International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 10 | Oct - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

# Sign Language Detection Using Deep Learning

Ashwini SP[1],Sneha D[2],Sumalatha T[2],Tanya D mulani[2],Varsha T S[2]
[1]Assistant Prof. ,Department of Computer Science and Engineering ,JNNCE ,Shimoga ,India
[2]UG Students Department of Computer Science and Engineering, JNNCE ,Shimoga ,India

***Abstract*—"Sign Language Detection Using Deep Learning" will determine the next step which is to develop a machine learning model capable of asserting sign, a medium of communication used by the deaf and hard of hearing. The suggested system would be in real time system where live sign gestures would be processed by the use of image processing. There would be the use of the classifiers to distinguish different signs and the output would be showing text. The aim of the system is the enhancement of the current system on this issue in the aspects of reaction speed and accuracy, through efficient algorithms, high-quality data sets and superior sensors. We expect that in our project we evolve a cognitive system that is responsive, robust, to such extent that it be applied in day-to-day tasks and operations by hearing or speech disabled individuals.**

**Keywords**—Sign Language Detection, Deep Learning, Image Processing, Real-Time Gesture Recognition, Assistive Technology, Deaf and Hard of Hearing Communication.

## 1. INTRODUCTION

In the age of technological innovation and development, real-time sign language detection deep learning has become a revolutionary approach in improving communication to deaf/ hard of hearing people. This is an innovation that can utilize computer vision and sophisticated deep learning tools to extract meaningful data from hands, fingers, and faces. Such systems translate visual signals into either speech or text and enable communication to be inclusive in a number of disciplines, including education, health services, and government services. The models of Deep Learning, such as Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN), are trained on an extensive gesture dataset in order to recognize and understand signs with high probability. The overall process of detecting the gesture quite typically includes as its steps the capturing of gesture data, preprocessing of the data, feature extraction, recognition of gestures, and real-time output production. The effectiveness of such systems is largely dependent upon the richness of the dataset and the optimality of the architecture of the model sign language detection technology has already played a huge role in a number of practices. In education it enables children with the hearing impairment to participate better in digital learning platforms. Within the healthcare sector it helps to improve communication amongst the medical staff and patients that use sign language, and can thus ensure improved delivery of medical care and elimination of misunderstandings. Our proposed system will expand these functions through precision and time response with using high-quality data-sets, superior algorithms, and superior hardware sensors. A future goal is to develop an intelligent and dependable mechanism that can be incorporated in other applications to assist persons who have speech or hearing difficulties.

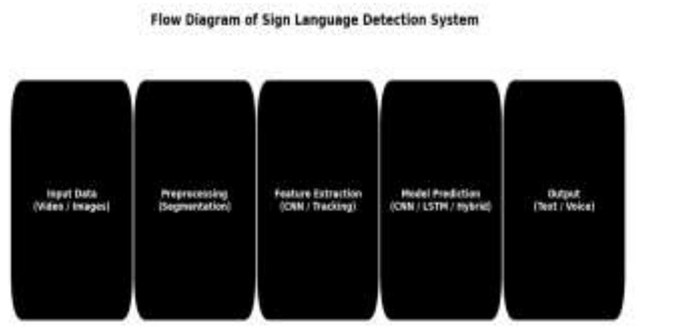


Fig 1:Flow Diagram of Sign Language Detection system

## 2. LITERATURESURVEY

### A. *Real-Time Sign Language Detection Using Deep Learning*

A real-time sign language recognition system developed by Pathak et.al (2022) [1] combines the computer vision circuit with the deep learning technique. They have applied SSD Mobile Net V2 and CNNs in their method in order to practice good gesture classification. The system will be structured into three main phases that include capturing the input using webcams, recognition of gestures using CNNs, and interpretation of the gestures into relevant sign language. The scientists created a collection of more than 2000 labeled gestures in diverse lighting and background settings related to 70-80 per cent accuracy. The system tackles such complications like variability in lighting, coincidence in gestures, and saturation through manipulative presetting methods and the use of supplementary inputs of data such as facial expressions and posture. The paper also investigates the implementation of mobile, and cloud-based platforms where they can be accessed by many people. Due to the necessity of region-specific sign language support, the authors believe it is best to develop it through collaboration and make it scalable and inclusive in terms of detection solutions

### B. *Real-Time Sign Language Detection Using CNNs*

The study of Maheshwari Chitampalli et.al (2023) [2] offered to develop a real-time sign language detection model supported by CNN with the idea in mind to improve communication between the persons with hearing loss. Video information is fed into the system which detects the hand gestures and converts them to text or voice with amazing 95 percent accuracy. Workflow includes a few steps with the assistance of preprocessing methods (such as noise reduction and background filters) to capture gesture detection, segment and classify. The system has use in the educational facility, online internet based communication, and mobile technology whereby the users experience a flexible and competent method of communicating. Ethical design is also highlighted as suggested by the researchers including the end-user response and facilitating accessibility. The future potential improvements of sign recognition are continuous sign recognition, lightweight models of wearable devices, and an enhancement to support different environmental conditions. Transfer learning and domain adaptation is employed by the study to enable the system to become more sturdy and flexible in real life applications

### *C. Real-Time Sign Language Recognition Using Deep Learning and Hand Tracking*

To detect a sign language, Monisha H. M. et.al (2023) [3] proposed a deep-learning algorithm, which is explored along with hand-tracking technologies. They are based on frameworks like Mediapipe and TensorFlow that they use to get high accuracy of gesture detection. The methods that preprocessing pipeline involves are gesture normalization, background removal, and data augmentation to improve performance under various conditions. To addressed shortcomings including gesture obstructing (gesture occlusion), rapid hand movements, and inconsistent hand shapes, the authors propose to apply depth-sensing cameras and also diverse datasets. Augmented Reality (AR) is also incorporated in order to give the interactive feedback and the real-time learning that can be effective in learning needs. Further functionality will comprise personalized learning processes through user practices and video conferencing-social network interface. The authors give importance to the social good of these technologies as non-obstructing communication barriers and making society more inclusive. Further improvements in the system efficiency and the scope of applicability are welcome through development of system research.

### *D. Indian Sign Language Recognition Using TensorFlow and Transfer Learning*

Srivastava et.al (2021) [4] have suggested to recognize an ISL in real time by applying objects detection API of TensorFlow with a transfer learning and setting SSD Mobile Net V2. The model uses webcam to capture hand gestures, converts them into text in English with an accuracy rate of 85.45 percent even though it is trained on a small data of 650 labeled images.It possesses phases such as data acquisition, carriage rejection, gesture division, and normalization procedures followed by collective extraction and classification with CNNs. The authors emphasized the possibility of expansion of the system to other regional and international sign languages with the help of using cross-lingual transfer learning. The uses are in the field of education, healthcare, legal systems and the public. The paper focuses on inclusivity, scale expansion and demands multi-actor collaboration of researchers and policymakers to achieve a broad scope of adoption. Future improvements can add bigger datasets, sentence level recognition, and enabling with advanced sensors (to have better tracking and real-time performance)..

### *E. Review of ISL Recognition Techniques Using Machine Learning*

A more comprehensive review of approaches to Indian Sign Language (ISL) recognition was proposed by Dessai and Naik (2022) [5] and focused not only on vision-only systems but also on sensor-based approaches. In the paper, algorithm efficiency is tested in the use of functions, like CNNs, SVM, and KNN in tasks, including in feature extractions, pre-processing, and the classification of items. Whereas most systems achieve over 90 percent accuracy, real-time performance is an ongoing challenge because of the environmental presence such as lighting, background clutter, and hand variation. The authors support data quality improvement, accelerating its processing, and the affordability of the tools, such as webcams, smartphones, etc. Another recommendation related to the paper is to combine the modalities of hand movements with facial expressions and body posture to provide a better rate of accuracy and context perceptions. CNNs are characterized by high feature extraction quality but need further intensive preprocessing in order to deal with the variability of the environment. It also proposes the hybrid version of CNN and SVM or KNN to enhance the accuracy of the study. The authors emphasize that to enhance performance further, it is critical to create datasets that would be diverse and take into account regional variations of the gesture. They additionally propose the transfer learning and domain adaptation to increase the system robustness under real life conditions. The paper concludes by suggesting the need to carry out more research to create larger, precise and inclusive ISL identification systems..

### *F. Real-Time Sign Language Recognition Using CNN and Transfer Learning*

Serai et al. (2017) [6] designed a system that recognizes sign language in real-time using a combination of deep learning and image processing so that more hearing and speech-impaired individuals can communicate. The system applies CNNs using the Sequential API but refers to Massey University dataset to train and adopts preprocessing approaches such as resizing, background subtraction in improving the accuracy in the dynamic environment. To augment the performance with comparatively small data, the authors train the models through transfer learning, such as GoogleNet and AlexNet. This method eliminates training time and improves the level of recognition. The system only works with single-hand gestures but will work on two-hand sign language and Indian Sign Language (ISL) in future. Future work involves compiling ISL-specific data sets, dealing with regional variations of gestures, and the possibility of hard-ware enhancement to achieve a better performance in real-time operation. The research has shown that deep learning and preprocessing can be effectively combined in order to facilitate reactive and effective gesture recognition systems, suitable to be used in the receptive, realistic world.

### *G. Real-Time Sign Language Detection Using LSTM and Action Recognition*

Sreyasi Dutta et.al (2023) [7] proposed a real time sign language recognition system which utilized the actions recognition techniques with the help of LSTM models. The system recognizes gesture sequences by capture via video citation, featuring noise reductions and normalization of the sequence of gestures, motion feature extraction using optical flow and utilis LSTM networks to complete their identification. LSTM plays the shining role in the learning of the sequential aspect of gestures, which helps it to

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 10 | Oct - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

directly know the actual recognition of the dynamic sign language. The real-time performance of the model was very high, hence applicable to assistive tools. The feature extraction and preprocessing of data increases accuracy of the models, as it progressively increases the quality of data and motion tracking. The paper raises issues like diversity of the data sets, computation performance, and the ability to adapt to new surroundings. It emphasizes the necessity of high quality datasets including regional differences and being optimized to be used in mobile and low-power devices. The paper finds that LSTM-based systems are highly promising inclusion technologies and that future enhancements must focus on collecting enlarged data sets, optimizing the system, and increasing its real-life applicability.

***H Survey on Sign Language Recognition Systems***

Ashok K. Sahoo et al. (2014) [8] provide a comprehensive survey of sign language recognition technology in order to optimize communications between the hearing impaired. There are notable main processes captured in the study namely, signature acquisition, preprocessing, feature extraction, and gesture classification as applied in both static and dynamic gestures. The technologies relating to hand tracking, facial expression recognition and gesture analysis are mentioned to discuss their contribution to the proper interpretation of signs. One of the main constraints identified is a deficiency of large-scale gesture corpus that limits vocabulary and generalization of the systems. The authors stress the necessity of a system that would generally be flexible to real-world environments and be able to respond to the continuous, natural gestures, not necessarily in controlled space. Possibility of combination of neural networks and Hidden Markov Models to enhance the recognition rate and establish the representation of time in gestures is also noted.

## Result :

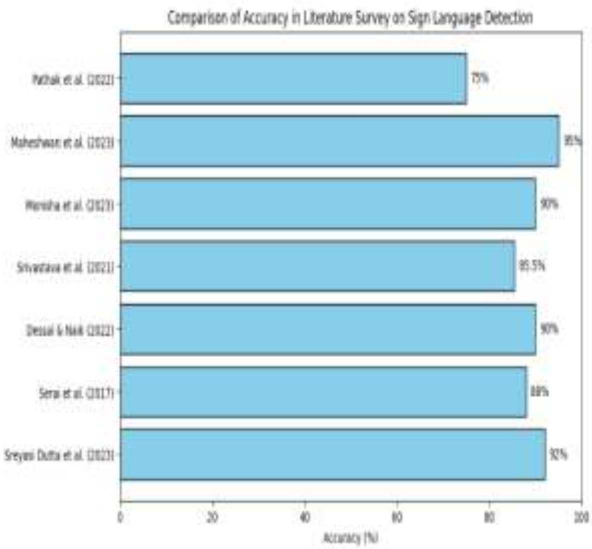


Fig 2:Comparision of Accuracy in Literature Survey

The horizontal bar chart in Fig. 1 presents a comparison of accuracy Levels reported in different research works on sign language Recognition the performance values range roughly from75%To 95% showing noticeable variations depending on the model and methodology applied. Pathak et al.2022 of reported a moderate accuracy about 75% where as in the Maheshwari et al.2023 achieved the highest performance At 95% using a CNN based framework. Studies by the of Monisha et.al 2023 and Dessai and Naik 2022 also reached Strong results, exceeding the 90% mark. This trend show that more recent approaches are pushing recommendation accuracy higher . Overall, the chart reflects progressing sign language detection system as deep learning hybrid technique continue to evolve.

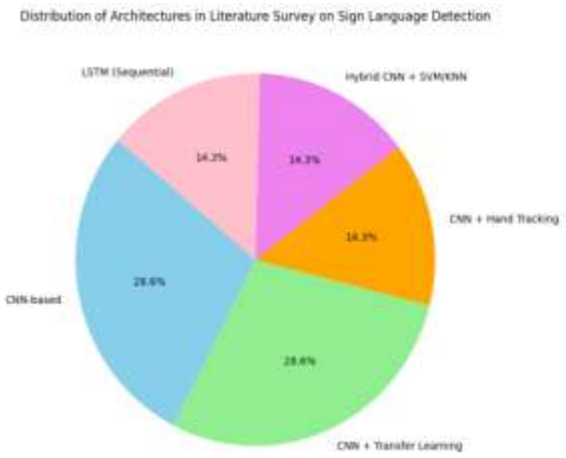


Fig 3: Distribution of Architectures in Literature Survey

The pie chart in Fig. 2 shows how different architectures have been used across research studies in sign language detection. Most works make use of CNN models, either directly or with transfer learning, which together cover the largest portion of the chart. A few studies explored hybrid methods like CNN combined with SVM/KNN, as well as CNN supported by hand-tracking techniques. One paper applied an LSTM model, emphasizing the role of sequence learning for gesture recognition. From this distribution, it is clear that CNN-based methods dominate the field, while other architectures are emerging as complementary approaches

| Reference/Author | Method/Approach | Key Features/Techniques | Findings/Inference |
|---|---|---|---|
| Pathak et al. (2022) | The system uses CNN and Mobile Net V2 to detect hand gestures from webcam input and match them to sign language symbols. | It includes a custom gesture dataset, preprocessing for lighting issues, and mobile/cloud support for wider use. | The model achieved 70–80% accuracy and can work in different conditions, with future improvements suggested for better accuracy and language coverage. |
| MaheshwariChitampalli et al. (2023) | The system captures hand gestures through video and uses CNN models to convert them into text or speech with 95% accuracy. | It applies preprocessing like noise reduction and background removal, uses a diverse dataset, and supports mobile and wearable integration. | The model performs well in varied conditions, with future goals focused on continuous gesture recognition and inclusive, user-driven design. |
| Monisha H. M. et al. (2023) | Combined Media pipe for hand tracking and TensorFlow-based deep learning for real-time sign language recognition. | Gesture normalization, background subtraction, data augmentation; proposed use of AR for feedback and skill training. | Achieved high accuracy; highlighted limitations due to occlusion and hand shape diversity; recommended depth sensors and adaptive learning for personalization. |
| S. Srivastava et al. (2021) | Used webcam-based real-time ISL recognition with TensorFlow's Object Detection API and SSD Mobile Net V2 via transfer learning | Background subtraction, CNN architecture fine-tuning, and gesture segmentation from 650 annotated images. | Achieved 85.45% accuracy; potential for multilingual sign language support; scalability proposed through cross-lingual transfer learning. |
| Dessai and Naik (2022) | Comparative review of sensor- and vision-based ISL recognition systems using traditional and deep learning classifiers. | Use of CNN, SVM, and KNN; emphasis on feature extraction, multimodal input (hand + face), and hybrid modeling. | Systems achieved >90% accuracy in ideal conditions; real-time performance remains limited; recommends large, inclusive datasets and affordable webcam-based solutions. |
| Serai et al. (2017) | Developed a CNN-based system using transfer learning (GoogleNet, AlexNet) for gesture classification with real-time output. | Image preprocessing (resizing, flipping, background subtraction), Keras API for model development, convex hull detection. | Demonstrated fast and accurate gesture recognition; proposed future support for ISL and dual-hand gestures for broader coverage. |
| Sreyasi Dutta et al. (2023) | Real-time sign detection using LSTM neural networks trained on motion features extracted from video using optical flow. | Temporal modeling of dynamic gestures, preprocessing for noise reduction and normalization, LSTM for sequence learning. | Showed strong real-time performance with high accuracy; emphasized dataset diversity and real-world adaptability for future improvements. |
| Ashok K. Sahoo et al. (2023) | Designed a multi-headed CNN that processes raw image data and hand landmarks in parallel for ASL gesture class. | Landmark-image fusion, data augmentation, dynamic learning rates, low-cost architecture using the "Finger Spelling A" dataset. | Achieved 98.98% accuracy; reduced dependency on expensive sensors; robust against cluttered backgrounds, suited for practical use. |
| Ahmed A. F. Youssef (2014) | Gesture recognition using neural, HMM. | Gesture acquisition, preprocessing, and classification; incorporation of hand and face tracking for semantic interpretation. | Stressed the need for large-vocabulary, real-world adaptable systems; highlighted hybrid models as a path forward for inclusivity. |

**TABLE I: Summary of Sign Language Detection** :

Despite performing effectively under controlled conditions, the system still encounters certain limitations. These include challenges such as sensitivity to variations in lighting, the complexity of different gesture patterns, and the scarcity of large and diverse datasets. Looking ahead, future improvements could focus on supporting dynamic gestures, enriching the dataset with more variety, and optimizing the model for use on mobile or wearable devices. Overall, this study establishes a solid groundwork for advancing assistive technologies aimed at promoting inclusivity and enhancing accessibility for individuals with hearing impairments.

## 3 .RECOMMENDATIONS AND FUTURE DIRECTION

To further enhance the performance and real-world applicability of the sign language detection system, several recommendations can be made. Expanding the dataset to include a broader variety of gestures, different sign languages, and diverse backgrounds will improve the model's robustness and generalization. Incorporating dynamic gesture recognition using models like LSTM or Transformers can enable the system to understand continuous sign sequences rather than static signs. Optimization for mobile devices and integration with wearables such as smartwatches or AR glasses will make the system more accessible and practical for daily use. Adding features like speech output alongside text can offer a more complete communication bridge. Engaging with the deaf community for user testing and feedback will ensure the system is user-friendly and addresses real communication needs. Future work could explore integrating the system into virtual and augmented reality platforms for educational and training purposes. Enhancing noise resistance and real-time processing under various lighting conditions will also boost performance. Incorporating emotion or facial expression recognition may add depth to the communication. Developing support for regional dialects and gestures can make the system more inclusive. Overall, continued research, user-centric design, and collaboration with accessibility experts can drive this technology toward becoming a powerful tool for social inclusion and assistive communication.

## 4. CONCLUSION

This project presents a practical deep learning–driven solution for real-time sign language recognition, designed to reduce the communication barrier between hearing-impaired individuals and the general population. By combining computer vision techniques with neural network models, the system is able to detect hand gestures and convert them into readable text, offering valuable support in sectors such as education and healthcare. The approach demonstrates promising accuracy and responsiveness under controlled conditions. Still, it faces certain challenges, including sensitivity to lighting changes, the complexity of gesture variations, and the limited availability of large, diverse datasets. Potential future enhancements may involve extending the model to handle dynamic gestures, expanding dataset diversity, and adapting the system for mobile or wearable platforms. In essence, this work provides a strong foundation for advancing assistive technologies that foster inclusivity and improve accessibility for people with hearing disabilities

## Reference

[1] A. Pathak, A. Kumar, P. Priyam, P. Gupta, and G. Chugh, “Real-Time Sign Language Detection,” *International Journal for Modern Trends in Science and Technology*, vol. 8, no. 01, pp. 32–37, 2022. doi: 10.46501/IJMTST0801006.

[2] M. Chitampalli, D. Takalkar, G. Pillai, P. Gaykar, S. Khubchandani, “Real-Time Sign Language Detection,” *International Research Journal of Modernization in Engineering Technology and Science*, vol. 5, no. 04, pp. 2983–2986, 2023. doi: 10.56726/IRJMETS36648.

[3] M. H. M. Monisha, B. S. Manish, R. R. Iyer, S. J. J. Siddarth, “Sign Language Detection and Classification Using Hand Tracking and Deep Learning in Real-Time,” *International Research Journal of Engineering and Technology (IRJET)*, vol. 10, no. 11, pp. 875–881, 2023. doi: 10.56726/IRJET2381.

[4] S. Srivastava, A. Gangwar, R. Mishra, and S. Singh, "Sign Language Recognition System using TensorFlow Object Detection API," Preprint, International Conference on Advanced Network Technologies and Intelligent Computing (ANTIC-2021), 2021, doi: 10.1007/978-3-030-96040-7_48.

[5] Dessai, S., & Naik, S. (2022). Literature Review on Indian Sign Language Recognition System. International Research Journal of Engineering and Technology (IRJET), Volume 9, Issue 7. Available at: www.irjet.net.

[6] Serai, D., Dokare, I., Salian, S., Ganorkar, P., & Suresh, A. (2017). Proposed System for Sign Language Recognition. 2017 International Conference on Computation of Power, Energy, Information and Communication (ICCPEIC). IEEE.

[7] Sreyasi Dutta, Adrija Bose, Sneha Dutta, Kunal Roy. "Sign Language Detection Using Action Recognition with Python." International Journal of Engineering Applied Sciences and Technology, Vol. 8, Issue 01, ISSN No. 2455-2143, Pages 61-67, Published Online May 2023.

[8] Refat Khan Pathan, Munmun Biswas, Suraiya Yasmin, Mayeen Uddin Khandaker, Mohammad Salman, Ahmed A. F. Youssef. "Sign Language Recognition Using the Fusion of Image and Hand Landmarks Through Multi-Headed Convolutional Neural Network." Scientific Reports, Vol. 13, Article 16975, 2023. DOI: 10.1038/s41598-023-43852-x.

[9] Ashok K. Sahoo, Gouri Sankar Mishra, Kiran Kumar Ravulakollu. "Sign Language Recognition: State of the Art." ARPN Journal of Engineering and Applied Sciences, Vol. 9, No. 2, February 2014.