Article

Breaking Barriers with AI: The Evolution and Challenges of Automated Sign Language Recognition

Mamta Joshi [1], Pranjul Khankriyal [1], Yashvi Chandola* [1], and Vivek Uniyal [1]

[1]Department of Computer Science & Engineering, Institute of Technology Gopeshwar, Uttarakhand

DOI: 10.47344/9z9bqn04

Abstract

Communication remains a significant challenge for individuals with hearing impairments and speech-related disabilities, especially when others are not familiar with sign language. Developing technologies that facilitate seamless communication for these individuals is crucial to promote equality for disabled people and accessibility for all. Sign language recognition systems have emerged as a promising solution, typically implemented using a hardware or software-based approach. Hardware solutions, such as sensor-equipped gloves, often pose usability and cost barriers, making them less appealing for widespread adoption. In contrast, software-driven approaches using artificial intelligence (AI), deep learning (DL) and machine learning (ML) offer a more practical and scalable alternative. This paper provides a complete review of recent developments in AI-based sign language recognition systems, with a particular attention towards deep learning architectures such as Convolution Neural Networks (CNNs). The aim is to evaluate current methodologies, highlight their strengths and limitations, and identify potential directions for future research to improve communication technologies for hearing-impaired people.

Keywords: Sign Language Recognition, Machine Learning, Deep Learning, Assistive Technology, Communication Accessibility.

I. Introduction

“As per National Institute on Deafness and Other Communication Disorder (NIH) approx. 7.7% of United States children of the age ranging between 3-17 has had disorder of voice, speech, language or swallowing in past year. Among these children around 67.7% have speech problem Hoffman HJ et al. (2015) [16].” An automated system to bridge the communication gap between

Email: mamtajoshi26dec@gmail.com
Email: shivamkhankriyal515@gmail.com
*Corresponding author: yashvi.chandola@gmail.com
Email: yashvi.chandola@gmail.com ORCID: 0000-0003-0601-7092
Email: vivekuniyal12@gmail.com ORCID: 0009-0002-8487-6351

Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.

hearing-impaired individuals and others holds great promise in the context of AI advancements. Joze et al. (2018) [27]. Some of the research work showed the important aspect of training the model with Spatiotemporal Convolution to process continuous frames in case of video analysis and sign language translation. Their proposed architecture includes R(2+1)D in which they are able to achieve admirable results comparatively to Sports1M, Kinetics, UCF101, and HMDB51 with accuracy of 73.3% which is by far best published result in Sports1M Tran et al, (2018) [1]. Mostly sign language detection is categorised into hardware-based approach and software-based approach. Also keeping in note, the fact that wearing the armband all the time to sign could be uncomfortable and having less amount of data for model training. So far, the authors have seen that majority of research work concludes towards improving the software models and the concern about the lack of dataset availability, which is essential in the accuracy of the trained model.

A. Variations in Sign Languages

As it is encountered till now there are around 138 to 300 distinct sign languages, which is the first language for beyond 72 million hearing impaired people all over world, as signs and gestures varies according to religions and countries Hoffman HJ et al. (2015) [16]. Most of the sign languages are very different from others, which makes model Training and testing Language dependent. Also, datasets for each language are different, and due to the differences between these languages there are limited resources to train the models for sign Language Translation, which always turns out to be a major concern because larger the dataset, the more accurate the trained model will be Joze et al. (2018) [27], Albanie et al. (2020) [28]. The Figure 1 shows the different sign language datasets. Some Sign Languages which are commonly used are ASL(American Sign Language used by around 2,50,000-5,00,000

Fig. 1: Different Sign Language Datasets

people), BSL(British Sign Language used by 1,50,000 people), ISL(Indian Sign Language used by around 1 million to 2.7 million), CSL (Chinese Sign Language used by around 4.2 million peoples as per in 2021), DGS(German Sign Language used by 250,000 people) etc Li, Y., Zhang et al. (2022) [29]. According to Ethnologue and other sources (Joshi et al. [17], 2024; Sridhar et al., 2020 [26]), an estimated 1.5 million sign language users in India were reported to use Indo-Pakistani Sign Language (IPSL) as of 2008, making it one of the major sign languages in the region, shared with Pakistan and characterized by its own linguistic structure, though regional variations may exist; more recent studies suggest the number of users may have increased due to greater awareness and inclusion efforts, though precise updated figures remain limited, highlighting the need for further research and policy support to recognize and promote IPSL in education and accessibility initiatives

The Table I shows the different datasets available to facilitate the operation of training sign language translation models. Phoenix-2014 is a widely used benchmark for continuous sign language recognition. It includes weather forecast videos paired with German Sign Language (DGS) gloss annotations. It is often used to evaluate sequential modelling and translation pipelines. The Chinese Sign Language dataset contains continuous sequences with varied vocabulary and sentence structures. It supports research on region-specific sign recognition and gesture dynamics Li, Y., Zhang et al. (2022) [29]. One of the largest publicly available datasets for CSLR based on British Sign Language (BSL-1K). It includes a broad vocabulary, more than 1000 BSL Signs and many signers, making it suitable for developing signer-independent models Albanie et al. (2020) [28]. The MS-ASL dataset covers isolated sign recognition across 1,000 classes and more than 25,000 annotated videos of American Sign Language. It includes different signer in real-life recording condition enabling Machine Learning Models to understand ASL more Joze et al. (2018) [27]. The ISL is a dataset on Indian Sign Language created and used for the motive of training the machine learning model and research motives on

Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.

TABLE I: Description of Sign Language Translation Datasets

Datasets	Language	Video Instances	Year	Type
Phoenix-2014 [2], [6]	German	6,841	2014	CSLR
CSL [2]	Chinese	25,000	2016	CSLR
BSL-1K [2], [6]	British	273,000	2020	CSLR
MS-ASL [4]–[6], [9]	American	25,513	2019	ISLR
ISL [10], [12], [13], [15]	Indian	18,863	2019	CSLR
GSL [4]	Greek	10,295	2021	CSLR

Note: CSLR: Continuous Sign Language Recognition, CSL: Chinese Sign Language dataset, BSL-1K: British Sign Language Dataset, MS-ASL: large-scale dataset in American Sign Language, ISL: Indian Sign Language Datasets, GSL: Greek Sign Language Datasets, ISLR: Isolated Sign Language Recognition.

Sign Language Translation Joshi et al. (2024) [17]. GSL is a large-scale dataset generated in Greek Sign Language which supports CSLR tasks. It is used for sign language recognition and translation Papadimitriou et al. (2024) [30]. In the Figure 2 the authors had categorised some of the widely used Sign Languages to show the differences between their very foundation. It can be observed in Figure 2(a) that ASL uses single hand to sign whereas in Figure 2(b) ISL and Figure 2(c) BSL uses both hands to sign. Also, almost every sign for alphabets in each language is different from others Albanie et al. (2020) [28].

B. Sign Language Translation

Sign Language Recognition for Models and Translation involves multiple stages to accurately identify sign gestures from the frame and translate these gestures into text or audio. Each step is very important on its own and plays an important role in maintaining the efficiency and accuracy of the translation system. Figure 3 shows the AI-Based sign language recognition system Cycle.

1) Data Collection and Preprocessing: In this process the raw video or image data of sign gestures is collected, they may vary in lighting or background noises to help in training the model in every possible criterion. Preprocessing involves operations like noise removal, background subtraction and gestures segmentation to prepare consistent inputs for the model and remove unnecessary overhead for the model Sridhar et al. (2020) [26]. The Figure 4 shows Steps for Preprocessing of collected data.

2) Feature Extraction and Selection: Appropriate features and gestures from the frame- including hand position, shape, orientation, facial expressions, and motion trajectories-are extracted. Classical methods like Histogram of oriented Gradients (HOG) and optical flow, while deep learning models extract features automatically using CNNs or 3D-CNNs Alzubaidi et al. (2021) [31], Al-Selwi et al. (2024) [32], Chen et al. (2021) [33].

3) Classification (Training and Testing): Extracted features are fed into classifiers (e.g., Random Forest, Neural networks or SVM) or end-to-end deep learning models. These systems are trained on labelled datasets and then tested to evaluate performance across various sign classes.

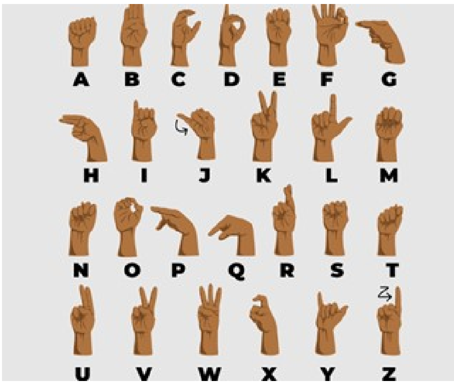
4) Sign-to-Text/Audio Translation: Recognized signs are mapped to their corresponding textual meaning. Using Natural Language Processing (NLP), grammatically correct sentences are generated and optionally converted to audio using text-to-speech (TTS) systems for real-time interaction Nadkarni et al. (2011) [20], Kumar et al. (2023) [21].

5) User Interface and Real-Time Application: A user-friendly interface enables deaf or hard-of-hearing individuals to interact with the system. Integration with cameras or mobile devices facilitates real-time translation in practical settings like classrooms, public services, or workplaces.

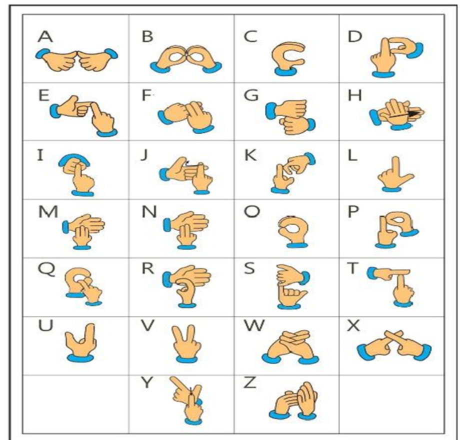
C. Challenges in AI- Based Sign Language Recognition

Many sign languages lack large, annotated datasets, especially for low-resource or regional languages. This scarcity limits the training capability of deep models and reduces generalizability Papadimitriou et al. (2024) [30]. Differences in hand size, speed, signing style, and facial expressions among users can reduce recognition accuracy. Models must be adaptable to signer-independent inputs. Unlike spoken languages, signs rely heavily on spatial grammar, body posture, and facial cues. Capturing and interpreting these nuances accurately requires multimodal fusion and temporal modelling Saunders et al. (2022) [24]. Deploying systems in real-world applications demands low latency and high frame-rate processing, which can be computationally intensive, especially on

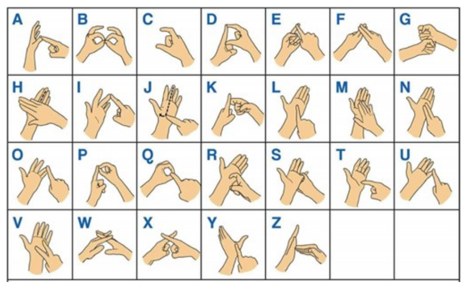
Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.



(a) American Sign Language (ASL)



(b) Indian Sign Language (ISL)



(c) British Sign Language (BSL)

Fig. 2: Comparison of sign language alphabets

edge devices or mobile platforms. Apart from these background clutters and uneven lighting can highly affect the efficiency of the model Kadam et al. (2020) [25]. The Figure 5 illustrates the Challenges in AI-Based Sign Language Recognition

II. Literature Review

Traditional methods for Sign language recognition (SLR) included sensor gloves which are hardware dependent, they are effective but has higher cost and may suffer the user from discomfort. Whereas with advancements in artificial intelligence (AI) and deep learning (DL), Sign Language Recognition has also evolved. Moreover, the focus of study has also shifted towards software-based approaches that utilize the concept of computer vision, neural networks and machine learning to interpret sign language. A major milestone in this domain is the introduction of spatiotemporal convolutional models, which drastically enhanced video-based action recognition. Tran et al. (2018) [1] proposed the R(2+1)D convolutional architecture in their paper, which is a model that separates spatial and temporal dimensions, demonstrating improved performance across several benchmark datasets like Sports1M and UCF101. This concept proves to be efficient for modelling dynamic hand gestures in continuous sign language recognition. Many thorough surveys have highlighted the fact of vast diversity among sign languages and the impact it has on dataset and sign language recognition system development. For instance, Madhiarasan and Roy (2022) [2] categorized sign languages by modality and regional variation, not taking accounts of the challenges for building a generalized models that can handle distinct grammar and gestures across ASL, BSL, ISL, and others. This leads to the requirement of the use of large, Sign language-specific datasets such as Phoenix-2014 for German Sign Language and MS-ASL for American Sign Language Joze et al. (2018) [27], Albanie

Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.

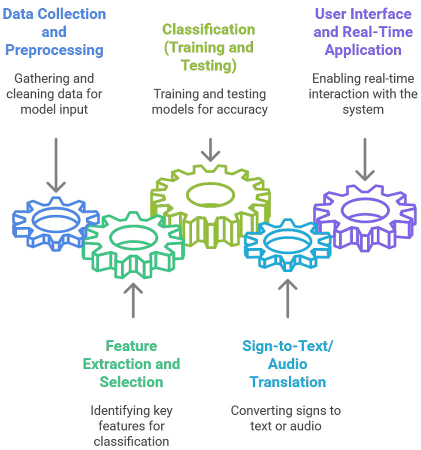


Fig. 3: AI-Based sign language recognition system Cycle

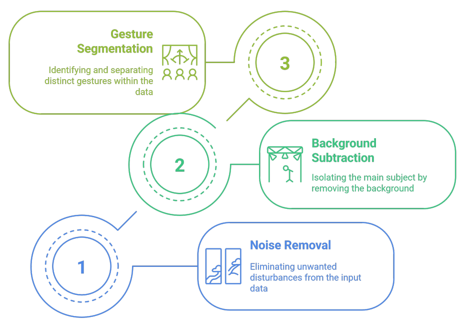


Fig. 4: Steps for Preprocessing of collected data

et al. (2020) [28], Li, Y., Zhang et al. (2022) [29]. To tackle the problem of processing complex multi-video input and user variability, Dignan et al. (2022) [3] developed a group-based recognition model that integrates multiple video streams to enhance classification accuracy. Their approach demonstrates that playing on diverse data perspectives can significantly affect the robustness of the system. Transformer-based models of Vaswani et al. (2017) [5] "Attention is All You Need," have also started to impact SLR architectures. The self-attention mechanism enables models to weigh temporal and spatial features more effectively and efficiently, which has proven useful in continuous sign translation Joze et al. (2018) [27]. Apart from architecture, Adeyanju et al. (2021) [6] critically evaluated machine learning approaches for SLR, identifying limitations in sign language recognition model such as overfitting, lack of performance in real-time, and difficulty in handling background clutter. They supported the use of hybrid models that integrates classical ML techniques with deep networks to reduce these concerns. A significant concern in the field is the scarcity of annotated datasets, especially for non-dominant languages. Subburaj and Murugavalli (2022) [12] noted that most public datasets are constrained in terms of signer variability, background complexity, and linguistic coverage. This data limitation hinders the development of truly scalable SLR systems Sridhar et al. (2020) [26]. Furthermore, several studies have underscored the importance of real-time processing and user adaptability. Najib (2025) [10] explored AI-driven models capable of real-time sign translation, incorporating speech synthesis and natural language processing (NLP) to fill the difference between gesture and spoken communication.

Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.

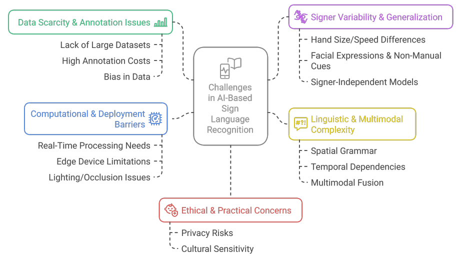


Fig. 5: Challenges in AI-Based Sign Language Recognition

The comparative Table II of SLR studies reveals several important patterns and trends that highlight the evolution of modelling strategies in the field. These results provide insights into researcher's decisions about input data, model settings, and application areas. Learning Rate controls how much the learning model will update its weight after each step. Batch size is defined as the number of samples processed before amending the model. Epoch means the model has seen all the training data once. Optimizer is termed as the algorithm that adjusts the weight. Dropout is a regularization technique to prevent overfitting as it randomly deactivates neurons during training. Learning rate schedule is a strategy in which the authors slowly increase at the beginning, it helps in stabilizing training in early stage of model training. the authors also observed that the demand for video-based input is growing, especially for the sign language recognition system models which are working on Continuous Sign Language Recognition (CSLR) Aloysius et al. (2020) [18]. Video-based methods have more hand movements, facial expression and temporal gestures in comparison to image-based methods. Work by Tran et al. (2018) [1] and Dignan et al. (2022) [3] provides that foundation for incorporating temporal features have greatly improved performance for dynamic signing Saunders et al. (2022) [24]. Most recent research has worked on deep learning models, particularly CNNs, RNNs, or CNN-LSTM combination models Al-Selwi et al. (2024) [32]. Typical hyperparameters are the batch sizes of 16–64, learning rates of approximately 0.001, and optimizers. These values were noticed to balance convergence speed with model stability, especially when the model is trained on medium-to-large datasets. Recent studies such as by Najib (2025) [10] and Papastratis et al. (2021) [4] indicate a growth in interest for the real time deployment of model for better usability and effectiveness. Lightweight models and inference speed are being top priority to consider, especially for mobile applications or assistive systems for deaf people and those who are hard-of-hearing.

A. *Classical Machine Learning Techniques*

The initial work in SLR were mostly based on hand-designed feature extraction and then machine learning classifiers. These methods needed manual design of features like hand orientation, trajectory, and position Adeyanju et al. (2021) [6], Oyeniran et al. (2020) [8].

- Support Vector Machines (SVMs) popularly used for gesture classification, SVMs showed high performance on limited datasets. They were particularly effective for isolated sign recognition tasks.
- Hidden Markov Models (HMMs) HMMs played a crucial role in modelling the temporal properties of gestures. They were able to capture sign transitions through time and were appropriate for sequential data such as CSLR.
- K-Nearest Neighbours (KNN) and Decision Trees were also used but often lacked scalability and robustness in complex, real-world settings.

While these models were computationally efficient, their dependence on manual features limited their generalizability across datasets and signers.

The Figure 6 shows Classical Machine Learning Techniques popularly used for sign language recognition.

The comparative Table III provides a summary about the studies that used classical machine learning methods in Sign Language Recognition system. It also stated the input type, task type, algorithms used for model and accuracy for these models. An accuracy of approximately 92 % has been achieved as highest among them.

B. *Deep Learning-Based Approaches*

With the rise of large datasets and computational power, deep learning transformed SLR by enabling automatic feature learning from raw image or video data Al-Qurishi et al. (2021) [22], Al-Selwi et al. (2024) [32]. The Figure 7 shows Deep Learning

Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.

TABLE II: Studies carried out based on hyper parameters and input type for Sign Language Translation

Investigators (Year)	Input Type	Task Type	Key Hyperparameters	Description
Tran et al. (2014) [1]	Video	CSLR	Model's Learning Rate: 0.01, number of Epochs: 200, Batch Size: 16, Optimizer: SGD, Dropout: 0.5	R(2+1)D CNN, tested on Sports1M, UCF101
Dignan et al. (2022) [3]	Multi-Video	CSLR	Batch Size: 32, Epochs: 150, Optimizer: Adam	Fusion of multiple streams improves performance
Papastratis et al. (2021) [4]	Video	CSLR	Learning Rate of model: 0.001, number of Epochs: 100, Optimizer: Adam	LSTM + CNN hybrid for gesture sequence learning
Vaswani et al. (2017) [5]	Text (MLP base)	–	Learning Rate Schedule: Warm-up, Dropout: S0.1	Introduces Transformer, foundational in attention-based SLR models
Amrutha & Prabu (2021) [7]	Image	ISLR	Epochs: 100, Optimizer: Adam, Batch Size: 64	Simple CNN on Indian sign dataset
Bhaumik et al. (2023) [9]	Video	CSLR	Learning Rate: 0.0001, Batch Size: 32	Deep CNNs + NLP module for sentence translation
Najib (2025) [10]	Video	CSLR	Batch Size: 16, Optimizer: AdamW, Epochs: 120	End-to-end translation, Transformer backbone
Sultan et al. (2022) [11]	Image	ISLR	Dropout: 0.3, Batch Size: 64, Epochs: 80	Comparative study with CNN, MLP
Hussain et al. (2023) [15]	Video	CSLR	Batch Size: 32, Model's Learning Rate: 0.001	ISL-focused CNN-RNN architecture

Note: ISLR: Isolated Sign Language Recognition, LSTM: Long Short-Term Memory, CSLR: Continuous Sign Language Recognition, MLP: Multilayer Perceptron, RNN: Recurrent Neural Networks, SPORTS1M: Larger Dataset of YouTube videos, UCF101: Action Recognition Dataset, CNN: Convolution Neural Networks, BLEU: Bilingual Evaluation Understudy (for sign-to-text translation).

Techniques popularly used for sign language recognition.

1) Convolutional Neural Networks (CNNs): CNNs became foundational for image-based gesture recognition. By extracting hierarchical spatial features, CNNs significantly improved recognition accuracy Subburaj and Murugavalli (2022) [12], Amrutha and Prabu (2021) [7], Al-Qurishi et al. (2021) [22], Alzubaidi et al. (2021) [31], Al-Selwi et al. (2024) [32].

- Used effectively for static hand signs (e.g., alphabet recognition in ASL).
- Capable of learning complex spatial patterns like finger positions and hand shapes.
- Limitations include poor temporal understanding when used alone on video sequences.

2) Recurrent Neural Networks (RNNs) and LSTMs: To model temporal features in sign sequences, RNNs and their other variations like Long Short-Term Memory (LSTM) networks were introduced Papastratis et al. (2021) [4], Al-Selwi et al. (2024) [32].

- Effective for learning sequential relationships in continuous video.
- Can be integrated with CNNs (CNN-LSTM models) to learn spatial-temporal features.
- Suffer from limitations like vanishing gradients and high training time for long sequences.

3) 3D CNNs and Spatiotemporal Models: 3D CNNs can operates on spatial and temporal features concurrently Tran et al. (2018) [1], Alzubaidi et al. (2021) [31], Chen et al. (2021) [33]. Notable models include:

- C3D, I3D, and R(2+1)D: These extract features from frame sequences and learn motion dynamics.

Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.

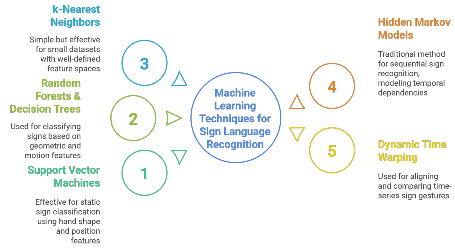


Fig. 6: Classical Machine Learning Techniques for SLR

TABLE III: Studies using Classical Machine Learning Approaches in Sign Language Recognition

Investigators	Input Type	Task Type	Algorithm Used	Accuracy
Adeyanju et al. (2021)	Image/Video	ISLR and CSLR	SVM, KNN, Decision Tree	80-92%
Amrutha and Prabu (2021) [7]	Image	ISLR	SVM	∼91%
Ozguizua et al. (2020)	Image/Video	ISLR	Naive Bayes, Decision Tree, KNN	78-88%
Nair and Bindu (2013)	Image	ISLR	Template Matching, HMM	∼80%
Sultan et al. (2022)	Image	ISLR	MLP, Decision Tree, KNN	85-89%
Hussain et al. (2023)	Image/Video	ISLR	SVM, Random Forest	∼90%

Note: Input type: static image frames or video sequences; Task type: ISLR (Isolated Sign Language Recognition) or CSLR (Continuous Sign Language Recognition).

- Particularly useful for CSLR, where the motion trajectory is vital.
- These models achieve higher accuracy but are computationally intensive.

The Table IV summarizes the studies that has improvised deep learning methods in Sign Language Recognition, it has also provided details about the input type, task type, model used and accuracy for these models. The maximum of 90 % accuracy has been reported approximately for the studies Al-Qurishi et al. (2021) [22].

4) Multimodal and Hybrid Models: Advanced systems often integrate multiple input types to improve robustness Dignan et al. (2022) [3]. Multimodal SLR combines vision (camera input), depth data (e.g., Kinect), and hand pose sensors. Hybrid Architectures use CNNs for feature extraction from image and LSTMs/Transformers for modelling of sequence. Some models integrate Natural Language Processing (NLP) for grammar-aware output and Text-to-Speech (TTS) systems for audio feedback. These approaches aim to replicate the richness of human communication by combining visual cues with linguistic and contextual processing.

The comparison Table V highlights the studies which uses multimodal or hybrid models in Sign Language Recognition, about their input type, task type, models used and their reported accuracy, an accuracy of 90 % approximately has been noted as highest for these models Sarhan et al. (2023) [19], Aloysius et al. (2020) [18].

5) Real-Time and Lightweight Models: Deploying SLR systems in real-world settings like mobile devices or AR glasses demands efficient models Najib (2025) [10], Papastratis et al. (2021) [4].

- Use of MobileNet, Tiny-YOLO, and quantized neural networks reduces computational cost Wang, W et al. (2020) [23].
- Techniques like frame skipping, model pruning, and knowledge distillation allow real-time inference without significant accuracy loss.

Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.

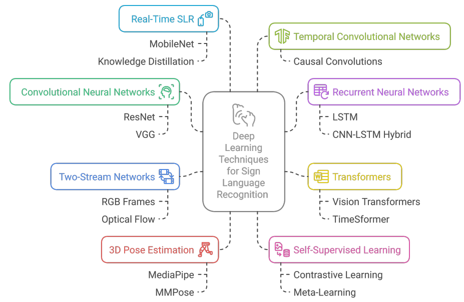


Fig. 7: Deep Learning Techniques for SLR

TABLE IV: Studies Using Deep Learning Techniques in Sign Language Recognition

Investigators	Input Type	Task Type	Model Used	Accuracy
Tran et al. (2018) [1]	Video	CSLR	R(2+1)D CNN	73.3% (Sports1M)
Dignan et al. (2022) [3]	Multi-Video	CSLR	Multi-stream CNN + Ensemble	∼85%
Papastratis et al. (2021) [4]	Video	CSLR	CNN + LSTM Hybrid	∼90%
Vaswani et al. (2017) [5]	Text (NLP)	–	Transformer	Foundational work
Bhaumik et al. (2023) [9]	Video	CSLR	Deep CNN + NLP	∼88%
Najib (2025) [10]	Video	CSLR	Transformer-based + TTS	∼89%
Ibrahim et al. (2020) [14]	Video	CSLR	Conceptual DL architectures	N/R
Subburaj and Murugavalli (2022) [12]	Video/Image	CSLR/ISLR	CNNs, RNNs (survey)	N/A

Note: Input type: video, multi-video streams or NLP inputs; Task type: ISLR (Isolated Sign Language Recognition) or CSLR (Continuous Sign Language Recognition); N/R: Not Reported; N/A: Not Applicable; TTS: Text-to-Speech.

Real-time models prioritize speed and energy efficiency, often at the expense of some precision, but are vital for practical adoption.

The Table VI provides the summary of studies which have used real-time and lightweight models in Sign Language Recognition illuminating the reported accuracy of the models used and their input type along with task types. A highest of 90% accuracy has reached approximately in the studies Sarhan et al. (2023) [19], Aloysius et al. (2020) [18].

The Table VII provides performance analysis of different Studies carried out in Sign Language Translation. The majority of high-performing models utilize video sequences as input, particularly for Continuous Sign Language Recognition (CSLR) Aloysius et al. (2020) [18], Sarhan et al. (2023) [19]. This trend reflects the importance of capturing motion dynamics and sequential dependencies in real-time signing, as seen in studies like those by Tran et al. (2018) [1] and Dignan et al. (2022) [3]. Static image inputs, used mainly in older or isolated sign models, tend to offer less contextual accuracy. Most studies prioritize accuracy as the primary metric, often accompanied by F1-score, precision, or recall. This dual focus ensures both correct classification and balance across varied sign classes. Some advanced systems, like Najib (2025) [10], also include BLEU scores for language-level evaluation, reflecting the evolution toward full sign-to-text translation pipelines. While only a few models (e.g., Najib (2025) [10], Hussain et al. (2023)

Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.

TABLE V: Studies Using Multimodal or Hybrid Models in SLR

Investigators	Input Type	Task Type	Model + Algorithm	Accuracy
Dignan et al. (2022) [3]	Multi-Video Streams	CSLR	Multi-view CNN Ensemble	∼85%
Papastratis et al. (2021) [4]	Video + Sequential	CSLR	CNN + LSTM Hybrid	∼90%
Bhaumik et al. (2023) [9]	Video + NLP	CSLR	Deep CNN + Sentence NLP	∼88%
Najib (2025) [10]	Video + TTS	CSLR	Transformer + TTS	∼89%
Subburaj and Murugavalli (2022) [12]	Video/Image + Motion	CSLR/ISLR	Vision + Temporal Fusion	N/A

Note: ISLR: Isolated Sign Language Recognition, CSLR: Continuous Sign Language Recognition, CNN: Convolutional Neural Networks, LSTM: Long Short-Term Memory, TTS: Text-to-Speech, N/A: Not Applicable.

TABLE VI: Studies Using Real-time and Lightweight Models in SLR

Investigators	Input Type	Task Type	Model Used	Accuracy
Dignan et al. (2022) [3]	Multi-Video	CSLR	Multi-stream CNN (optimized)	∼85%
Papastratis et al. (2021) [4]	Video	CSLR	CNN + LSTM (low-latency)	∼90%
Najib (2025) [10]	Video + TTS	CSLR	Transformer (real-time)	∼89%
Hussain et al. (2023) [15]	Image/Video	ISLR	SVM + Fast Features	∼90%

Note: ISLR: Isolated Sign Language Recognition, CSLR: Continuous Sign Language Recognition, TTS: Text-to-Speech.

[15]) are explicitly designed for real-time processing, this capability is gaining importance. Real-time readiness is still constrained by computationally intensive demands, particularly with models using Transformers or 3D CNNs. Nevertheless, more recent studies increasingly focus on speed optimization in order to facilitate mobile and wearable deployment Wang, W et al. (2020) [23], Chen et al. (2021) [33].

III. Challenges and Limitations

Even though there are rapid advancement in deep learning and multimodal modelling systems, AI-based Sign Language Recognition system still faces significant challenges. These challenges include both the linguistic complexity (how complicated a particular sign language is?) and technical constraints of certain systems. The Figure 8 shows the major challenges and limitations in sign language recognition.

A. Dataset Scarcity and Lack of Standardization

One challenge that has been persistent in SLR research is the shortage of availability of large, different, and annotated datasets. Many datasets that are widely used such as MS-ASL, RWTH-PHOENIX, and BSL-1K are language-specific and often lack required signer diversity to train the learning model, real-world variability, or consistent annotation protocols for the accuracy and efficiency of model Madhiarasan and Roy (2022) [2], Subburaj and Murugavalli (2022) [12]. The lack of standardization in datasets makes it difficult to create a specific term for models to follow to generalize performance across languages Joze et al. (2018) [27], Albanie et al. (2020) [28], Li, Y., Zhang et al. (2022) [29], Papadimitriou et al. (2024) [30].

B. Co-Articulation and Temporal Segmentation

In continuous sign language recognition (CSLR), signs are not neatly segmented a s they are interconnected and gestures flow into one another, this is termed as a phenomenon known as co-articulation. Detecting the boundaries between individual signs

Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.

TABLE VII: Performance of Studies for Sign Language Translation

Authors (Year)	Dataset Used	Metrics Reported	Real-Time	Signer Dependency
Tran et al. (2018) [1]	Sports1M, UCF101	Top-1:73.3%, Top-5	No	Dependent
Dignan et al. (2022) [3]	Custom (multi-video)	Acc:∼85%, Prec/Rec	Partial	Mixed
Papastratis et al. (2021) [4]	RWTH-PHOENIX	Acc, F1-score	Yes (Proto)	Independent
Vaswani et al. (2017) [5]	N/A	BLEU, Accuracy	Yes	N/A
Amrutha & Prabu (2021) [7]	Custom ISL	Acc:∼91%	No	Dependent
Bhaumik et al. (2023) [9]	MS-ASL	Acc:∼88%, TTS	Yes	Independent
Najib (2025) [10]	Custom Transformer	Acc, BLEU, Latency	Yes (Edge)	Independent
Sultan et al. (2022) [11]	ISL, MS-ASL	Acc, F1-score	No	Dependent
Nair & Bindu (2013) [13]	ISL	Acc:∼80%	No	Dependent
Hussain et al. (2023) [15]	ISL	Acc:∼90%, Speed	Yes	Mixed

Note: CSLR: Continuous SL Recognition, ISLR: Isolated SL Recognition, BLEU: Bilingual Evaluation Understudy, MS-ASL: American SL Dataset, ISL: Indian SL, RWTH-PHOENIX: German SL Dataset.

without creating segments in manually is a very difficult task Ibrahim et al., (2020) [14]. This issue gets even more complicated by variations in speed of signing between different individuals, expression, and regional styles Sarhan et al. (2023) [19], Saunders et al. (2022) [24].

C. Generalization to Unseen Signers and Environments

One if the major constraints for the Models is that they are trained on a limited set of signers and they often fail to generalize in recognizing and differentiating the unseen users due to differences in hand shapes, motion patterns, body proportions, and signing styles Papastratis et al., (2021) [4]. Additionally, models trained in controlled settings like good lighting, no background clutter etc. perform less efficiently in unrestricted environments with changing backgrounds, lighting, and occlusions Najib (2025) [10].

D. Real-Time Processing and Computational Constraints

Deploying SLR systems in real-world applications such as assistive tools or mobile devices needs models which are lightweight and that can process gestures in real time with minimal or no time difference. However, high-performing models made of architectures like 3D CNNs and Transformers often demands for considerable computational resources, which is commonly not available and makes them unsuitable for edge deployment without significant optimization for real-world use Dignan et al. (2022) [3], Chen et al. (2021) [33].

E. Ethical, Cultural, and Accessibility Considerations

SLR systems must be compatible to the cultural diversity in sign languages as gestures, grammars, syntax, signs all can vary widely across regions. For example, the authors take Indian sign Language is Completely contrasting in sign gestures, and grammar from American sign language, so a model specifically trained on American Sign Language (ASL), is not transferable to Indian Sign

Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.

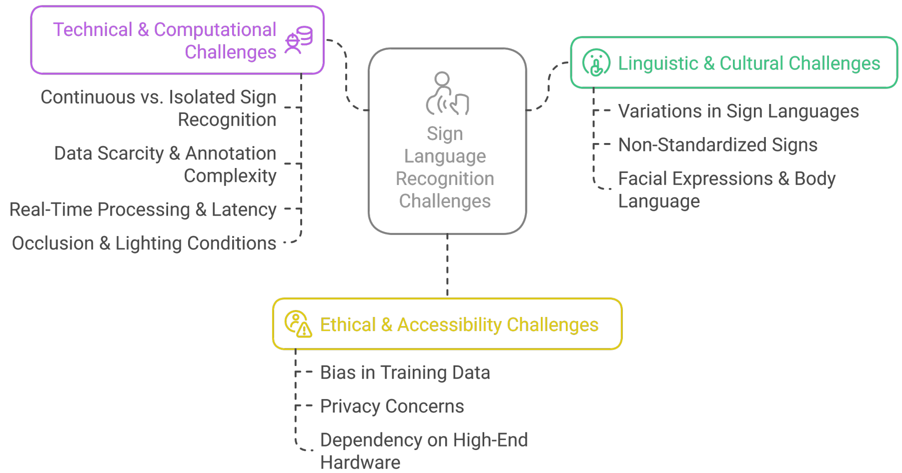


Fig. 8: Challenges and Limitations of Sign Language Recognition

Language (ISL) or British Sign Language (BSL). Moreover, most of the datasets does not consider the representation from disabled, elderly, or non-native signers which can further lead to inefficiency when they try to use it, all these raise concerns about fairness and accessibility Hussain et al. (2023) [15]. There are also privacy and ethical issues about the collection of video data, footage or real- time video capture for training, what if the data is collected without consent Adeyanju et al. (2021) [6]. The Figure 9 shows the Ethical, Cultural, and Accessibility Considerations in sign language recognition.

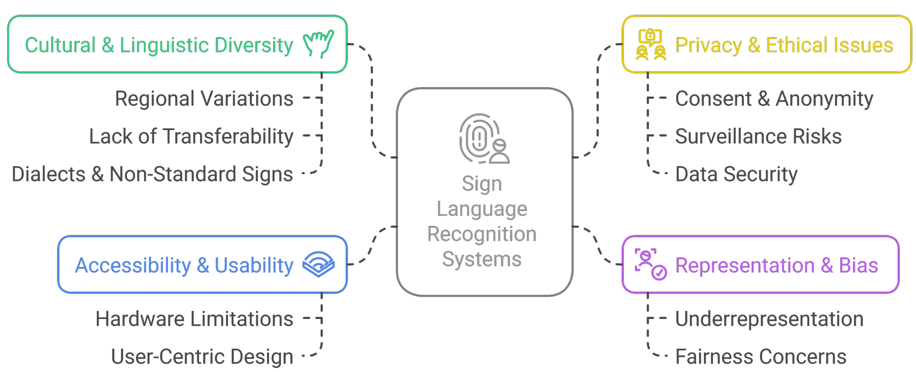


Fig. 9: Ethical and Cultural Considerations in SLR Systems

IV. Future Directions and Research Opportunities

As sign language recognition systems are continuously evolving, there are many promising directions for future research and scopes. The main motive of these scopes is to overcome existing limitations, to enhance model generalizability, accessibility and real time implementation without constraints. The Figure 10 shows the Future Directions and Research Opportunities in sign language recognition.

Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.

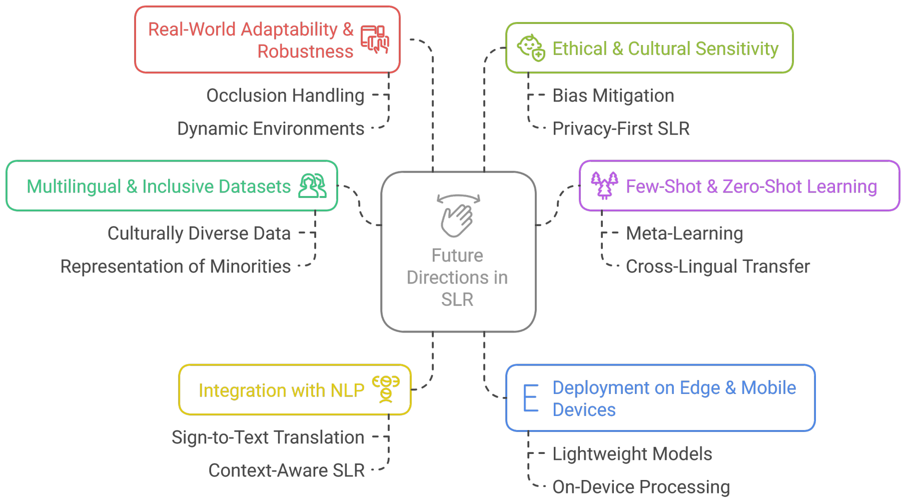


Fig. 10: Future Directions and Research Opportunities in SLR

A. *Expansion of Multilingual and Inclusive Datasets*

An important step toward creating better SLR systems is the creation of larger, multilingual, and demographically diverse datasets for model training. Future datasets should include signers of various ages, abilities, and physical conditions to guarantee fair, accessible and inclusive models. Additionally, some languages that majorly lacks annotated datasets such as African, Southeast Asian, and indigenous sign languages should be considered to annotate as they are largely unexplored Madhiarasan and Roy (2022) [2], Subburaj and Murugavalli (2022) [12], Hussain et al. (2023) [15], Joze et al. (2018) [27], Albanie et al.(2020) [28], Li, Y., Zhang et al. (2022) [29], Papadimitriou et al.(2024) [30].

B. *Few-Shot and Zero-Shot Learning Approaches*

To remove the dependency of models on large annotated datasets, future models can try adopting the few-shot and zero-shot learning paradigms. These approaches work on the simple strategy i.e. to recognize new gestures or signs with minimal possible or no training data. Combining this with techniques like meta-learning and contrastive learning can help our models to generalize to unseen sign classes with better efficiency as per the current level Adeyanju et al. (2021) [6], Najib (2025) [10] Kadam et al. (2020) [25].

C. *Integration with Natural Language Processing (NLP)*

As our current SLR systems can translate signs to individual words or phrases, future research can also focus on context-aware sentence generation using advanced NLP models. Embedding syntactic and grammatical analysis can allow systems to produce natural, grammatically correct output, which will enhance their practical use in real-time translation tools Bhaumik et al. (2023) [9], Najib (2025) [10], Nadkarni et al. (2011) [20], Kumar et al. (2023) [21].

D. *Deployment on Edge and Mobile Devices*

Developing models that are lightweight and efficient which can be deployed on smartphones, AR glasses, and other edge devices without any specific system requirements is a key research direction. Optimizing the already available deep learning models using pruning, quantization, and knowledge distillation techniques will play crucial role in achieving real-time translation of Sign Language without depending on cloud infrastructure Papastratis et al., (2021) [4], Najib (2025) [10].

Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.

E. Real-World Adaptability and Robustness

Models that are to be trained and evaluated should be trained under real-world conditions, such as different lighting condition, occlusions, and complex backgrounds which will help in training the model that provides better efficiency. Furthermore, adaptive models that can learn from user feedback and self-correct in dynamic environments can be implemented to improve long-term usability and reliability Dignan et al. (2022) [3], Ibrahim et al., (2020) [14].

F. Ethical and Cultural Sensitivity in Model Design

As Sign Language Recognition systems are being deployed more widely, developers must not compromise with the cultural, ethical, and accessibility considerations in their design or datasets. This includes respecting the linguistic identity of Deaf communities, ensuring privacy in data collection, and adding sign language users in the development cycle of the model and datasets as well to co-design inclusive technologies Hussain et al. (2023) [15], Adeyanju et al. (2021) [6].

V. Conclusion

This review paper provides a comprehensive discussion about Technologies used in Artificial Intelligence (AI) based Sign Language Recognition Systems, particularly about Machine Learning and Deep Learning approaches. From early classical machine learning models to deep learning models including CNNs, LSTMs, 3D-CNNs, Spatiotemporal models and real-time and lightweight model with MobileNet, the field has taken significant steps in improving gesture recognition, translation accuracy and accessibility for the users Alzubaidi et al. (2021) [31], Al-Selwi et al. (2024) [32], Chen et al. (2021) [33], Nadkarni et al. (2011) [20], Kumar et al. (2023) [21], Kadam et al. (2020) [25]. The comparative analysis shows that video-based inputs and Continuous Sign Language Recognition (CSLR) have become the preferred choice of research and development due to their ability to capture dynamic gestures and real-time interaction. Models like R(2+1)D CNNs, CNN-LSTM hybrids, and attention-based transformers have achieved notable accuracy, especially when trained in vast and diverse datasets. Despite all the achievements these systems remain highly influenced by the quality of datasets, signer variability, and deployment constraints. Apart from these, the major challenges that stays persistent are data scarcity for low-resource sign language, generalizability for unseen signers and real-time processing limitations. Looking ahead, AI-based Sign Language Recognition System holds tremendous amount of potential with real-time lightweight models, few-shot learning techniques and integration with NLP and TTS technologies.

VI. Funding and Conflicts of Interest

- This study was conducted without any financial support from external funding agencies or organizations.
- The researchers involved in this work have no competing interests or financial affiliations to declare that might influence the research findings or their interpretation.

References

[1] D. Tran, H. Wang, L. Torresani, J. Ray, Y. LeCun, and M. Paluri, "A closer look at spatiotemporal convolutions for action recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6450–6459.

[2] M. Madhiarasan and P. P. Roy, "A comprehensive review of sign language recognition: Different types, modalities, and datasets," *arXiv preprint arXiv:2204.03328*, 2022.

[3] C. Dignan, E. Perez, I. Ahmad, M. Huber, and A. Clark, "An AI-based approach for improved sign language recognition using multiple videos," *Multimedia Tools and Applications*, vol. 81, no. 24, pp. 34525–34546, 2022. DOI: 10.1007/s11042-022-13076-8.

[4] I. Papastratis, C. Chatzikonstantinou, D. Konstantinidis, K. Dimitropoulos, and P. Daras, "Artificial intelligence technologies for sign language," *Sensors*, vol. 21, no. 17, p. 5843, 2021. DOI: 10.3390/s21175843.

[5] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[6] I. A. Adeyanju, O. O. Bello, and M. A. Adegboye, "Machine learning methods for sign language recognition: A critical review and analysis," *Intelligent Systems with Applications*, vol. 12, p. 200056, 2021. DOI: 10.1016/j.iswa.2021.200056.

[7] K. Amrutha and P. Prabu, "ML based sign language recognition system," in *2021 International Conference on Innovative Trends in Information Technology (ICITIIT)*, 2021, pp. 1–6. DOI: 10.1109/ICITIIT51526.2021.9399601.

Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.

[8] O. A. Oyeniran, J. O. Oyeniyi, K. A. Sotonwa, and A. O. Ojo, ”Review of the application of artificial intelligence in sign language recognition system,” *International Journal of Engineering and Artificial Intelligence*, vol. 1, no. 4, pp. 20–25, 2020.

[9] R. Bhaumik, S. Patra, D. Chakraborty, S. Basack, P. Mazumder, and P. Das, ”Sign Language Detection and Translation,” *Journal of Mines, Metals and Fuels*, pp. 607–613, 2023.

[10] F. M. Najib, ”Sign language interpretation using machine learning and artificial intelligence,” *Neural Computing and Applications*, vol. 37, no. 2, pp. 841–857, 2025. DOI: 10.1007/s00521-024-09884-8.

[11] A. Sultan, W. Makram, M. Kayed, and A. A. Ali, ”Sign language identification and recognition: A comparative study,” *Open Computer Science*, vol. 12, no. 1, pp. 191–210, 2022. DOI: 10.1515/comp-2022-0241.

[12] S. Subburaj and S. Murugavalli, ”Survey on sign language recognition in context of vision-based and deep learning,” *Measurement: Sensors*, vol. 23, p. 100385, 2022. DOI: 10.1016/j.measen.2022.100385.

[13] A. V. Nair and V. Bindu, ”A review on Indian sign language recognition,” *International Journal of Computer Applications*, vol. 73, no. 22, 2013.

[14] N. B. Ibrahim, H. H. Zayed, and M. M. Selim, ”Advances, challenges and opportunities in continuous sign language recognition,” *Journal of Engineering and Applied Sciences*, vol. 15, no. 5, pp. 1205–1227, 2020.

[15] A. Hussain, N. Saikia, and C. Dev, ”Advancements in Indian sign language recognition systems: Enhancing communication and accessibility for the deaf and hearing impaired,” *Asian Journal of Electrical Sciences*, vol. 12, no. 2, pp. 37–49, 2023.

[16] H. J. Hoffman, C. M. Li, K. E. Bainbridge et al., ”Voice, speech, and language problems in the U.S. pediatric population: The 2012 National Health Interview Survey (NHIS),” *International Journal of Epidemiology*, vol. 44, p. i260, 2015. DOI: 10.1093/ije/dyv096.273.

[17] A. Joshi, R. Mohanty, M. Kanakanti, A. Mangla, S. Choudhary, M. Barbate, and A. Modi, ”iSign: A Benchmark for Indian Sign Language Processing,” *arXiv preprint arXiv:2407.05404*, 2024.

[18] N. Aloysius and M. Geetha, ”Understanding vision-based continuous sign language recognition,” *Multimedia Tools and Applications*, vol. 79, no. 31, pp. 22177–22209, 2020. DOI: 10.1007/s11042-020-09169-x.

[19] N. Sarhan and S. Frintrop, ”Unraveling a decade: a comprehensive survey on isolated sign language recognition,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 3210–3219. DOI: 10.1109/ICCV51070.2023.00300.

[20] P. M. Nadkarni, L. Ohno-Machado, and W. W. Chapman, ”Natural language processing: an introduction,” *Journal of the American Medical Informatics Association*, vol. 18, no. 5, pp. 544–551, 2011. DOI: 10.1136/amiajnl-2011-000464.

[21] Y. Kumar, A. Koul, and C. Singh, ”A deep learning approaches in text-to-speech system: a systematic review and recent research perspective,” *Multimedia Tools and Applications*, vol. 82, no. 10, pp. 15171–15197, 2023. DOI: 10.1007/s11042-023-15489-5.

[22] M. Al-Qurishi, T. Khalid, and R. Souissi, ”Deep learning for sign language recognition: Current techniques, benchmarks, and open issues,” *IEEE Access*, vol. 9, pp. 126917–126951, 2021. DOI: 10.1109/ACCESS.2021.3110912.

[23] W. Wang, Y. Li, T. Zou, X. Wang, J. You, and Y. Luo, ”A novel image classification approach via dense-MobileNet models,” *Mobile Information Systems*, vol. 2020, p. 7602384, 2020. DOI: 10.1155/2020/7602384.

[24] B. Saunders, N. C. Camgoz, and R. Bowden, ”Signing at scale: Learning to co-articulate signs for large-scale photo-realistic sign language production,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 5141–5151. DOI: 10.1109/CVPR52688.2022.00508.

[25] S. Kadam and V. Vaidya, ”Review and analysis of zero, one and few shot learning approaches,” in *Intelligent Systems Design and Applications*, 2020, pp. 100–112. DOI: 10.1007/978-3-030-16657-1_10.

[26] A. Sridhar, R. G. Ganesan, P. Kumar, and M. Khapra, ”Include: A large-scale dataset for indian sign language recognition,” in *Proceedings of the 28th ACM International Conference on Multimedia*, 2020, pp. 1366–1375. DOI: 10.1145/3394171.3413528.

[27] H. R. V. Joze and O. Koller, ”Ms-asl: A large-scale data set and benchmark for understanding american sign language,” *arXiv preprint arXiv:1812.01053*, 2018.

[28] S. Albanie, G. Varol, L. Momeni, T. Afouras, J. S. Chung, N. Fox, and A. Zisserman, ”BSL-1K: Scaling up co-articulated sign language recognition using mouthing cues,” in *Computer Vision–ECCV 2020*, 2020, pp. 35–53. DOI: 10.1007/978-3-030-58580-8_3.

[29] Y. Li, Y. Zhang, Z. Zhao, L. Shen, W. Liu, W. Mao, and H. Zhang, ”CSL: A large-scale Chinese scientific literature dataset,” *arXiv preprint arXiv:2209.05034*, 2022.

[30] K. Papadimitriou, G. Sapountzaki, K. Vasilaki, E. Efthimiou, S. E. Fotinea, and G. Potamianos, ”A large corpus for the recognition of Greek Sign Language gestures,” *Computer Vision and Image Understanding*, vol. 249, p. 104212, 2024. DOI: 10.1016/j.cviu.2024.104212.

Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.

[31] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, O. Al-Shamma, J. Santamaría, M. A. Fadhel, M. Al-Amidie, and L. Farhan, ”Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,” *Journal of Big Data*, vol. 8, no. 1, p. 53, 2021. DOI: 10.1186/s40537-021-00444-8.

[32] S. M. Al-Selwi, M. F. Hassan, S. J. Abdulkadir, A. Muneer, E. H. Sumiea, A. Alqushaibi, and M. G. Ragab, ”RNN-LSTM: From applications to modeling techniques and beyond—Systematic review,” *Journal of King Saud University-Computer and Information Sciences*, vol. 36, no. 2, p. 102068, 2024. DOI: 10.1016/j.jksuci.2024.102068.

[33] C. F. R. Chen, R. Panda, K. Ramakrishnan, R. Feris, J. Cohn, A. Oliva, and Q. Fan, ”Deep analysis of CNN-based spatio-temporal representations for action recognition,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 6165–6175. DOI: 10.1109/CVPR46437.2021.00610.

[34] V. Uniyal, S. Panchpuri, and G. Kamboj, ”A Survey of Privacy-Handling Techniques and Algorithms for Data Mining,” *HCTL Open International Journal of Technology Innovations and Research*, vol. 8, 2014. ISBN: 978-1-62951-499-4.

[35] V. Uniyal, S. Panchpuri, and G. Kamboj, ”A Survey of Security Techniques and Algorithms for Data Mining,” *HCTL Open International Journal of Technology Innovations and Research*, vol. 8, 2014. ISBN: 978-1-62951-499-4.

[36] S. Joshi, R. Shah, Y. Chandola, and V. Uniyal, ”Classification of Brain MRI Images using End-to-end Trained AlexNet & End-to-end Pre-trained MobileNet,” *International Journal of Research Publication and Reviews*, vol. 5, no. 7, pp. 205–218, 2024.

[37] Y. Chandola, V. Uniyal, Y. Bachheti, N. Lakhera, and R. Rawat, ”Data Augmentation Techniques Applied to Medical Images,” *International Journal of Research Publication and Reviews*, vol. 5, no. 7, pp. 483–501, 2024.

[38] Y. Chandola, V. Uniyal, R. Rawat, and A. Dhaundiyal, ”Transformative Impact of Cloud Computing on the Gaming Industry,” *International Journal of Research Publication and Reviews*, vol. 5, no. 7, pp. 465–482, 2024.

[39] Y. Chandola, J. Virmani, H. S. Bhadauria, and P. Kumar, *Deep Learning for Chest Radiographs: Computer-Aided Classification*. Elsevier, 2021. ISBN: 9780323901840.

[40] Y. Chandola, J. Virmani, H. Bhadauria, and P. Kumar, ”Introduction,” in *Deep Learning for Chest Radiographs*, 2021, pp. 1–33. DOI: 10.1016/B978-0-323-90184-0.00003-5.

[41] Y. Chandola, J. Virmani, H. Bhadauria, and P. Kumar, ”Review of related work,” in *Deep Learning for Chest Radiographs*, 2021, pp. 35–57. DOI: 10.1016/B978-0-323-90184-0.00010-2.

[42] Y. Chandola, J. Virmani, H. Bhadauria, and P. Kumar, ”Methodology adopted for designing of computer-aided classification systems for chest radiographs,” in *Deep Learning for Chest Radiographs*, 2021, pp. 59–115. DOI: 10.1016/B978-0-323-90184-0.00008-4.

[43] Y. Chandola, J. Virmani, H. Bhadauria, and P. Kumar, ”End-to-end pre-trained CNN-based computer-aided classification system design for chest radiographs,” in *Deep Learning for Chest Radiographs*, 2021, pp. 117–140. DOI: 10.1016/B978-0-323-90184-0.00011-4.

[44] Y. Chandola, J. Virmani, H. Bhadauria, and P. Kumar, ”Hybrid computer-aided classification system design using end-to-end CNN-based deep feature extraction and ANFC-LH classifier for chest radiographs,” in *Deep Learning for Chest Radiographs*, 2021, pp. 141–156. DOI: 10.1016/B978-0-323-90184-0.00007-2.

[45] Y. Chandola, J. Virmani, H. Bhadauria, and P. Kumar, ”Hybrid computer-aided classification system design using end-to-end Pre-trained CNN-based deep feature extraction and PCA-SVM classifier for chest radiographs,” in *Deep Learning for Chest Radiographs*, 2021, pp. 157–166. DOI: 10.1016/B978-0-323-90184-0.00005-9.

[46] Y. Chandola, J. Virmani, H. Bhadauria, and P. Kumar, ”Lightweight end-to-end Pre-trained CNN-based computer-aided classification system design for chest radiographs,” in *Deep Learning for Chest Radiographs*, 2021, pp. 167–183. DOI: 10.1016/B978-0-323-90184-0.00001-1.

[47] Y. Chandola, J. Virmani, H. Bhadauria, and P. Kumar, ”Hybrid computer-aided classification system design using lightweight end-to-end Pre-trained CNN-based deep feature extraction and ANFC-LH classifier for chest radiographs,” in *Deep Learning for Chest Radiographs*, 2021, pp. 185–196. DOI: 10.1016/B978-0-323-90184-0.00009-6.

[48] Y. Chandola, J. Virmani, H. Bhadauria, and P. Kumar, ”Hybrid computer-aided classification system design using lightweight end-to-end Pre-trained CNN-based deep feature extraction and PCA-SVM classifier for chest radiographs,” in *Deep Learning for Chest Radiographs*, 2021, pp. 197–204. DOI: 10.1016/B978-0-323-90184-0.00006-0.

[49] Y. Chandola, J. Virmani, H. Bhadauria, and P. Kumar, ”Comparative analysis of computer-aided classification systems designed for chest radiographs: Conclusion and future scope,” in *Deep Learning for Chest Radiographs*, 2021, pp. 205–209. DOI: 10.1016/B978-0-323-90184-0.00004-7.

Received: June 11, 2025. Reviewed: November 23, 2025. Accepted: December 9, 2025.