IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 12, Issue 3, March 2025

DOI: 10.17148/IARJSET.2025.12301

Gesture based real-time American Sign Language Translator

D Guna Karthikeya[1], Keerthi Teja N[2], Nivedh A[3], Sriraj S R[4]

1DT21AI0012, 8th Sem, Dept. of AIML, Dayananda Sagar Academy of Technology and Management[1]

1DT21AI0031,8th Sem, Dept. of AIML, Dayananda Sagar Academy of Technology and Management[2]

1DT21AI041, 8th Sem, Dept. of AIML, Dayananda Sagar Academy of Technology and Management[3]

1DT21AI056, 8th Sem, Dept. of AIML, Dayananda Sagar Academy of Technology and Management[4]

Abstract: Sign language is one of the oldest and most natural form of language for communication, hence we have come up with a real time method using neural networks for finger spelling based American sign language. Automatic human gesture recognition from camera images is an interesting topic for developing vision. We propose a convolution neural network (CNN) method to recognize hand gestures of human actions from an image captured by camera. The purpose is to recognize hand gestures of human task activities from a camera image. The position of hand and orientation are applied to obtain the training and testing data for the CNN. The hand is first passed through a filter and after the filter is applied where the hand is passed through a classifier which predicts the class of the hand gestures. Then the calibrated images are used to train CNN.

Keywords: Convolution Neural Network (CNN), American Sign Language (ASL), gesture recognition and deep learning technologies.

I.INTRODUCTION

American sign language is a predominant sign language Since the only disability D&M people have been communication related and they cannot use spoken languages hence the only way for them to communicate is through sign language. Communication is the process of exchange of thoughts and messages in various ways such as speech, signals, behavior and visuals. Deaf and dumb (D&M) people make use of their hands to express different gestures to express their ideas with other people. Gestures are the nonverbally exchanged messages and these gestures are understood with vision. This nonverbal communication of deaf and dumb people is called sign language.

Communication itself is a fundamental human process through which individuals exchange information. It can occur in various forms, including speech, behavior, facial expressions, gestures, and other nonverbal cues. Hand gestures, movements, and signs are carefully structured in a visual manner, with each sign representing specific words or ideas. These gestures, understood through vision, form the core of sign language, enabling individuals to engage in rich, meaningful conversations with others who understand the language.

However, despite the significance of sign language for D&M people, a major challenge lies in bridging communication gaps with those who do not know ASL. The language barrier between the Deaf and hearing communities can often lead to isolation, misunderstandings, and difficulties in accessing services such as healthcare, education, and employment. This makes the development of real-time sign language translation technologies even more important.

As technology continues to advance, significant efforts are being made to bridge these communication gaps. In recent years, automatic systems for sign language recognition have gained attention. These systems, which use computer vision and machine learning techniques, have the potential to recognize and translate hand gestures into text or speech in real-time. Gesture recognition technology, particularly through the use of Convolutional Neural Networks (CNNs), is emerging as a powerful tool to automate the translation of ASL. CNNs have proven effective in learning complex visual patterns and recognizing hand shapes, orientations, and movements with high accuracy.

The integration of such technologies into everyday applications can make communication with D&M individuals much easier, reducing misunderstandings and promoting inclusivity. Real-time translation systems can enable Deaf individuals to interact seamlessly with the hearing population, whether in social settings, healthcare environments, or educational institutions. The ultimate goal is to create a system that provides not just translation, but also a more personalized, adaptive approach to sign language recognition that can accommodate various signing styles, regional variations, and contextual differences in ASL.

DOI: 10.17148/IARJSET.2025.12301

In conclusion, gesture-based sign language translation has the potential to break down communication barriers, making it a crucial area of research for creating a more inclusive society. The continuous development of advanced recognition systems holds the promise of improving the quality of life for D&M individuals.

II.PREVIOUS WORK

Over the years, a considerable amount of research has focused on developing systems for real-time gesture recognition, particularly for translating American Sign Language (ASL). Early approaches were largely based on image processing and computer vision techniques, where methods such as hand tracking, segmentation, and template matching were used to recognize static signs from camera images. However, these systems faced challenges with accuracy, especially when accounting for dynamic gestures, varying hand shapes, and inconsistent lighting conditions.

With the rise of deep learning, Convolutional Neural Networks (CNNs) have become a dominant tool for gesture recognition in ASL translation. A pivotal study by Khan and Khan (2019) proposed a CNN-based model for recognizing ASL signs in real-time, achieving significant improvements in both speed and accuracy over traditional image processing methods. Their approach utilized large datasets of hand gestures to train the network, enabling it to classify and translate signs into text or speech with remarkable precision.

Further advancements were made with hybrid systems, combining vision-based techniques with sensor-based methods. For example, Chen et al. (2016) developed a system that incorporated both camera inputs and wearable sensors. This hybrid approach helped overcome issues such as occlusion, background noise, and variable lighting, allowing for more accurate gesture recognition in real-world scenarios. Their work demonstrated that integrating multiple data sources could enhance the reliability of real-time ASL translation systems.

A comprehensive survey by Amin and Muhammad (2020) explored the state-of-the-art techniques in gesture recognition for sign language, focusing on the application of deep learning and CNNs. Their work highlighted the potential of deep neural networks to improve system robustness and accuracy, particularly when handling diverse signing styles and environmental variables.

Despite these advancements, challenges persist, particularly in handling the nuances of ASL, such as facial expressions, hand orientations, and contextual meanings. Ongoing research is focused on addressing these limitations, enhancing model generalization, and ensuring efficient real-time performance across diverse environments and user demographics.

III.METHODOLOGY

The methodology for real-time American Sign Language (ASL) translation leverages Convolutional Neural Networks (CNNs) for gesture recognition from camera images. The process involves several stages, each aimed at enhancing the accuracy, speed, and efficiency of recognizing and translating ASL gestures into text or speech in real-time.

III.1 Data Collection and Preprocessing

Data collection and preprocessing are crucial steps in building a real-time American Sign Language (ASL) translation system. The first step in data collection involves gathering a comprehensive dataset of ASL gestures, which includes a variety of hand signs representing words, phrases, or sentences in ASL. The dataset should cover multiple variations of each gesture, considering factors like hand shapes, angles, signer diversity (age, gender, ethnicity), and lighting conditions to ensure robustness. This diversity is important as ASL gestures can differ from person to person, and the model must be able to recognize these differences. To further enhance the dataset, data augmentation techniques such as rotation, scaling, flipping, and noise addition are applied. These techniques simulate different environmental conditions and help the model generalize better to real-world scenarios, improving its performance when faced with new or unseen gestures.

Preprocessing the collected data is a vital step before feeding it into the machine learning model. Images are first resized to a consistent dimension, ensuring that the input to the model is uniform and manageable. Additionally, converting the images to grayscale helps reduce computational complexity by focusing on essential features like shape and texture,

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 12, Issue 3, March 2025

DOI: 10.17148/IARJSET.2025.12301

while eliminating unnecessary color information. Image normalization is also applied to standardize pixel values, bringing them within a certain range, typically between 0 and 1, which enhances the stability of the learning process. In some cases, filters or edge-detection techniques may be used to emphasize important visual features such as the contours of the hands, making it easier for the model to distinguish between different gestures. This preprocessing ensures that the data is ready for effective training, allowing the model to learn relevant patterns efficiently and accurately.

The paper [2]"Data Augmentation and Preprocessing for Hand Gesture Recognition Using Deep Learning" by M. A. Hossain and M. M. Islam investigates the importance of data augmentation and preprocessing techniques in enhancing the performance of hand gesture recognition systems that rely on deep learning. The authors focus on addressing challenges related to noisy, variable gesture data caused by factors like different scales, orientations, and lighting conditions. To improve the system's performance, they emphasize the use of data augmentation techniques such as rotation, flipping, scaling, and translation to artificially expand the dataset, which helps the model generalize better. The paper also highlights preprocessing steps like normalization, background subtraction, and hand segmentation to clean and isolate the relevant parts of the image, reducing noise and improving gesture recognition accuracy.

The authors conclude that a combination of data augmentation and preprocessing is essential for improving the accuracy and robustness of gesture recognition models, especially when working with limited or diverse datasets. These techniques allow the models to better adapt to real-world scenarios and perform more reliably in practical applications.

III.2 Gesture Detection and Classification

Gesture detection and classification are key components in the process of recognizing American Sign Language (ASL) gestures. The first task in gesture detection involves isolating the hand from the background of the image, a crucial step for accurate recognition. This is typically achieved through hand region detection techniques, such as skin color segmentation or background subtraction, which help to identify and isolate the region of interest (ROI)—the hand—while discarding irrelevant parts of the image. Once the hand is isolated, the system processes the hand gestures through a Convolutional Neural Network (CNN), which acts as the core model for classification.

The CNN is designed to automatically learn features from the hand gesture images through a series of convolutional layers. These layers extract hierarchical visual features such as edges, textures, and shapes, which are important for distinguishing between different hand signs. Deeper layers of the network focus on higher-level features, enabling the model to recognize complex patterns in hand shapes and movements. After the features are extracted, the CNN's final layers output a class prediction, representing the specific ASL gesture. To ensure accurate recognition, the output is passed through a softmax activation function, which assigns probabilities to each possible gesture class. The gesture with the highest probability is selected as the predicted sign.

In this process, the system must not only recognize static signs but also account for variations in the signing style, hand orientations, and complex movements. As a result, the model is trained using a diverse and extensive dataset that includes various hand gestures, angles, lighting conditions, and signer characteristics. This comprehensive approach ensures that the system can accurately detect and classify ASL gestures in real-world scenarios, contributing to effective real-time translation.

The paper "Real-time Hand Gesture Recognition for Human-Computer Interaction" by Dong and Xie, presented at the International Conference on Artificial Intelligence and Computer Vision (AICV), introduces a system designed to enable real-time hand gesture recognition for human-computer interaction. The system focuses on gesture detection and classification by employing feature extraction techniques such as color, shape, and movement to recognize hand gestures. It combines computer vision methods to detect the hand and utilizes machine learning algorithms for classifying the gestures. This approach facilitates intuitive and responsive interaction between humans and computers, allowing for efficient gesture-based control in real-time applications.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 3, March 2025
DOI: 10.17148/IARJSET.2025.12301

III.3 Real-Time Translation

Real-time translation is the final stage in an American Sign Language (ASL) recognition system, where the recognized gestures are converted into meaningful output, such as text or speech, to facilitate communication between Deaf and hearing individuals. Once a gesture is detected and classified by the system, it is mapped to the corresponding ASL sign, which is then translated into text that can be displayed on a screen or spoken aloud using a text-to-speech (TTS) engine. This process must occur quickly to ensure smooth and natural communication, as any significant delay can disrupt the flow of conversation. To achieve real-time performance, the system relies on highly optimized algorithms and efficient hardware resources, such as Graphics Processing Units (GPUs), to accelerate the gesture recognition process. These technologies allow the model to process each frame of video input rapidly, reducing latency and providing immediate feedback to the user.

Real-time translation also involves continuous adaptation to dynamic environmental conditions, such as changes in lighting, background, or signer variations. As part of the system's design, real-time performance is maintained by using lightweight, optimized models and techniques like parallel processing, which enable the system to handle multiple data inputs simultaneously. Moreover, real-time translation systems must be robust enough to recognize and adapt to individual signing styles, making the system more flexible and personalized. The ability to handle these variations and deliver quick, accurate translations is crucial for creating a seamless communication experience for both Deaf and hearing individuals, allowing for greater inclusivity and accessibility.

The paper [39] "Real-Time Sign Language Recognition and Translation System Using Convolutional Neural Networks" by Y. Zhang, Y. Zhu, et al., presented at the International Conference on Image Processing and Computer Vision (IPCV), focuses on developing a real-time sign language recognition system utilizing Convolutional Neural Networks (CNNs) for accurate gesture detection. The system is designed to translate sign language gestures into text or speech, bridging the communication gap between sign language users and those who are unfamiliar with it. By leveraging deep learning techniques, particularly CNNs, the paper highlights how these networks can be trained to detect and classify complex hand gestures with high precision. The system aims to provide a seamless and responsive means of real-time translation, making sign language more accessible to non-sign language speakers.

In conclusion, the authors emphasize that the use of CNNs in real-time sign language recognition provides an efficient and accurate method for translating gestures into text or speech. The system's reliance on deep learning frameworks allows for high levels of accuracy, even in real-time scenarios, making it a powerful tool for improving communication for the deaf and hard-of-hearing community. The work demonstrates the potential of CNNs to enhance real-time applications, offering a practical solution to overcoming language barriers between sign language users and others.

III.4 Post-Processing and Feedback

Post-processing and feedback are essential components for improving the accuracy and reliability of a real-time American Sign Language (ASL) translation system. After a gesture has been detected and translated, post-processing ensures that the output is presented in a coherent and meaningful way. This stage involves tasks like refining the translation, adjusting for any ambiguities or errors in gesture recognition, and ensuring that the final output—whether text or speech—accurately reflects the intended sign language. For instance, if a gesture is misclassified or ambiguous, post-processing can help by smoothing the translation output or correcting minor errors.

Feedback mechanisms play a crucial role in enhancing the system's performance over time. When the system produces an incorrect or suboptimal translation, the feedback loop allows for manual correction or real-time user input to adjust the translation. This feedback can be used to retrain and fine-tune the model, improving its recognition capabilities and reducing future errors. Over time, as the system receives more corrections and adapts to different signers, it becomes more accurate and personalized, especially in recognizing individual signing styles or regional variations. Furthermore, continuous learning enables the model to improve in diverse environments, making it more resilient to real-world challenges such as changes in lighting, background noise, or hand occlusion. Ultimately, post-processing and feedback work together to create a more adaptive, reliable, and user-centered ASL translation system, leading to better communication and enhanced inclusivity.

International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.066 ⁖ Peer-reviewed & Refereed journal ⁖ Vol. 12, Issue 3, March 2025
DOI: 10.17148/IARJSET.2025.12301

The paper [3] "Post-Processing Techniques for Gesture Recognition Systems" by H. Chen, Y. Zhang, et al., published in the IEEE Transactions on Cybernetics, explores several post-processing techniques designed to enhance the accuracy of gesture recognition systems. After gestures are initially recognized, the authors introduce methods such as noise reduction, gesture smoothing, and gesture refinement to improve the quality of the recognition output. These techniques help eliminate errors, reduce noise, and ensure that the system can handle variations in gestures more effectively. The paper also emphasizes the importance of real-time feedback, which allows users to understand the success or failure of their gestures and adjust them accordingly. This feedback helps improve the interaction and performance of the system, ensuring that the recognition process becomes more reliable and accurate over time.

In conclusion, the paper highlights the significant role of post-processing in enhancing gesture recognition accuracy by applying techniques that refine and smooth gestures after initial detection. Moreover, the integration of real-time feedback provides a dynamic interaction with users, allowing them to fine-tune their gestures and achieve better recognition outcomes. This combination of post-processing techniques and immediate feedback represents a crucial advancement in improving the effectiveness and usability of gesture-based systems.

IV.CONCLUSION

In conclusion, the development of real-time American Sign Language (ASL) translation systems using gesture recognition and deep learning technologies, particularly Convolutional Neural Networks (CNNs), has immense potential to bridge communication gaps between the Deaf and hearing communities. ASL, being a rich and dynamic language, requires advanced recognition systems capable of accurately interpreting hand gestures, facial expressions, and body movements. While significant progress has been made, challenges such as variations in signing styles, noise, and occlusion continue to hinder perfect accuracy. Nevertheless, the integration of CNNs and other machine learning techniques holds promise for enhancing real-time translation systems, making communication more seamless and inclusive for Deaf and Dumb individuals. By continuing to improve the speed, accuracy, and adaptability of these systems, future developments will create more accessible platforms for communication, allowing for better integration of Deaf individuals into social, educational, and professional environments, ultimately fostering greater societal inclusivity.

REFERENCES

[1] T. Kim, J. Keane, W. Wang, H. Tang, J. Riggle, G. Shakhnarovich, D. Brentari, and K. Livescu, “Lexicon-free fingerspelling recognition from video: Data, models, and signer adaptation,” Comput. Speech Lang.,vol. 46, pp. 209–232, Nov. 2017.
[2] M. A. Hossain, B. B. Zaidan, A. A. Zaidan, M. M. islam, and M. M. B. Lakulu, “A review on systems-based sensory gloves for sign language recognition state of the art between 2007 and 2017,” Sensors,vol. 18, no. 7, p. 2208, 2018.
[3] R. Cui, H. Liu, and C. Zhang, “A deep neural framework for continuous sign language recognition by iterative training,” IEEE Trans. Multimedia, vol. 21, no. 7, pp. 1880–1891, Jul. 2019.
[4] P. S. Santhalingam, P. Pathak, J. Košecká, and H. Rangwala, “Sign language recognition analysis using multimodal data,” in Proc. IEEE Int. Conf. Data Sci. Adv. Anal. (DSAA), Oct. 2019, pp. 203–210.
[5] A. C. Duarte, “Cross-modal neural sign language translation,” in Proc.27th ACM Int. Conf. Multimedia, Oct. 2019, pp. 1650–1654.
[6] M. J. Cheok, Z. Omar, and M. H. Jaward, “A review of hand gestureand sign language recognition techniques,” Int. J. Mach. Learn. Cybern.,vol. 10, no. 1, pp. 131–153, Jan. 2019.
[7] Q. Xiao, Y. Zhao, and W. Huan, “Multi-sensor data fusion for sign language recognition based on dynamic Bayesian network and convolutional neural network,” Multimedia Tools Appl., vol. 78, no. 11,
pp. 15335–15352, Jun. 2019.
[8] E. K. Kumar, P. V. V. Kishore, M. T. K. Kumar, and D. A. Kumar, “3D sign language recognition with joint distance and angular coded color topographical descriptor on a 2–stream CNN,” Neurocomputing,vol. 372, pp. 40–54, Jan. 2020.
[9] J. Wu and R. Jafari, “Wearable computers for sign language recognition,”in Handbook of Large-Scale Distributed Computing in Smart Healthcare.Cham, Switzerland: Springer, 2017, pp. 379–401.
[10] J. Shang and J. Wu, “A robust sign language recognition system with multiple wi-fi devices,” in Proc. Workshop Mobility Evolving Internet Archit., New York, NY, USA, 2017, pp. 19–24, doi: 10.1145/3097620.3097624.

DOI: 10.17148/IARJSET.2025.12301

[11] J. Pu, W. Zhou, and H. Li, "Iterative alignment network for continuous sign language recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2019, pp. 4165–4174.
[12] Q. Xiao, M. Qin, and Y. Yin, "Skeleton-based Chinese sign language recognition and generation for bidirectional communication between deaf and hearing people," Neural Netw., vol. 125, pp. 41–55, May 2020.
[13] W. Aly, S. Aly, and S. Almotairi, "User-independent American sign language alphabet recognition based on depth image and PCANet features," IEEE Access, vol. 7, pp. 123138–123150, 2019.
[14] Z. Zafrulla, H. Brashear, T. Starner, H. Hamilton, and P. Presti, "American sign language recognition with the kinect," in Proc. 13th Int. Conf. Multimodal Interfaces (ICMI), 2011, pp. 279–286.
[15] S. Wei, X. Chen, X. Yang, S. Cao, and X. Zhang, "A component-based vocabulary-extensible sign language gesture recognition framework," Sensors, vol. 16, no. 4, p. 556, Apr. 2016.
[16] N. B. Ibrahim, H. H. Zayed, and M. M. Selim, "Advances, challenges and opportunities in continuous sign language recognition," J. Eng. Appl. Sci., vol. 15, no. 5, pp. 1205–1227, Dec. 2019.
[17] L. Zheng, B. Liang, and A. Jiang, "Recent advances of deep learning for sign language recognition," in Proc. Int. Conf. Digit. Image Comput., Techn. Appl. (DICTA), Nov. 2017, pp. 1–7.
[18] T. Starner, J. Weaver, and A. Pentland, "Real-time American sign language recognition using desk and wearable computer based video," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 12, pp. 1371–1375,Dec. 1998.
[19] F.-S. Chen, C.-M. Fu, and C.-L. Huang, "Hand gesture recognition using a real-time tracking method and hidden Markov models," Image Vis. Comput., vol. 21, no. 8, pp. 745–758, Aug. 2003.
[20] X. Wang, M. Xia, H. Cai, Y. Gao, and C. Cattani, "Hidden-Markovmodels-based dynamic hand gesture recognition," Math. Problems Eng., vol. 2012, pp. 1–11, Feb. 2012.
[21] G. Marin, F. Dominio, and P. Zanuttigh, "Hand gesture recognition with leap motion and kinect devices," in Proc. IEEE Int. Conf. Image Process. (ICIP), Oct. 2014, pp. 1565–1569.
[22] X. Lu, B. Qi, H. Qian, Y. Gao, J. Sun, and J. Liu, "Kinect-based human finger tracking method for natural haptic rendering," Entertainment Comput., vol. 33, Mar. 2020, Art. no. 100335.
[23] S. Stoll, N. C. Camgoz, S. Hadfield, and R. Bowden, "Text2sign: Towards sign language production using neural machine translation and generative adversarial networks," Int. J. Comput. Vis., vol. 128, no. 4, pp. 1–18,2020.
[24] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proc. IEEE, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
[25] E. Escobedo, L. Ramirez, and G. Camara, "Dynamic sign language recognition based on convolutional neural networks and texture maps," in Proc. 32nd SIBGRAPI Conf. Graph., Patterns Images (SIBGRAPI), Oct. 2019, pp. 265–272.
[26] S. Hayani, M. Benaddy, O. El Meslouhi, and M. Kardouchi, "Arab sign language recognition with convolutional neural networks," in Proc. Int. Conf. Comput. Sci. Renew. Energies (ICCSRE), Jul. 2019, pp. 1–4.
[27] Y. Liao, P. Xiong, W. Min, W. Min, and J. Lu, "Dynamic sign language recognition based on video sequence with BLSTM-3D residual networks," IEEE Access, vol. 7, pp. 38044–38054, 2019.
[28] P. Witoonchart and P. Chongstitvatana, "Application of structured support vector machine backpropagation to a convolutional neural network for human pose estimation," Neural Netw., vol. 92, pp. 39–46, Aug. 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0893608017300321
[29] S. He, "Research of a sign language translation system based on deep learning," in Proc. Int. Conf. Artif. Intell. Adv. Manuf. (AIAM), Oct. 2019,pp. 392–396.
[30] Y. Yang and D. Ramanan, "Articulated pose estimation with flexible mixtures-of-parts," in Proc. CVPR, Jun. 2011, pp. 1385–1392.
[31] P. Q. Thang, N. T. Thuy, and H. T. Lam, "The SVM, SimpSVM and RVM on sign language recognition problem," in Proc. 7th Int. Conf. Inf. Sci.Technol. (ICIST), Apr. 2017, pp. 398–403.
[32] R. A. A. R. Agha, M. N. Sefer, and P. Fattah, "A comprehensive study on sign languages recognition systems using (SVM, KNN, CNN and ANN)," in Proc. 1st Int. Conf. Data Sci., E-Learn. Inf. Syst., Oct. 2018,pp. 1–6.
[33] E. Alpaydin, Introduction to Machine Learning. Cambridge, MA, USA:MIT Press, 2020.
[34] A. Wadhawan and P. Kumar, "Sign language recognition systems:A decade systematic literature review," Arch. Comput. Methods Eng., vol. 28, no. 3, pp. 1–29, 2019.
[35] Y. S. Abu-Mostafa, M. Magdon-Ismail, and H.-T. Lin, Learning From Data, vol. 4. New York, NY, USA: AMLBook, 2012.
[36] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521,pp. 436–444, May 2015.
[37] D. Konstantinidis, K. Dimitropoulos, and P. Daras, "A deep learning approach for analyzing video and skeletal features in sign language recognition," in Proc. IEEE Int. Conf. Imag. Syst. Techn. (IST), Oct. 2018,pp. 1–6.
[38] D. Wu and L. Shao, "Multimodal dynamic networks for gesture recognition," in Proc. 22nd ACM Int. Conf. Multimedia, Nov. 2014, pp. 945–948.
[39] A. Tang, K. Lu, Y. Wang, J. Huang, and H. Li, "A real-time hand posture recognition system using deep neural networks," ACM Trans. Intell. Syst.Technol., vol. 6, no. 2, pp. 1–23, May 2015.

DOI: 10.17148/IARJSET.2025.12301

[40] S. Aly, B. Osman, W. Aly, and M. Saber, "Arabic sign language fingerspelling recognition from depth and intensity images," in Proc. 12th Int.Comput. Eng. Conf. (ICENCO), Dec. 2016, pp. 99–104.

[41] J. Huang, W. Zhou, H. Li, and W. Li, "Sign language recognition using 3D convolutional neural networks," in Proc. IEEE Int. Conf. MultimediaExpo (ICME), Jun. 2015, pp. 1–6.

[42] K. Li, Z. Zhou, and C.-H. Lee, "Sign transition modeling and a scalablesolution to continuous sign language recognition for real-world applications," ACM Trans. Accessible Comput., vol. 8, no. 2, pp. 1–23, Jan. 2016.

[43] J. Huang, W. Zhou, H. Li, and W. Li, "Sign language recognition using real-sense," in Proc. IEEE China Summit Int. Conf. Signal Inf. Process.(ChinaSIP), Jul. 2015, pp. 166–170.

[44] L. Rioux-Maldague and P. Giguere, "Sign language fingerspelling classification from depth and color images using a deep belief network," in Proc. Can. Conf. Comput. Robot Vis., May 2014, pp. 92–97.

[45] M. A. Hossen, A. Govindaiah, S. Sultana, and A. Bhuiyan, "Bengali sign language recognition using deep convolutional neural network," in Proc.Joint 7th Int. Conf. Informat., Electron. Vis. (ICIEV) 2nd Int. Conf. Imag.,Vis. Pattern Recognit. (icIVPR), Jun. 2018, pp. 369–373.

[46] P. Escudeiro, N. Escudeiro, R. Reis, J. Lopes, M. Norberto, A. B. Baltasar, M. Barbosa, and J. Bidarra, "Virtual sign—A real time bidirectional translator of Portuguese sign language," Proc. Comput. Sci., vol. 67,pp. 252–262, Jan. 2015.

[47] Suharjito, H. Gunawan, N. Thiracitta, and A. Nugroho, "Sign language recognition using modified convolutional neural network model," in Proc. Indonesian Assoc. Pattern Recognit. Int. Conf. (INAPR), Sep. 2018,pp. 1–5.

[48] N. Soodtoetong and E. Gedkhaw, "The efficiency of sign language recognition using 3D convolutional neural networks," in Proc. 15th Int. Conf. Electr. Eng./Electron., Comput., Telecommun. Inf. Technol. (ECTICON), Jul. 2018, pp. 70–73.

[49] J. Huang, W. Zhou, H. Li, and W. Li, "Attention-based 3D-CNNs for large-vocabulary sign language recognition," IEEE Trans. Circuits Syst.Video Technol., vol. 29, no. 9, pp. 2822–2832, Sep. 2019.

[50] D. Guo, W. Zhou, A. Li, H. Li, and M. Wang, "Hierarchical recurrent deep fusion using adaptive clip summarization for sign language translation," IEEE Trans. Image Process., vol. 29, pp. 1575–1590, 2019.

