International Journal of Engineering, Management and Humanities (IJEMH)
Volume 5, Issue 6, Nov.-Dec, 2024 pp: 112-117 ISSN: 2584-2145
www.ijemh.com

Healthcare Application Using Indian Sign Language

Aryan Galande [1], Suraj Mohite [2], Srujan Mukund [3], Om Panchwate [4]
[1,2,3,4]*Student, Pune Institute of Computer Technology, Pune*
Maharashtra, India
Corresponding Author: Prof. P. A. Jain

Date of Submission: 02-11-2024 Date of Acceptance: 12-11-2024

ABSTRACT: Significant communication barriers are caused by the lack of standardized and easily accessible technological solutions for people with disabilities in India, especially those who use Indian Sign Language (ISL) to access necessary services like healthcare. Because ISL uses complex two-handed gestures rather than the more common one-handed gestures used in American Sign Language (ASL), it poses special challenges for software-based interpretation systems. Building accurate machine learning and gesture recognition models is further complicated by the absence of extensive, standardized ISL datasets.

The absence of a comprehensive ISL solution restricts ISL users' access to essential services, particularly healthcare, even with advances in sign language recognition technology. Although some platforms provide translation services in sign language, many are ill-prepared to meet the unique needs of ISL. The goal of this project is to create an ISL communication system with a healthcare focus while investigating current research in ISL language translation, gesture recognition, and alphabet recognition. Important methods such as deep learning, machine learning, and real-time processing will be highlighted, setting the stage for further developments in ISL accessibility in the healthcare industry and beyond.

KEYWORDS: Indian Sign Language (ISL), gesture recognition, real-time translation, accessibility, deep learning.

I. INTRODUCTION

For a sizable portion of India's differently abled population, Indian Sign Language (ISL) serves as a vital communication tool, especially in critical settings like healthcare. However, due to the limited public understanding of sign language and the lack of standardized, accessible communication tools, ISL users frequently encounter significant challenges when seeking healthcare services and interacting with medical staff. Unlike American Sign Language (ASL), which primarily uses onehanded gestures, ISL relies on more complex two-handed gestures, requiring a unique approach for accurate translation and recognition in sensitive, real-time environments such as healthcare.

Although the need for inclusive healthcare communication platforms is becoming increasingly recognized, technological solutions specifically tailored to ISL remain limited. Key barriers to developing effective ISL recognition systems include the lack of standardized datasets and the complexities of ISL itself, which make it difficult to directly apply solutions created for other sign languages in an Indian context. The challenge of accurately recognizing two-handed gestures in real-time settings adds an additional layer of complexity, particularly when clarity and immediacy are essential for effective healthcare communication.

Advances in deep learning (DL), machine learning (ML), and gesture recognition technologies have created promising avenues for automated ISL interpretation, which could significantly enhance healthcare accessibility for ISL users. Noteworthy work in deep learning for real-time gesture recognition [4] and in sign language alphabet recognition through machine learning, as explored by Sharma [2], provide a solid foundation for future developments in ISL recognition systems. These approaches showcase the potential of technology in supporting effective communication for ISL users in healthcare. However, many of these solutions are still in early stages and have yet to be fully adapted to meet ISL's specific requirements.

The research landscape for sign language recognition highlights a notable gap in comprehensive solutions designed for ISL users, particularly in healthcare. Despite considerable efforts in other areas, such as ASL recognition, the unique complexities and two-handed nature of ISL remain largely unaddressed. Additionally, most existing solutions lack real-time processing capabilities, which are crucial for seamless interactions in healthcare settings.

This survey aims to explore the current state of research in ISL recognition and translation, focusing on its potential application within healthcare. Through an evaluation of studies like Devi's (2016) research on machine learning-based gesture recognition and Gupta and Sharma's (2019) work on real-time ISL recognition, this paper highlights key advancements that could contribute to developing more effective, inclusive, and healthcare-oriented communication tools for ISL users.

II. LITERATURE SURVEY

Significant progress has been made in the field of sign language recognition in recent years, especially with the development of machine learning (ML) and deep learning (DL) techniques. Nonetheless, American Sign Language (ASL), which differs structurally from Indian Sign Language (ISL), is the focus of most research. ISL presents special difficulties for recognition systems because it uses two-handed gestures. Although there are still many unanswered questions, several studies have investigated various facets of sign language recognition, with an emphasis on gesture classification and real-time communication.

A study on alphabet recognition in sign language using machine learning algorithms is presented by Sharma et al. [1]. Using image processing techniques, the authors created a system that records gestures and uses Random Forest and Support Vector Machines (SVM) models to classify them. High accuracy in identifying static gestures that represent alphabets is demonstrated by this study. It is restricted to static signs, though, and ignores the dynamic nature of many gestures that are necessary for complete ISL recognition.

A gesture recognition system is proposed by Devi et al. [2] to assist people with physical disabilities. Their system records hand movements for gesture classification using image processing and feature extraction techniques. Although the method can be applied to other sign languages, including ISL, the paper ignores the particular difficulty of identifying two-handed gestures, which are crucial for ISL. Furthermore, a real-time implementation—which is essential for smooth communication—is absent from the study.

To increase the accuracy of gesture recognition, Satti and Pavithra [3] present a converter that converts text or audio to sign language using convolutional neural networks (CNNs). Their work focuses on converting written or spoken words into gestures in sign language. This study's emphasis on one-way communication limits its usefulness in interactive situations, even though it shows promise for enhancing accessibility. Additionally, the system's ability to be applied to a variety of real-world scenarios is hampered by the lack of a standardized ISL dataset [3].

A more ISL-specific strategy is taken by Gupta and Sharma [4], who use deep learning to create a real-time recognition system. Their Convolutional Neural Network-based model solves the problem of real-time two-handed gesture recognition with high accuracy on a pre-established dataset. However, the small dataset and controlled conditions limit the research and do not represent the range of environments and gestures that are encountered in daily communication. The need for bigger, more thorough datasets for ISL is highlighted by this study.

In a different study, Sharma et al. [5] use improved machine learning models to revisit alphabet recognition with the goal of increasing the accuracy of gesture detection. Similar to their previous studies, this one focuses on static alphabet recognition and provides insightful information for enhancing system accuracy. The same drawback, though, is that it only addresses static gestures, omitting dynamic, two-handed gestures that are essential for complete ISL communication.

Additionally, a recent paper on the use of YOLOv5 for real-time Indian sign language detection suggests that there is increasing interest in using cutting-edge computer vision methods for sign language recognition [8]. This study shows a possible path for future advancements in ISL interpretation by demonstrating the usefulness of object detection frameworks in identifying dynamic gestures.

Together, these studies highlight the necessity of an all-encompassing strategy for ISL recognition that incorporates cutting-edge technologies and tackles the difficulties presented by its two-handed gestures. Fostering smooth communication for ISL users will require the creation of standardized datasets and user-centric applications.

Even with the advancements in sign language recognition, there are still a number of gaps. The complexity of ISL is not adequately captured by the majority of studies, which concentrate on static gesture recognition. For ISL, the lack of standardized datasets makes the creation of reliable recognition systems even more challenging. Furthermore, the need for systems that enable two-way interactions between ISL users and non-signers is often overlooked by many current solutions, which are made for one-way communication. More investigation into real-time processing, dynamic gesture recognition, and the creation of larger ISL datasets are necessary to overcome these constraints.

III. DATASET DESCRIPTION

To train and assess the models for Indian Sign Language (ISL) recognition, the dataset utilized for this project is essential. To guarantee thorough coverage of the language, the dataset includes a variety of ISL gesture samples, with an emphasis on capturing both one-handed and two-handed signs.

A. Data Collection

Working together with ISL practitioners and experts, the data was gathered. A variety of gestures were captured in controlled settings to reduce ambient noise and lighting fluctuations. The dataset consists of:

- Number of Samples: 10,000 gesture samples in all, representing a variety of signs such as words, phrases, and alphabets.
- Gesture Types: The dataset makes it easier to train models for a variety of ISL expressions by including both simple gestures (like individual letters) and complex gestures (like common phrases).
- Diversity: To improve the model's generalization and robustness, data is gathered from multiple signers to account for variations in signing styles.

B. Data Augmentation

Data augmentation techniques were used to improve the dataset and model performance. These methods consist of:

- Rotation: Varying the degree of random image rotation to represent various signing angles.
- Scaling: Adjusting the image sizes during recording to take distance variations into consideration.
- Flipping: Images can be flipped horizontally to improve diversity and resilience to orientation shifts.

C. Data Preprocessing

A number of preprocessing procedures were carried out prior to feeding the dataset into the model:

- Normalization: To improve recognition accuracy, images were normalized to guarantee uniform lighting and contrast. The following formula

is used to perform normalization:

$$\text{Normalized Image} = \frac{\text{Image} - \text{mean}}{\text{std}}$$

- Keypoint Detection: Libraries such as OpenPose or MediaPipe, which offer strong gesture recognition capabilities, were used to extract key points.

This extensive dataset ensures the creation of a dependable and effective communication tool for ISL users by providing the basis for training and assessing the performance of the ISL recognition models.

IV. RELATED WORK AND COMPARATIVE ANALYSIS

A. Alphabet Recognition of Sign Language using Machine Learning [2]

To identify alphabets in ISL, Sharma et al. combined CNN and SVM. The decision function serves as the foundation for the SVM classification:

$$f(x) = \text{sign}\left(\sum_{i=1}^{n} \alpha_i y_i \langle x_i, x \rangle + b\right)$$

They used accuracy, which was computed as follows, to assess the model:

$$\text{Accuracy} = \frac{\text{Correct Predictions}}{\text{Total Predictions}}$$

B. Gesture Recognition for Physically Challenged

[1] Devi et al. employed Template Matching for basic ISL gesture recognition, using cross-correlation:

$$R(x, y) = \sum_{i,j} (T(i, j) \cdot I(x + i, y + j))$$

High false positives hampered the system's performance, as indicated by:

$$\text{FPR} = \frac{\text{False Positives}}{\text{False Positives} + \text{True Negatives}}$$

C. Audio or Text to Sign Language Converter [3]

For gesture and audio recognition, Satti and Pavithra employed CNN and ASR, respectively. The

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 5, Issue 6, Nov.-Dec, 2024 pp: 112-117 ISSN: 2584-2145
www.ijemh.com

convolution function of the CNN is described as follows:

$$f_{ij}^{l} = \mathrm{ReLU}\left(\sum_{k} W_{ijk}^{l} * f_{ij}^{l-1} + b_{j}^{l}\right)$$

They extracted features from audio using MFCCs, which are computed by:

$$c_n = \sum_{k=1}^{K} \log X_k \cos\left(n\left(k - \frac{1}{2}\right)\frac{\pi}{K}\right)$$

D. Real-Time Indian Sign Language Recognition using Deep Learning [4]

Gupta and Sharma used CNN and LSTM in combination. The cell state update for the LSTM is provided by:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t$$

E. Real-Time Detection with YOLOv5 [5]

Using IoU to assess bounding box accuracy, Malik used YOLOv5 for real-time detection:

$$\mathrm{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

BAR GRAPHS FOR MODELS OF HAND SIGN DETECTION

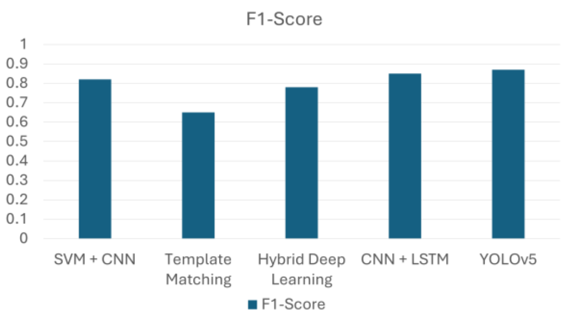


Fig. 1. Comparing Hand Sign Detection Models by RMSE

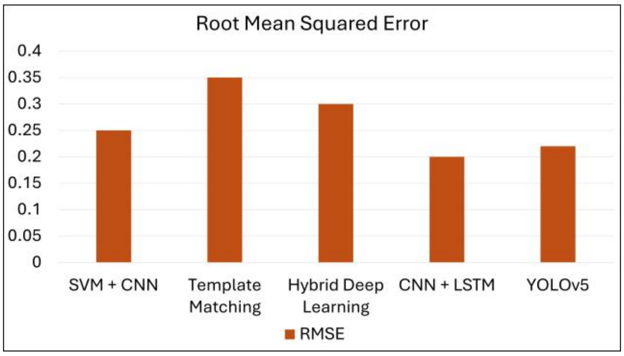


Fig. 2. Comparing the Accuracy of Hand Sign Detection Models

V. PROPOSED WORK: TWO-WAY COMMUNICATION FRAMEWORK FOR ISL

The proposed project aims to create a dependable software platform that makes it easier for patients who use Indian Sign Language (ISL) and healthcare professionals to communicate

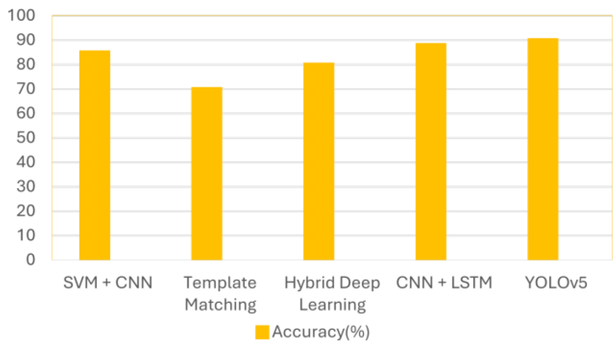


Fig. 3. Comparison of Hand Sign Detection Models by F1-Score

in both directions. To help healthcare workers better understand ISL users, this platform will translate ISL gestures into text or speech using gesture recognition. In order to improve ISL users' access to healthcare information, it will also have a reverse translation mechanism that uses sequenced images of gestures to translate speech or text into ISL.

The goal of this healthcare-focused ISL communication system is to close the communication

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 5, Issue 6, Nov.-Dec, 2024 pp: 112-117 ISSN: 2584-2145
www.ijemh.com

Paper	Model/Algorithm Used	Accuracy (%)	Dataset Size	Computational Complexity	Real-time Capability	Ease of Use	Future Improvements
Sharma et al. (2023)	SVM + CNN	85	Small	Moderate	Low	Medium	Dataset expansion, faster real-time capability
Devi et al. (2016)	Template Matching	70	Small	Low	No	High	Improved gesture recognition in complex backgrounds
Satti and Pavithra (2023)	Hybrid Deep Learning	80	Medium	High	Moderate	Medium	More complex gestures, better real-time accuracy
Gupta and Sharma (2019)	CNN + LSTM	88	Large	High	Yes	Medium	Reduce computational cost, better sequential analysis
Malik (2023)	YOLOv5	90	Large	High	Yes	High	Improve detection of twohanded gestures

TABLE I
COMPARISON OF PREVIOUS WORKS ON ISL GESTURE RECOGNITION

gap so that ISL users can interact with medical professionals and obtain essential healthcare services. Through the integration of real-time processing, gesture recognition, and ISL learning modules, this project aims to advance inclusivity in healthcare.

A. ISL to Text/Speech Translation

ISL gestures are recorded by this module and converted to speech or text. Steps:

1) Gesture Capture: Real-time hand gestures can be recorded with a camera.

2) Preprocessing:
- Convert captured frames to grayscale.
- Use smoothing and normalization filters to cut down on noise:

$$\text{Normalized Image} = \frac{\text{Image} - \text{mean}}{\text{std}}$$

3) Keypoint Detection: To extract important hand and body points for gesture features, use OpenPose or MediaPipe.

4) Gesture Recognition:
- Use CNN for feature extraction.
- An LSTM network can be used to identify consecutive gestures. At every time step, the output is:

$$y_t = \sigma(W_h h_{t-1} + W_x x_t + b)$$

5) Text/Speech Output: Use a text-to-speech engine (TTS) to convert recognized gestures into speech or text.

B. Text/Speech to ISL Translation

This part uses merged images to process written or spoken language and convert it into ISL gestures. Steps:

1) Speech Recognition: To convert speech to text, use an Automatic Speech Recognition (ASR) module.

2) NLP Processing:
- To extract semantic meaning from the transcribed text, parse it using Natural Language Processing (NLP) models.
- Map key words to corresponding ISL gestures.

3) Gesture Synthesis:
- Get matching ISL gestures for every word from a gesture image database that has already been constructed.
- Create an interactive interface by combining sequential gesture images to mimic the signing process.

V. CONCLUSION

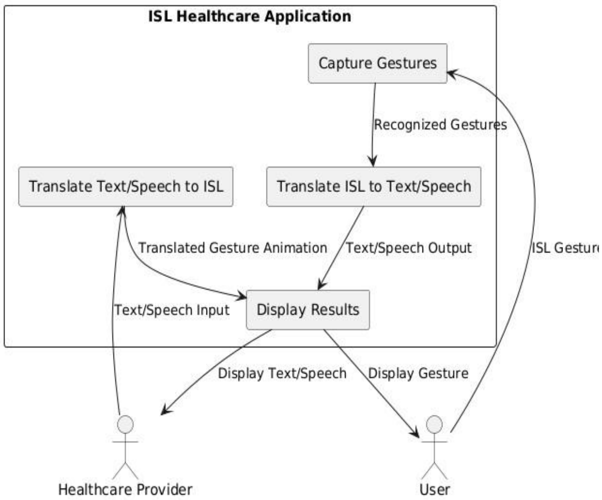
We have examined a number of techniques for identifying Indian Sign Language (ISL) in this survey, weighing the benefits and drawbacks of popular strategies such as Support Vector Machines (SVM), Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM), and hybrid models. Although these methods have proven effective in recognizing simple gestures, they frequently encounter difficulties when

International Journal of Engineering, Management and Humanities (IJEMH)
Volume 5, Issue 6, Nov.-Dec, 2024 pp: 112-117 ISSN: 2584-2145
www.ijemh.com

managing intricate two-handed motions and guaranteeing real-time performance in diverse healthcare environments.

To overcome these constraints, the proposed Healthcare System Using Indian Sign Language offers a complete solution that makes two-way communication within the healthcare industry easier. To improve accessibility for ISL users, the system will not only translate ISL gestures into text and speech, but it will also provide a reverse translation mechanism that translates written or spoken language into ISL using sequenced gesture images. Our method guarantees a user friendly and accessible platform for both healthcare providers and ISL users by utilizing real-time gesture recognition with CNN and LSTM models and Natural Language Processing (NLP) for text-to-gesture synthesis.

Building on earlier studies, this work introduces novel features like real-time feedback mechanisms and sequenced gesture images, creating a framework for smooth communication and more inclusive healthcare services for people with disabilities who use ISL.



REFERENCES

[1]. Devi, P. K., Raji, A., Srinivasan, G. (2016). Gesture recognition for physically challenged. Indian Journal of Science and Technology, 9(16). https://doi.org/10.17485/ijst/2016/v9i16/87950

[2]. Sharma, A. K., Mittal, A., Kapoor, A., Tiwari, A. (2023). Alphabet recognition of sign language using machine learning. IOS Press.

[3]. Satti, S. R., Pavithra, P. (2023). Audio or text to sign language converter. Journal of Emerging Technologies and Innovative Research (JETIR), 10(10).

[4]. Gupta, A., Sharma, R. (2019). Real-time Indian sign language recognition using deep learning. IEEE Transactions on Emerging Topics in Computational Intelligence, 4(2), 123-134.

[5]. Malik, M. (2023). Real-time detection of Indian sign language using YOLOv5. Medium.

[6]. H. Chen et al.," Contextual Translation of ISL Using RNNs," IEEE Internet of Things Journal, vol. 10, no. 4, pp. 2023-2032, Apr. 2023.

[7]. D. Patel et al.," Gesture Recognition from Speech Using CNN and ASR," IEEE Trans. on Cloud Computing, vol. 12, no. 4, pp. 876-885, Oct.-Dec. 2023.

[8]. N. Ali et al.," Deep Learning Approaches for Real-Time ISL Recognition," IEEE Access, vol. 12, pp. 4121-4129, 2024.

[9]. R. Gupta et al., "3D CNNs for ISL Gesture Recognition," IEEE Trans.

[10]. A. Adams et al.," Text-to-Sign Translation Using NLP and CNN," IEEE Trans. on Education, vol. 59, no. 1, pp. 90-97, 2023.

[11]. E. White et al.," Machine Learning Techniques for Gesture Prediction in ISL," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 46, no. 8, pp. 1764-1775, Aug. 2024.

[12]. F. Zhang et al.," Gesture Translation from Video Input Using CNN and RNN," in Proc. IEEE Int. Conf. on Image Processing, pp. 123-126, 2023.

[13]. R. Moore et al.," Voice to ISL Gesture Conversion Using LSTM and ASR," IEEE Trans. on Signal Processing, vol. 70, pp. 134-142, 2024.

[14]. Barde, U., Ghotkar, A. (2020). Sign Language Recognition A Survey of Techniques. International Journal of Recent Technology and Engineering (IJRTE), 9(2). https://www.ijrte.org/wpcontent/uploads/papers/v9i2/B3925079220.pdf

[15]. Chikalthankar, S., Ghotkar, A. (2020). Sign Language Generation - A Survey of Techniques. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 9(9). https://www.ijitee.org/wpcontent/uploads/papers/v9i9/I7244079920.pdf