Interdisciplinary Journal of AI, Machine Learning & Data Science (IJAIMLDS)

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

Article - e002

A Comprehensive Review of Machine Learning and Deep Learning Techniques for Sign Language Recognition.

Shivani Gupta[1], **Rakhi Jha**[2], **Dr. Sumit Jain**[3]

[1]Research Scholar, Sage University, Indore, India
[2]Assistant professor Parul University,Vadodara Gujrat, India
[3]Associate professor, Sage University Indore, India

Received: 20/02/2026 **Revision Received**: 12/03/2026 **Accepted**: 22/03/2026

ABSTRACT

Sign language serves as the primary means of communication for the deaf and hard-of-hearing community. However, communication barriers still exist between sign language users and the general population. Recent advances in machine learning (ML) and deep learning (DL) have enabled automated sign language recognition (SLR) systems capable of translating gestures into text or speech. This paper presents a systematic review of ML- and DL-based techniques used for SLR across various languages, including American Sign Language and Arabic Sign Language. The review examines classical machine learning approaches, convolutional neural networks (CNNs), hybrid architectures such as CNN-LSTM and CNN-HMM, and video-based models including 3D CNNs. Comparative analysis indicates that deep learning methods significantly outperform traditional techniques due to their ability to automatically extract spatial and temporal features. However, challenges such as limited datasets, signer variability, computational complexity, and real-time deployment constraints remain significant. This review identifies research gaps and discusses future directions, including multimodal learning, lightweight architectures, and multilingual datasets, to improve the scalability and applicability of SLR systems.

This systematic review analyzes 30 research papers published between 2021 and 2025 using the PRISMA methodology.

KEYWORDS: Sign Language Recognition, Machine Learning, Deep Learning, CNN, LSTM, Gesture Recognition, Computer Vision.

1. INTRODUCTION

Sign language plays a crucial role in enabling communication for deaf and hard-of-hearing individuals worldwide. It relies on a combination of hand gestures, facial expressions, and body movements to convey linguistic meaning. Despite its importance, the majority of the population is not familiar with sign language, which creates communication barriers in daily interactions, education, and public services.

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

Traditionally, human interpreters have been used to bridge this communication gap. However, interpreters are not always available, and their services may be costly. Advances in computer vision, machine learning, and deep learning have enabled the development of automated sign language recognition systems that can translate gestures into text or speech.

Early research in gesture recognition relied on handcrafted features and statistical models such as Hidden Markov Models (HMMs). Although these approaches provided a foundation for SLR systems, they were limited in their ability to handle complex visual patterns and variations in gesture execution. With the rise of deep learning, architectures such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid models have significantly improved recognition accuracy.

Despite these advances, several challenges remain. SLR systems often require large, annotated datasets, substantial computational resources, and robust methods capable of handling real-world variability. Therefore, a comprehensive review of existing approaches is necessary to understand the progress made in this field and identify future research directions.

The main contributions of this review are as follows:

- Provide a structured overview of machine learning and deep learning techniques used in sign language recognition.
- Compare classical methods, deep learning architectures, and hybrid frameworks.
- Identify key challenges and research gaps in existing SLR systems.
- Highlight future research directions for developing scalable and efficient SLR solutions.

2. REVIEW METHODOLOGY

This study follows a systematic literature review approach to analyze recent advancements in sign language recognition (SLR) using machine learning and deep learning techniques.

2.1 Literature Search Strategy

Relevant research papers were collected from major academic databases, including IEEE Xplore, SpringerLink, ScienceDirect, and Google Scholar.

The search queries used include combinations such as: (“sign language recognition” AND “deep learning”), (“gesture recognition” AND “CNN”), and (“sign language” AND “LSTM” OR “transformer”).

2.2 Time Range and Database Coverage

The review covers studies published between **2021 and 2025**, ensuring inclusion of recent developments in SLR research.

2.3 Inclusion Criteria

The following criteria were used to select relevant studies:

- Peer-reviewed journal and conference papers
- Studies focusing on ML/DL-based SLR techniques
- Papers including datasets, models, and experimental evaluation

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

2.4 Study Selection Process

A systematic literature search was conducted using major academic databases including IEEE Xplore, SpringerLink, ScienceDirect, and Google Scholar. The search process followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework to ensure a structured and transparent selection of relevant studies.

In the identification stage, a total of 110 research articles were retrieved based on the defined search queries. After removing duplicate and irrelevant records, 75 articles were retained for further screening.

During the screening stage, titles and abstracts of the selected papers were evaluated to determine their relevance to machine learning and deep learning approaches in sign language recognition. This process resulted in 45 articles being shortlisted for full-text review.

In the eligibility stage, full-text articles were carefully assessed based on inclusion and exclusion criteria, including methodological clarity, dataset usage, and experimental validation.

Finally, in the inclusion stage, 30 research papers were selected for detailed analysis in this systematic review.

The complete study selection process is illustrated in Fig. 1, which presents the PRISMA flow diagram including identification, screening, eligibility, and inclusion phases.

2.5 Quality Assessment Criteria

The selected papers were evaluated based on methodological clarity, dataset usage, model performance, and relevance to sign language recognition.

2.6 Study Selection and Bias Reduction

The initial search identified **110 papers**, which were reduced to **75** after removing duplicates. After title and abstract screening, **45 papers** were selected for full-text review. Finally, **30 papers** were included in this review. To reduce selection bias, multiple databases were used and only peer-reviewed, experimentally validated studies were considered. The selection process is illustrated in **Fig. 1**.

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

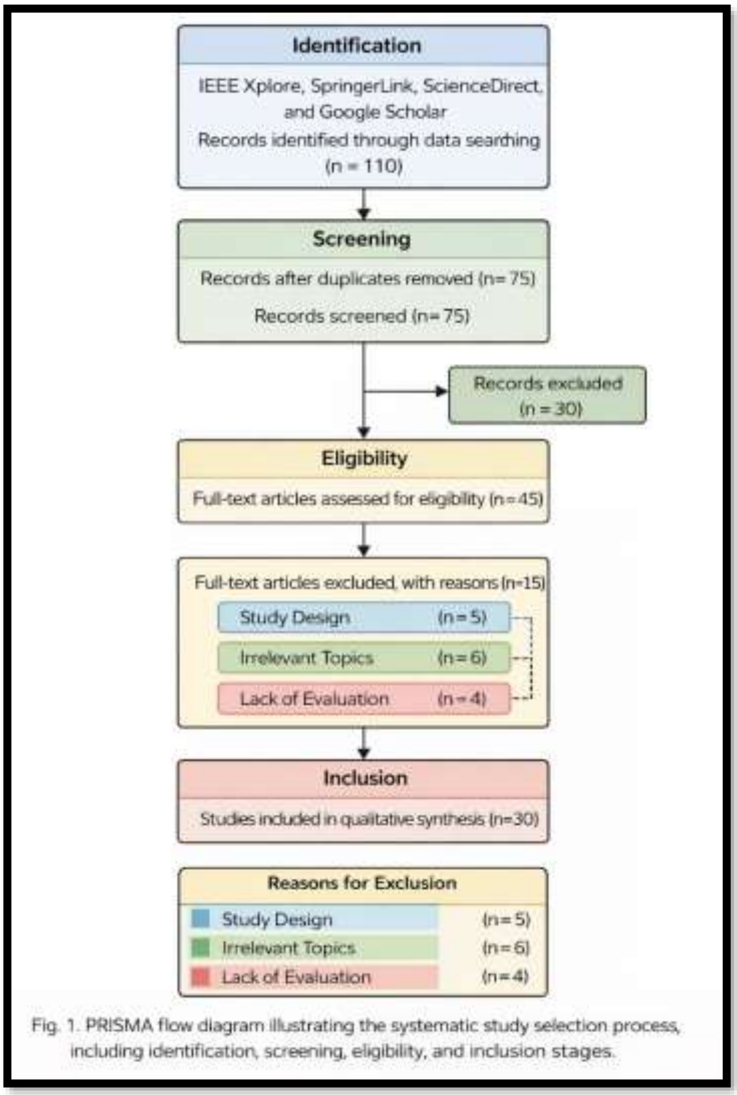

Fig. 1. PRISMA flow diagram illustrating the systematic study selection process, including identification, screening, eligibility, and inclusion stages.

3. LITERATURE REVIEW

Sign language recognition (SLR) has gained significant attention due to advancements in computer vision and deep learning techniques. Early approaches relied on traditional machine learning methods using handcrafted features, which struggled with complex gesture variations.

With the emergence of deep learning, Convolutional Neural Networks (CNNs) have become widely used for extracting spatial features from gesture images, improving recognition accuracy [1], [7], [14]. However, CNN-based models are limited in capturing temporal dependencies required for dynamic gesture recognition.

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

To address this limitation, hybrid models such as CNN–LSTM have been proposed, combining spatial feature extraction with temporal sequence modeling. These models have demonstrated improved performance for continuous sign language recognition tasks [15], [20], [29].

Recent advancements include transformer-based architectures, which use self-attention mechanisms to capture long-range dependencies in gesture sequences. These models have shown promising results on large-scale datasets such as WLASL [16], [19]. Additionally, pose estimation techniques have been used to extract skeletal keypoints, improving efficiency and reducing computational complexity [3], [18].

Despite these advancements, challenges such as limited datasets, high computational cost, and real-world variability still remain. Moreover, there exists a trade-off between accuracy and computational efficiency across different SLR approaches, which limits real-time deployment.

Table 1. Summary of Key Studies in Sign Language Recognition

Author	**Method**	**Dataset**	**Key Contribution**	**Limitation**
Buttar et al. [1]	Hybrid Deep Learning	Sign gesture dataset	Recognition of static and dynamic signs	Limited dataset diversity
Jindal et al. [7]	CNN	Image dataset	Effective spatial feature extraction	Difficulty with dynamic gestures
Zhang et al. [14]	Deep CNN	Gesture dataset	Improved gesture classification accuracy	High computational cost
Kumari et al. [15]	Attention-based LSTM	Video dataset	Temporal modeling of dynamic gestures	Large training time
Paul et al. [20]	CNN–LSTM	Gesture video dataset	Real-time gesture recognition	Requires large training data
Mansour et al. [16]	Vision Transformer	Arabic SL dataset	Captures long-range dependencies	High model complexity
Myagila [18]	Pose estimation + DL	Continuous video dataset	Key point-based gesture detection	Accuracy depends on pose estimation
Jayanthi et al. [29]	CNN + LSTM	Gesture dataset	Sentence prediction from gesture sequences	Limited vocabulary

Interdisciplinary Journal of AI, Machine Learning & Data Science (IJAIMLDS)
ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

4. TAXONOMY OF SIGN LANGUAGE RECOGNITION TECHNIQUES

4.1 Classical Machine Learning Approaches

This subsection discusses early methods used for sign language recognition, including algorithms such as Hidden Markov Models (HMM), Support Vector Machines (SVM), k-Nearest Neighbours (k-NN), and other feature-based classifiers. These approaches typically rely on handcrafted features extracted from images or video frames. While they are computationally efficient, their performance is limited when handling complex gesture variations and dynamic environments.

4.2 Deep Learning-Based Approaches

Deep learning models have significantly improved the performance of sign language recognition systems. Convolutional Neural Networks (CNNs) are widely used for extracting spatial features from hand gestures, while architectures such as Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks capture temporal dependencies in sequential gesture data.

4.3 Hybrid Architectures

Hybrid models combine multiple techniques to leverage the strengths of different algorithms. For example, CNN–LSTM models integrate spatial feature extraction with temporal sequence modeling, while CNN–HMM frameworks combine deep learning representations with statistical sequence modeling methods. These architectures provide improved recognition accuracy for continuous sign language recognition tasks.

4.4 Video-Based and 3D CNN Models

To capture spatiotemporal information from gesture sequences, researchers have proposed 3D convolutional neural networks. These models process multiple video frames simultaneously, enabling the system to understand motion patterns and dynamic gestures more effectively than traditional 2D CNN approaches.

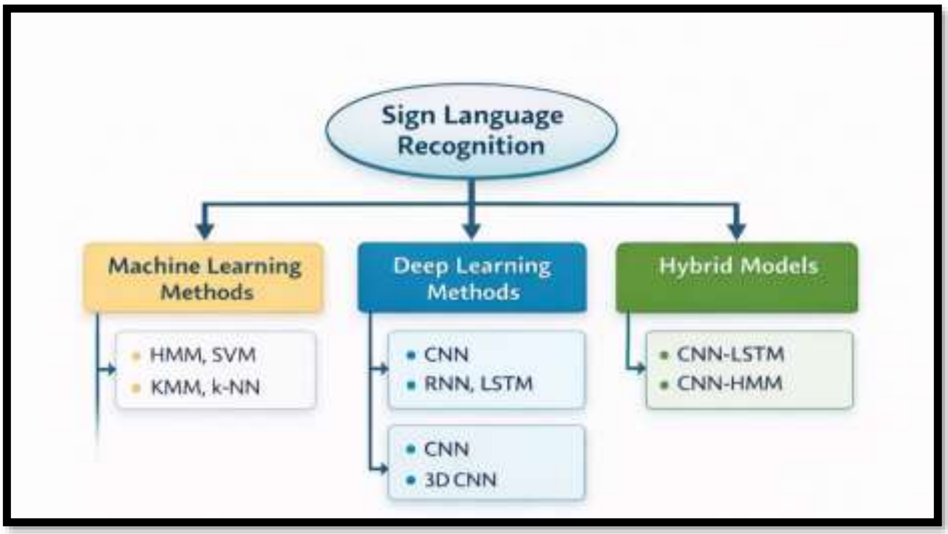

Fig. 2. Taxonomy of sign language recognition techniques.

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

5. COMPARATIVE ANALYSIS OF SIGN LANGUAGE RECOGNITION TECHNIQUES

Sign language recognition techniques have evolved from traditional machine learning approaches to advanced deep learning models. Early methods based on SVM and HMM relied on handcrafted features and were limited in handling complex gestures [3], [9].

Deep learning approaches, particularly CNNs, improved performance by automatically extracting spatial features from gesture images [1], [7], [14]. However, CNN models alone are insufficient for capturing temporal dependencies in dynamic gestures.

Hybrid models such as CNN–LSTM overcome this limitation by combining spatial and temporal learning, making them suitable for continuous sign recognition [15], [20], [29]. However, these models require higher computational resources and longer training time.

Transformer-based models further improve sequence modeling by capturing long-range dependencies using attention mechanisms [16], [19]. Despite their high accuracy, these models are computationally expensive and less suitable for real-time applications.

Overall, there is a trade-off between model accuracy, computational complexity, and real-time deployment capability in SLR systems.

Table 2. Comparison of Sign Language Recognition Techniques

Technique	Advantage	Reference
Machine Learning	**Low computational cost**	**[3], [9]**
CNN	**Strong spatial feature extraction**	**[1], [7], [14]**
CNN–LSTM	**Handles dynamic gestures**	**[15], [20]**
Transformer	**Captures long-range dependencies**	**[16], [19]**

CNN models are most suitable for static gesture recognition tasks where only spatial features are required. CNN–LSTM models perform better for dynamic gesture recognition involving temporal sequences. Transformer-based models are preferred for large-scale datasets requiring long-range dependency modeling, although they are less suitable for real-time applications due to high computational cost.

6. PERFORMANCE EVALUATION OF SLR MODELS

Performance evaluation plays a critical role in assessing the effectiveness of sign language recognition (SLR) systems. Different models are evaluated using various performance metrics such as accuracy, precision, recall, F1-score, and computational efficiency. These metrics help determine how well a model can recognize gestures under different conditions, including variations in datasets, environmental factors, and signer differences.

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

Early SLR systems based on traditional machine learning techniques relied heavily on handcrafted feature extraction methods. While these approaches demonstrated moderate accuracy, their performance was limited when dealing with complex gesture patterns and large-scale datasets. With the emergence of deep learning, models such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid architectures have significantly improved recognition performance.

Deep learning models are capable of automatically extracting hierarchical spatial and temporal features from gesture data, which allows them to achieve higher accuracy compared to classical machine learning approaches. However, the performance of these models also depends on the quality and size of the dataset used for training.

6.1 Evaluation Metrics

Several evaluation metrics are commonly used to assess the performance of SLR systems:

- Accuracy: Measures the proportion of correctly classified gestures among all predictions.
- Precision: Indicates the proportion of correctly predicted positive observations among all predicted positives.
- Recall: Measures the ability of the model to correctly identify all relevant gesture instances.
- F1-score: The harmonic mean of precision and recall, providing a balanced evaluation metric.
- Computational Complexity: Evaluates the processing time and resources required for training and inference.

These metrics provide a comprehensive assessment of both model effectiveness and computational efficiency

6.2 Comparative Performance of SLR Models

Table 3 presents a comparison of different SLR models based on commonly reported performance metrics and datasets used in various studies.

Table 3. Performance Comparison of Sign Language Recognition Models

Model	Dataset Type	Accuracy (Approx.)	Advantages	Limitations
HMM / SVM	Static gesture datasets	80–88%	Simple implementation, low computational cost	Limited feature representation
CNN	Image-based SLR datasets	90–94%	Strong spatial feature extraction	Weak temporal modeling
CNN–LSTM	Video-based gesture datasets	93–96%	Captures temporal dependencies	Higher training complexity
CNN–	Continuous SLR	91–94%	Effective sequence	Complex

Interdisciplinary Journal of AI, Machine Learning & Data Science (IJAIMLDS)
ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

HMM	datasets		modeling	architecture
3D CNN	Large video datasets	95–97%	Spatio-temporal feature learning	High computational requirements

Source: Performance results adapted from studies in [1], [14], [15], [16], [20], [29].

6.3 Discussion

The comparative analysis shows that deep learning-based approaches outperform traditional machine learning techniques in terms of recognition accuracy and robustness. CNN-based models are highly effective for static gesture recognition, while hybrid architectures such as CNN–LSTM and CNN–HMM perform better in dynamic and continuous sign recognition scenarios.

Transformer-based models further improve performance by capturing long-range dependencies in gesture sequences; however, they require large datasets and high computational resources. Therefore, selecting an appropriate model depends on the trade-off between accuracy, computational cost, and real-time deployment requirements.

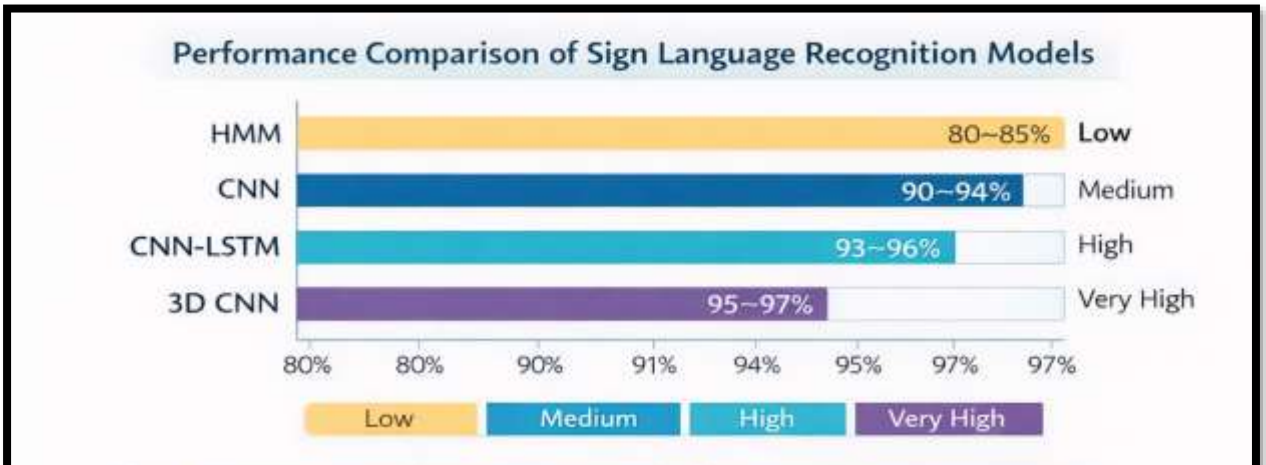

Fig. 3. Comparative performance analysis of different sign language recognition models based on reported accuracy.

These findings highlight the importance of selecting appropriate models based on application requirements, dataset size, and computational constraints.

7. DATASETS USED IN SIGN LANGUAGE RECOGNITION SYSTEMS

Datasets play a crucial role in the development and evaluation of sign language recognition (SLR) systems. The performance of machine learning and deep learning models largely depends on the availability of large, diverse, and well-annotated datasets. These datasets contain images or videos of hand gestures, facial expressions, and body movements that represent different signs used in sign languages.

Most SLR datasets are categorized into **static gesture datasets** and **dynamic video-based datasets**. Static datasets contain individual images representing specific hand gestures, while dynamic datasets consist of video sequences capturing continuous sign language movements.

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

Deep learning models typically require large-scale video datasets to effectively learn spatial and temporal patterns.

Several publicly available datasets have been widely used by researchers for training and evaluating SLR models. These datasets differ in terms of language type, dataset size, gesture complexity, and recording conditions.

7.1 Publicly Available Sign Language Datasets

Table 4. Commonly Used Datasets for Sign Language Recognition

Dataset	Language	Data Type	Number of Classes	Description
ASL Alphabet Dataset	American Sign Language	Image	26	Contains images representing alphabets of ASL
WLASL	American Sign Language	Video	2000+	Large-scale word-level dataset for SLR
RWTH-PHOENIX-Weather	German Sign Language	Video	Continuous	Dataset used for continuous sign language translation
CSL Dataset	Chinese Sign Language	Video	500+	Contains multiple signers performing Chinese signs
ISL Dataset	Indian Sign Language	Image/Video	Varies	Dataset focusing on Indian sign language gestures

7.2 Characteristics of SLR Datasets

SLR datasets vary in several important characteristics, including dataset size, gesture complexity, number of signers, and environmental conditions. Large datasets with multiple signers and varied backgrounds provide better training data for deep learning models. However, many datasets are collected in controlled environments, which limits their applicability to real-world scenarios.

Another important aspect is the inclusion of **temporal information**. Video-based datasets capture the motion patterns of gestures, making them more suitable for training models such as CNN–LSTM or 3D CNN architectures.

7.3 Challenges Related to SLR Datasets

Despite the availability of several datasets, significant challenges remain in the development of comprehensive datasets for SLR systems:

- **Limited dataset size:** Many datasets contain a small number of samples, which restricts model generalization.

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

- **Lack of multilingual datasets:** Most datasets focus on a single sign language such as American Sign Language.
- **Limited signer diversity:** Many datasets include a small number of participants, reducing robustness across different users.
- **Controlled recording conditions:** Many datasets are recorded in controlled environments with simple backgrounds, which may not reflect real-world conditions.

To address these challenges, future research should focus on developing **large-scale, multilingual, and real-world datasets** that capture the diversity and complexity of sign language communication.

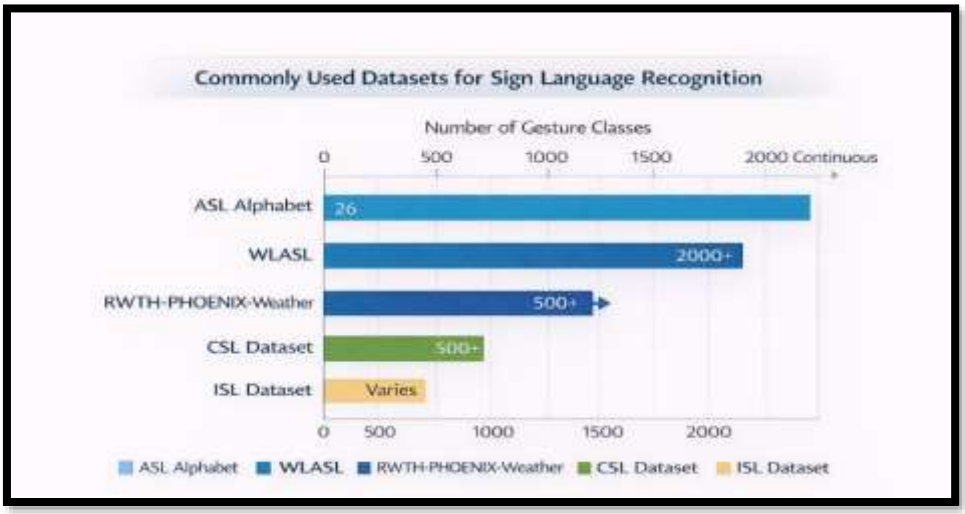

Fig. 4. Commonly used datasets for sign language recognition.

8. CHALLENGES IN SIGH LANGUAGE RECOGNITION

Despite significant advancements in machine learning and deep learning techniques, several challenges remain in developing reliable and scalable sign language recognition (SLR) systems. These challenges affect the performance, generalization, and real-world deployment of automated SLR systems.

8.1 Limited Availability of Large-Scale Datasets

One of the major challenges in SLR research is the lack of large and diverse datasets. Many sign languages, particularly regional languages such as Indian Sign Language (ISL), have limited publicly available annotated datasets. Deep learning models require large amounts of labelled data for effective training, and the absence of such datasets restricts model performance and generalization.

8.2 Signer Variability

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

Different individuals perform sign gestures with variations in speed, hand orientation, motion trajectory, and facial expressions. These variations introduce inconsistencies in gesture representation, making it difficult for recognition models to generalize across different users. Handling signer-independent recognition remains a significant research challenge.

8.3 Environmental Variations

Real-world environments introduce challenges such as changes in lighting conditions, background clutter, camera angles, and occlusion. Many existing SLR systems are trained in controlled environments, and their performance often degrades when deployed in real-world settings.

8.4 Computational Complexity

Deep learning architectures such as CNN-LSTM and 3D CNN models achieve high recognition accuracy but require substantial computational resources for training and inference. This makes it difficult to deploy such systems on mobile devices or embedded platforms where computational resources are limited.

8.5 Continuous Sign Recognition

Most existing studies focus on recognizing isolated gestures. However, real-world sign language communication consists of continuous sequences of gestures forming sentences. Continuous sign recognition requires models capable of understanding temporal dependencies and segmenting gesture sequences accurately, which remains a complex problem.

8.6 Lack of Multilingual and Cross-Dataset Generalization

Many SLR systems are designed for specific sign languages such as American Sign Language (ASL). However, sign languages vary significantly across regions. Models trained on one dataset often fail to generalize to other sign languages due to differences in gesture structures and vocabulary.

Summary of Key Challenges

Table 5. Summary of Key Challenges

Challenge	**Impact on SLR Systems**
Limited datasets	Restricts deep learning model training
Signer variability	Reduces model generalization
Environmental variations	Affects recognition accuracy
High computational cost	Limits real-time deployment
Continuous sign recognition	Difficult sequence modeling
Multilingual differences	Limits cross-language applicability

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

9. FUTURE RESEARCH DIRECTIONS

Although significant progress has been made in sign language recognition (SLR) using machine learning and deep learning techniques, several research opportunities remain for improving the performance, scalability, and real-world applicability of SLR systems. Future research should focus on addressing current limitations related to datasets, model efficiency, multimodal learning, and deployment challenges.

9.1 Development of Large-Scale Multilingual Datasets

One of the major limitations in current SLR research is the lack of large and diverse datasets covering multiple sign languages. Future research should focus on developing multilingual datasets that include various sign languages such as American Sign Language (ASL), Indian Sign Language (ISL), Arabic Sign Language, and others. Large-scale datasets with multiple signers and real-world recording conditions can significantly improve the robustness of deep learning models.

9.2 Lightweight Deep Learning Models for Real-Time Applications

Deep learning models such as CNN–LSTM and 3D CNN architectures provide high recognition accuracy but often require substantial computational resources. Future research should explore lightweight neural network architectures that can run efficiently on mobile devices and edge computing platforms. Model compression, pruning techniques, and efficient network architectures can help achieve real-time sign language recognition.

9.3 Multimodal Sign Language Recognition

Sign language communication involves not only hand gestures but also facial expressions and body movements. Future SLR systems should incorporate multimodal learning techniques that combine visual hand gestures with facial expression recognition and body posture analysis. Integrating multiple modalities can significantly improve recognition accuracy and contextual understanding.

Additionally, multimodal approaches combining visual data with skeletal key points and sensor-based inputs have shown improved robustness in real-world environments [8], [18].

9.4 Transformer-Based Architectures

Transformer-based architectures have recently gained attention in sign language recognition due to their ability to model long-range dependencies using self-attention mechanisms. Vision Transformers (ViT) and hybrid transformer models have demonstrated improved performance on large-scale datasets such as WLASL [16], [19]. However, these models require large training datasets and high computational resources, which limits their applicability in real-time systems.

9.5 Cross-Dataset Generalization and Transfer Learning

Many existing models perform well only on the datasets used for training. Future research should focus on developing models capable of generalizing across multiple datasets and sign

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

languages. Transfer learning techniques can help leverage knowledge from large datasets to improve recognition performance in smaller datasets.

9.6 Explainable AI for Sign Language Recognition

Deep learning models often operate as black-box systems, making it difficult to interpret their decision-making processes. Future research should explore explainable AI techniques that provide insights into how SLR models recognize gestures. This can improve trust, reliability, and usability of SLR systems in real-world applications.

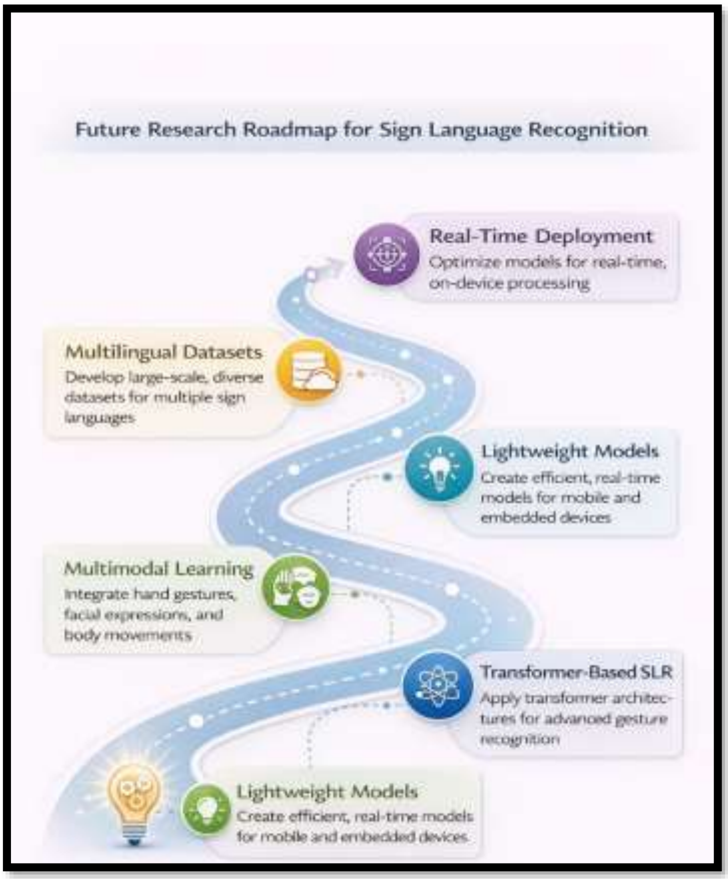

Fig. 5. Future research roadmap for sign language recognition.

10. CONCLUSTION

Sign language recognition (SLR) has emerged as an important research area aimed at improving communication accessibility for the deaf and hard-of-hearing community. This review presented a comprehensive analysis of machine learning and deep learning techniques used for automated sign language recognition. The study categorized existing approaches into classical machine learning methods, deep learning models, and hybrid architectures, highlighting their advantages and limitations.

The comparative analysis revealed that deep learning models, particularly convolutional neural networks and hybrid architectures such as CNN–LSTM, have significantly improved recognition accuracy by effectively capturing spatial and temporal features from gesture data.

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

However, several challenges remain, including limited dataset availability, signer variability, computational complexity, and difficulties in real-time deployment.

This review also analyzed commonly used datasets in SLR research and identified key research gaps, such as the lack of multilingual datasets, limited cross-dataset generalization, and insufficient research on explainable AI techniques. Addressing these challenges is essential for developing scalable and robust SLR systems.

Future research should focus on developing lightweight deep learning architectures, incorporating multimodal learning approaches, and leveraging transformer-based models for improved temporal modeling. Additionally, the creation of large-scale multilingual datasets and real-time deployment solutions will play a crucial role in advancing sign language recognition technologies.

Overall, this review provides a comprehensive overview of the current state of SLR research and highlights promising directions for future studies aimed at improving accessibility and communication through intelligent gesture recognition systems.

ACKNOWLEDGMENTS

The authors declare that no financial or institutional support was received for this research.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

REFERENCES:

[1] A. M. Buttar, U. Ahmad, A. H. Gumaei, A. Assiri, M. A. Akbar, and B. F. Alkhamees, “Deep learning in sign language recognition: A hybrid approach for the recognition of static and dynamic signs,” *Mathematics*, vol. 11, no. 17, p. 3729, Aug. 2023.

[2] V. Ravikiran, “Real-time sign language recognition and translation using MediaPipe and LSTM-based deep learning,” *International Journal of Computer Applications*, vol. 187, no. 25, pp. 10–14, Jul. 2025, doi:10.5120/ijca2025925415.

[3] A. Moryossef, I. Tsochantaridis, R. Aharoni, S. Ebling, and S. Narayanan, “Real-time sign language detection using human pose estimation,” *arXiv preprint arXiv:2008.04637*, 2020.

[4] N. Adaloglou et al., “A comprehensive study on deep learning-based methods for sign language recognition,” *IEEE Transactions on Multimedia*, vol. 24, pp. 1750–1762, 2022.

[5] S. Daga, A. Dusane, and D. Bobby, “With you – Indian sign language detection and alert system,” in *Proc. Int. Conf.*, Pune, India, 2023.

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

[6] S. Kurundkar, A. Joshi, A. Thaploo, S. Auti, and A. Awalgaonkar, "Real-time sign language detection," in *Proc. IEEE Conf.*, 2023.

[7] N. Jindal, N. Yadav, N. Nirvan, and D. Kumar, "Sign language detection using convolutional neural network (CNN)," in *Proc. IEEE World Conf. Applied Intelligence and Computing*, 2022.

[8] H. Kim et al., "Smart gloves-based triboelectric nanogenerator for sign language detection," *Micro and Nano Systems Letters*, vol. 13, p. 11, 2025.

[9] P. Verma and K. Badli, "Real-time sign language detection using TensorFlow, OpenCV and Python," *International Journal for Research in Applied Science & Engineering Technology*, vol. 10, no. 5, May 2022.

[10] A. Deep, A. Litoriya, A. Ingole, V. Asare, S. M. Bhole, and S. Pathak, "Real-time sign language detection and recognition," 2023.

[11] S. Misal and U. Gaikwad, "Sign language detection using machine learning," *International Journal of Innovative Science and Research Technology*, vol. 9, no. 8, Aug. 2024.

[12] B. D. Khuat, D. T. Phung, H. T. T. Pham, A. N. Bui, and S. T. Ngo, "Vietnamese sign language detection using MediaPipe," in *Proc. 10th Int. Conf. Software and Computer Applications (ICSCA)*, Kuala Lumpur, Malaysia, 2021.

[13] 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, p. 1780, 2022.

[14] J. Zhang et al., "Sign language recognition based on dual-path background erasing deep convolutional neural network," *Scientific Reports*, vol. 14, 2024.

[15] D. Kumari, S. Gupta, and P. Singh, "Isolated video-based sign language recognition using attention-based LSTM networks," *Electronics*, vol. 13, no. 7, 2024.

[16] I. Mansour et al., "Hybrid deep learning models for Arabic sign language recognition using vision transformers," *Informatics*, vol. 9, no. 11, 2025.

[17] S. T. Abd Al-Latief, "Deep learning for sign language recognition: A comprehensive review," *Journal of Intelligent Systems and Internet of Things*, 2024.

[18] K. Myagila, "Framework for detecting and recognizing sign language from continuous videos using pose estimation," *Computers & Electrical Engineering*, 2025.

[19] A. Alkhoraif et al., "Ensemble transformer-based word-level sign language recognition using WLASL dataset," *Artificial Intelligence Review*, 2025.

[20] S. K. Paul et al., "An Adam-based CNN and LSTM approach for real-time sign language recognition," *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 1, 2023.

[21] Y. S. N. Rao et al., "Dynamic sign language recognition and translation through deep learning: A systematic review," *Journal of Theoretical and Applied Information Technology*, 2024.

ISSN: 3139-3527 | April-June-Issue, Vol. 1, No. 2 (2026) | DOI: 10.66261/56qhw573

[22] J. Rheiner et al., "High-accuracy CPU real-time American Sign Language recognition using deep learning," *Computers & Electrical Engineering*, 2025.

[23] L. Madhan et al., "Sign language recognition in real time: A deep learning framework for human–computer interaction," *Journal of Information Systems Engineering and Management*, 2025.

[24] I. M. Violet et al., "A comprehensive survey on recent advances and challenges in sign language recognition," *Artificial Intelligence Review*, 2025.

[25] D. Li, C. Rodriguez, X. Yu, and H. Li, "Word-level deep sign language recognition from video using the large-scale WLASL dataset," *arXiv preprint arXiv:1910.11006*, 2019.

[26] H. Luqman et al., "SignVLM: A pre-trained large video model for sign language recognition," 2025.

[27] K. Govindharajalu and P. Anand, "Indian sign language recognition utilizing BERT and LSTM models," *International Journal of Computational and Experimental Science and Engineering*, vol. 11, no. 1, 2025.

[28] M. Verma et al., "Sign language detection and recognition using image processing and deep learning techniques," 2025.

[29] P. Jayanthi, R. K. Sathia Bhama Ponsy, and B. Madhubalasri, "Sign language recognition using deep CNN with normalized keyframe extraction and prediction using LSTM," *Journal of Scientific & Industrial Research*, vol. 82, no. 7, pp. 745–755, Jul. 2023.

[30] P. Rawat, P. Kumar, V. K. Tamta, and A. Kumar, "A comprehensive approach to Indian sign language recognition: Leveraging LSTM and MediaPipe holistic for dynamic and static hand gesture recognition," *EAI Endorsed Transactions on AI and Robotics*, vol. 4, 2025, doi:10.4108/airo.8693.