Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

A Comprehensive Review of Hand Gesture Recognition: Vision-Based vs. Wearable Sensor Approach

Aaliyah Mohammed Thabit AL-Hisnawi [1*], Mohammad Khalaf Rahim Al-juaifari [2]
[1](University of Kufa /College of Education/department of Computer Science/Najaf, Iraq).
E-mail: aaliyahm.alhisnawi@student.uokufa.edu.iq
[2] (University of Kufa / College of Medicine,Najaf, Iraq)
E-mail: mohammad.aljuaifari@uokufa.edu.iq

Article's Information

Received: 22.12.2025
Accepted: 30.03.2026
Published: 31.03.2026

Keywords:
Gesture Recognition, Machine Learning, Deep Learning, Vision data, Wearable sensor.

Abstract

This comprehensive review systematically compares vision-based and wearable sensor approaches for hand gesture recognition (HGR), addressing a critical gap in existing literature that often treats these modalities in isolation. While previous surveys have predominantly focused on either vision-based methods or sensor-based approaches separately, this review provides an integrated analysis of both paradigms, highlighting their complementary strengths and limitations across the entire HGR pipeline: from data acquisition and preprocessing to feature extraction and classification. We examine diverse datasets including UCI MYO Thalami and RGB-based collections, analyzing preprocessing techniques (data augmentation, noise reduction, normalization) and both traditional machine learning (SVM, ANN, KNN) and deep learning methods (CNN, RNN, LSTM). Our comparative analysis reveals that sensor-based methods excel in controlled environments with precise motion capture, while vision-based approaches offer greater usability at the cost of environmental sensitivity. This review contributes a unified framework for selecting appropriate HGR technologies based on application requirements, computational constraints, and user needs, particularly in healthcare rehabilitation, virtual reality, and assistive technologies for hearing-impaired communities.

https://doi.org/10.46649/fjiece.v5.1.10a.31.3.2026
*Corresponding author: aaliyahm.alhisnawi@student.uokufa.edu.iq

1. INTRODUCTION

Approximately 360 million people worldwide, including 72 million children, are hearing impaired-mute, according to data from the World Federation of the Hearing Impaired and the World Health Organisation [1]. According to RDM, a significant portion of the population with speech and hearing impairments has trouble with traditional literacy [2]. For this reason, the primary means of communication for the mute and hearing impaired is gesture, particularly in sign language, which

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

incorporates manual expressions such as forms, movements, and facial contortions in passing information. However, like most aspects of human communication, sign language does have its flaws. These are; The need to involve most of the hand, restriction to the number of signs used, and the difficulty that is associated with this style of communication [3]. In the meantime, gesture is postcode as a purposeful and intentional movement of hands, fingers, face, arms, body, or head for communication other than verbal conversations [4], [5]. Hand movements in this context are classified into two distinct categories: "static" and "dynamic." There is another type of a gesture, a static gesture, where unlike in dynamic gesture, the way the hands form the specific gesture is not especially important. However, special emphasis must be placed on the hand which is then called a gestural hand. On the other hand, distance and trajectory are both related to dynamic properties as a part of the shape of the hand and resistance about which dynamic hand gestures are contingent upon being fundamental parts of the gesture and an element of human motion understanding. However, the dynamics of this task get even harder due to higher differentiation in hand shapes and the existence of considerable inferences between fingers, which creates a major problem for identifying dynamic hand motion with the help of single-camera video sensors. Some of these limitations greatly hamper the performance of video-based hand gesture detection [6]. Several advanced depth sensors such as the Microsoft Kinetic sensor and LMC have emerged during the last couple of decades to greatly improve object segmentation and three-dimensional hand motion detection capabilities. These elements include gesture recognition, hand feature identification, data acquisition, and hand localization using the features that have been recognized [7]. Traditional approaches to capturing data involve the use of the colour camera, which has been identified to be useful for tasks such as gesture recognition [8]. On the other hand, the concept of encouraging machines to do what we want has been one of the defining goals ever since the advent of machinery [9]. However, as was realized in previous studies [10], there are shortcomings inherent in the existing methods. A major problem arises by the virtue of flexibility and diversity of human gesture. For example, even a simple gesture may be executed in a slightly different way even within the same trial; as such, the recognition of such motions is complex and requires a high amount of flexibility. Furthermore, limitations are seen in the efforts to categorize all human gestural varieties with the help of narrow schemes provoking the search for more delicate and sensitive approaches [11]. The search for an efficient hand localization solution is additionally challenged by the fact that people's movements are very complex and diverse. Several factors, including the position of the hand, the position of objects in the scene, temporal aspect of the gesture add a complexity that goes along with precise hand localization [12]. This review addresses gaps in existing literature by providing: (1) integrated analysis of both vision and sensor approaches, (2) comprehensive coverage of ML to DL evolution, and (3) application-specific selection guidelines.

2. BACKGROUND

As a result, researchers are forced to seek out complex algorithms to not only handle the complexity of a variety of gestures but also to explicitly take into account the timing and relevance aspects of natural human interaction. Thus, the question of occlusions in hand localization still remains the most significant factor of current research. Occlusions, in which one object conceals another object, make hand localization challenging due to the added complexity [13]. The vision-based method involves acquiring gesture picture data through a camera that is followed by picture processing to determine motions [14]. This system is established to be used by the user intuitively and no extra equipment which the user is required to put on. However, it is difficult to develop this technology since there are complex and detailed calculations in the formulation of the algorithms on features and movements [15]. Further, it is also vulnerable to the issues associated with variations in the lighting environments [14, 16, 17]. However, the sensor-based technique needs a sensory glove device to track the finger flexing, position, and motion. The data glove strategy enables precise,

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

responsive, and flexible motion capture [5]. However, this approach imposes a very stringent condition on the structure of the hand and there is a certain amount of uneasiness [18]. Notably, it does away with the requirement of Preprocessing as well as segmentation [19]. The vision system used in the Glove Track has been developed to complement a previously developed hybrid-based technology that combines vision and glove-based approaches on top of sensor readings in order to improve the quality of the visual data. Nevertheless, the use of this method has been limited due to the cost and computational overheads associated with the complete system which forms a scarcity of literature on this subject [20]. The difference in the required vision and sensor systems also emphasizes the main trade-off between user experience and computational expense [3]. To systematically organize the diverse methodologies in hand gesture recognition, Figure 1 presents a comprehensive taxonomy that categorizes approaches based on their underlying sensing technology.

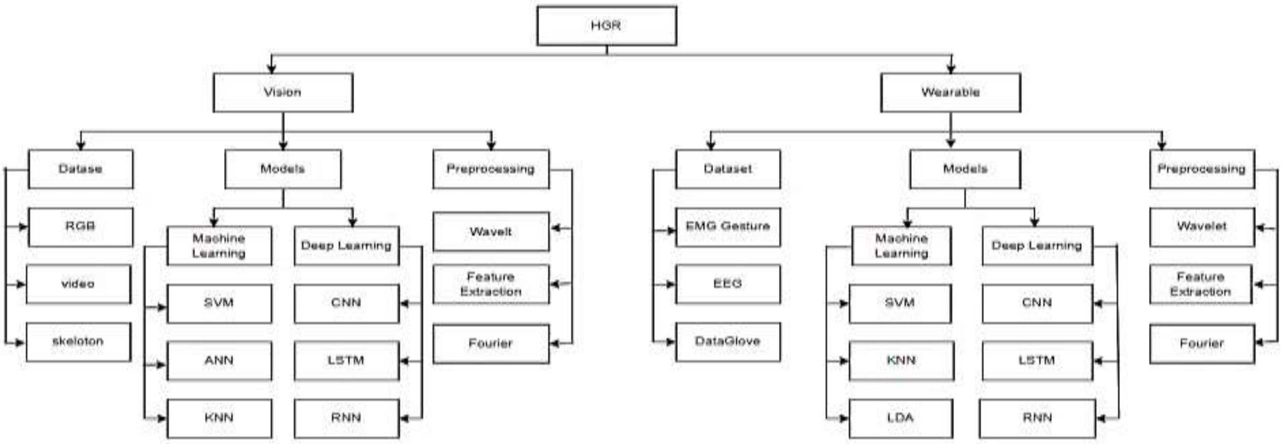

Figure 1. Hand gesture recognition methods

2.1 VISION-BASED METHODS

typically involve a number of steps, including data collection, Preprocessing, segmentation, feature extraction, and classification [21]. Static gesture recognition analyses individual image frames, whereas dynamic sign languages analyse continuous video frames. The primary difference between vision-based and sensor-based methods is how they collect data [6]. The workflow of vision-based systems follows a standardized processing pipeline, as depicted in Figure 2, which illustrates how raw visual data undergoes transformation through multiple computational stages.

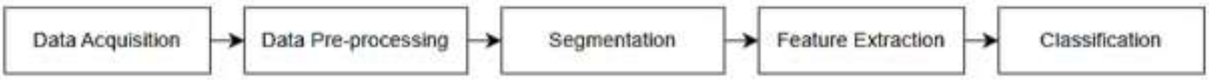

Figure 2. Vision-based recognition process

2.1.1. RGB-MODALITY

Based on HGR concerns still images or video analysis to detect signs and dynamic gesture recognition. RGB datasets, primarily acquired through cameras, is easy to capture and conveys highly descriptive appearance information of scene context [22], [23] This can be either static or dynamic where a single image can encapsulate the meaning of the dynamic gesture and where the number of frames required to fully describe the dynamic gesture [24], [25].

2.1.2. 3D SKELETON MODALITY

Based on HGR Much research has been done on individual gesture recognition but not necessarily in isolation. In the case of skeleton-based dynamic gesture recognition, there are mainly two broad

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

classifications of the techniques, that is the handcrafted feature-based techniques and the deep learning-based techniques. Some of them involve handcrafted features such as skeletal joint coordinates, joint geometric features, and motion features and were used in early works. For example, Yang and Tian [26] developed an inertial feature descriptor of dynamic information of each joint using temporal and spatial differences of skeletal joints. This is then passed through a non-parametric Naive Bayes Nearest Neighbor (NBNN) classifier to identify different actions. Hand direction was defined from the wrist and palm joints by De Smedt et al [27], [28], hand shape was defined using grouped joint coordinates, while hand motion was described from the joint velocities; a linear SVM was used for gesture classification. Although these traditional handcrafted feature-based methods are small and effective, they cannot well recognize the complicated gestures.

2.2 DATA ACQUISITION

The data set description utilising a PC with a Bluetooth receiver to record patterns and a MYO Thalami bracelet worn on the user's forearm. Eight sensors that are evenly distributed over the forearm and concurrently gather myographic signals are included in the bracelet. The signals are transmitted to a PC via a Bluetooth link. We show 36 subjects' raw EMG data as they do a series of motionless hand movements. The subject makes six (seven) fundamental gestures in each of the two sets. Every gesture was made for three seconds, with a three-second break in between Each column contains approximately 40,000–50,000 samples with at least 30,000 guaranteed for analysis.
Systems for gesture recognition that are implemented based on data gathering with the use of vision methods employ sequences of pictures. These systems make use of a variety of image-capturing devices, including webcams, video cameras, stereo cameras, infrared cameras, and more sophisticated active approaches like Kinetic and LMC (Light Measuring Camera). Using methods based on vision, the data that is being recorded and captured using these devices is used to understand and identify different gestures. 3D cameras include stereo cameras, Kinetic, and LMC which facilitate depth and vision information capturing [29].

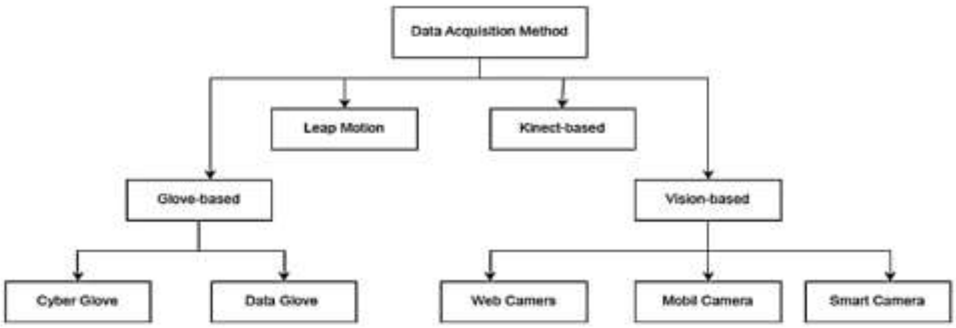

Figure 3. Data acquisition methods.

As Figure 3 on hand shows**,** Comparison of data acquisition modalities in HGR, highlighting 2D RGB cameras versus 3D depth-sensing technologies (stereo cameras, Kinect, LMC) for improved robustness in complex environments. These device cameras improve the robustness and preciseness of the recognition system because the device cameras are capable of recognizing the space orientation and hand movements [30]. The type of capturing device chosen depends on the precision wanted for a particular application and its environment [31].

3. DATA PREPROCESSING

Preprocessing is the process through which the inputs of images/videos are modified to enhance the system's performance. Median filters and Gaussian filters are some of the most traditional noise-reduction techniques applied to photos or videos. However, it is also noteworthy that in some works [16], only median filtering was used during the per-processing stage. At the same time, the application of morphological investigations is done to eliminate undesirable information. For instance, author [31] applied a sequence where the input image is transformed into a binary image through threshold operation. Subsequently, Gaussian and median filters were used to reduce the level of noise within the image. After that, morphological operations were used as the main per-processing step as a result of previous work. Processed images in some research investigations get to reductions in the image resolution before they advance to other phases. According to reference [6], it was identified that making the input image pixels smaller has a positive correlation with computational performance. The research [31] provided important findings for the study that included the time taken to process images at each of the down-sampled resolution levels. The study indicated that the reduction of measurements on the scale to a number that can be divided evenly by 64 was the most effective way; that way, processing time was cut by a whopping 43.8%. Furthermore, in text processing through histogram equalization, the contrast, of input images taken under varying environmental conditions, is improved. There is also the technique of standardizing the brightness and the measure of illumination on the various images in order to enable them to be consistent regardless of the locations they are working to be used in. These preprocessing techniques noted herein serve to illustrate the extent to which the preparation of input signals and noise elimination play a key role in enhancing the quality and feasibility of gesture recognition systems [32].

4. Segmentation

In image processing, when an area is separated into different recognizable zones then this process is known as image segmentation. This stage is very crucial in gesture recognition particularly when trying to differentiate the Region-of-Interest (ROI) from the rest of the image. Segmentation techniques are generally classified into two main types: contextual and non-contextual. Contextual segmentation takes into account the spatial relations of objects in an image; most of them for example use tools like edge detection to distinguish between regions and others. Non-contextual segmentation clumps the same pixels together about their overall features disregarding their location. These approaches assist GR to differentiate parts of an image that contain valuable data to understand gestures; it improves the identification of particular gestures [33].

4.1 Colour Segmentation Method

To clarify, colour segmentation normally involves using one or another colour space like an RGB, Cb-Cr HSV, or HSI [34]. The difficult task of providing reliable colour segmentation is aggravated by the problems related to sensitivity to lighting, type of camera, and skin colour tone [33]. This Research has reported the use of HSV colour space due to the high contrast between the skin colour of the palm and that of the arm to make it easier to segment the palm from the arm [35]. Research [36] for example focuses on the segmentation of the face and hand using the HSV colour. However, the author [37] carries out the skin-color segmentation on the basis of the RGB colour by applying the criterion of as for the method of distinguishing different skin tones, this invention compares $R > G > B$ with the existing skin-color samples. More questions like a study [38]. The work [39] disinvests a normalized RGB against the non-uniformity vulnerability of RGB solely. Thus, the methods on the basis of skin colour distribution and categorized by using skin-color models have been suggested as a strategy to overcome the shortcomings attributable to fixed skin colour thresholds. The skin–colour segmentation was carried out by the author by adopting the YCbCr colour [40]. In [41], the Gaussian model using the YCbCr colour space is used to differentiate between skin pixels and the background. Furthermore, the author [42] uses a system that is similar to [41] but in the new model, the author adopted a Gaussian model and not the histogram model. The authors in [43] propose a dynamic method for modelling skin colour by including measuring elements into both locally trained and globally trained skin models. This approach leads to

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

the development of an adaptable skin colour model. The objective of these developments in skin colour segmentation approaches is to improve precision and flexibility in various circumstances.

4.2 Other Segmentation Method

The author [44] introduced a segmentation method that depends on the difference between the background pictures that has been proven to be very efficient in a complex backdrop situation. The method begins by applying the outs threshold technique to images. This is followed by employing the '3sprincipal' method whereby the aim is to merchandise the variance of the different classes to the maximum. In their work, the author [41] proposed a Hand-Tracking-Segmentation (HTS) framework in their Research. For histogram generation of skin pixels, they used the continuous-adaptive-mean-shift (Cam Shift) algorithm in the HSV colour space. The histogram is then used to determine the right threshold value to be used in the segmentation process. The further steps include the analysis of a canny edge detection followed by the utilization of dilation and erosion. Finally, there will be an application of an edge traversal technique to distinguish between the gesture of the hand and the context in which the gesture is being made.

5. Feature-Extraction

Feature extraction therefore means the process of reduction of important facets of input data into easily manageable sets of features. This process is important as the real world contains a lot of noise which needs to be filtered so that appropriate pattern recognition and data analysis can be achieved as is the case with the raw input data presented. In the case of gesture recognition, the features that need to be extracted must be relevant to the input of hand movements. Such features need to be made in a brief manner in order to differentiate the gesture being classed from other gestures. These can be classified as PCA, and LDA [45].

5.1 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a mathematical tool comprising of the orthogonal transformation in an endeavour to transform the set of usually correlated variables in the observations into another set referred to as the principal components A training set containing M images, expressed by a S dimensional vector, PCA aims at determining a subspace of this dimension This method of data dimensional or reduction is highly beneficial in the sense that it is able to condense the data, into manageable optimum features Many times it is used as a technique for Differentiation where an array of potential-correlated variables is brought into a definite set of irrelevant main components [46].

5.2 Linear Discriminant Analysis (LDA)

While LDA and PCA methodologies try to achieve the purpose of identifying linear transformations of features that adequately explain the inherent characteristics of the data samples, the samples are transformed into low-dimensional space. As described in [47], LDA enhances the degree of separation between various classes of articles by revealing a transformation of features that provide equally valuable discriminant ability for each class. While PCA on the other hand is used to look for the first principal component of the spread of characteristics and does not account for variability within classes. Linear Discriminant Analysis (LDA) has a two-fold role here; it serves as a linear classifier and at the same time is used to reduce dimensional. In [48], the author used the PCA and LDA features from gesture categories and achieved as low as 26% using PCA with a perfect 100% for LDA. The under-performance of Principal Component Analysis (PCA) can be attributed to the problem of over-fitting. Similarly, another work done in [49] investigated the accuracy of both PCA and LDA by employing five over-fitting different categories and 50 input images. The study also showed that PCA managed to reach 60 % accuracy while LDA have a

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

slightly higher of 62% accuracy. Author [50] adopted LDA on sign language in Arabic language based on specifications that are similar to the one explained in reference [51].

5.3 Feature-Extraction in Frequency-Domain

Feature extraction in the frequency domain entails using methods such as the Cosine Transform, Fourier Transform, and Wavelet Transform to convert input data from the time domain to the frequency domain. According to the authors in reference [52], Fourier Descriptors (FD) have advantageous size-invariant properties. Additionally, FD exhibits rotation in-variance, which means that rotation-induced changes in hand motions only result in a phase shift. The most effective way to reduce noise is to eliminate high-frequency components since localised high-frequency oscillations are typically caused by noise and quantisation errors. Conversely, contour-based features such as FD, Wavelet Descriptors (WD), and B-spline may have performance problems, particularly when fingers are curled inward, which results in the loss of contour qualities, according to [53]. Region-based features, like the Principal Curvature-Based Region detector (PCBR), on the other hand, make use of semi-local structural information such edges and curvilinear shapes. This method overcomes the drawbacks of contour-based techniques by exhibiting resilience to changes in intensity, colour, and shape. For example, Haar basis functions up to level two are used in the 2-D Wavelet Packet Decomposition (WPD-2) in [54], which efficiently uses high-frequency channels with a lot of information. The system effectively detects 23 static using a hybrid feature extraction method covered in [55], which combines PCBR, WPD-2, and Convexity flaws. This hybrid outperforms combinations involving only two features when employing KNN and SVM classifiers. Similar positive outcomes are reported in [56], where discrete wavelet Transform (DWT) features are extracted for the classification of 23 static Persian Sign Language (PSL).

6. Classification

In Classification in gesture recognition using Machine learning both the supervised as well as unsupervised learning methods are employed. In supervised learning the system focuses on labelled data because the objective is to define relations the system can use to make outputs on new unlabelled data [57, 58].

However, unsupervised learning is making useful inferences with minimum supervision or even without it at all or at least with little mention of it [59].

6.1. Machine Learning

6.1.1 Support Vector Machine

One type of supervised machine learning that aids in classifying the input feature space is Support Vector Machines (SVM). To obtain the highest margins on the side of the separating hyperplane, SVM calculates optimisation procedures [60]. The best least squares fit to the data is found for two hyper-planes. The superiority of linear kernel SVM over nonlinear Gaussian kernels is supported by studies pertaining to gesture databases. Specifically, when the number of gestures was increased to 25 ESL, the confirmation of linear SVM classification with the set of 14 ESL decreased from 99.2% to 82.3%. Extensive research on feature extraction using Scale-Invariant Feature Transform (SIFT) has produced encouraging results. Following that, the features undergo quantisation using K-means to obtain clustering, mapping into the Bag-of-Features (BoF), and classification using a Support Vector Machine (SVM). In contrast, another technique known as Proximal SVM (PSVM) avoids the necessity of inequality in the conventional SVM by using equality constraints instead. PSVM, which is widely used in many scientific fields, including [61], is capable of handling multiple categories and has a 91% attribute identification rate while classifying 30 TSL. Based on multi-dimensional classification, authors of [62] have confirmed the usefulness of employing the nonlinear SVM to evaluate data in comparison with linear SVM. Similar results were obtained in another

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

investigation [63], where the authors ascribed 99% accuracy to the identification after extracting SIFT features from 30 ASL cases with seven training photographs each.

6.1.2 Artificial Neural Networks (ANN)

ANN are, therefore, information processing systems that emulate the behaviour of biological neural networks with reference to their performance parameters. Specifically, ANN mimic the structure and function of neurons, and how they connect and work in order to allow learning and signal processing in receptive contexts in machines [63]. The topology that links neurones in various layers, the weights given to these connections, and the function that regulates each neuron's response are the three key components that characterise an ANN. Neurones get information in the form of x1, w1, x2, w2,..., where the weights represent the permeability level. The function of a neuron is described as a nonlinear transformation of weighted inputs, following the technique by Author [64], a system implementing gesture recognition was trained using Artificial Neural Networks (ANN) and was able to detect fifteen distinct gestures. A dataset of 7392 gesture signals was used for this. They were able to attain an average classification accuracy of 97.01% by using a single artificial neural network (ANN) with 45 input nodes, 14 output nodes, and two hidden layers. A neural network that combines gesture recognition and fuzzy learning (GRFNN) [65]. Fuzzy control was incorporated into learning parameters to increase precision and reduce the necessity for pre-selection of a training pattern suite. The algorithm identified 36 American Sign Language (ASL) motions with an average forecast accuracy of 93.19%. The Time Delay Neural Network (TDNN) is designed to function effectively in the presence of continuous input data. However, the feed-forward Multi-Layered-Perceptron Neural Network (MLPNN) is capable of classifying non-linear: distinct data.

6.1.3. K-Nearest Neighbour (KNN)

A popular supervised technique the K-Nearest Neighbours (KNN) technique is a probability model where the rule obtained follows a non-parametric approach through probable data points of the input data. For the data, the class is determined from the majority class among the k nearest neighbours of the data. The similarity measure is often calculated using the Euclidean distance, when averages of shapes of the training data to be used to transform the testing data are derived, the Euclidean distance is calculated concerning testing data point to the training data points. The testing data is labelled based on the majority classes of the KNN nearest training data points. The research [66] took a step to perform a comparison between the KNN algorithm and the parametric Bayes classifier, where the result that can be derived is that the KNN was superior to the Bayes classifier. Using KNN in another study [67], 30 test images were classified for the 26 gestures. According to the results presented, the overall accuracy was a very high 90%. However, other research using similar train and test data sizes has shown that while SVM has a higher overall accuracy,

KNN has comparatively a lower one. However, KNN is advantageous because it is computationally efficient and easy to implement [68]. Table1 shows the result of machine learning.

TABLE 1 Result Algorithm of machine learning

Year Ref.	Method	Dataset	Advantage	Disadvantage	accuracy
2021[61]	SVM	ArSL, 30 nstances	Effective for non-linear eparations	Requires more omputational resources	99%
2021[63]	ANN	Dataset with 392 gesture ignals	High classification ccuracy, handles complex atterns	Requires large training ata, prone to over-itting.	97.01%
2020[68]	KNN	30 test images, 6 gestures	Easy to implement, omputationally efficient	Lower accuracy ompared to SVM	90%

6.2 Deep Learning

Model complexity is an important consideration in Developing and deploying deep learning models for tasks such as gesture recognition.

6.2.1. Convolutional Neural Networks (CNN)

DL models have attracted the attention of researchers to effective and generalized HGR with large-scale datasets because, due to their problems with ML algorithms, such algorithms are hampered. Miah et al. applied a preprocesses segmentation and augmentation model based on CNN BenSignNet, which in the BdSL38 data-set and ASL data-set showed impressive accuracy 93.00 percent and 99.00 percent respectively [69]. In the same light DL models presented outstanding levels of recognition accuracies over Chinese, Arabic and Japanese sign languages [70]. Wang ET AL. [71] address the architecture of the CRNN (Convolution Recurrent Neural Network) that involves both convolution layer to extract features, and recurrent layer to model accordance. Then, the Bi-LSTM model presents semantic dependencies both leftwards and rightwards, and it recorded 98.80 percent in single gestures. Tao ET AL. used a DL algorithm on RGB-headed HGR and obtained an outstanding accuracy of 99.9 percent on leave-one-out assessments on the 24 indicators. The YOLO5 is a model suggested by Jain ET Al. [73] recorded an admirable precision of 92.00 percent with Danish SLR. Moreover [74]. A CNN-based recognition system of hand gestures as well. Miah et al. [75] used multi-stage deep neural network to improve the effectiveness of features used and demonstrated good performance accuracy with a wide range of hand gestures.

6.2.2. Recurrent Neural Networks (RNNs)

Pigou et al. [76] used an end-to-end neural network works such as bidirectional recurrence, and temporal convolution are effectively included to enhance performance. In the more recent studies, Molchanov et al. suggested the R3DCNN method for gesture recognition [77]. This method uses a 3D CNN architecture for decomposing the short-term video sequences with an RNN for long-term video sequences. Is particularly suitable for the effective captures of spatial as well as temporal features.
Nature of RNNs. In its place, we observed the use of back-propagation through time which could lead to gradients growing very small so that the network is not easily formed. To learn long-term dependencies effectively. Consequently, as with many RNN that attempt to keep track of the history effect over long sequences, for this reason, encounter difficulties in achieving precise This is about recognizing dynamic gestures which are a kind of gestures that need long-term. Contextual information. To solve this problem Zhu ET AL. have developed an LSTM-based DL system [78]. Additionally, they used 2D CNN's to increase the effectiveness of achieving that goal. as belonging to the feature extraction portion of the feature selection procedure. Over 90% was cited by Rastgoo ET AL. [79] is recognition accuracy in hand SLR using LSTM on the RKSPERSIANSIGN datasets. As shown in Table 2 Algorithms in Deep Learning.

6.2.3. Long Short-Term Memory (LSTM)

De Smedt et al. [80] leverage the inherent structure of hand topology for skeleton-based HGR. Methods using RNN, especially LSTM units, combined with CNN's are notably more effective than traditional ML and standalone CNN models. Konstantinos ET AL. [81] demonstrate the enhanced precision of SLR by integrating HMM with CNN's and bidirectional RNN incorporating LSTM units. Lai et al. integrated a CNN with an RNN DL model for recognizing skeleton-based hand gestures, achieving 85.61% accuracy on the DHG 14 gesture datasets [82]. Han ET AL. proposed a two-stream method combining RGB and skeleton data using KLSTM-3D Res-net to improve recognition rates [83]. Han ET AL. also highlighted the enhanced spatial temporal feature.

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

Extraction by the fusion of Res-net and LSTM networks. Nunez et al. integrated CNN and LSTM features for extracting temporal features from 3D pose estimation, reporting 99.00% accuracy, underscoring the superiority of RNN, especially LSTM [84]. To improve the performance, Chen et al. extracted articulated finger movement and hand movement features, then concatenated the features and fed them into the RNN and reported 84.68% accuracy for 14 classes and 80.32% for 28 classes on the DHG datasets[85].

Table 2 shows the result of deep learning models or method

Year Ref.	**Method**	**Advantages**	**Disadvantages**	**Accuracy**	**Dataset**
2018[82]	CNN with RNN	Good accuracy on skeleton data	Requires preprocessing	85.61%	DHG14
2022[69]	CNN (BenSignNet)	High accuracy with pr-	Complexity of pr-processing	93.00%	BdSL38, ASL
2017[85]	RNN with concatenated finger and hand movement features	Multi-feature concatenation for better accuracy	Time-consuming training process	84.68% (14 classes), 80.32% (28 classes)	DHG
2020[79]	LSTM	Good performance on skeleton-based	Limited to specific datasets	90%	RKSPERS IANSIGN

7. Sensor Based HGR

Due to its ever-lowering cost and lower safety concerns compared to invasive sensing techniques, wearable hand gesture devices have been widely used in the decoding of human movement for intention recognition. Wearable technology that incorporates hand gesture recognition (HGR) sensors offers innovative ways to enhance the quality of life for individuals with disabilities. [86] These wearable sensors include surface electromyography (EMG) to command.

Human arm and hand joint kinematics are continuously estimated and movement classes are identified using statistical and predictive learning techniques such as maximal a posteriori and expectation maximisation (EM). [88] Deep learning (DL) methods, which use intricate hierarchical neural networks without taking a preexisting knowledge into consideration, are also frequently used for hand gesture classification in order to make greater use of the sensory data. [89] Recently, researchers have created a variety of wearable upper-limb devices, such as wrist/armbands and wearable data gloves. [90] Data gloves, which are fitted with sensors to identify hand movements such as finger bending and abduction, wrist flexion and abduction, etc., are another way to obtain signals [91].

A. Electromyography (EMG)

Electromyography (EMG) is an unavoidable means to record the electrical signals of muscle tissues. It is a method that involves the use of electrodes that are fixed to the body's surface or can be inserted into the muscles. The study of the author [92] used a combination of 6-axis accelerometer input and 10-channel EMG signals associated to the user's hand. The best accuracy of 93.1% was obtained by using the fuzzy k-means clustering for classification of 72 dynamic Chinese Sign Language (CSL) movements. The combination of accelerometer with electromyography (EMG) data presents a demonstration of the potential advantage in combining a range of sensor modalities in improving the accuracy of gesture recognition. In a similar vein, Author [93] investigated the use of EMG sensors placed on the user's arm to record finger motion. By combining Bayes and the KNN classifier linearly, this experiment successfully achieved a 94% accuracy in 20 different gestures classification. This study demonstrates the excellent potential of EMG sensor in recording complex muscle movement for gesture recognition. In literature 94 the extraction of EMG pattern signatures for different movements were investigated and Artificial Neural Networks (ANNs) were used to classify the signal according to certain features. Importantly, the results of this study illustrate the feasibility of EMG-based classification systems to differentiate various muscle activities, supporting the possibility of differentiating more subtle gesture patterns. The Myo armband, which includes the IMU and EMG sensors, has become a full instrument for GSR. In research conducted by [95] the authors used the Myo armband to identify 20 classes of Libras, that is, the American Sign Language. Using the Support Vector Machine classifier, the mean accuracy achieved a remarkable 98.6%, indicating that both IMU and EMG data can be usefully fused for achieving robust gesture recognition. This fusion of vision-based and EMG data illustrates the synergy offered by multi-modal sensor integration to enhance gesture recognition accuracy. Also, the fusion of SEMG with Caber glove was investigated in [96] to recognize the flexion of all Five fingers. ICA and PCA were employed for the reduction of computational burden. The research presented the possibility of utilizing surface EMG coupled with glove input for finger movement recognition, achieving a good success rate of 90% with LDA. SUMMARIES Taken together, these studies demonstrate the multiple facets of EMG usage for gesture recognition; its versatility to record muscle activity patterns associated to complex hand and finger movements. Integrating EMG with other sensor modalities including accelerometer or depth cameras can also increase the richness of data inputs, leading to more robust and accurate GR systems. As technology continues to evolve, the application of EMG in multi-modal sensor arrangements shows potential in promoting the development of gesture recognition and human-computer interaction [97].

Data preprocessing of sensor data involves the cleaning or filtering of data that is to be input into the recognition system. preprocessing is important in this case since it can enhance the outcome of the recognition system. Here are some of the commonly used techniques for sensor data preprocessing:

1) Noise Removal: This is because, in sensor data, there is the possibility of noise and hence the recognition system is less accurate. The techniques utilized for noise removal include mean filtering median filtering and wavelet filtering and from this [98-99].

2) Signal Amplification: There are several occasions that the spectrometer sensor yields a weak signal that is usually hard to capture minimal changes. Various techniques comprise gain adjustment, signal averaging, and resampling of signal can be employed for enhancing the sensor signal quality [100][101].

B. Electroencephalography (EEG)

EEG signal analysis involved utilizing nine different feature types: cD1, cD2, cD3, cD4, cD5, cA5, Energy, Entropy, and standard deviation. These features were used for processing EEG signals and detecting signals corresponding to hand movements involved in sign language from the subject's brain activity. In the research by Al-Anbary and Al-Qaraawi, [102], out of the multidimensional and temporally varying EEG signals, a 1D representation was obtained called EEG features. Bursts A and C primarily represented visual information, for the extraction of matching image descriptors that could subsequently be used for automated categorization. Furthermore, in [103], spike-related features were used by deploying a temporal contrast coding methodology where measured analogy signals are mapped to spiking analogy or spike streams. These spike streams which include both positive and negative spike streams were used as features for classification. AlQattan and Sepulveda [104] have proven good results in the classification of the EEG signals using SVM and LDA algorithms endowed with an accuracy rate of 75% based on the Entropy of feature kind. In the EEG-based BCI in SLR, the algorithms SVM and LDA gave 75% accuracy. Chaves [105] proposed DL models and reached 89.03 % in semantic image classification.

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)
ISSN -2708-3985

C. Data Glove

Someone noticed this was even better at counting than the old-fashioned mechanical pedometers and the IMU sensors (gyroscopes and accelerometer) used by gesture and sign language recognition systems employing so-called data gloves was born. These sensors are used for gathering the orientation, angular movement, and acceleration values information [106]. Using some data glove, the fingers flexion data can be measured through flex sensors. 6 The V PL-Data glove (a glove with flex/fibre transducers). This is used to measure flex angles and orientation data with high accuracy. In [107], 16 streams of raw data were obtained from V PL-Data glove, and the movements for both of the hands were classified into ten basic gestures. The FMNN was trained using the movement patterns for recognition of 25-words in KSL with an accuracy of 85\%. More simply put, we successfully used an advanced neural network, the FMNN, to read complex hand signals obtained with a V PL-Data glove. This approach proved to be highly accurate in recognizing a wide variety of gestures of the Korean Sign Language. Besides the use of data gloves, second research [108] focused on recognition and interpretation of 250 words in Taiwanese Sign Language. These include finger flexion features derived from the Data Glove. . These features were used as inputs to Hidden Markov Models (HMMs) and used to recognise 51 poses, 6 orientations and 8 movement types. It is worth mentioning that a 100% ratio was achieved in terms of accuracy of the classification; the 100% of the targeted items fell on their corresponding category. The authors also conducted experiments with single-hand movements, short sentences, and long sentences of 250 words, with the results of 89.5%, 70.4%, and 81.6% accuracy, respectively. In another [108], a work of researchers [109] employed electromyography to build possible myoelectric prosthetic hands FIGURE 6. V PL-Data Glove. For disabled people. The system achieved a high accuracy of 91.64% in detecting ten different dynamic hand gestures by means of HMM. Together, these results stress the capabilities of data gloves for recording fine-scale hand movements and pointing gestures, indicating their suitability for implementing a reliable SLR. Apart from the hardware aspects, the selection of recognition algorithms also has a strong impact on measure performance of gesture recognition systems. The use of Fuzzy Min-max Neural Networks (FMNN) and Hidden Markov Models (HMM) in the above research works demonstrates the superior capability of these techniques in processing and recognizing the complex data stream derived from data gloves. Moreover, the work has demonstrated some measure of the system's generalization capability across new sign languages, leading to its potential for greater wider-spread use and inclusivity [110]. Table 3 Algorithms Implemented with Wearable Sensor HGR

Table 3 Algorithms Used in Wearable Sensor HGR

Ref.	Year	Dataset	Method	Wearable Sensor	Accuracy
[93]	2023	20 gestures	Bayes and KNN	EMG sensors	94%
[92]	2022	72 dynamic Chinese Sign Language (CSL) movements	Fuzzy K-means clustering	6-axis accelerometer, 10-channel EMG	93.1%
[95]	2022	20 American Sign Language (ASL) classes	Support Vector Machine (SVM)	Myo armband (IMU + EMG)	98.6%
[96]	2022	Finger flexion classification	Linear Discriminant Analysis (LDA)	sEMG and Cyber Glove	90%
[97]	2021	25 Korean Sign Language (KSL) gestures	Fuzzy Min-max Neural Network (FMNN)	VPL-Data Glove	85%
[102]	2021	Brain signals for sign language movements	SVM and LDA	EEG	75%

Compared with the traditional ML algorithms, DL has the following comparative advantages for the wearable sensor-based gesture recognition task, such as

1) Improving accuracy and robustness in detecting gestures;
2) Learning deep features in an unsupervised manner as opposed to manually selecting and extracting [111].
3) Management of confidence with respect to the estimation of the cross-modality features for hybrid modalities tasks [112].

It can learn from raw data without handcrafted features It removes some of the data preprocessing that usually comes with ML. [113] Yet, note that DL methods for HGR recognition do not necessarily offer \better" recognition performance in comparison to traditional methods. There are several aspects that need to be taken into account to decide whether DL should be applied, such as the number of datasets, the complexity of task, and real-time computational cost [114].

Deep learning (DL) methods (like convolution neural networks (CNN) and artificial neural networks (ANN)) can manage input and unstructured data and perform automatic extraction of features to weaken obvious dependence of the trainer. As ANN is adaptive, so the connectionist approaches have the ability to include learning in a data-rich environment. Different connection models are used in HGR research, such as multi-layer perception (MLP), time delay neural network (TDNN) and radial basis function neural network (RBFN). [115] CNN, however, needs to have far more data inputs to obtain its high accuracy rate.

8. Conclusion

HGR has emerged as a vital enabler of intuitive human-computer interaction, with transformative potential in domains such as healthcare rehabilitation, virtual reality, and assistive technology. This review provides a systematic comparison of vision-based and sensor-based HGR methodologies, evaluating their performance across the full processing pipeline—from data acquisition and preprocessing to feature extraction and classification. While vision-based systems offer high accessibility and user convenience, they remain susceptible to environmental variations and occlusion. Conversely, wearable sensor-based approaches deliver precise motion tracking but are often constrained by ergonomic and usability concerns. A critical observation across the reviewed literature is the prevalent reliance on limited and homogeneous datasets, such as the UCI MYO Thalami dataset, which includes only 36 subjects and may not adequately represent population diversity or real-world conditions. To advance the field, future research must prioritize: the creation of larger, more diverse datasets; the development of efficient, lightweight architectures for real-time deployment; the adoption of standardized evaluation frameworks; the integration of ethical design and accessibility considerations; and enhanced interdisciplinary collaboration among computer scientists, linguists, and the disability community. Addressing these challenges will be essential for building next-generation HGR systems that are robust, inclusive, and practically viable.

References

[1] M. S. Amin, S. T. H. Rizvi, and M. M. Hossain, ‘‘A comparative review on applications of different sensors for sign language recognition,’’ J. Imag., vol. 8, no. 4, p. 98, Apr. 2022, doi: 10.3390/jimaging8040098.

[2] A. Sultan, W. Makram, M. Kayed, and A. A. Ali, "Sign language identification and recognition: A comparative study," Open Comput. Sci., vol. 12, no. 1, pp. 191–210, May 2022, doi: 10.1515/comp-2022-0240.

[3] Z. R. Saeed, Z. B. Zainol, B. B. Zaidan, and A. H. Alamoodi, ‘‘A systematic review on systems-based sensory gloves for sign language pattern recognition: An update from 2017 to 2022,’’ IEEE Access, vol. 10, pp. 123358–123377, 2022, doi: 10.1109/ACCESS.2022.3219430.

[4] T. Song, H. Zhao, Z. Liu, H. Liu, Y. Hu, and D. Sun, ‘‘Intelligent human hand gesture recognition by local–global fusing quality-aware features,’’ Future Gener. Comput. Syst., vol. 115, pp. 298–303, Feb. 2021, doi: 10.1016/j.future.2020.09.013.

[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] 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.

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

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] J. Bora, S. Dehingia, A. Boruah, A. A. Chetia, and D. Gogoi, "Realtime 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.

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

[17] 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.

[18] Q. Han, Z. Huangfu, W. Min, T. Ding, and Y. Liao, "Sign language recognition 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.

[19] 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.

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

[21] 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.

[22] S. Jain, ‘‘ADDSL: Hand gesture detection and sign language recognition on[annotated Danish sign language,’’ 2023, arXiv:2305.09736.

[23] S. Güney and M. Erkuş, ‘‘A real-time approach to recognition of Turkish sign language by using convolutional neural networks,’’ Neural Comput. Appl., vol. 34, no. 5, pp. 4069–4079, Mar. 2022.

[24] J. Zhang, Q. Wang, Q. Wang, and Z. Zheng, ‘‘Multimodal fusion framework based on statistical attention and contrastive attention for sign language recognition,’’ IEEE Trans. Mobile Comput., vol. 23, no. 2, pp. 1431–1443, Feb. 2024.

[25] M. S. Amin, S. T. H. Rizvi, A. Mazzei, and L. Anselma, ‘‘Assistive data glove for isolated static postures recognition in American sign language using neural network,’’ Electronics, vol. 12, no. 8, p. 1904, Apr. 2023.

[26] X. Yang and Y. Tian, ‘‘Effective 3D action recognition using EigenJoints,’’ J. Vis. Commun. Image Represent., vol. 25, no. 1, pp. 2–11, Jan. 2014.

[27] Q. De Smedt, H. Wannous, and J.-P. Vandeborre, ‘‘Skeleton-based dynamic hand gesture recognition,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Jun. 2016, pp. 1206–1214.

[28] D. Q. Smedt, H.Wannous, and J. P. Vandeborre, ‘‘Shrec’17 track: 3D hand gesture recognition using a depth and skeletal dataset,’’ in Proc. 3DOR- 10th Eurographics Workshop 3D Object Retrieval, 2017, pp. 1–6.

[29] 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.

[30] A. S. Al-Shamayleh, R. Ahmad, M. A. M. Abushariah, K. A. Alam, and N. Jomhari, "A systematic literature review on vision-based gesture recognition techniques," Multimedia Tools Appl., vol. 77, no. 21, pp. 28121–28184, Nov. 2018.

[31] T. Kapuscinski and K. Inglot, ‘‘Vision-based gesture modeling for signed expressions recognition,’’ Proc. Comput. Sci., vol. 207, pp. 1007–1016, Jan. 2022.

[32] 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.

[33] 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.

[34] S. Subburaj and S. Murugavalli, "Survey on sign language recognition in the context of vision-based and deep learning," Meas., Sensors, vol. 23, Oct. 2022, Art. no. 100385.

[35] X. Li, X. Zhang, X. Chen, X. Chen, and A. Liu, ‘‘Cross-user gesture recognition from sEMG signals using an optimal transport assisted student-teacher framework,’’ Comput. Biol. Med., vol. 165, Oct. 2023, Art. no. 107327.

[36] 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 handworn motion sensors,’’ Inf. Sci., vol. 606, pp. 722–741, Aug. 2022.

[37] M. Marais, D. Brown, J. Connan, and A. Boby, "An evaluation of hand-based algorithms for sign language recognition," in Proc. Int. Conf. Artif. Intell., Big Data, Comput. Data Commun. Syst. (icABCD), Aug. 2022, pp. 1–6.

[38] N. Nayan, D. Ghosh, and P. M. Pradhan, ‘‘An unsupervised learning approach to handle movement epenthesis in continuous sign language recognition,’’ in Proc. 17th Int. Conf. Control, Autom., Robot. Vis. (ICARCV), Dec. 2022, pp. 862–867.

[39] Q. M. Areeb and M. Nadeem, ‘‘Deep learning based hand gesture recognition for emergency situation: A study on Indian sign language,’’ in Proc. Int. Conf. Data Anal. Bus. Ind. (ICDABI), Oct. 2021, pp. 33–36.

[40] B. Shi, X. Chen, Z. He, and R. Han, "Development of magnetic-sensor based hand gesture recognition system for sign language," in Proc. IEEE 6th Int. Electr. Energy Conf. (CIEEC), May 2023, pp. 2302–2305, doi: 10.1109/CIEEC58067.2023.10167092.

[41] R. Kushwaha, G. Kaur, and M. Kumar, "Hand gesture based sign language recognition using deep learning," in Proc. 3rd Int. Conf. Secure Cyber Comput. Commun. (ICC), May 2023, pp. 293–297, doi: 10.1109/ICSCCC58608.2023.10176912.

[42] J. S. Lee, M. Seo, S.W. Kim, and M. Choi, ‘‘Fetal QRS detection based on convolutional neural networks in noninvasive fetal electrocardiogram,’’ in Proc. 4th Int. Conf. Frontiers Signal Process. (ICFSP), Poitiers, France, Sep. 2018, pp. 75–78.

[43] 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.

[44] T. Raghuveera, R. Deepthi, R. Mangalashri, and R. Akshaya, "A depth-based Indian sign language recognition using Microsoft Kinect," S¯adhan¯a, vol. 45, Jan. 2020, Art. no. 34.

[45] Y. Yang, J. Li, B. Li, and Y. Zhang, ‘‘MDHandNet: A lightweight deep neural network for hand gesture/sign language recognition based on micro-Doppler images,’’ World Wide Web, vol. 25, no. 5, pp. 1951–1969, Sep. 2022.

[46] G. Kumar and P. K. Bhatia, "A detailed review of feature extraction image processing systems," in *Proc. 4th Int. Conf. Adv. Comput.Commun. Technol.*, Feb. 2014, pp. 5–12.

[47] 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.

[48] K. Padmanand and P.-C. Lim, ‘‘Malaysian sign language recognition using 3D hand pose estimation,’’ in Proc. Int. Conf. Digit. Transformation Intell. (ICDI), Dec. 2022, pp. 214–218.

[49] L. Ivanska and T. Korotyeyeva, ‘‘Mobile real-time gesture detection application for sign language learning,’’ in Proc. IEEE 17th Int. Conf. Comput. Sci. Inf. Technol. (CSIT), Nov. 2022, pp. 511–514.

[50] A. S. Ghotkar, R. Khatal, S. Khupase, S. Asati, and M. Hadap, "Hand gesture recognition for Indian sign language," in Proc. Int. Conf. Comput. Commun. Inform., Jan. 2012, pp. 10–20.

[51] E. Ronando and A. R. Hakim, "Hand gesture recognition system for recognizing of Indonesian sign language number using the attentional convolutional network," in Proc. Int. Conf. Energy, Power, Environ., Control, Comput. (ICEPECC), Mar. 2023, pp. 1–6.

[52] IEEE Standard Criteria for the Protection of Class 1E Power Systems and Equipment in Nuclear Power Generating Stations, Standard 741-1986, 1986.

[53] G. Kumar and P. K. Bhatia, ‘‘A detailed review of feature extraction in image processing systems,’’ in Proc. 4th Int. Conf. Adv. Comput. Commun. Technol., Feb. 2014, pp. 5–12, doi: 10.1109/ACCT.2014.74.

[54] 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.

[55] H. Hameed, M. Usman, M. Z. Khan, A. Hussain, H. Abbas, M. A. Imran, and Q. H. Abbasi, ‘‘Privacy-preserving British sign language recognition using deep learning,’’ in Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Jul. 2022, pp. 4316–4319, doi: 10.1109/EMBC48229.2022.9871491.

[56] T. Madhumitha, G. S. Geethika, and V. Radhesyam, ‘‘Real-time recognition of Indian sign language using OpenCV and deep learning,’’ in Proc. 4th Int. Conf. Emerg. Technol. (INCET), May 2023, pp. 1–4, doi: 10.1109/INCET57972.2023.10170080.

[57] Y. Shi, Y. Li, X. Fu, M. Kaibin, and M. Qiguang, ‘‘Review of dynamic gesture recognition,’’ Virtual Reality Intell. Hardw., vol. 3, no. 3, pp. 183–206, Jun. 2021.

[58] M. F. Wahid, R. Tafreshi, M. Al-Sowaidi, and R. Langari, "Subject independent hand gesture recognition using normalization and machine learning algorithms," J. Comput. Sci., vol. 27, pp. 69–76, Jul. 2018.

[59] I. A. Adeyanju, O. O. Bello, and M. A. Adegboye, ‘‘Machine learning methods for sign language recognition: A critical review and analysis,’’ Intell. Syst. With Appl., vol. 12, p. 56, Nov. 2021.

[60] R. Sreemathy, M. Turuk, S. Chaudhary, K. Lavate, A. Ushire, and S. Khurana, ‘‘Continuous word level sign language recognition using an expert system based on machine learning,’’ Int. J. Cognit. Comput. Eng., vol. 4, pp. 170–178, Jun. 2023.

[61] J. Nayak, B. Naik, P. B. Dash, A. Souri, and V. Shanmuganathan, ‘‘Hyperparameter tuned light gradient boosting machine using memetic firefly algorithm for hand gesture recognition,’’ Appl. Soft Comput., vol. 107, Aug. 2021, Art. no. 107478.

[62] V. Adithya and R. Rajesh, ‘‘Hand gestures for emergency situations: A video dataset based on words from Indian sign language,’’ Data Brief, vol. 31, Aug. 2020, Art. no. 106016.

[63] F. Wang, R. Hu, and Y. Jin, ‘‘Research on gesture image recognition method based on transfer learning,’’ Proc. Comput. Sci., vol. 187, pp. 140–145, Jan. 2021.

[64] 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.

[65] V. R. Ivanova, A. S. Ivanova, and L. V. Fetisov, ‘‘The development of an automated station for group soldering of the led lines,’’ in Proc. 14th Int.Sci.-Tech. Conf. Actual Problems Electron. Instrum. Eng. (APEIE), Oct. 2018, pp. 336–338.

[66] S. Bharti, A. Balmik, and A. Nandy, ‘‘Novel error correction-based key frame extraction technique for dynamic hand gesture recognition,’’ Neural Comput. Appl., vol. 35, no. 28, pp. 21165–21180, Oct. 2023.

[67] Y. Jiang, M. Zhao, C. Wang, F. Wei, K. Wang, and H. Qi, "Diver's hand gesture recognition and segmentation for human-robot interaction on AUV," Signal, Image Video Process., vol. 15, no. 8, pp. 1899–1906, Nov. 2021.

[68] C. Dignan, E. Perez, I. Ahmad, M. Huber, and A. Clark, ‘‘An AI-based approach for improved sign language recognition using multiple videos,’’ Multimedia Tools Appl., vol. 81, no. 24, pp. 34525–34546, Oct. 2022.

[69] A. S. M. Miah, J. Shin, M. A. M. Hasan, and M. A. Rahim, ‘‘BenSignNet: Bengali sign language alphabet recognition using concatenated segmentation and convolutional neural network,’’ Appl. Sci., vol. 12, no. 8, p. 3933, Apr. 2022.

[70] S.-i. Ito, M. Ito, and M. Fukumi, ‘‘Japanese sign language classification using gathered images and convolutional neural networks,’’ in Proc. IEEE 2nd GlobalConf. Life Sci. Technol. (LifeTech), Mar. 2020, pp. 349–350.

[71] Y.Wang, Z. Hao, X. Dang, Z. Zhang, and M. Li, ‘‘UltrasonicGS: A highly robust gesture and sign language recognition method based on ultrasonic signals,’’ Sensors, vol. 23, no. 4, p. 1790, Feb. 2023.

[72] W. Tao, M. C. Leu, and Z. Yin, ‘‘American sign language alphabet recognition using convolutional neural networks with multiview augmentation and inference fusion,’’ Eng. Appl. Artif. Intell., vol. 76, pp. 202–213, Nov. 2018.

[73] S. Jain, ‘‘ADDSL: Hand gesture detection and sign language recognition on annotated Danish sign language,’’ 2023, arXiv:2305.09736.

[74] S. Güney and M. Erkuş, ‘‘A real-time approach to recognition of Turkish sign language by using convolutional neural networks,’’ Neural Comput. Appl., vol. 34, no. 5, pp. 4069–4079, Mar. 2022.

[75] A. S. M. Miah, M. A. M. Hasan, J. Shin, Y. Okuyama, and Y. Tomioka, ‘‘Multistage spatial attention-based neural network for hand gesture recognition,’’ Computers, vol. 12, no. 1, p. 13, Jan. 2023

[76] L. Pigou, A. van den Oord, S. Dieleman, M. Van Herreweghe, and J. Dambre, ‘‘Beyond temporal pooling: Recurrence and temporal convolutions for gesture recognition in video,’’ Int. J. Comput. Vis., vol. 126, nos. 2–4, pp. 430–439, Apr. 2018.

[77] P. Molchanov, X. Yang, S. Gupta, K. Kim, S. Tyree, and J. Kautz, ‘‘Online detection and classification of dynamic hand gestures with recurrent 3D convolutional neural networks,’’ in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 4207–4215.

[78] Y. Zhu, Z. Lan, S. Newsam, and A. Hauptmann, ‘‘Hidden two-stream convolutional networks for action recognition,’’ in Proc. Asian Conf. Comput. Vis. Cham, Switzerland: Springer, 2018, pp. 363–378.

[79] R. Rastgoo, K. Kiani, and S. Escalera, ‘‘Hand sign language recognition using multi-view hand skeleton,’’ Expert Syst. Appl., vol. 150, Jul. 2020, Art. no. 113336.

[80] Q. De Smedt, H. Wannous, and J.-P. Vandeborre, ‘‘Heterogeneous hand gesture recognition using 3D dynamic skeletal data,’’ Comput. Vis. Image Understand., vol. 181, pp. 60–72, Apr. 2019.

[81] D. Konstantinidis, K. Dimitropoulos, and P. Daras, ‘‘Sign language recognition based on hand and body skeletal data,’’ in Proc. 3DTV-Conf., True Vis. Capture, pp. 1–4, Thessaloniki, Greece, July 2018

[82] K. Lai and S. N. Yanushkevich, ‘‘CNN+RNN depth and skeleton based dynamic hand gesture recognition,’’ in Proc. 24th Int. Conf. Pattern Recognit. (ICPR), Aug. 2018, pp. 3451–3456.

[83] A. 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.

[84] J. C. Núñez, R. Cabido, J. J. Pantrigo, A. S. Montemayor, and J. F. Vélez, ‘‘Convolutional neural networks and long short-term memory for skeleton-based human activity and hand gesture recognition,’’ Pattern Recognit., vol. 76, pp. 80–94, Apr. 2018.

[85] X. Chen, H. Guo, G. Wang, and L. Zhang, ‘‘Motion feature augmented recurrent neural network for skeleton-based dynamic hand gesture recognition,’’ in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2017, pp. 2881–2885.

[86] P. Maceira-Elvira, T. Popa, A.-C. Schmid, and F. C. Hummel, "Wearable technology in stroke rehabilitation: Towards improved diagnosis and treatment of upper-limb motor impairment in stroke patients," *Journal of Neuro Engineering and Rehabilitation*, vol. 16, no. 1, p. 1, 2019

[87] J. Pan, Y. Li, Y. Luo, X. Zhang, X. Wang, D. L. T. Wong, C.-H. Heng, C.-K. Tham, and A. V.-Y. Thean, "Flexible and stretchable microelectrodes integrated with wearable sensors for monitoring human electrophysiology during daily activities," ACS Sensors, vol. 6, no. 11, pp. 4156–4165, 2021.

[88] Y. Pei and X. Guo, "Artificial intelligence in wearable devices for healthcare: A review," vol. 4, no. 1, p. 2100046, 2022.

[89] A. Moin, A. Zhou, A. Rahimi, A. Menon, S. Benatti, G. Alexandrov, S. Tamakloe, J. Ting, N. Yamamoto, Y. Khan, F. Burghardt, L. Benini, A. C. Arias, J. M. Rabaey, Nat. Electron. 2020, 4, 54.

[90] M. Dong, B. Fang, J. Li, F. Sun, and H. Liu, "Wearable sensors for human activity monitoring: A comprehensive review," *Journal of Engineering in Medicine*, vol. 235, no. 1, pp. 117–130, 2020.

[91] O. Ozioko and R. Dahiya, "Recent advancements in wearable intelligent systems: Sensors, materials, and applications," vol. 3, no. 4, p. 2100091, 2021.

[92] A. Boukdir, M. Benaddy, A. Ellahyani, O. E. Meslouhi, and M. Kardouchi, "3D gesture segmentation for word-level Arabic sign language using large-scale RGB video sequences and autoencoder convolutional networks,", vol. 16, no. 8, pp. 2055–2062, Nov. 2022.

[93] S. K. Singh and A. Chaturvedi, "A reliable and efficient machine learning pipeline for American sign language gesture recognition using EMG sensors," *Multimedia Tools Appl.*, vol. 82, no. 15, pp. 23833–23871, Jun. 2023.

[94] J. P. Vásconez, L. I. Barona López, Á. L. V. Caraguay, and M. E. Benalcázar, "A comparison of EMG-based hand gesture recognition systems based on supervised and reinforcement learning,", vol. 123, Aug. 2023, Art. no. 106327.

[95] A. Kasapbaşi, A. E. A. Elbushra, O. Al-Hardanee, and A. Yilmaz, "DeepASLR: A CNN-based human-computer interface for American sign language recognition for hearing-impaired individuals,", vol. 2, Jan. 2022, Art. no. 100048.

[96] D. Avola, L. Cinque, A. Fagioli, G. L. Foresti, A. Fragomeni, and D. Pannone, "3D hand pose and shape estimation from RGB images for keypoint-based hand gesture recognition," *Pattern Recognit.*, vol. 129, Sep. 2022, Art. no. 108762.

[97] A. H. Alrubayi, M. A. Ahmed, A. A. Zaidan, A. S. Albahri, B. B. Zaidan, O. S. Albahri, A. H. Alamoodi, and M. Alazab, "A pattern recognition model for static gestures in Malaysian sign language based on machine learning techniques, vol. 95, Oct. 2021, Art. no. 107383.

[98] G. Latif, N. Mohammad, J. Alghazo, R. AlKhalaf, R. AlKhalaf, ArASL: Arabic Alphabets Sign Language Dataset, Data Brief. 23,2019,103777.

[99] R. Saif, M. Ahmad, S.Z.H. Naqvi, S. Aziz, M.U. Khan, M. Faraz, Multi-Channel EMG Signal analysis for Italian Sign Language Interpretation, in: 2022 International Conference on Emerging Trends in Smart Technologies (ICETST), 2022: pp. 1–5.

[100] G.L. Sukruth, V.K. Bp, M.R. Tejas, K. Rithvik, Enhancing Collaborative Interaction with the Augmentation of Sign Language for the Vocally Challenged, Journal of Advanced …. 1 (2023) 2010.

[101] F. Hamid Yousuf, A. Bushnaf Alwarfalli, I. Ighneiwa, Arabic Sign Language Recognition System by Using Surface Intelligent EMG Signal, in: The 7th International Conference on Engineering & MIS 2021, Association for Computing Machinery, New York, NY, USA, 2021: pp. 1–6.

[102] A. H. Al-Anbary and S. M. Al-Qaraawi, ‘‘A proposed sign language model for speechless persons using EEG signals,’’ Iraqi J. Inf. Commun. Technol., vol. 4, no. 3, pp. 23–29, Dec. 2021.

[103] Y. Ai and B. Rajendran, ‘‘A convolutional spiking network for gesture recognition in brain-computer interfaces,’’ 2023, arXiv:2304.11106.

[104] D. AlQattan and F. Sepulveda, "Towards sign language recognition using EEG-based motor imagery brain-computer interface," in Proc. 5th Int. Winter Conf. Brain-Comput. Interface (BCI), Jan. 2017, pp. 5–8.

[105] A. B. Chaves, ‘‘Personalized image classification from EEG signals using deep learning,’’ B.S. thesis, Escola Tecnica d'Enginyeria de Telecomunicacio de Barcelona, Universitat Politecnica de Catalunya, Barcelona, Spain, 2017.

[106] A. R. Patil and S. Subbaraman, "A spatiotemporal approach for vision-based hand gesture recognition using Hough transform and neural network," Signal, Image Video Process., vol. 13, no. 2, pp. 413–421, Mar. 2019, doi: 10.1007/s11760-018-1370-1.

[107] R. Ferhat and F. Z. Chelali, ‘‘Textural feature descriptors for a static and dynamic hand gesture recognition system,’’ Multimedia Tools Appl., vol. 83, no. 3, pp. 8165–8187, Jan. 2024, doi: 10.1007/s11042-023- 15410-0.

[108] S. Das, S. K. Biswas, and B. Purkayastha, ‘‘Automated Indian sign language recognition system by fusing deep and handcrafted feature,’’ Multimedia Tools Appl., vol. 82, no. 11, pp. 16905–16927, May 2023, doi: 10.1007/s11042-022-14084-4.

[109] Z. Zhang, Q. Shen, and Y. Wang, "Electromyographic hand gesture recognition using a convolutional neural network with multi-attention," Biomed Signal Process Control, vol. 91, May 2024, Art. no. 105935, doi: 10.1016/j.bspc.2023.105935.

[110] A. A. Zare and S. H. Zahiri, "Recognition of a real-time signer independent static Farsi sign language based on Fourier coefficients amplitude," Int. J. Mach. Learn. Cybern., vol. 9, no. 5, pp. 727–741, May 2018, doi: 10.1007/s13042-016-0602-3.

[111] Y. Yu, X. Chen, S. Cao, X. Zhang, and X. Chen, "Wearable hand gesture recognition system based on electromyography and accelerometer signals," vol. 24, no. 6, pp. 1310–1319, 2020. doi: 10.1109/JBHI.2020.2992437.

[112] C. Du, L. Zhang, X. Sun, J. Wang, and J. Sheng, "Gesture recognition using a wearable system based on surface electromyography and accelerometer signals,", vol. 8, pp. 144610–144619, 2020. doi: 10.1109/ACCESS.2020.3015729.

[113] S. Bhushan, M. Alshehri, I. Keshta, A. K. Chakraverti, J. Rajpurohit, and A. Abugabah, "A survey of deep learning methods for gesture recognition in sign language systems," vol. 11, no. 6, p. 968, 2022. doi: 10.3390/electronics11060968.

[114] J. Kim, S. Mastnik, and E. André, "EMG-based hand gesture recognition for real-time biosignal interfacing,", pp. 30–39, 2008. doi: 10.1145/1378773.1378779.

[115] Y. Wu, D. Jiang, J. Duan, X. Liu, R. Bayford, and A. Demosthenous, "A wearable and reconfigurable sensor system for hand gesture recognition,", IEEE, Piscataway, NJ, pp. 1–5, 2018. doi: 10.1109/ISCAS.2018.8351185.