WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

A Comparative Study on Helping Hearing-Impaired and Mutes Using Sign Language Translator

Alen Joy[1], Aryan B[2], Balasubramanian G[3], Richu Suresh[4], Dr. Lino Abraham Varghese[5]

[1234] 4th Year Student, [5] HOD & Professor

Department of Computer Science and Engineering, Ilahia College of Engineering And Technology Mulavoor 686673 India

A B S T R A C T

This paper introduces a real-time sign language translator aimed at facilitating communication for individuals who are hearing impaired or mute. The translator employs advanced deep learning techniques to accurately interpret sign language gestures and translate them into spoken or written language in real-time. Through the utilization of cutting-edge deep learning models, including Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), the system can effectively analyze and understand the intricate movements inherent in sign language. This technology enables comprehensive communication between individuals who use sign language and those who do not, breaking down communication barriers and promoting inclusivity in various settings such as conversations, presentations, and emergency situations. The real-time functionality of the translator enhances accessibility, fostering greater engagement and participation for individuals with hearing impairments or who are mute in society. This technique uses sign language as hand gesture and we use cameras to detect their hand signs using object detection techniques and appropriate output is obtained. Human computer interaction between people also increases thereby connecting all specially-abled people together.

Keywords: Human–computer interaction, sign language, hand gesture recognition, classification, object detection.

1. Introduction

Motion acknowledgment innovation is progressing consistently, encouraging critical effects on human-PC cooperation. Hand gesture recognition systems are becoming more accurate and efficient as AI and computer vision progress, driving innovation across industries. For instance, in augmented simulation, coordinating hand motion acknowledgment with haptic criticism and eye-following advancements could make more vivid encounters. Users may be able to interact with virtual objects through eye movements and tactile feedback as well as control them with their hands thanks to this integration. Likewise, in mechanical technology, joining hand signal acknowledgment with regular language handling and feeling acknowledgment could empower robots to grasp the client's aim and close to home state. This could work with additional caring and regular communications, helping fields like medical services and client assistance. Besides, motion acknowledgment innovation is probably going to see expanded use in wellbeing delicate fields like medical services and car. In independent vehicles, for example, hand motion acknowledgment could guarantee smooth and safe human-vehicle connection, while in medical care, signal based points of interaction could assist clinical experts with getting to patient data or control clinical gadgets without compromising sterility. In training, coordinating motion acknowledgment with versatile learning frameworks could customize instructive encounters further. Dissecting an understudy's signals during recreations could assist with changing trouble levels or give explicit input, improving the growing experience. By and large, progressions in motion acknowledgment innovation offer colossal potential for improving communications with PCs and advanced gadgets across different spaces. By empowering more normal, natural, and vivid encounters, these advancements can upgrade efficiency, imagination, and availability in exceptional ways. As scientists and architects keep on pushing limits, the fate of signal acknowledgment seems promising and groundbreaking.

1.1 Sign Language Recognition Using Artificial Intelligence

Real-time sign language translation using AI is a groundbreaking advancement that aims to bridge communication gaps between sign language users and Real-time sign language translation using AI represents a significant breakthrough in bridging communication barriers between sign language users and those unfamiliar with it. This application of artificial intelligence relies on advanced algorithms and machine learning to instantly interpret sign language gestures, converting them into spoken language or text. Through the utilization of computer vision and deep learning techniques like Convolutional Neural Networks (CNNs) and recurrent models, the technology accurately recognizes and analyzes the subtle movements and expressions inherent in sign language, ensuring precise translations. The potential impact of this innovation is vast, promising to greatly enhance accessibility and communication for the deaf and hard-of-hearing communities. By enabling seamless communication with others, it promotes inclusivity across various domains such as education, healthcare, customer service, and social interactions. This empowerment allows sign language users to fully participate in diverse environments and activities. However, real-time sign language translation using AI faces inherent challenges. Variability in signing styles and regional dialects poses difficulties in creating universally accurate models, potentially leading to misinterpretations. The complexity of sign language gestures, especially capturing subtle movements in real-time, requires sophisticated algorithms that may struggle with rapid or nuanced gestures. Additionally, limited diversity in training datasets and reliance on contextual cues and facial expressions further complicate accurate interpretation. Addressing privacy concerns related to camera usage for real-time recognition is crucial, necessitating secure and privacy-respecting implementations. Moreover, the continuous evolution of sign languages and the hardware requirements for real-time translation pose ongoing challenges that require diverse datasets and concerted efforts to enhance adaptability, accuracy, and inclusivity across various contexts and activities. Despite these challenges, the advancement of real-time sign language translation using AI holds immense promise for transforming communication and fostering greater inclusion for all.

1.2 Challenges In Deep Learning Techniques

Continuous gesture-based communication interpretation for the conference hindered utilizing profound learning experiences various difficulties. One critical deterrent includes the different marking styles and provincial lingos inside communications through signing, making it complex to adjust profound learning models for exact understanding. Besides, making extensive and differed datasets that envelop the extensive variety of communication via gestures articulations is asset concentrated. Moreover, proficient and improved calculations are vital for ongoing handling, taking into account computational asset requirements. Misinterpretations might emerge, particularly with fast or unobtrusive developments, possibly upsetting consistent correspondence. Also, guaranteeing framework precision across various settings, like low-light circumstances or complex foundations, adds intricacy. To overcome these obstacles, ongoing research and development is required to improve real-time sign language translation accessibility for the hearing-impaired community, diversify datasets, and refine deep learning models.

2. Literature Review

Numerous investigations have delved into the complexities and progressions in gesture recognition and interpretation, striving to enhance communication for individuals with hearing impairments. One notable study conducted by Qazi Muhammed Areeb et al. [1] is dedicated to enabling real-time communication through the recognition and interpretation of sign language using smartphones equipped with cameras. Their approach harnesses advanced deep learning techniques such as YOLO (You Only Look Once) and Convolutional Neural Networks (CNNs) to accurately detect and track sign language gestures in video streams, with a particular focus on emergency situations. Through the deployment of state-of-the-art algorithms, their system adeptly translates these sign language gestures into textual representations, thus significantly improving accessibility and inclusivity for the hearing-impaired community.

Jii Yan Han et al. introduce an innovative solution aimed at enhancing speech perception for individuals using hearing aids. Their method revolves around a smartphone application designed to ameliorate speech sounds captured by a microphone and subsequently relay them to the hearing aid via Bluetooth [II]. Central to their approach is the utilization of a Deep Neural Network (DNN) for speech enhancement, coupled with feature extraction facilitated by a deep denoising autoencoder (DDAE). This sophisticated system offers users the advantage of real-time speech enhancement delivered through an intuitive and user-friendly interface. By leveraging this technology, the accessibility of communication is significantly improved, particularly for those individuals with partial hearing impairment, thereby fostering greater inclusivity and ease of interaction in various social and environmental settings

Yanying Zhang et al. delve into an extensive examination of image processing deep learning algorithms, including Particle Swarm Optimization (PSO), Scale-Invariant Feature Transform (SIFT), and Convolutional Neural Networks (CNN) [III]. Through their meticulous analysis, they scrutinize the efficacy of different digital image processing techniques, shedding light on the unparalleled effectiveness of CNN in tasks related to image recognition and feature extraction. Their research serves as a compelling testament to the pivotal role that deep learning plays in advancing the field of image processing. By emphasizing the superiority of CNN over traditional methods, Zhang et al. underscore the transformative potential of deep learning algorithms in tackling complex image analysis tasks with unparalleled accuracy and efficiency. This elucidation not only enriches our understanding of modern image processing methodologies but also highlights the burgeoning significance of deep learning techniques in reshaping the landscape of visual data analysis and interpretation.

Muneer Al-Hammadi and colleagues introduce a groundbreaking approach centered around deep learning principles aimed at recognizing sign language gestures, ultimately fostering improved communication between the hearing and hearing-impaired communities [IV]. Their innovative framework incorporates the use of 3D Convolutional Neural Networks (3DCNN) for video modeling, specifically targeting the intricate nuances of hand signals inherent in sign language. Through the strategic application of deep learning methodologies, they significantly enhance the accuracy and efficiency of sign language recognition. By harnessing the power of neural networks, Al-Hammadi et al. not only advance the technological frontier of sign language interpretation but also contribute to the broader societal goals of inclusivity and accessibility. Their pioneering work stands as a testament to the transformative potential of deep learning in bridging communication gaps and promoting a more inclusive society for individuals with hearing impairments.

Bhat Geetalaxmi Jiram and their team embark on a significant endeavor, employing Convolutional Neural Networks (CNNs) and deep learning methodologies to tackle the challenges associated with real-time sign language translation and capture [V]. Their research underscores the critical importance of addressing the complexity surrounding user training, identifying it as a crucial research gap. By emphasizing the necessity of effective user training protocols, they highlight the pivotal role it plays in ensuring the accuracy of sign language recognition systems. In their exploration, Jiram et al. shed light on the intricate process of training users to interact seamlessly with sign language translation technology. They stress the significance of tailored training approaches that account for individual differences in signing styles and gestures. Through their findings, they advocate for the development of user-friendly training modules that can adapt to diverse user needs and preferences. By acknowledging the challenges associated with user training, Jiram et al. underscore the importance of human-centric design principles in the development of sign language translation systems. Their research not only advances the technical capabilities of CNN-based models but also highlights the need for inclusive and user-centered approaches in technology design. Ultimately, their work paves the way for more effective and accessible sign language translation solutions that can better serve the needs of both hearing and hearing-impaired communities

Hamza Luqman and his collaborators introduce a pioneering approach in their research [VI], unveiling a hierarchical sign learning module meticulously crafted to tackle the nuances arising from variations in sign samples across different signers, specifically focusing on isolated sign language recognition. At the heart of their innovative framework lies a sophisticated deep learning network, strategically designed to incorporate interconnected networks aimed at capturing the intricate spatiotemporal characteristics inherent in sign gestures. Central to their methodology is the recognition that sign language, much like spoken language, exhibits considerable variability depending on factors such as regional dialects, individual signing styles, and contextual factors. To address this challenge, Luqman et al. propose a hierarchical structure within their learning module, enabling the system to adapt and learn from diverse sign samples encountered during training. By leveraging the power of deep learning, their network can effectively discern subtle variations in sign gestures, ultimately contributing to more accurate and robust gesture-based communication recognition. This breakthrough not only enhances the performance of sign language recognition systems but also holds significant promise for fostering more inclusive communication environments for individuals with hearing impairments.

Gautam S Bhat and his collaborators introduce a groundbreaking solution in their research [VII], unveiling a real-time speech enhancement method tailored specifically for users of hearing aids. Their innovative approach revolves around the implementation of a multi-objective learning Convolutional Neural Network (CNN) model, ingeniously integrated into a smartphone application. This application serves as a practical and accessible tool for individuals seeking to improve their hearing experience in various environments. Central to their methodology is the recognition of the pervasive challenge posed by background noise in everyday communication scenarios. By harnessing the power of deep learning, Bhat et al. have developed a system capable of effectively reducing noise in speech signals, thereby enhancing the overall quality and intelligibility of speech for users of hearing aids. Their real-time speech enhancement method represents a significant advancement in the field, offering a practical solution to a longstanding problem faced by individuals with hearing impairments. By seamlessly integrating their CNN model into a smartphone application, Bhat and his team have democratized access to state-of-the-art speech enhancement technology, empowering users to engage more confidently and comfortably in conversations across a wide range of noisy environments

In their recent literature review, Panneer Selvam E and colleagues delve into the forefront of gesture recognition technology, particularly emphasizing its pivotal role in addressing the pressing need for improved accessibility and communication channels for individuals with hearing impairments [VIII]. Through a thorough examination of current advancements, they underscore the significance of ongoing research efforts aimed at refining gesture recognition systems to cater to the unique needs of this community. By highlighting the latest strides in neural network models and gesture recognition techniques, Selvam et al. illuminate the evolving landscape of this field, showcasing the relentless pursuit of inclusivity and user-centric design. Their review underscores the transformative potential of gesture recognition technology in dismantling communication barriers and fostering greater accessibility for individuals with hearing impairments. Overall, their comprehensive overview provides valuable insights into the state-of-the-art in gesture recognition technology, reaffirming the importance of continued research and development to ensure that these technologies effectively serve the needs of diverse user populations, particularly those with hearing impairments.

In their comprehensive literature review, Ashin and colleagues undertake a meticulous examination of Convolutional Neural Networks (CNNs), focusing on their definition and practical applications within the realm of image classification [IX]. Central to their analysis is a dedicated effort to harness modern technologies, with a particular emphasis on leveraging widely adopted libraries such as TensorFlow, Keras, and Theano to implement advanced multi-image classification systems. The review delves deep into the theoretical foundations of CNNs, providing a thorough exploration of key concepts including convolutional layers, pooling layers, and fully connected layers. Through clear and concise exposition, Ashin et al. establish a robust understanding of these fundamental components, laying a solid groundwork for comprehending the intricate workings of CNNs in the context of image classification tasks. Furthermore, the review transcends theoretical discourse to offer practical insights into the implementation of CNN-based models. By harnessing the capabilities of TensorFlow, Keras, and Theano – three prominent deep learning libraries – the authors demonstrate how researchers and practitioners can effectively leverage these tools to develop resilient image classification systems capable of handling diverse datasets of varying complexities.

B Natarajan and colleagues tackle the intricate challenges surrounding gesture recognition and interpretation in their recent research [X]. Central to their contribution is the introduction of a novel system that integrates MediaPipe with a hybrid model, aiming to enhance both accuracy and visual quality in gesture recognition. Their innovative approach not only addresses the technical complexities inherent in gesture recognition but also emphasizes the importance of visual fidelity in communication. By leveraging the combined capabilities of MediaPipe and their hybrid model, Natarajan et al. demonstrate a significant improvement in gesture recognition accuracy and the generation of visually appealing videos. The effectiveness of their model is particularly noteworthy as it serves to bridge the commnication gap between individuals who are hearing impaired and the hearing world.

3. Comparison Table

In alignment with the objectives of our project, our team conducted a meticulous literature review, analyzing 10 IEEE papers. Subsequently, we performed a comparative table, contrasting the aforementioned papers with our proposed framework. The analysis encompasses various aspects, including image enhancement, motion identification, feature extraction, sign language recognition, object detection, two-way communication involving sign-to-text and speech-to-sign language, multilingual capability, cost-effectiveness, CNN training, and accuracy. Our proposed system distinguishes itself from the competition due to these features.

Table1: Comparison Table

FEATURE	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	PROPOSED SYSTEM
IMAGE ENHANCEMENT	Capturing only	No	Yes	Yes	Medium	Yes	No	No	Yes	Yes	Yes
GESTURE IDENTIFICATION	Yes	No	No	Yes	Yes	Yes	No	Yes	No	Yes	Yes
IMAGE EXTRACTION	Inefficient	No	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
SIGN LANGUAGE RECOGNITION	Yes	No	No	Yes	Yes	Yes	No	Yes	No	Yes	Yes
OBJECT DETECTION	No	No	Yes	No	Yes	No	No	No	Yes	No	Yes
TWO WAY COMMUNICATION	Only sign to text	Yes It is possible via hearing aid	No	Yes	No	Yes	Yes It is possible by normal earphones	No	No, It is used for object detection	No	Yes
MULTILINGUAL CAPABILITY	American Sign Language only	It only amplifies the sound from the mic	Used for object detection	American Sign Language only	American Sign Language only	Yes	It enhances the sound around the person	American and British Language only	Used for object detection	Yes	Yes
COST EFFICIENCY	Medium	Low	High	Low	Medium	Medium	High	Low	Medium	High	High
CNN TRAINING	Yes	Yes	Yes	Yes	Yes	DMN, AMN, SRN	Yes	Yes	Yes	Yes	Yes
ACCURACY	Less	Medium	Medium	High	Less	High	High	High	High	High	High

3.1 Overview

Continuous correspondence by means of motion translation using profound learning is a state-of-the-art application that tackles the force of complex brain organizations to connect correspondence boundaries between people utilizing communication via gestures and the individuals who may not be capable in it. Counterfeit brain networks are prepared to perceive examples and make forecasts utilizing enormous datasets in profound learning, a subset of AI. Deep learning algorithms like Convolutional Neural Networks (CNNs) and recurrent models are used to analyze and interpret the intricate movements and expressions that are inherent in sign language gestures in the context of sign language interpretation. The interaction normally includes the utilization of PC vision strategies to catch and comprehend the visual data related with communication via gestures. Diverse datasets containing a wide variety of sign language expressions and signals are used to train the deep learning model. These datasets empower the model to become familiar with the perplexing mappings between visual information sources and comparing etymological implications, permitting it to make precise forecasts continuously situations. The capacity of deep learning to adapt and generalize to a variety of signing styles and dialects is one significant advantage of using it for real-time sign language interpretation. The various leveled nature of profound brain networks empowers them to learn progressive highlights, catching both fine subtleties and dynamic portrayals of gesture-based communication signals. When dealing with the variety of sign languages, this adaptability is absolutely necessary.

4. Conclusion

In summary, the advent of real-time sign language translation, powered by deep learning algorithms, signifies a monumental stride in assistive technology tailored for individuals with hearing impairments. Through the utilization of sophisticated deep learning models like Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), this innovative translator excels in accurately decoding and converting sign language gestures into spoken or written language instantaneously. Such a breakthrough holds immense promise in dismantling communication barriers, thereby granting individuals who are hearing impaired unprecedented access to seamless interaction in various social and professional settings.

By seamlessly integrating deep learning techniques into the translation process, this technology not only ensures accuracy and efficiency but also fosters inclusivity and empowerment. The real-time capabilities of the translator enable individuals with hearing impairments to actively participate in conversations, presentations, and critical situations, thereby enhancing their overall engagement and integration within society. Furthermore, as research and development in this field progress, ongoing refinement and enhancement of real-time sign language translation systems will continue to amplify accessibility and promote greater inclusivity for individuals with hearing impairments worldwide. Thus, the journey towards achieving comprehensive and effective communication solutions for the hearing-impaired community is propelled forward by the remarkable advancements in deep learning-based sign language translation technology. In summary, real-time sign language recognition using deep learning and CNNs holds significant promise for the hearing-impaired community. It represents a significant step towards breaking communication barriers, fostering inclusivity, and improving the overall quality of life for individuals with hearing impairments. Innovation and research in this area must continue, as they will lead to more accessible, accurate, and adaptable solutions that will truly enhance the lives of people who use sign language to communicate.

References

[I]. Q. M. Areeb, Maryam, M. Nadeem, R. Alroobaea and F. Anwer, "Helping Hearing-Impaired in Emergency Situations: A Deep Learning-Based Approach," in IEEE Access, vol. 10, pp. 8502-8517, 2022, doi: 10.1109/ACCESS.2022.3142918.

[II]. J. Y. Han, W. Z. Zheng, R. J. Huang, Y. Tsao and Y. H. Lai, "Hearing aids APP design based on deep learning technology," 2018 11th International Symposium on Chinese Spoken Language Processing (ISCSLP), Taipei, Taiwan, 2018, pp. 495-496, doi: 10.1109/ISCSLP.2018.

[III]. Y. Zhang and X. Zheng, "Development of Image Processing Based on Deep Learning Algorithm," 2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), Dalian, China, 2022, pp. 1226-1228, doi: 10.1109/IPEC54454.2022.9777479.

[IV]. M. Al-Hammadi et al., "Deep Learning-Based Approach for Sign Language Gesture Recognition With Efficient Hand Gesture Representation," in IEEE Access, vol. 8, pp. 192527-192542, 2020, doi: 10.1109/ACCESS.2020.3032140

[V] B. G. Jairam and D. Ponnappa, "Gesture Based Virtual Assistant For Deaf-Mutes Using Deep Learning Approach," 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2023, pp. 1-7, doi: 10.1109/ICACCS57279.2023.10112986.

[VI]. 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.

[VII]. G. S. Bhat, N. Shankar and I. M. S. Panahi, "Automated machine learning based speech classification for hearing aid applications and its real-time implementation on smartphone," 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 2020, pp. 956-959, doi: 10.1109/EMBC44109.2020.

[VIII] .S. E. Panneer and M. Sornam, "Recent Advances in Sign Language Recognition using Deep Learning Techniques," 2022 6th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2022, pp. 1261-1265, doi: 10.1109/ICOEI53556.2022.9777104.

[IX]. A. Ashin, P. D. Rathika, Y. Mahavidhya, N. Hemapriya and S. S. Gowri, "Image Classification in the Era of Deep Learning," 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), Coimbatore, India, 2021, pp. 1-5, doi: 10.1109/ICAECA52838.2021.9675614

[X]. B. Natarajan et al., "Development of an End-to-End Deep Learning Framework for Sign Language Recognition, Translation, and Video Generation," in IEEE Access, vol. 10, pp. 104358-104374, 2022, doi: 10.1109/ACCESS.2022.3210543