camera input through preprocessing, feature extraction, sign recognition, text generation, and speech output.

Fig. 2. Data flow Diagram



4.2 UML Diagrams

UML diagrams provide a visual representation of the proposed system. They describe users, system functions, processing activities, data structures, and interactions between different components.

Fig. 3. Activity Diagram of the proposed Indian Sign Language Text/Speech Translation system.

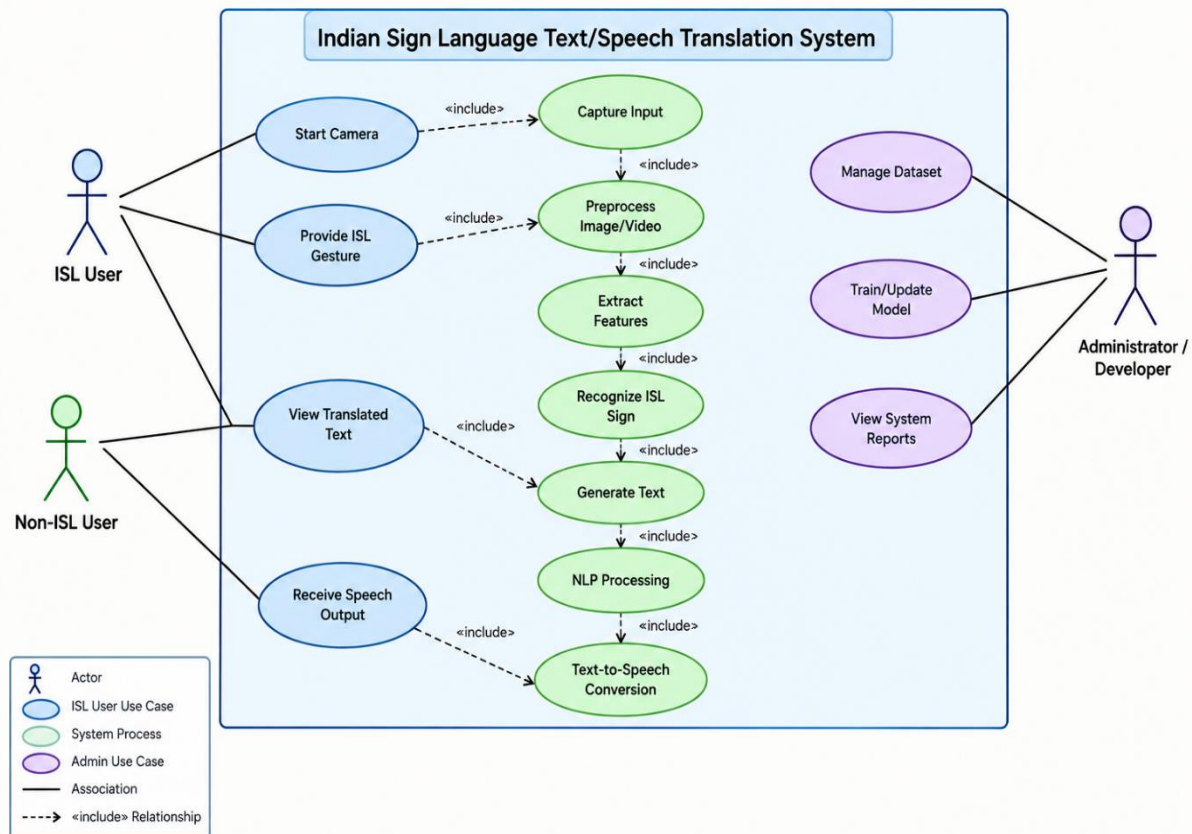


4.3 Proposed Use Cases

The proposed system mainly involves the **ISL User** and **Non-ISL User**. The ISL User performs sign gestures through a camera. The system captures and processes the gestures, recognizes the ISL signs, and converts them into text. The

generated text can then be processed using NLP and converted into speech. The Non-ISL User can read the text or listen to the generated speech.

Fig. 4. Use case diagram of ISL text/speech translation system



V. EXPECTED RESULTS AND DISCUSSION

The proposed Indian Sign Language Text/Speech Translation system is expected to provide a simple and accessible method for converting ISL gestures into understandable text and speech. Instead of depending completely on manual interpretation, the system will use camera input, image processing, feature extraction, and a machine learning or deep learning model to recognize the given ISL gesture. The recognized sign will then be converted into text and, where required, into speech using a text-to-speech component. This can make communication between ISL users and people who do not understand sign language easier.

The effectiveness of the proposed system can be evaluated using different performance measures such as recognition accuracy, precision, recall, F1-score, recognition time, and error rate. The system can be tested using a suitable collection of ISL images or video samples. Different conditions such as lighting, background, hand position, camera angle, and variations between users can also be considered during testing. For continuous sign translation, the quality of the generated sentence or phrase can additionally be evaluated using suitable language-translation measures.

The expected results can be compared with the performance reported by existing ISL recognition and translation approaches. A pilot implementation with selected ISL gestures can provide practical information about recognition performance, response time, and usability. The results obtained during testing can be used to identify incorrectly recognized signs and improve the training data, preprocessing methods, or recognition model.

Overall, the proposed system is expected to demonstrate the feasibility of using computer vision, deep learning, NLP, and text-to-speech technologies for ISL translation. It is expected to provide both text and speech output and create a foundation for future development of continuous and multilingual ISL translation systems.

5.1 Expected Benefits

The proposed Indian Sign Language Text/Speech Translation system is expected to provide several benefits for improving communication accessibility. It can provide translated text and speech without requiring the other person to understand ISL. This may be useful in educational institutions, healthcare environments, offices, public services, and other everyday communication situations.

The system can provide a faster and more convenient method for communicating basic ISL information. The use of a camera as the input device can make the system easier to use because users do not necessarily need specialized hardware. The combination of text and speech output can also support different communication requirements.

Another expected benefit is the possibility of extending the system with additional ISL signs, phrases, and Indian languages. Continuous ISL translation research, such as the ISLTranslate dataset, shows the potential for moving beyond individual sign recognition toward sentence-level translation.

The system can also provide a foundation for future mobile and web-based applications. Improvements in recognition models, datasets, preprocessing techniques, and language-processing methods may further improve the usefulness of the proposed system.

5.2 Limitations

Although the proposed Indian Sign Language Text/Speech Translation system can help reduce communication barriers, it has some limitations. The recognition accuracy depends on the quality and size of the training dataset. If the dataset contains a limited number of signs or does not represent sufficient variations between users, the system may not recognize some gestures correctly.

Environmental conditions can also affect recognition performance. Changes in lighting, background, camera position, hand orientation, distance from the camera, and speed of gesture movement may result in incorrect recognition. Real-time recognition may also require sufficient processing power, particularly when deep learning models are used.

Another limitation is that the proposed system may initially support only a selected set of ISL signs, words, or phrases. Continuous sign language translation is more complex because the system must understand a sequence of signs and produce meaningful language. The availability of suitable continuous ISL datasets is therefore an important consideration.

The generated speech also depends on the accuracy of the recognized text. If the sign is incorrectly recognized, the resulting text and speech may also be incorrect. Therefore, the proposed system should be considered as a technological communication-support system rather than a complete replacement for professional human interpretation in all situations. Finally, the system may require further testing with different users and real-world environments before it can be considered suitable for large-scale use. Additional work is also required for multilingual output, complex sentences, regional variations in signing, and improved real-time performance.

VI. CONCLUSION

Effective communication is an important part of social inclusion and education, especially for people who use Indian Sign Language (ISL). The proposed Indian Sign Language Text/Speech Translation system provides a way to reduce communication barriers by converting ISL gestures into understandable text and speech. The system uses camera input, image or video preprocessing, feature extraction, machine learning or deep learning techniques, sign recognition, text generation, and text-to-speech conversion. By combining these processes into one system, it can provide a more convenient method for communicating with people who may not understand ISL.

The proposed system also provides a foundation for further improvements in the future. Additional features such as continuous sign recognition, multilingual text and speech output, mobile applications, improved datasets, real-time translation, and context-aware recognition can be integrated to make the system more effective. Before implementing the system on a larger scale, it should be tested with different users, signing styles, backgrounds, lighting conditions, and real-world environments. The performance can be evaluated using measures such as accuracy, precision, recall, F1-score, recognition time, and error rate. Feedback from ISL users and other users can help identify limitations and improve the system according to practical communication requirements.

VII. FUTURE SCOPE

The proposed Indian Sign Language Text/Speech Translation system has good potential for further development as Artificial Intelligence, Computer Vision, and Natural Language Processing technologies continue to improve. In the future, the system can be enhanced to recognize continuous ISL sentences instead of only individual signs or words. This can help provide more natural and meaningful translation during real-time communication.

The system can also be extended to support multiple Indian languages such as Marathi, Hindi, and other regional languages. Text generated from recognized ISL gestures can be translated into the selected language and converted into speech. A mobile application can also be developed so that users can access the translation system using smartphones and cameras.

Further improvements can include the use of larger and more diverse ISL datasets containing different users, hand shapes, movements, facial expressions, and signing conditions. Deep learning models can be improved to increase recognition accuracy and reduce translation time. The system can also be designed to work under different lighting conditions, backgrounds, camera angles, and distances.

In addition, Natural Language Processing (NLP) can be used to improve sentence formation and provide more meaningful translations. Context-aware and emotion-aware recognition can also be explored to understand the meaning of signs more effectively. Cloud-based deployment can allow the system to be updated with new signs and models and accessed from different devices. These future enhancements can make the proposed system more accurate, accessible, real-time, and useful for communication between ISL users and non-ISL users.

REFERENCES

- [1]. Khetam, S. S., Murkute, K., Surve, S., Bhunje, R., Raut, P., & Kumar, A. (2025). Indian Sign Language to Text/Speech Translation: A Deep Learning. *International Journal of Sciences and Innovation Engineering (IJSCE)*, Vol. 2, Issue 4, pp. 11–15.
- [2]. Bontha, S. S. Y., Dastagiri, P., Sandeep, P., Reddy, T. S. S., & Bhanu, P. Y. (2025). Indian Sign Language to Text/Speech Translation. *International Journal of Scientific Development and Research (IJSER)*, Vol. 10, Issue 4, pp. a628–a632.
- [3]. Nithish, D., Naveen, M., Harish, S., AjithKumar, M., & Maheshwari, K. (2025). AI Based Indian Sign Language Translator. *International Journal of Scientific Development and Research (IJSER)*, Vol. 10, Issue 5, pp. c28–c31.
- [4]. Ingole, S., & Bakal, J. (2025). Interpretation of Indian Sign Language to Text and Speech to Communicate with Speech and Hearing Impaired Community. *Procedia Computer Science*, Vol. 258, pp. 1980–1992.
- [5]. Puneekar, A. C., Nambiar, S. A., Laliya, R. S., & Salian, S. R. (2020). A Translator for Indian Sign Language to Text and Speech. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, Vol. 8, Issue 6, pp. 1640–1646.
- [6]. Joshi, A., Agrawal, S., & Modi, A. (2023). ISLTranslate: Dataset for Translating Indian Sign Language. *Findings of the Association for Computational Linguistics: ACL 2023*, pp. 10466–10475.
- [7]. Brahmavar, A., Mandalam, A. V., Umesh, A., Monali, B., Agarwal, P., et al. (2025). Real-Time Speech Generation of Regional Languages Using Indian Sign Language. *Computer and Electrical Engineering*, 2025.
- [8]. Chouthe, A. U., & Railkar, P. N. (2026). Real-time Emotion and Context-Aware Indian Sign Language Translation System. *International Journal of Scientific Development and Research (IJSER)*, Vol. 11, Issue 5, pp. b704–b712.