iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** XI **Month of publication:** November 2024

DOI: https://doi.org/10.22214/ijraset.2024.65539

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue XI Nov 2024- Available at www.ijraset.com

A Review on Advances in Indian Sign Language Recognition: Techniques, Models, and Applications

Ms.Rashmi J[1], Mr.Saurav Sahani[2], Mr.Pulkit Kumar Yadav[3]
Dept. of Computer Science and Engineering Sri Venkateshwara College of Engineering, Bengaluru – 562157

Abstract: ***Sign language serves as a crucial medium of communication for individuals with hearing and speech impairments, yet it presents barriers for those unfamiliar with its nuances. Recent advancements in artificial intelligence, computer vision, and deep learning have paved the way for innovative sign language recognition (SLR) systems. This paper provides a comprehensive survey of cutting-edge approaches to real-time sign language recognition, with a specific focus on Indian Sign Language (ISL). Various methodologies, including Convolutional Neural Networks (CNNs), Hidden Markov Models (HMMs), and transfer learning, are explored to evaluate their performance in recognizing static hand poses and dynamic gestures. Systems leveraging grid-based features, skin color segmentation, and feature extraction techniques such as Histogram of Oriented Gradients (HOG) are examined for their accuracy and efficiency. Furthermore, the integration of advanced frameworks, including pre-trained models like ResNet-50 and novel pipelines combining gesture reclassification with text-to-gesture synthesis, is discussed. This survey emphasizes the role of machine learning models, from Random Forests to state-of-the-art Large Language Models (LLMs), in addressing linguistic variability and enabling cross-linguistic translations between sign languages. By consolidating these diverse techniques, this study aims to provide valuable insights into the development of robust, real-time SLR systems that enhance accessibility, social integration, and inclusivity for the hearing-impaired community.***

Keywords: Sign Language Recognition (SLR), Indian Sign Language (ISL), American Sign Language (ASL), Gesture Recognition, Convolutional Neural Networks (CNNs), Hidden Markov Models (HMMs), Large Language Models (LLMs), Grid-Based Feature Extraction, Histogram of Oriented Gradients (HOG), k-Nearest Neighbors (k-NN), Random Forest Classifier, Text-to-Gesture Synthesis, Transfer Learning, RIFE-Net, Natural Language Processing (NLP), Assistive Technology, Real-Time Processing, Cross-Linguistic Translation, Human-Computer Interaction (HCI), Accessibility and Inclusivity.

I. INTRODUCTION

As the population in densely populated nations like India continues to grow, an increasing number of individuals are experiencing hearing impairments, highlighting the urgent need for innovative solutions to bridge communication gaps. For the deaf and hard-of-hearing community, one of the most significant barriers to social interaction and accessibility is the lack of effective communication tools. Sign languages, such as Indian Sign Language (ISL), play a vital role in enabling people with hearing impairments to express themselves and engage with the world around them. However, the understanding of ISL remains limited among the general public, leading to significant challenges in communication. Traditional sign language interpretation methods, while useful, often involve high costs, reliance on skilled interpreters, and are not always accessible in real-time, further exacerbating the communication gap.

This research aims to address these challenges by exploring the potential of cutting-edge technologies, such as machine learning and deep learning, to create an efficient and accessible ISL translation system. Specifically, the study focuses on utilizing Convolutional Neural Networks (CNNs) for the real-time recognition and translation of ISL gestures. The proposed system is designed to use a smartphone camera to capture ISL gestures, which are then processed by a server to convert the captured hand poses into meaningful text. The system is trained to recognize 33 distinct hand poses in ISL, including 10 digits and 23 letters, with an emphasis on improving accuracy and speed in gesture recognition. Notably, gestures representing letters like 'h', 'j', and 'v' are considered in the training phase, each having distinct representations in ISL.

©IJRASET: All Rights are Reserved

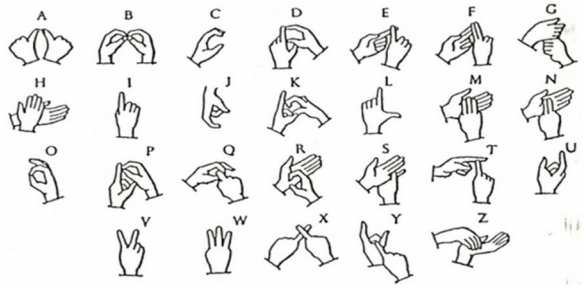


Figure 1. Indian Sign Language [11]

While existing systems for sign language translation focus on interpreting gestures from American Sign Language (ASL) or other regional sign languages, there is a notable lack of attention to ISL, which is the primary mode of communication for millions of people in India. To bridge this gap, this research focuses specifically on ISL, developing a real-time recognition system that uses CNN-based algorithms to identify hand gestures accurately. The system, capable of recognizing one-handed gestures in the initial phase, can be extended to handle more complex, two-handed gestures in the future. It also emphasizes scalability, ensuring the model can adapt to diverse dialects of ISL. Additionally, the solution aims to integrate seamlessly with existing communication platforms to facilitate real-time interactions.

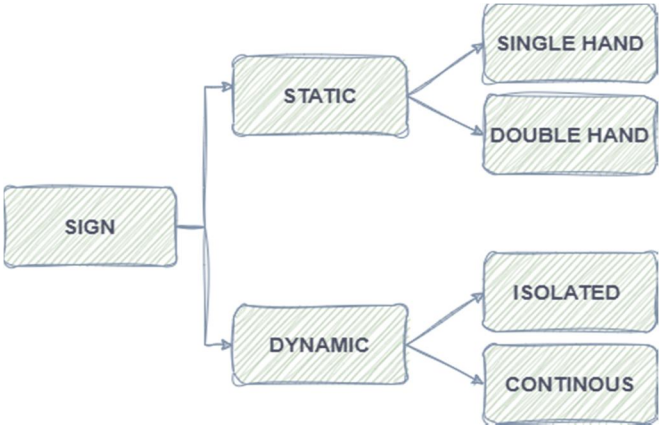


Figure 2. Classification of Sign Language [9]

The solution described in this research eliminates the discomfort associated with traditional glove-based sign language recognition systems and offers portability, making it suitable for widespread use. Through the integration of advanced technologies, the system achieves an accuracy rate exceeding 90%, ensuring that the translations are reliable and precise. By focusing on ISL, the research aims to improve communication for individuals with hearing impairments in India, providing them with a tool that enhances social integration and reduces the barriers they face in day-to-day interactions.
Additionally, the study outlines the development of an Android application designed to facilitate real-time translation of ISL gestures, allowing users to communicate seamlessly with others who do not understand the language. This system represents a significant step forward in making communication more inclusive and accessible for the deaf and hard-of-hearing community. The paper also discusses the experimental results of implementing these techniques and explores future directions, including the expansion of the system to handle more complex gestures and the potential integration of other sign languages. Ultimately, the research highlights the promise of machine learning and deep learning technologies in transforming communication for the deaf community, offering a path toward greater inclusivity and social integration.

II. LITERATURE SURVEY

Recognition and Translation of Sign language recognition and translation systems have seen significant development in recent years due to their vital role in enabling communication for the deaf and mute. Various methodologies and challenges in this domain have been identified, focusing on improving segmentation, feature extraction, classification, and translation techniques.

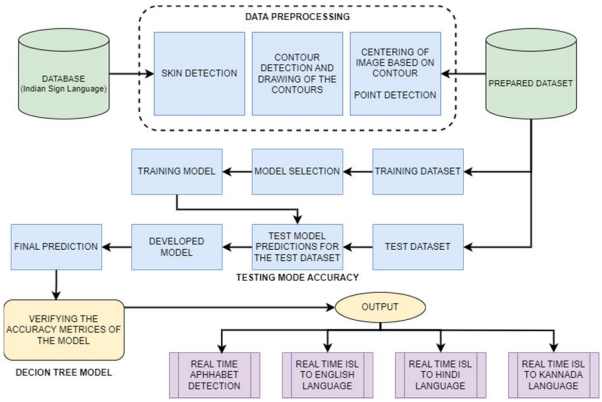


Figure 3. Proposed Methodology Flowchart

[12]Artificial Neural Network-Based ISL Recognition Adithya et al. proposed a method utilizing YCbCr color space for segmentation and applied distance transformation and Fourier descriptors for feature extraction. Their work achieved effective recognition using “central moments” for classification, demonstrating the potential of simple yet efficient feature extraction techniques. Similarly, [21] Sood et al. presented "AAWAAZ," which employed HSV histograms and the Harris algorithm for feature extraction, showcasing advancements in dataset feature matching and improving the overall accuracy of Indian Sign Language (ISL) recognition systems. [22] Halder and Tayade (2021) in a recent study that designed a real-time vernacular sign language recognition system using media pipe and machine learning propose the possibility of real-time applications in this field .[17] Sharma et al. introduced KNN classifiers and similarity measures like Cosine Similarity and Correlation, emphasizing the flexibility of traditional classifiers when combined with optimized similarity measures.[18] Sahoo et al. integrated Sobel edge detectors with neural networks for better edge feature recognition, showcasing how edge detection techniques can complement neural networks to enhance the feature representation of sign language data.

[15] Sign Language Recognition Using Deep Learning Rao et al. highlighted the role of Natural Language Toolkit (NLTK) for linguistic processing in sign translation, underlining the importance of combining linguistic tools with neural networks to achieve better translation accuracy. [13] Mubashira and James introduced Transformer Networks for video-to-text translation, achieving notable success in sequence modeling, especially in capturing the temporal dependencies between frames. This innovation demonstrated the potential of Transformer architectures in handling complex sequential data. In contrast,[20] Shin et al. demonstrated the application of Transformer-Based Neural Networks for real-time recognition, emphasizing their universality across various sign languages. Their work showcased the scalability of Transformers in real-time scenarios, paving the way for broader adoption of deep learning models in practical applications.

[14] Multimodal Approaches and Sensor-Based Recognition Pathan et al. developed a fusion approach combining image and hand landmarks using multi-headed convolutional neural networks, achieving superior results by leveraging the complementary nature of multiple modalities. This method demonstrated how integrating multimodal inputs can enhance recognition accuracy and robustness. Hardware solutions like Microsoft Kinect, SignAloud gloves, and MotionSavvy tablets have also emerged as notable techniques. These devices provide precise motion tracking and gesture recognition capabilities, yet their reliance on additional hardware limits portability and accessibility. While these systems address certain aspects of SLR, they underscore the need for more portable and cost-effective solutions.

[15] Advances in Indian Sign Language (ISL) Research Studies by Sharma and Singh employed Graph Convolutional Neural Networks (G-CNNs), surpassing the performance of traditional models like ResNet and VGG, particularly in capturing spatial and relational features within ISL data. These approaches aim to bridge gaps in Indian Sign Language (ISL) recognition using transfer learning, thereby reducing the dependency on large-scale labeled datasets. [19]Bansal et al. demonstrated the potential of modified AlexNet and VGG-16 architectures for static gesture recognition, achieving a remarkable 99.82% accuracy. Such advancements underscore the promise of tailoring existing deep learning models for ISL-specific applications, making these solutions more accessible to diverse user groups.

Emerging Techniques in Sign Language Translation Recent innovations include SIGNLLM, a large-scale multilingual SLP model generating skeletal poses for sign languages, representing a significant step toward unifying recognition and translation across multiple sign languages.

Fang's introduction of gloss-free methods reduces annotation requirements, improving system usability and scalability. These methods address the challenges of data scarcity and annotation costs, making them more practical for real-world applications. Gong emphasized the integration of Large Language Models (LLMs) to enhance translation accuracy for complex sign sequences, leveraging their contextual understanding and ability to generalize across varied input scenarios.

Limitations and Challenges despite these advancements, several challenges persist. High computational demands of deep learning models often limit their deployment in real-time applications or on resource-constrained devices. Dependence on annotated datasets, which lack heterogeneity and cultural inclusivity, hinders the generalizability of current models. Additionally, sensitivity to noise and variations in real-world environments remains a significant obstacle, affecting recognition accuracy in dynamic settings. In conclusion, while significant strides have been made in recognizing and translating sign languages, ongoing research focuses on enhancing real-time recognition, improving dataset diversity, and integrating multimodal and transfer learning techniques for robust solutions. Addressing these challenges will ensure that SLR systems become more accessible, inclusive, and practical for everyday use.

[7] This paper proposes a trainable deep learning network for isolated sign language recognition using accumulative video motion. The network consists of three networks: dynamic motion network (DMN), accumulative motion network (AMN), and sign recognition network (SRN). The DMN stream uses key postures to learn spatiotemporal information about signs, while the AMN stream generates an accumulative video motion frame. The extracted features are fused with the DMN features to be fed into the SRN for learning and classification of signs. The proposed approach is successful for isolated sign language recognition, especially for recognizing static signs. The approach was evaluated on the KArSL-190 and KArSL-502 Arabic sign language datasets, with the obtained results outperforming other techniques by 15% in the signer-independent mode. The proposed approach also outperformed state-of-the-art techniques on the Argentinian sign language dataset LSA64. Sign language is the primary communication medium for persons with hearing impairments, and the study aims to improve the recognition of sign languages using accumulative video motion.

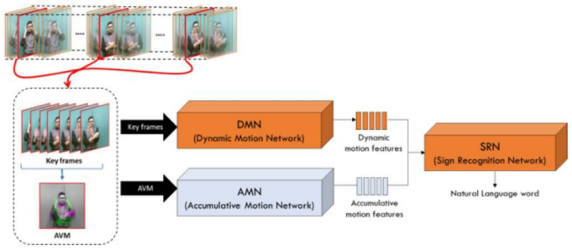


Figure 4. Framework of Sign Language Recognition system [7]

[8] Sign language is a common way of communication for people with hearing and/or speaking impairments. AI-based automatic systems for sign language recognition are desirable as they can reduce barriers between people and improve Human-Computer Interaction (HCI) for the impaired community. Automatically recognizing sign language is still an open challenge due to its complex structure to convey messages. The key role is played by isolated signs, which refer to single gestures carried out by hand movements. In the last decade, research has improved the automatic recognition of isolated sign language from videos using machine learning approaches.

The study proposes an advanced convolution-based hybrid Inception architecture to improve the recognition accuracy of isolated signs. The main contributions are to enhance InceptionV4 with optimized backpropagation through uniform connections. An ensemble learning framework with different Convolution Neural Networks has been introduced and exploited to further increase the recognition accuracy and robustness of isolated sign language recognition systems. The effectiveness of the proposed learning approaches has been proven on a benchmark dataset of isolated sign language gestures. The experimental results demonstrate that the proposed ensemble model outperforms sign identification, yielding higher recognition accuracy (98.46%) and improved robustness. Sign language is a gesture-based medium of communication and can be divided into two basic categories: static and dynamic. Static sign language mainly consists of digits, an alphabet, and some common words, while dynamic sign language is a complete foam of sign language to communicate proper meaning. Isolated sign gestures are used to express words, whereas continuous sign gesture is a visual representation of the language.In conclusion, the proposed deep learning model helps to reduce the communication barrier for impaired communities by identifying sign gestures.

The proposed novel ensemble approach is best suitable for the identification of Indian isolated sign language, which includes 11 distinct signs, based on classification efficiency and loss.

[9]Sign language is a common way of communication for people with hearing and/or speaking impairments. AI-based automatic systems for sign language recognition are desirable as they can reduce barriers between people and improve Human-Computer Interaction (HCI) for the impaired community. Automatically recognizing sign language is still an open challenge due to its complex structure to convey messages. The key role is played by isolated signs, which refer to single gestures carried out by hand movements. In the last decade, research has improved the automatic recognition of isolated sign language from videos using machine learning approaches.

The study proposes an advanced convolution-based hybrid Inception architecture to improve the recognition accuracy of isolated signs. The main contributions are to enhance InceptionV4 with optimized backpropagation through uniform connections. An ensemble learning framework with different Convolution Neural Networks has been introduced and exploited to further increase the recognition accuracy and robustness of isolated sign language recognition systems. The effectiveness of the proposed learning approaches has been proven on a benchmark dataset of isolated sign language gestures. The experimental results demonstrate that the proposed ensemble model outperforms sign identification, yielding higher recognition accuracy (98.46%) and improved robustness. Sign language is a gesture-based medium of communication and can be divided into two basic categories: static and dynamic. Static sign language mainly consists of digits, an alphabet, and some common words, while dynamic sign language is a complete foam of sign language to communicate proper meaning. Isolated sign gestures are used to express words, whereas continuous sign gesture is a visual representation of the language.In conclusion, the proposed deep learning model helps to reduce the communication barrier for impaired communities by identifying sign gestures. The proposed novel ensemble approach is best suitable for the identification of Indian isolated sign language, which includes 11 distinct signs, based on classification efficiency and loss.

[10] Sign language recognition (SLR) is an important field of research aimed at bridging the communication gap between the deaf community and hearing individuals. The World Health Organization (WHO) estimates that by 2021, 430 million people worldwide required medical rehabilitation for hearing loss, a number projected to reach 2.5 billion by 2050. Despite their unique way of communicating through sign language, the deaf community often faces integration challenges due to language barriers. Consequently, SLR has become essential for fostering inclusivity, enabling hearing individuals to understand the intentions of the deaf. This review examines traditional and deep learning-based SLR methods, exploring their techniques, datasets, challenges, and potential research directions.Traditional SLR methods often rely on sensor-based systems like data gloves, which capture hand gestures and body movements, or image processing techniques using cameras and color gloves. These systems extract features from gestures and apply temporal models such as Hidden Markov Models (HMM) and Conditional Random Fields (CRF) to recognize sequences of signs. While these methods laid the foundation for SLR, they are constrained by environmental factors such as lighting conditions, occlusion, and scalability issues when applied to large datasets. Traditional approaches are particularly effective for controlled environments but struggle in real-world applications.

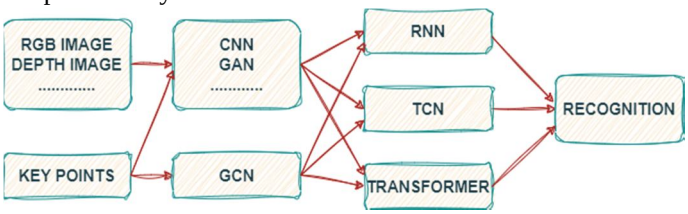


Figure 5. The process of the sign language recognition based on deep learning. [10]

The advent of deep learning has transformed SLR, enabling the recognition of both manual components, such as hand shapes and orientations, and non-manual components, like eye gaze and facial expressions. Deep learning-based SLR involves pre-processing, feature extraction, temporal modeling, and classification using advanced architectures such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Transformer-based models. These methods have shown remarkable performance in both isolated sign language recognition (ISLR) and continuous sign language recognition (CSLR). The ability to leverage larger datasets has significantly improved recognition accuracy and robustness in real-world scenarios.Despite these advancements, challenges remain in SLR. Signer independence—ensuring models perform well across diverse individuals—continues to be a significant hurdle. Environmental factors such as lighting variability and occlusions also impact recognition performance.

Moreover, the computational complexity of deep learning models raises concerns about scalability and deployment in resource-constrained settings.In conclusion, while SLR has advanced considerably with deep learning methods outperforming traditional approaches, practical applications remain challenging. Future research must focus on developing lightweight, scalable models that address signer independence and real-world variability. This will ensure that SLR can effectively bridge communication gaps, fostering inclusivity and improving the quality of life for the deaf community.

III. CONCLUSION

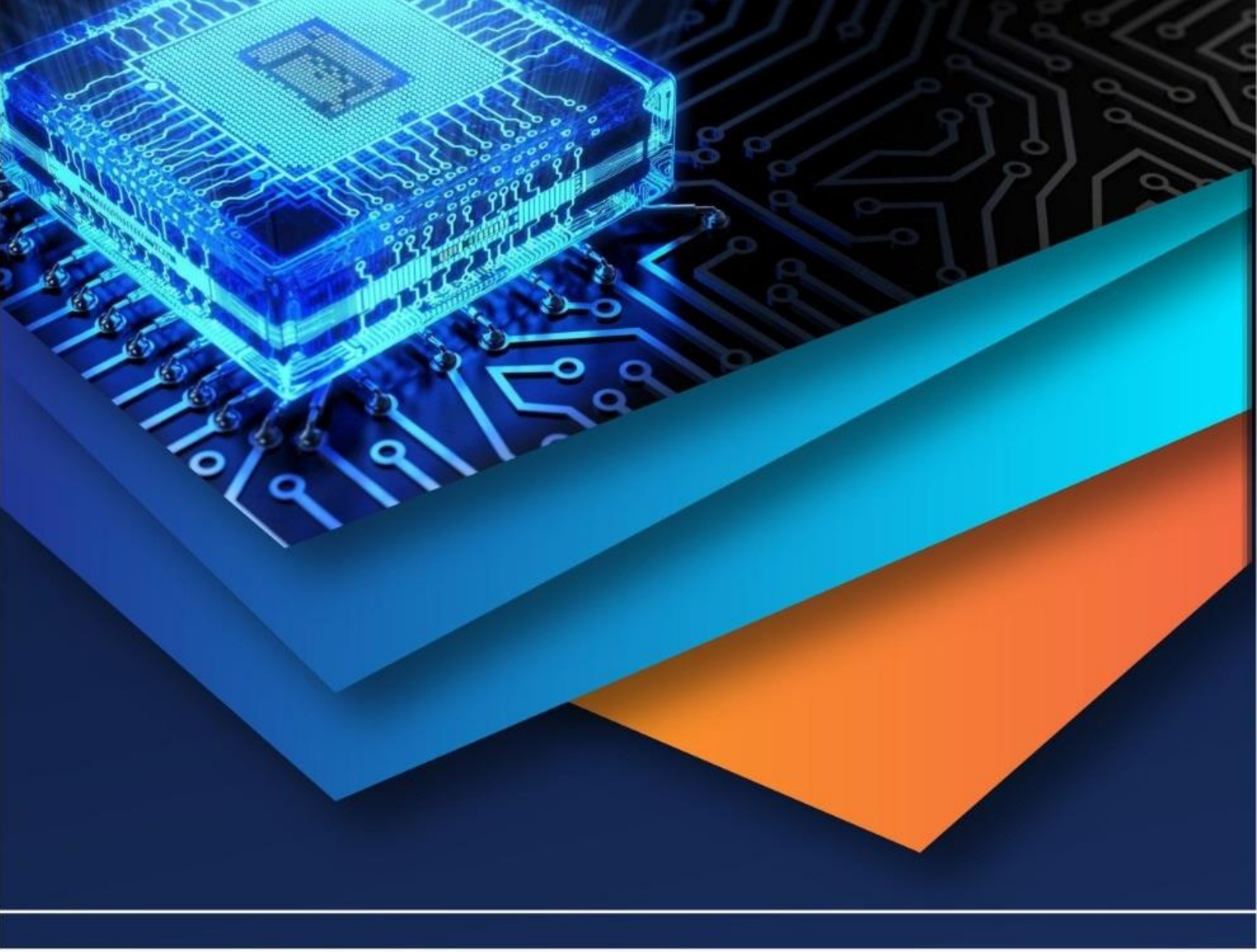
Sign language recognition and translation have undergone significant advancements, driven by the need to bridge communication gaps for the deaf and mute community. Techniques such as segmentation, feature extraction, classification, and translation have seen continuous improvements, leveraging artificial neural networks, deep learning, and multimodal approaches. Methods like YCbCr color space, Sobel edge detection, and HSV histograms have enhanced feature extraction, while advanced models, including CNNs, RNNs, G-CNNs, and Transformers, have improved recognition accuracy and robustness. Real-time systems employing Transformer networks and MediaPipe have demonstrated practical applications, while sensor-based hardware solutions, though effective, face limitations in portability and accessibility. Emerging approaches, like SIGNLLM and gloss-free methods, address data scarcity and annotation costs, offering scalability across diverse languages. Multimodal techniques combining image, hand landmark, and motion data have further enhanced recognition by leveraging complementary modalities. Studies have also demonstrated the effectiveness of tailored architectures like AlexNet and InceptionV4 in achieving near-perfect recognition accuracy for isolated sign gestures. Despite notable achievements, challenges persist, including computational demands, dependence on annotated datasets, environmental sensitivity, and signer independence, limiting real-world applicability. Cultural inclusivity in datasets remains a major hurdle, as existing resources often lack diversity to generalize across global contexts. Furthermore, issues like scalability for continuous sign language recognition and accurate modeling of non-manual features, such as facial expressions, continue to impede progress. Research must also focus on lightweight models for resource-constrained environments, reducing hardware dependence to ensure universal accessibility. The integration of linguistic tools like NLTK and large language models has shown promise in improving translation accuracy, further broadening the scope of these systems. Addressing these challenges through scalable, real-time, and culturally inclusive models will ensure SLR systems become robust, accessible, and transformative, fostering inclusivity and significantly improving the quality of life for impaired communities worldwide.

REFERENCRES

[1] Real-time Indian Sign Language (ISL) Recognition. Kartik Shenoy, Tejas Dastane, Varun Rao, Devendra Vyavaharkar. . 9th ICCCNT 2018 July 10-12, 2018, IISC, Bengaluru Bengaluru, India
[2] Real Time Indian Sign Language Recognition using Convolutional Neural Network. Snehal Hon, Manpreet Sidhu, Sandesh Marathe, Tushar A. Rane. 2024 IJNRD |Volume 9, Issue 2 February 2024|ISSN: 2456-4184 |IJNRD.ORG
[3] Enhanced Sign Language Translation between American Sign Language (ASL) and Indian Sign Language (ISL) Using LLMs Malay Kumar, S. Sarvajit Visagan, Tanish Sarang Mahajan, and Anisha Natarajan. arXiv:2411.12685v1 [cs.CL] 19 Nov 2024
[4] Saini B, Venkatesh D, Chaudhari N, et al. (2023) A comparative analysis of Indian sign language recognition using deep learning models. Forum for Linguistic Studies 5(1): 197–222. DOI: 10.18063/fls.v5i1.1617.
[5] Pathak, Aman & Kumar|priyam|priyanshu, Avinash & Chugh, Gupta|gunjan & Ijmtst, Editor. (2022). Real Time Sign Language Detection. International Journal for Modern Trends in Science and Technology. 8. 32-37. 10.46501/IJMTST0801006.
[6] Rokade, Yogeshwar & Jadav, Prashant. (2017). Indian Sign Language Recognition System. International Journal of Engineering and Technology. 9. 189-196. 10.21817/ijet/2017/v9i3/170903S030.
[7] H. Luqman, "An Efficient Two-Stream Network for Isolated Sign Language Recognition Using Accumulative Video Motion," in IEEE Access, vol. 10, pp. 93785-93798, 2022, doi: 10.1109/ACCESS.2022.3204110.
[8] D. R. Kothadiya, C. M. Bhatt, H. Kharwa and F. Albu, "Hybrid InceptionNet Based Enhanced Architecture for Isolated Sign Language Recognition," in IEEE Access, vol. 12, pp. 90889-90899, 2024, doi: 10.1109/ACCESS.2024.3420776.
[9] D. R. Kothadiya, C. M. Bhatt, T. Saba, A. Rehman and S. A. Bahaj, "SIGNFORMER: DeepVision Transformer for Sign Language Recognition," in IEEE Access, vol. 11, pp. 4730-4739, 2023, doi: 10.1109/ACCESS.2022.3231130.
[10] T. Tao, Y. Zhao, T. Liu and J. Zhu, "Sign Language Recognition: A Comprehensive Review of Traditional and Deep Learning Approaches, Datasets, and Challenges," in IEEE Access, vol. 12, pp. 75034-75060, 2024, doi: 10.1109/ACCESS.2024.3398806.
[11] https://www.1specSialplace.com/2021/02/11/all-about-indian-sign-language/
[12] Adithya, V. & Vinod, P.R. & Gopalakrishnan, Usha. (2013). Artificial neural network based method for Indian sign language recognition. 1080-1085. 10.1109/CICT.2013.6558259.
[13] M. N and A. James, "Transformer Network for video to text translation," 2020 International Conference on Power, Instrumentation, Control and Computing (PICC), Thrissur, India, 2020, pp. 1-6, doi: 10.1109/PICC51425.2020.9362374.

[14] R. K. Pathan, et al., ”Sign Language Recognition Using the Fusion of Image and Hand Landmarks Through Multi-Headed Convolutional Neural Network,” Scientific Reports, vol. 13, 2023, p. 16975. doi: 10. 1038/s41598-023-43852-x
[15] V. Singh and A. Sharma, ”A Review on Sign Language Recognition Techniques,” Journal of Computer Science and Applications, vol. 19, no. 1, pp. 22–30, 2021.
[16] P. S. Rao, et al., ”Multiple Languages to Sign Language Using NLTK,” International Journal of Scientific Research in Science and Technology, vol. 10, no. 2, pp. 12–17, Mar.–Apr. 2023. [Online]. Available: https: //doi.org/10.32628/IJSRST2310189
[17] P. Sharma, et al., ”Translating Speech to Indian Sign Language Using Natural Language Processing,” Future Internet, vol. 14, no. 9, p. 253, 2022. doi: 10.3390/fi14090253
[18] A. Sahoo, G. Mishra, and K. Ravulakollu, ”Sign Language Recognition: State of the Art,” ARPN J. Eng. Appl. Sci., vol. 9, 2014, pp. 116–134.
[19] R. Bansal and S. Bansal, ”A Review on Sign Language Recognition Techniques,” International Journal of Advanced Research in Computer Science, vol. 11, no. 2, pp. 123-128, Feb. 2020.
[20] J. Shin, et al., ”Korean Sign Language Recognition Using Transformer Based Deep Neural Network,” Applied Sciences, vol. 13, no. 5, p. 3029, 2023. doi: 10.3390/app13053029
[21] Sood, Anchal & Mishra, Anju. (2016). AAWAAZ: A communication system for deaf and dumb. 620-624. 10.1109/ICRITO.2016.7785029.
[22] A. Halder and A. Tayade, ”Real-time Vernacular Sign Language Recog nition Using MediaPipe and Machine Learning,” Int. J. Res. Proj. Res., vol. 2, no. 3, 2021. [Online]. Available: www.ijrpr.com

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)