ISSN 2277-2685
IJESR/April-June. 2025/ Vol-15/Issue-2s/1065-1074
Mrs. Y. Uma *et. al.,* / International Journal of Engineering & Science Research

INDIAN SIGN LANGUAGE TO TEXT SPEECH TRANSLATION

Mrs. Y. Uma, M. Indupriya, J. Deepak, V. Kranthi Kumar, G. Madhumith,

Assistant Professor,Dept of IT,TKR College of Engineering & Technology, Meerpet,Hyderabad,

yacharamuma@gmail.com

B Tech Student, Dept of IT, TKR College of Engineering and Technology, Meerpet, Hyderabad

indupriya9144@gmail.com,Jaidideepakreddi@gmail.com,valabojukranthi@gmail.com,reddiemadhu6@gmail.com

Abstract: Communication between hearing and non-hearing individuals has always posed a challenge, especially for those who are unable to speak or hear. Indian Sign Language (ISL) plays a crucial role in enabling members of the hearing-impaired community to express themselves and interact with others. However, a major barrier persists in translating ISL gestures into a format understandable by non-signing individuals, particularly in real-time settings.

With recent advancements in computer vision and deep learning, it is now possible to automate the recognition of ISL gestures, significantly improving the communication experience. Computer vision techniques, such as wrist and hand detection, are employed to isolate hand movements from images or video frames, regardless of varying lighting conditions or camera angles. These gestures are then processed by deep learning models—primarily Convolutional Neural Networks (CNNs)—which are capable of learning and identifying intricate spatial patterns within the gesture data.

Key words- *American Sign Language (ASL),Convolutional Neural Networks (CNN),Deep Learning,Computer Vision,Sign Language Recognition,Wrist Recognition*

1. INTRODUCTION

Communication is one of the most essential aspects of human interaction. However, individuals with speech or hearing impairments often face significant challenges in expressing themselves and connecting with others. Indian Sign Language (ISL) serves as a vital communication tool for the deaf and hard-of-hearing community, enabling them to engage in meaningful social interactions. Unfortunately, there are considerable barriers to communication between ISL users and those who are unfamiliar with the language. The absence of widespread knowledge of ISL, coupled with the limited availability of trained interpreters, creates communication gaps in various settings. This project aims to address these challenges by developing an automated ISL recognition system that converts ISL gestures into text, fostering more inclusive and accessible communication for individuals with hearing impairments.

The key benefit of this project is its potential to reduce the dependency on human interpreters, making communication faster, more reliable, and available at all times. It empowers individuals with hearing impairments to communicate independently, enhancing their autonomy and social integration. Moreover, it fosters mutual understanding between hearing and non-hearing communities, creating a more inclusive and empathetic society.

In conclusion, the proposed automated ISL recognition system represents a significant step toward promoting inclusivity and independence for individuals with speech or hearing impairments. By breaking down the communication barriers that many hearing-impaired individuals face, the system enables them to participate fully in society, whether in education, the workplace, or social settings. Ultimately, this project aims to contribute to a world where accessibility and communication are universal, allowing individuals of all abilities to engage with the world around them.

2. METHODOLOGY

Related Work:

Requirement Gathering and analysis – All possible requirements of the system to be developed are captured in this phase and

ISSN 2277-2685
IJESR/April-June. 2025/ Vol-15/Issue-2s/1065-1074
Mrs. Y. Uma *et. al.,* / International Journal of Engineering & Science Research

documented in a requirement specification document.

System Design – the requirement specifications from first phase are studied in this phase and the system design is prepared. This system design helps in specifying hardware and system requirements and helps in defining the overall system architecture.

Implementation – with inputs from the system design, the system is first developed in small programs called units, which are integrated in the next phase. Each unit is developed and tested for its functionality, which is referred to as Unit Testing.

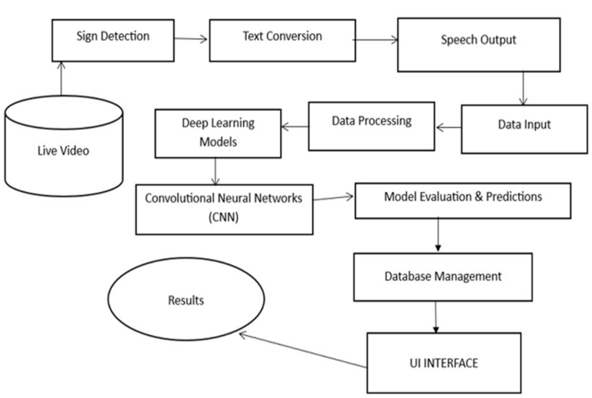
Integration and Testing – All the units developed in the implementation phase are integrated into a system after testing of each unit. Post integration the entire system is tested for any faults and failures.

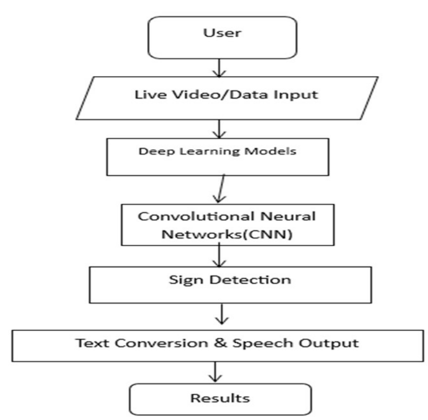
Deployment of system – Once the functional and non-functional testing is done; the product is deployed in the customer environment or released into the market.

Maintenance – There are some issues which come up in the client environment. To fix those issues, patches are released. Also, to enhance the product some better versions are released. Maintenance is done to deliver these changes in the customer environment.

.

3. IMPLEMENTATION





ISSN 2277-2685
IJESR/April-June. 2025/ Vol-15/Issue-2s/1065-1074
Mrs. Y. Uma *et. al.*, / International Journal of Engineering & Science Research

4. DATASETCOLLECTION

Training gestures: ISL Translate is a large-scale dataset developed to support research in translating Indian Sign Language (ISL) into English. It contains approximately 31,000 sentence and phrase pairs, where each ISL video clip is paired with an English translation. The dataset was curated with the goal of improving machine translation systems for ISL and includes pose features extracted using MediaPipe Holistic, which captures 3D key points of body, hand, and face movements. It is divided into training, validation, and test splits with 80%, 10%, and 10% of the data respectively. To ensure quality, a subset of 291 video-text pairs was validated by a certified ISL instructor. Benchmarking experiments using a transformer-based model showed moderate performance, with metrics such as a BLEU-4 score of 6.09 and a word error rate of 61.88%, indicating the complexity of ISL translation and the potential for further advancements.

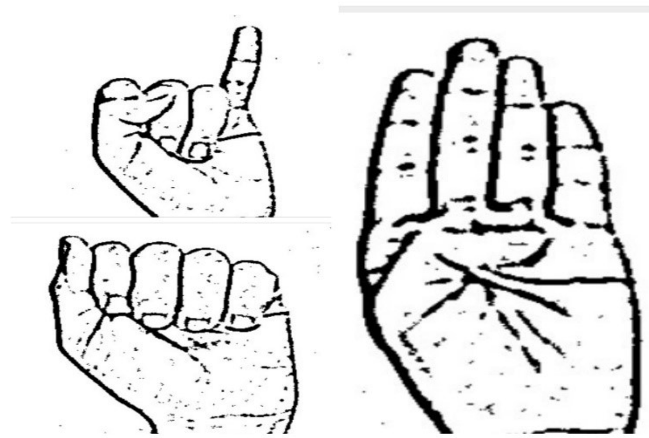


Fig 1 Classification dataset

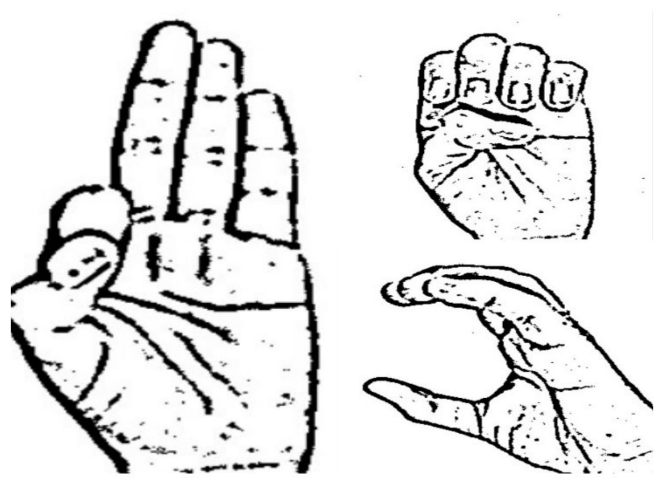


Fig 2 Detection dataset

5. ALGORITHMS

VGG and ResNet are two landmark CNN architectures that reflect different design philosophies in deep learning for computer vision. VGG, introduced in 2014, emphasizes architectural simplicity by using a series of 3×3 convolutional filters stacked on top of each other, followed by max pooling and fully connected layers. Its depth (VGG-16 and VGG-19) allows it to extract complex image features, and its uniform structure makes it easy to implement and modify. However, VGG networks are computationally expensive and require significant memory due to the large number of parameters, especially in the fully connected layers.

On the other hand, ResNet, which won the ILSVRC 2015 competition, brought a major innovation to deep learning with the concept of residual learning. Traditional deep networks often suffer from degradation problems, where adding more layers leads to worse training accuracy due to vanishing gradients. ResNet addresses this by introducing skip connections—shortcut paths that allow gradients to flow directly across layers. These residual blocks enable training of extremely deep networks, such as ResNet-152, without performance loss. This made ResNet a major advancement, not just in classification, but also in tasks like object detection (e.g., in Faster R-CNN), segmentation (e.g., in Mask R-CNN), and even in non-vision tasks.

In practice, VGG is often used as a strong baseline or a feature extractor in transfer learning due to its straightforward structure and pre-trained availability. It's especially useful when model interpretability or modularity is needed. In contrast, **R**esNet is preferred for high-performance applications and real-world deployment, as it offers better accuracy with fewer parameters and faster convergence during training. Additionally, ResNet's modular residual blocks have inspired the design of many newer architectures such as DenseNet and EfficientNet

6. EXPERIMENTAL RESULTS

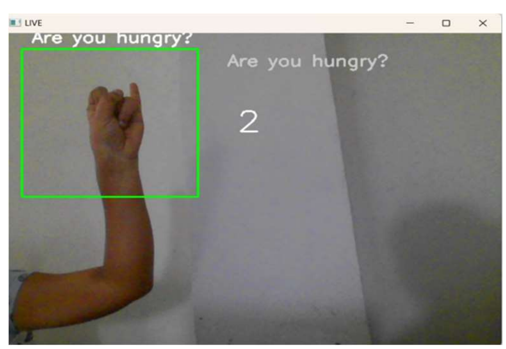
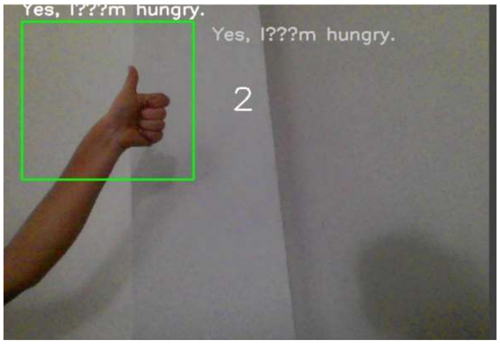


Fig 3 sign detection



The automated ISL recognition system translates gestures into text in real-time, providing an immediate and accessible form of communication for hearing-impaired individuals. Once recognized, the text can be displayed on a screen or converted into speech using text-to-speech (TTS) technology.

7. CONCLUSION

The robust performance of the system underscores its potential for deployment in real-world scenarios such as educational settings, healthcare facilities, and public service platforms. Furthermore, the solution is designed to be scalable and adaptable, allowing for future enhancements and expansion to include more complex gestures or continuous signing. With its high reliability and ease of implementation, this ISL recognition system serves as a valuable tool for empowering individuals with speech or hearing impairments—ultimately.

8. FUTURE SCOPE

Expanding the system to recognize dynamic gestures, including letters J and Z, will enhance its comprehensiveness. Incorporating real-time video recognition using continuous sign language recognition (CSLR) techniques can further improve practical usability. Additionally, integrating facial expression and body movement analysis could provide more contextual information for accurate sentence-level interpretation. Implementing multi-language sign recognition to accommodate different regional variations of sign language would broaden its accessibility. Moreover, optimizing the system for mobile and low-resource devices through model compression and quantization techniques can enhance its usability in real-world applications. Future developments may also explore incorporating natural language processing (NLP) to convert sign language gestures directly into spoken language for seamless communication. By addressing these aspects, the system can evolve into a more robust, inclusive, and user-friendly assistive technology

REFERENCES

1. Alsharif, B., Altaher, A. S., Altaher, A., Ilyas, M., & Alalwany, E. (2023). Deep Learning Technology to Recognize American Sign Language Alphabet. *Sensors*, *23*(18), 7970. https://doi.org/10.3390/s23187970
2. R. G. Rajan and M. Judith Leo, "American Sign Language Alphabets Recognition using Hand Crafted and Deep Learning Features," *2020 International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore, India, 2020, pp. 430-434, doi: 10.1109/ICICT48043.2020.9112481.
3. L. Li, D. Liu, C. Shen and J. Sun, "American Sign Language Recognition Based on Machine Learning and Neural Network," *2022 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE)*, Guangzhou, China, 2022, pp. 452-457, doi: 10.1109/MLISE57402.2022.00096.
4. Ali Salem Altaher, Chiron Bang, Bader Alsharif, Ahmed Altaher, Munid Alanazi, Hasan Altaher, Hanqi Zhuang,Mamba vision models: Automated American sign language recognition,Franklin Open,Volume 10,2025,100224,ISSN 2773-1863,https://doi.org/10.1016/j.fraope.2025.100224.
5. M. Taskiran, M. Killioglu and N. Kahraman, "A Real-Time System for Recognition of American Sign Language by using Deep Learning", *2018 41st International Conference on Telecommunications and Signal Processing (TSP)*, pp. 1-5, 2018.
6. S.I. Ito, M. Ito and M. Fukumi, "A Method of Classifying Japanese Sign Language using Gathered Image

Generation and Convolutional Neural Networks", *2019 IEEE Intl Conf on Dependable Autonomic and Secure Computing Intl Conf on Pervasive Intelligence and Computing Intl Conf on Cloud and Big Data Computing Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*, pp. 868-871, 2019.

7. V. Anuja and Bindu Nair, "A Review on Indian Sign Language Recognition", *International Journal of Computer Applications (0975 –8887)*, vol. 73, no. 22, pp. 33-38.
8. G. Ramya and N. S. Naik, "Implementation of telugu speech synthesis system", *2017 International Conference on Advances in Computing Communications and Informatics (ICACCI)*, pp. 1151-1154, 2017.
9. M. R. Reddy et al., "Transcription of Telugu TV news using ASR", *2015 International Conference on Advances in Computing Communications and Informatics (ICACCI)*, pp. 1542-1545, 2015.
10. D. Pubadi et al., "A focus on codemixing and codeswitching in Tamil speech to text", *2020 8th International Conference in Software Engineering Research and Innovation (CONISOFT)*, pp. 154-165, 2020.
11. "Voice and speech recognition in Tamil language", *2017 2nd International Conference on Computing and Communications Technologies (ICCCT)*, pp. 288-292, 2017.
12. S. C. Sajjan and C. Vijaya, "Continuous Speech Recognition of Kannada language using triphone modeling", *2016 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET)*, pp. 451-455, 2016.
13. S. Sawant and M. Deshpande, "Isolated Spoken Marathi Words Recognition Using HMM", *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, pp. 1-4, 2018.
14. L. B. Babu, A. George, K. R. Sreelakshmi and L. Mary,

"Continuous Speech Recognition System for Malayalam Language Using Kaldi", *2018 International Conference on Emerging Trends and Innovations In Engineering And Technological Research (ICETIETR)*, pp. 1-4, 2018.

15. S. Bansal and S. S. Agrawal, "Development of Text and Speech Corpus for Designing the Multilingual Recognition System", *2018 Oriental COCOSDA - International Conference*
16. Viraj Shinde, Tushar Bacchav, Jitendra Pawar, Mangesh Sanap - Hand Gesture Recognition System Using Camera, International Journal of Engineering Research & Technology (IJERT) Vol.3 Issue1, January-2014 IJERTISSN:2278-0181
17. Dave Rupit, Patel Jwoli - Hand Gesture Recognition for Computer Vision Technology, International Journal of Engineering Research & Technology (IJERT) Vol. 3 Issue 1, January - 2014 IJERT ISSN: 2278-0181
18. Mayur V. Gore, Dr. Anil R. Karwankar - Human Computer Interaction using Hand Gesture Recognition, International Journal of Engineering Research & Technology (IJERT) Vol. 3 Issue 4, April - 2014 IJERT ISSN: 2278-0181
19. Arjunlal, Minu Lalitha Madhavu - A Survey on Hand Gesture Recognition and Hand Tracking, International Journal of Scientific Engineering and Applied Science (IJSEAS) - Volume-2, Issue-1, January 2016 ISSN: 2395-3470
20. Sachin Bhat, Amruthesh M, Ashik, Chidanand Das, Sujith - Translating Indian Sign Language to text and voice messages using flex sensors, International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 5, May 2015
21. Mr. Girme R. B, Prof. Marathe V. R. - Real Time Indian Sign Language Recognition System using SVM Classifier, International Journal of Advance Engineering and Research Development Volume 5, Issue 01, January -2018
22. Mahesh Kumar N B - Conversion of Sign Language into Text, International Journal of Applied

Engineering Research ISSN 0973-4562 Volume 13, Number 9 (2018) pp. 7154-7161

23. 23. Joyeeta Singha, Karen Das - Recognition of Indian Sign Language in Live Video,International Journal of Computer Applications (0975 – 8887) Volume 70– No.19, May 2013

24. 24. Er. Vandana Soni, Dr. Tripti Arjariya - Survey Paper of Indian Sign Language Gesture System Weighted Euclidean Distance Based, International Journal of Science and Research (IJSR) ISSN (Online): 2319-7064