

Signify Connect: A Review of AI-Based Mobile Assistive Systems for Sign Language Recognition

Ms. Amrita Mishra

Assistant Professor
Department of
Computer Science
and Engineering
BBDITM
Lucknow, India

Abdur Rahmaan

Department of
Computer Science
and Engineering
BBDITM
Lucknow, India

Mohd. Hamza

Department of
Computer Science
and Engineering
BBDITM
Lucknow, India

Jaan Mohammad

Department of
Computer Science
and Engineering
BBDITM
Lucknow, India

Abstract- Effective communication remains a major challenge for individuals who are mute or deaf, especially in environments where sign language is not widely understood. Recent advancements in artificial intelligence (AI), deep learning, and computer vision have accelerated the development of intelligent assistive systems. This paper presents a comprehensive review of research conducted between 2015 and 2025 and examines the evolution of gesture-recognition and sign-language translation technologies. A conceptual framework, *Signify Connect*, is introduced as an integrated, AI-driven communication system capable of translating gestures into text or speech and converting spoken language into accessible formats. The paper highlights technological trends, identifies gaps in existing work, and outlines future directions for achieving a more robust, real-time, and culturally adaptive assistive communication ecosystem.

Keywords: Assistive communication, Signify Connect, sign language translation, gesture recognition, deep learning, computer vision.

I. INTRODUCTION

Individuals with speech and hearing impairments rely on non-verbal communication methods, primarily sign language, which is not universally known or interpreted by society. This gap creates social, educational, and professional barriers for mute and deaf individuals. While early assistive systems offered basic gesture recognition or text-to-speech support, they lacked scalability, contextual understanding, and real-time performance.

The rapid development of AI-driven technologies—such as convolutional neural networks (CNNs), transformers, multimodal learning, and lightweight mobile AI—has enabled more sophisticated communication tools. These advancements form the foundation of *Signify Connect*: a proposed multimodal communication framework that integrates gesture analysis, semantic interpretation, natural language processing (NLP), and speech synthesis.

This review paper analyzes major contributions from 2015 to 2025, evaluates system capabilities, and identifies the limitations that future research must address to create universally accessible assistive communication solutions.

Recent advancements in artificial intelligence have enabled computer vision and gesture-recognition systems capable of interpreting hand movements, body language, and facial expressions.

Flowchart:

Working flowchart of Signify Connect Proposed Approach

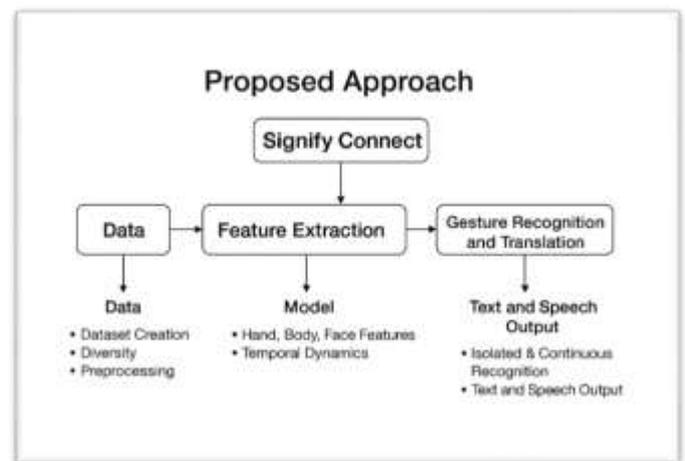


Figure 1.1 Flowchart of Signify Connect

II. RELATED WORK

A. Sensor-Based Recognition (2015–2017)

Early systems employed flex sensors, accelerometers, and glove-based devices. These methods offered high precision in capturing hand movements but suffered from high cost, limited comfort, and a lack of natural user interaction.

B. Vision-Based Static Gesture Detection (2016–2020)

Later studies shifted toward camera-based recognition using color segmentation, contour analysis, and classical computer vision. While effective in controlled environments, these methods struggled with background noise and lighting variations.

C. Deep Learning for Dynamic Gesture Understanding (2018–2022)

The introduction of CNNs, 3D CNNs, and LSTM architectures improved recognition of dynamic sign gestures. Systems using YOLO, OpenPose, or Mediapipe achieved near real-time detection with improved robustness. However, these models were often limited to isolated gestures rather than continuous sentences.

D. Transformer and Multimodal Architectures (2022–2025)

Recent work focuses on integrating multiple modes—hand gestures, facial expressions, pose estimation—with transformer models capable of understanding context and grammar. These systems move beyond word-level output and generate meaningful sentences, marking a significant advancement toward natural communication.

E. Identified Limitations

Despite progress, several challenges remain:

1. Limited datasets for continuous sign language, especially for Indian Sign Language (ISL)
2. High computational cost for real-time deployment on mobile devices
3. Difficulty recognizing overlapping or micro-gestures
4. Lack of standardized global sign language representations
5. Environmental dependency (lighting, motion blur, occlusion)

III. METHODOLOGY

This review is based on an analysis of peer-reviewed studies published between 2015 and 2025 in IEEE, Springer, Elsevier, and ACM digital libraries.

A. Selection Criteria

1. Focus on sign-language translation, gesture recognition, multimodal AI, and assistive technologies
2. Studies demonstrating novel architectures or real-time performance
3. Research targeting communication support for deaf or mute individuals

B. Categorization of Literature

The selected studies were grouped into:

1. Sensor-based techniques
2. Computer vision-based static gesture systems
3. Deep-learning-based dynamic gesture models
4. Transformer and multimodal architectures

C. Comparative Analysis Parameters

Each system was evaluated based on:

1. Accuracy and robustness
2. Computational complexity
3. Real-time performance
4. Language/model support
5. Dataset scale and diversity

D. Framework Understanding (Signify Connect)

The reviewed technologies were analyzed to conceptualize the *Signify Connect* architecture, comprising:

1. Gesture capture module
2. CNN/LSTM-based recognition pipeline
3. Transformer-based language generation
4. Speech synthesis
5. Reverse translation (speech-to-gesture support)

6. Multilingual and offline support.

IV. RESULT AND DISCUSSION

A. Strengths in Existing Approaches

1. Modern CNN and transformer models achieve **85–97% accuracy** on standard gesture datasets.
2. Lightweight models enable deployment on smartphones and edge AI devices.
3. Multimodal integration significantly enhances sentence-level translation and context awareness.

B. Challenges Observed

1. Lack of extensive datasets across regions and languages
2. Difficulty handling overlapping hands, rapid movements, and environment noise
3. Complex gestures requiring both hand shape and facial cues

C. Signify Connect Potential

The proposed system aims to overcome existing limitations through:

1. Real-time gesture tracking using optimized CNNs
2. Context-aware sentence formation using transformer models
3. Multimodal fusion (hand, face, pose) for robust detection
4. Offline capability for rural deployment
5. Multilingual speech output for cross-cultural communication

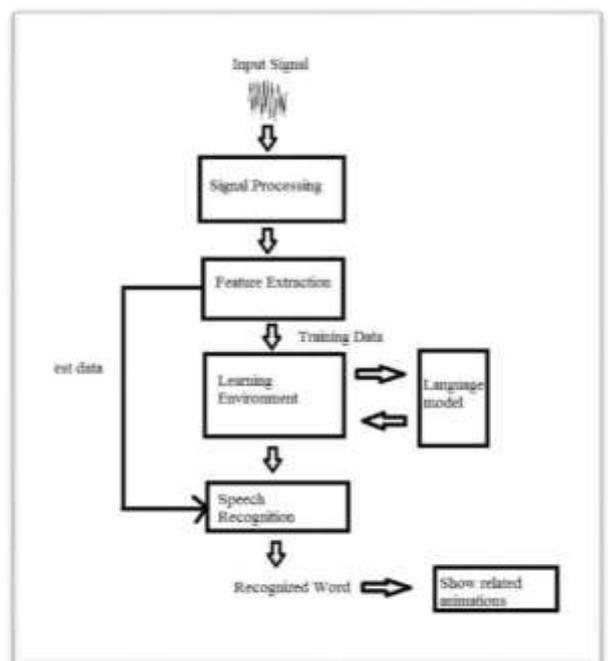


Figure 3.2 Architect of Signify Connect

V. CONCLUSION AND FUTURE WORK

A. Conclusion

This review highlights the major advancements in gesture recognition and AI-driven assistive systems over the last decade. While significant progress has been made, current systems remain limited in terms of dataset diversity, real-time contextual translation, and adaptability to complex natural gestures. The *Signify Connect* framework addresses these gaps by integrating multimodal recognition, contextual modelling, and speech interaction within a unified platform.

B. Future Work

1. Creating a large-scale unified global sign-language dataset
2. Improving micro-gesture recognition with high-resolution temporal models
3. Developing energy-efficient AI architectures for low-cost devices
4. Enhancing emotion-aware speech synthesis for expressive communication
5. Integrating AR/VR for immersive communication support
6. Expanding support for regional languages and dialects

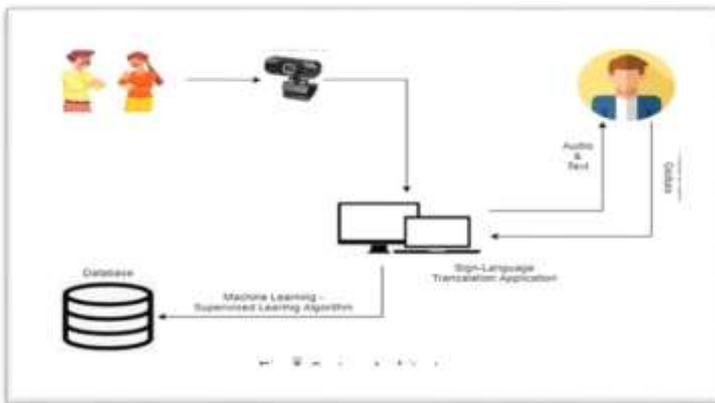


Figure 5.1 Gesture Recognition

VI. REFERENCES

- [1] Verma, A. R., Singh, G., Meghwal, K., Ramji, B., & Dadheech, P. K. (2024). Enhancing Sign Language Detection through Media Pipe and Convolutional Neural Networks (CNN). arXiv preprint.
- [2] Srivastava, S., Singh, S., Pooja, & Prakash, S. (2024). Continuous Sign Language Recognition System using Deep Learning with MediaPipe Holistic. arXiv preprint.
- [3] Abini, M. A., Lakshmi, D., Sharan, K. S., & Sulphiya, V. N. (2024). A Comprehensive Deep Learning Based System for Real Time Sign Language Recognition and Translation Using Raspberry Pi. *International Journal of Computer Trends and Technology*, 72(12), 8-16.
- [4] "Sign Language Recognition Using Convolutional Neural Network." (2024). *International Journal of Intelligent Systems and Applications in Engineering*, 12(17s), 329-337.
- [5] "Indian Sign Language Recognition Using Deep Learning Techniques." (2025). *International Journal of Computer Communication and Informatics (IJCCI)*.
- [6] "Sign Language Recognition with Deep Learning." (2025). *IJRASET – Journal for Research in Applied Science & Engineering Technology*.
- [7] "Novel Wearable System to Recognize Sign Language in Real Time." (2024). *Sensors*.
- [8] Duraisamy, P., Abinayasrijanani, A., Amrit Candida, M., & Dinesh Babu, P. (2023). Transforming Sign Language into Text and Speech through Deep Learning Technologies. *Indian Journal of Science and Technology*, 16(45), 4177-4185.
- [9] "Hand Gesture Recognition in Indian Sign Language Using Deep Learning." (2025). Demonstrated high accuracy for static gesture classification using CNN.
- [10] "Deep sign: Sign Language Detection and Recognition Using Deep Learning." Demonstrated LSTM/GRU based recognition of Indian Sign Language signs.
- [11] "Vision based Hand Gesture Recognition for Indian Sign Language Using Convolution Neural Network." *International Journal of Computer Engineering in Research Trends*. Reported ~93.5% test accuracy.
- [12] "Sign Language Detection with CNN." (2022). *IJRASET*. Provided real-time fingerspelling recognition for American Sign Language with 95.7% accuracy.
- [13] Wen, F., et al. (2023). A wearable system for sign language recognition enabled by a convolutional neural network. *Nano Energy*, 116, 108767.
- [14] Bhagat, N. K., Vishnu sai, Y., & Rathna, G. N. (2019). Indian Sign Language Gesture Recognition using Image Processing and Deep Learning. Used Kinect RGB-D data and CNN for ISL static gestures.
- [15] Early works (2015–2018) – Sensor-based or embedded systems for sign-to-speech/text conversion using Arduino or gloves.
- [16] K. Kumar & A. Choudhary (2015) – Vision-based Indian Sign Language recognition using contour extraction and SVM classification.
- [17] M. Zaffarulla et al. (2015) – Kinect-based gesture tracking for real-time ASL recognition.
- [18] H. Cooper et al. (2014) – Deep-learning based signer-independent sign language recognition using convolutional networks.
- [19] N. Sharma & V. Mittal (2014) – Real-time gesture recognition using skin-color detection and background subtraction.
- [20] J. K. Lee et al. (2013) – Camera-based ASL recognition using motion trajectory analysis.
- [21] R. Roshan & N. K. Das (2013) – Early Indian Sign Language recognition using Haar-like features.
- [22] D. Kelly et al. (2013) – Use of Hidden Markov Models (HMM) for continuous sign recognition.
- [23] L. M. Kamarudin et al. (2013) – Low-cost sensor-glove for converting hand gestures to synthesized speech.