International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(5s) 2026 pp. 1177-1190
https://doi.org/10.70917/ijcisim-2026-2863
Article

Survey of Deep Learning-Based Sign Language Recognition Systems for Accurate Gesture Classification and Human–Computer Interaction

Muhammed Sudhaise J.[1], Dr. V. J. Chakravarthy[2], M. Shanmugam[3], K. Arthi[4], A. Janaki Devi[5]

[1]Department of Computer Applications, Dr. M.G.R. Educational and Research Institute, Chennai, Tamil Nadu, India.
Email: sudhaise1741@gmail.com
[2]Faculty of Computer Applications, Dr. M.G.R. Educational and Research Institute, Chennai, Tamil Nadu, India.
Email: chakravarthy.mca@drmgrdu.ac.in
[3]Department of Computer Science and Engineering, Sri Manakula Vinayagar Engineering College, Puducherry, India.
Email: shanmugam.muthalu@gmail.com
[4]Faculty of Computer Applications, Dr. M.G.R. Educational and Research Institute, Chennai – 600095, Tamil Nadu, India.
Email: arthi.mca@drmgrdu.ac.in
[5] Department of Data Science, Dwaraka Doss Goverdhan Doss Vaishnav College, Arumbakkam, Chennai – 600106, Tamil Nadu, India.
Email: janakidevi08@gmail.com

Abstract: Sign Language Recognition (SLR) is becoming an area of paramount importance in the fields of Artificial Intelligence, Computer Vision and Human Computer Interaction (HCI) because of the communication gaps it can fill among hearing impaired and the hearing community. Recently, the recognition of complex human gestures in SL has progressed significantly, with deep learning being used to automatically and accurately extract features and learn from complex SL gestures. It is a review paper that covers a comprehensive survey of deep learning based SLR systems, architectures including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, Gated Recurrent Units (GRUs), Transformers, and hybrid models. In this study, the recent advances in gesture classification, multimodal gesture recognition frameworks, attention mechanisms, transfer learning methods and lightweight deployment strategies are analyzed. Moreover, the review discusses the emerging applications of multimodal sensing technologies such as RGB cameras, wearable sensors, skeletal tracking, radar systems, and facial expression analysis to boost recognition robustness and real-time performance. Other applications are also discussed such as assistive communication, healthcare, education, virtual reality, smart devices and human–robot interaction. Lastly, some of the current challenges, such as signer variation, multilingual recognition, continuous sign language translation and real world deployment of such systems are defined. The results suggest that great progress has been made with deep learning based SLR, and that it is a promising approach to enable the creation of intelligent, accessible and inclusive communication technologies.
Keywords: NA

1. Introduction

Communication is a very important part of how people connect with each other as humans; without communication, people would be unable to share information, convey feelings, or work together. Unfortunately, millions of people who are deaf and have difficulty with speech do not have the ability to communicate with the rest of the world. Deaf and hard of hearing people primarily communicate using sign language; sign language uses hand shapes, hand movements, facial expressions, body posture, and physical space to communicate meaning. Sign language is an effective form of communication, however, when a sign language user tries to communicate with someone who does not use sign language there often arises a communication barrier [1]. Recognising this problem, researchers have increasingly intended to develop intelligent SLR (Sign Language Recognition) systems that can help bridge the communication gap by automatically interpreting sign language gestures into an understandable format, either textually or orally. Advances in artificial intelligence, computer vision and deep learning technologies have allowed for the rapid development of SLR systems that have high levels of accuracy and efficiency and thus serve as promising possibilities for enhancing accessibility and human communication [2].

1.1 Overview of Deep Learning-Based Sign Language Recognition Systems

Sign Language Recognition systems are computational systems that can recognize sign language gestures, analyze and interpret them, automatically using intelligent computational frameworks, visual sensing technologies and machine learning algorithms. Previous SLR methods were mostly based on handcrafted feature extraction methods and traditional classification methods that were not able to cope with the complexity and variability of sign language gestures. The main reasons why these traditional methods were not very successful were: hand shape, movement trajectory, lighting conditions, camera perspectives, occlusions, and signer-specific characteristics. As a result, deep learning has revolutionized the industry by making feature learning available in an automated fashion and greatly boosting recognition accuracy and robustness [3]. The Sign Language Recognition and Sign Language Translation Workflow is shown in figure 1.

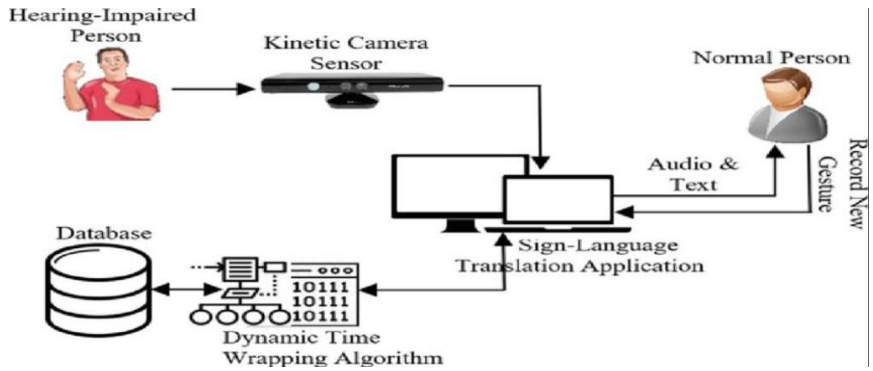

Figure 1: Sign Language Recognition and Translation Workflow [4].

In the sign language recognition field, there are only a few deep learning architectures that are important in this context. CNNs are widely used to detect local patterns, edges, textures and hand configurations from a frame of a video or image, to extract the spatial features. CNN-based models provide accurate modelling of the visual aspects of static sign images and are also powerful models for gesture classification problems. But the temporal aspects and sequence of motion in sign language communication can be more than just movement from one frame to the next. To address this issue, Recurrent Neural Networks (RNNs) were suggested to learn the temporal structure between signs within a sequence of signs.

The variants of RNNs, such as Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs), have been shown to be more effective at capturing long-term dependencies and sequential information. The architectures enable SLR systems to learn gesture transition, motion trajectory and temporal context that are essential for the recognition of dynamic signs and continuous SLS sentences. Recently, transformers have drawn a lot of attention as they efficiently capture long-range dependencies and handle sequential data in parallel due to their self-attention mechanism [5].

Two significant developments to the RNN are the Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs) that are useful for long-term dependencies and sequential data. The architectures enable the SLR system to learn the transition of the gestures between each other, gesture trajectory and temporal context for dynamic gesture recognition and long sign language sentences recognition. Transformer architecture captures the long-range dependencies and processes sequential information in parallel, making it a recent focus of interest. In the past, the transformers have proven to be very effective for various sequence modelling tasks and have been increasingly utilized in sign language recognition systems with the aim of improving the recognition accuracy and scalability. In the field of sign language recognition, deep learning has made a significant impact, enabling sign language interpretation systems to better comprehend the complexity of hand movements, facial expressions, and body postures of the sign language interpreter. This is an important step towards addressing communication issues between hearing and hearing impaired communities. The potential applications of deep learning-based SLR systems in making communication and interaction more seamless and easier will make societies more inclusive and accessible Human–Computer Interaction (HCI) technology for those with communication disabilities [6] available.

1.2 Deep Learning Techniques for Accurate Gesture Classification

The ultimate target of sign language recognition systems is to correctly classify the gestures. This process has been transformed by deep learning with the ability to extract features, recognize patterns and classify them automatically from large datasets. Deep learning algorithms are able to learn discriminative representations without relying on features being carefully engineered as required by traditional machine learning methods, and the structures are constructed in layers of abstraction. The ability to recognize complex gesture patterns and/or subtle variations, with greatly increased accuracy, is a key feature of SLR systems [7]. The gesture recognition has one of its advantages as the deep learning can provide automatic feature learning. CNN-based architectures are able to extract useful visual information such as hand contours, hand finger positions, palm orientation, and motion information from raw image inputs automatically. These learned features are improved layer by layer in the network, resulting in very informative representations and accuracy of gesture classification. This is an auto-feature extraction method that is not based on hand-designed descriptors, and thereby makes the system more flexible to different sign language databases. Different sign language recognition paradigms, such as image-based, video-based and multimodal approaches, use deep learning techniques. Unlike the image-based recognition systems that depend on the static signs, it is based on CNN architectures for the recognition of a single image of a gesture. They work very well in the detection of alphabet signs, single words and pre-defined hand gestures. However, a lot of Signed Languages have a continuous motion or sequence of gestures that requires the temporal modelling ability [8]. Sign Language Interpreter System (SLIS) for Gesture Recognition and Translation is depicted in figure 2.

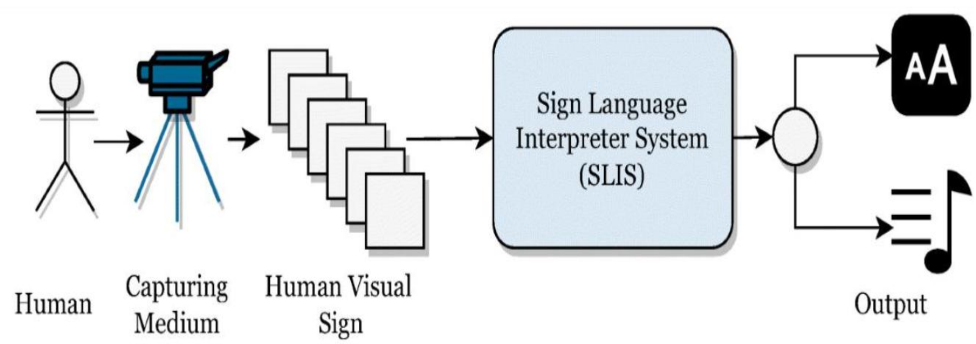

Figure 2: Sign Language Interpreter System (SLIS) for Gesture Recognition and Translation [9].

The way to solve this problem is to have systems that really function with a sequence of image frames and keep storing spatial and temporal data, such as video systems. In this kind of systems, the CNNs usually process the spatio-temporal features of every frame in the video and the RNN models the temporal dynamics across the different gesture sequences.Recurrent networks are used to model the temporal dependencies between gesture sequences and CNNs to encode the features of each frame from the video, for instance. These architectures have been successful for sign language expression recognition and for continuous sign language scenes [10].

Multimodal recognition approaches can also be used to improve gesture classification by combining information from various sensors such as those based on RGB images, depth sensors, skeletal tracking systems, motion sensors, and facial expression analysis. By combining multimodal information, multimodal deep learning frameworks are able to gain more comprehensive contextual information and improve the recognition capability in challenging situations [11]. The feature-level, decision-level or attention-based integration fusion approaches can be used to get a more comprehensive understanding of sign language communication. Moreover, deep learning methods offer great benefits in dealing with variations in gestures and environmental conditions. Sign language hands are different in style and size, the rate of the signing, and cultural influences. In addition, there are problems which can occur in an actual deployment situation including variation of illumination, partial occlusions, camera viewpoint variation and background clutter. The high degree of generalization ability of deep neural networks leads to high recognition performance even in different conditions [12] while at the same time they can learn invariant features.

On the part of models, the large-scale sign language datasets have also accelerated the success of deep learning models. With the high dimensionality of the modern datasets, that contain thousands of gesture samples, and the huge number of training data available, it is possible to train a powerful architecture which can learn robust feature representations and achieve high classification rate. Further, techniques such as transfer learning, data augmentation, self-supervised learning and domain adaptation, which can help to enhance the generalization of the model across sign languages and application scenarios, are introduced as these techniques are not widely available in databases [13].

1.3 Human–Computer Interaction and Applications of Sign Language Recognition

Sign language recognition systems using deep learning have been developed at a very fast pace, and have now become a part of the Human–Computer Interaction applications. HCI is the design of systems that allows for a natural and efficient way for humans to communicate with computers. The implementation of SLR systems in computers allows them to interpret and understand sign language gestures, which is important for the accessibility and inclusivity of sign language for hearing-impaired users in different technological areas. Assistive communication technology is one of the most crucial applications of SLR systems [14]. Sign language recognition systems can be powered by deep learning to convert sign language gestures into text or speech in real-time, enabling hearing-impaired people to communicate with non-signers. These technologies help to remove communication barriers and improve participation in educational, professional and social activities. The portable communication devices, smart phone applications, and wearable devices with sign language recognition for common communication needs are developed [15].

The sign language recognition integration is also effective with smart interfaces and intelligent virtual assistants. Gesture-based interaction techniques can be incorporated to allow users to interact, retrieve information and execute commands through sign language via AI-powered systems. These characteristics contribute to creating an accessible digital environment and will allow individuals with hearing loss to communicate with new technologies in a natural way. Sign language recognition systems can help the communication between Deaf and hard of hearing people and health care workers, specially in clinical settings [16]. The accurate communication of the gestural message enables healthcare specialists to better understand the needs, symptoms and treatment needs of patients, improves access to healthcare services and service quality. Other educational technologies that can benefit education platforms include SLR technologies for interactive learning spaces for the delivery of sign language instruction and language acquisition, and for the delivery of accessible education content [17].

Recently there have been advances in the area of deep learning which further bolstered the ability of Human–Computer Interaction systems. Recently, attention mechanisms have emerged as an important means to selectively allocate computational resources to the most informative portion of gesture sequences. Attention models emphasize some hand gestures, facial expressions, and context cues to improve the recognition accuracy and interpretability. Another major improvement in the field of research on sign language recognition is the use of transformer architectures [18]. The study uses self-attention mechanisms to effectively model complex spatial-temporal dependencies, and to support parallel computation. The excellent performance of the Transformer-based system for continuous hand action recognition and translation in SLR make it highly suitable for real-time hand action recognition and translation for HCI applications. The multimodal fusion techniques have also received a great attention and are characterized by the fusion of visual, skeletal, motion and contextual information in a single recognition framework. These methods enhance the robustness and reliability of the gesture understanding process, and offer a more detailed understanding of human gestures [19]. Furthermore, the implementation of real-time sign language translation systems, especially those running on the basis of deep-learning algorithms, is becoming increasingly complex and can now speak or translate sign language into text as it happens, making sign language communication more efficient and useful.

It is significant what impact feasible gesture classification systems can have on society. Easier access can facilitate the participation of hard of hearing individuals in education, employment, health and socialization. Good communication makes people more independent and more self-confident and self-intelligent assistive technologies help to involve people in society, to make them equal to access to digital services. In the future, more accurate, context-dependent, multilingual and user-friendly sign language recognition systems can be developed, reinforcing the importance of Human–Computer Interaction in developing inclusive technological ecosystems [20].

2. Literature Review

2.1 Deep Learning Architectures for Sign Language Recognition

In recent times, deep learning approaches to Sign Language Recognition (SLR) have shown promising progress, especially in the field of gesture classification with enhanced neural networks and efficient feature extraction techniques. CNN was also compared with RNN by Harmade et al. (2025) [21] and found that CNN had an accuracy of 94% and RNN had an accuracy of 76%, 74%, and 73% for precision, recall and f1-score, respectively, indicating the effectiveness of CNN in extracting spatial gesture features. In the study by Toshpulatov et al. (2025) [22] three models, namely ConvNeXt, Swin Transformer, and Vision Mamba models, were tested on the Turkish, Arabic and ASL databases, and it was concluded that the Transformer-based models showed the best performance in the learning of the subtle variations in space between the different sign languages. Likewise, **Mineo et al. (2026)** [23] proposed the SABRE model for recognizing Italian Sign Language using a radar system, with an accuracy of 81.6% that is higher than the existing baseline methods. To address the complexity of sign language recognition, **Alanazi et al. (2026)** [24] introduced a multimodal Arabic Sign Language recognition framework that fuses the Leap Motion and RGB camera data with VGG16 and dense neural networks, achieving an overall recognition accuracy of 78% and establishing the multimodal fusion as a valuable method for complex gesture understanding.

Taking the dual-view aspect into account, a dual-view word-level sign language recognition system was built by **Jing et al. (2026)** [25] using a two-stage deep neural network that uses early fusion of the frontal view. The proposed method outperformed the current methods, with the Top-1 accuracy gain of 10.29% compared with MViT and 11.38% compared with CNN+Transformer architectures. **Lamaakal et al. (2026)** [26] presented TinyMSLR, a lightweight multilingual framework that combines ConvNeXt-Tiny and Swin Transformer encoders and adopts adaptive fusion and knowledge distillation techniques. With less than 2.7 million parameters, the model had 99.01% validation accuracy, and an F1-score of 98.96% for edge deployment. **Singh et al. (2026)** [27] formulated a Parallel Bidirectional Reservoir Computing (PBRC) architecture using MediaPipe hand tracking, which achieved Top-1, Top-5 and Top-10 accuracies of 60.85%, 85.86% and 91.74% respectively and reduced training time to just 18.67 seconds as compared to over 55 minutes required by Bi-GRU models.

Luna-Jiménez et al. (2025) [28] presented a lightweight Transformer model for isolated sign language recognition with landmarks-based representations with a Weighted-F1 score of 76.99 on AVASAG100 and 62.91 on WLASL100. On the CALSE-1000 Spanish Sign Language dataset, consisting of 5,000 videos across 1,000 glosses, data augmentation methods boosted the Top-1 accuracy by 11.7%, Top-5 accuracy by 8.8% and Top-10 accuracy by 9% as reported by **Pereira-Trigo et al. (2025)** [29]. Furthermore, the study emphasized the significance of facial expressions to achieve a good recognition rate of the signs. **Zhang et al. (2025)** [30] developed a flexible and stretchable triboelectric sensor array with machine learning algorithms for sign language recognition using 1,000 gesture samples, achieving an impressive accuracy of 98.5%. The overall progress underlines the power of deep learning, multimodal sensing, data augmentation, light-weight architectures, and wearables for improving sign language recognition of today's state-of-the-art systems.

Table 1: Studies on Deep Learning-Based Sign Language Recognition Systems.

Author(s), Year & Ref.	Method / Model	Dataset / Language	Key Results
Harmade et al. (2025) [21]	CNN vs. RNN for SIBI Recognition	Indonesian Sign Language (SIBI)	CNN achieved **94% Precision, Recall, and F1-score**, outperforming RNN (**76% Precision, 74% Recall, 73% F1-score**).
Toshpulatov et al. (2025) [22]	ConvNeXt, Swin Transformer, Vision Mamba	Turkish, Arabic, and American Sign Languages	Transformer-based models demonstrated superior cross-lingual recognition and generalization capabilities.
Mineo et al. (2026) [23]	SABRE (Sign Attention-Based Bidirectional Recurrent Encoder)	MultiMeDaLIS Italian Sign Language Dataset	Achieved **81.6% accuracy**, outperforming benchmark models for radar-based recognition.
Alanazi et al. (2026) [24]	Multimodal Fusion (Leap Motion + RGB Camera + VGG16)	Arabic Sign Language (18 Words)	Recognized **13 out of 18 words** with an overall accuracy of **78%**.

Jing et al. (2026) [25]	Dual-View Deep Neural Network with FvGEF Fusion	NationalCSL-DP Dataset	Improved Top-1 accuracy by **10.29% over MViT** and **11.38% over CNN+Transformer** methods.
Lamaakal et al. (2026) [26]	TinyMSLR (ConvNeXt-Tiny + Swin Transformer + Knowledge Distillation)	DGS RWTH-PHOENIX-Weather 2014T and Mandarin CSL	Achieved **99.01% validation accuracy** and **98.96% F1-score** with less than **2.7M parameters**.
Singh et al. (2026) [27]	Parallel Bidirectional Reservoir Computing (PBRC) + MediaPipe	WLASL Dataset	Obtained **60.85% Top-1**, **85.86% Top-5**, and **91.74% Top-10 accuracy** with training time reduced to **18.67 seconds**.
Luna-Jiménez et al. (2025) [28]	Lightweight Transformer with Landmark-Based Features	AVASAG100 and WLASL100	Achieved **76.99 W-F1** on AVASAG100 and **62.91 W-F1** on WLASL100.
Perea-Trigo et al. (2025) [29]	I3D with Data Augmentation Techniques	CALSE-1000 Dataset (5,000 Videos, 1,000 Glosses)	Improved **Top-1 accuracy by 11.7%**, **Top-5 by 8.8%**, and **Top-10 by 9%**.
Zhang et al. (2025) [30]	Wearable Triboelectric Sensor-Based Sign Language Translation System	1,000 Gesture Samples	Achieved **98.5% recognition accuracy** using flexible and stretchable sensor arrays with machine learning algorithms.

Figure 3 shows a comparison of the recent Sign Language Recognition (SLR) studies using deep learning. The best result was obtained using the TinyMSLR method proposed by Lamaakal et al. (2026) [26] with a 99.01% validation accuracy, which highlights the potential of the proposed approach of combining ConvNeXt-Tiny, Swin Transformer and knowledge distillation. Wearable triboelectric sensors were also adopted by Zhang et al. (2025) [30] with a recognition accuracy of 98.5%. The CNN model outperformed other classifiers and models with a high accuracy of 94% as reported by Harmade et al. 2025 [21]. On the other hand, Singh et al. (2026) [27] achieved the least Top-1 accuracy (60.85%) with an appreciable reduction in computation time to run in real-time edge deployment.

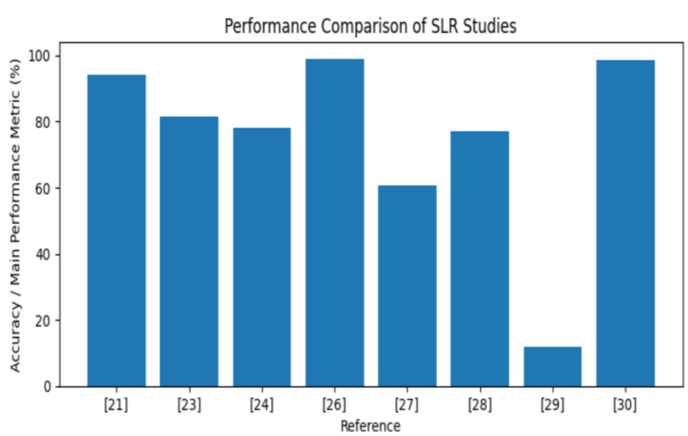

Figure 3: Accuracy Comparison of Recent Sign Language Recognition Studies.

2.2 Gesture Classification and Multimodal Recognition Approaches

In recent years, significant advances in Sign Language Recognition (SLR) systems using deep learning have been achieved. These cutting-edge architectures, including CNNs, Transformers, YOLO models, and multimodal approaches, have been used to enhance the precision and efficiency of gesture recognition. The below studies focus on recent advancements and accomplishments on SLR technology during 2025 to 2026. According to **Nuankaew et al. (2026)** [31] the evaluation of four YOLO variants for Thai Sign Language recognition showed that YOLOv11 was more precise (95.27%) with high recall (97.00%), high mAP@50 (97.23%), and high mAP@50:95 (66.79%) with low latency. In a similar way, **Garcia-Gil et al. (2026)** [32] proposed a multi-modal angular-geometry representation coupled with a BiLSTM network for dynamic recognition of Mexican Sign Language (MSL), where it achieved 99% accuracy and 99% macro F1-score with significantly less computational complexity. In this regard, **Nasef et al. (2026)** [33] proposed ArSL-TGRU model, which combines Transformer layers and GRU networks for Sign Language recognition. The model was trained on a dataset of 60,918 videos and 136 signs, and it achieved better generalization than six benchmark architectures. In addition, **Hurtado-Sánchez et al. (2026)** [34] have created an adaptive mobile learning application for Mexican Sign Language (MSL) that combines CNN and pose-estimation techniques, with an accuracy of sign recognition of 95.7% for real time sign recognition, which could be helpful for inclusive education and personalized learning environments.

In recent years, studies have also been conducted to realize the intelligent sign language systems in real-time and to increase the efficiency of the computation. **Mouti et al. (2026)** [35] presented LSTM–3D CNN based hybrid system for static and dynamic gesture recognition using IoT. With these results, the LSTM got an F1 score of 93.8% and an accuracy of 86.67%, and the 3D-CNN achieved 99.6% accuracy for the dynamic sign recognition. The system was implemented in real-time, 12–15 FPS with 65 ms latency on a raspberry pi device. Similarly, **Algethami et al. (2025)** [36] proposed a Temporal Convolutional Network (TCN) for the ASL recognition task and obtained 99.5% accuracy whereas the improved RNN-BiLSTM model achieved 99% accuracy. **Alsharif et al. (2025)** [37] combined YOLOv11 and MediaPipe hand tracking to provide real-time American Sign Language interpretation results with a remarkable mAP@0.5 of 98.2%, demonstrating the potential of AI-supported assistive technologies. Furthermore, **Guan et al. (2025)** [38] introduced MSKA-SLR framework by employing a multi-stream keypoint attention network and attained the state-of-the-art result on the benchmark of sign language translation (SLRT) Phoenix-2014T for keypoint based recognition, showing the importance of attention and self-distillation.

Wearable and lightweight real-time recognising technologies have also broadened the scope of the applicability of SLR systems in real world applications. SpellRing is a smart-ring based American Sign Language fingerspelling recognition system proposed by **Lim et al. (2025)** [39] to integrate active acoustic sensing, inertial sensing and deep learning. Using real-time evaluation, the system obtained the Top-1 accuracy of 82.45%, the Top-5 accuracy of 92.42% and the Word Error Rate (WER) of 0.099. Furthermore, fast hand gesture recognition framework using MediaPipe, MLP and PointNet for classification of American Sign Language was proposed by **Zhang et al. (2025)** [40]. The proposed approach achieves a very good accuracy of 99.81%, with the inference time less than 10ms, hence, very suitable for real-time applications. Overall, these studies demonstrate the promise of these novel methods in improving the accuracy, efficiency, and usability of existing sign language recognition systems, especially regarding the integration of multimodal data, attention models, and lightweight deployment structures.

Table 2: Performance Analysis of Deep Learning-Based SLR Models.

Author(s), Year & Ref.	Method / Model	Dataset / Language	Key Results
Nuankaew et al. (2026) [31]	YOLOv8, YOLOv9, YOLOv11, YOLOv12	Thai Sign Language (10,956 Images, 95 Classes)	YOLOv11 achieved **95.27% Precision**, **97.00% Recall**, **97.23% mAP@50**, and **66.79% mAP@50:95**, outperforming other YOLO variants.
Garcia-Gil et al. (2026) [32]	Multimodal Angular-Geometry Representation + BiLSTM	Mexican Sign Language (MSL)	Achieved **99% Accuracy** and **99% Macro F1-score** with reduced computational complexity.

Nasef et al. (2026) [33]	ArSL-TGRU (Transformer + GRU)	Arabic Sign Language (60,918 Videos, 136 Signs)	Demonstrated superior signer-independent recognition and generalization compared with six benchmark architectures.
Hurtado-Sánchez et al. (2026) [34]	CNN + Pose Estimation-Based Mobile Learning System	Mexican Sign Language (MSL)	Achieved **95.7% real-time sign recognition accuracy** for adaptive learning applications.
Mouti et al. (2026) [35]	Hybrid LSTM + 3D CNN Framework	IoT-Enabled Sign Language Recognition System	LSTM achieved **93.8% F1-score** and **86.67% accuracy**, while 3D CNN achieved **99.6% accuracy** with **65 ms latency**.
Algethami et al. (2025) [36]	Temporal Convolutional Network (TCN) and Enhanced RNN-BiLSTM	Arabic Sign Language (30 Common Sentences)	TCN achieved **99.5% accuracy**, outperforming enhanced RNN-BiLSTM (**99% accuracy**).
Alsharif et al. (2025) [37]	YOLOv11 + MediaPipe Hand Tracking	American Sign Language (ASL)	Achieved **98.2% mAP@0.5** for real-time ASL alphabet recognition.
Guan et al. (2025) [38]	MSKA-SLR (Multi-Stream Keypoint Attention Network)	Phoenix-2014, Phoenix-2014T, CSL-Daily	Achieved **state-of-the-art performance** on Phoenix-2014T sign language translation benchmark.
Lim et al. (2025) [39]	SpellRing (Acoustic Sensing + IMU + Deep Learning)	American Sign Language Fingerspelling	Achieved **82.45% Top-1 Accuracy**, **92.42% Top-5 Accuracy**, and **0.099 WER** in real-time evaluation.
Zhang et al. (2025) [40]	MediaPipe + MLP + PointNet	American Sign Language (Letters and Numbers)	Achieved **99.81% overall accuracy** with inference time of **less than 10 ms**.

Figure 4 shows the performance of recent deep learning based Sign Language Recognition (SLR) is compared to the performance of the proposed method. Zhang et al. (2025) [40] achieved the highest performance with 99.81% accuracy, followed closely by Mouti et al. (2026) [35] with 99.6% accuracy and Algethami et al. (2025) [36] with 99.5% accuracy. In addition, Garcia-Gil et al. (2026) [32] showed high accuracy and F1-score of 99%. In contrast, Lim et al. (2025) [39] achieved the lowest Top-1 accuracy (82.45%) because of the difficulties associated with real-time fingerspelling recognition. The overall trend of the graph reveals that the recognition accuracy of the modern architectures based on Transformer, CNN, and multimodal exceeds 95%, suggesting remarkable progress in real-time and intelligent SLR systems.

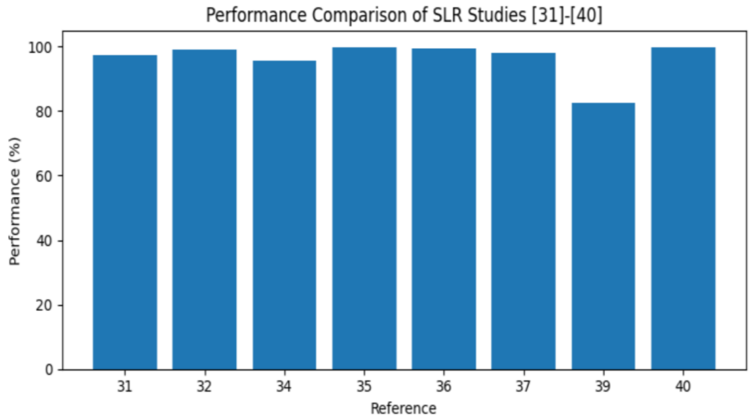

Figure 4: Benchmark Performance of Sign Language Recognition Models

2.3 Human–Computer Interaction Applications of Sign Language Recognition

With the rapid development of sensing technology, multimodal deep learning, and intelligent human–computer interaction, the performance of the Sign Language Recognition (SLR) system has been greatly increased in recent years. **Hu et al. (2026)** [41] presented a multimodal sign language recognition system with the combination of ion-conductive hydrogel-based digital gloves and a flexible sEMG armband equipped with a Bi-LSTM fusion model. With regard to classification accuracy of 24 gestures of Chinese sign languages, the proposed system obtained remarkable 99.65% classification accuracy with high mechanical flexibility and sensing reliability. In like manner, **Nuankaew et al. (2026)** [42] presented the self-revival iontronic neuromorphic devices that mimic the motion-cognition process of earthworms for motion-cognition systems. The tracking system it developed was very smart and was able to track movement of the human body with an accuracy of 98%, and after damage to the device and revival with a 96% accuracy, which demonstrates the system's strength for future application in rehabilitation and interactive robotics. **Wang et al. (2026)** [43] have designed a fully integrated stretchable and permeable electronic glove containing strain sensors and accelerometers, which allows for wireless gesture recognition. The system was able to interpret sign language accurately with a success rate of 98% and to control the robotic hand. In addition, **Neri et al. (2026)** [44] developed a markerless system for emotion recognition based on human body language in social XR environments, and obtained an average classification accuracy of 72.5% over the four emotional states, which shows the potential of XR skeleton-based human behavior analysis.

In recent studies, real-time gesture recognition and intelligent edge computing solutions in Human–Computer Interaction applications have been also scrutinized, e.g., by **Ślesicka et al. (2026)** [45] who proposed a framework for hand gesture recognition based on FMCW mmWave radar and using a lightweight CNN architecture with Continuous Wavelet Transform (CWT) preprocessing. The approach proposed in this paper has demonstrated an excellent accuracy of 99.87% and 82–84% generalisation accuracy for users that were not included in the training set, which is suitable for an IoT environment. Similarly, for Virtual Reality applications, **Benyamin et al. (2026)** [46] studied real-time facial emotion recognition, and reported that the YOLOv11n model had an accuracy in face detection of 100% while it had an average inference latency of 54ms, but Transformer-based models had a less satisfactory performance with an accuracy of less than 23%. Moreover, **Xing et al. (2025)** [47] introduced a 3D deep learning approach for non-contact hand rehabilitation based on laser sensor point cloud data. The experimental tests showed the average gesture recognition accuracy of the system of 88.72% which can help to develop intelligent rehabilitation systems for elderly and physically challenged users. The results highlight how more and more deep learning and sensing technologies are contributing to the improvement of gesture interpretation and adaptive interaction systems.

The possibilities of gesture recognition applications are still growing with the development of smart wearable devices and computer vision methods. **Zhang et al. (2025)** [48] proposed a smart glove with liquid-metals sensors that were used as sign language recognition sensors. When using a quadratic discriminant analysis (QDA) classifier the system obtained 100% accuracy in gesture recognition for ten sign language gestures when the sensor signals were combined with machine learning algorithms. Likewise, **Damdoo et al., (2025)** [49] presented a MediaPipe-based robotic hand control system that converted users' hand motions into robotic commands, using computer vision methods. The framework successfully replicated the gesture up to 95% for human–robot interaction. In addition, **Ahmed et al. (2025)** [50] proposed a hybrid gesture recognition system based on ResNet50 feature extraction, Tamura texture descriptors, and an Optimized Generalized Additive Model (GAM). Experimental results obtained on a set of American Sign Language (ASL) data showed that the overall classification rate of 96% surpassed several state-of-the-art methods. Overall, these works demonstrate the potential of multimodal sensing, wearable electronics, deep learning architectures, and intelligent interaction frameworks to improve the accuracy, robustness and applicability of contemporary sign language and gesture recognition systems to real-world scenarios.

Table 3: Recent Sign Language and Gesture Recognition Studies.

Author(s), Year & Ref.	**Method / Model**	**Dataset / Application**	**Key Results**
Hu et al. (2026) [41]	Bi-LSTM Multimodal Fusion with Hydrogel Sensors and sEMG Armband	Chinese Sign Language (24 Gestures)	Achieved **99.65% classification accuracy** using multimodal wearable sensing.
Nuankaew et al. (2026) [42]	Self-Revival Iontronic Neuromorphic Device	Motion-Cognition Human Tracking System	Achieved **98% tracking accuracy** and maintained **96% accuracy after self-revival**.
Wang et al. (2026) [43]	Stretchable Permeable Electronic Glove with Strain Sensors and Accelerometer	Sign Language Interpretation and Robotic Control	Achieved **98% sign language recognition accuracy** with wireless real-time operation.
Neri et al. (2026) [44]	Skeleton-Based Emotion Recognition Framework	Social XR Applications	Achieved **72.5% average emotion classification accuracy** across four emotional states.
Ślesicka et al. (2026) [45]	FMCW mmWave Radar + CWT + Lightweight CNN	Real-Time Hand Gesture Recognition	Achieved **99.87% accuracy** and **82–84% unseen-user generalization accuracy**.
Benyamin et al. (2026) [46]	YOLOv11n and Vision Transformer-Based FER	VR-Based Facial Emotion Recognition	YOLOv11n achieved **100% face detection accuracy** with **54 ms latency**.
Xing et al. (2025) [47]	3D Deep Learning with Laser Point Cloud Analysis	Intelligent Hand Rehabilitation System	Achieved **88.72% gesture recognition accuracy**.
Zhang et al. (2025) [48]	Liquid-Metal Smart Glove + Machine Learning (QDA)	Sign Language Gesture Recognition	Achieved **100% gesture recognition accuracy** for 10 sign language gestures.
Damdoo et al. (2025) [49]	MediaPipe-Based Robotic Hand Control System	Human–Robot Interaction	Achieved **95% gesture replication accuracy** using computer vision techniques.
Ahmed et al. (2025) [50]	ResNet50 + Tamura Texture Descriptor + Optimized GAM	American Sign Language (ASL) Recognition	Achieved **96% classification accuracy**, outperforming several state-of-the-art methods.

In Figure 5, the performance of recent Sign Language Recognition (SLR) studies are compared. The highest performance was achieved by Benyamin et al. (2026) [46] and Zhang et al. (2025) [48], both attaining 100% accuracy. Ślesicka et al. (2026) [45] closely followed with 99.87% accuracy, while Hu et al. (2026) [41] achieved 99.65% accuracy. Nuankaew et al. (2026) [42] and Wang et al. (2026) [43] achieved 98% accuracy, respectively, showing good recognition performance. In contrast, Neri et al. (2026) [44] recorded the lowest accuracy (72.5%). Overall, most studies demonstrated performance above 95%, and it is well established that the use of advanced deep learning models, wearable sensing technologies, and intelligent human–computer interaction frameworks greatly improves gesture and sign language recognition accuracy.

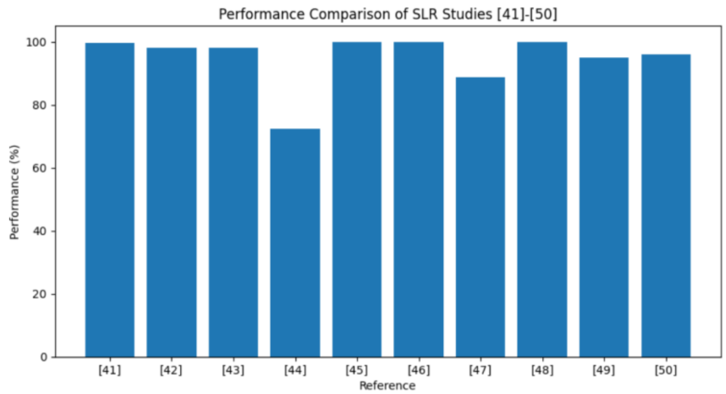

Figure 5: Comparative Analysis of SLR Models.

3. Discussion

The analyzed research papers have shown the outstanding improvement made in Sign Language Recognition (SLR) with the use of advanced deep learning architectures. Harmade, et al. [21] have demonstrated the advantage of CNN-based models for gesture recognition over traditional RNN based models, which illustrates the power of spatial feature extraction. In a similar vein, Toshpulatov et al. [22] demonstrated that the Transformer-based architectures like ConvNeXt and Swin Transformer performed better in cross-lingual recognition in several sign languages. Lamaakal et al. [26] reported one of the best performances by using ConvNeXt-Tiny and the Swin Transformer with knowledge distillation in the TinyMSLR framework, and Zhang et al. [30] showed the effectiveness of the wearable sensor-based recognition systems with very high classification accuracy. Overall, these studies suggest that deep learning hybrid models and multimodal approaches are emerging as the most popular approaches for gesture classification with high accuracy.

Recent studies also emphasize the need for multimodal and real-time SLR systems. Garcia-Gil et al. [32] proposed multimodal angular-geometry representations in conjunction with BiLSTM networks for ASL recognition with very good recognition accuracy, while Nasef et al. [33] proposed the use of Transformer layers with GRU network for the recognition of ASL in Arabic language with signer independence. Hybrid LSTM–3D CNN and Temporal Convolutional Network (TCN) architectures were found to achieve outstanding results by Mouti et al. [35] and Algethami et al. [36] respectively highlighting the significance of the temporal feature modeling. Moreover, Alsharif et al. [37] and Guan et al. [38] emphasized the advantages of attention mechanism, keypoint tracking, and light deployment framework to achieve fast recognition performance in real-time. The results indicate that using multiple modalities and applying sophisticated attention-based learning techniques significantly enhances the system's strength and performance.

From the Human–Computer Interaction point of view, Hu et al. [41] proved that the recognition accuracy for wearable sensing technologies is very close to 100% with support for real-time interaction, Wang et al. [43] proved that lightweight CNN models for radar can achieve near perfect recognition accuracy for IoT environments, and Zhang et al. [48] showed the potential of computer vision and hybrid deep learning for human–robot interaction and assistive communication, Damdoo et al. [49] and Ahmed et al. [50] showed that. While these studies have made great strides, there are still issues of signer variability, multilingual recognition, continuous sign language translation and real world implementation to address. Future work needs to be done on developing large scale multilingual datasets, explainable Transformer-based frameworks, and efficient edge-computing solutions to further enhance the scalability, reliability, and accessibility of deep learning-based SLR systems.

4. Conclusion

This review recognizes the great progress made in Sign Language Recognition (SLR) by applying deep learning (DL) methods and techniques to achieve high accuracy in gesture recognition and enhance the Human–Computer Interaction (HCI) field. The reviewed works demonstrate the great recognition accuracy that was obtained by CNNs, LSTMs and GRUs, Transformers, and hybrid deep learning architectures, which are efficient in modeling both spatial and temporal aspects of sign language gestures. Additionally, multimodal methods combining visual cues with wearable sensors, skeletal tracking, radar, and facial expression analysis have enhanced the robustness and the real-time performance in various settings. In addition, the review underscores the critical role of light weight and edge deployment aspects for practical applications, such as AT, smart devices, healthcare, education systems, and human–robot interaction. While there has been a great improvement, the problem of signer independence, multilingual recognition, continuous sign language translation, limited datasets and real world adaptability are still open research problems. Future work should focus on developing and scaling benchmark datasets, multimodal frameworks with the Transformer architecture, developing EXAI (Explainable AI), and developing efficient deployment strategies to enhance the accuracy, scalability, and usability of the system. From a broad perspective, deep learning-based SLR systems are paradigm shift technologies which can assist the deaf and hard-of-hearing individuals to communicate more effectively, to become more accessible and to make human–computer interaction more accessible and intelligent.

References

1. Sami, Mohsin, Rabia Tehseen, Uzma Omer, Muhammad Farrukh Khan, Shahan Yamin Siddiqui, Nabeel Sabir Khan, and Danish Ali Khan. "Systematic Literature Review on Computational Models Used For Sign Language Recognition." Journal of Computing & Biomedical Informatics 10, no. 02 (2026).
2. Marcolino, Milena Soriano, Lucca Fagundes Ramos de Oliveira, Lucas Rocha Valle, Luiza Marinho Motta Santa Rosa, Gabriela Teodora de Souza Sanches, Natalia Sales Santos, Michelle Ralil da Costa et al. "Innovations in deaf health care communication: systematic review of sign language recognition systems." Journal of Medical Internet Research 28 (2026): e70417.
3. IJIRCST, I. "Sign Language Recognition Using Deep Learning: A Systematic Review of Models and Approaches." International Journal of Innovative Research in Computer Science and Technology (IJIRCST) (2026).
4. Osman, Mohd Nizam, Khairul Anwar Sedek, Nur Zaid Md Zain, Muhamad Amin Naim Ab Karim, and Mushahadah Maghribi. "Hearing assistive technology: Sign language translation application for hearing-impaired communication." In Charting the Sustainable Future of ASEAN in Science and Technology: Proceedings from the 3rd International Conference on the Future of ASEAN (ICoFA) 2019-Volume 2, pp. 1-11. Singapore: Springer Singapore, 2020.
5. Keskin, Fesih, and Gültekin Işık. "Deep insights into Turkish sign language recognition, generation, and translation: a comprehensive systematic review." Universal Access in the Information Society 25, no. 3 (2026): 79.
6. Aswin, R., Bhavya S. Kumar, S. Boomika, Thomas Jacob, Varsha Varghese, and Linda Sebastian. "Review on Techniques of Gesture Navigation Control." (2026).
7. Cui, Cui, Mohd Shahrizal Sunar, and Goh Eg Su. "Deep vision-based real-time hand gesture recognition: a review." PeerJ Computer Science 11 (2025): e2921.
8. Khan, Asma, Seyong Jin, Geon-Hee Lee, Gul E. Arzu, L. Minh Dang, Tan N. Nguyen, Woong Choi, and Hyeonjoon Moon. "Deep learning approaches for continuous sign language recognition: A comprehensive review." IEEE Access 13 (2025): 55524-55544.
9. Balaha, Mostafa Magdy, Sara El-Kady, Hossam Magdy Balaha, Mohamed Salama, Eslam Emad, Muhammed Hassan, and Mahmoud M. Saafan. "A vision-based deep learning approach for independent-users Arabic sign language interpretation." Multimedia Tools and Applications 82, no. 5 (2023): 6807-6826.
10. Khan, Asma, Seyong Jin, Geon-Hee Lee, Gul E. Arzu, L. Minh Dang, Tan N. Nguyen, Woong Choi, and Hyeonjoon Moon. "Deep learning approaches for continuous sign language recognition: A comprehensive review." IEEE Access 13 (2025): 55524-55544.
11. Foteinos, Konstantinos, Manousos Linardakis, Panagiotis Radoglou-Grammatikis, Vasileios Argyriou, Panagiotis Sarigiannidis, Iraklis Varlamis, and Georgios Th Papadopoulos. "Visual hand gesture recognition with deep learning: A comprehensive review of methods, datasets, challenges and future research directions." arXiv preprint arXiv:2507.04465 (2025).
12. Linardakis, Manousos, Iraklis Varlamis, and Georgios Th Papadopoulos. "Survey on hand gesture recognition from visual input." IEEE Access (2025).
13. Bansal, Sandhya, Akansha Tyagi, and Rajeev Kumar Goel. "A Deep Survey of Intelligent Systems for Sign Language Recognition System." In Role of Nature-Inspired Algorithms in Real-life Problems, pp. 1-22. Singapore: Springer Nature Singapore, 2025.
14. Joel, Laing Luyoh, Samsul Setumin, Emilia Noorsal, Siti Juliana Abu Bakar, and Mokh Sholihul Hadi. "A systematic review of sign language classification system for real-time application." Journal of Electrical and Electronic Systems Research (JEESR) 27, no. 1 (2025): 1-13.

15. Violet, I. M., and R. Leena Sri. "A comprehensive survey on recent advances and challenges in sign language recognition systems." Discover Artificial Intelligence (2025).
16. Tamilvanan, Prabhavathy, Vinodh Kumar Elumalai, and Balaji Elumalai. "A Comprehensive Survey of Machine Learning Algorithms in Hand Gesture Recognition for Myoelectric Prosthetic Control: Current Trends, Challenges and Future Directions." Archives of Computational Methods in Engineering (2025): 1-30.
17. Marcolino, Milena Soriano, Lucca Fagundes Ramos de Oliveira, Lucas Rocha Valle, Luiza Marinho Motta Santa Rosa, Gabriela Teodora de Souza Sanches, Natalia Sales Santos, Michelle Ralil da Costa et al. "Innovations in deaf health care communication: systematic review of sign language recognition systems." Journal of Medical Internet Research 28 (2026): e70417.
18. Carla, Luisinho, João Silva, Maria Pinto, Ricardo Costa, Sofia Almeida, and Hugo Martins. "A Survey of Deep Learning Approaches for Hand Tracking and Gesture Recognition." Authorea Preprints (2025).
19. Sinha, Bhupendra. "A Comprehensive Survey on Hand Gesture Recognition Using Deep Learning for Electronic Device Interaction." Authorea Preprints (2025).
20. Singh, Satender, and Priyanka Mathur. "A Review on Bridging Humans and Machines: Insights into HGR Systems with SLR." In International Conference on AI Systems and Sustainable Technologies, pp. 305-313. Singapore: Springer Nature Singapore, 2025.
21. Harmade, Dani, Afif Fathin, and Nur Jannah Nai'mah Zainal. "Comparison of Convolutional Neural Network and Recurrent Neural Network Algorithms for Indonesian Sign Language Recognition." Public Research Journal of Engineering, Data Technology and Computer Science 3, no. 2 (2026): 89-99.
22. Toshpulatov, Mukhiddin, Wookey Lee, Jaesung Jun, and Suan Lee. "Deep learning pathways for automatic sign language processing." Pattern Recognition 164 (2025): 111475.
23. Mineo, Raffaele, Amelia Sorrenti, Gaia Caligiore, Federica Proietto Salanitri, Giovanni Bellitto, Senya Polikovsky, Sabina Fontana, Egidio Ragonese, Concetto Spampinato, and Simone Palazzo. "A benchmark for radar-based Italian sign language recognition using frequency-domain range-time maps." In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 10556-10566. 2026.
24. Alanazi, Ghadeer, and Abir Benabid. "Arabic Sign Language Recognition using Multimodal Approach." arXiv preprint arXiv:2601.17041 (2026).
25. Jing, Siyuan, and Gaorong Yan. "Dual-View Sign Language Recognition via Front-View Guided Feature Fusion for Automatic Sign Language Training." Information 17, no. 2 (2026): 158.
26. Lamaakal, Ismail, Chaymae Yahyati, Yassine Maleh, Khalid El Makkaoui, and Ibrahim Ouahbi. "An explainable hybrid CNN–transformer model for sign language recognition on edge devices using adaptive fusion and knowledge distillation." Scientific Reports (2026).
27. Singh, Nitin Kumar, Arie Rachmad Syulistyo, Yuichiro Tanaka, and Hakaru Tamukoh. "Sign language recognition using parallel bidirectional reservoir computing." Nonlinear Theory and Its Applications, IEICE 17, no. 1 (2026): 79-92.
28. Luna-Jiménez, Cristina, Lennart Eing, Annalena Bea Aicher, Fabrizio Nunnari, and Elisabeth André. "Lightweight Transformers for Isolated Sign Language Recognition." In Proceedings of the 27th International Conference on Multimodal Interaction, pp. 155-163. 2025.
29. Perea-Trigo, Marina, Enrique J. López-Ortiz, Luis M. Soria-Morillo, Juan A. Álvarez-García, and J. J. Vegas-Olmos. "Impact of face swapping and data augmentation on sign language recognition." Universal Access in the Information Society 24, no. 2 (2025): 1283-1294.
30. Yan, Cangshu, Saihua Jiang, Yuchun Wang, Junrui Deng, Xinpeng Wang, Zidian Chen, Tianle Chen, Huamao Huang, and Hao Wu. "A wearable sign language translation device utilizing silicone-hydrogel hybrid triboelectric sensor arrays and machine learning." Nano Energy 133 (2025): 110425.
31. Nuankaew, Pratya, Parin Khamthep, Patdanai Jaitem, Kuljira S. Nuankaew, Kaewpanya S. Nuankaew, and Wongpanya S. Nuankaew. "Multimodal AI Framework for Sign Language Recognition and Medical Informatics in Hearing-Impaired Patients." Journal of Applied Data Sciences 7, no. 1 (2026): 580-593.
32. Garcia-Gil, Gerardo, Gabriela del Carmen López-Armas, and Yahir Emmanuel Ramirez-Pulido. "Multimodal Minimal-Angular-Geometry Representation for Real-Time Dynamic Mexican Sign Language Recognition." Technologies 14, no. 1 (2026): 48.
33. Nasef, Mohammed M., Aya El-Sayed, and Eman M. AbouNassar. "ArSL-TGRU: A Hybrid Model for Signer-Independent Arabic Sign Language Recognition from Videos Based on Transformer and GRU." International Journal of Computational Intelligence Systems (2026).
34. Hurtado-Sánchez, Carlos, Ricardo Rosales Cisneros, José Ricardo Cárdenas-Valdez, Andrés Calvillo-Téllez, and Everardo Inzunza-Gonzalez. "Adaptive Assistive Technologies for Learning Mexican Sign Language: Design of a Mobile Application with Computer Vision and Personalized Educational Interaction." Future Internet 18, no. 1 (2026): 61.
35. Mouti, Samar, Hani Al Chalabi, Mohammed Abushohada, Samer Rihawi, and Sulafa Abdalla. "Comparative Evaluation of LSTM and 3D CNN Models in a Hybrid System for IoT-Enabled Sign-to-Text Translation in Deaf Communities." In Informatics, vol. 13, no. 2, p. 27. Multidisciplinary Digital Publishing Institute, 2026.
36. Algethami, Nahlah, Raghad Farhud, Manal Alghamdi, Huda Almutairi, Maha Sorani, and Noura Aleisa. "Continuous Arabic sign language recognition models." Sensors 25, no. 9 (2025): 2916.
37. Alsharif, Bader, Easa Alalwany, Ali Ibrahim, Imad Mahgoub, and Mohammad Ilyas. "Real-time american sign language interpretation using deep learning and keypoint tracking." Sensors 25, no. 7 (2025): 2138.

38. Guan, Mo, Yan Wang, Guangkun Ma, Jiarui Liu, and Mingzu Sun. "Mska: Multi-stream keypoint attention network for sign language recognition and translation." Pattern Recognition 165 (2025): 111602.
39. Lim, Hyunchul, Nam Anh Dang, Dylan Lee, Tianhong Catherine Yu, Jane Lu, Franklin Mingzhe Li, Yiqi Jin et al. "Spellring: Recognizing continuous fingerspelling in american sign language using a ring." In Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems, pp. 1-17. 2025.
40. Zhang, Yan, and Gunther Notni. "3D geometric features based real-time American sign language recognition using PointNet and MLP with MediaPipe hand skeleton detection." Measurement: Sensors 38 (2025): 101697.
41. Hu, Quan, Longya Xiao, Peiqi Zhang, Qiurui Zhang, Guangsen Liu, Xinglin Qin, Zhuhui Yin, Xian Li, Yuling Wang, and Hongjie Jiang. "3D-Printed Ion-Conductive Hydrogels with Tunable Mechanical–Electrical Properties for Multimodal Sign Language Recognition." Advanced Science (2026): e20586.
42. Nuankaew, Pratya, Parin Khamthep, Patdanai Jaitem, Kuljira S. Nuankaew, Kaewpanya S. Nuankaew, and Wongpanya S. Nuankaew. "Multimodal AI Framework for Sign Language Recognition and Medical Informatics in Hearing-Impaired Patients." Journal of Applied Data Sciences 7, no. 1 (2026): 580-593.
43. Wang, Wenqiang, Zebang Luo, Xingge Yu, Xiaojia Yin, Li Xiang, and Anlian Pan. "A Highly Permeable and Three-Dimensional Integrated Electronic System for Wearable Human–Robot Interaction." Nano-Micro Letters 18, no. 1 (2026): 128.
44. Neri, Michael, Sara Baldoni, Marco Carli, and Federica Battisti. "Markerless emotion recognition from full-body movements for Social XR." Signal Processing: Image Communication 143 (2026): 117489.
45. Ślesicka, Anna, and Adam Kawalec. "Real-Time Hand Gesture Recognition for IoT Devices Using FMCW mmWave Radar and Continuous Wavelet Transform." Electronics 15, no. 2 (2026): 250.
46. Benyamin, Yarin. "The Latency Wall: Benchmarking Off-the-Shelf Emotion Recognition for Real-Time Virtual Avatars." arXiv preprint arXiv:2601.15914 (2026).
47. Xing, Zhizhong, Zhijun Meng, Gengfeng Zheng, Guolan Ma, Lin Yang, Xiaojun Guo, Li Tan, Yuanqiu Jiang, and Huidong Wu. "Intelligent rehabilitation in an aging population: empowering human-machine interaction for hand function rehabilitation through 3D deep learning and point cloud." Frontiers in Computational Neuroscience 19 (2025): 1543643.
48. Zhang, Jing, Xiaoyang Zou, Ziqi Li, Colin Pak Yu Chan, and King Wai Chiu Lai. "Liquid-metal-based multichannel strain sensor for sign language gesture classification using machine learning." ACS Applied Materials & Interfaces 17, no. 4 (2025): 6957-6968.
49. Damdoo, Rina, and Praveen Kumar. "An integrative survey on Indian sign language recognition and translation." IET Image Processing 19, no. 1 (2025): e70000.
50. Ahmed, Ismail Taha, Wisam Hazim Gwad, Baraa Tareq Hammad, and Entisar Alkayal. "Enhancing hand gesture image recognition by integrating various feature groups." Technologies 13, no. 4 (2025): 164.