JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A SURVEY ON INDIAN SIGN LANGUAGE TRANSLATING SYSTEM

[1]Mr.Vitraag Shete, [2]Mr.Yogeshwar Prabhu, [3]Ms.Kusuma, [4]Ms.Sameeksha, [5]Prof.Rithika.

[1]Student, [2]Student, [3]Student, [4]Student, [5]Assistant Professor
[1]Computer Science and Business Systems,
[1]Srinivas Institute of Technology, Mangalore, India

Abstract: Indian Sign Language (ISL) plays a vital role in bridging communication gaps between the hearing-impaired community and the hearing population in India. However, widespread technological support for automated ISL translation is limited, which hampers effective communication and inclusion. This paper presents an exhaustive survey of contemporary ISL translating systems, emphasizing both vision-based and sensor-based methodologies. We thoroughly analyze various machine learning techniques, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid architectures, which have been applied for spatial and temporal feature extraction from ISL gestures. This study discusses the challenges faced by current systems such as dataset insufficiency, signer variability, and the complexity of dynamic sign recognition. Moreover, we explore emerging trends and future research directions aimed at developing scalable, accurate, and user-friendly ISL translation solutions. Our work provides a comprehensive foundation for researchers and practitioners interested in advancing ISL recognition technology to enhance social integration and accessibility for the deaf community in India.

***IndexTerms* - Indian Sign Language, Gesture Recognition, Deep Learning, Assistive Technology, Human-Computer Interaction.**

I. INTRODUCTION

Language forms the cornerstone of human interaction, enabling individuals to convey ideas, emotions, knowledge, and experiences. For millions of people worldwide who are deaf or hard of hearing, sign languages serve as their primary mode of communication. In India alone, an estimated 60 million people experience some form of hearing impairment, rendering Indian Sign Language (ISL) an essential linguistic and cultural medium for daily communication.

ISL is a rich and expressive language system characterized by a unique combination of hand shapes, movements, orientations, facial expressions, and body postures. Unlike spoken languages, ISL is highly visual-spatial, relying on intricate hand gestures and facial cues to convey complex meanings, grammatical structures, and emotions. The language is deeply rooted in the cultural and social contexts of its users, which vary regionally across India, making the development of universal translation tools especially challenging.

Despite its significance, ISL has not yet been fully embraced by mainstream society, largely due to a lack of awareness, limited educational resources, and the scarcity of proficient interpreters. These barriers create communication difficulties for ISL users in accessing education, healthcare, employment, and social services, thereby limiting their opportunities and quality of life.

With the rapid advancements in Artificial Intelligence (AI) and computer vision, there is immense potential to develop automated systems capable of translating ISL gestures into textual or spoken language in real time. Such systems can serve as assistive tools, fostering greater inclusion, accessibility, and autonomy for the deaf community. While considerable research has been conducted on American Sign Language (ASL) and other sign languages globally, ISL-specific research remains in its nascent stages.

This paper aims to fill this gap by providing a detailed survey of existing ISL translation systems, categorizing them by their technological approaches, evaluating their methodologies, discussing challenges, and highlighting opportunities for future development. By offering a comprehensive overview, we intend to guide researchers and developers towards building robust, scalable, and user-friendly ISL translation solutions.

Communication is a fundamental aspect of human interaction, yet millions of people around the world who are deaf or hard of hearing face daily challenges in expressing themselves and understanding others. Sign language serves as a vital means of communication for the deaf community. However, the lack of widespread understanding of sign language among the general population often leads to communication barriers, social exclusion, and limited access to essential services.

Indian Sign Language (ISL) is the primary language used by the deaf community in India. Despite its importance, ISL is not widely recognized or understood outside the deaf community. This gap underscores the need for a system that can translate sign language into text or speech in real time, enabling seamless communication between sign language users and non-signers.

The objective of this project is to design and develop an Indian Sign Language Translating System that can recognize hand gestures and convert them into readable or audible output. Utilizing technologies such as computer vision, machine learning, or deep learning, the system aims to identify signs and translate them accurately to bridge the communication gap. This system not only promotes inclusivity but also supports efforts toward digital accessibility and empowerment of the deaf community in India.

II. LITERATURE REVIEW

The domain of Sign Language Recognition (SLR) has grown substantially over the last decade, driven by the necessity to create accessible communication aids for the hearing-impaired population globally. SLR techniques primarily focus on automatically interpreting gestures into text or speech, facilitating communication with non-signers. Broadly, the methodologies can be divided into vision-based, sensor-based, and hybrid systems, each with distinct advantages and limitations.

Vision-Based Systems: Vision-based sign language recognition systems rely on cameras—standard RGB, depth sensors, or multi-camera arrays—to capture visual data representing hand gestures and body movements. These systems analyze video frames to recognize signs without requiring the signer to wear any devices, thus offering a non-intrusive user experience.Typical processing pipelines involve:

Preprocessing: Includes background subtraction, hand segmentation, and noise filtering to isolate the hands and relevant regions from the scene. Techniques like skin color detection and histogram backprojection assist in this step. **Feature Extraction:** Deep learning models, particularly Convolutional Neural Networks (CNNs), are highly effective at extracting spatial features from images, learning representations of hand shapes, contours, and finger positions. These models can learn from large datasets to discriminate between visually similar signs.

Temporal Modeling: Since many ISL gestures involve movement sequences rather than static postures, temporal modeling is crucial. Recurrent Neural Networks (RNNs), including Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU), capture dependencies over time, enabling dynamic gesture recognition.

Classification: The extracted features are classified into specific sign categories using softmax layers or other classifiers.Vision-based systems are gaining popularity due to their flexibility and minimal hardware requirements, but they face challenges related to varying lighting conditions, occlusions, cluttered backgrounds, and signer diversity (such as differences in skin tone, hand size, and signing speed). These factors affect recognition accuracy, necessitating robust preprocessing and model generalization.

Sensor-Based Systems Sensor-based recognition employs wearable devices embedded with motion and flex sensors to capture precise hand and finger movements. Devices such as sensor gloves integrate multiple sensors that measure finger bend angles, acceleration, angular velocity, and hand orientation. The quantitative sensor data enables highly accurate and reliable gesture detection, even under varying environmental conditions. Advantages of sensor-based systems include: **Precision:** Detailed sensor data leads to high accuracy in recognizing both static and dynamic signs **Environmental Robustness:** Less affected by lighting or background changes. **Real-Time Performance:** Sensors provide immediate motion data, reducing processing delays. However, these systems have practical drawbacks: **User Comfort:** Wearing sensor gloves can be cumbersome and may limit natural hand movements. **Cost and Accessibility:** Specialized hardware can be expensive and not readily available to all users. **Calibration:** Regular calibration is required to ensure data accuracy. **Portability:** Bulkiness restricts everyday use. Because of these factors, sensor-based systems are often confined to controlled environments such as research labs or clinical settings. **Hybrid Systems** To combine the complementary strengths of vision-based and sensor-based methods, hybrid systems integrate both data streams, leveraging sensor fusion algorithms to improve gesture recognition accuracy and robustness. For example: Visual data provides spatial context and global hand posture.nSensor data supplies detailed motion and finger articulation information.

Hybrid systems typically employ data fusion techniques at the feature or decision level, utilizing ensemble learning or multimodal deep learning frameworks. These approaches show improved recognition rates, especially for complex and dynamic signs, but require sophisticated synchronization mechanisms and increased computational resources. Their complexity can limit portability and real-time responsiveness.

Notable Indian Sign Language Translating Systems In recent years, several initiatives and research projects have focused on developing ISL translation systems: **DeepSign ISL:** This system utilizes CNN architectures trained on ISL alphabet datasets, achieving high accuracy in recognizing isolated static signs. The model incorporates data augmentation techniques to improve generalization. **ISLRTC Resources:** The Indian government's Indian Sign Language Research and Training Centre (ISLRTC) has developed extensive video dictionaries and learning modules, promoting ISL education and awareness. **Mobile Real-Time Recognition:** Leveraging frameworks such as Google's MediaPipe for efficient hand tracking combined with lightweight neural networks, researchers have developed mobile-friendly applications capable of real-time ISL alphabet recognition. **Sign2Speech Systems:** These models translate ISL sequences into natural language text and then generate synthesized speech, providing an interface for deaf individuals to communicate seamlessly with non-signers. Despite these promising developments, real-time translation of continuous ISL sentences, involving grammatical nuances and non-manual cues, remains an ongoing challenge.

III. RESEARCH METHODOLOGY

The research methodology underlying ISL translation systems involves systematic approaches to data acquisition, model development, and evaluation, tailored to the complexities of sign language gestures. **Categorization of ISL Translation Systems** This survey categorizes ISL translation systems into: **Vision-Based Approaches:** Systems that rely purely on image or video data captured through cameras. They emphasize non-intrusiveness and ease of deployment but require robust visual processing to overcome environmental noise. **Sensor-Based Approaches:** Systems utilizing wearable hardware to collect precise motion and articulation data. They prioritize accuracy but suffer from usability constraints due to the need for specialized devices. **Hybrid Approaches:** Systems combining both modalities to leverage spatial and temporal information comprehensively. **Data Acquisition and Dataset Utilization** One of the foremost challenges in ISL translation is the limited availability of comprehensive, annotated datasets that capture the diversity of ISL signs and signer variations.

Common datasets used include: **ISL Alphabet Dataset:** Contains thousands of images representing static ISL alphabets captured under varying conditions. It provides a basis for training static sign recognition models. **ISL-SLR Dataset:** A video dataset capturing dynamic ISL signs and gestures, facilitating research into temporal recognition. **Custom Datasets:** Many researchers create bespoke datasets by recording native ISL signers performing controlled sign sequences or sentences. These datasets require meticulous annotation, often involving linguistic experts. Data augmentation techniques such as rotation, scaling, and background variation are widely used to artificially increase dataset size and improve model robustness.

Theoretical Framework ISL translation typically follows a multi-stage pipeline: **Input Acquisition:** Capturing raw ISL gestures using cameras or sensors. **Preprocessing:** Enhancing the raw data by filtering noise, segmenting hands, and normalizing input to standard formats. **Feature Extraction:** Applying deep learning models to extract spatial (shape, texture) and temporal (movement) features. **Classification:** Assigning recognized signs to predefined ISL categories using supervised learning algorithms. **Output Generation:** Translating the recognized signs into readable text or synthesized speech, potentially applying natural language processing to maintain semantic coherence. Advanced frameworks also incorporate signer adaptation and transfer learning to accommodate individual variability.

Tools, Algorithms, and Technologies Modern ISL translation systems leverage state-of-the-art tools and methodologies: **MediaPipe Hands:** Google's real-time hand tracking solution efficiently detects and tracks hand landmarks with low latency. **Deep Learning Frameworks:** TensorFlow and PyTorch are predominantly used for developing CNNs, RNNs, and hybrid networks tailored for ISL gesture recognition. **Transfer Learning:** Pretrained models such as MobileNet, ResNet, or EfficientNet are fine-tuned on ISL datasets to accelerate convergence and improve performance with limited data. **Sensor Data Processing:** Algorithms analyze multi-dimensional sensor data streams, employing techniques such as Kalman filtering, feature normalization, and principal component analysis (PCA).

The development of an Indian Sign Language (ISL) Translating System involves a multidisciplinary approach, combining aspects of computer vision, machine learning, and linguistic analysis. The following research methodology was adopted to ensure systematic design and implementation of the system:

Problem Definition The first step was to identify the communication gap between the hearing population and the Indian deaf community. The primary goal was to build a system capable of recognizing ISL gestures and translating them into readable or audible formats.

Data Collection A dataset of Indian Sign Language gestures was collected, either from publicly available sources or created manually through video/image capture of volunteers performing ISL signs. Each gesture corresponds to an alphabet, word, or phrase in ISL. Multiple samples per gesture were gathered to account for variations in hand size, speed, and background.

Data Preprocessing Videos or images were processed to extract frames and isolate hand gestures. Techniques such as grayscale conversion, background subtraction, resizing, and normalization were applied. Key points (like fingertips and joints) were extracted using tools like MediaPipe or OpenCV.

Feature Extraction Visual features were extracted using: Computer Vision techniques (e.g., contour detection, edge detection). Deep Learning models such as Convolutional Neural Networks (CNNs) for automatic feature learning.

Model Training and Classification Various classification models were evaluated, including: Support Vector Machines (SVM) K-Nearest Neighbors (KNN) CNN, RNN, and LSTM models for static and dynamic gesture recognition The model was trained on a labeled dataset and validated using cross-validation methods.

System Design and Integration A modular architecture was implemented: Input Module (camera interface) Processing Unit (gesture recognition engine) Output Module (display of translated text and/or text-to-speech) Testing and Evaluation The system was tested under varying conditions (lighting, background, hand orientation). Performance metrics such as accuracy, precision, recall, and F1-score were used to evaluate the model. Usability testing was conducted with real users to evaluate effectiveness and interface quality.

Iterative Improvements Feedback was incorporated to improve gesture recognition accuracy. Noise handling, gesture segmentation, and model tuning were carried out to optimize performance

IV. RESULTS AND DISCUSSION

A comprehensive evaluation of ISL translation systems reveals insights into their strengths, limitations, and areas for improvement.

Performance of Vision-Based Systems Vision-based systems show promising accuracy, especially for static alphabet recognition, often achieving above 90% accuracy in controlled environments. Incorporation of temporal models like LSTM enables dynamic gesture recognition, though results vary depending on dataset size and signer diversity. However, such systems remain sensitive to lighting changes, occlusions, and complex backgrounds. Variations in individual signing styles, hand shapes, and speeds further affect robustness.

Sensor-Based System Efficacy Sensor-based systems achieve superior accuracy in gesture recognition due to precise motion data but face practical usability issues. The need for hardware limits mass adoption, and sensor calibration is required regularly. **Advantages of Hybrid Approaches** Hybrid systems improve robustness by fusing visual and sensor data, outperforming individual modalities under challenging conditions. However, they require sophisticated hardware setups and increase computational load, posing challenges for real-time mobile deployment.

Key Challenges Identified Dataset Limitations: Insufficient large-scale, annotated ISL datasets limit model training and generalization. Signer Variability: Differences in signing speed, style, and physical characteristics create recognition inconsistencies. Non-Manual Features: Facial expressions and body language, essential for ISL grammar and meaning, are often neglected. Real-Time Processing: Balancing accuracy with low-latency performance remains challenging. Standardization: Lack of benchmark datasets and uniform evaluation protocols hampers comparison across systems.

V. CONCLUSION AND FUTURE WORK

This survey comprehensively reviewed existing Indian Sign Language translation systems, highlighting advances in vision-based, sensor-based, and hybrid approaches. While significant strides have been made in static sign recognition, dynamic and continuous ISL translation continues to pose research challenges.

Future research directions include: **Building Large, Diverse Datasets:** Creating standardized, annotated datasets covering a broad range of signs, signer demographics, and contexts. **Developing Lightweight Models:** Designing efficient, mobile-friendly deep learning architectures to enable real-time translation on portable devices. **Incorporating Multimodal Cues:** Integrating facial expressions and body posture recognition to capture ISL's grammatical richness. **Community Involvement:** Collaborating closely with deaf communities for system design, evaluation, and usability studies. **Standardization and Benchmarking:** Establishing open benchmarks and evaluation metrics for consistent comparison.

By focusing on these areas, the development of inclusive, practical, and robust ISL translation systems can accelerate, promoting social inclusion and accessibility for millions of hearing-impaired individuals across India.

The Indian Sign Language Translating System developed in this project marks a significant step toward bridging the communication gap between the deaf community and the hearing population. By recognizing and translating hand gestures into text or speech, the system enhances accessibility and promotes inclusivity. The project demonstrates that machine learning and computer vision techniques can be effectively utilized to interpret static and dynamic gestures of Indian Sign Language with reasonable accuracy. While the system performs well under controlled conditions, it also highlights challenges such as variability in lighting, background noise, gesture speed, and differences in individual signing styles. Despite these challenges, the prototype validates the potential of automated sign language translation for real-world applications.

. Real-Time Continuous Gesture Recognition: Expanding the system to handle continuous signing (rather than isolated gestures) for fluid conversation. Incorporation of Facial Expressions and Body Posture: Including non-manual signals (like facial expressions) which are essential components of Indian Sign Language grammar. Mobile and Web Integration: Developing lightweight versions of the system for deployment on smartphones and web platforms for greater accessibility. Multi-user and Diverse Dataset Training: Expanding the training dataset to include variations across different users, regions, and sign dialects to improve generalization. Speech-to-Sign Translation: Adding the reverse functionality where spoken language can be translated into animated sign language, enabling two-way communication. Support for Regional Sign Variations: Incorporating regional dialects and signs used across different parts of India to make the system more comprehensive.

While the current system demonstrates the feasibility of translating Indian Sign Language (ISL) into text or speech, there are several areas that warrant further exploration and improvement. One major direction is the transition from isolated gesture recognition to continuous sign language translation, which would allow the system to understand and interpret complete sentences and real-time conversations. Integrating multimodal data such as facial expressions, head movements, and body posture will also enhance accuracy, as these non-manual signals are essential components of ISL grammar.Another important area is the development of bi-directional systems that not only translate sign language into text or speech but also convert spoken or written language back into sign language using 3D animated avatars or augmented reality. Optimizing the system for mobile and wearable devices would improve its usability in daily life, allowing users to communicate on the go. Additionally, building a larger and more diverse dataset that includes regional sign variations and real-world signing scenarios will enhance the system's generalization and robustness. Incorporating natural language understanding (NLU) can help the system grasp the context and semantics of ISL more accurately, as ISL structure differs from spoken Indian languages. Personalization is another promising direction, where the system adapts to the user's individual signing style for improved accuracy and user experience. Moreover, improving the user interface for accessibility, including features like voice commands, subtitles, and pictorial cues, will make the system more inclusive. Finally, integrating the system into public services such as healthcare, education, and government facilities can significantly contribute to social inclusion and accessibility for the deaf community. Future developments must also consider ethical implications, ensuring that user privacy, data security, and consent are maintained throughout system usage.

VI. ACKNOWLEDGMENT

We sincerely thank the Indian Sign Language Research and Training Centre (ISLRTC) for their continuous support and valuable resources. We also appreciate the efforts of the ISL community and researchers whose contributions have been instrumental in advancing ISL technology. Our gratitude extends to Assistant Professor Ms. Rithika for her expert guidance and encouragement throughout this project, and to Mr. Shailesh Shetty, Head of the Department, for providing academic leadership and support.

I would like to express my sincere gratitude to all those who have supported and guided me throughout the development of this project on the Indian Sign Language Translating System. First and foremost, I extend my heartfelt thanks to my mentor/guide, Ms.Rithika , for their invaluable guidance, encouragement, and continuous support throughout the project. Their insights and expertise were instrumental in shaping the project. I would also like to thank Mr.Shailesh Shetty, for providing me with the necessary resources and environment to carry out this work successfully.

A special thanks to the members of the deaf and hard-of-hearing community whose insights and feedback helped me better understand the practical challenges and requirements of a real-time sign language translator. I am also grateful to my family and friends for their constant encouragement and support during the entire duration of this project. Finally, I acknowledge the efforts of all researchers and developers whose prior work in the field of sign language recognition laid the foundation for this project.

REFERENCES

[1] Tao, T., et al. (2024). Sign Language Recognition: A Comprehensive Review of Traditional and Deep Learning Approaches, Datasets, and Challenges. IEEE Access.

[2] Shanableh, T. (2024). Motion Image with Positioning Information and Deep Learning for Continuous Arabic Sign Language Recognition in Signer Dependent and Independent Modes. IEEE Access.

[3] Chamnongthai, K. (2023). IDF-Sign: Addressing Inconsistent Depth Features for Dynamic Sign Word Recognition. IEEE Access.

[4] Elakkiya, R., & Subramaniya Swamy, V. (2022). Development of an End-to-End Deep Learning Framework for Sign Language Recognition, Translation, and Video Generation. IEEE Access.

[5] Shin, J. (2024). Hand Gesture Recognition for Multi-Culture Sign Language Using Graph and General Deep Learning Network. IEEE Access.

[6] Al-Qurishi, M. (2021). Deep Learning for Sign Language Recognition: Current Techniques, Benchmarks, and Open Issues. IEEE Access.

[7] Al Abdullah, B. A. (2024). Advancements in Sign Language Recognition: A Comprehensive Review and Future Prospects. IEEE Access.

[8] Choi, W. (2025). Deep Learning Approaches for Continuous Sign Language Recognition: A Comprehensive Review. IEEE Access.

[9] Geetha, M. (2022). Optimized Multi-Modal Conformer-Based Framework for Continuous Sign Language Recognition. IEEE Access.

[10] Aurelrius, E. (2025). Transfer Learning and Custom Loss Applied to Transformer-Based Text Translation for Sign Language Animated Subtitles. IEEE Access.

[11] Abdul Kadir, M. (2024). MediSign: An Attention-Based CNN-BiLSTM Approach of Classifying Word-Level Signs for Patient-Doctor Interaction in Hearing Impaired Community. IEEE Access.

[12] Abid, A. (2021). A Crowdsourcing-Based Framework for the Development and Validation of Machine-Readable Parallel Corpus for Sign Languages. IEEE Access.

[13] Muhammad, G. (2020). Deep Learning-Based Approach for Sign Language Gesture Recognition with Efficient Hand Gesture Representation. IEEE Access.

[14] Zdravevski, E. (2022). Technological Solutions for Sign Language Recognition: A Scoping Review of Research Trends, Challenges, and Opportunities. IEEE Access