# A Survey on Recognizing Sign Language Using Deep Learning

**Shegar Sneha Ramdas[1], Lagad Atul Ganesh[2], Modhave Samiksha Prabhakar[3], Dumbre Sakshi Rajendra[4]**
[1, 2, 3, 4] Samarth group of Institutions College of Engineering, Belhe, Pune

***Abstract***- *Effective communication hinges on the exchange of information, ideas, and emotions, necessitating a shared language. For individuals with speech impairments, sign language serves as a vital means of communication, but a lack of understanding can pose significant challenges when interacting with those without impairments. To address this, a machine learning-based model can be developed to recognize various sign language motions and convert them into a universal language, facilitating interactions with speech-impaired individuals. While existing Sign Language Recognition systems have limitations in recognizing multiple languages, this article introduces a real-time Sign Language Recognition system that employs transfer learning with TensorFlow and utilizes a webcam to detect and interpret sign language gesture.*

***Keywords***- Deep Learning, Gesture Recognition, English Sign Language, Sign Language Recognition (SLR), Computer Vision

## I. INTRODUCTION

Sign language serves as the primary mode of communication for over 70 million individuals who are deaf or hard of hearing, according to the World Federation of the Deaf [11]. The usage of sign language varies across regions and is influenced by factors such as the prevalence of deafness and hearing loss, cultural attitudes toward sign language, and the availability of resources and services related to sign language. Despite its crucial role, individuals who use sign language encounter substantial challenges in many communities. In numerous countries, deaf and hard-of-hearing individuals face limited access to education and the issue of language deprivation, as they often don't have the opportunity to develop their first language or receive adequate education in their preferred sign language. This arises from the fact that sign languages are not officially recognized, and there is a shortage of qualified teachers to provide sign language instruction. Consequently, there is often limited funding and resources allocated to sign language education and support services.

In India, Sign Language serves as the primary channel for exchanging information for around 18 million hearing-impaired individuals, according to India's National Association of the Deaf [12][13]. Indian Sign Language (ISL) is the main medium for information exchange by hearing-impaired individuals in India. However, ISL faces substantial barriers to recognition and acceptance in various Indian states, primarily due to the lack of standardization and uniformity in its use. Different regions have significant differences in the signs used, grammar, and syntax, making communication between sign language users from different states a challenge. While ISL is used by many deaf and hearing-impaired individuals in India, grammar, and syntax, providing a rich source of linguistic and cultural diversity that should be recognized and preserved. However, despite the emergence of distinct sign languages, individuals who use sign language still face communication challenges when interacting with those who do not understand sign language, primarily due to the limited knowledge and understanding of sign languages compared to spoken languages.

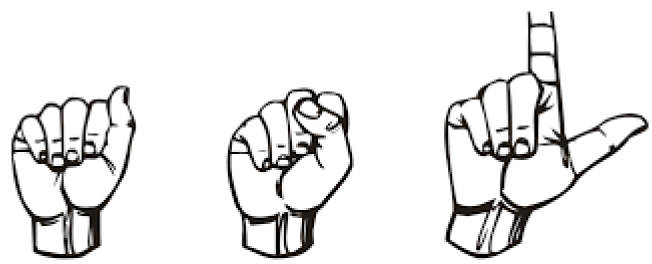

Fig 1: Example of Sign Language

## II. RELATED WORKS

Effective strategies for data collection and classification have been developed by numerous authors. When categorizing previous research based on data collection methods, two distinct categories emerge: direct measurement techniques and vision-based approaches. Direct measurement methods involve the use of sensors, motion-tracking devices, or gloves to gather motion data, enabling precise tracking of fingers, hands, and other body parts. This data acquisition is instrumental in the development of Sign Language detection techniques. In contrast, vision-based SLR techniques rely on RGB photographs to capture geographic and temporal information, a critical aspect of these methods. Before classifying hand areas as gestures, the majority of vision-based algorithms first aim to monitor and extract hand regions.

Since 2020, extensive research has been conducted to detect sign language using deep learning techniques, including long short-term memory networks (LSTMs), recurrent neural networks (RNNs), and convolutional neural networks (CNNs).

Castro et al. pioneered a deep learning algorithm for recognizing signs in Indian, Brazilian, and Korean Sign Languages, with the aim of enhancing communication between the deaf and non-deaf populations [1]. Their model combines various data types for the classification task, including segmented hand images, facial features, hand movement speed and distance, and RGB frames with artificially generated depth maps. The architecture is a multi-stream network designed to consider all this information. The evaluation of their model on a publicly available Brazilian Sign Language dataset revealed an accuracy of 0.91 plus or minus 0.07 and an f1-score of 0.90 plus or minus 0.08, underscoring the effectiveness of this approach in recognizing signs in different sign languages.

Jain et al. introduced a sign language recognition system using CNNs and Support Vector Machines (SVM) [2]. They conducted various preprocessing techniques and experimented with different kernels for SVM and various filter sizes for single and double-layer CNNs. The findings showed that the single-layer CNN achieved an accuracy of 97.344%, while the double-layer CNN achieved an even higher accuracy of 98.581%.

In their research, Sharma et al. presented a deep learning model for recognizing gesture-based sign language [3]. Their model, based on a Convolutional Neural Network (CNN) architecture, was trained and evaluated on two distinct datasets: an Indian Sign Language (ISL) dataset containing 2150 images and an American Sign Language (ASL) dataset. The proposed model outperformed two other CNN models, VGG-11 and VGG-16, achieving remarkable accuracies of 99.96% and 100% for the ISL and ASL datasets, respectively. Additionally, the model exhibited robustness to rotation and scaling transformations and outperformed other techniques in terms of reliability and efficiency.

Ankit et al. developed a desktop software application that utilizes a computer's webcam to capture sign gestures in American Sign Language (ASL) and provides real-time translation into text and speech [4]. The software employed a Convolutional Neural Network (CNN) for gesture recognition, transforming the text into speech for audio output, effectively creating a finger-spelling sign language translator using the CNN approach.

Liao et al. proposed a novel approach for sign language recognition, aiming to overcome challenges faced by existing methods [5]. The new method, known as the BLSTM-3D residual network (B3D ResNet), combines bidirectional LSTM networks with deep 3D residual ConvNet. The method consists of three steps: localizing the hand in video frames to reduce computational complexity, extracting comprehensive features from video sequences, and classifying the video order to accurately identify sign language. Experiments on DEVISIGN_D and SLR Dataset datasets demonstrated an approximate recognition accuracy of 89.8% on DEVISIGN_D and 86.9% on SLR Dataset. The method also excelled in recognizing complex hand gestures and achieving improved accuracy for 500 vocabularies in Chinese hand sign language recognition.

## III. MAJOR CHALLENGES

Despite the extensive research conducted in the field of sign language recognition, there remain several unresolved questions that continue to intrigue researchers. In the context of English Sign Language (MSL), there are specific issues and challenges that demand attention:

Variability in Signs and Grammar: MSL is a natural language that has organically evolved within the deaf community over time. Consequently, it exhibits variations in signs and grammar. These disparities make it challenging to establish a standardized form of MSL for effective detection and recognition.

Limited Availability of MSL Data: The scarcity of publicly accessible MSL data poses a substantial obstacle in developing and training sign language recognition models. This shortage hampers researchers' capacity to amass the necessary data to gain a comprehensive understanding of MSL complexities, subsequently impacting the performance and reliability of MSL recognition systems.

Nuanced Signs and Subtleties: MSL incorporates nuanced signs and subtleties that can prove intricate for machine learning models to grasp. These subtleties can create difficulties for models in distinguishing between similar signs and accurately identifying the correct one. Furthermore, the intricate nature of many MSL signs poses a challenge in comprehending the relationships between various sign components and the underlying conveyed meaning. This diversity and complexity collectively present a significant hurdle in developing precise sign language recognition systems.

## III. LITRATURESURVEY

1. Real-time Sign Language Finger Spelling Recognition:

This work focuses on recognizing static finger spelling in American Sign Language (ASL) without the need for hand-held gloves or sensors. It captures finger gestures using depth maps and continuously converts them to text or voice.

To expand on this, you can mention that depth maps are a 3D representation of the environment, which helps in capturing the depth information of the signer's hand shapes. Convolutional Neural Networks (CNNs) are a type of deep learning model commonly used for image recognition tasks.

2. Design of a Communication Aid for Physically Challenged:

This system is implemented in MATLAB and consists of two phases: a training phase and a testing phase, using feed-forward neural networks. It's aimed at assisting individuals with physical disabilities in communication.

To elaborate further, you can discuss the challenges associated with MATLAB's efficiency and limitations, and how other programming languages or tools may offer more efficient alternatives. Also, integrating concurrent attributes could refer to the difficulty of managing multiple features or input sources simultaneously.

3. Sign Language Interpreter System for Deaf and Dumb Individuals:

This system can recognize 20 out of 24 static ASL alphabets, with difficulties in recognizing the letters A, M, N, and S due to occlusion issues. It uses a limited number of images.

To provide additional information, you can mention that occlusion issues can occur when one part of the hand or fingers obscures another, making it challenging for the system to accurately identify the gestures. You might also discuss potential solutions to improve recognition, such as using more diverse training data or employing advanced image processing techniques.

4. Survey on Sign Language Recognition:

This survey discusses sign language recognition, particularly in the context of vision-based and deep learning approaches. It mentions the evolution from classifying static signs and alphabets to a more comprehensive system.

To expand on this, you can provide an overview of the evolution of sign language recognition technology, emphasizing the role of deep learning models like CNNs and recurrent neural networks (RNNs) in improving accuracy and extending the system's capabilities to include dynamic signs and full sentences in sign language.

5. Recent Advances in Sign Language Recognition: A Literature Review:

This literature review provides an overview of recent advancements in the field of sign language recognition. It covers a wide range of topics, including techniques, challenges, and applications in the context of sign language interpretation. The review aims to highlight the progress made in sign language recognition systems, with a focus on both static and dynamic sign recognition.

6. Natural Language Processing in Healthcare: A Comprehensive Literature Review:

This comprehensive literature review provides an extensive overview of the applications, challenges, and recent advancements in the field of Natural Language Processing (NLP) in healthcare. NLP techniques have gained prominence in managing, extracting insights from, and utilizing the vast amount of textual data in the healthcare domain.

The integration of NLP in healthcare has opened up new horizons for data-driven healthcare management. This review explores the multifaceted applications of NLP, including clinical documentation, electronic health records (EHRs), medical literature analysis, telemedicine, and disease surveillance. It emphasizes the critical role of NLP in enhancing healthcare delivery and decision-making.

## IV. DISCUSSION

In this paper, a survey on sign language recognition is presented and various techniques have been studied and analysed for the same. In recognition process, segmentation plays a crucial part in which skin region is separated from the background which usually affects the recognition accuracy. Besides segmentation, classification also depends on the feature extraction techniques which performs dimensionality reduction and reduces the computation cost. Study of various classification techniques concludes that deep neural network

(CNN, Inception model, LSTM) performs better than traditional classifiers such as KNN and SVM.

## REFERENCES


[1] G. Z. de Castro, R. R. Guerra, and F. G. Guimarães, “Automatic translation of sign language with multi-stream 3D CNN and generation of artificial depth maps,” Expert Systems with Applications, vol. 215, p. 119394, Apr. 2023, doi: 10.1016/j.eswa.2022.119394.

[2] Jain, V., Jain, A., Chauhan, A. et al. American Sign Language recognition using Support Vector Machine and Convolutional Neural Network. Int. j. inf. tecnol. 13, 1193–1200 (2021). https://doi.org/10.1007/s41870-021-00617-x

[3] S. Sharma and S. Singh, “Vision-based hand gesture recognition using deep learning for the interpretation of sign language,” Expert Systems with Applications, vol. 182, p. 115657, Nov. 2021, doi: 10.1016/j.eswa.2021.115657.

[4] Ankit Ojha, Ayush Pandey, Shubham Maurya, Abhishek Thakur, Dr. Dayananda P, 2020, Sign Language to Text and Speech Translation in Real Time Using Convolutional Neural Network, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) NCAIT – 2020 (Volume 8 – Issue 15)

[5] Y. Liao, P. Xiong, W. Min, W. Min and J. Lu, "Dynamic Sign Language Recognition Based on Video Sequence With BLSTM-3D Residual Networks," in IEEE Access, vol. 7, pp. 38044-38054, 2019, doi: 10.1109/ACCESS.2019.2904749.

[6] Dhanjal, A.S., Singh, W. An automatic machine translation system for multi-lingual speech to Indian sign language. Multimed Tools Appl 81, 4283–4321 (2022). https://doi.org/10.1007/s11042-021- 11706-1

[7] Srivastava, Sharvani, Amisha Gangwar, Richa Mishra, and Sudhakar Singh. “Sign Language Recognition System Using TensorFlow Object Detection API.” Sign Language Recognition System Using TensorFlow Object Detection API | SpringerLink, February 18, 2022. https://doi.org/10.1007/978-3-030-96040-7_48.

[8] Sharma, Ashish, Nikita Sharma, Yatharth Saxena, Anuraj Singh, and Debanjan Sadhya. “Benchmarking Deep Neural Network Approaches for Indian Sign Language Recognition - Neural Computing and Applications.” SpringerLink, October 24, 2020. https://doi.org/10.1007/s00521-020-05448-8.

[9] GAO, W, G FANG, D ZHAO, and Y CHEN. “A Chinese Sign Language Recognition System Based on SOFM/SRN/HMM.” Pattern Recognition 37, no. 12 (December 2004): 2389–2402. https://doi.org/10.1016/s0031-3203(04)00165-7.

[10] TensorFlow. “Transfer Learning and Fine-Tuning | TensorFlow Core,” n.d. https://www.tensorflow.org/tutorials/images/transfer_learning.

[11] Our Work - WFD. “Our Work - WFD,” n.d. https://wfdeaf.org/ourwork/.

[12] National Association of the Deaf, n.d. http://nadindia.org/.

[13] Brill R. 1986. The Conference of Educational Administrators Serving the Deaf: A History. Washington, DC: Gallaudet University Press.

[14] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791

[15] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510-4520, doi: 10.1109/CVPR.2018.00474..

[16] L. K. Hansen and P. Salamon, "Neural network ensembles," in IEEE Transactions on PatternAnalysis and Machine Intelligence, vol. 12, no. 10, pp. 993-1001, Oct. 1990, doi: 10.1109/34.58871.

[17] Kang, Byeongkeun, Subarna Tripathi, and Truong Q. Nguyen. ”Real- time sign language fingerspelling recognition using convolutional neural networks from depth map.” arXiv preprint arXiv: 1509.03001 (2015).

[18] Pradeep Kumar, Himaanshu Gauba, Partha Pratim Roy, and Debi Prosad Dogra. “Coupled HMM-based multi-sensor data fusion for sign language recognition”. In: Pattern Recognition Letters 86 (2017), pp. 1–8.

[19] Y. Liu, X. Yao, Ensemble learning via negative correlation, Neural Networks, Volume 12, Issue 10,1999, Pages 1399-1404, ISSN 0893-6080,https://doi.org/10.1016/S0893-6080(99)00073-8.

[20] Ankit Ojha, Ayush Pandey, Shubham Maurya, Abhishek Thakur, Dr. Dayananda P, 2020, Sign Language to Text and Speech Translation in Real Time Using Convolutional Neural Network, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) NCAIT – 2020 (Volume 8 – Issue 15)

[21] \newline Akash. ASL Alphabet. url: https://www.kaggle.com/grassknoted/asl-alphabet. (accessed: 24.10.2018).

[22] G. Z. de Castro, R. R. Guerra, and F. G. Guimarães, “Automatic translation of sign language with multi-stream 3D CNN and generation of artificial depth maps,” Expert Systems with Applications, vol. 215, p. 119394, Apr. 2023, doi: 10.1016/j.eswa.2022.119394.

[23] D. Metaxas. Sign language and human activity recognition, June 2011. CVPR Workshop on Gesture Recognition.

[24] M. Ranzato. Efcient learning of sparse representations with an energy-based model, 2006. Courant Institute of Mathematical Sciences.

[25] S. Sarkar. Segmentation-robust representations, matching, and modeling for sign language recognition, June 2011. CVPR Workshop on Gesture Recognition, Co-authors: Barbara Loeding, Ruiduo Yang, Sunita Nayak, Ayush Parashar.

[26] P. Y. Simard. Best practices for convolutional neural networks applied to visual document analysis, August 2003. Seventh International Conference on Document Analysis and Recognition.

[27] X. Teng. A hand gesture recognition system based on local linear embedding, April 2005. Journal of Visual Languages and Computing.