Sign Language Recognition: A Real-Time Solution for Inclusive Communication

Manisha More[1], Kunal Bavdhane[2], Vedant Bartakke[3], Sushant Bari[4], Ved Batra[5] and Tanay Bargir[6]
{manisha.more1@vit.edu[1], kunal.bavdhane24@vit.edu[2], vedant.bartakke24@vit.edu[3], sushant.bari24@vit.edu[4], ved.batra24@vit.edu[5], tanay.bargir24@vit.edu[6] }

Department of Engineering, Sciences and Humanities (DESH), Vishwakarma Institute of Technology, Upper Indira Nagar Bibwewadi, Pune, Maharashtra, India.[1, 2, 3, 4, 5, 6]

Abstract. Some individuals face challenges due to hearing and speech impairment. The sign language plays a vital role in the communication of these people. But most of the people in our surroundings are unfamiliar with sign language, and here is the barrier of communication. Although various tools exist to facilitate communication many of them are quite expensive, with less accuracy or fail to provide real-time solutions. Our project aims to address this gap by creating an accessible platform that translates Indian sign language gestures, captured by camera, into English text. This solution enables seamless communication between sign language users and others who do not understand the sign language and breaks down the barriers in personal and professional interactions.

Keywords: Indian sign language, Machine language, Real-time processing, Sign language recognition, Hand Gesture Recognition.

1 Introduction

According to the WHO (World Health Organization) report, over 466 million people are speech or hearing impaired, and 80% of them are semi-illiterate or illiterate [1]. Sign language is a fundamental mode of communication for individuals with hearing and speech impairments [2]. However, a significant barrier exists due to the lack of widespread understanding of sign language among the general population speech-impaired community [3]. The research aims to contribute to the advancement of assistive technology by providing an intuitive and scalable approach for sign language translation, ultimately fostering greater inclusivity in society. This communication gap limits the ability of sign language users to interact seamlessly in personal, social, and professional settings.

There are Several solutions have been developed to translate sign language into text. However, these solutions have high cost, low accuracy, or the unable to provide real-time recognition. To address these challenges, this research presents a real-time sign language recognition system tailored for Indian Sign Language (ISL) by using CNN.

The proposed system used machine learning techniques and computer vision for sign language gestures capturing and interpretation [4]. The model is trained by a dataset of gestures, with landmarks that are from Mediapipe Pose. A Random Forest classifier is employed to train data to recognize gestures with high accuracy [5]. By the real-time processing immediate translation of sign gestures into English text, there will be no problem for smooth communication between individuals with and without knowledge sign language [6].

ICITSM-Part II 2025, April 28-29, Tiruchengode, India

DOI 10.4108/eai.28-4-2025.2357973

By giving efficient machine learning algorithms and real-time processing, this system offers a cost-effective, accurate, and accessible solution to bridge the communication gap between individuals with and without knowledge sign language

2 Related works

Sign Language Recognition (SLR) has supported deep learning and computer vision for a long time. Various means to improve accuracy and performance are being discovered. One study has introduced a multi-stream convolutional network. This model has combined skeletal and facial expressions with motion data. Thus, it's doing a great job handling issues like background noise and occlusion [7]. As time has passed, SLR has moved beyond simple gesture recognition to full Sign Language Translation (SLT). Early models had a lot of trouble understanding the complexity of sign language. Thus, some newer methods like Neural Machine Translation (NMT) were introduced. These methods have made sign-to-text conversion better. Thus, helping to deal with grammar and word order differences [8]. Another approach used transfer learning with VGG16. It reached a really considerable accuracy and even worked in a mobile app. Thus, supporting real-time sign recognition [9]. Another study used a Leap Motion Controller (LMC) as well as Dynamic Time Warping (DTW) for gesture tracking. It could track precise movements, but it struggled with real-time processing [10].

A detailed review of SLR helped researchers divide sign gestures into two types: manual and non-manual. It highlighted the factors like the environment and dataset quality that affect accuracy [11]. There are hybrid models that combine visual and motion data. Such models have improved ISL translation. But dataset limitations are still an issue [12]. Another study created an SLR system in MATLAB to recognize 26 hand gestures. It worked well but struggled with real-time complexity [13]. Another study focused on converting spoken language into sign videos using Neural Machine Translation (NMT). The system had high accuracy, but it struggled with video quality and recognizing facial expressions [14]. Transformers have also been used in SLR to recognize and translate signs at the same time. One model achieved state-of-the-art results but had trouble recognizing names and numbers [15]. Researchers have explored different machine learning techniques for ISL classification. An ANN-based model using feature extraction reached a considerable accuracy of 94.37%. Thus, beating SVM classifiers [16]. A Tamil script-based ISL recognition system analysed fingertip positions. It performed well but found some complex signs difficult to classify [17]. Table 1 show the literature review.

Table. 1. Literature Review.

Review Paper	Technologies Used	Limitations and Gaps
Sign language recognition using graph and general deep neural network based on large scale dataset [18]	1. Multi-stream Graph Convolutional Network (GCAR) 2. Large-scale datasets (WLASL, PSL, MSL, ASLLVD)	1. Focused primarily on recognition rather than translation 2. Limited information on real-time efficiency and deployment
Deep learning for sign language recognition: Current techniques,	1. Deep Learning for SLR and Sign Language Translation (SLT)	1. Language complexities in sign language is still a challenge.

benchmarks, and open issues [19]	2. Neural Machine Translation (NMT)	2. Real-time processing speed not discussed
Sign language to text conversion in real time using transfer learning [20]	1. Transfer Learning (VGG16 pre-trained model). 2. Flutter-based mobile application.	1. Relies mostly on pre-trained ImageNet features**.** 2. No discussion on handling continuous sign language recognition
Conversion of sign language to text and speech	1. Leap Motion Controller (LMC) for skeletal tracking 2. Dynamic Time Warping (DTW) for gesture classification	1. Real-time performance is limited. 2. Lack of deep learning integration (relies on traditional DTW). 3. Issues with overlapping figures.
A comprehensive review of sign language recognition: Different types, modalities, and datasets [21]	1. SLR modalities and datasets 2. Manual and non-manual signs	1. Less discussion on real-time implementation. 2. Quality of dataset affects the accuracy of model.
A simple multi-modality transfer learning baseline for sign language translation [22]	1. CNNs and hybrid models 2. Multimodal Learning for continuous sign recognition	1. Limited dataset and challenge in real time processing.
Conversion of Sign Language into Text	1. MATLAB-based image processing 2. Linear Discriminant Analysis (LDA)	1. LDA is outdated compared to modern deep learning 2. Limited to 26 hand gestures without any real time implementation.
Sign Language Production using Neural Machine Translation and Generative Adversarial Networks [23]	1. Neural Machine Translation (NMT) 2. Generative Adversarial Networks (GANs)	1. Video resolution and gesture recognition issues
Sign Language Transformers: Joint End-to-End Sign Language Recognition and Translation [22]	1. Transformer-based model 2. PHOENIX14T dataset	1. Difficulty in recognizing named items and numbers. 2. Dataset size limits.
Indian Sign Language Recognition System	1. Artificial Neural Network 2. Support Vector Machine	1. SVM has low performance compared to deep learning models
Real-Time Indian Sign Language Recognition System to Aid Deaf-mute People [24]	1. Tamil script-based ISL recognition	1. Struggles with complex gestures 2. Limited scalability for other sign languages

3 Methodology

The working of real time sign language recognition system.

Data Acquisition: For real-time prediction initially, we need a dataset from which our system can predict sign language. For the data set, first of all, collect video frames in real time. Basically, video frames are extracted with landmarks across the whole body. This landmark extraction is done through MediaPipe Pose [25]. This detects 33 landmarks on the body. The extracted landmarks from each frame are saved in a CSV file with the corresponding action label. To work our system more accurately, the data should be collected in different lighting conditions and from different users [26].

Preprocessing: Once the dataset is collected, there is a need to enhance the quality and consistency of the dataset. We need to convert frames that are collected from video from BGR to RGB format. The landmark coordinates, which are extracted and saved in CSV files, are normalized. By normalizing these coordinates, the dataset becomes more robust and adaptable across different conditions.

Model Training: After pre-processing the data, the next step is to train the model. Basically, in CSV files, it contains landmark coordinates, visibility, and action labels [27]. which are further loaded, concatenated, and processed. To train the pre-processed dataset, we used the Random Forest Classifier (RFC) algorithm. There are multiple decision trees in RFC based on the landmark features. Each tree makes votes (predictions). The class of tree that has majority votes is selected as the final votes (predictions). During the training process, the model is trained on 80% of the dataset. The remaining 20% of the dataset is used for testing. After all this new data of all gestures is saved to a .pkl file, which is for further deployment. The flowchart for our program is in fig 1.

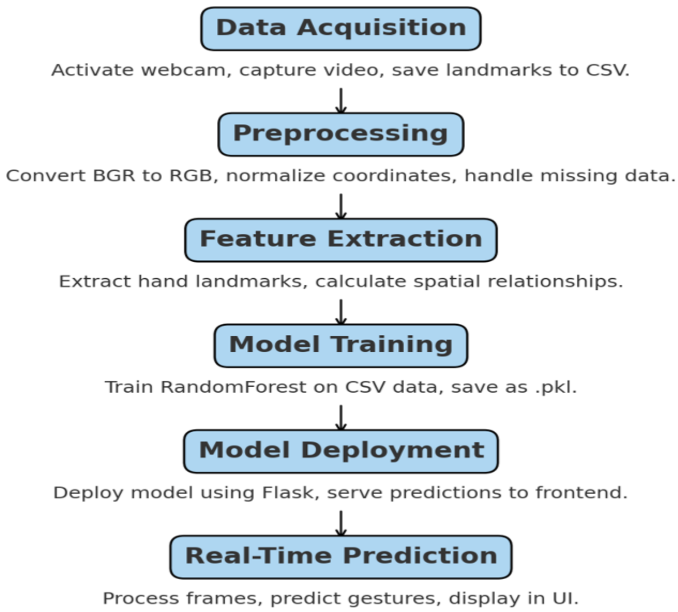

Fig. 1. System Flow Diagram.

Model Deployment: Once the model is trained, we have to deploy this trained model using a Flask-based server. The server is deployed such that it accesses the webcam feed. From the front end, video frames are passed to the back end simultaneously, and predicted gestures are passed to the front end.

Real-Time Prediction: The real-time prediction is done through a web-based interface. While real-time prediction further processes are followed. Such as through web-based interface frames extracted from video, they are passed through the trained model. Flask serves the necessary data from the backend, and the prediction of the corresponding sign is displayed. The dynamic updating of prediction is handled by JavaScript [28]. This methodology ensures an efficient workflow from data collection to real-time prediction.

4 Results and Evaluation

Fig 2 shows the results show that the system effectively translates Indian Sign Language (ISL) into written English instantly. The system uses Mediapipe for landmark extraction and a Random Forest Classifier to understand the gestures. To prove the effectiveness and accuracy of the proposed system, we have carried out a number of experiments.

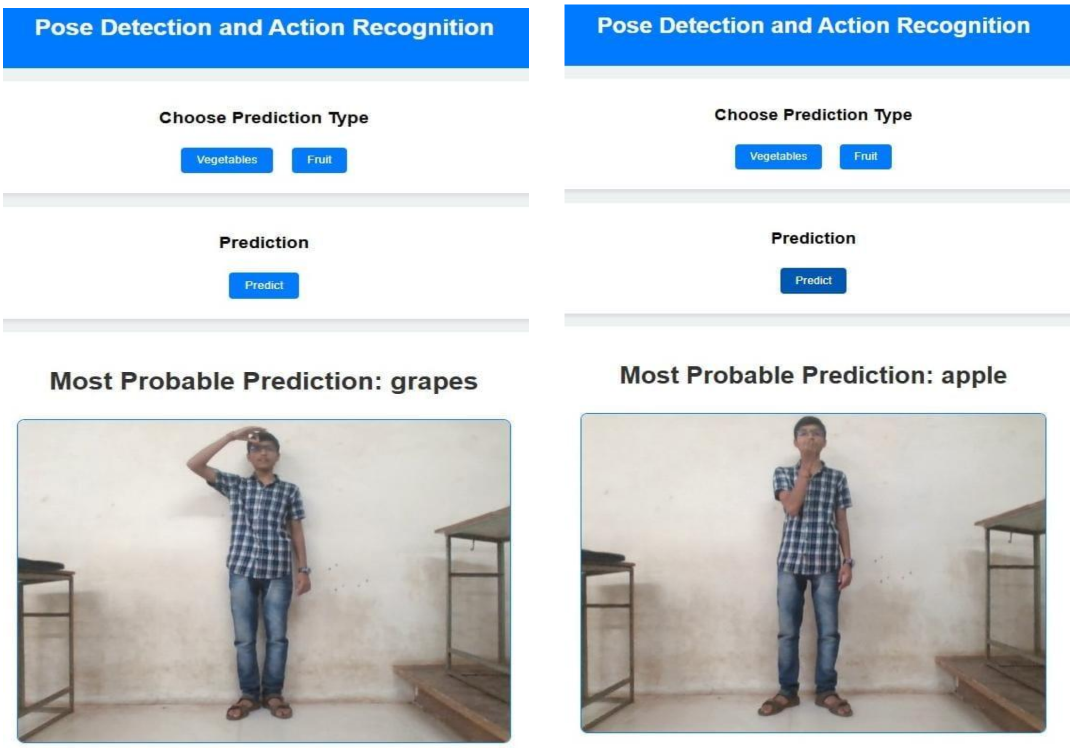

Fig. 2. Sign Language “grapes” and Sign Language “apple”

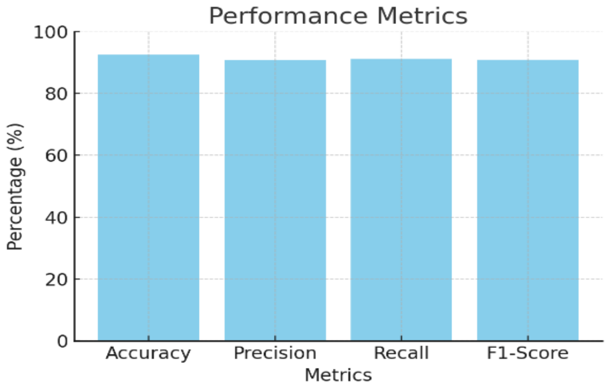

Fig. 3. Performance Metrics.

Performance Metrics Overview: The performance of the Random Forest Classifier (RFC) was assessed using standard metrics [29]. Accuracy, precision, recall and F1-score are assessed. The metrics were computed using our test dataset and are showed in Figure 3.

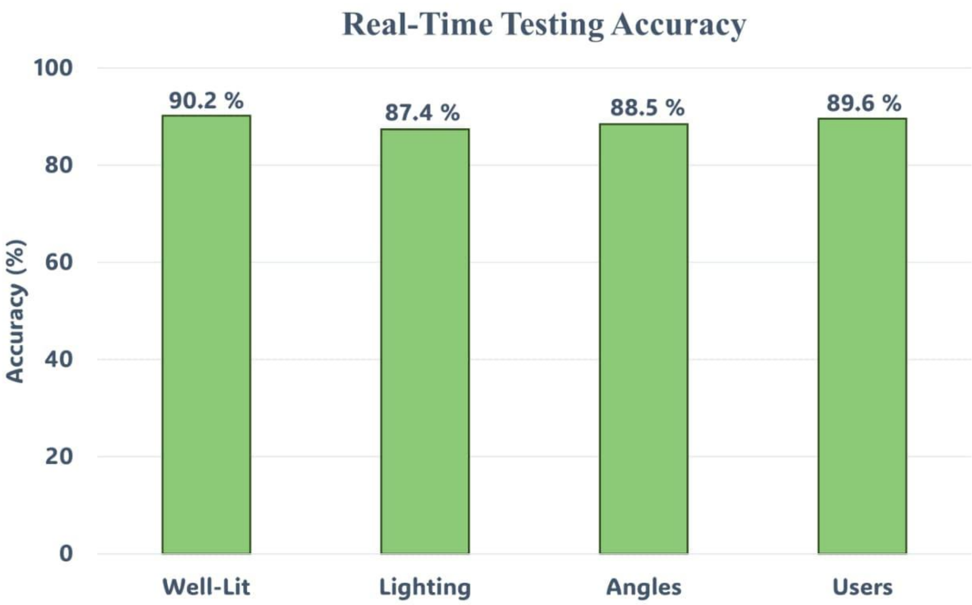

Fig. 4. Accuracy Graph.

Real-Time Testing Results: To test the model's real-time performance, testing was done and observed under various environmental conditions. The results that are depicted in Fig 4, indicate the model's strength across different scenarios.

The testing conditions and the respective accuracy were as follows: A environment with good light achieved 90.2%, dim lighting resulted in 87.4%, side angles showed 88.5% and testing across different users recorded 89.6%. Figure 5 illustrates these performance variations. Thus, highlighting the importance of consistent lighting for optimal results.

5 Discussion

The proposed real-time sign language recognition system has illustrated promising results; several avenues for future research and development exist to further enhance its capabilities and broaden its applicability. The following represent the scope for future work:

To improve the system's performance in low light, development in image preprocessing techniques needs to be done. Techniques such as adaptive histogram equalization or low-light image enhancement algorithms are instances.

The current system's training data consists of (ISL) Indian Sign Language gestures. The dataset can be enlarged to include additional gestures. Even hand movement variants and data from a variety of users can be included in order to increase generality and accuracy.

The accuracy and user experience of the system can be improved. It can be done by integrating a feedback mechanism that allows users to report regarding the misclassified gestures in real-time. This feedback can be used to regularly update and retrain the mode. Thus, helping to increase model’s adaptability for different users.

6 Conclusion

The proposed system is for continuous Indian Sign Language recognition. It successfully addresses the communication gap between deaf and mute people. The system gives high accuracy and real-time performance with the use of Mediapipe. The system extracts continuous gestures into isolated gestures. It also shows how many frames an isolated gesture we will have. This system translates Indian sign language into English text, which is very helpful for deaf individuals. It can help them in personal, social, and professional situations. It also offers a significant contribution to inclusivity across personal, social, and professional domains.

References

[1] W. H. Organization. (2018) Deafness and hearing loss. [Online]. Available: https://www.who.int/news-room/ fact-sheets/detail/deafness-and-hearing-loss

[2] K. Anitha, R. N. Karthick, K. C. Varsheni and A. Kalaiselvi, "Gesture based Sign Language Recognition System," *2023 2nd International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, Coimbatore, India, 2023, pp. 1-6, doi: 10.1109/ICAECA56562.2023.10199750.

[3] 9 P. Vyavahare, S. Dhawale, P. Takale, V. Koli, B. Kanawade, and S. Khonde, "Detection and Interpretation of Indian Sign Language Using LSTM Networks," J. Intell Syst. Control, vol. 2, no. 3, pp. 132-142, 2023. https://doi.org/10.56578/jisc020302

[4] S, Ajay & Potluri, Ajith & Mohan George, Sara & R, Gaurav & S, Anusri. (2021). Indian Sign Language Recognition Using Random Forest Classifier. 1-6. 10.1109/CONECCT52877.2021.9622672.

[5] V. Kumar, R. Sreemathy, M. Turuk, J. Jagdale and A. Agarwal, "Real-time Indian Sign Language Recognition using Skeletal Feature Maps," 2023 14th International Conference on Computing

Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1-6, doi: 10.1109/ICCCNT56998.2023.10308342.
[6] Khartheesvar, G & Kumar, Mohit & Yadav, Arun Kumar & Yadav, Divakar. (2023). Automatic Indian sign language recognition using MediaPipe holistic and LSTM network. Multimedia Tools and Applications. 83. 1-20. 10.1007/s11042-023-17361-y.
[7] A.S. M. Miah, M. A. M. Hasan, S. Nishimura, and J. Shin, "Sign language recognition using graph and general deep neural network based on large scale dataset," *IEEE Access*, vol. 12, pp. 34553-34568, Mar. 2024, doi: 10.1109/ACCESS.2024.3372425.
[8] M. Al-Qurishi, T. Khalid, and R. Souissi, "Deep learning for sign language recognition: Current techniques, benchmarks, and open issues," *IEEE Access*, vol. 9, pp. 126917-126934, Sep. 2021, doi: 10.1109/ACCESS.2021.3110912.
[9] S. Thakar, S. Shah, B. Shah, and A. V. Nimkar, "Sign language to text conversion in real time using transfer learning," *arXiv preprint arXiv:2211.14446v2*, Dec. 2022
[10] G. U. R, S. Madhyastha, V. A. M, and P. B. G, "Conversion of sign language to text and speech," *Int. J. Adv. Res. Dev.*, vol. 2, no. 3, pp. 42, 2017.
[11] M. Madhiarasan and P. P. Roy, "A comprehensive review of sign language recognition: Different types, modalities, and datasets," *J. LaTeX Class Files*, vol. XX, no. X, pp. 1-XX, Apr. 2022, arXiv:2204.03328.
[12] Y. Chen, F. Wei, X. Sun, Z. Wu, and S. Lin, "A simple multi-modality transfer learning baseline for sign language translation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Apr. 2022, pp. 5120–5121, arXiv:2204.00434.
[13] M. K. N. B., "Conversion of Sign Language into Text," *Int. J. Appl. Eng. Res.*, vol. 13, no. 9, pp. 7154-7161, 2018.
[14] S. Stoll, N. C. Camgoz, S. Hadfield, and R. Bowden, "Sign Language Production using Neural Machine Translation and Generative Adversarial Networks," in *Proc. British Machine Vision Conference (BMVC)*, 2018.
[15] N. C. Camgoz, O. Koller, S. Hadfield, and R. Bowden, "Sign Language Transformers: Joint End-to-End Sign Language Recognition and Translation," in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 10023-10033.
[16] Y. I. Rokade and P. M. Jadav, "Indian Sign Language Recognition System," *International Journal of Engineering and Technology (IJET)*, vol. 9, no. 3S, pp. [page numbers], July 2017. doi: 10.21817/ijet/2017/v9i3/170903S030.
[17] P. S. Rajam and G. Balakrishnan, "Real-Time Indian Sign Language Recognition System to Aid Deaf-Dumb People," *2011 IEEE International Conference on Computer Science and Automation Engineering (CSAE)*, vol.pp.2011. doi: 10.1109/CSAE.2011.
[18] H L, G., Flammini, F., Srividhya, S., M L, C., & Selvam, S. (Eds.). (2024). Computer Science Engineering: Proceedings of the 1st International Conference on Computing and Intelligent Information Systems (ICCIIS 2024), Bangalore, India, 19-20th April, 2024 Volume 1 (1st ed.). CRC Press. https://doi.org/10.1201/9781032711157
[19] MUHAMMADAL-QURISHI, (Member, IEEE), THARIQ KHALID, AND RIAD SOUISSI, "Deep Learning for Sign Language Recognition: Current Techniques, Benchmarks, and Open Issues" Research and Innovation Division, Research Department, Elm Company, Riyadh 12382, Saudi Arabia
[20] Pooja Bagane, Muskaan Thawani, Prerna Singh, Raasha Ahmad, Rewaa Mital, Obsa Amenu Jebessa, "Breaking the Silence: An innovative ASL to Text Conversion System Leveraging Computer Vision & Machine Learning for Enhanced Communication", International Journal of Intelligent Systems and Applications in Engineering, ISSN:2147-6799
[21] Manoharan, Madhiarasan & Roy, Partha. (2022). A Comprehensive Review of Sign Language Recognition: Different Types, Modalities, and Datasets. 10.48550/arXiv.2204.03328.
[22] Yutong Chen Fangyun Wei Xiao Sun Zhirong Wu Stephen Lin Tsinghua University Microsoft Research Asia, "ASimple Multi-Modality Transfer Learning Baseline for Sign Language Translation"

[23] Fabrizio Nunnari, Cristina España-Bonet, Eleftherios Avramidis, "A Data Augmentation Approach for Sign-Language-To-Text Translation In-The-Wild"

[24] T. Anusha, Deepti Ravi Kumar, Nandha Kishore, Rahul Raj D, Raswanth E A, Samyuktha Sreekanth, "COMPUTER VISION-ENABLED GESTURE RECOGNITION FOR DIFFERENTLY-ABLED PEOPLE" , International Research Journal of Engineering and Technology (IRJET) , e-ISSN: 2395-0056 , p-ISSN: 2395-0072

[25] K. Tripathi, N. Baranwal, and G. C. Nandi, "Continuous Indian Sign Language Gesture Recognition and Sentence Formation," in *Procedia Computer Science*, vol. 54, pp. 523-531, 2015, doi: 10.1016/j.procs.2015.06.060.

[26] F. Ansari, N. Ansari, M. Khanche, and I. Shaikh, "Real Time Sign Language to Text Conversion," *Int. J. Res. Appl. Sci. Eng. Technol. (IJRASET)*, vol. 11, no. IV, pp. 3430, Apr. 2023. doi: 10.22214/ijraset.2023.50962.

[27] M. Al-Qurishi, T. Khalid, and R. Souissi, "Deep learning for sign language recognition: Current techniques, benchmarks, and open issues," *IEEE Access*, vol. 9, pp. 126917-126934, Sep. 2021, doi: 10.1109/ACCESS.2021.3110912.

[28] N. Pugeault and R. Bowden, *"Spelling It Out: Real-time ASL Fingerspelling Recognition,"* in *Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW)*, Barcelona, Spain, 2011, pp. 1114-1119.

[29] N. V. Jagan Mohan, R., Chandra Sekhar, V., & M. N. S. S. V. K. R. Gupta, V. (Eds.). (2024). Algorithms in Advanced Artificial Intelligence: ICAAAI-2023 (1st ed.). CRC Press. https://doi.org/10.1201/9781003529231