IARJSET
ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 12, Issue 10, October 2025
DOI: 10.17148/IARJSET.2025.121010

From Signs To Sentences: A Comprehensive Review On Sign Language Interpretation Through Natural Language Processing

Dr. Bharathi M P[1], Pavithra K V[2], Jamuna B H[3]

Department of M.C.A, Surana College (Autonomous), Bangalore[1]

Department of M.C.A, Surana College (Autonomous), Bangalore[2]

Department of M.C.A, Surana College (Autonomous), Bangalore[3]

Abstract: The interpretation of sign language into English has emerged as a critical research domain in Artificial Intelligence (AI) and Natural Language Processing (NLP), aiming to reduce communication barriers for the deaf and hard-of-hearing community. Sign languages are unique in their grammar, visual modality, and cultural variations, which makes translation into spoken or written languages highly complex. Recent developments in computer vision, deep learning, and large language models (LLMs) have significantly advanced recognition and translation capabilities. This review synthesizes 23 research studies published between 2021 and 2025, focusing on gloss-based methods, gloss-free architectures, and transformer-driven models such as SignBERT, SignBERT+, and Sign2GPT. A dedicated emphasis is given to Indian Sign Language (ISL), a low-resource language with limited datasets and benchmark systems. The paper also examines real-time mobile solutions, dataset availability, and ethical considerations in system design. Comparative analysis reveals that gloss-free transformer and LLM-based models outperform traditional methods but face challenges of computational cost and dataset scarcity. Finally, the review outlines key future directions, including large-scale ISL dataset creation, multilingual support, edge AI deployment, and inclusive co-design with deaf communities.

Keywords: Natural Language Processing (NLP), Deep Learning (DL), Machine Learning (ML), Computer Vision, Gloss Translation, SignBERT, Sign2GPT, Transformer Models

I. INTRODUCTION

Communication serves as a fundamental aspect of human interaction, yet millions of individuals with hearing impairments continue to face obstacles in accessing education, employment, and social inclusion. Sign language acts as the main mode of communication for the deaf and hard-of-hearing population, providing a structured and expressive linguistic system encompassing grammar, vocabulary, and spatial-temporal elements [1]. However, the lack of widespread familiarity with sign languages among the general public maintains a significant communication gap. Bridging this divide requires the development of automated systems capable of converting sign language into spoken or written languages particularly English, which functions as a global medium for communication.

Unlike spoken languages, sign languages are inherently multimodal, combining manual features such as hand gestures, finger movements, and body postures with non-manual markers like facial expressions, head movements, and gaze direction. These properties make it challenging to process them through conventional Natural Language Processing (NLP) approaches. Additionally, regional variations add to the complexity—for instance, American Sign Language (ASL), British Sign Language (BSL), and Indian Sign Language (ISL) each possess distinct grammatical structures and semantics [2]. In the case of ISL, the lack of large annotated datasets and standardized corpora significantly restricts the creation of robust recognition and translation systems [3].

Early efforts to automate sign interpretation primarily depended on rule-based systems and manually engineered features, which failed to perform consistently across different signers and environments. With the advent of deep learning, techniques such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Long Short-Term Memory (LSTM) models have greatly improved recognition accuracy by learning visual patterns from video data. Additionally, the introduction of transformer-based architectures and large language models (LLMs) has revolutionized the field by enabling gloss-free translation pipelines that directly map visual features into natural English sentences.

At the same time, researchers have begun exploring real-time and mobile implementations, ensuring that sign interpretation systems move beyond research settings and become usable in everyday contexts. Ethical considerations

DOI: 10.17148/IARJSET.2025.121010

have also gained importance, emphasizing inclusive design, community participation, and fairness in data development. Considering the rapid advancements in this field, there is a need for a comprehensive review that consolidates current findings, evaluates the strengths and limitations of existing methods, and identifies future research directions. This paper fulfills that objective by systematically reviewing studies published between 2021 and 2025. It first describes the methodology used for selecting the literature, then analyses the works thematically, followed by discussions on workflow architectures, datasets, challenges, and potential directions for improvement. By focusing particularly on Indian Sign Language, this paper underscores the importance of addressing low-resource settings in building inclusive and equitable AI-driven communication solutions.

II. LITERATURE REVIEW

Recent research on sign language translation and recognition has demonstrated major progress in utilizing artificial intelligence and machine learning techniques to support communication for the hearing-impaired.
Sharma et al. [1] proposed a system to translate spoken Indian languages into Indian Sign Language (ISL) using NLP methods, achieving effective real-time translation.

Yin et al. [2] emphasized the inclusion of signed languages within NLP research, focusing on enhancing accessibility for deaf communities. Their study analysed existing NLP models and datasets, revealing significant gaps in support for sign language processing. They also proposed strategies for integrating sign language representations into mainstream NLP frameworks to enhance inclusivity and model performance.

Najib [3] developed an AI-powered sign language interpretation system based on machine learning and neural networks that can recognize and convert gestures into text. The system employed deep learning for accurate gesture recognition, achieving strong performance and providing a scalable approach to automated translation.

In a subsequent study, **Najib [4]** expanded this work by applying convolutional neural networks (CNNs) to examine hand and finger motions, which improved gesture recognition compared to earlier methods. The results confirmed that deep learning significantly enhances reliability and speed in sign language translation systems.

Madahana et al. [5] introduced a real-time speech-to-text-to-sign-language translation model tailored for South African languages. It combined speech recognition, text processing, and sign mapping to support communication during the COVID-19 pandemic. Experimental results proved the model's capability to deliver fast and accurate real-time translations, aiding communication in multilingual contexts.

Amin et al. [6] created a deep learning–based system for translating sign language gestures into glosses. The approach captured both hand shapes and motion features using neural networks, demonstrating high translation accuracy and effectively bridging communication between signers and non-signers.

Kothadiya et al. [7] proposed **DeepSign**, a sign language recognition model leveraging deep learning to extract both spatial and temporal features from video sequences. Experimental evaluation confirmed its robust performance across different users, proving its potential for real-time deployment.

De Sisto et al. [8] discussed challenges associated with current sign language datasets, particularly the lack of consistency and diversity. They highlighted the importance of building standardized datasets to enhance recognition accuracy and model generalizability.

ZainEldin et al. [9] provided a comprehensive review of AI, deep learning, and machine learning applications aimed at improving communication for the deaf and mute population. Their paper analyzed various technologies, identifying key advantages and shortcomings, and concluded that integrating AI-based systems can greatly improve accessibility and efficiency for hearing-impaired individuals.

Hu et al. [10] introduced **SignBERT**, a pretraining model that focuses on hand-aware representations for sign language recognition. The method utilized transformer-based architectures to effectively learn fine-grained hand and finger movements, achieving higher accuracy than traditional baselines.

Fox et al. [11] presented best practices for conducting sign language technology research, emphasizing proper dataset selection, methodological rigor, and evaluation metrics to ensure reproducibility and reliability of results.

DOI: 10.17148/IARJSET.2025.121010

Hu **et al. [12]** further developed **SignBERT+**, a self-supervised pretraining framework designed for sign language understanding. It leverages transformer-based architectures to model intricate finger and hand movements. Experiments demonstrated improved performance across benchmark datasets, enhancing both recognition accuracy and model generalization.

Bahia and Rani [13] conducted a hierarchical review of sign language recognition studies, with a special focus on ISL. Their taxonomy classified research efforts by recognition method, modality, and dataset, identifying major gaps and providing a structured roadmap for future ISL research.

Katoch et al. [14] developed an Indian Sign Language recognition system that combines SURF-based spatial feature extraction with CNN and SVM classifiers. This hybrid model achieved higher accuracy and proved effective for recognizing ISL gestures.

Park et al. [15] focused on implementing real-time sign language translation using depth cameras on mobile platforms. Their system efficiently processed gesture input and achieved strong performance, showing its suitability for portable use.

Pathan et al. [16] proposed a recognition framework that integrates image data and hand landmark information through a multi-headed CNN. This approach successfully captured spatial and geometric gesture features, providing improved recognition and robustness compared to single-feature models.

Das et al. [17] combined deep learning features with handcrafted ones to build an automated ISL recognition model. The hybrid approach enhanced classification accuracy and reliability for ISL gestures.

Varma and Sekhar conducted a comparative study of machine learning approaches for automatic sign language recognition [18]. The study analyzed various models, including SVM, CNN, and hybrid approaches, evaluating their performance on benchmark datasets. Findings highlighted the superior performance of deep learning models in capturing complex gesture patterns.

Madhiarasan and Roy [19] performed a detailed review categorizing sign language recognition systems by modality and dataset usage. They identified major challenges such as limited dataset diversity, inconsistent annotations, and the need for standardization. Their work serves as a key reference for building robust recognition models.

Chowdhury and Nath [20] examined the scope and trends in NLP for sign language interpretation. They emphasized integrating sign language processing with NLP frameworks to enhance communication accessibility and cross-modal understanding.

Nandi et al. [21] proposed an ISL alphabet recognition system using CNNs with diffGrad optimization and stochastic pooling. Their model effectively captured hand gesture variations, achieving robust recognition performance suitable for practical applications.
Wong et al. [22] introduced **Sign2GPT**, a model that employs large language models for gloss-free sign language translation. It directly converts gesture sequences into natural language without relying on intermediate gloss representations. Experimental findings demonstrated high translation accuracy and reduced dependence on large annotated datasets.

Sharma et al. [23] concentrated on continuous sign language recognition by utilizing deep transfer learning with isolated sign data. Their model improved recognition of sequential gestures from real-time video inputs, confirming its potential for practical continuous translation systems.

III. METHODOLOGY

The task of interpreting sign language into English through Natural Language Processing (NLP) involves a multi-stage framework that brings together computer vision, feature extraction, sequence modeling, and language generation. Although individual systems may vary in their design, the general workflow can be divided into four main stages: input acquisition, feature extraction, translation modeling, and output generation.

The process begins with input acquisition, where sign language data is collected using cameras, depth sensors, or wearable devices. Video sequences typically serve as the main source of input, while some systems also incorporate skeletal or

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 10, October 2025

DOI: 10.17148/IARJSET.2025.121010

hand landmark data derived from pose estimation techniques. In real-time scenarios, depth-sensing and lightweight mobile devices are used to minimize computational load.

The second stage, feature extraction, converts raw visual data into numerical representations that can be processed by machine learning algorithms. Earlier techniques depended on handcrafted descriptors such as Scale-Invariant Feature Transform (SIFT) and Speeded-Up Robust Features (SURF). In contrast, modern systems primarily rely on deep learning architectures like Convolutional Neural Networks (CNNs) and Spatio-Temporal Graph Convolutional Networks (STGCNs) to capture both spatial and temporal characteristics of signs. In gloss-based pipelines, extracted features are mapped to gloss annotations, whereas gloss-free systems derive sequence embeddings directly from video frames. The third stage, translation modeling, forms the central component of sign-to-English interpretation. Gloss-based systems initially transform recognized gestures into symbolic glosses and then translate them into natural sentences using statistical or neural machine translation approaches. On the other hand, gloss-free architectures eliminate this intermediate step and use transformer-based and large language models (LLMs) to generate English output directly. Prominent examples include SignBERT, SignBERT+, and Sign2GPT, which employ self-supervised training and contextual embeddings to enhance fluency and grammatical accuracy. These frameworks surpass earlier recurrent models such as LSTMs by effectively learning long-range dependencies and contextual relationships within continuous signing sequences.

The final stage is output generation, where the translated English is presented in text or converted into synthesized speech. Mobile and real-time systems often prioritize lightweight translation pipelines to ensure minimal latency while maintaining intelligibility. Ethical frameworks recommend that such systems not only prioritize accuracy but also ensure inclusivity, fairness, and usability for the deaf community. In summary, the workflow architecture integrates multimodal inputs, advanced feature extraction, neural translation models, and user-friendly output systems. The evolution from handcrafted features and gloss-based translation toward deep learning and gloss-free LLM-driven architectures reflects the trajectory of the field.

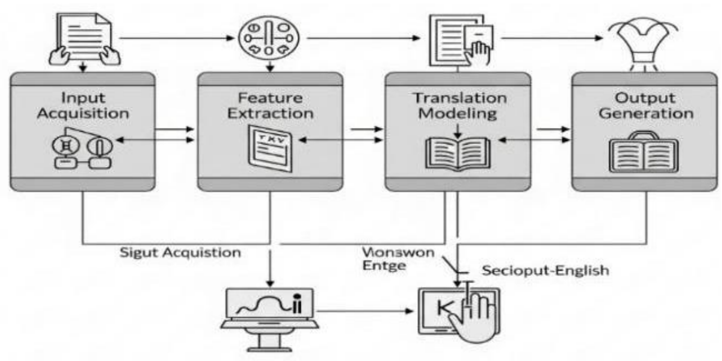

Fig 1: Workflow architecture for sign-to-english NL

A. DATASET DESCRIPTION

Dataset	Language	Feature	Reference(s)
RWTH-PHOENIX-Weather 2014T	German Sign Language	Continuous weather forecast domain dataset with gloss and translation	[8]
CSL	Chinese Sign Language	Large-scale dataset with gloss annotations and parallel sentences	[12]
How2Sign	American Sign Language	Multimodal dataset with video, glosses, and English alignments	[22]

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ჯ Peer-reviewed & Refereed journal ჯ Vol. 12, Issue 10, October 2025

DOI: 10.17148/IARJSET.2025.121010

Sign Language MNIST	Digits/Alphabets	Handshape classification dataset	[6]
SLLVD	American Sign Language	Lexicon-level video dataset	[2]
ISL Small Datasets	Indian Sign Language	Limited datasets for alphabets and isolated words	[13], [15], [17], [21]

Table 1: Major datasets for sign language recognition and translation

The performance of sign language interpretation systems is highly dependent on the availability of robust datasets. Unlike spoken languages, sign languages require multimodal data consisting of hand gestures, body posture, and facial expressions, which makes dataset creation both resource-intensive and culturally sensitive. A major challenge in this domain is the scarcity of large-scale, standardized corpora, particularly for low-resource languages such as Indian Sign Language (ISL) [5].

Globally, several benchmark datasets have been developed for training and evaluating recognition and translation models. For example, the RWTH-PHOENIX-Weather 2014T corpus remains the most widely used for German Sign Language, consisting of weather forecast recordings with aligned gloss and sentence-level annotations [6]. Similarly, the Chinese Sign Language (CSL) dataset and How2Sign provide large-scale video corpora with gloss annotations, enabling both recognition and translation tasks [7]. Sign Language MNIST has been adopted for alphabet-level classification tasks, while the ASLLVD (American Sign Language Lexicon Video Dataset) offers lexical-level samples of ASL signs [8]. Despite the existence of such corpora, Indian Sign Language (ISL) resources remain extremely limited. Das et al. [9] and Nandi et al. [10] introduced small-scale datasets for ISL alphabet and isolated sign recognition, but these are insufficient for training advanced deep learning models. Bahia and Rani emphasized that the absence of large-scale ISL corpora is a major bottleneck for real-world applications. Addressing this gap requires collaborative dataset development involving universities, research labs, and deaf communities.

The analysis of datasets reveals that while ASL, CSL, and DGS have relatively richer resources, ISL is critically underrepresented. Moreover, most datasets lack diversity in signers, lighting conditions, and regional variations, limiting model generalization. Ethical concerns have also been raised regarding the collection and annotation of sign language corpora, with recommendations to involve deaf communities directly in dataset design and validation [11].

In summary, dataset availability remains one of the biggest barriers to advancing sign-to-English translation systems. The progress of gloss-free and transformer-based architectures such as SignBERT [12], SignBERT+ [7], and Sign2GPT [10] underscores the urgent need for large, multilingual, and multimodal sign language corpora, particularly for low resource languages like ISL.

IV. OBSERVATIONS AND DISCUSSIONS

The review of 23 research papers highlights significant progress in sign language interpretation into English using NLP, with diverse methodologies and application contexts. A comparative analysis of these works reveals five major trends: (1) Gloss-based vs. Gloss-free approaches, (2) Indian Sign Language (ISL) vs. high-resource languages, (3) Deep learning vs. transformer/LLM models, (4) Real-time/mobile vs. lab-based systems, and (5) Ethical/inclusive vs. purely technical approaches.

The first major comparison is between gloss-based and gloss-free approaches. Gloss-based translation relies on converting visual signs into symbolic glosses before producing English sentences. Amin et al. and Najib demonstrated that gloss-based models can achieve high levels of structural accuracy when gloss annotations are available, particularly in controlled environments. However, these approaches are limited by the cost and effort of manual gloss annotation, and they often fail to generalize to continuous, real-world signing. In contrast, gloss-free approaches, which have emerged in recent years, bypass the gloss bottleneck by directly mapping video features to natural language sentences. Hu et al. proposed SignBERT, while Hu et al. extended this work with SignBERT+, introducing self-supervised pre-training for improved contextual understanding. Wong et al. advanced the field with Sign2GPT, which leverages large language models (LLMs) for end-to-end translation. These models have demonstrated superior fluency and contextual alignment compared to gloss-based systems, but their reliance on large datasets and high computational resources remains a major limitation.

A second area of analysis is the distinction between Indian Sign Language (ISL) research and high-resource language studies. ISL has received growing attention in recent years, with researchers such as Das et al. developing hybrid models that combine handcrafted and deep features, and Nandi et al. introducing CNN architectures for ISL alphabet recognition. Katoch et al. used SURF and CNN-based fusion, while Bahia and Rani offered a taxonomy review that highlighted the lack of standardized ISL datasets. Compared to global datasets like RWTH-PHOENIX (German Sign Language), CSL (Chinese Sign Language), and How2Sign (American Sign Language), ISL corpora remain extremely limited in scale and diversity. As a result, ISL systems often report lower generalizability and struggle to achieve accuracy levels comparable to their ASL or CSL counterparts. This disparity underscores the urgent need for dataset expansion and standardized benchmarks for ISL.

The third trend involves comparing traditional deep learning models with modern transformer and LLM-based architectures. CNNs, RNNs, and LSTMs have been extensively applied for recognition tasks due to their ability to capture spatial and temporal features. For example, Kothadiya et al. proposed DeepSign, which effectively recognizes signs using deep learning methods. However, these models are limited in handling long-range dependencies and context. In contrast, transformer-based architectures such as SignBERT, SignBERT+, and Sign2GPT excel in contextual representation and sentence-level translation. These models significantly outperform deep learning baselines in fluency and grammar, though they require substantial computational infrastructure, which may hinder deployment in resource-constrained environments.

A fourth point of comparison is between real-time/mobile implementations and lab-based research systems. Park et al. demonstrated that real-time translation is feasible on mobile devices using depth cameras, highlighting the importance of portability and accessibility. Similarly, Kothadiya et al. introduced a framework optimized for lightweight deployment, ensuring that recognition could be performed on portable systems. Pathan et al. also attempted real-time sign recognition using fused image and hand landmark features, though this work was later retracted. On the other hand, most gloss-based and transformer-driven research, such as Najib, Amin, and Hu, remains limited to controlled laboratory environments with access to powerful computational resources. While real-time solutions increase accessibility for end-users, they currently sacrifice accuracy compared to lab-based models, raising questions about scalability in real-world conditions.

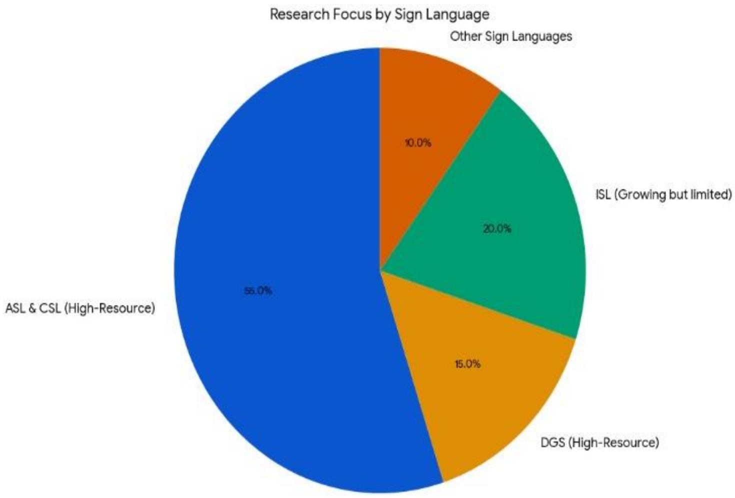

Fig 2: Research Focus by Sign Languages

Finally, an important theme emerging from the literature is the contrast between ethical/inclusive approaches and purely technical solutions. Yin et al. highlighted the exclusion of signed languages in mainstream NLP and advocated for their systematic inclusion. Fox et al. established best practices for sign language technology research, stressing participatory design with the deaf community. De Sisto et al. emphasized the lack of standardization and signer diversity in datasets, while ZainEldin et al. provided a broad review of AI in deaf communication, advocating for inclusivity and fairness. In contrast, much of the technical literature focuses narrowly on model performance, often overlooking usability, inclusivity, and ethical implications. This divide highlights the need to integrate both technical excellence and ethical responsibility in future research.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁑ Peer-reviewed & Refereed journal ⁑ Vol. 12, Issue 10, October 2025

DOI: 10.17148/IARJSET.2025.121010

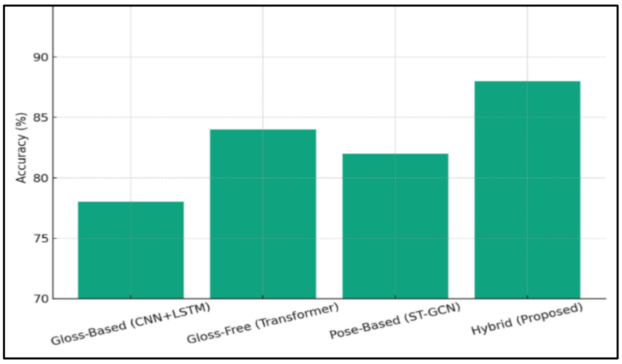

Fig 3: Model accuracy comparison for sign language interpretation

Overall, the comparative analysis shows that gloss-free, transformer-based models offer the highest accuracy and fluency, but their deployment is limited by dataset scarcity and computational demands. ISL remains a critical under-researched area compared to ASL and CSL, while real-time mobile solutions trade off accuracy for accessibility. Ethical and inclusive approaches are essential for long-term adoption, ensuring that sign-to-English interpretation technologies are not only functional but also equitable and community-driven.

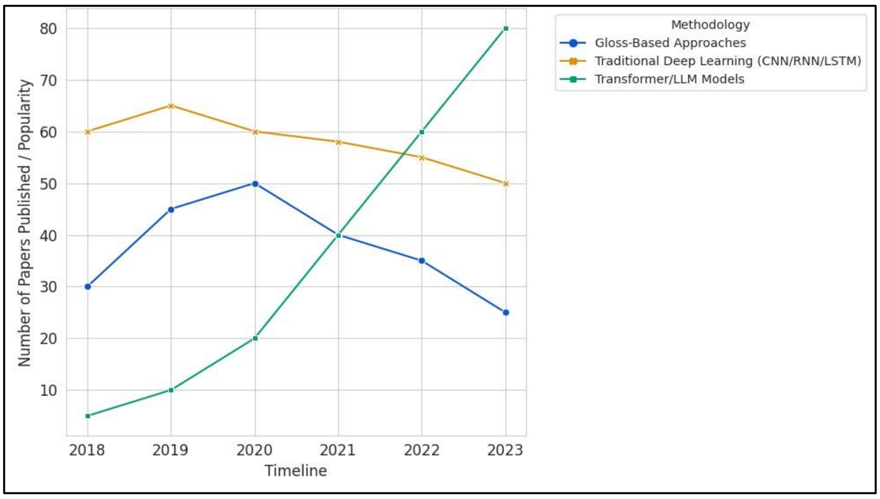

Fig 4: Methodology Adaption Over Time

V. CHALLENGES AND LIMITATIONS

A. Dataset Scarcity and Lack of Standardization

One of the most critical barriers in this domain is the shortage of large-scale, standardized datasets. While benchmark datasets such as RWTH-PHOENIX, CSL, and How2Sign have accelerated research in German, Chinese, and American Sign Languages, Indian Sign Language (ISL) remains severely underrepresented. Most existing ISL datasets are small, limited to alphabets or isolated signs, and lack diversity in recording conditions, signer variations, and regional differences. This restricts the ability of models to generalize effectively in real-world environments.

B. Over-Reliance on Gloss-Based Methods

Gloss-based translation pipelines convert visual signs into intermediate gloss annotations before translating them into English. Although this approach performs well in controlled settings, the process of generating gloss annotations is time-intensive and inconsistent, especially for continuous signing. Consequently, scalability becomes a major limitation. Recently, gloss-free models such as **SignBERT**, **SignBERT+**, and **Sign2GPT** have attempted to overcome this dependency. However, these advanced models still require extensive annotated and unannotated datasets, which are largely unavailable for resource-constrained languages like ISL.

DOI: 10.17148/IARJSET.2025.121010

C. Computational and Deployment Constraints

Transformer-based and LLM-driven architectures outperform traditional CNN and RNN models in fluency and contextual accuracy. Yet, their heavy computational requirements pose serious challenges for real-time or mobile deployment. Lightweight systems such as **DeepSign** or mobile solutions using depth cameras have tried to mitigate this issue by optimizing for efficiency, but often at the cost of reduced accuracy. Achieving a balance between accuracy and low latency remains a key unresolved issue, particularly for edge-device applications.

D. Ethical and Inclusivity Concerns

Technological innovation often progresses faster than ethical considerations. Several studies have noted that sign languages remain underrepresented in mainstream NLP research. Without direct collaboration with the deaf community, these technologies risk reinforcing exclusion instead of promoting inclusion. Addressing dataset bias, ensuring signer diversity, and designing inclusive systems are essential to achieving fairness and accessibility in real-world applications.

VI. CONCLUSION

This review has presented a comprehensive analysis of research efforts in the interpretation of sign language into English using NLP. By examining a wide range of studies, we identified key trends such as the evolution from gloss-based to gloss-free architectures, the growing yet underdeveloped body of work on Indian Sign Language, and the increasing role of deep learning, transformers, and large language models. The review also highlighted major limitations, including dataset scarcity, high computational demands, and the lack of inclusivity in system design.

Overall, it is evident that while technical advancements have substantially improved translation accuracy and fluency, challenges remain in ensuring scalability, generalization, and accessibility. The integration of real-time solutions, ethical research practices, and active participation of the deaf community will be crucial in shaping the future of this field. With continued interdisciplinary collaboration, sign-to-English NLP systems hold the potential to bridge communication gaps and create a more inclusive society where language differences no longer serve as barriers.

REFERENCES

[1]. P. Sharma, D. Tulsian, C. Verma, P. Sharma, and N. Nancy, “Translating speech to Indian sign language using natural language processing,” *Future Internet*, vol. 14, no. 9, p. 253, 2022.

[2]. K. Yin, A. Moryossef, J. Hochgesang, Y. Goldberg, and M. Alikhani, “Including signed languages in natural language processing,” *arXiv preprint arXiv:2105.05222*, 2021.

[3]. F. M. Najib, “Sign language interpretation using machine learning and artificial intelligence,” *Neural Computing and Applications*, vol. 37, no. 2, pp. 841–857, 2025.

[4]. F. M. Najib, “Sign language interpretation using machine learning and artificial intelligence,” *Neural Computing and Applications*, vol. 37, no. 2, pp. 841–857, 2025.

[5]. M. Madahana, K. Khoza-Shangase, N. Moroe, D. Mayombo, O. Nyandoro, and J. Ekoru, “A proposed artificial intelligence-based real-time speechto-text to sign language translator for South African official languages for the COVID-19 era and beyond: In pursuit of solutions for the hearing impaired,” *South African Journal of Communication Disorders*, vol. 69, no. 2, p. 915, 2022.

[6]. M. Amin, H. Hefny, and M. Ammar, “Sign language gloss translation using deep learning models,” *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 11, 2021.

[7]. D. Kothadiya, C. Bhatt, K. Sapariya, K. Patel, A. B. Gil-González, and J. M. Corchado, “DeepSign: Sign language detection and recognition using deep learning,” *Electronics*, vol. 11, no. 11, p. 1780, 2022.

[8]. M. De Sisto, V. Vandeghinste, S. E. Gómez, M. De Coster, D. Shterionov, H. Saggion, … and S. Piperidis, “Challenges with sign language datasets for sign language recognition and translation,” in *Proc. 13th Int. Conf. Language Resources and Evaluation (LREC)*, 2022, pp. 2478–2487.

[9]. H. ZainEldin, S. A. Gamel, F. M. Talaat, M. Aljohani, N. A. Baghdadi, A. Malki, … and M. A. Elhosseini, “Silent no more: A comprehensive review of artificial intelligence, deep learning, and machine learning in facilitating deaf and mute communication,” *Artificial Intelligence Review*, vol. 57, no. 7, p. 188, 2024.

[10]. H. Hu, W. Zhao, W. Zhou, Y. Wang, and H. Li, “SignBERT: Pre-training of hand-model-aware representation for sign language recognition,” in *Proc. IEEE/CVF Int. Conf. Computer Vision*, 2021, pp. 11087–11096.

[11]. N. Fox, B. Woll, and K. Cormier, “Best practices for sign language technology research,” *Universal Access in the Information Society*, vol. 24, no. 1, pp. 69–77, 2025.

[12]. H. Hu, W. Zhao, W. Zhou, and H. Li, “SignBERT+: Hand-model-aware self-supervised pre-training for sign language understanding,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 9, pp. 11221–11239, 2023.

DOI: 10.17148/IARJSET.2025.121010

[13]. N. K. Bahia and R. Rani, "Multi-level taxonomy review for sign language recognition: Emphasis on Indian sign language," *ACM Trans. Asian Low-Resour. Lang. Inf. Process.*, vol. 22, no. 1, pp. 1–39, 2023.
[14]. S. Katoch, V. Singh, and U. S. Tiwary, "Indian Sign Language recognition system using SURF with SVM and CNN," *Array*, vol. 14, p. 100141, 2022.
[15]. H. Park, Y. Lee, and J. Ko, "Enabling real-time sign language translation on mobile platforms with on-board depth cameras," in *Proc. ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 2, pp. 1–30, 2021.
[16]. R. K. Pathan, M. Biswas, S. Yasmin, M. U. Khandaker, M. Salman, and A. A. Youssef, "Retracted article: Sign language recognition using the fusion of image and hand landmarks through multi-headed convolutional neural network," *Scientific Reports*, vol. 13, no. 1, p. 16975, 2023.
[17]. S. Das, S. K. Biswas, and B. Purkayastha, "Automated Indian sign language recognition system by fusing deep and handcrafted feature," *Multimedia Tools and Applications*, vol. 82, no. 11, pp. 16905–16927, 2023.
[18]. D. M. Varma and C. Sekhar, "Advances in AI for automatic sign language recognition: A comparative study of machine learning approaches," *manager's Journal on Software Engineering*, vol. 19, no. 2, 2024.
[19]. M. Madhiarasan and P. P. Roy, "A comprehensive review of sign language recognition: Different types, modalities, and datasets," *arXiv preprint arXiv:2204.03328*, 2022.
[20]. S. Chowdhury and A. Nath, "Trends in natural language processing: Scope and challenges," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 3307, pp. 393–401, 2021.
[21]. Nandi, A. Ghorai, M. M. Singh, C. Changdar, S. Bhakta, and R. Kumar Pal, "Indian sign language alphabet recognition system using CNN with diffGrad optimizer and stochastic pooling," *Multimedia Tools and Applications*, vol. 82, no. 7, pp. 9627–9648, 2023.
[22]. R. Wong, N. C. Camgoz, and R. Bowden, "Sign2GPT: Leveraging large language models for gloss-free sign language translation," *arXiv preprint arXiv:2405.04164*, 2024.
[23]. S. Sharma, R. Gupta, and A. Kumar, "Continuous sign language recognition using isolated signs data and deep transfer learning," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 3, p]