IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 11, April 2026
Impact Factor: 8.2

Bridging Silence: Sign Language Recognition Using Deep Learning and Computer Vision

Shivam Bhardwaj[1], Deepanshu Bhardwaj[2], Vikas Chaudhary[3], Deepak Bisht[4], Shagun Malhotra[5]
UG Scholars, Computer Science & Engineering Department[1,2,3,4]
Assistant Professor, Computer Science & Engineering Department[5]
Sunder Deep Engineering College, Ghaziabad, India

Abstract: *This article is about using deep learning and computer vision to help with sign language, like recognizing it and translating into words or sounds. It seems important because so many people who are deaf or can't speak clearly struggle with communication every day, and this kind of tech could make things easier for them. The main part involves processing videos in real time with something called Convolutional Neural Networks, or CNNs. These networks look at hand gestures from the video feed and figure out what they mean, then output text or even audio. I think it works by mixing image processing techniques with machine learning to spot patterns in the gestures. That way, the translations come out pretty accurate, at least from what the study says. One thing that stands out is how this could apply in actual situations. Like in schools for teaching, or in hospitals for better patient care, even public places or just general assistive devices. It feels like it has potential there, but the study points out some real hurdles too. Changing lights can mess up the recognition, or busy backgrounds make it hard to isolate the hands. Plus, datasets are often small, which limits how well it learns. And there is ethical stuff, like making sure its accessible to everyone and reliable enough not to fail when it matters. The results look promising overall; it shows good accuracy in tests. But they mention needing better data quality down the line, and speeding up the real-time part so it doesn't lag. I might be missing some details here, but that's the gist.*

Keywords: Deep Learning, Convolutional Neural Network (CNN), Computer Vision, Gesture Recognition, Image Processing, Human-Computer Interaction, and Sign Language Recognition

I. INTRODUCTION

Sign language is an important way for people who are deaf or have trouble speaking to communicate. It uses body language, hand gestures, and facial expressions to get the message across. But when you work with people who don't know it well, communication problems can happen, especially in public services, healthcare, and education. Accessible solutions are very important because they affect millions of people around the world. Automated solutions are needed because regular interpreters are either too expensive or not available.

Advances in AI, machine learning, and computer vision, especially CNNs, have made it possible to accurately detect gestures. The suggested system uses a webcam to record movements, image analysis and classification to figure out what they mean, and then speech and text to turn them into words. It is affordable, simple to use, and encourages diversity by enhancing accessibility and communication for people with speech and hearing impairments.

II. LITERATURE REVIEW

For people who can't hear or speak, sign language is a very important way to talk to others. A lot of people don't get it, though, which makes it hard to talk to each other. Researchers have created systems that can turn sign language into text or speech using deep learning and computer vision [1]. This field has gotten a lot of attention in the last few years because of improvements in artificial intelligence (AI) that make real-time and accurate translation possible [2]. In the past, systems that could recognize sign language used sensors, like gloves and motion trackers. These methods were

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 11, April 2026
Impact Factor: 8.2

costly and not easy to use [2]. Most modern systems use vision-based methods, which means that cameras record hand movements and deep learning models analyse them. These systems are more useful and used more often today [3]. Methods for Deep Learning CNNs are commonly utilized for image processing tasks. CNNs help identify hand shapes and gestures from pictures when recognizing sign language. Research indicates that CNN-based models can attain exceptional accuracy (up to 99%) in gesture recognition [1]. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are used to understand data that comes in a sequence, like hand movements in videos. For real-time systems, combining CNN with LSTM makes them work better [2]. Hybrid Models: New research uses a mix of methods (CNN + LSTM + optimization algorithms) to make things faster and more accurate [4]. Deep learning stuff needs a ton of data to actually work well during training. Like, you see these models' using pictures and videos of hand gestures all the time, [2], that kind of thing. It seems straightforward at first. But then there's this big issue with not having enough varied datasets for sign languages. Especially the ones that are more local or regional, they just don't have as much out there.

III. METHODOLOGY

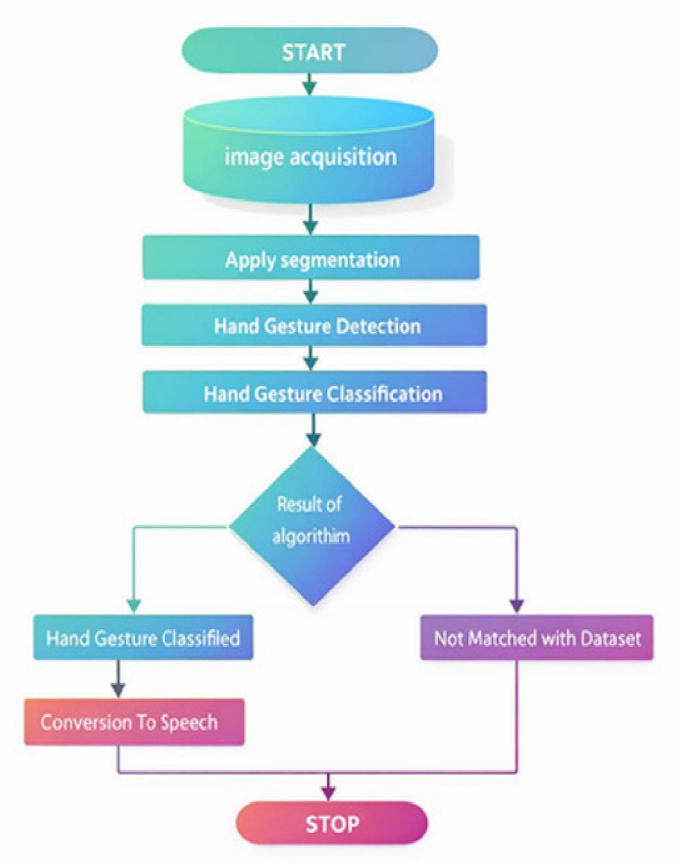

Fig. 1: Flow Chart for Proposed System

The suggested sign language recognition system uses a structured pipeline to translate hand motions into text and speech in real time.

1. Acquisition of Images: A webcam that records continuous video and divides it into frames is used to record hand motions. These frames facilitate dynamic gesture recognition and serve as input for real-time processing.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 11, April 2026 Impact Factor: 8.2

2. Pre-processing Images: To increase quality and guarantee consistent input for better performance, captured photos are improved by resizing, grayscale conversion, noise reduction, and normalizing.
3. Hand Detection and Segmentation: In order to eliminate background noise and increase recognition accuracy, the system isolates the hand region using methods like skin detection and thresholding.
4. Extraction of Features: Important gesture elements are automatically extracted by a CNN, which captures patterns and shapes without human interaction.
5. Gesture Recognition using CNN: The CNN classifies gestures using trained layers, accurately predicting gesture labels from input frames.
6. Text Generation: Recognized gestures are converted into corresponding text for user understanding.
7. Speech Generation: Finally, text is transformed into speech using a TTS system, enabling effective communication.

IV. RESULTS

The implemented system successfully recognizes hand gestures and converts them into text and speech in real time with satisfactory accuracy. The CNN model demonstrates consistent performance under controlled lighting and simple backgrounds. Minor variations in gesture orientation and environmental conditions slightly affect prediction accuracy. Overall, the system proves effective as a low-cost assistive tool, showing potential for practical deployment in communication scenarios involving hearing-impaired individuals.

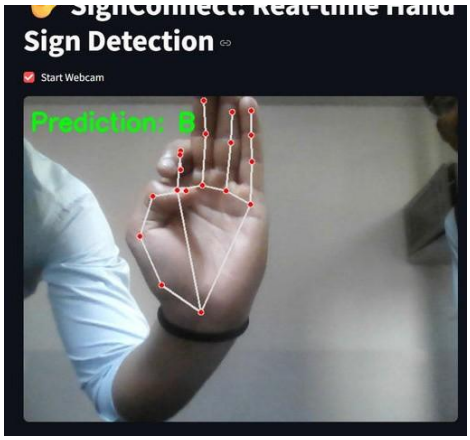

Fig. 2: GUI Screenshot 1

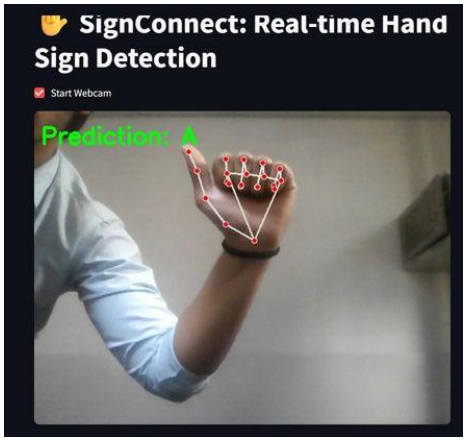

Fig. 3: GUI Screenshot 2

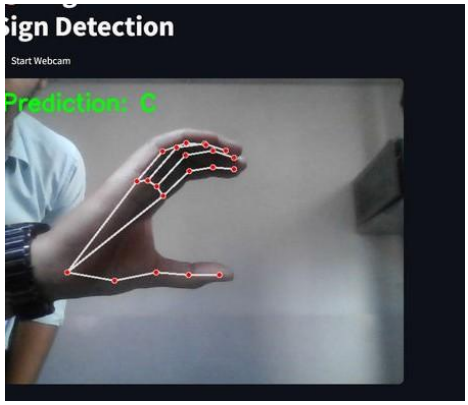

Fig. 4: GUI Screenshot 3

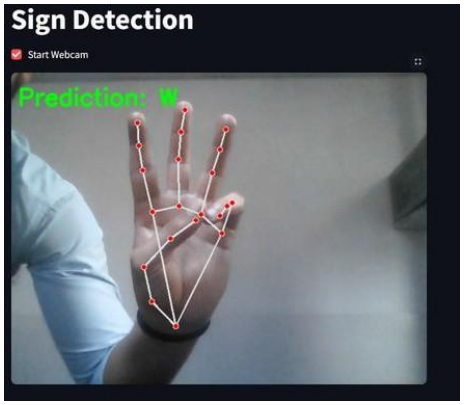

Fig. 5: GUI Screenshot 4

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 11, April 2026
Impact Factor: 8.2

V. CONCLUSION

The implemented system successfully recognizes hand gestures in order to close the communication gap between hearing-impaired people and the broader public; this research proposes a real-time sign language recognition and translation system. The system may work well in controlled settings, but things like changes in light, background complexity, and the variety of datasets can all affect how well it works. To make things more accurate and make sure they work well in changing situations, these problems need to be fixed.

In the end, this study shows how assistive technologies based on deep learning can help with communication that includes everyone. The suggested system not only helps improve sign language recognition, but it also lays the groundwork for future improvements in intelligent human-computer interaction systems. Continued improvements in model efficiency, dataset size, and the ability to process data in real time can make these kinds of systems even more useful and impactful in everyday situations.

VI. AUTHORS AND AFFILIATIONS

Mr. Shivam Bhardwaj is pursuing his B.Tech degree in Computer Science and Engineering from Sunder Deep Engineering College, Ghaziabad, affiliated to Dr. A.P.J. Abdul Kalam Technical University. His areas of interest include machine learning, artificial intelligence, Python programming, and Data science.

Mr. Deepanshu Bhardwaj is pursuing his B.Tech degree in Computer Science and Engineering from Sunder Deep Engineering College, Ghaziabad, affiliated to Dr. A.P.J. Abdul Kalam Technical University. His areas of interest include Machine Learning, Artificial Intelligence and Data Science.

Mr. Vikas Chaudhary is pursuing his B.Tech degree in Computer Science and Engineering from Sunder Deep Engineering College, Ghaziabad, affiliated to Dr. A.P.J. Abdul Kalam Technical University. His areas of interest include machine learning, artificial intelligence and Python programming.

Mr. Deepak Bisht is pursuing his B.Tech degree in Computer Science and Engineering from Sunder Deep Engineering College, Ghaziabad, affiliated to Dr. A.P.J. Abdul Kalam Technical University. His areas of interest include machine learning and artificial intelligence and IOT.

Mr. Shagun Malhotra completed his B.Tech from Guru Gobind Singh Indraprastha University, New Delhi and M.Tech from Maharshi Dayanand University, Rohtak. He is doing his PhD from Delhi Skill & Entrepreneurship University, New Delhi. He has got rich industrial experience of working for various private as well as government sectors. He has authored various research articles in many international journals. Currently, he is serving as Assistant Professor in Sunder Deep Engineering College, Ghaziabad. His fields of interest include artificial intelligence, machine learning, and IOT.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to Sunder Deep Engineering College (SDGI Global University), for providing the necessary facilities and a supportive academic environment to successfully carry out this research work. We are deeply thankful to our respected faculty members and project guides for their continuous support, valuable guidance, and encouragement throughout the development of this project. Their insights and suggestions played a significant role in the successful completion of this work. We also acknowledge the contributions of the researchers whose work has been referred to in this study, as their research provided a strong foundation for our project.

REFERENCES

[1] Chabukswar, R. R., Chavan, P. V., & Oza, K. S. (2025). Sign Language Recognition using Deep Learning. International Journal on Advanced Computer Engineering and Communication Technology, 14(1), 102–107. https://doi.org/10.65521/ijacect.v14i1.201

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 11, April 2026
Impact Factor: 8.2

[2] Pandey, Anandshankar. (2025). Real-Time Sign Language Recognition Using Deep Learning and Computer Vision. 10.13140/RG.2.2.30247.05285.
[3] Razieh Rastgoo, Kourosh Kiani, Sergio Escalera, Sign Language Recognition: A Deep Survey, Expert Systems with Applications, Volume 164, 2021, 113794, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2020.113794.
[4] Baihan, A., Alutaibi, A.I., Alshehri, M. et al. Sign language recognition using modified deep learning network and hybrid optimization: a hybrid optimizer (HO) based optimized CNNSa-LSTM approach. Sci Rep 14, 26111 (2024). https://doi.org/10.1038/s41598-024-76174-7
[5] Arsan, Taner & Ulgen, Oguz. (2015). Sign Language Converter. International Journal of Computer Science & Engineering Survey. 6. 39-51. 10.5121/ijcses.2015.6403.
[6] V. N. T. Truong, C. Yang and Q. Tran, "A translator for American sign language to text and speech," 2016 IEEE 5th Global Conference on Consumer Electronics, 2016, pp. 1-2, doi: 10.1109/GCCE.2016.7800427.