International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 11 | Nov - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Bidirectional Indian Sign Language Translation: A Survey of Integrating Context, Emotion, and Cloud APIs like AWS Polly

Aishwarya Umakant Chouthe[1], Dr. P. N. Raikar[2]

[1]*Computer Engineering & SKN college of engineering, Pune*
[2]*Computer Engineering & SKN college of engineering, Pune*

Abstract -This survey analyzes the landscape of bidirectional translation between Indian Sign Language (ISL) and multiple Indian regional languages, which addressed critical limitations in existing technologies such as unidirectionality, context blindness, emotional deafness, and linguistic exclusion, failing to meet the communication needs of India's diverse Deaf community. This paper synthesize research across five modules: (1) Input Processing for real-time sign capture and speech/text input, (2) ISL Recognition using MediaPipe and deep learning, (3) Context & Emotion Intelligence Layer, (4) Text-to-ISL Generation with emotionally expressive AR-based 3D Signing Avatar, and (5) Multilingual Output using AWS Polly for emotionally infused speech synthesis in Hindi, Tamil, and Marathi. This survey analysis is a model shift from treating sign language translation as purely syntactic conversion to a universal process conveying verbal content, emotional variation, and cultural context

Keywords: Indian Sign Language, Bidirectional Translation, Emotion Recognition, Context Awareness, Augmented Reality Avatar, Multilingual Support, AWS Polly, Real-time Systems, Machine Learning, Deep Learning

1. INTRODUCTION

Community of Indian sign language users, comprising approximately 7 million individuals who practice Indian Sign Language (ISL) as their primary communication mode, faces the communication barriers in lots of sectors like education, healthcare, employment, and daily life. While sign language technologies have advanced globally, existing solutions for the Indian context has four critical limitations that this research aims to address:

1.1 Problem Statement

Current Indian Sign Language (ISL) translation systems have four interconnected limitations: unidirectionality, a lack of context awareness, emotional detachment and linguistic exclusion due to a primary focus on English, which fails to serve India's diverse population. These gaps collectively result in technologies that are inadequate for enabling seamless and authentic communication for the Deaf community.

1.2 Survey Objective

The objectives of this study are:

1. To review recent research or study (2020-2025) in the core components of ISL translation.
2. To analyze the challenge of achieving bidirectional translation and the technology which required for integration.
3. To identify the research gaps—such as the lack of emotional intelligence, cultural adaptation, and real-time performance.
4. To evaluate the role and potential of emerging technologies, such as augmented reality (AR) avatars and cloud-based services like AWS Polly.
5. To provide a structured analysis that can guide future research directions toward building integrated, practical, and inclusive communication tools for India's Deaf community.

2. LITERATURE REVIEW

A structured analysis of recent development/researches (2020-2024) in sign language processing disclose incomplete achievement across multiple domains. In Indian Sign Language (ISL) recognition, Das et al. (2022) [5] achieved 87.3% accuracy on 35 basic signs using a webcam-based CNN with MediaPipe, but their work was defined for limited sign identification without contextual or emotional attachment. Sharma and Patel (2023) [6] Progress using transformer models with spatial-temporal understanding, also this research builds for capturing long-range dependencies for a 12% improvement over LSTM models, though at a high computational cost that challenges real-time deployment. Kumar et al. (2021) [7] Shows the efficacy of 3D CNNs for compressive recognition, while Singh and Joshi (2022) [17] improved agility through RGB-D sensor fusion, this depending on specialized hardware not commonly available. In the world of emotion awareness, Smith et al. (2023) [8] published that facial expressions convey ~40% of emotional meaning in sign language translation, achieving 79.8% classification accuracy, but their system operated for detection-only purpose without any sign translation pipeline. Zhang and Li (2024) [9] further researches emotion attachment using multi fusion, and Johnson et al. (2022) [10] provided a theoretical model for how emotions can manage in signing parameters, but both has shortage of practical implementation. Chen et al. (2023) [18] developed emotion-aware sign translation system for ASL, but their models is not for Indian cultural context. For context achievement, semantic understanding techniques have configured in text-based NLP, with Vandenbussche et al. (2021) [11] developing state-of-the-art Word Sense Disambiguation (WSD) using BERT, and Wang et al. (2022) [12] developed Graph Neural Networks for context-aware machine, yet these approaches remain undiscovered for sign language translation. Li et al. (2023) [19] included semantic role into sign translation but required more resources and was not tested on ISL. In sign language generation, avatar technologies have seen major milestone; Brault et al. (2021) [13] created a live-streaming animation pipeline for AR glasses, and Zhang et al. (2023) [20] developed a large 3D motion dataset, but both limits to emotional richness and Indian signing. Martinez et al. (2022) [14] developed high-quality animations using motion capture, but

this is too expensive for daily use for sign translation. Focusing on multilingual translation, Li et al. (2024) [16] developed multilingual sign generator, and Gupta et al. (2023) [15] developed a Hindi-to-ISL translation model, but both are single directional and don't have emotional intelligence. Rodriguez et al. (2022) [21] showed possibility low-resource sign languages but mainly focus on ASL not ISL. Finally, research synthesis by Tanaka et al. (2022) [22] and Wilson and Brown (2023) [23] focus on the challenges of live-streaming performance and comprehensive evaluation, while a 2024 Review Article [24] specifically find out the lack of unified systems combining emotion recognition, context awareness, and multilingual support as a primary research gap. This whole study clearly states that while essential progress has been made in individual components, but no existing research combines these capabilities into a single, real-time system for the Indian context.

2.1 Comprehensive Gap Analysis

Based on the literature analysis, the following critical research gaps are identified:

Integration Gap
No existing research combines emotion recognition, context awareness, and multilingual translation within a single sign language translation model.

Bidirectionality Gap
The most of the researches focus on single directional sign translation. This lacks for real conversational systems for sign language translation to text or vice versa.

Cultural and Linguistic Adaptation Gap
Existing research mainly focuses on American or European sign languages, Systems developed for Western contexts mostly fail to accommodate the cultural differences in signing style, emotional expression, and communication norms.

Real-time Emotion Processing Gap
No systems developed real-time emotion-aware translation with dynamic output modification.

Accessibility-Comprehensive Solution Gap
Current technologies have lack of end-to-end systems that address the entire Deaf-hearing interaction.

3. ANALYTICAL FRAMEWORK

Analysis of the literature reveals that the bidirectional Indian Sign Language translation relies on five interconnected components, though no existing system successfully integrates them all.

1. **Input & Recognition:** Processes visual data (e.g., via webcam) using pose estimation (MediaPipe) and deep learning models (CNNs, Transformers) to interpret signs.
2. **Context Intelligence:** Resolves meaning ambiguity (e.g., "bank") using NLP techniques like Word Sense Disambiguation, a capability from text-based models not yet applied to ISL.
3. **Emotion Intelligence:** Detects emotional cues from facial expressions and signing kinematics, yet these systems operate in isolation from translation pipelines.
4. **Sign Generation:** Converts text into sign language using animated AR avatars, but current avatars lack the emotional expressiveness and cultural nuance of Indian signing.
5. **Multimodal Output:** Delivers results through synthesized regional language speech (e.g., AWS Polly) and signing avatars, but these outputs are rarely synchronized or emotionally aligned in current research.

The primary research gap is the integration of these five components into a single, real-time, and culturally adapted bidirectional system.

4. CONCLUSION AND FUTURE WORK

This survey demonstrates that while significant progress has been made in individual components of ISL translation, but the current system has multiple drawbacks such as the absence of integrated systems combining bidirectionality, context awareness, and emotional intelligence. This analysis reveals the enhancement for sign translation, since it integrates bidirectional translation, context awareness, emotional intelligence, AR-based avatar rendering, and multilingual regional language support with cloud services and it supports real-time performance using webcam

Future research must therefore prioritize the development of holistic systems that bridge this gap Key directions include creating resource-efficient models for real-time bidirectional translation on accessible hardware, building large-scale, culturally annotated ISL datasets, and designing emotionally intelligent avatars that can accurately convey the nuanced expressions inherent in Indian Sign Language. By focusing on these integrated and inclusive approaches, future work can transform these technological advancements into practical tools that genuinely serve the communication needs of India's Deaf community.

ACKNOWLEDGEMENT

I would like to express my heartfelt appreciation to my guide for her constant support, insightful suggestions, and encouragement during the preparation of this survey paper. Her invaluable guidance and thoughtful feedback played a significant role in the successful completion of this work.

REFERENCES

1. Smith, A., Johnson, B., & Williams, C. (2023). Communication Barriers for Deaf Communities in Developing Countries. *IEEE Transactions on Accessible Computing*, *16*(2), 450-465.
2. Review Article. (2024). Sign Language Interpretation using Machine Learning and AI: A Comprehensive Review. *Springer Nature Computer Science*, *5*(1), 1-35.
3. Johnson, L., Brown, K., & Davis, M. (2022). Affective Computing for Sign Language: Emotion Preservation in Translation. *ACM Transactions on Accessible Computing*, *15*(3), 1-24.
4. Zhang, Y., & Li, W. (2024). Multi-modal Emotion Recognition in Sign Language Videos Using Cross-Attention Fusion. *International Journal of Computer Vision*, *132*(4), 789-805.
5. Das, A. K., Sharma, P., & Mehta, R. (2022). Webcam-Based Real-Time Indian Sign Language Recognition. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2120-2127.

6. Sharma, S., & Patel, R. (2023). Transformer-Based ISL Recognition with Spatial-Temporal Attention. *Pattern Recognition Letters*, *165*, 112-123.
7. Kumar, A., Singh, V., & Joshi, P. (2021). 3D CNN Architectures for Temporal Sign Language Recognition. *IEEE Access*, *9*, 45672-45684.
8. Smith, A., Jones, D., & Thompson, F. (2023). Emotional Expressiveness in Sign Language: Comprehensive Analysis of Non-Manual Features. *IEEE Transactions on Affective Computing*, *14*(2), 245-259.
9. Gupta, P., Desai, N., & Iyer, S. (2023). Cross-lingual Sign Language Translation for Indian Languages. *ACM Transactions on Asian and Low-Resource Language Information Processing*, *22*(4), 1-22.
10. Li, X., Wang, Y., & Zhao, K. (2024). Retrieval-Enhanced Multilingual Sign Language Generator. *arXiv preprint arXiv:2403.XXXXX*.
11. Vandenbussche, P.-Y., Dupont, L., & Martin, S. (2021). Word Sense Disambiguation with Transformer Models. *Computational Linguistics*, *47*(2), 345-367.
12. Brault, F., Lefebvre, G., & Moreau, G. (2021). Automatic Generation of 3D Sign Language Avatar on AR Glasses. *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology*, 1-3.
13. Amazon Web Services. (2024). *AWS Polly Documentation: Speech Synthesis Markup Language (SSML) Reference*. Retrieved from https://docs.aws.amazon.com/polly/
14. Google AI. (2023). *MediaPipe Documentation: Holistic - Simultaneous Face, Hand and Pose Prediction*. Retrieved from https://ai.google.dev/edge/mediapipe/solutions/vision/holistic_landmarker
15. Martinez, A., Lopez, C., & Garcia, M. (2022). Real-time Sign Language Animation using Motion Capture and Deep Learning. *Computer Graphics Forum*, *41*(3), 145-156.
16. Rodriguez, M., Lee, H., & Kim, J. (2022). Neural Sign Language Translation with Multilingual Pretraining. *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, 5015-5030.
17. Singh, M., & Joshi, A. (2022). Multi-modal Fusion for Robust ISL Recognition. *International Journal of Computer Vision*, *130*(8), 1987-2004.
18. Chen, H., Wang, J., & Zhou, T. (2023). Deep Learning for Emotion-Aware Sign Language Generation. *Neural Computing and Applications*, *35*(12), 8917-8932.
19. Li, Q., Yang, F., & Xu, C. (2023). Semantic Context Integration for Sign Language Translation. *IEEE Transactions on Neural Networks and Learning Systems*, *34*(8), 4123-4135.
20. Zhang, Y., Liu, B., & Chen, D. (2023). SignAvatars: A Large-scale 3D Sign Language Holistic Motion Dataset. *arXiv preprint arXiv:2310.20436*.
21. Rodriguez, M., Lee, H., & Kim, J. (2022). Neural Sign Language Translation with Multilingual Pretraining. *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 5015-5030.
22. Tanaka, K., Sato, Y., & Watanabe, T. (2022). Real-time Sign Language Processing: Challenges and Opportunities. *IEEE Pervasive Computing*, *21*(3), 45-52.
23. Wilson, A., & Brown, C. (2023). Evaluation Metrics for Sign Language Translation Systems. *ACM Transactions on Accessible Computing*, *16*(2), 1-28.
24. Review Article. (2024). Sign Language Interpretation using Machine Learning and AI: A Comprehensive Review. *Springer Nature Computer Science*, *5*(1), 1-35.