



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

A futuristic digital interface with a glowing Earth globe in the center. The interface is overlaid on a background of a person's hands reaching out towards the globe. The interface includes various data visualizations such as bar charts, line graphs, and pie charts. Text elements include "BUSINESS", "NETWORK SEARCH", "LORDING 100%", "MEDIA", and "WORLD". The overall color scheme is blue and purple with a glowing effect.

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 6, June 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

 9940 572 462

 6381 907 438

 ijirset@gmail.com

 www.ijirset.com

Hand Gesture Identification for Artificial Intelligence through Sign Translation

P. Sai Krishna¹, P. Kanishka², S.V. Harini³, K. Kasthuri⁴, S. Poovarasi⁵

Assistant professor, Department of ECE, Mahendra Institute of Technology, Tamil Nadu, India¹

Final Year Students, Department of ECE, Mahendra Institute of Technology, Namakkal, Tamil Nadu, India^{2,3,4,5}

ABSTRACT: The translation of hand signals used in sign language into spoken or written language is known as sign language recognition. It involves analysing and deciphering hand and finger gestures using computer vision and machine learning algorithms, then mapping those actions to corresponding words or phrases. There are various uses for sign language recognition, including facilitating communication between hearing and deaf persons, increasing accessibility for those with hearing loss, and improving sign language interpreting education and training programmes. The way we engage with people who primarily communicate through sign language may change significantly if this technology is used. Sign language can be recognized using a variety of techniques, such as sensor-based techniques, computer vision techniques, and hybrid techniques that combine both. Computer vision algorithms use visual data from cameras to identify hand movements and gestures. Sensor-based methods capture motion and record data by attaching sensors to the hands or fingers. Consequently, it may be possible to investigate various approaches to sign language recognition and predict a better method with a greater accuracy rate.

KEYWORDS: Artificial intelligence, Hand gestures, Sign language, Deep learning, Neural networks

I. INTRODUCTION

Sign language recognition is a technological game-changer that facilitates communication between the Deaf and hard of hearing populations and the hearing world. For many people, sign language is their preferred method of communication since it is a complex and expressive form of expression that uses hand gestures, facial emotions, and body movements [11][12]. Nonetheless, the multiplicity of languages poses distinct difficulties for automatic recognition of sign language. The development of sign language recognition systems has been made possible in recent years by developments in computer vision and machine learning [13][14]. These systems have great potential for helping the Deaf community use technology, access learning materials, and engage more fully in a world where spoken language is the primary form of communication. With the potential to promote equality of opportunity, improve inclusivity, and transform communication, this technology represents an exciting and crucial area for research and development. Computer vision and machine learning techniques are important foundations for the creation of systems for recognising sign language. [16, 15]. These systems use cameras to record and analyse a signer's hand, facial, and body actions. The cameras then translate the gestures into written or audible output. [18][17] Even though technology is still in its infancy, it has come a long way and is used in many different sectors, such as:

Education: Recognition of sign language can be integrated into educational tools and platforms to create a more inclusive learning environment [19][20]. It enables Deaf students to access content and communicate with teachers and peers effectively.

Accessibility: Sign language recognition software can help people communicate in public areas like transit stations and government buildings, ensuring that Deaf individuals can access information and services independently [21][22].

Healthcare: In healthcare settings, the technology can be used to bridge the communication gap between Deaf patients and medical professionals, allowing for more accurate and empathetic care [23][24].

Telecommunication: Sign language recognition can be applied to video conferencing and real-time communication platforms, enabling Deaf individuals to participate in virtual meetings and conversations [25][26].

Assistive Devices: Wearable devices with sign language recognition capabilities can serve as personal interpreters, improving everyday communication for Deaf individuals [27].

As research and development in sign language recognition continue to advance, there is an increasing focus on enhancing the accuracy and robustness of these systems, supporting various sign languages and regional dialects [28]. Furthermore, creating affordable and user-friendly applications that integrate seamlessly into daily life remains a critical goal [29]. Fig 1 [Sign Language MNIST on Kaggle] shows the gestures for alphabets

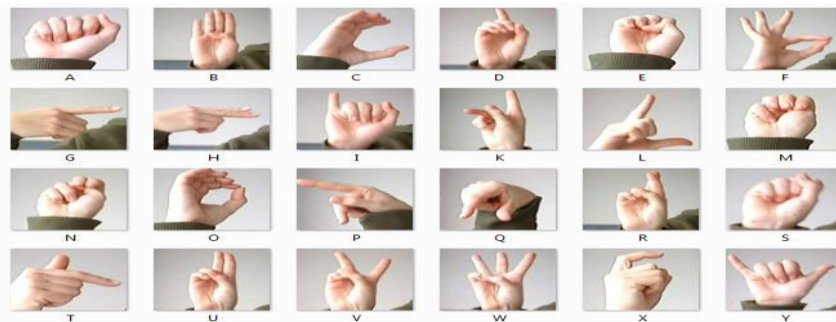


Figure 1: Sign language for alphabets.

II. RELATED WORK

Farman shah, et.al,...[1] suggested for the bare-handed recognition of 36 static alphabets. The videos on sign language are the source of the dataset. Four vision-based characteristics are then extracted: edge-oriented histogram, robust features, a histogram of directed gradients, and local binary patterns with accelerated processing. Support vector machines (SVMs) with multiple kernels learning (MKL) are used to classify each of the extracted characteristics separately then implemented the basic the multi-class SVM using the binary SVM using a one-to-all strategy. To determine the ultimate recognition, a vote method is implemented, and the accuracy, precision, recall, and F-score of the suggested technique are used as performance metrics. The simulation results show promise when compared to current methods. People who are deaf all around the world engage and communicate through sign language. This type of communication is carried out using various arm/body movements, face expressions, and hand gestures. The sign language recognition system aims to enhance the deaf community's capacity for public interaction. The human-computer interface (HCI) is provided by a highly structured symbolic set. Sign language is an extremely useful tool for communication, and millions of deaf people use it every day to express themselves and communicate with one another throughout the world. Through this facilitation and assistance, deaf individuals are encouraged and empowered to integrate into society and be a healthy part of it.

Shikharsharma, et.al,...[2] carried out Sign language is an organized hand gesture system that integrates messages and visual inputs to facilitate communication. Sign language provides the deaf and speech-impaired community with practical skills for everyday engagement. Several bodily parts, In sign language, the fingers, hand, arm, head, body, and facial expression are all employed. to convey information. Nonetheless, the hearing community does not use sign language much, and even fewer can comprehend it. This presents a real communication hurdle that hasn't been completely resolved to this day between the deaf population and the general public. More and more cutting-edge technologies are becoming available, like Kinect, LMC, and EMG, which can more easily record gesture data. Pre-processing techniques that are frequently employed include shrinking photos before moving on to further stages and using the median and Gaussian filters. One of the most used segmentation techniques is skin color segmentation. The color spaces CIE Lab, YCbCr, and HSV are typically more resilient to lighting conditions. More recent studies increase segmentation performance by combining a number of additional spatial features and modeling techniques.

Nikolas adaloglou, et.al,...[3] outlined a study that aimed to shed light on the detection of sign language by assigning glosses to non-segmented video streams. For this task, two new sequence training criteria are provided; these are well-known from the scene text recognition domain and voice. Additionally, a wide range of pre-training programs are covered in detail. At last, a fresh A dataset for Greek sign language is created using RGB+D. The duty of interpreting a signer's glosses from video footage is known as sign language recognition, or SLR. Despite a substantial quantity of

work in the topic of SLR, comprehensive experimental research is severely lacking. Furthermore, the majority of articles neither publish their code nor report results across all datasets. As a result, experimental findings in the field of SL are frequently uninterpretable and replicable. Since spoken languages are heard by the ear and articulated with the mouth, they operate through the "vocal - auditory" channel. Said languages are the source of all writing systems, or they serve as images of them. Sign languages (SLs) are distinct from other languages because they operate on the "corporal - visual" channel, which is generated with the body and seen with the eyes. SLs are frequently used by the Deaf populations and are not international. Since they emerge naturally whenever Deaf people have the chance to get together and interact, these languages are considered natural languages. This research presents a comparative experimental evaluation of approaches for computer vision-based sign language recognition. A comprehensive evaluation on several publicly available datasets is carried out by utilizing the state-of-the-art deep neural network techniques in this domain.

Ageloskratimenos, et.al,...[4] examined the procedure for removing hand, face, and three-dimensional body posture components in order to recognise sign language. These attributes were compared to open posture primary concepts, the most popular method for extracting 2D skeleton properties and characteristics from raw RGB frames and their optical flow. A state-of-the-art Deep Learning architecture that is used for action and sign language recognition is equipped with these capabilities. The trials showed the advantage of SMPL-X features due to the thorough and qualitative feature extraction in the three previously designated zones of interest. Furthermore, use SMPL-X to emphasise how crucial it is to combine all three of these sites in order to achieve the greatest results for SLR. Further experiments in other independent datasets with a larger signer count and a more diverse environment will be conducted in the future to further investigate 3D body, face, and hand extraction for SLR.

Ilhamelouariachi, et.al,...[5] three elements were used to create a sign gesture: non-manual elements like body position, facial expressions, and finger spelling when words are spelled out in alphabetical order, and manual characteristics, which are hand motions. When an idea, like names, technical terminology, or foreign words, lacks a distinct sign, it becomes important to fingerspell. The intricacy of sign language, which conveys meaning through a variety of means including words, body posture, facial expressions, and finger spellings, makes it challenging to recognize signs. The ability to collect RGB and Depth (RGB-D) data has allowed recent technology breakthroughs, such as the Kinect sensor, to open up new possibilities for human-computer interaction and sign language. For feature extraction, the RGB and Depth pictures are handled separately in the conventional methods.

Dongxu li, et.al[6] offer a novel method to improve the performance of sign language recognition models by utilising cross-domain information from the subtitled sign news videos. By coarsely aligning isolated signs and news signs via combined training, it is suggested that class centroids be stored in a typical memory for use in both offline and online inference. Next, for every isolated sign, our model learns a domain-invariant description. Use the temporal attention method to highlight features unique to a class and conceal features shared by other classes according to the description for domain invariance.

Ozgemercanoglusincan, et.al[7]summed up the ChaLearn LAP Isolated Signer Independent Large Scale Signer Challenge. In the two computational tracks (RGB and RGB+D), more than 190 participants took up the challenge; in all, more than 1.5K contributions were made. Reviewed and discussed the challenge concept, top-winning solutions, and outcomes. Interestingly, the challenge activity and results across the several tracks showed that the RGB data is now getting more attention from the academic community in this field than the RGB+Depth data. Pose estimation, hand/face identification, transfer learning, external data, fusion/ensemble of modalities, and many methods to represent spatiotemporal information were all combined into the most effective strategies. Certain signs, such as government and Ataturk, just have a small change in the hand's orientation. In the Ataturk sign, the index finger touches the face; in the government sign, it touches the under eye. The hands-on certain signs, such "school" and "soup," only move slightly differently.

Boháček, Matyáš, and Marek Hruz, et.al [8] investigated the use of Transformer for the isolated SLR challenge. In order to get reasonable results, previous research addressing this topic have either depended on pre-trained backbones or used a computationally intensive method, which is hindered by the usage of manually created pose feature representations, which lowers the dimensionality. Moreover, prior systems were unable to properly take use of any augmentations or normalizations given to visual data beyond the standard ones. Provide a creative way to use Transformer for this assignment. Apply SL linguistics knowledge to develop novel data augmentation strategies

tailored to the SL, as well as a robust normalization technique, as our model relies on body pose sequence representations. And used two datasets for isolated SLR to validate the methodology achieved overall state of the art results for the LSA64 and set state of the art results in the pose-based model category for the WLASL. Additionally carried out a performance analysis contrasting our model with the I3D baseline, demonstrating how much less demanding the recently suggested architecture is and how effectively it generalizes, even on extremely tiny training sets.

Katoch, Shagun, et.al [9] provided a framework for creating a large, diversified, and dependable real-time alphabet (A-Z) and digit (0-9) recognition system for Indian Sign Language. The authors have used camera photos to identify signs in these images rather than depending on pricey tools like gloves or the Kinect. The correctness of the result is also covered in this study. Real-time, accurate, and quick judgement on ISL sign identification is required to enable communication between those with hearing or speech impairments and the able-bodied. More signs in various languages spoken in other countries can be added to the collection in the future, resulting in a more useful framework for real-time applications. Simple phrases and expressions may be generated using the approach for both continuous and isolated identification tasks. For real-time applications, reaction time acceleration is essential.

Kothadiya, Deep, et al [10] The study that was presented aimed to develop an offline system for the recognition of sign language. The signer's data was gathered using a vision-based approach. One of the main features of the study is the system's ability to detect and recognise the words in IISL2020 (our own dataset). For every one of the eleven terms in the IISL2020 collection, there are more than 1100 video examples including 16 study subjects, both male and female. Natural lighting, orientation, background correction, gloves, and other accessories were taken into consideration when creating the IISL2020 dataset. Unlike other sign language datasets, the scientists did not use any external aids, such sensors or smart gloves, to identify hand motions. The accuracy of the model could be increased in subsequent studies by creating new datasets under optimal circumstances, shifting the camera's angle, and even utilizing wearable technology.

Table 1 provides the overall details about literature survey papers.

S.NO	TITLE	TECHNIQUES	MERITS	DEMERITS
1	Sign Language Recognition Using Multiple Kernel Learning: A Case Study of Pakistan Sign Language	Multiple kernel learning (MKL) in support vector machine (SVM)	Provide the results for features sets	Outliers can be occurred
2	ASL-3DCNN: American sign language recognition technique using 3-D convolutional neural networks	Convolutional Neural Networks (CNNs)	Dynamic gesture recognition system	Does not support real time applications
3	A Comprehensive Study on Deep Learning-based Methods for Sign Language Recognition	DNN based approach	Image based sign recognition	Time complexity is high
4	Independent sign language recognition with 3d body, hands, and face reconstruction	Deep Learning architecture	Skeleton based approach	Accuracy is less
5	Explicit quaternion krawtchouk moment invariants for finger-spelling sign language recognition	Quaternion Krawtchoukmoments (QKMs)	Video based finger spelling data analysis	irrelevant outcomes in the presence of background complexity and noisy situations.
6	Transferring cross-domain knowledge for video sign language recognition	Transfer learning techniques	Video sign recognition	Computational complexity is high
7	Chalearn lap large scale signer independent isolated sign language recognition	RGB to depth analysis	Image based analysis system	Training datasets are high

	challenge: design, results and future research			
8	Sign pose-based transformer for word-level sign language recognition	Pose-based Transformer	Reduce the dimensionality	Only support image datasets
9	Indian Sign Language recognition system using SURF with SVM and CNN	SURF (Speeded Up Robust Features)	Custom dataset-based analysis	Misclassification can be occurred
10	Deepsign: Sign Language Detection and Recognition Using Deep Learning	LSTM model	Achieve higher accuracy in recognition	Time complexity can be high

III. EXISTING METHODOLOGIES

Sign language is a common communication tool used by deaf-dumb people [30]. Diverse hand forms, motions, orientations, and facial expressions combine to create the multitude of hand gestures that make up sign language [31]. Of the 466 million people with hearing loss worldwide, 34 million are youngsters’ people who are labelled as "deaf" have minimal or no sense of hearing. They converse with each other via sign language. Different sign languages are used by people in different places of the world [32]. They are extremely few in number when compared to spoken languages. Finger gesture detection efforts in the current system have been limited due to a lack of datasets and variations in sign language with locale [33][34]. The initial efforts towards bridging the communication gap between normal people and deaf and dumb people are being taken by an ongoing initiative that employs Indian sign language [50]. A useful expansion of current research into words and everyday expressions could help the deaf and hard of hearing connect with the outside world more quickly and easily, as well as further the development of autonomous systems that can comprehend and assist them [35][36] Research in Indian Sign Language lags behind that of its American equivalent due to a lack of standard databases. [37].

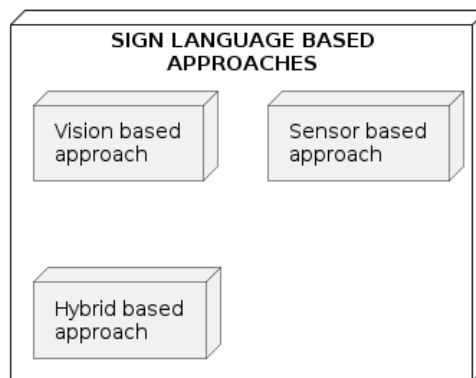


Fig 2. Sign language-based approaches

The primary method by which vision-based systems obtain the necessary input data is through cameras. The main advantage of using a camera is that it reduces the building expenses of the system and does away with the need for sensors in sensory gloves [38][39]. Because web cameras can produce blur, most laptops employ high-spec cameras, which are reasonably inexpensive [40] [41].

Using a specific type of instrumented glove that has many sensors installed, such as flexion (or bend) sensors, accelerometers (ACCs), proximity sensors, and abduction sensors, is an alternative way to collect gesture-related data [42][43]. These sensors measure the wrist's orientation (roll, pitch, and yaw), the abduction between fingers, and the finger bend angles. Different degrees of freedom (DoF) may be obtained with such gloves, depending on how many sensors are included into the glove. [44][45].The third technique for gathering raw gesture data uses a hybrid strategy that blends camera- and glove-based technology. To improve overall accuracy and precision, this method employs mutual error elimination [46][47]. Unfortunately, because of the setup costs and computational overheads, not much

research has been done in this area. However, when combined with hybrid tracking methodology, augmented reality systems yield encouraging results [48][49].

IV. PROPOSED METHODOLOGIES

For the Deaf and Hard of Hearing communities, sign language is an essential form of communication. Building a reliable CNN-based sign language recognition system can improve accessibility and communication. In order to recognize and understand sign language motions for real-time translation and communication, this suggested effort would make use of deep learning techniques. Sign language is a vital means of communication for this group. The primary objective of the study is to develop an accurate and efficient Convolutional Neural Network (CNN) model that can quickly recognise and comprehend gestures in sign language. The task will entail developing an intuitive application that enables quick identification and interpretation of sign language gestures. To guarantee that the model can accommodate a wider range of signers and sign languages, the research will also investigate how well the model adapts to these variations. To enhance the model's performance, techniques including data augmentation and fine-tuning will be applied. Examining ways to implement the system in assistive technology and educational settings is part of the long-term objective, which will eventually help the Deaf and Hard of Hearing community communicate and be more accessible.

4.1 CNN ALGORITHM STEPS

Data Collection:

- Collect a dataset of sign language gestures. This dataset should consist of images or video frames of individuals making various sign language signs. Ensure that the dataset is diverse and well-labelled.

Data Preprocessing:

- Preprocess the dataset by resizing images to a consistent size, normalizing pixel values, and potentially applying data augmentation techniques like rotation, scaling, and brightness adjustment to increase dataset diversity.
- Data Splitting:
- Separate the dataset into sets for testing, validation, and training. The model is trained on the training set, hyperparameter tweaking is done on the validation set, and the model's ultimate performance is assessed on the testing set.

Model Architecture:

- Create the architecture for CNN. Convolutional layers are usually used for feature extraction, and fully linked layers are used for classification. Enough depth in the design is required to capture intricate spatial patterns.
- Convolutional Layers:
- Implement convolutional layers with appropriate filter sizes and strides. These layers learn to detect features and patterns within the sign language gestures.
- Activation Functions:
- Apply activation functions, such as ReLU (Rectified Linear Unit), after each convolutional layer to introduce non-linearity into the model.

Pooling Layers:

- Insert pooling layers, like max-pooling, to minimise the feature maps' spatial size while maintaining essential information. Pooling helps make the model more robust to variations in hand positions.
- Flatten Layer:
- After convolutional and pooling layers, add a flatten layer to convert the 2D feature maps into a 1D vector for input to the completely linked layers.

Fully Connected Layers:

- Incorporate completely linked layers for classification. These layers should have neurons equal to the number of distinct signs want to recognize. Use an appropriate activation function for the output layer (e.g., softmax for multi-class classification).

Loss Function:

- Select an appropriate loss function, like categorical cross-entropy, to calculate the discrepancy between the training data's expected and actual signs.

Optimization Algorithm:

- Select an optimization algorithm (e.g., Adam, SGD) to minimize the loss function. Set hyperparameters like learning rate and momentum.

Training:

- Train the CNN model using the training dataset. This involves forward and backward passes to update the model's weights through multiple epochs.

Validation:

- Validate the model's training results on the validation set monitor metrics like accuracy and loss to tune hyperparameters and prevent overfitting.

Testing:

- After training, evaluate the model on the testing set to assess its generalization and accuracy.
- Fine-Tuning (Optional):
- Fine-tune hyperparameters or the model architecture to improve performance if necessary.

Deployment:

- Deploy the trained CNN for real-time sign language recognition, possibly integrating it into an application or device with a user-friendly interface.
- These steps provide a framework for building a CNN for sign language recognition, and they can be adjusted based on the specific requirements of sign language recognition

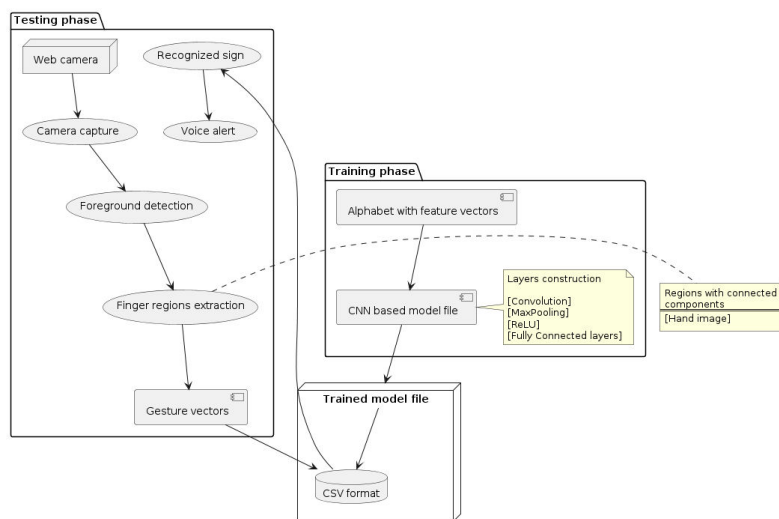


Fig 3: Proposed Block diagram

The suggested work can be divided into two phases, such as the training and testing phases, as shown in this diagram. CNN method is used to train vector files with features during the training phase. During the testing phase, the user's hand is captured by the camera to identify the regions of the hand and fingers. Ultimately, gesture values are categorized and a verbal notice regarding detected signs is provided.

V. CONCLUSION

Through speaking, everyone is able to understand one another and communicate ideas very powerfully. In order to help blind people communicate, our project aims to bridge the gap by adding a low-cost computer to the mix. This will enable sign language to be recognized, recorded, and converted into speech. This article identifies the handmade motions using an image processing approach. This application illustrates a modern integrated system developed for those with hearing difficulties. The user's data collection can be facilitated by the camera-based zone of interest. According to this survey, real-time streaming based on vision and CNN algorithm is more efficient than current sensor and hybrid approaches. Hybrid approaches based on sensors are expensive and challenging to install.

REFERENCES

1. Shah, Farman, et al. "Sign language recognition using multiple kernel learning: A case study of Pakistan sign language." *Ieee Access* 9 (2021): 67548-67558.
2. Sharma, Shikhar, and Krishan Kumar. "ASL-3DCNN: American sign language recognition technique using 3-D convolutional neural networks." *Multimedia Tools and Applications* 80.17 (2021): 26319-26331.
3. Adaloglou, Nikolaos M., et al. "A comprehensive study on deep learning-based methods for sign language recognition." *IEEE Transactions on Multimedia* (2021).
4. Kratimenos, Agelos, Georgios Pavlakos, and Petros Maragos. "Independent sign language recognition with 3d body, hands, and face reconstruction." *ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2021.
5. Elouariachi, Ilham, et al. "Explicit quaternion krawtchouk moment invariants for finger-spelling sign language recognition." *2020 28th European Signal Processing Conference (EUSIPCO)*. IEEE, 2021.
6. Li, Dongxu, et al. "Transferring cross-domain knowledge for video sign language recognition." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020.
7. Sincan, OzgeMercanoglu, et al. "Chalearn LAP large scale signer independent isolated sign language recognition challenge: Design, results and future research." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2021.
8. Boháček, Matyáš, and Marek Hruz. "Sign pose-based transformer for word-level sign language recognition." *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. 2022
9. Katoch, Shagun, Varsha Singh, and Uma Shanker Tiwary. "Indian Sign Language recognition system using SURF with SVM and CNN." *Array* 14 (2022): 100141.
10. Kothadiya, Deep, et al. "Deepsign: Sign language detection and recognition using deep learning." *Electronics* 11.11 (2022): 1780.
11. A. Wadhawan and P. Kumar, "Deep learning-based sign language recognition system for static signs," *Neural Comput. Appl.*, vol. 32, no. 12, pp. 7957–7968, Jun. 2020.
12. M. Jiang, J. Dong, D. Ma, J. Sun, J. He, and L. Lang, "Inception spatial temporal graph convolutional networks for skeleton-based action recognition," in *Proc. Int. Symp. Control Eng. Robot. (ISCR)*, Feb. 2022, pp. 208–213.
13. P. Das and A. Ortega, "Symmetric sub-graph spatio-temporal graph convolution and its application in complex activity recognition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2021, pp. 3215–3219.
14. Z. Huang, X. Shen, X. Tian, H. Li, J. Huang, and X.-S. Hua, "Spatio-temporal inception graph convolutional networks for skeleton-based action recognition," in *Proc. 28th ACM Int. Conf. Multimedia*, Oct. 2020, pp. 2122–2130.
15. A. A. Hosain, P. SelvamSanthalingam, P. Pathak, H. Rangwala, and J. Kosecka, "Hand pose guided 3D pooling for word-level sign language recognition," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2021, pp. 3429–3439.
16. J. Li, C. Xu, Z. Chen, S. Bian, L. Yang, and C. Lu, "HybrIK: A hybrid analytical-neural inverse kinematics solution for 3D human pose and shape estimation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 3383–3393.
17. D. Li, C. R. Opazo, X. Yu, and H. Li, "Word-level deep sign language recognition from video: A new large-scale dataset and methods comparison," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2020, pp. 1459–1469.
18. A. Tunga, S. V. Nuthalapati, and J. Wachs, "Pose-based sign language recognition using GCN and BERT," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. Workshops (WACVW)*, Jan. 2021, pp. 31–40.
19. M. Boháček and M. Hruz, "Sign pose-based transformer for word-level sign language recognition," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis. Workshops (WACVW)*, Jan. 2022, pp. 182–191
20. H. Luqman and E.-S.-M. El-Alfy, "Towards hybrid multimodal manual and non-manual Arabic sign language recognition: MArSL database and pilot study," *Electronics*, vol. 10, no. 14, p. 1739, Jul. 2021.
21. Z. Liu, L. Pang, and X. Qi, "MEN: Mutual enhancement networks for sign language recognition and education," *IEEE Trans. Neural Netw. Learn. Syst.*, early access, May 25, 2022, doi: 10.1109/TNNLS.2022.3174031.
22. Y. C. Bilge, R. G. Cinbis, and N. Ikizler-Cinbis, "Towards zero-shot sign language recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 1, pp. 1217–1232, Jan. 2023.

23. H. Park, Y. Lee, and J. Ko, "Enabling real-time sign language translation on mobile platforms with on-board depth cameras," *Proc. ACM Interact., Mobile, Wearable Ubiquitous Technol.*, vol. 5, no. 2, pp. 1–30, 2021.
24. C. C. dos Santos, J. L. A. Samatelo, and R. F. Vassallo, "Dynamic gesture recognition by using CNNs and star RGB: A temporal information condensation," *Neurocomputing*, vol. 400, pp. 238–254, Aug. 2020.
25. J. Imran and B. Raman, "Deep motion templates and extreme learning machine for sign language recognition," *Vis. Comput.*, vol. 36, no. 6, pp. 1233–1246, Jun. 2020.
26. O. M. Sincan and H. Y. Keles, "Using motion history images with 3D convolutional networks in isolated sign language recognition," *IEEE Access*, vol. 10, pp. 18608–18618, 2022.
27. S. Jiang, B. Sun, L. Wang, Y. Bai, K. Li, and Y. Fu, "Skeleton aware multimodal sign language recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2021, pp. 3413–3423.
28. O. M. Sincan, J. C. S. Jacques, S. Escalera, and H. Y. Keles, "ChaLearn LAP large scale signer independent isolated sign language recognition challenge: Design, results and future research," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2021, pp. 3472–3481.
29. O. M. Sincan and H. Y. Keles, "AUTSL: A large scale multi-modal Turkish sign language dataset and baseline methods," *IEEE Access*, vol. 8, pp. 181340–181355, 2020.
30. J. Wang, "Deep high-resolution representation learning for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 10, pp. 3349–3364, Oct. 2020.
31. Y.-F. Song, Z. Zhang, C. Shan, and L. Wang, "Stronger, faster and more explainable: A graph convolutional baseline for skeleton-based action recognition," in *Proc. 28th ACM Int. Conf. Multimedia*, Oct. 2020, pp. 1625–1633.
32. D. Li, X. Yu, C. Xu, L. Petersson, and H. Li, "Transferring cross-domain knowledge for video sign language recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 6205–6214.
33. B. Subramanian, B. Olimov, S. M. Naik, S. Kim, K.-H. Park, and J. Kim, "An integrated mediapipe-optimized GRU model for Indian sign language recognition," *Sci. Rep.*, vol. 12, no. 1, pp. 1–16, Jul. 2022.
34. J. Rasley, S. Rajbhandari, O. Ruwase, and Y. He, "DeepSpeed: System optimizations enable training deep learning models with over 100 billion parameters," in *Proc. 26th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2020, pp. 3505–3506.
35. Obi, Yulius, et al. "Sign language recognition system for communicating to people with disabilities." *Procedia Computer Science* 216 (2023): 13-20.
36. Kim, Hyeon-Jun, and Soo-Whang Baek. "Application of wearable gloves for assisted learning of sign language using artificial neural networks." *Processes* 11.4 (2023): 1065.
37. Das, Soumen, Saroj Kr Biswas, and BiswajitPurkayastha. "Automated Indian sign language recognition system by fusing deep and handcrafted feature." *Multimedia Tools and Applications* 82.11 (2023): 16905-16927.
38. Das, Sunanda, et al. "A hybrid approach for Bangla sign language recognition using deep transfer learning model with random forest classifier." *Expert Systems with Applications* 213 (2023): 118914.
39. Chen, Xuena, et al. "Sign language gesture recognition and classification based on event camera with spiking neural networks." *Electronics* 12.4 (2023): 786.
40. Gangrade, Jayesh, and Jyoti Bharti. "Vision-based hand gesture recognition for Indian sign language using convolution neural network." *IETE Journal of Research* 69.2 (2023): 723-732.
41. Damaneh, MahinMoghbeli, FarahnazMohanna, and Pouria Jafari. "Static hand gesture recognition in sign language based on convolutional neural network with feature extraction method using ORB descriptor and Gabor filter." *Expert Systems with Applications* 211 (2023): 118559.
42. Nandi, Utpal, et al. "Indian sign language alphabet recognition system using CNN with diffGrad optimizer and stochastic pooling." *Multimedia Tools and Applications* 82.7 (2023): 9627-9648.
43. Wolfe, Rosalee, et al. "Special issue on sign language translation and avatar technology." *Universal Access in the Information Society* (2023): 1-3.
44. Venugopalan, Adithya, and Rajesh Reghunadhan. "Applying Hybrid Deep Neural Network for the Recognition of Sign Language Words Used by the Deaf COVID-19 Patients." *Arabian Journal for Science and Engineering* 48.2 (2023): 1349-1362.
45. Balaha, Mostafa Magdy, et al. "A vision-based deep learning approach for independent-users Arabic sign language interpretation." *Multimedia Tools and Applications* 82.5 (2023): 6807-6826.
46. Alzubaidi, Mohammad A., MwaffaqOtoom, and Areen M. Abu Rwaq. "A novel assistive glove to convert arabic sign language into speech." *ACM Transactions on Asian and Low-Resource Language Information Processing* 22.2 (2023): 1-16.
47. Patil, Kiratey, GayatriPendharkar, and G. N. Gaikwad. *American sign language detection*. Diss. CALIFORNIA STATE UNIVERSITY, NORTHRIDGE, 2023.
48. Bora, Jyotishman, et al. "Real-time assamese sign language recognition using mediapipe and deep learning." *Procedia Computer Science* 218 (2023): 1384-1393.
49. Núñez-Marcos, Adrián, Olatz Perez-de-Viñaspre, and GorkaLabaka. "A survey on Sign Language machine translation." *Expert Systems with Applications* 213 (2023): 118993.

50. Zheng, Jiangbin, et al. "Cvt-slr: Contrastive visual-textual transformation for sign language recognition with variational alignment." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details