JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A METHODOLOGICAL AND STRUCTURAL REVIEW OF HAND GESTURE RECOGNITION: AN UPDATE FROM 2018 TO 2025

[1] **Siddangouda Hosamani,** [2] **Dr. Jithendra P. R. Nayak**

[1] Lecturer Selection Grade 1, Electronics and Communication Engineering Department, Government Polytechnic Hubli, Karnataka, India, sidduhosamani@yahoo.com

[2] Research Professors in Computer Science and Engineering, Srinivas Institute of Technology, Mukka, Srinivas University, Mangaluru, Karnataka, India, jithendraatmanphdom@gmail.com

ABSTRACT: With the development of today's technology, and as humans tend to naturally use hand gestures in their communication process to clarify their intentions, hand gesture recognition is considered to be an important part of Human Computer Interaction (HCI), which gives computers the ability of capturing and interpreting hand gestures, and executing commands afterwards. The aim of this study is to perform a systematic literature review for identifying the most prominent techniques, applications and challenges in hand gesture recognition. To conducts a systematic review of existing literature review on hand gesture recognition, with a particular focus on existing methods that address the application of vision, sensor, and hybrid-based methods in the context of hand gesture recognition. This systematic review covers the period from 2018 to 2025, making use of prominent databases including IEEE Xplore, Science Direct, Scopus, and Web of Science. The chosen articles were carefully examined according to predetermined criteria for inclusion and disqualification. Our main focus was on evaluating the hand gesture representation, data acquisition, and accuracy of vision, sensor, and hybrid-based methods for recognizing hand gestures. The results of this paper can be summarized as the following; the accuracy of discernment in scenarios that rely on the specific signer varies from 64% to 98%, with an average of 87.9% among the studies that were analyzed. On the other hand, in situations where the signer's identity is not important, the accuracy of recognition ranges from 52% to 98%, with an average of 79% based on the research analyzed. The problems observed in continuous gesture identification highlight the need for more research efforts to improve the practical feasibility of vision-based gesture recognition systems. The findings also indicate that the size of the dataset continues to be a significant obstacle to hand gesture detection. Hence, this study seeks to provide a guide for future research by examining the academic motivations, challenges, and recommendations in the developing field of sign language recognition. The paper will discuss the gesture acquisition methods, the feature extraction process, the classification of hand gestures, the applications that were recently proposed, the challenges that face researchers in the hand gesture recognition process, and the future of hand gesture recognition. We shall also introduce the most recent research from the year 2016 to the year 2018 in the field of hand gesture recognition for the first time.

***Index Terms* -** Sign language recognition (SLR), vision-based hand gesture, hand gesture recognition (HGR), sensor-based hand gesture, hybrid-based hand gesture, classification, feature extraction.

I. INTRODUCTION

In our daily interactions, non-verbal communication plays a crucial role, conveying approximately 65% of human messages, compared to verbal communication, which accounts for only 35% [1], [2]. Nowadays, people increasingly use gestures to control daily devices such as televisions, computers, fans, and air-conditioning systems. Non-verbal communication includes body gestures like head movements, facial expressions, nodding, shaking the head, mouth movements, winking, eye gaze direction, body movements, arm gestures, and hand gestures. Effective HGR methods ensure robust human-computer interaction (HCI), offering alternatives to traditional tools like mice and keyboards [3]. Hand gestures are essential in daily activities, and automatic HGR is vital for natural nonverbal communication. SLR is particularly important in bridging the gap between hearing impaired and general communities by translating hand movements into speech or text, which is invaluable for communication, education, and rehabilitation, especially without a human interpreter. Despite extensive research on static and dynamic HGR systems for various applications, several challenges remain. These challenges include adapting to diverse inputs like environmental noise, signer variations, and language differences [4].

Meanwhile, gesture, defined as deliberate and expressive bodily motions involving the hands, fingers, face, arms, body, or head, serve as a prominent form of non-verbal communication in human interactions [5], [6]. Within the technology domain, tools of authority find heightened preference for leveraging hand gestures over alternative forms of gestural communication. This preference is substantiated by the facilitation of navigation and sustenance within the technology environment through the use of hand gestures

[7]. Consequently, an imperative arises for the employment of appropriate approaches in gesture recognition to interpret human hand gestures within machine learning settings. Gesture recognition, in this context, pertains to the identification of class labels from videos or images featuring gestures executed by users. This recognition capability assumes paramount significance in discerning and responding to the nuanced intricacies of hand gestures within the virtual domain [8].

In the domain of virtual reality (VR), the application of gesture recognition models is pervasive across various fields [9]. For instance, author [10] proposed an application for computer mouse control, using an algorithm and specific hand features to optimize performance and enhance user comfort. In a different context, author [11] presented a method for automatic gesture recognition intended to recognize hand gestures during virtual reality (VR) training in crane rigging operations. Similarly. Proposing a universal approach, author [12] recommended a standardized set of gestures for VR interaction, drawing inspiration from the versatility observed in desktop computing mouse control across diverse applications.

Hand movements in this context are classified into two distinct categories: ‘‘static’’ and ‘‘dynamic.’’ A static-gesture is similar to a signature, in which the precise hand movements do not contribute significantly to the gesture. Instead, the hand itself is of utmost importance [13]. On the contrary, dynamic hand gestures are contingent upon both the shape and motion of the hand, constituting essential components of the gesture and a critical aspect of human motion perception. Nevertheless, the complexity of this task is further complicated by the greater variety in hand shapes and significant interferences found between fingers, which poses a significant difficulty for accurately capturing dynamic hand movements utilizing single-camera video sensors. The performance of video-based hand gesture detection is greatly limited by these limitations [14].

SLR represents a critical domain of investigation with the overarching objective of ameliorating communication barriers for the hearing impaired-mute community as mentioned. This area of research seeks to employ computer vision technology to translate sign language gestures into either textual or spoken formats [15]. Its multidisciplinary nature integrates elements of computer science, artificial intelligence, and linguistics to address the intricate challenges posed by the swift and highly co articulated motions inherent in sign language. The complexity of recognizing gesture sequences within sign language compounds the difficulty of this endeavour [16].

Furthermore, the purview of sign language recognition extends beyond general applications to encompass specific domains such as finger spelling recognition and hand pose recognition. These domain-specific advancements cater to the diverse linguistic needs inherent in different sign languages. Consequently, the focus of research lies in the creation of SL systems that are designed specifically for particular SL, such as Arabic SL and Indigenous People SL [17]. This underscores the imperative of adopting language specific approaches to account for the nuanced intricacies of different sign languages. These language-specific endeavours not only enhance the inclusivity of sign language recognition but also contribute to a more comprehensive understanding of the diverse linguistic modalities within the hearing impaired mute community. Moreover, the integration of contextual and contextualized information within sign language recognition systems represents an area of burgeoning interest. The integration of contextual indicators, including body language and facial expressions, is designed to enhance the precision and resilience of the recognition procedure. This holistic approach acknowledges the multimodal nature of sign language communication, recognizing that gestures are embedded within a broader context of non-verbal expression [18].

Thus, the aim of this research is to perform a thorough examination of current methods used for hand gesture recognition, more specifically for sign language recognition (SLR), in order to gain a present state of hand gesture recognition systems in this field. Furthermore, this study attempts to present a roadmap for the technological evolution of SLR systems, delineating their features and elucidating the existing limitations of current technology. The methodology used in this study adheres to the principles outlined in Noraini’s paper [8], Zinah’s paper [3] and other pertinent research publications, thereby ensuring a systematic approach to both the Systematic Literature Review and the Systematic Mapping Study conducted herein.

II. PROBLEM BACKGROUND

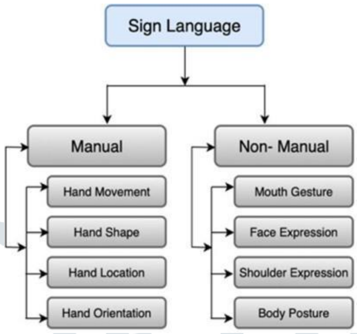
HGR is a technology that translates hand movements in sign language into text or speech. It can be divided into vision-based and device-based systems based on how they capture hand gestures [19], [20]. Vision-based HGR systems offer a more natural interaction experience as users do not need to wear any cumbersome devices. These systems find wider applications in outdoor settings due to their ease of use. However, challenges arise in handling dynamic sign language datasets containing both isolated and continuous gestures. Existing research focuses on recognizing isolated gestures, limiting their practical applicability. More robust feature extraction and discrimination methods are necessary to enhance vision-based systems. A temporal modelling-based HGR system is also needed. Due to numerous applications, HGR has sparked significant research interest, as highlighted in numerous review papers [21].

In 2013, Chen et al. surveyed the HGR methodology, vision-based, depth-based and glove-based approach [22]. Check et al. surveyed the state-of-the-art approach used in the hand gesture-based recognition system in 2019. They summarized the hand feature recognition system’s pre-processing, segmentation, augmentation and classification technique [23]. Aloysius et al. surveyed only vision-based video or continuous sign language recognition (CSLR system) in 2020 [24]. By including the dynamic dataset based HGR, Wadhawan et al. surveyed academic literature spanning from 2007 to 2017, where they included six key dimensions such as dataset collection approach, different types of signs based on time, mode of sign, one-hand or two-hand sign, classification approach and rates of recognition in 2021 [25]. Ratsgoo et al. surveyed vision-based HGR for SLR in 2021 [26]. Recent surveys have further expanded the field, such as Jain et al. provided a comprehensive review of DL approaches for HGR in 2022, focusing on advancements in model architectures and applications in human-computer interaction [26].

In 2024, In light of the existing communication obstacles, numerous researchers have suggested to implement of SLR system, presenting a promising solution to address challenges experienced by individuals with hearing impairments and the broader community. As an example, in a study by [27], an unsupervised learning methodology was suggested as a potential solution to the issue of hand movement in continuous sign language. By implementing vision-based modelling, the establishment of such a system, known as a Sign Language Recognition (SLR) system, is intended to improve the identification and comprehension of sign language gestures. The principal objective is to foster inclusiveness and eradicate enduring communication obstacles that exist between individuals with hearing impairments and the broader community.

Within the domain of sign language, every expression consists of two essential parts: the manual and nonmanual elements. The manual aspects encompass factors like hand movement, orientation, location, and shape, while the non-manual elements include body-posture, mouth gestures, and facial-expressions [3]. It is noteworthy that the primary means of conveying signs predominantly relies on the manual components, as depicted in Figure 4.

Figure 4. Sign Language comprises two integral parts.



The vision-based method entails capturing gesture picture data using a camera and then utilizing image-processing technology to identify motions [28]. This system is designed to be easily used by the user, eliminating the necessity for the user to wear any additional equipment. However, the development of this technology is hard, requiring sophisticated and extensive calculations for the formulation of algorithms in feature and movement recognition [30]. In addition, it is prone to problems related to fluctuations in lighting conditions [28], [29], [30]. In contrast, the sensor-based technique requires the use of a sensory-glove-device to accurately measure finger bending, hand position, and movement. The strategy utilizing a data glove achieves superior precision, rapid reactivity, and improved manoeuvrability [5]. Nevertheless, this approach places a strict limitation on the structure of the hand, leading to a certain level of discomfort [31]. Significantly, it removes the need for pre-processing and segmentation [32].

III. REVIEW OF LITERATURE ON VISION-BASED

In vision-based, the procedure generally consists of many phases: data gathering, pre-processing, segmentation, feature extraction, and classification. These stages are categorized as shown in figure 7 [33]. Static gesture recognition involves analyzing individual image frames, while dynamic sign languages include analyzing continuous video frames. The main distinction between vision-based approaches and sensor-based approaches lies in their respective techniques for acquiring data [34]. The methods and approaches employed by researchers in the field of vision-based gesture-recognition are examined in the sections that follow.

Figure 7. Vision-based recognition process.

A. DATA ACQUISITION

Gesture recognition systems that rely on vision-based data acquisition utilize sequences of images. These systems use various image-capturing devices, including video cameras, webcams, stereo cameras, infrared cameras, and more sophisticated active methods such as Kinect and LMC (Light Measuring Camera). Vision-based approaches analyze the visual information captured by these devices to interpret and recognize gestures [34]. Stereo cameras, Kinect, and LMC are 3D cameras that capture depth and visual data.

B. IMAGE PRE-PROCESSING

Image pre-processing is the phase where image/video inputs are altered to improve the system's performance. Commonly used methods for reducing noise in captured photos or videos involve using median filters and Gaussian filters. It is worth mentioning that in certain research cases [35], just median filtering has been used during the pre-processing phase. In addition, morphological techniques are widely used to remove undesirable information. For example, author [36] utilized a sequence in which the input image initially converted into a binary image using a thresholding technique. Then, gaussian and median filters were applied to eliminate any noise present in the image. Following then, morphological operations were utilized as an essential pre-processing phase.

C. SEGMENTATION

In the domain of image processing, image segmentation pertains to the procedure of partitioning images into discrete and recognizable components [37-40]. This step focuses on isolating the Region-of-Interest (ROI) from the rest of the image. Segmentation techniques are broadly categorized into two types: contextual and non-contextual. Contextual segmentation involves examining spatial relationships between features, often employing techniques like edge detection. In contrast, non-contextual segmentation ignores spatial connections and groups pixels based on overall characteristics [41].

D. FEATURE-EXTRACTION

Feature extraction involves the transformation of significant elements within input data into condensed sets of feature vectors [41]. This process is essential in pattern recognition and data analysis, as it helps distill relevant information from the raw input, facilitating more efficient and effective processing. Within the context of gesture recognition, the extracted features should include pertinent information from the input of hand movements. These features should be presented in a concise way that distinguishes the gesture being classed from other gestures. These can be classified as PCA, LDA, and FEFD.

E. CLASSIFICATION

Machine learning classification involves the use of supervised and unsupervised algorithms [42]. In supervised machine learning, the system undergoes training to identify distinct patterns in input data, and this obtained knowledge is then utilized to make predictions about future data. This method entails using a collection of pre-existing labelled training data to deduce a function that helps identify patterns in new, unlabeled data [43]. In contrast, unsupervised machine learning is specifically designed to extract meaningful information from datasets in which the input data does not have a labelled response [43]. Here is the example that scholars have used.

1) SUPPORT VECTOR MACHINE

Support Vector Machines (SVM) is a supervised machine learning approach designed to identify the optimal hyperplane for categorizing input data points. SVM achieves the maximum margin around the separating hyperplane by employing optimization approaches [44]. Two hyperplanes, which accurately describe the data, are identified. The research conducted on gesture databases provide evidence that linear kernel SVM outperforms non-linear Gaussian kernels. Specifically, the accuracy of linear SVM classification with a set of 14 ESL declined from 99.2% to 82.3% as the number of gestures increased to 25 ESL. Encouraging outcomes have been attained through significant investigations exploring the application of Scale-Invariant Feature Transform (SIFT) in feature extraction. Subsequently, these features undergo quantization via K-means clustering, and the mapping into a Bag-of-Features (BoF) is performed, followed by classification using Support Vector Machines (SVM).

2) ARTIFICIAL NEURAL NETWORK

Artificial Neural Networks (ANNs) are computational systems designed to process information, mirroring the performance characteristics observed in biological neural networks. In essence, ANNs simulate the way biological neurons interconnect and communicate to enable learning and information processing in a machine context [44]. ANN can be defined by three essential parameters: the arrangement of connections between different layers of neurons, the given weights for these connections, and the activation function that determines the behaviour of each neuron.

3) K-NEAREST NEIGHBOR (KNN)

The K-Nearest Neighbours (K-NN) method is a non-parametric statistical approach for classifying input data based on the majority vote from its neighbouring data points. The class assigned to the data is determined by the most prevalent class among its k closest neighbours. The literature under consideration reflects a global perspective on sign language recognition, encompassing various languages. American Sign Language (ASL) takes precedence with 56 instances, followed by Arabic, Indigenous People, Chinese, Malay, German, Turkish, Brazilian, and other languages. This highlights the cross-cultural applicability and importance of creating recognition systems tailored to different linguistic nuances. The upcoming Table 1 will demonstrate the literature review of gesture recognition. It will discuss the focus of each proposed methodology, followed by discussions.

Table 1. Literature review of gesture recognition.

REF	UTHORS	YEAR	FOCUS
[45]	Gao et al.	2022	Dynamic Hand Gesture Recognition
[45]	Ding & Zheng	2022	Depth-sensor-based Hand Gesture Recognition
[45]	Wang et al.	2020	Continuous Hand Gesture Recognition Method
[46]	Jhansi*	2020	Hand Calculator System using CNN
[46]	J et al	2020	Hand Gesture Recognition with ML Algorithms
[46]	Du et al	2018	Gesture Recognition Based on Depth Info
[47]	Nanani et al	2018	Real-time Hand Gesture Recognition System
[48]	Prabhu & Sasikala	2018	Survey on Hand Gesture Recognition Systems
[48]	Kolhe et al	2017	Part-Based Hand Gesture Recognition
[49]	Gao et al	2017	Static Hand Gesture Recognition
[50]	Cheng et al	2016	3D Hand Gesture Recognition
[51]	Rahman & Afrin	2013	Multiclass Support Vector Machine
[52]	Sarkar et al.	2013	Hand Gesture Recognition Systems Survey
[53]	Ibraheem & Khan	2012	Various Gesture Recognition Technologies

The discussion of this body of work illuminates a clear trajectory from simple recognition in controlled environments to sophisticated, real-time interaction in complex settings. The integration of multimodal data, advancements in sensor technology, and the application of cutting-edge machine learning algorithms have transformed vision-based hand gesture recognition into a dynamic field with significant implications for how humans interact with machines. As the technology continues to mature, future research will likely tackle the remaining challenges of generalizability across diverse user populations and environments, as well

as the seamless integration of these systems into a broader range of applications, from assistive technologies to immersive virtual reality experiences.
On the other hand, Table 2 is set to review the literature on vision-based dynamic gesture recognition. It categorizes the body of work into four principal sections: classification, feature extraction, segmentation, and scope as static table. The organization of the table aims to facilitate an in-depth examination of the different strategies utilized in the identification of static gestures, offering an insightful look at their distinct features and the extent of their implications in various applications.
As it can be seen from this table, the array of research from 2018 to 2025 on vision-based dynamic gesture

Table 2. Literature review of vision based static gesture recognition.

REFERENCES	AUTHORS	YEAR	CLASSIFICATION	FEATURE EXTRACTION	SEGMENTATION	SCOPE
[53]	Gao et al.	2022	Dynamic Hand Gesture	3D Pose Estimation	Human-Robot Interaction	Human-Robot Interaction
[53]	Rajbonshi et al.	2022	Real-Time Hand Gesture	Static & Dynamic Gesture	Media Players	Media Players
[54]	Ding & Zheng	2022	RGB-D Depth-sensor-based	Deep Learning	Shadow Effect Removal	Smart Gesture Communication
[54]	Neethu et al.	2021	SVM-based	Distance Transform	Autonomous Vehicles	Autonomous Vehicle Applications
[55]	Krish et al.	2020	Machine Learning	Data Acquisition	Pre-processing	Human-Computer Interaction
[55]	Shah et al.	2019	Vision-based	Hand normalization	Manual Placement	Hand Gesture Recognition
[56]	Jiang et al.	2018	Gesture Recognition	Depth Information	CNN	Natural Communication
[57]	Liu et al	2018	Real-Time Hand Gesture	Data Gloves	Vision-based	Hand Gesture Recognition System
[57]	Cheng et al.	2016	3D Hand Gesture	Hand modelling	Gesture recognition	Gesture Recognition
[58]	Gao et al.	2017	Static Hand Gesture	CNNs	Space HRI	Human-Computer Interaction
[58]	Xu et al.	2013	Sparse Representation	Kinect-based	Hand Gesture Recognition	Human-Computer Interaction
[59]	Raj et al.	2012	FPGA-based	Gesture recognition	Artificial Intelligence	Computing and Image Processing
[60]	Zhang & Yun	2010	Robust Gesture Recognition	Distance recognition	Skin-colour Segmentation	Gesture Recognition
[61]	Pavlovic et al.	1997	Visual interpretation	Gesture representation	Feature extraction	Human-Computer Interaction

Recognition showcases remarkable progress and diversity in methodologies, reflecting the interdisciplinary nature of the field. Initially, methods were rooted in manual feature labelling and basic machine vision techniques, as depicted in Pavlović et al.'s 1997 seminal work, which focused on the visual interpretation of hand gestures for human-computer interaction. These efforts laid the groundwork for the integration of more sophisticated machine learning algorithms and sensor technologies. By 2013, the field had advanced to incorporate featured-based segmentation methods using neural networks, indicative of the growing trend towards artificial intelligence in gesture recognition. The subsequent years saw a significant tilt towards multimodal and sensor-based approaches, especially with the use of CNNs for gesture recognition, such as the two-stream CNN framework for American Sign Language recognition in 2015 and 2019. These models leveraged multimodal data fusion and deep learning to interpret complex gesture dynamics more effectively.
Future research will likely continue to push the boundaries of what's possible with vision-based hand gesture recognition, delving into areas such as three-dimensional modelling, gesture prediction, and the integration of haptic feedback to provide a more tactile experience. As these technologies mature, we can expect them to become increasingly inte- grated into our everyday lives, changing the way we interact with our digital environments.

VIII. CONCLUSION

The current landscape highlights a notable lack in persons' proficiency in using sign language to communicate with the hearing impaired. It is essential to rectify this deficiency in order to promote social engagement with this community. This research primarily investigates sign language recognition (SLR) within the field of hand gesture recognition. This study conducts a systematic analysis of the literature to thoroughly examine the current status, difficulties, reasons, and suggestions related to the recognition of sign language in the field of hand gestures.
The focus of our systematic literature review was to investigate several methods for recognizing sign language using vision-based, sensor-based, and hybrid-based approaches specifically for hand motions. Vision-based methods lever- age visual information, sensor-based approaches acquire data from sensors embedded in gloves, capturing parameters like bend, hand orientation, and rotation. This sensor-based method demonstrates resilience to environmental conditions, ensuring more accurate data by mitigating factors such as performer location and background conditions. However, it is acknowledged that the sensor-based approach has its drawbacks, being perceived as cumbersome and bulky due to the requirement of wearing multiple boards and sensors for precise sign capture.

Nevertheless, the authors suggests that the trajectory of future hand gesture recognition studies should encompass the translation of non-manual signs, an aspect often overlooked in existing works. Moreover, expanding datasets to encom- pass a more extensive lexicon, particularly dynamic words, is identified as a critical imperative. While prevailing models predominantly focus on isolated sign language recognition, a forthcoming change in perspective requires tackling the difficulties of ongoing SLR. To this end, the study underscores the significance of employing deep learning algorithms for sign classification, positing their potential to enhance the efficacy of gesture recognition. A crucial recommendation put forth is the reduction in the size of hardware used in glove systems to enhance their conformability and mobility. This adjustment aligns with the evolving landscape of wearable technology, emphasizing the importance of unobtrusive and user-friendly designs. Lastly, the authors propose a future investigation into the trade-off between device robustness and sensitivity, acknowledging the need for striking an optimal balance to enhance the overall effectiveness and user experience of sign language recognition systems. These recommendations collectively chart a course for future research supports, aiming to bridge existing gaps and elevate the capabilities of sign language recognition technologies.

REFERENCES

[1] S. M. Miah, J. Shin, M. A. M. Hasan, Y. Okuyama, and A. Nobuyoshi, ‘‘Dynamic hand gesture recognition using effective feature extraction and attention based deep neural network,’’ in Proc. IEEE 16th Int. Symp. Embedded Multicore/Many-Core Syst.-Chip (MCSoC), Kuala Lumpur, Malaysia, Dec. 2023, pp. 241–247.

[2] A. S. M. Miah, M. A. M. Hasan, Y. Tomioka, and J. Shin, ‘‘Hand gesture recognition for multi-culture sign language using graph and general deep learning network,’’ IEEE Open J. Comput. Soc., vol. 5, pp. 144–155, 2024.

[3] D. Sarma and M. K. Bhuyan, ‘‘Methods, databases and recent advancement of vision-based hand gesture recognition for HCI systems: A review,’’ Social Netw. Comput. Sci., vol. 2, no. 6, p. 436, Nov. 2021.

[4] Q. Gao, U. E. Ogenyi, J. Liu, Z. Ju, and H. Liu, ‘‘A two-stream CNN framework for American sign language recognition based on multimodal data fusion,’’ in Proc. 19th U.K. Workshop Comput. Intell., Portsmouth, U.K. Cham, Switzerland: Springer, 2019, pp. 107–118.

[5] M. S. Al-Samarraay, M. M. Salih, M. A. Ahmed, A. A. Zaidan, O. S. Albahri, D. Pamucar, H. A. AlSattar, A. H. Alamoodi, B. B. Zaidan, K. Dawood, and A. S. Albahri, ‘‘A new extension of FDOSM based on Pythagorean fuzzy environment for evaluating and benchmarking sign language recognition systems,’’ Neural Comput. Appl., vol. 34, no. 6, pp. 4937–4955, Mar. 2022, doi: 10.1007/s00521-021-06683-3.

[6] R. E. Nogales and M. E. Benalcázar, ‘‘Hand gesture recognition using machine learning and infrared information: A systematic literature review,’’ Int. J. Mach. Learn. Cybern., vol. 12, no. 10, pp. 2859–2886, Oct. 2021, doi: 10.1007/s13042-021-01372-y.

[7] H. Zahid, M. Rashid, S. Hussain, F. Azim, S. A. Syed, and A. Saad, ‘‘Recognition of Urdu sign language: A systematic review of the machine learning classification,’’ PeerJ Comput. Sci., vol. 8, p. e883, Feb. 2022, doi: 10.7717/peerj-cs.883.

[8] N. Mohamed, M. B. Mustafa, and N. Jomhari, ‘‘A review of the hand gesture recognition system: Current progress and future directions,’’ IEEE Access, vol. 9, pp. 157422–157436, 2021, doi: 10.1109/ACCESS.2021.3129650.

[9] S. Wang, A. Wang, M. Ran, L. Liu, Y. Peng, M. Liu, G. Su, A. Alhudhaif, F. Alenezi, and N. Alnaim, ‘‘Hand gesture recognition framework using a lie group based spatio-temporal recurrent network with multiple hand-worn motion sensors,’’ Inf. Sci., vol. 606, pp. 722–741, Aug. 2022, doi: 10.1016/j.ins.2022.05.085.

[10] A. Harichandran and J. Teizer, ‘‘Automated recognition of hand gestures for crane rigging using data gloves in virtual reality,’’ in Proc. Int. Symp. Autom. Robot. Construction, Int. Assoc. Autom. Robot. Construct. (IAARC), 2022, pp. 304–311, doi: 10.22260/isarc2022/0043.

[11] A. H. Hoppe, D. Klooz, F. van de Camp, and R. Stiefelhagen, ‘‘Mouse-based hand gesture interaction in virtual reality,’’ in Proc. Int. Conf. Hum.-Comput. Interact., 2023, pp. 192–198, doi: 10.1007/978-3-031-36004-6_26.

[12] M. J. Cheok, Z. Omar, and M. H. Jaward, ‘‘A review of hand gesture and sign language recognition techniques,’’ Int. J. Mach. Learn. Cybern., vol. 10, no. 1, pp. 131–153, Jan. 2019, doi: 10.1007/s13042-017-0705-5.

[13] C. Chansri and J. Srinonchat, ‘‘Hand gesture recognition for Thai sign language in complex background using fusion of depth and color video,’’ Proc. Comput. Sci., vol. 86, pp. 257–260, Jan. 2016, doi: 10.1016/j.procs.2016.05.113.

[14] A. Sharma, A. Mittal, S. Singh, and V. Awatramani, ‘‘Hand gesture recognition using image processing and feature extraction techniques,’’ Proc. Comput. Sci., vol. 173, pp. 181–190, Jan. 2020, doi: 10.1016/j.procs.2020.06.022.

[15] E.-S.-M. El-Alfy and H. Luqman, ‘‘A comprehensive survey and taxonomy of sign language research,’’ Eng. Appl. Artif. Intell., vol. 114, Sep. 2022, Art. no. 105198, doi: 10.1016/j.engappai.2022.105198.

[16] V. Adithya and R. Rajesh, ‘‘A deep convolutional neural network approach for static hand gesture recognition,’’ Proc. Comput. Sci., vol. 171, pp. 2353–2361, Jan. 2020, doi: 10.1016/j.procs.2020.04.255.

[17] N. M. Bhiri, S. Ameur, I. Alouani, M. A. Mahjoub, and A. Ben Khalifa, ‘‘Hand gesture recognition with focus on leap motion: An overview, real world challenges and future directions,’’ Expert Syst. Appl., vol. 226, Sep. 2023, Art. no. 120125, doi: 10.1016/j.eswa.2023.120125.

[18] X. Chu, J. Liu, and S. Shimamoto, ‘‘A sensor-based hand gesture recognition system for Japanese sign language,’’ in Proc. IEEE 3rd Global Conf. Life Sci. Technol., Mar. 2021, pp. 311–312, doi: 10.1109/LifeTech52111.2021.9391981.

[19] A. Maitrey, V. Tyagi, K. Singhal, and S. Gupta, ‘‘A framework for sign language to speech conversion using hand gesture recognition method,’’ in Proc. Int. Conf. Comput. Intell., Commun. Technol. Netw. (CICTN), Apr. 2023, pp. 333–338, doi: 10.1109/CICTN57981. 2023.10140730.

[20] D. Hatibaruah, A. K. Talukdar, and K. K. Sarma, ‘‘A static hand gesture based sign language recognition system using convolutional neural networks,’’ in Proc. IEEE 17th India Council Int. Conf. (INDICON), Dec. 2020, pp. 1–6, doi: 10.1109/INDICON49873.2020.9342405.

[21] N. Buckley, L. Sherrett, and E. Lindo Secco, ‘‘A CNN sign language recognition system with single & double-handed gestures,’’ in Proc. IEEE 45th Annu. Comput., Softw., Appl. Conf. (COMPSAC), Jul. 2021, pp. 1250–1253, doi: 10.1109/COMPSAC51774.2021.00173.

[22] L. Yuan, W. Qi, K. Cai, C. Li, Q. Qian, and Y. Zhou, ‘‘Gesture recognition device based on cross reticulated graphene strain sensors,’’ J. Mater. Sci., Mater. Electron., vol. 32, no. 7, pp. 8410–8417, Apr. 2021, doi: 10.1007/s10854-021-05448-x.

[23] R. T. Johari, R. Ramli, Z. Zulkoffli, and N. Saibani, ‘‘A systematic literature review on vision-based hand gesture for sign language translation,’’ Jurnal Kejuruteraan, vol. 35, no. 2, pp. 287–302, Mar. 2023, doi: 10.17576/jkukm-2023-35(2)-03.

[24] L. I. B. López, F. M. Ferri, J. Zea, Á. L. V. Caraguay, and M. E. Benalcázar, ‘‘CNN-LSTM and post-processing for EMG-based hand gesture recognition,’’ Intell. Syst. With Appl., vol. 22, Jun. 2024, Art. no. 200352, doi: 10.1016/j.iswa.2024.200352.

[25] D. H. Neiva and C. Zanchettin, ‘‘Gesture recognition: A review focusing on sign language in a mobile context,’’ Expert Syst. Appl., vol. 103, pp. 159–183, Aug. 2018, doi: 10.1016/j.eswa.2018.01.051.

[26] S. M. M. M. Mahin, M. R. Islam, and S. M. M. Ahsan, ‘‘Phrase level Bangla sign language recognition using keypoints from hand gesture video,’’ in Proc. Int. Conf. Next-Gener. Comput., IoT Mach. Learn. (NCIM), Jun. 2023, doi: 10.1109/ncim59001.2023.10212460.

[27] B. Sundar and T. Bagyammal, ‘‘American sign language recognition for alphabets using MediaPipe and LSTM,’’ Proc. Comput. Sci., vol. 215, pp. 642–651, Jan. 2022, doi: 10.1016/j.procs.2022.12.066.

[28] C. K. M. Lee, K. K. H. Ng, C.-H. Chen, H. C. W. Lau, S. Y. Chung, and T. Tsoi, ‘‘American sign language recognition and training method with recurrent neural network,’’ Expert Syst. Appl., vol. 167, Apr. 2021, Art. no. 114403, doi: 10.1016/j.eswa.2020.114403.

[29] D. K. Singh, ‘‘3D-CNN based dynamic gesture recognition for Indian sign language modeling,’’ Proc. Comput. Sci., vol. 189, pp. 76–83, Jan. 2021, doi: 10.1016/j.procs.2021.05.071.

[30] N. F. Attia, M. T. F. S. Ahmed, and M. A. M. Alshewimy, ‘‘Efficient deep learning models based on tension techniques for sign language recognition,’’ Intell. Syst. With Appl., vol. 20, Nov. 2023, Art. no. 200284, doi: 10.1016/j.iswa.2023.200284.

[31] S. Siddique, S. Islam, E. E. Neon, T. Sabbir, I. T. Naheen, and R. Khan, ‘‘Deep learning-based Bangla sign language detection with an edge device,’’ Intell. Syst. With Appl., vol. 18, May 2023, Art. no. 200224, doi: 10.1016/j.iswa.2023.200224.

[32] H. F. T. Ahmed, H. Ahmad, K. Narasingamurthi, H. Harkat, and S. K. Phang, ‘‘DF-WiSLR: Device-free Wi-Fi-based sign language recognition,’’ Pervas. Mobile Comput., vol. 69, Nov. 2020, Art. no. 101289, doi: 10.1016/j.pmcj.2020.101289.

[33] S. Subburaj and S. Murugavalli, ‘‘Survey on sign language recognition in context of vision-based and deep learning,’’ Meas., Sensors, vol. 23, Oct. 2022, Art. no. 100385, doi: 10.1016/j.measen.2022.100385

[34] B. Mallik, M. A. Rahim, A. S. M. Miah, K. S. Yun, and J. Shin, ‘‘Virtual keyboard: A real-time hand gesture recognition-based character input system using LSTM and mediapipe holistic,’’ Comput. Syst. Sci. Eng., vol. 48, no. 2, pp. 555–570, 2024.

[35] M. A. Rahim, A. S. M. Miah, A. Sayeed, and J. Shin, ‘‘Hand gesture recognition based on optimal segmentation in human-computer interaction,’’ in Proc. 3rd IEEE Int. Conf. Knowl. Innov. Invention (ICKII), Aug. 2020, pp. 163–166.

[36] L. Chen, F. Wang, H. Deng, and K. Ji, ‘‘A survey on hand gesture recognition,’’ in Proc. Int. Conf. Comput. Sci. Appl., Dec. 2013, pp. 313–316.

[37] A. Wadhawan and P. Kumar, ‘‘Sign language recognition systems: A decade systematic literature review,’’ Arch. Comput. Methods Eng., vol. 28, no. 3, pp. 785–813, May 2021.

[38] N. Aloysius and M. Geetha, ‘‘Understanding vision-based continuous sign language recognition,’’ Multimedia Tools Appl., vol. 79, nos. 31–32, pp. 22177–22209, Aug. 2020.

[39] R. Rastgoo, K. Kiani, and S. Escalera, ‘‘Sign language recognition: A deep survey,’’ Expert Syst. Appl., vol. 164, Feb. 2021, Art. no. 113794.

[40] M. J. Cheok, Z. Omar, and M. H. Jaward, ‘‘A review of hand gesture and sign language recognition techniques,’’ Int. J. Mach. Learn. Cybern., vol. 10, no. 1, pp. 131–153, Jan. 2019.

[41] R. Jain, R. K. Karsh, and A. A. Barbhuiya, ‘‘Literature review of visionbased dynamic gesture recognition using deep learning techniques,’’ Concurrency Comput., Pract. Exper., vol. 34, no. 22, Oct. 2022, Art. no. e7159.

[42] J. Qi, L. Ma, Z. Cui, and Y. Yu, ‘‘Computer vision-based hand gesture recognition for human-robot interaction: A review,’’ Complex Intell. Syst., vol. 10, no. 1, pp. 1581–1606, Feb. 2024.

[43] T. Tao, Y. Zhao, T. Liu, and J. Zhu, ‘‘Sign language recognition: A comprehensive review of traditional and deep learning approaches, datasets, and challenges,’’ IEEE Access, vol. 12, pp. 75034–75060, 2024.

[44] T. Kapuscinski and K. Inglot, ‘‘Vision-based gesture modeling for signed expressions recognition,’’ Proc. Comput. Sci., vol. 207, pp. 1007–1016, Jan. 2022, doi: 10.1016/j.procs.2022.09.156.

[45] Y. Obi, K. S. Claudio, V. M. Budiman, S. Achmad, and A. Kurniawan, ‘‘Sign language recognition system for communicating to people with disabilities,’’ Proc. Comput. Sci., vol. 216, pp. 13–20, Jan. 2023, doi: 10.1016/j.procs.2022.12.106.

[46] T. Ozcan and A. Basturk, ‘‘Transfer learning-based convolutional neural networks with heuristic optimization for hand gesture recognition,’’ Neural Comput. Appl., vol. 31, no. 12, pp. 8955–8970, Dec. 2019, doi: 10.1007/s00521-019-04427-y.

[47] A. Sharma, N. Sharma, Y. Saxena, A. Singh, and D. Sadhya, ‘‘Bench- marking deep neural network approaches for Indian sign language recognition,’’ Neural Comput. Appl., vol. 33, no. 12, pp. 6685–6696, Jun. 2021, doi: 10.1007/s00521-020-05448-8.

[48] Q. Han, Z. Huangfu, W. Min, T. Ding, and Y. Liao, ‘‘Sign language recog- nition based on skeleton and SK3D-residual network,’’ Multimedia Tools Appl., vol. 83, no. 6, pp. 18059–18072, Jul. 2023, doi: 10.1007/s11042- 023-16117-y.

[49] N. K. Bahia and R. Rani, ‘‘Multi-level taxonomy review for sign language recognition: Emphasis on Indian sign language,’’ ACM Trans. Asian Low- Resource Lang. Inf. Process., vol. 22, no. 1, pp. 1–39, Jan. 2023, doi: 10.1145/3530259.

[50] M. Yasen and S. Jusoh, ‘‘A systematic review on hand gesture recognition techniques, challenges and applications,’’ PeerJ Comput. Sci., vol. 5, p. e218, Sep. 2019, doi: 10.7717/peerj-cs.218.

[51] N. Musthafa and C. G. Raji, ‘‘Real time Indian sign language recognition system,’’ Mater. Today, Proc., vol. 58, pp. 504–508, Jan. 2022, doi: 10.1016/j.matpr.2022.03.011.

[52] J. Bora, S. Dehingia, A. Boruah, A. A. Chetia, and D. Gogoi, ‘‘Real- time assamese sign language recognition using MediaPipe and deep learning,’’ Proc. Comput. Sci., vol. 218, pp. 1384–1393, Jan. 2023, doi: 10.1016/j.procs.2023.01.117.

[53] H. Liang, L. Fei, S. Zhao, J. Wen, S. Teng, and Y. Xu, ‘‘Mask- guided multiscale feature aggregation network for hand gesture recog- nition,’’ Pattern Recognit., vol. 145, Jan. 2024, Art. no. 109901, doi: 10.1016/j.patcog.2023.109901.

[54] H. Kolivand, S. Joudaki, M. S. Sunar, and D. Tully, ‘‘An implementation of sign language alphabet hand posture recognition using geometrical features through artificial neural network (Part 2),’’ Neural Comput. Appl., vol. 33, no. 20, pp. 13885–13907, Oct. 2021, doi: 10.1007/s00521- 021-06025-3.

[55] 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, doi: 10.1007/s00371-019-01725-3.

[56] Y. Lecun, Y. Bengio, and G. Hinton, ‘‘Deep learning,’’ Nature, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[57] X. Jiang, S. C. Satapathy, L. Yang, S.-H. Wang, and Y.-D. Zhang, ‘‘A survey on artificial intelligence in Chinese sign language recognition,’’ Arabian J. Sci. Eng., vol. 45, no. 12, pp. 9859–9894, Dec. 2020, doi: 10.1007/s13369-020-04758-2.

[58] M. Kowdiki and A. Khaparde, ‘‘Automatic hand gesture recognition using hybrid meta-heuristic-based feature selection and classification with dynamic time warping,’’ Comput. Sci. Rev., vol. 39, Feb. 2021, Art. no. 100320, doi: 10.1016/j.cosrev.2020.100320.

[59] B. K. Chakraborty, D. Sarma, M. K. Bhuyan, and K. F. MacDorman, ‘‘Review of constraints on vision-based gesture recognition for human– computer interaction,’’ IET Comput. Vis., vol. 12, no. 1, pp. 3–15, Feb. 2018, doi: 10.1049/iet-cvi.2017.0052.

[60] A. S. Al-Shamayleh, R. Ahmad, M. A. M. Abushariah, K. A. Alam, and N. Jomhari, ‘‘A systematic literature review on vision based ges- ture recognition techniques,’’ Multimedia Tools Appl., vol. 77, no. 21, pp. 28121–28184, Nov. 2018, doi: 10.1007/s11042-018-5971-z.

[61] S. Sharma and S. Singh, ‘‘Vision-based hand gesture recognition using deep learning for the interpretation of sign language,’’ Expert Syst. Appl., vol. 182, Nov. 2021, Art. no. 115657, doi: 10.1016/j.eswa.2021.115657.