**JOURNAL PAPERS**

**1) Real-Time Sign Language Recognition and Translation Using Deep Learning Techniques**

Tazyeen Fathima et al. (2024) address the challenge of accurate sign language recognition and translation in real-time applications, focusing on improving accessibility and communication for deaf and hearing-impaired individuals. Despite advances in deep learning, Sign Language Recognition (SLR) systems still face challenges in terms of accuracy and visual quality, while Sign Language Translation (SLT) systems are hampered by limited language comprehension datasets. The paper presents an innovative approach for sign language recognition and conversion to text using a custom dataset containing 15 different classes, each containing 70-75 images. The proposed solution uses the YOLOv5 architecture, a state-of-the-art Convolutional Neural Network (CNN), combined with data preprocessing using Roboflow for dataset management. The methodology includes custom dataset creation, data preprocessing and labeling, YOLOv5 model architecture selection for real-time processing capabilities, model training with iterative procedures, fine-tuning and optimization of hyperparameters, and comprehensive evaluation and deployment with user interface implementation. The authors achieve impressive mAP values (average accuracy) of 92% to 99% for each of the 15 classes, demonstrating the model's effectiveness in promoting inclusive and accessible communication platforms. However, the approach has limitations including validation limited to the custom dataset, potential scalability issues to larger sign vocabulary sets, dependency on consistent lighting and background conditions for optimal performance, and lack of evaluation on diverse demographic groups. This research is particularly relevant to automated sign language systems as it provides insights into effective real-time processing using state-of-the-art object detection architectures, dataset creation methodologies, and deployment strategies for accessibility applications.

**Link:** <https://irjaeh.com/index.php/journal/article/view/33/26>

**2) A Comprehensive Deep Learning Based System for Real Time Sign Language Recognition and Translation Using Raspberry Pi**

Abini M.A et al. (2024) tackle the challenge of developing a cost-effective, portable, and accessible sign language translation system for deaf and hearing-impaired individuals. The main problems addressed include the significant communication gap between deaf individuals and the general public, the time-consuming nature of manual sign language interpretation, and the need for expensive specialized hardware in existing systems. This work presents a novel method for translating sign language into spoken language using a Raspberry Pi 3 and the MobileNet-V2 deep learning model. The methodology involves using a standard camera to capture images of hand gestures, processing them using the MobileNet-V2 architecture for classification, implementing transfer learning techniques for model training, and integrating text-to-speech software for output conversion. The system was trained on a large dataset of sign language movements and achieved an accuracy of 99.52% on the validation set. Key contributions include the introduction of a low-cost, portable solution using Raspberry Pi 3 hardware, demonstration of effective transfer learning application for sign language recognition, achievement of high accuracy rates with minimal computational requirements, and successful integration of multiple components (camera, processing, text-to-speech). The approach has limitations such as restriction to specific gesture sets in the training data, dependency on consistent lighting conditions for optimal performance, limited evaluation on diverse user groups and environments, and potential challenges in scaling to larger vocabulary sets. This research is highly relevant to accessible technology development as it provides a practical framework for implementing cost-effective sign language recognition systems using edge computing devices and demonstrates the feasibility of deploying deep learning solutions in resource-constrained environments.

**Link:** <https://doi.org/10.14445/22312803/ijctt-v72i12p102>

**3) Real-Time Vision-Based Indian Sign Language Translation Using Deep Learning Techniques**

Subham Pandey et al. (2025) address the challenge of developing a comprehensive real-time sign language translation system specifically for Indian Sign Language (ISL), focusing on both isolated and continuous gesture recognition. The main problems tackled include the limited availability of robust ISL recognition systems, challenges in real-time processing of complex gesture sequences, difficulties in handling spatiotemporal gesture modeling, and the need for non-intrusive translation systems that don't require external sensors. The paper proposes a vision-based approach using state-of-the-art deep learning architectures including CNN (Convolutional Neural Networks), LSTM (Long Short-Term Memory) networks, and Transformer-based encoder-decoder models. The methodology incorporates data preprocessing techniques such as Dynamic Time Warping (DTW) for gesture sequence augmentation and normalization, utilization of custom ISL and public ASL datasets, implementation of multimodal fusion techniques, and development of pose-language alignment mechanisms. The system achieves remarkable performance with a Transformer-based model outperforming other approaches, reaching a BLEU score of 0.74 and classification accuracy of 96.1%. Additional contributions include the development of a desktop application enabling real-time ISL-to-English translation at 18 FPS, comprehensive evaluation using multiple metrics (precision, recall, F1-score, BLEU, ROUGE, CER, WER), validation through ablation studies demonstrating benefits of multimodal fusion, and advancement of non-intrusive sign language translation technology. However, the system has limitations including dependency on quality of input datasets, computational requirements for real-time processing, potential challenges in handling diverse signing styles and speeds, and limited evaluation across different demographic groups. This research is particularly significant for the deaf and hard-of-hearing (DHH) community as it demonstrates a robust, scalable approach to sign language translation and provides a foundation for developing more inclusive communication technologies.

**Link:** <https://doi.org/10.55524/ijircst.2025.13.3.6>

**4) Real-Time Sign Language Recognition and Translation using MediaPipe and LSTM-Based Deep Learning**

Ravikiran V. (2025) addresses the communication challenges faced by people with speech and hearing impairments, focusing on developing an accessible and cost-effective sign language recognition system. The main problems include the difficulty in communication between speech-impaired individuals and others, the need for expensive hardware in existing solutions, and limitations in real-time gesture recognition accuracy. The paper presents a Sign Language Gesture Recognition and Translation system designed to help people with speech impairments interact easily with others by recognizing five important sign language gestures: Yes, No, I Love You, Thank You, Hello, and OK, converting them to clear text messages in real-time. The methodology utilizes a Long Short-Term Memory (LSTM) neural network combined with Python, OpenCV, and MediaPipe for accurate and detailed hand tracking. The system effectively processes video frames from a webcam using computer vision techniques, detects hand movements accurately using MediaPipe for 21 hand landmark identification, employs LSTM for temporal pattern recognition in gesture sequences, and requires only hand movements in front of a standard webcam making it cost-effective and user-friendly. The approach achieves impressive results with 100% training accuracy, 98-99% validation accuracy, 96.6% real-time testing accuracy, and inference time ranging between 0.35-0.45 seconds per gesture. Key contributions include the development of a lightweight, accessible system requiring minimal hardware, successful integration of MediaPipe with LSTM for temporal gesture analysis, achievement of high accuracy with real-time processing capabilities, and demonstration of practical applicability for daily communication assistance. Limitations include restriction to only six basic gestures, dependency on consistent webcam input quality, limited evaluation across diverse user groups and environmental conditions, and challenges in scaling to larger sign language vocabularies. This research is particularly relevant for assistive technology development as it provides a practical framework for implementing cost-effective sign language recognition systems and demonstrates the potential for AI-assisted communication tools to enhance accessibility and independence for speech-impaired individuals.

**Link:** <https://doi.org/10.5120/ijca2025925415>

**5) Smart Glove-Based Sign Language Recognition and Translation Systems**

Multiple researchers (2022-2025) have developed smart glove-based systems for real-time sign language detection and translation, addressing the communication gap between sign language users and non-signers. The primary problems include the need for non-intrusive, wearable solutions for continuous sign language recognition, challenges in accurately capturing finger movements and hand orientations, difficulties in real-time processing of sensor data, and the requirement for portable, battery-operated devices. These research projects present smart glove prototypes equipped with multiple sensors including five flex sensors for finger movement detection, MPU-6050 or similar gyroscope sensors for hand orientation and motion tracking, Arduino microprocessors for data collection and processing, and wireless communication modules (WiFi/Bluetooth) for data transmission. The methodology involves sensor data fusion techniques, Recurrent Neural Network (RNN) with Long Short-Term Memory (LSTM) for gesture pattern recognition, integration of text-to-speech engines for audio output, and focus on American Sign Language (ASL) gestures corresponding to letters A-Z and numbers 0-9. Various implementations have achieved different performance levels, with some systems reaching up to 99% accuracy for specific gesture sets, successful real-time translation capabilities with minimal latency, effective wireless communication between glove and processing devices, and practical demonstration of wearable assistive technology. Key contributions include the development of portable, wearable sign language translation devices, successful integration of multiple sensor types for comprehensive gesture capture, demonstration of effective machine learning approaches for sensor data classification, and practical solutions for everyday communication assistance. However, these systems face limitations such as restriction to predefined gesture sets, dependency on proper sensor calibration and maintenance, potential battery life constraints for extended use, limited durability of wearable components, and challenges in handling complex sentence-level sign language communication. This research area is particularly significant for advancing wearable assistive technologies and provides practical solutions for enhancing communication accessibility for the deaf and hard-of-hearing community.

**Links:**

* <https://doi.org/10.46632/jdaai/4/1/28>
* <https://github.com/WaniaKhance/Smart-Glove-Sign-Language-Translator>
* <https://repository.londonmet.ac.uk/9015/>

**6) Deep Learning-Based American Sign Language Recognition Systems**

Various researchers (2022-2024) have developed advanced deep learning systems for American Sign Language (ASL) recognition, focusing on improving accuracy and real-time performance. The main challenges addressed include achieving high accuracy in distinguishing between similar hand gesture letters, developing efficient models for real-time deployment on various hardware platforms, handling dataset limitations and ensuring generalization across diverse users, and creating robust systems that work under different lighting and background conditions. These studies utilize advanced deep learning architectures including ResNet-50, EfficientNet, AlexNet, ConvNext, VGG-16, MobileNetV2, and custom CNN models. The methodologies involve transfer learning techniques with pre-trained models, comprehensive data augmentation strategies, five-fold cross-validation for performance measurement, and optimization using Adam optimizers with carefully tuned hyperparameters. Remarkable results have been achieved with ResNet-50 reaching 99.98% accuracy, EfficientNet achieving 99.95% accuracy, AlexNet and ConvNext both attaining 99.51% accuracy, and various MobileNet implementations reaching 94-98% accuracy for embedded applications. Significant contributions include establishment of new performance benchmarks for ASL alphabet recognition, demonstration of effective transfer learning applications in sign language domains, development of lightweight models suitable for mobile and embedded deployment, and comprehensive comparative analysis of different deep learning architectures. However, these systems face limitations including restriction to static sign alphabet recognition (letters and numbers), limited dataset diversity affecting model generalization, computational requirements for high-accuracy models, and challenges in transitioning from controlled to real-world environments. This research is crucial for advancing computer vision applications in assistive technology and provides foundational work for developing more comprehensive sign language communication systems.

**Links:**

* <https://pmc.ncbi.nlm.nih.gov/articles/PMC10535774/>
* <https://ijisae.org/index.php/IJISAE/article/view/3920>
* <https://doi.org/10.22214/ijraset.2024.60728>

**7) CNN and LSTM Hybrid Approaches for Continuous Sign Language Recognition**

Recent research (2024-2025) has focused on developing hybrid CNN-LSTM architectures for continuous sign language recognition, addressing the challenges of temporal sequence modeling in sign language interpretation. The primary problems include accurately modeling temporal dependencies in continuous signing, handling variable-length gesture sequences, achieving real-time processing while maintaining high accuracy, and developing systems that can handle both isolated and continuous sign language recognition. These studies implement hybrid architectures combining Convolutional Neural Networks (CNNs) for spatial feature extraction with Long Short-Term Memory (LSTM) networks for temporal sequence modeling. The methodology includes preprocessing techniques using MediaPipe Holistic for comprehensive landmark extraction (hands, face, body), dynamic time warping (DTW) for sequence normalization, attention mechanisms for focusing on relevant gesture components, and end-to-end training approaches for integrated recognition systems. Performance results demonstrate the effectiveness of these hybrid approaches with systems achieving 88-95% accuracy for continuous sign recognition, successful real-time processing capabilities at 15-30 FPS, effective handling of variable-length sequences, and robust performance across different signers and environments. Key contributions include advancement of continuous sign language recognition technology, successful integration of spatial and temporal modeling approaches, development of comprehensive evaluation frameworks using multiple metrics, and demonstration of practical applicability in real-world scenarios. Limitations include computational complexity requiring high-performance hardware, challenges in handling rapid signing speeds and transitions between gestures, dependency on quality landmark detection for optimal performance, and limited evaluation on diverse sign language vocabularies and dialects. This research is particularly important for developing more natural and comprehensive sign language communication systems that can handle the complexity of real-world sign language usage.

**Links:**

* <https://doi.org/10.47001/irjiet/2025.iccis-202523>
* <https://www.ijraset.com/research-paper/empowering-communication-with-cnn-sign-language-recognition>

**8) Multi-Modal and Sensor Fusion Approaches for Enhanced Sign Language Recognition**

Advanced research (2024-2025) has explored multi-modal approaches combining various sensing technologies and deep learning architectures for improved sign language recognition accuracy. The main challenges addressed include integrating multiple data modalities (visual, sensor, depth) for comprehensive gesture understanding, developing robust systems that work across different environmental conditions, achieving high accuracy while maintaining computational efficiency, and creating systems that can adapt to individual signing styles and preferences. These studies implement sophisticated sensor fusion techniques combining RGB cameras with depth sensors, flex sensors with IMU sensors, MediaPipe holistic features with custom CNN architectures, and multi-stream neural networks for processing different data types simultaneously. The methodology involves advanced preprocessing pipelines for multi-modal data alignment, attention mechanisms for cross-modal feature fusion, ensemble learning approaches for improved robustness, and comprehensive training strategies using large-scale datasets. Results demonstrate significant improvements with multi-modal systems achieving 97-99% accuracy, successful deployment on edge devices including Raspberry Pi and Jetson Nano, effective real-time processing with multi-modal inputs, and enhanced robustness to environmental variations. Major contributions include advancement of multi-modal learning approaches for sign language recognition, successful demonstration of edge computing applications for assistive technology, development of comprehensive evaluation frameworks considering multiple performance metrics, and practical validation in real-world deployment scenarios. However, these systems face limitations including increased system complexity and cost due to multiple sensors, higher computational requirements for multi-modal processing, challenges in maintaining synchronization across different data streams, and potential reliability issues with multiple hardware components. This research represents the cutting edge of sign language recognition technology and provides important insights for developing next-generation assistive communication systems.

**Links:**

* <https://doi.org/10.3390/electronics13163233>
* <https://doi.org/10.1007/s11042-018-6565-5>
* <https://journal.esrgroups.org/jes/article/view/6153>

9) **Computer vision-based hybrid efficient convolution for isolated dynamic sign language recognition**

**Prothoma Khan Chowdhury, Kabiratun Ummi Oyshe, Muhammad Aminur Rahaman, Tanoy Debnath, Anichur Rahman, Neeraj Kumar. Neural Computing and Applications, 2024.**

**Problem Addressed**

This study tackles the challenge of Isolated Dynamic Sign Language Recognition (IDSLR), which involves recognizing sign gestures comprising sequences of image frames, each potentially showing multiple linguistic features, in environments with cluttered backgrounds and varying illumination. IDSLR is important for accessibility and inclusion, enabling deaf and hearing-impaired people to communicate fully in different aspects of life, such as social relationships, education, and work. The main limitations in the field are dataset diversity (backgrounds, skin tone, age, lighting conditions), robustness to environmental variation, and misclassification due to ambiguity and similar gestures.

**Methodology**

The authors propose a novel "Hybrid Efficient Convolution" (HEC) model, which ensembles a pre-trained EfficientNet-B3 with custom dense layers (256 units), batch normalization, and dropout, followed by an output layer optimized for gesture classification. The HEC architecture leverages EfficientNet-B3 features and enhances adaptability using a custom fully connected layer for task-specific learning. Key regularization methods—weight decay and dropout—are employed to prevent overfitting.  
A new dataset, BdSL\_OPA\_23\_GESTURES, was curated for the study, comprising **6000 video clips** of 100 isolated dynamic Bangla sign language words, with each word captured by 20 different people and across five different background types (solid, ambiguous-indoor, ambiguous-outdoor, cluttered-indoor, cluttered-outdoor), incorporating a wide range of lighting conditions and signers' ages (13–50 years). Data augmentation (random flips, rotations, shifts, zoom) was used to further enhance model robustness.  
The system preprocesses each video by extracting frames and applies the HEC for feature extraction and classification. Architecture specifics include depth-wise separable and point-wise convolutions, and a linear bottleneck to manage computational complexity and capture diverse visual features.

**Contributions**

* **Novel hybrid architecture** (HEC) integrating EfficientNet-B3 and a new dense layer, effectively capturing rich spatiotemporal gesture features even in noisy environments.
* **Creation of the BdSL\_OPA\_23\_GESTURES dataset:** 6000 annotated video clips from 20 signers in five complex backgrounds with varying lighting conditions, age, and gender, providing significant diversity lacking in prior datasets.
* **Superior performance:** On the new dataset, HEC achieves **93.17% accuracy**, outperforming established pre-trained CNNs (AlexNet, ResNet, InceptionV3, ViT, etc. all ≤84.2%) under the same testing conditions.
* **Efficiency:** HEC demonstrates faster inference (20 ms per gesture) and lower parameter count (10.9M) than common alternatives, enabling feasible real-time deployment.
* **Public resource:** Both the model and dataset are made openly available for the research community, supporting further advances in sign language technology.

**Limitations**

* **Misclassification of complex gestures:** A few specific signs, especially longer or visually similar ones, show higher misclassification rates.
* **Vocabulary and dataset scope:** The dataset, while diverse, includes only 100 words, and only dynamic isolated gestures—not continuous sign sentences.
* **Background diversity:** Although five types are included, backgrounds may still not capture the full variety encountered in real-world deployments.
* **Further generalizability:** The focus is on Bangla Sign Language; while the approach is extensible, tested results apply only to this context.

**Research Focus**

The study is notable for focusing on practical deployment challenges: cluttered/ambiguous backgrounds, illumination variation, and diverse signers. This addresses key limitations of many earlier SLR datasets and systems. The combination of robust hybrid CNN architectures with realistic, open-access data resources pushes the field toward more inclusive and accurate automated sign language recognition.

**Link**

<https://doi.org/10.1007/s00521-024-10258-3>

**10)Next-Gen Dynamic Hand Gesture Recognition: MediaPipe, Inception-v3 and LSTM-Based Enhanced Deep Learning Model**

Yaseen, Oh-Jin Kwon, Jaeho Kim, Sonain Jamil, Jinhee Lee, Faiz Ullah (2024)  
*Electronics* 13(16), 3233

**Problem Addressed**

Dynamic hand gesture recognition (HGR) poses challenges due to the temporal nature and variable length of gesture sequences, high-dimensional spatiotemporal data, and efficiency constraints especially in real-time applications like drone control and AR/VR. Existing models struggle with efficiently capturing temporal dependencies while maintaining computational feasibility. Furthermore, varying backgrounds and lighting conditions reduce robustness.

**Methodology**

The authors propose a novel triple-layer hybrid architecture combining:

* MediaPipe (v0.10.14) for hand landmark detection and ROI extraction, significantly reducing input dimensionality and background noise;
* Inception-v3 CNN for extracting 2D spatial features from each ROI-cropped frame;
* An LSTM network to capture temporal dynamics for sequence classification.

This model processes variable-length sequences by reducing 3D feature maps into 1D vectors, which lowers computational cost and maintains temporal context. Data augmentation and transfer learning were employed to improve robustness and efficiency.

**Contributions**

* Integration of MediaPipe for efficient and lightweight ROI hand region extraction, enabling reduced resolution inputs and faster processing.
* Novel dimensionality reduction from 3D spatial-temporal data to low-dimensional vectors for manageable training and inference.
* Demonstrated enhanced accuracy, achieving over 89.7% while maintaining real-time performance and reducing computational load compared to baseline models.
* Outperformed state-of-the-art methods on established datasets with complex backgrounds and lighting variations.

**Limitations**

* The model currently works on a limited subset of gestures from the Depth\_Camera\_Dataset (6 classes) and may require extension to larger gesture vocabularies for real-world utility.
* Performance is subject to quality of input video and landmark detection accuracy by MediaPipe.
* Potential issues handling extreme background clutter or occlusion remain.

**Relevance**

This research advances dynamic gesture recognition, vital for accessible interfaces requiring non-contact communication (e.g., sign language recognition, human-computer interaction). The lightweight and efficient architecture proposed suits embedded or mobile platforms and use in varied environmental conditions.

**Link**

<https://doi.org/10.3390/electronics13163233>

Here is the summarized literature review entries for the two IJRASET papers (dco.pdf and doc2.pdf) in the format consistent with your sample literature review document:

**Real Time Sign Language Recognition Using Deep Learning**

K. Mahimanvitha and Dr. M. Arathi (2023)  
*International Journal for Research in Applied Science & Engineering Technology* Volume 11, Issue IX, September 2023

**Problem Addressed:**  
Communication barriers faced by speech and hearing impaired individuals due to limited sign language understanding by others. Existing systems focused mainly on static signs or alphabets, lacking robust real-time dynamic gesture recognition.

**Methodology:**  
Uses MediaPipe Holistic to extract comprehensive keypoints (33 pose, 468 face, 21 per hand) from live video frames. A deep Long Short-Term Memory (LSTM) neural network is trained on these temporal features to classify dynamic British Sign Language gestures. Real-time gesture recognition is implemented via OpenCV with a Streamlit-based GUI for displaying recognized text.

**Contributions:**

* Dynamic recognition of British Sign Language actions (hello, thanks, I love you) achieving ~96.5% accuracy;
* Real-time implementation with webcam input;
* Integration of MediaPipe landmark detection with temporal deep learning for robust classification;
* User-friendly interface using Streamlit to display recognized signs as text in real-time.

**Limitations:**

* Limited gesture vocabulary (three phrases);
* Small dataset size of 30 video sequences per gesture;
* Challenges distinguishing visually similar gestures;
* Model accuracy depends on lighting and pose quality.

**Link:**  
<https://doi.org/10.22214/ijraset.2023.55621>

**Deep Learning - Based Real Time Sign Language Translator Using YOLO**

Noorul Moufica M et al. (2025)  
*International Journal for Research in Applied Science & Engineering Technology* Volume 13, Issue V, May 2025

**Problem Addressed:**  
Bridging communication gap between hearing-impaired individuals and the hearing community by providing a real-time sign language translator supporting alphabets, numbers, and sentence-level gestures.

**Methodology:**  
A modular real-time system combining YOLOv8 object detection for alphabets (A–Z) and numbers (0–9), plus sentence gesture recognition (32 common phrases). Speech-to-text integration is included for voice input transcription. The system is deployed on a web platform using Flask, Python, OpenCV, MediaPipe, and Ultralytics YOLOv8, supporting translation into Tamil and Hindi with multilingual output and responsive UI controls.

**Contributions:**

* Real-time alphabets and sentence-level gesture recognition with ~95% accuracy;
* Speech-to-text translation integrated with gesture recognition;
* Multilingual output support (English, Tamil, Hindi);
* User-friendly web UI with live detection, screenshot, and recording features;
* Low latency, suitable for practical assistive communication scenarios.

**Limitations:**

* Currently supports static and predefined sentence gestures, not continuous dynamic gestures;
* Recognition accuracy affected by lighting, hand variations, and environmental factors;
* Speech recognition performance drops in noisy environments;
* Real-time constraints on low-spec hardware without GPU acceleration.

**Link:**  
<https://doi.org/10.22214/ijraset.2025.70838>

**Comparative Analysis Table**

| **Sl No.** | **Paper/System** | **Problem Addressed** | **Methodology** | **Contributions** | **Limitations** |
| --- | --- | --- | --- | --- | --- |
| **1** | Real-Time Sign Language Recognition Using YOLOv5 | Accurate real-time sign language recognition with limited datasets and computational challenges | YOLOv5 architecture, custom dataset (15 classes, 70-75 images each), Roboflow preprocessing | mAP values 92-99% across 15 classes, real-time processing capability, accessible technology framework | Limited to 15 gesture classes, dependency on lighting conditions, lack of diverse demographic evaluation |
| **2** | Raspberry Pi MobileNet-V2 System | Cost-effective, portable sign language translation with expensive hardware requirements | MobileNet-V2, Raspberry Pi 3, transfer learning, camera integration, text-to-speech conversion | 99.52% validation accuracy, low-cost portable solution, successful edge computing implementation | Specific gesture set limitations, lighting dependency, limited scalability evaluation |
| **3** | Vision-Based ISL Translation | Comprehensive ISL recognition for isolated and continuous gestures without external sensors | CNN-LSTM-Transformer hybrid, DTW preprocessing, multimodal fusion, pose-language alignment | 96.1% accuracy, 18 FPS real-time processing, BLEU score 0.74, non-intrusive approach | Dataset quality dependency, computational requirements, limited demographic evaluation |
| **4** | MediaPipe LSTM Recognition | Accessible communication for speech-impaired individuals with minimal hardware requirements | LSTM neural networks, MediaPipe hand tracking, OpenCV processing, 21 hand landmarks | 100% training accuracy, 96.6% real-time accuracy, 0.35-0.45s inference time, cost-effective solution | Limited to 6 gestures, webcam quality dependency, restricted vocabulary scope |
| **5** | Smart Glove Systems | Wearable solution for continuous sign language recognition and translation | Flex sensors, MPU-6050, Arduino, RNN-LSTM, wireless communication, sensor fusion | Up to 99% accuracy, real-time translation, portable wearable design, practical deployment | Predefined gesture limitations, sensor calibration requirements, battery life constraints |
| **6** | Deep Learning ASL Recognition | High-accuracy static sign alphabet recognition with transfer learning | ResNet-50, EfficientNet, AlexNet, transfer learning, data augmentation, cross-validation | 99.98% accuracy (ResNet-50), benchmark performance, comprehensive architecture comparison | Static gesture limitation, dataset diversity challenges, computational requirements |
| **7** | CNN-LSTM Hybrid Systems | Continuous sign language recognition with temporal sequence modeling | Hybrid CNN-LSTM, MediaPipe Holistic, attention mechanisms, DTW normalization | 88-95% continuous recognition, real-time processing, temporal dependency modeling | Computational complexity, rapid signing challenges, landmark detection dependency |
| **8** | Multi-Modal Sensor Fusion | Enhanced recognition through multiple sensing modalities and comprehensive gesture understanding | Multi-sensor integration, ensemble learning, attention mechanisms, edge computing deployment | 97-99% multi-modal accuracy, edge device deployment, environmental robustness | System complexity, higher costs, synchronization challenges, hardware reliability |
| **9** | Computer vision-based hybrid efficient convolution for isolated dynamic sign language recognition | Isolated dynamic gesture recognition in varied backgrounds/lighting; diversity and robustness in Bangla sign language classification. | EfficientNet-B3 + custom FC layer, augmentation, 6000 videos in five background types, open dataset | Outperformed all baselines with 93.17% accuracy, new diverse dataset, fast and efficient model, open code and data. | Errors in specific gestures, 100-word scope, isolated not continuous, Bangla-only validation. |
| **10** | Next-Gen Dynamic Hand Gesture Recognition: MediaPipe, Inception-v3 and LSTM-Based Enhanced Deep Learning Model (Yaseen et al., 2024) | Challenges in recognizing dynamic hand gestures due to temporal complexity, variable sequence lengths, and computational costs | A triple-layer hybrid model using MediaPipe for ROI extraction, Inception-v3 CNN for spatial features, and LSTM for temporal sequence classification | Efficient dimensionality reduction, leveraging MediaPipe to improve feature quality and reduce computational cost, achieving ~90% accuracy on Depth\_Camera dataset | Limited gesture vocabulary (6 classes), sensitivity to occlusion, tested on a relatively small dataset |
|  | Real Time Sign Language Recognition Using Deep Learning (K. Mahimanvitha & Dr. M. Arathi, 2023) | Communication barrier for speech/hearing impaired due to limited understanding of sign language; lack of dynamic gesture recognition | MediaPipe Holistic for pose, face, hand keypoint extraction; LSTM model trained on temporal keypoints; real-time recognition via OpenCV and Streamlit UI | High accuracy (~96.5%) on dynamic British Sign Language phrases (hello, thanks, I love you); real-time webcam processing with user interface | Limited to three gestures; small dataset; difficulties distinguishing visually similar gestures; environmental sensitivity |
|  | Deep Learning-Based Real Time Sign Language Translator Using YOLO (Noorul Moufica M et al., 2025) | Communication gap for deaf/mute by recognizing alphabets, numbers, sentence gestures, integrating speech-to-text and translation | YOLOv8 object detection for static alphabets (A-Z) and numbers (0-9); sentence gesture recognition (32 phrases); speech recognition via Google API; Flask web UI | Real-time recognition with ~95% accuracy; multilingual translation (English/Tamil/Hindi); supports video recording, screenshots, voice input; practical user interface | Supports only static/predefined sentence gestures; impact of lighting and noise on accuracy; real-time processing constraints on low-spec hardware |

**Here are the summarized tables for the two IJRASET papers on real-time sign language recognition and translation, formatted consistently with your sample literature review:**

| **Sl No** | **Paper / Tool** | **Problem Addressed** | **Methodology** | **Contributions** | **Limitations** |
| --- | --- | --- | --- | --- | --- |
| **21** |  |  |  |  |  |
| **22** |  |  |  |  |  |

**If desired, these can be directly integrated into your existing literature review table. Let me know if you want me to help with that.**

1. [**https://ppl-ai-file-upload.s3.amazonaws.com/web/direct-files/attachments/16311898/61514319-51c7-4785-9cd9-c715fb39493c/sample-literature-review-doc.pdf**](https://ppl-ai-file-upload.s3.amazonaws.com/web/direct-files/attachments/16311898/61514319-51c7-4785-9cd9-c715fb39493c/sample-literature-review-doc.pdf)
2. [**https://ppl-ai-file-upload.s3.amazonaws.com/web/direct-files/attachments/16311898/90ae4671-6205-4057-ad69-220a93bf6ae7/doc.pdf**](https://ppl-ai-file-upload.s3.amazonaws.com/web/direct-files/attachments/16311898/90ae4671-6205-4057-ad69-220a93bf6ae7/doc.pdf)
3. [**https://ppl-ai-file-upload.s3.amazonaws.com/web/direct-files/attachments/16311898/7c323a1d-4e82-4125-9269-381f969c6ed7/doc2.pdf**](https://ppl-ai-file-upload.s3.amazonaws.com/web/direct-files/attachments/16311898/7c323a1d-4e82-4125-9269-381f969c6ed7/doc2.pdf)

**Research Gaps and Future Directions**

Based on the comprehensive analysis of these research papers, several key insights emerge that shape the future of sign language recognition technology. The reviewed literature demonstrates significant progress in accuracy and real-time processing capabilities, with modern systems achieving recognition rates exceeding 95% across various approaches. However, several critical gaps remain unaddressed, including the limitation of most systems to static or isolated gesture recognition rather than continuous sign language communication, the lack of comprehensive evaluation across diverse demographic groups and environmental conditions, insufficient focus on sentence-level and contextual sign language understanding, and limited integration of facial expressions and body language components crucial for complete sign language comprehension.

Future research directions should focus on developing more comprehensive systems that can handle continuous, contextual sign language communication, implementing robust evaluation frameworks that consider diverse users and real-world conditions, advancing multi-modal approaches that integrate visual, sensor, and contextual information for enhanced accuracy, and creating scalable solutions that can adapt to different sign languages and individual signing styles. The ultimate goal remains the development of truly inclusive communication technologies that can bridge the gap between sign language users and the broader community, enabling seamless and natural interaction across all communication modalities.

1. <https://ppl-ai-file-upload.s3.amazonaws.com/web/direct-files/attachments/16311898/61514319-51c7-4785-9cd9-c715fb39493c/sample-literature-review-doc.pdf>
2. <https://irjaeh.com/index.php/journal/article/view/33/26>
3. <https://www.ijcttjournal.org/archives/ijctt-v72i12p102>
4. <https://doi.org/10.55524/ijircst.2025.13.3.6>
5. <https://www.ijcaonline.org/archives/volume187/number25/ravikiran-2025-ijca-925415.pdf>
6. <https://www.ijraset.com/best-journal/sign-language-gesture-detection-using-cnn>
7. <https://repository.londonmet.ac.uk/9015/>
8. <https://pmc.ncbi.nlm.nih.gov/articles/PMC10535774/>
9. <https://www.ijraset.com/research-paper/empowering-communication-with-cnn-sign-language-recognition>
10. <https://newsroom.ucla.edu/releases/glove-translates-sign-language-to-speech>
11. <https://ijisae.org/index.php/IJISAE/article/view/3920>
12. <https://www.ijraset.com/best-journal/sign-language-recognition-with-deep-learning>
13. <https://github.com/WaniaKhance/Smart-Glove-Sign-Language-Translator>
14. <https://www.techscience.com/CMES/v139n3/55626/html>
15. <https://www.ijraset.com/best-journal/bridging-the-communication-gap-a-comprehensive-study-on-sign-language-detection-system>
16. <https://journal.esrgroups.org/jes/article/view/6153>
17. <https://www.sciencedirect.com/science/article/abs/pii/S0045790624002672>
18. <https://www.ijraset.com/research-paper/sign-language-recognition-using-cnn>
19. [https://www.brightsignglove.com](https://www.brightsignglove.com/)
20. <https://www.sciencedirect.com/science/article/pii/S2215016124003534>
21. <https://www.ijraset.com/research-paper/sign-language-detection-using-cnn-model>
22. <https://sist.sathyabama.ac.in/sist_naac/documents/1.3.4/1822-b.e-cse-batchno-68.pdf>
23. <https://www.nature.com/articles/s41598-024-76174-7>
24. <https://www.ijraset.com/research-paper/indian-sign-language-recognition-using-cnn-inception-v3-a-review>
25. <https://projecthub.arduino.cc/ayooluwa98/sensor-glove-for-sign-language-translation-c14ae7>