International Journal of Computer Applications (0975 - 8887)
Volume 187 - No.2, May 2025

Hand Gesture Recognition Systems: A Review of Methods, Datasets, and Emerging Trends

Md Selim Sarowar
CRC Research Centre
Dhaka, Bangladesh

Nur E Jannatul Farjana
Department of Computer Science and Engineering
International University of Business Agriculture and Technology,
Dhaka, Bangladesh

Md. Asraful Islam Khan
Department of Computer Science and Engineering
International University of Business Agriculture and Technology,
Dhaka, Bangladesh

Md Abdul Mutalib
Department of Computer Science and Engineering
North East University Bangladesh

Syful Islam
Department of Computer Science and Engineering
Barisal Information Technology College

Mohaiminul Islam
Department of Holistic Systems
Simula Research Laboratory

ABSTRACT

Hand gestures are a powerful method of communication that serve as a bridge between humans and computers, enabling intuitive interaction. Hand Gesture Recognition (HGR) systems aim to support this vision but face several challenges such as gesture irregularity, illumination variation, background interference, and computational complexity. This study evaluates 252 peer-reviewed articles published between 1995 and 2024, with a focus on input modalities, algorithmic approaches, benchmark datasets, application domains, and system-level challenges such as automation, scalability, generalization, and real-time performance.The evolution of HGR methods is categorized chronologically, beginning with early rule-based models, progressing through classical machine learning techniques such as SVM, KNN, and HMM, and advancing to deep learning frameworks including CNNs, RNNs, LSTMs, 3D CNNs, and Graph Convolutional Networks (GCNs). In recent years, hybrid and pretrained architectures including LSTM+3DCNN, MAE+ST-GCN, and Transformer-based models have been proposed to address existing limitations and improve performance. Various input modalities have been explored, including RGB image and video data, depth sensors, skeletal tracking, IMU, and EMG signals. Widely adopted benchmark datasets include SHREC, DHG-14/28, and NVGesture. A temporal classification framework is introduced to segment the progression of HGR technologies across decades. The study highlights key trends, technological advancements, and unresolved challenges, offering insights that may guide the development of accurate, efficient, and user-centric HGR systems, particularly in mobile and embedded computing contexts.

General Terms

Human-Computer Interaction, Machine Learning

Keywords

Hand Gesture Recognition, Deep Learning, LSTM, Multimodal Fusion, Lightweight Architectures

[†]These authors contributed equally to this work.

1. INTRODUCTION

An interface is required to communicate between humans and machines and it should also be user-friendly. Speech and hand gestures are the most common interface. People use several controlling devices such as joysticks and remote controls to operate the machines. Many application systems use these devices like software interface

control, gaming, virtual environments, automatic television control, smart home interactive control, and sign language interpretation [1]. Humans generally communicate using hand gestures, and the community of people with hearing impairments uses sign language, a natural form of hand gesture communication [2]. The use of hand gestures in communication is successful for humans, and efforts are currently being made to replicate this effectiveness in computer vision systems [3].
However, Hand Gesture Recognition (HGR) systems face challenges in achieving high accuracy because of background interference. To address this problem, researchers studied to find out various methods. Still, HGR systems face complexities in implementation in real-life applications. Therefore, a comprehensive analysis of the HGR system is essential. In this paper, we reviewed more than 250 articles, tracing the evolution of the HGR system across different eras and highlighting their findings and challenges. In Figure 1, we have represented the total papers year-wise used in this paper.

1.1 Context and Background

The interaction between humans and computers is one of the most significant evolutions indicating that they adapt to each other in different situations [4]. Human gesture recognition (HGR) is used in human-robot interaction to build intuitive user interfaces [1, 5]. Many researchers tried implementing various models to improve human-computer interaction using hand gestures [1]. The main steps in vision-based recognition include data acquisition, hand region segmentation, feature extraction, and Gesture classification[1].
In Figure 2, the general flowchart of a HGR system is shown. Several sensors such as data gloves, leap motion, vision sensors, depth sensors, etc. are available to develop a hand gesture recognition (HGR) system [6]. These are used to acquire the dataset of multiple modalities. Then the dataset is sent for appropriate pre-processing and feature extraction. After selecting the best feature, we need a classifier to classify the hand gesture correctly [7].

1.2 Research Motivation

Gesture recognition holds great possibilities for improving HCI, with applications ranging from sign language recognition to virtual reality [8]. Despite progress, challenges remain, particularly in making systems work reliably in real-world conditions beyond controlled environments [9]. Hand gesture recognition (HGR) is an essential human-machine interface for several applications, such as assistive technologies, smart home systems, virtual worlds, and gaming [10, 11]. High accuracy and reliable performance are still difficult to achieve, though, because of problems including background interference and a variety of hand behaviors. This study reviews more than 200 publications and identifies important issues to close the gap between theoretical developments and real-world usability in HGR systems. To improve human-computer interaction, the objective is to create HGR systems that are more precise, flexible, and easy to use.

1.3 Research Contribution

This review paper provides a comprehensive analysis of the evolution and advancements in Hand Gesture Recognition (HGR) systems by reviewing over 200 research papers. The key contributions of this work are as follows:
Systematic Review: We present a detailed year-wise review of HGR-related studies, highlighting significant developments, methodologies, and trends.
Framework Analysis: The framework for HGR systems of different eras is discussed.
Technology Overview: Various sensors and modalities used for HGR systems, including data gloves, vision sensors, and depth sensors, are analyzed to provide insights into their strengths and limitations.
Challenges and Gaps: We identify persistent challenges such as background interference, variability in gestures, and real-world implementation complexities, providing a foundation for future research.

1.4 Research Question

Through this research, we have tried to identify this question below
RQ#1: How have HGR systems developed over time concerning its datasets, input modalities, and techniques and how has this evolution taken place?
To answer this question (RQ#1), we have reviewed the number of papers carefully and sorted their findings in tabular form in this paper.

1.5 Structure of the Paper

The structure of the paper is as shown in figure 3. Section II investigates the methodological review of early-stage methods, in section III, we have discussed the traditional ML-related papers. In section IV, we analyzed the deep learning era of HGR systems. In section V, we presented the current methods and trends in the HGR system. We proceeded in section VI to discuss the applications of the HGR system. In section VII, we identified the challenges and their potential solutions. Finally, we concluded our paper in section VIII.

2. EARLY STATE OF HGR (1995-2010)

The development of the hand gesture recognition (HGR) system initially were relied on rule-based and handcrafted features [12]. In most of the scenarios before 2010, rule-based systems and handcrafted feature extraction methods were commonly used. Additionally, most of them were heavily dependent on predefined algorithms and manual selection of features for the interpretation of gestures. For Example, geometric features like finger lengths, angles, and contours are extracted manually to distinguish features [13]. For the segmentation of the HGR, computer vision techniques such as edge detection, and background subtraction played a pivotal role in isolating the hand region from the background [14]. Moreover, data acquisition in this era was often done by using data globes with sensors that captured the movement and position of the hand gesture with poor accuracy [15] [16].
However, the dependence on such hardware made the systems more bulky and less user-friendly. One of the main weaknesses of these earlier works is that neither are they flexible nor powerful. Since, the features were handpicked, the generalization performance of the systems was not very good across different users, lighting conditions, and backgrounds [17]. Moreover, the computational costs associated with processing these handcrafted features in real-time were too high, and hence the systems could hardly be applied in realistic scenarios [18].
Despite these challenges, the foundational work in this era set the stage for future advancements. Limitations of the rule-based and handcrafted feature systems provided insights that set the stage for more adaptive and learning-based approaches. The ground was be-

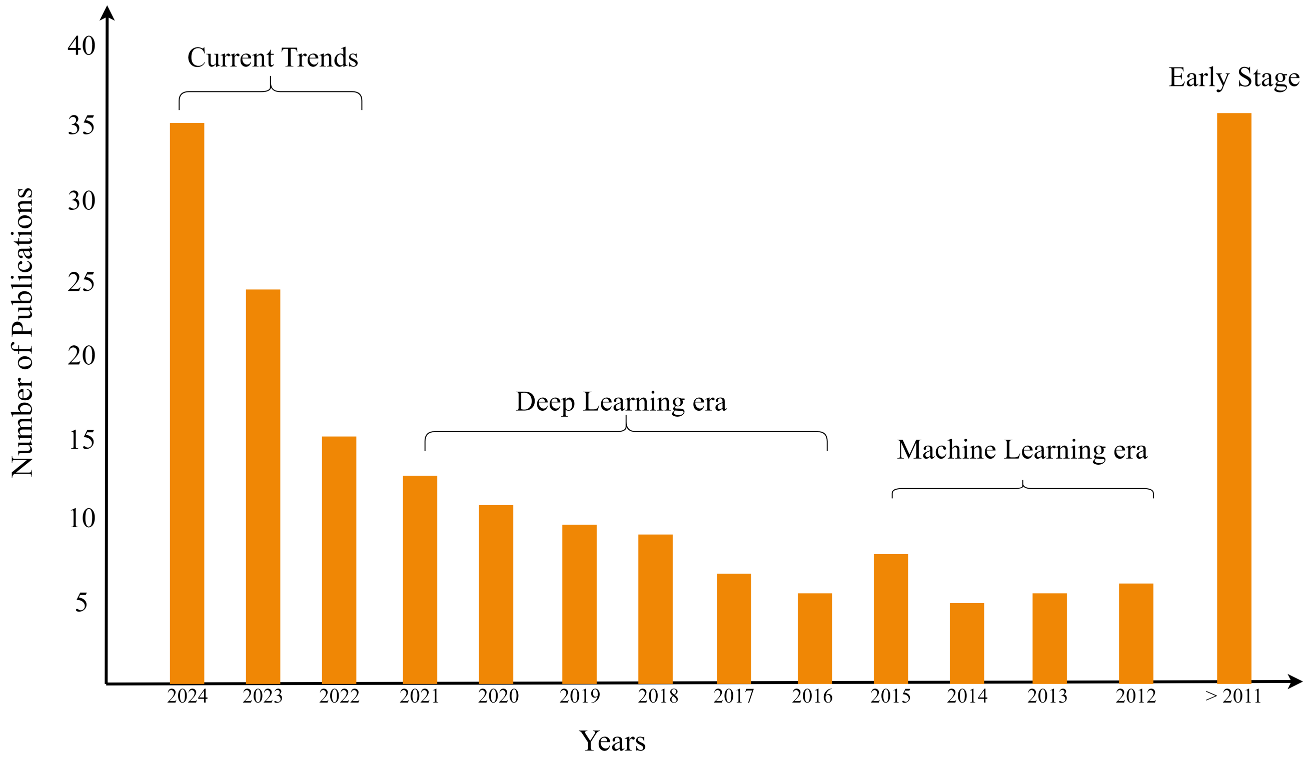

Fig. 1. Number of year-wise publication used in our review paper.

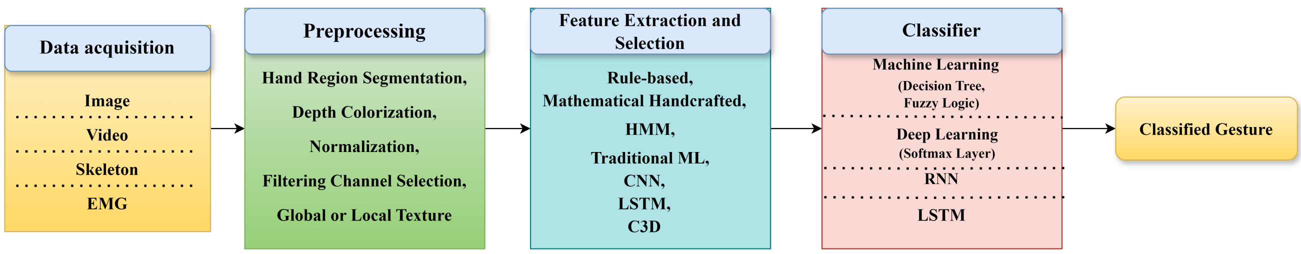

Fig. 2. Flowchart of a Common HGR System.

ing laid at this point for integrating machine learning and, eventually deeper learning techniques in hand gesture recognition. Before 2010, most of the systems of hand gesture recognition were dependent on manual feature extraction and a rule-based algorithm. While these approaches had rather reasonable performance in controlled conditions, they were not flexible or robust enough for wider applicability, thus signifying the further need for more advanced and adaptive techniques over the years [19].

2.1 Input modalities

Datasets are the key to developing and evaluating HGR systems. They serve as benchmarks for testing algorithms, training machine learning models, and validating the effectiveness of recognition techniques. Over these years, a variety of datasets have been developed by researchers depending on targeted applications, ranging from static gesture classification to dynamic sequence analysis, including applications within SLR, virtual reality, and human-computer interaction. These datasets vary in characteristics, starting from the type of data, such as images, videos, or sensor readings, to the number of gesture classes, variety in subjects, and the number of samples [31]. Earlier datasets were quite limited in their scope, usually containing small gesture classes and constrained diversity, which are good for a controlled experiment but not effective for real-world applications. Recently, more effort has been put into the development of more complete datasets with a wider range of gestures, different subjects, and challenging environmental conditions to better mimic real-world settings.

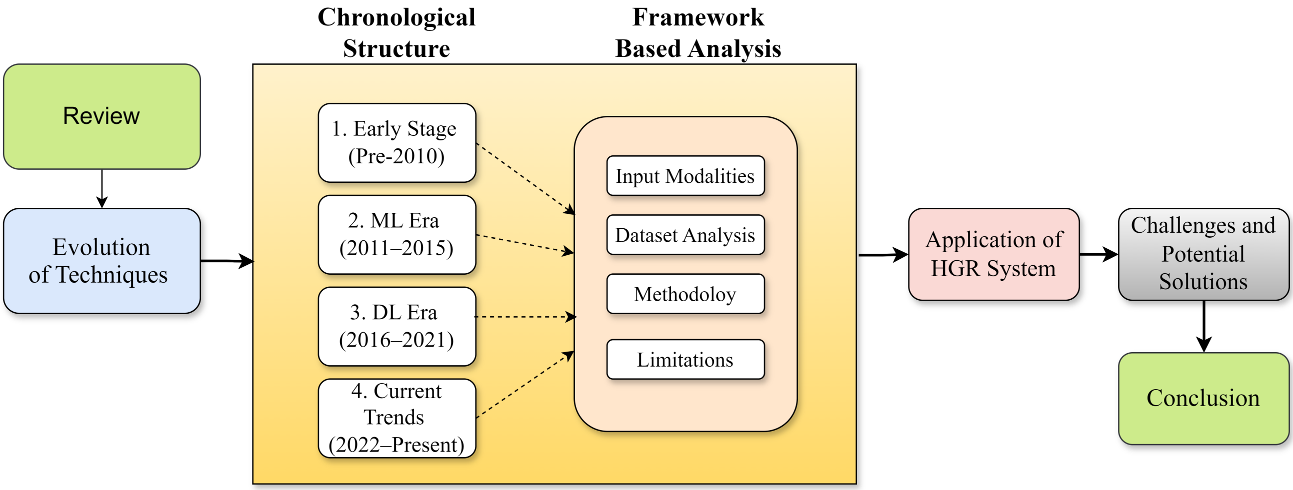

Fig. 3. Structure of the Paper.

Table 1. Gesture Recognition Datasets and Their Limitations

Authors	**Year**	**Dataset Name**	**Dataset Type**	**Subject**	**Gesture**	**Class**	**Sample**	**Limitation**
Chen et al. [20]	2003	-	Image	-	20	-	1200	Does not use different classes for HGR
Alon et al. [21]	2009	ASL	Video	7	8	-	60	There is no sufficient training dataset
Quattoni et al. [22]	2006	Hand Gesture	Video	-	6	5	-	Participant and gesture size were too small
Gupta et al. [23]	2001	ASL	Image	-	10	10	-	Limited gesture size
Shengli Zhou et al. [24]	2009	Hand Gesture	Image	30	-	-	200	Relatively small sample size for the training
Kolsch et al. [25]	2004	Hand Gesture	Image	20	6	-	2300	Limited to diverse posture
Lee et al. [26]	2002	-	-	-	-	-	370	Lack of detailed dataset description
Ming-Hsuan Yang et al. [27]	2002	-	Image	5	18	6	230	Small dataset size and limited to specific gestures
Suk et al. [28]	2010	-	Video	7	10	-	490	A small number of subjects and gestures
Elmezain et al. [29]	2008	-	Video	20	10	-	200	Limited to a small number of gestures
Lee and Hong [30]	2010	-	Image	-	30	4	240	Lack of dataset diversity

Chen et al. [20] proposed an image-based dataset consisting of 20 classes of gestures, with 1200 samples. While it provided the basic foundation for gesture recognition, the lack of diversity in the classes of gestures again made it inefficient for complex HGR systems. Alon et al.[21] used a video-based American Sign Language dataset that consisted of gestures from 7 subjects across 8 classes. However, this dataset had limited training samples, hence restricting scalability and performance in real-world applications. Moreover, Shengli Zhou et al.[24] presented an image-based dataset of 30 gesture classes and 200 samples. While relatively small in size, this database, despite its modest diversity, was insufficient to train HGR models. Kolsch et al.[25] proposed an image dataset that includes samples for 6 classes by 20 subjects and comprised a total of 2300 samples. Its use for general applications has, however been limited because this large-scale dataset possesses poor variation in posture. Other datasets are more geared toward increasing variety in the type of gestures and samples. Such is the case with the image-based dataset presented by Ming-Hsuan Yang et al.[27], including 18 classes of gestures with a total of 230 samples, whereas Suk et al. [28] and Elmezain et al.[29] presented video-based datasets of 10 gesture classes but had limitations concerning subjects and sample variation, which influences their generalizability. Though these datasets indeed laid a very vital ground for research in gesture recognition, their limitations in diversity, sample size, and subject representation bring up the need for more inclusive and expansive datasets. Larger, more diverse, and annotated datasets will be critical in moving HGR systems forward for practical applications.

2.2 Algorithmic Approach

2.2.1 Hidden Markov Model (HMM) for HGR. Hidden Markov Models (HMMs) have proven highly effective for dynamic hand gesture recognition, primarily due to their ability to model sequential data [32]. These probabilistic models represent gestures as sequences of feature vectors extracted from video frames, making them particularly suitable for analyzing temporal variations in hand movements [33]. In this model, pre-processing techniques such as skin color-based segmentation and morphological operations are often used to isolate the hand regions and enhance the trajectories of the image [34, 35]. Moreover, feature extraction methods also range from Hu invariant moments, hand orientation, position, velocity, shape, and 8-directional chain codes [36] [37] [38]. Speeded-up robust Features have also been used because of their efficiency in finding hand features accurately [33]. HMM method training typically involves the Baum-Welch algorithm and often employs a Left-Right Banded topology, which is well-suited for modeling the progressive nature of gestures [33]. Furthermore, recognition is done by algorithms such as Forward or Viterbi, which calculate the likelihood of new gesture sequences matching the trained models [34] [35]. In some cases, alternative approaches such as state sequence analysis have been explored to minimize computational complexity [34]. Applications such as sign language recognition, human-computer interaction, and virtual reality interfaces demonstrate high accuracy in systems based on HMM. For example, more than 90% recognition rates for isolated gestures were reported by various researchers like Shrivastava et al. and Milu et al [33] [36]. However, despite this success, there are still considerable challenges regarding complex and natural gesture recognition in unconstrained environments and remains a very active field of research.

2.2.2 Rule-based HGR. Rule-based models rely on predefined rules and logical conditions to classify hand gestures based on extracted features such as hand shape, position, or motion patterns. These systems are straightforward, interpretable, and computationally efficient which makes them suitable for applications with simple and static gestures [46] [47]. A common approach in rule-based systems involves mapping features to particular gestures by the use of heuristics such as ”if-else” conditions or decision trees. For example, a ”swipe left” gesture can be recognized by detecting horizontal hand movement from right to left within a certain velocity range. Feature extraction covers not only static attributes, such as the position of fingertips and hand shapes but also dynamic attributes like the direction and speed of movement. Some feature-advanced implementations make use of fuzzy logic for spatiotemporal gesture recognition, providing a greater degree of flexibility by allowing more variations [48] [49]. Recognition accuracies have been reported within a range from 80.5% to 98.5% for different types of gestures, proving the efficiency of rule-based systems in controlled conditions Despite their merits, rule-based models begin to suffer with dynamic gestures, noise, or when the set of different gestures is too big or complex to define and manage intricate rules. Nevertheless, they are especially suited for cases where computational efficiency, interpret ability, and low dependence on training data are desired. These models have found their successful application in controlling home appliances, recognizing simple gestures, and improving human-computer interaction [50] [51] [48].

3. MACHINE LEARNING ERA (2011-2015)

We looked at 35 articles between 2011 and 2015. Out of the 35 publications, we discovered that 24 were classified using the conventional ML model, 1 using deep learning [67], and 5 using hybrid models. We received 21 papers with vision-based input modalities, 6 with sensor-based input modalities, and 3 with multimodal input modalities. Furthermore, we discovered that the majority of articles published during this period of time concentrated on creating applications that include robotic control, natural interactions in virtual and augmented reality settings, sign language recognition, and game control. The studies will be categorized in this part according to their datasets, applications, algorithmic techniques, input modalities, and restrictions. In addition, attempts are made to visualize the different model factors.

3.1 Methodological Approaches

The term ”Traditional ML” describes a group of algorithms that create prediction models for classification, regression, clustering, and other tasks using manually designed features and structured data. These strategies include k-nearest neighbors, support vector machines, decision trees, and linear regression. Traditional machine learning relies significantly on domain expertise for feature extraction and works effectively with smaller datasets. In contrast to deep learning, which automatically discovers patterns from raw data. When structured data is involved model interpret-ability is a top concern. It works especially well.
In the machine learning era, most of the papers used tradition machine learning models for classification and vision-based input modalities, and some of the papers built hybrid models by combining traditional ML methods. [68] This paper used three HMM for the hand coordinates, accelerations, and angles grouped and fed into three HMM classifier gestures to produce a classification. In the table-7 included hybrid models and their respective results. [69] used RGB and depth video for training and classification it used SVM that improved accuracy significantly. [70] utilizes an inertial sensor and depth sensor for input and for classification it utilized HMM that achieved an accuracy of 93%. In the table-5, table-6 mentioned various factor of dataset that have been used along with traditional machine learning methods. Table 6 differs from table-5 by input modalities. Table 6 utilized sensor-based input modality where table-5 used vision-based modality. Some of the algorithmic approaches of the machine learning era will be discussed in this section.

3.2 Key Algorithms and Findings

3.2.1 Support Vector Machines (SVM). SVM is most widely used algorithm for classification. It is a supervised learning algorithm that works by finding the optimal hyperplane that best separates the data points into different classes[71, 72]. This paper used PCA for feature extraction and then SVM for classification and they achieved an accuracy of 99.6%. The paper [73][19] used SVM for classification and they gained 98.24% accuracy.

3.2.2 Hidden Markov Models (HMM). A Hidden Markov Model is a graph-based model [75]. It works by converting gestures into sequential symbols. This model is frequently used for classification problems. Separate HMMs are trained for each gesture, class, or sequence using the corresponding training data and then Compute the likelihood of the given observation sequence for all trained HMMs. after that select the HMM (gesture/class) with the highest likelihood as the result [76, 77]. In this study, several papers used this model and got impressive results such as Pradeep and Nevatia et al[78] used the HMM classifier and got 100% accuracy for moving gestures. Xiaoyan Wang, Ming Xia, et al [79] used a hidden markov model and they achieved 98% accuracy. [80] and [81] also utilized HMM and got pretty good accuracy.

Table 2. Summary of HMM-based Gesture Recognition and their Performance

Authors	Year	Input Type	Dataset Type	Method	Accuracy
Chen et al. [20]	2003	Image	-	HMM	85%
Chen et al. [37]	2007	Image	-	AdaBoost	90%
Alon et al. [21]	2008	Video	ASL	CDPP	79%
Murakami et al. [38]	1991	-	-	RNN	94%
Schlomer et al. [39]	2008	Image	-	HMM	90%
Quattoni et al. [22]	2006	Image	-	HCRF	85.25%
Shengli Zhou et al. [24]	2009	Image	-	HMM	92.86%
Kjeldsen et al. [40]	1996	-	-	CPN, HMM	93.63%
Waldherr et al. [41]	2000	-	-	Neural Network (NN)	97%
Kolsch et al. [25]	-	Image	ASL	HMM, ML	95.42%
Starner et al. [42]	1996	-	-	HMMs	91.90%
Vogler et al. [43]	2001	Image	ASL	PaHMMs	93.27%
Starner et al. [42]	1998	Image	ASL	HMM	92%
Suk et al. [28]	2010	Image	-	DBN	84%
Elmezain et al. [29]	2008	Image	-	HMM	95.7%
Gaus et al. [44]	2012	Image	MSL	HMM	83%
Elmezain et al. [45]	2009	Image	Digit Sign	HMM	94%
Mahmoud et al. [46]	2010	Image	Arabic Digit	HMM	95.87%
Liu et al. [47]	2004	Image	English Alphabet	HMM	90%

Table 3. Summary of Rule-based Gesture Recognition and their Performance

Authors	Year	Input Type	Dataset Type	Method	Classifier	Accuracy
Hachaj et al. [52]	2013	Image	Static Pose	Rule-based	GDL	80.5%
Riad et al. [53]	2014	Image	ArSL	Rule-based	Geometric	95.3%
Craven et al. [54]	1997	Image	-	Rule-based	-	82.7%
McGlaun et al. [55]	2004	Image	-	Rule-based	-	93.7%
Ren et al. [56]	2013	Image	Hand Gesture	Rule-based	FEMD	93.2%
Leon-Garza et al. [57]	2022	-	-	Rule-based	-	96.4%
Cui et al. [58]	2000	Image	Hand Sign	Rule-based	-	93.2%
Bauer et al. [59]	2000	Video	German Sign Language	Rule-based	-	91.7%
Hachaj et al. [60]	2012	Image	Movement Sequences	Rule-based	GDL	80.5%
Hachaj et al. [61]	2014	Image	Body Poses	Rule-based	GDL	85.6%
McGlaun et al. [62]	2004	Video	Head Gestures	Rule-based	Template-matching	93.7%
Bedregal et al. [63]	2006	Data Glove	LIBRAS	Fuzzy Rule-based	Fuzzy Logic	92.5%
Lech et al. [64]	2012	Camera	Dynamic Hand Gestures	Fuzzy Rule-based	Fuzzy Logic	89.3%
Chen et al. [65]	2000	Image	Hand Gestures	Rule-based	Decision Tree	85.7%
Rautaray et al. [66]	2005	Camera	Hand Gestures	Rule-based	Fixed Threshold	90.1%
Ohn-Bar et al. [19]	2004	Video	Automotive Interfaces	Rule-based	Geometric Features	88.9%

3.3 Datasets and Applications

One essential component for training a model is a dataset. It has a significant impact on both model accuracy and reduction [77]. The accuracy of models can be affected by several variables, including the number of samples, sample data fluctuation, dataset type, data preparation, and more [82]. Key dataset attributes utilized in the machine learning era are listed in Tables 5 and 6. The majority of the research in this time was on hand gesture detection based on vision-based input modalities. In the input system, we discovered three papers that exploited multi-modality. [69]. RGB and depth video data were combined in a vision-based input, and inertial and depth sensors were combined in sensor-based inputs [70] and [68].

3.4 Limitations

The papers adopt a standard machine learning approach in this period. A small number of articles used multimodalities in the datasets, while the majority used SVM, HMM, and Adabost for classification. Few samples were used in the majority of the dataset to train the models. Additionally, the dataset lacked diversity.

4. DEEP LEARNING ERA (2016-2021)

From 2016 to 2021, researchers started using deep learning-based models to recognize hand gestures effectively. In this era, they focused on CNNs, RNNs, and LSTM models to generalize the HGR system. Traditional ML algorithms face limitations in feature extraction and handling large datasets. DL addresses those challenges by extracting features and improving efficiency. CNNs use dropout and convolutional layers to avoid overfitting, create feature detectors, and use receptive fields for weight [107]. Again, sometimes we need human muscle tension data to analyze. In this case, EMG can be used to interpret data, and RNN with CNNs is used for the

Table 4. Comparison of Traditional Machine Learning and Deep Learning

Feature	Traditional ML	Deep Learning	Limitation/Advantage
Feature Extraction	Manual	Automated	DL reduces human effort but requires domain expertise for ML
Data Requirement	Small to Medium	Large	DL needs extensive data; ML works with limited samples
Interpretability	High	Low	ML models are transparent; DL acts as a "black box"
Complexity	Simple Algorithms	Complex Architectures	DL demands high computational resources
Data Type	Structured (Tabular)	Unstructured (Images, Text)	DL excels with raw, high-dimensional data

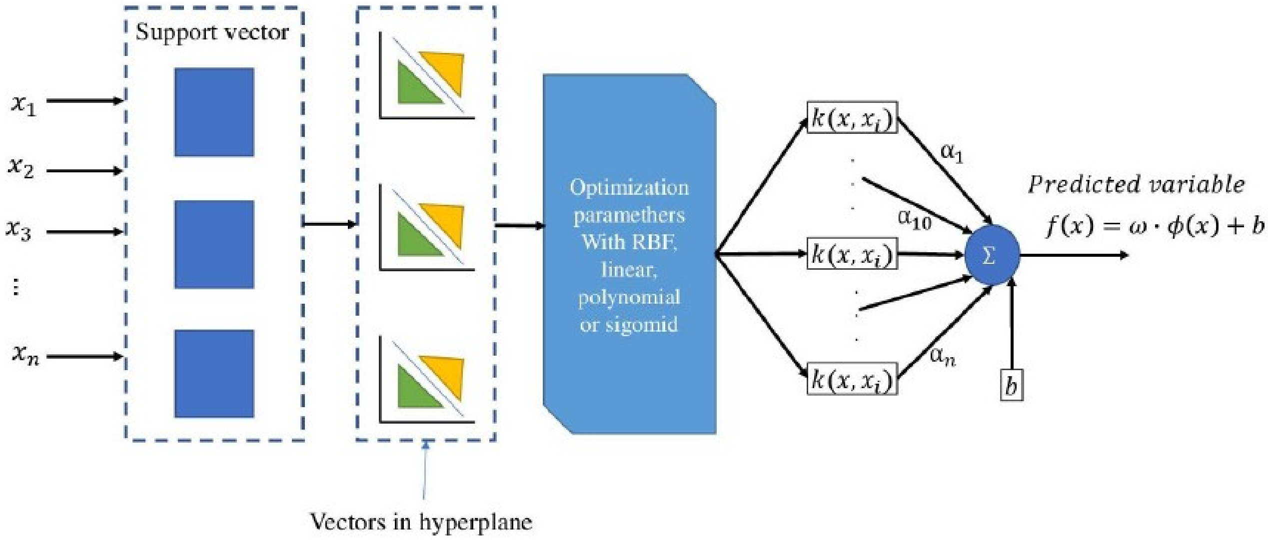

Fig. 4. SVM model architecture [74].

recognition [108]. In this section, we will discuss the input modalities, dataset, methodology, applications, and challenges of this era.

4.1 Datasets

Datasets are crucial to evaluate the model's performance. So, in our research, we tried to find the benchmark datasets used mostly in traditional machine learning and deep learning eras. Most researchers used publicly available datasets to test their system architecture. ASL and Chalearn 2016 are the two most used datasets. Other used datasets include the National University of Singapore (NUS) Dataset [109], Chinese Sign Language (CSL) Dataset [110], InterSegHands Dataset [111], DHG-14/28 [112] etc. Table 1 represents the comprehensive discussion of datasets used in this era.

4.2 Input Modalities

The input modality is important for hand gesture recognition as it defines the data type used in any model. It has a direct connection with a system's accuracy and applications. The main modalities many researchers used are RGB-based, Skeleton, Depth, Audio, EMG, EEG, and Fusion data. In Table 1, we have discussed the datasets of multiple modalities and their brief description.

4.3 Methodological Approaches

The introduction of Machine Learning (ML) and Deep Learning (DL) technologies has led to noticeable improvements in Hand Gesture Recognition (HGR) systems [122, 123]. These models have transformed the recognition and interpretation of gestures, allowing for extremely reliable and accurate systems that can be used in various fields, including virtual reality, assistive technologies, human-computer interaction, and sign language recognition [124]. This section will discuss the methodological approaches used in this era.

4.3.1 Traditional ML Approach. X. Zhang et al. developed a new HMM algorithm combined with DTW to improve the efficiency and accuracy of dynamic hand gesture recognition [125]. Yi Li et al. [126] discussed an integrated dynamic hand gesture recognition model based on the improved DTW (Dynamic Time Warping) algorithm that has a significant impact on the efficiency of dynamic trajectory analysis. They preprocessed the video, extracted characteristics, and then normalized them into a sequence template as part of their technique [126]. The gesture recognition result met the category with the least deviation from the template after the generated sequence was compared to templates in the training set. Nevertheless, the DDW method struggles with big data sets and intricate movements and does not use statistical methods during training.

Table 5. Comprehensive Analysis of Hand Gesture Recognition Studies

Author	Dataset Name	Type	Classes	Samples	Feature Extraction	Classifier	Accuracy
Mandeep Kaur Ahuja et al. [83]	-	Static	5	80	PCA	Template matching	91.25
Dipak Kumar Ghosh et al. [71]	Danish/International Hand Alphabet Dataset	Static	25	1000	LCS + Block-based	SVM	99.50
Alaa Barkoky et al. [84]	Persian Sign Language (PSL)	Static	-	300	-	Thinning methods	96.62
Hui Li et al. [85]	Kinect custom dataset	Static	6	-	HOG	AdaBoost	93.50
Ran Wang et al. [86]	-	Dynamic	5	600	-	Geometric algorithms	90.30
M. M. Gharasuie et al. [87]	-	Dynamic (video)	10	600	-	Hidden Markov Models (HMMs)	99.17
Hsiang Yueh Lai et al. [88]	-	Dynamic	9	330	-	Convex hull method	95.10
Alisha Pradhan et al. [89]	Real-time + static gesture system	-	-	-	-	Convex hull algorithm	-
Heba M. Gamal et al. [90]	Cambridge Hand Gesture Dataset	Static	4	70	Fourier descriptors	SVM, KNN, Euclidean	62.5–98.75
Srinivas Ganapathyraju [91]	Webcam-based real-time input	Dynamic	4	-	-	Convex hull algorithm	-
Xianghua Li et al. [92]	-	Static	5	500	Zhang-Suen thinning	Geometric features	82.93
Zhong Yang et al. [93]	-	-	18	1800	-	HMM	96.67
Lalit K. Phadtare et al. [94]	HamNoSys Dataset	Static	-	40	-	Shape Context + Plane Fit	-
D.K. Vishwakarma et al. [95]	Marcel-Triesch Dataset	-	-	-	-	-	97.47
Lukas Prasuhn et al. [96]	American Sign Language (ASL)	Static	19	1425	HOG	L2 Distance + Brute Force	-
Rajat Shrivastava [97]	-	Dynamic (video)	5	-	Hu moments + orientation	HMM	-
Bhumika Pathak et al. [98]	-	Dynamic	22	-	Key frame extraction	MSVM	-

Table 6. Traditional ML for Classification and Sensor-Based Input Modalities

Author	Dataset Name	Type	Classes	Samples	Feature Extraction	Classifier
H. Seyedarabi et al. [99]	-	Dynamic	10	-	-	Hidden Markov Models (HMMs)
S. Dai et al. [73]	-	Static	36	3600	-	Support Vector Machines (SVM)
Y. Wang et al. [100]	-	Dynamic	5	-	-	Linear Discriminant Functions
S. Dai et al. [101]	-	Static	16	1600	-	Bayesian Neural Network (BNN)
B. Luo et al. [102]	-	Dynamic	36	300	-	Support Vector Machines (SVM)

4.3.2 CNN-Based Models. As ML-based models could not show better performance due to some limitations, researchers tend to use CNN-based architectures for better performance and efficiency. Adithya V. et al. [109] proposed a CNN model to recognize static hand gestures. They used RGB images and utilized the fully connected layers to classify the gesture properly, ending with an accuracy of more than 94% [128]. Osama Mazhar at al. utilized RGB vision and dynamic gestures dataset to evaluate their model performance [129]. Figure 6 Shows the typical architecture of a CNN model, and on figure 7. represents the working procedure of a CNN model.

Table 7. Hybrid Model for Classification and Vision-Based Input Modalities

Author	Dataset Name	Dataset Type	Classes	Samples	Feature Extraction	Classification	Accuracy
S. Imran et al. [103]	-	Static	-	50	-	SVM-based classification	95
M. Rahmati et al. [104]	-	Dynamic (Video)	12	-	HOG	Fuzzy SVM + Dynamic Bayesian Networks (DBN)	90
R. Kapoor et al. [105]	ASL	Static	18	-	-	Self-Organizing Maps (SOMs)	92
C. Yang et al. [106]	-	Dynamic	-	-	Histogram of Oriented Gradients (HOG)	AdaBoost + Hidden Markov Models (HMMs)	-

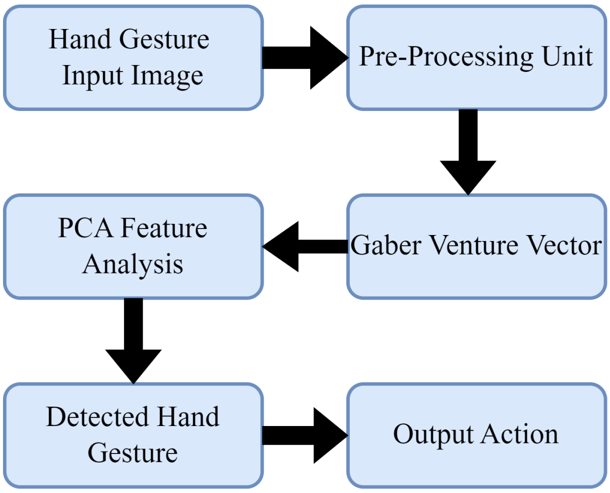

Fig. 5. Block Diagram of ML Approach [127].

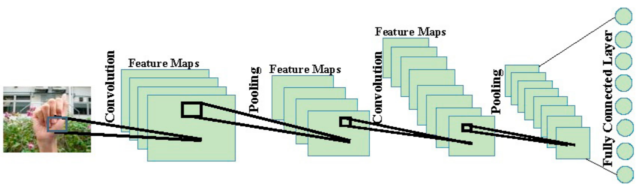

Fig. 6. Typical Architecture of CNN [128].

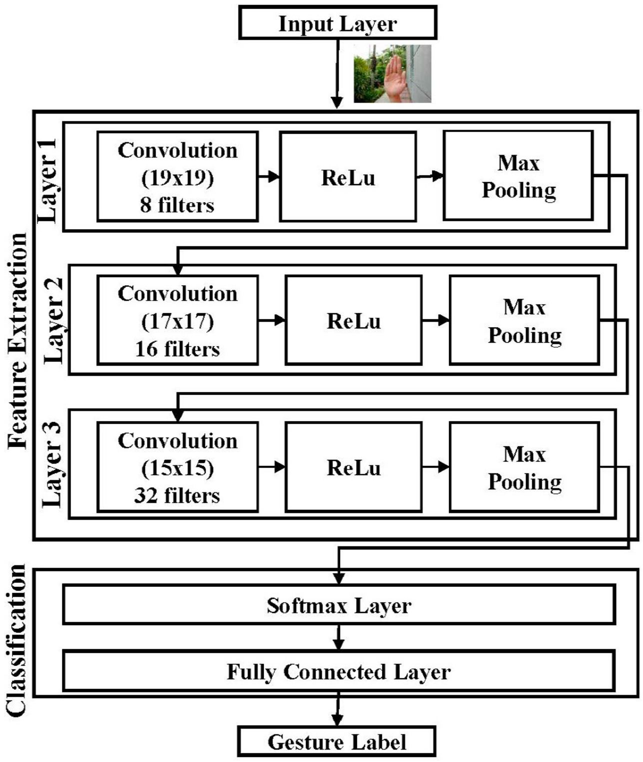

Fig. 7. The schematic representation of the CNN model for hand posture recognition [128].

4.3.3 RNN-based Models. Recurrent Neural Network (RNN) is a deep learning architecture that can be trained to process sequential data such as words, sentences, and time series into specific sequential data output [130]. It mimics the human brain performing sequential data transformations such as translating text from one language to another. Philipp Koch at al. proposed an RNN architecture based on accelerometer data since accelerometers may also be readily included in mobile devices. This tiny network performs far better than state-of-the-art hand gesture detection techniques that depend on multi-modal data, according to experiments conducted on three datasets [131, 130]. Many researchers use RNN with CNN or LSTM to get better performance which we will discuss in the fusion model part.

4.3.4 Fusion-Based Models. Kenneth Lai et al. proposed a model combining convolutional neural networks (CNN) and recurrent neural networks (RNN) to automate hand gesture detection using depth and skeletal data [112, 130]. CNN extracts significant spatial information from depth pictures, while RNN identifies movement patterns for individual joints. They achieved an accuracy of 85.46%. To discover hand muscle signals that may be utilized to operate drones, Ray Antonius at al. proposed a CNN-RNN neural network technique. Myo, an 8-channel EMG device, and 14,000 datasets of nine distinct movements are used in the procedure [108]. Using a drone kit to test the trained models on a drone, the average positive detection rate for each gesture was 96.60%. This high positive recognition rate may pave the way for wearable technology that allows for real-time human control. To learn both short-term and

Table 8. Brief Description of Datasets of Multiple Modalities from 2016–2021

Modality Type	Dataset Name	Year	Type	No. of Subjects	No. of Classes	Total Samples
RGB Image	OUHAND	2016	ASL	23	10	3000
RGB Image	NTU Datasets	2018	Digits	10	10	1000
RGB Image	ASL-10	2020	SL	22	10	2800
RGB Image	MUD	2021	ASL	-	36	2520
RGB Image	HGR 1	2021	ASL	12	25	899
RGB Image	Marcel	2021	ASL	-	6	5531
RGB Image	NUS II	2021	ASL	40	10	2000
RGB Video	ChaLearn LAP IsoGD [113]	2016	Hand Gesture	21	249	47933
RGB Video	DVS Gesture [114]	2016	Hand Gesture	29	11	1342
RGB Video	NVGesture [115]	2016	Hand Gesture	-	25	1532
RGB Video	IsoGD	2017	Multiple	-	13	47933
RGB Video	CSL-Daily [116]	2021	Chinese	10	2000	20654
RGB Video	SIGNUM [117]	2021	German	25	1230	15075
RGB Video	BOBSL [118]	2021	British	85	395	47551
RGB Video	LSFB [119]	2021	French, Belgian	100	6883	85132
Skeleton Data	ASLLVD [120]	2019	ASL	-	-	9748
Skeleton Data	[120]	2019	General	-	-	2800
Skeleton Data	MSRA [120]	2019	General	-	-	76500
Skeleton Data	PSL	2020	Pakistani	-	-	2700
Skeleton Data	AUTSL [121]	2020	Turkey	-	-	38336
EMG Data	DB1	2019	HGR	27	-	-
EMG Data	DB2	2014	HGR	40	-	-
EMG Data	DB3	2014	HGR	11	-	-
EMG Data	DB4	2020	HGR	10	-	-
EMG Data	DB5	2021	HGR	10	-	-
EMG Data	DB6	2019	HGR	10	-	-
EMG Data	DB7	2016	HGR	22	-	-
EMG Data	DB8	2020	HGR	12	-	-
EMG Data	DB9	2018	HGR	77	-	-
EMG Data	DB10	2017	HGR	45	-	-
EMG Data	Arm Band	2019	HGR	6	-	-
EMG Data	EMG High Density	2021	HGR	41	-	256

long-term characteristics from video inputs, Wenjin Zhang et al. [132] introduced a unique deep-learning network for hand gesture identification that integrates various modules. The network divides the input video into groups of frames, chooses frames at random, and extracts features using a convolutional neural network [133]. evaluated on well-known hand gesture datasets, the model yields competitive results and demonstrates resilience when evaluated on an enlarged dataset with a greater variety of hand movements [134]. In recent years, LSTMs have demonstrated exceptional performance in a few domains of sequence data processing. Researchers have started using LSTMs with CNN and RNN architectures. Figure 4 shows the block diagram of an LSTM.

5. CURRENT TRENDS IN HGR (2022–PRESENT)

Hand Gesture Recognition has seen rapid advancements in recent years, driven by innovations in machine learning, multimodal integration, and lightweight model architectures [163].

5.1 Different Architecture of Current HGR system

5.1.1 Transformer Models. Transformers, such as Vision Transformers (ViT) and Graph Vision Transformers (GViT), have been widely adopted for modeling complex hand gesture sequences. Their self-attention mechanism captures spatial and temporal dependencies, enabling higher accuracy in dynamic gesture recognition. Recent models like GestFormer incorporate Multiscale Wavelet Pooling to enhance efficiency [164, 165].

Table 9. Summary of Methodology Used from 2016–2021

Input Modality	Author	Year	Dataset	Feature Model	Classifier	Accuracy (%)
RGB Image	Wenjin Tao et al. [135]	2018	ASL	CNN	CNN	84.80
RGB Image	C. Bhuvaneshwari et al. [136]	2019	ASL, Indian SL	LSTM	Fully Connected Layer	-
RGB Image	Nada B. Ibrahim et al. [137]	2018	Arabic SL	-	HMM	97.00
RGB Image	Shin-ichi Ito et al.	2020	Japanese SL	CNN	MVSM	94.30
RGB Image	[109]	2020	ASL	CNN	Softmax	99.96
Video	Oliveira et al. [138]	2018	Iris SL	PCA	PCA	95.00
Video	Jing-Hao Sun et al. [139]	2018	SKIG	C3D+LSTM	Softmax	98.60
Video	Lionel Pigou et al. [140]	2018	Montalbano	RNN	Softmax	67.71
Video	Du Jiang et al. [141]	2019	ChaLearn LAP IsoGD	ResC3D	Softmax	50.93
Video	Prachi Sharma et al. [142]	2020	ChaLearn LAP IsoGD	C3D+Pyramid	Softmax	49.20
Video	Qing Gao et al. [143]	2020	ASL	2S-CNN	Softmax	92.00
Video	Zhimin Gao et al. [144]	2020	SKIG	R3DCNN+RNN	Softmax	100.00
Video	Kayo Yin et al. [145]	2020	PHOENIX14-T, ASLG-PC12	STMC-Transformer	Bi-LSTM, CTC	96.60
Video	Zhenxing Zhou et al. [116]	2021	Kinetics-400, HSL	HOG	(3+2+1)D CNN	94.60
Skeleton Image	De Smedt et al. [146]	2016	ASL	HOG	Fisher Kernel, SVM	86.86
Skeleton Image	De et al. [147]	2016	DHGD	SoCJ + HoHD + HoWR	Softmax	80.00
Skeleton Image	Boulahia [148]	2017	DHGD	Boulahia	Softmax	80.48
Skeleton Image	Liu et al. [149]	2017	ASL	LSTM	LSTM	96.30
Skeleton Image	Konstantinidis et al. [150]	2018	Argentinian SL	VGG-19 Network	CNN, RNN, LSTM	98.09
Skeleton Image	Juan C. Nunez et al. [151]	2018	DHGD	CNN+LSTM	Softmax	74.19
Skeleton Image	Yan et al. [152]	2018	DHGD	STA-GCN	Softmax	87.10
Skeleton Image	Ma et al. [153]	2018	DHGD	NIUKF-LSTM	Softmax	80.44
Skeleton Image	W. Wei et al. [154]	2019	AUSTL	Multi-Stream GCN	CNN	99.00
Skeleton Image	Hengyang Si et al. [155]	2019	Shrec	CNN+LSTM	Softmax	89.52
Skeleton Image	Jiang et al. [156]	2021	Turkish SL	SSTCN	CNN, LSTM	98.53
Skeleton Image	Jiang et al. [157]	2021	Chinese SL	SL-GCN, SSTCN, 3DCNN	GEM	99.81
Skeleton Image	Rastgoo et al. [158]	2021	Persian SL	3DCNN	2DCNN, 3DCNN, LSTM	99.80
EMG Data	Zhang et al. [159]	2019	American SL	RNN	RNN	89.60
EMG Data	Cote-Allard et al. [160]	2019	American SL	CNN	CNN	97.81
EMG Data	Wei et al. [161]	2019	American SL	CNN	DL	90.00
EMG Data	Lee et al. [162]	2021	American SL	RMS, VAR, MAV, SSC, ZC, WL	ANN	94.00

5.1.2 Real-Time Systems. Lightweight architectures, such as FGDSNet and HGR-Lite, have been developed for edge devices, ensuring high accuracy and minimal latency. These models prioritize computational efficiency without compromising performance, making them suitable for real-time applications in robotics and gaming [166, 167].

5.1.3 Multimodal Fusion. Combining multiple data modalities, including RGB, depth, and skeletal data, has significantly improved the robustness of HGR systems. Models like MF-HAN integrate these inputs through hierarchical self-attention mechanisms, achieving superior performance even in challenging environments [168, 169].Integrating RGB, depth, and sensor data has enhanced the robustness of gesture recognition systems [168, 169]. MoviNet, a two-stream architecture for wrist-worn cameras, achieved an impressive 98.48% accuracy in cross-subject evaluations [170].

5.1.4 Deep Learning Architectures. Deep learning remains the backbone of HGR research. Novel CNN-based models such as CMLG-Net [171] and lightweight CNNs for thermal video-based gestures [172] have demonstrated exceptional accuracy and robustness, achieving 97% in challenging datasets. Transformer-based architectures like MVTN have shown exceptional performance in handling hand variations, achieving a 98.61% accuracy rate [173].

5.1.5 EMG-Based Models. Shin et al. [174] introduced a multi-stream approach for sEMG signal recognition, achieving 94.31% (DB1) and 98.96% (DB9) on Ninapro datasets, significantly improving gesture recognition in prosthetics and robotics [174]. Other works explored diverse domain feature enhancement for EMG-based gesture recognition, reaching 97.43% accuracy with KNN [175].

5.1.6 Hybrid Systems. Hybrid approaches combining vision-based and sensor-based data have gained prominence. Examples include multimodal fusion in gesture recognition [176], transformer-based hybrid systems such as TY-Net [177], and CNN-LSTM frameworks for dynamic gestures [166, 178].

5.2 Specialized Applications

Thermal video recognition, as explored by Birkeland et al., highlights the potential for field testing in hard-of-hearing communities [172]. Novel methodologies like sterile training techniques with radar systems improve static gesture recognition rates up to 95% [179].

5.3 Input Modalities

The input modalities for hand gesture recognition systems play a crucial role in determining their robustness, adaptability, and performance. HGR systems can be broadly categorized into three types based on their input data sources:

5.3.1 Vision-Based Input. Vision-based systems utilize cameras to capture RGB or depth images of hand gestures. These systems excel at capturing spatial details and are widely used in gaming and virtual reality applications. However, they face challenges in low-light conditions and complex backgrounds [180, 176].

5.3.2 Sensor-Based Input. Sensor-based systems use wearable devices such as electromyography (EMG) sensors, force sensors, or inertial measurement units (IMUs) to capture gesture data. These systems are highly accurate for fine-grained gestures but require users to wear specialized hardware [181, 182].

5.3.3 Multimodal Input. Multimodal systems combine visual and sensor-based data to enhance accuracy and adaptability. By leveraging complementary strengths, they address the limitations of single-modality systems. Examples include integrating RGB and depth data or combining EMG signals with visual inputs [168, 183].

Table 10. Comparison of Input Modalities in HGR Systems

Input Modality	Advantages	Challenges
Vision-Based	High spatial resolution; no hardware dependency.	Sensitive to lighting and background variations.
Sensor-Based	Accurate for fine-grained gestures; wearable and portable.	Requires specialized hardware; user-dependent variability.
Multi-modal	Robust performance across conditions; improved accuracy.	Higher computational complexity; integration challenges.

5.4 Algorithmic Approaches

Hand Gesture Recognition (HGR) has seen the application of diverse algorithmic approaches, ranging from traditional machine learning techniques to cutting-edge deep learning and hybrid models [184]. Each approach offers unique benefits and limitations, as discussed below.

5.5 Traditional Machine Learning

Traditional ML methods, such as Support Vector Machines (SVMs), K-nearest neighbors (KNNs), and Random Forests, have been widely used in early HGR systems [185]. These methods rely on handcrafted feature extraction and work well for small-scale datasets. However, they often struggle to generalize in complex, real-world scenarios [186, 187].
Advantages:

—Simple and computationally efficient [186].
—Performs well on small datasets [187].

Challenges:

—Limited scalability for large-scale datasets [186].
—Reliance on manual feature engineering [187].

5.6 Deep Learning

Deep learning approaches leverage neural networks, such as Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and Transformers, to automatically learn features from raw data [188]. This paradigm shift has significantly improved the accuracy and generalization of HGR systems [165, 189].
Advantages:

—High accuracy and scalability [189].
—Automated feature extraction [166].

Challenges:

—Computationally intensive [165].
—Requires large annotated datasets [189].

5.6.1 Hybrid Models. Hybrid models combine traditional ML and deep learning techniques to achieve the best of both worlds [190]. For instance, hybrid architectures may use ML-based classifiers on deep features extracted by CNNs or use multimodal fusion techniques for improved robustness [176, 191].
Advantages:

—Improved robustness and accuracy [191].

Table 11. Feature Extraction and Classifier Techniques in HGR

Study	Feature Extraction	Classifier	Highlights
Birkeland et al. [172]	Lightweight CNN	Softmax	Real-time thermal gesture recognition for healthcare applications.
Shin et al. [174]	Multi-Stream CNN Features	Multi-Class SVM	98.96% accuracy on Ninapro DB9.
Nguyen et al. [170]	MoviNet Architecture	Dual-Stream Fusion	98.48% accuracy in dynamic gestures with wrist-worn cameras.
Garg et al. [173]	Transformer Features	Transformer Encoder	High accuracy in video gesture recognition.
Zhang et al. [171]	Local-Global CNN	Custom Classifier	Robust hand gesture authentication with low error rates.

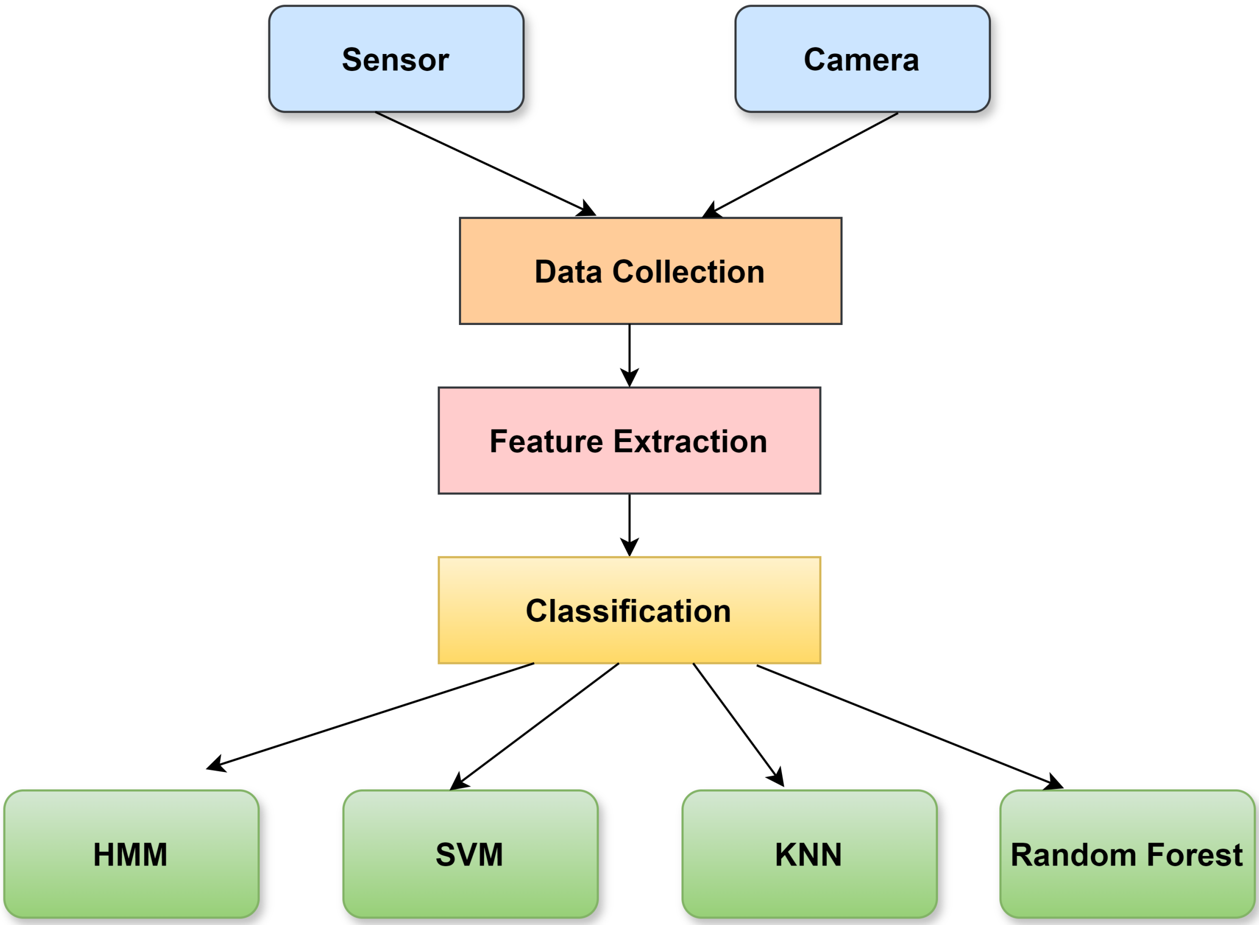

Fig. 8. The Workflow of Traditional Machine Learning

—Capable of handling multimodal inputs [176].

Challenges:

—Higher complexity in design and training [176].

—Increased computational requirements [191].

5.7 Algorithm Usage in HGR

Various algorithms have been employed in Hand Gesture Recognition (HGR) systems across traditional machine learning, deep learning, and hybrid approaches. Table 12 provides an overview of the most commonly used algorithms in these categories, along with their approximate usage percentages as reported in the literature.

5.8 Comprehensive Dataset Analysis

The choice of dataset significantly influences the performance and generalizability of hand gesture recognition systems. Benchmark datasets provide the foundation for training and evaluating hand gesture recognition (HGR) models. However, their characteristics,

Table 12. Algorithm Usage in Hand Gesture Recognition Approaches

Approach	**Algorithm**	**Usage (%)**
Traditional ML	SVM [186]	15%
	Random Forest [192, 182]	10%
	KNN [175, 193]	8%
	Neural Networks (NN) [194]	Unspecified
Deep Learning	CNN [166, 181, 172]	30%
	RNN [189]	12%
	Transformer [165, 177]	18%
	VGG [195]	Unspecified
Hybrid Models	CNN-LSTM [178]	10%
	DeReFNet [196]	Unspecified
	Multimodal Fusion [176, 197]	7%
	CNN-SVM [191]	5%
Other	Graph Convolutional Networks (GCN) [198, 199]	5%
	Attention-Based Networks [200, 199]	10%
	GCN + MHSA [199]	Unspecified

Table 13. Comparison of Algorithmic Approaches in Hand Gesture Recognition

Approach	**Advantages**	**Challenges**	**Best Use Case**
Traditional ML	Computationally efficient; works well on small datasets.	Limited scalability; requires manual feature engineering.	Simple gestures in controlled environments [186, 187].
Deep Learning	High accuracy; automated feature extraction.	Computationally expensive; requires large datasets.	Complex gestures in dynamic environments [166, 189].
Hybrid Models	Combines strengths of ML and DL; handles multimodal inputs.	Higher design complexity; increased computational cost.	Multimodal gestures in real-world scenarios [176, 191].

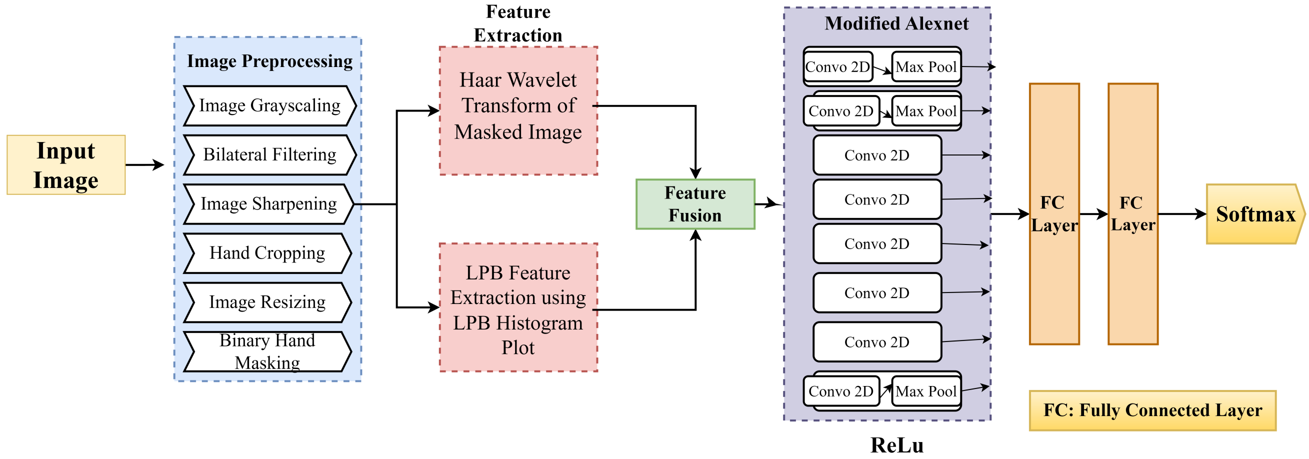

Fig. 9. Working Procedure of Traditional Machine Learning for HGR [186].

limitations, and inherent gaps significantly impact system performance. This section provides an updated analysis of key datasets, focusing on their characteristics, limitations, and benchmark metrics such as accuracy and loss.

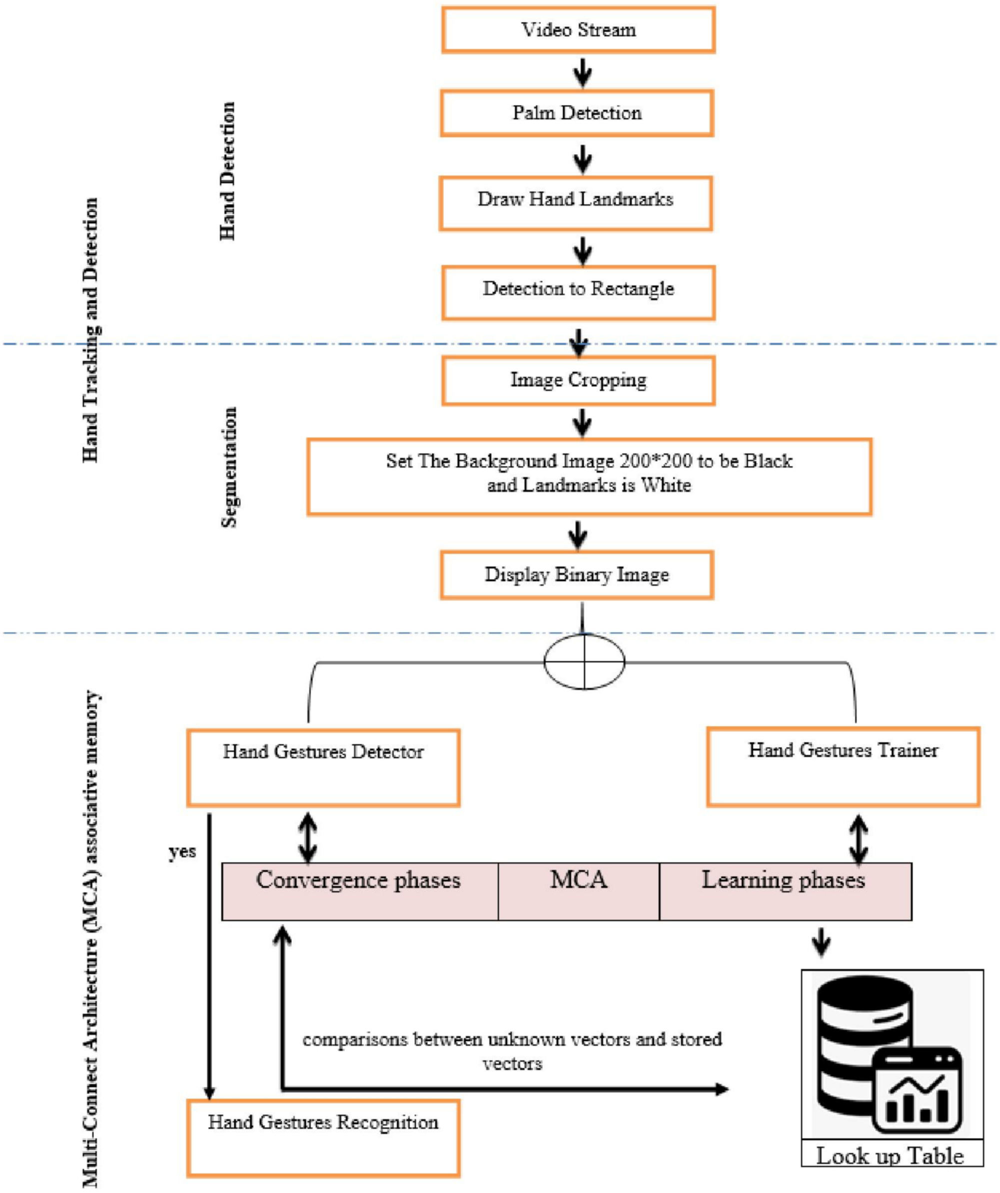

Fig. 10. The Workflow of Hybrid Model for HGR System[191].

5.9 Key Characteristics and Gaps

The development of robust and accurate Hand Gesture Recognition (HGR) systems heavily relies on the datasets used for training and evaluation [201]. Each dataset exhibits unique characteristics that align with specific algorithmic needs while exposing critical gaps that limit the generalization and scalability of HGR models.

—**Transformer Models** Vision Transformers (ViT) and Graph Vision Transformers (GViT) have become prominent in the HGR domain due to their ability to model complex gesture sequences effectively. These models leverage self-attention mechanisms to capture spatial and temporal dependencies, which are essential for recognizing dynamic gestures. However, their performance is constrained by the limited size and diversity of available datasets. Most datasets used in transformer-based HGR systems consist of predefined gestures performed in controlled environments, leading to challenges in adapting to real-world scenarios where gestures vary significantly across individuals, lighting, and backgrounds [164, 165]. Additionally, the computational requirements of these models can pose barriers to their deployment in real-time systems.

—**Real-Time Systems** Lightweight architectures, such as FGDSNet and HGR-Lite, focus on achieving high accuracy with minimal computational overhead, making them ideal for resource-constrained devices. These models are highly optimized for applications such as robotics and gaming, where latency is critical. Despite their advantages, a significant gap exists in the availability of real-time annotated datasets. Current datasets often fail to account for dynamic environments and natural variations in user behavior, such as speed, occlusion, and background clutter. Moreover, real-time systems face difficulties in scaling to diverse user populations, as the majority of datasets are limited to a specific demographic or small sample size [167, 183].

—**Multimodal Fusion** The integration of multiple data modalities, including RGB images, depth maps, and sensor data (e.g., EMG, IMU), has proven effective in enhancing the robustness and adaptability of HGR systems. Multimodal fusion addresses the limitations of single-modality systems by leveraging complementary information from diverse data sources. For instance, combining visual data with sensor-based inputs can improve recognition in scenarios with poor lighting or occlusions. However, multimodal datasets often lack standardized benchmarks and consistent annotation practices, making it difficult to compare the performance of different approaches [202]. Furthermore, the computational complexity associated with processing multiple modalities can hinder real-time deployment [168, 169]. There is also a need for datasets that include more complex gestures and interactions to reflect real-world applications, such as assistive technologies and augmented reality systems.

—**Ethical and Privacy Concerns** While not directly related to the technical characteristics of datasets, ethical and privacy considerations play a significant role in the development and usage of HGR datasets [203]. Many datasets involve video recordings of individuals, raising concerns about consent, anonymity, and data security. Addressing these issues is crucial to ensure the ethical deployment of HGR systems in applications such as surveillance, healthcare, and personal assistance.

—**Dataset Diversity and Scalability:** A recurring challenge across all dataset categories is the lack of diversity in gesture types, user demographics, and environmental conditions. Most existing datasets are limited in terms of geographic and cultural representation, gesture complexity, and interaction scenarios. Expanding the diversity of datasets to include gestures from different cultures, age groups, and contexts will significantly enhance the generalization of HGR systems.

—**Real-World Applicability:** Finally, a critical gap lies in the limited testing of HGR systems in real-world conditions. Most datasets are collected in controlled environments, which do not reflect the challenges encountered in practical applications, such as varying lighting, occlusions, and noise. Developing datasets that replicate real-world scenarios is essential to improve the robustness and reliability of HGR systems [204].

5.10 Feature Extraction and Classifier

Feature extraction and classification are two critical components of Hand Gesture Recognition (HGR) systems [205]. The performance of HGR models heavily depends on the ability to extract meaningful features and accurately classify gestures [133]. This subsection highlights common techniques and their applications in HGR.

5.10.0.1 Feature Extraction. Feature extraction transforms raw input data into a compact and informative representation, enabling models to focus on the most relevant aspects of hand gestures [206]. Commonly used methods include:

—**CNN-based Features** Convolutional Neural Networks (CNNs) are widely used for extracting spatial features from RGB and depth images, offering high accuracy for static and dynamic gestures [166, 207].

—**Attention Mechanisms:** Hierarchical and self-attention mechanisms enhance feature extraction by focusing on key regions and temporal dependencies [168].

—**Hybrid Techniques** Techniques such as combining spatial and temporal features (e.g., CNN-LSTM hybrids) improve recognition for complex gestures involving motion [208, 181].

—**Handcrafted Features:** Traditional methods like Local Binary Patterns (LBP) and Haar wavelets are efficient for small datasets but lack scalability for real-world applications [209].

5.10.0.2 Classifier. The classifier is responsible for mapping extracted features to predefined gesture classes. Different classification techniques are used based on the dataset and application:

—**Softmax Classifier** Common in deep learning models, softmax is effective for multi-class classification tasks [181].

—**Support Vector Machine (SVM):** Often paired with handcrafted features or CNN-extracted features, SVMs are robust for binary and multi-class classification [210].

—**Bayesian Classifiers:** These probabilistic models are effective in fusing multimodal data and improving classification robustness [191].

—**Custom Architectures:** Novel classifiers such as Associative Memory and Gated Attention Networks have shown promise in improving performance for specific applications [183, 167].

5.11 Dataset Accuracy, Loss, and Gaps

Evaluating the performance of Hand Gesture Recognition (HGR) models requires robust datasets with clearly defined metrics such as accuracy, loss, and identified gaps. This subsection provides a detailed analysis of benchmark datasets, highlighting their accuracy, loss values, the best-performing models, and the critical gaps that limit their utility in real-world applications.

—**Accuracy:** Accuracy is a key metric for assessing the effectiveness of HGR models [252]. Higher accuracy indicates better alignment between model predictions and actual gestures. For instance, NVGesture achieved an impressive accuracy of 98.1% using the GestFormer model, demonstrating its suitability for dynamic gesture recognition in controlled environments [165].

—**Loss:** Loss measures the model's error during training and evaluation. Lower loss values signify better optimization of the model. For example, the SHREC17 dataset reported a loss of 0.045 with the ViT model, highlighting its ability to learn complex spatial and temporal dependencies [164].

—**Gaps:** Despite advancements, several datasets exhibit limitations, such as insufficient gesture diversity, lack of real-world annotations, and challenges in handling dynamic or fine-grained gestures. Addressing these gaps is crucial for enhancing the generalizability of HGR systems.

Table 14. Feature Extraction and Classifier Details from Selected Studies from 2022–Present (Part 1)

Author(s)	**Year**	**Feature Extraction**	**Classifier**
Barona Lopez et al. [166]	2024	CNN-LSTM	Post-Processing Algorithm
Sharma et al. [211]	2024	Quantized CNN	Softmax
Balaji and Prusty [168]	2024	Hierarchical Self-Attention	Multimodal Fusion
Zhang et al. [181]	2023	Multi-Attention Mechanisms	Softmax
Farid et al. [210]	2024	SSD-CNN with Deep Masks	SVM
Tang et al. [212]	2024	3D Printable Sensor Features	Graph Neural Network (GNN)
Shin et al. [208]	2024	Temporal and Spatial Features	CNN-TCN Hybrid
Wang et al. [213]	2024	Virtual Dimension Increase	Separability Feature Vector (SFV)
Awaluddin et al. [164]	2024	Hybrid Image Augmentation	CNN
Garg et al. [165]	2024	Multiscale Wavelet Pooling	Transformer
Sen et al. [214]	2024	Optimized YOLOv5 Features	Bayesian Classifier
Bhaumik et al. [169]	2023	Spatial Feature Attention	Custom Deep Network
Sarma et al. [180]	2023	Attention-based Semantic Segmentation	VGG16 and C3D
Mahmud et al. [176]	2023	Quantized Depth and Skeleton	CRNN
Sunanda et al. [209]	2024	Haar Wavelet Transform + LBP	Modified AlexNet
Zhou et al. [167]	2024	Cosine Similarity Attention	FGDSNet
Fadel et al. [183]	2024	Multi-Connect Architecture	Associative Memory
Mohammadi et al. [215]	2023	RGB Image Processing	YOLOv4
Bamani et al. [216]	2024	HQ-Net for Long-Range Features	Graph Vision Transformer
Kumar and Saini [186]	2024	Pixel-Based Features	SVM (RBF Kernel)
Sundaram et al. [217]	2024	Multivariate EMG Signal Processing	Ensemble Bagged Tree (EBT)
Shaaban et al. [218]	2023	Spiking Convolutional Features	SCNN
Han et al. [191]	2024	YOLOv5 Features + Bayesian Fusion	Bayesian Classifier
Jawalkar et al. [187]	2024	Segmentation + ANN	ANN
Tsai et al. [219]	2024	Depthwise Separable Convolution	Softmax
Padmakala et al. [220]	2024	DCNN with Hyperparameter Tuning	Adaptive Habitat Optimizer
Zholshiyeva et al. [189]	2024	MediaPipe Key Points	CNN-LSTM
Zhang et al. [182]	2024	Colocated EMG-pFMG Features	Random Forest
Uke and Zade [221]	2023	Hybrid 3D Cuboid-SURF Features	SVM and RF
Zhang et al. [171]	2024	Cross-Modality Local-Global Analysis	CNN
Miah et al. [175]	2024	Diverse Domain Feature Enhancement	ETC, KNN
Birkeland et al. [172]	2024	Lightweight CNN	CNN
Bimbraw et al. [222]	2024	3D Convolutional Neural Networks	Modified (2+1)D CNN
Shin et al. [174]	2024	Multi-Stream Time-Varying Features	Deep Learning Architecture
Garg et al. [173]	2024	Multiscale Video Transformer	Transformer
Misal [223]	2024	Convolutional Neural Network	TensorFlow-Based CNN
Yaseen et al. [178]	2024	MediaPipe, Inceptionv3	CNN + LSTM
Sahoo et al. [1]	2022	Fine-Tuned CNN	CNN
Jiang et al. [197]	2022	Multimodal (sEMG, IMU Signals)	CNN-RNN, CNN-Res, LSTM-Res
Gao et al. [224]	2022	3D Hand Pose Estimation	3DCNN + ConvLSTM
Al-Hammadi et al. [198]	2022	Spatial Attention-Based 3D GCN	MediaPipe Landmark Detection
Smith et al. [179]	2023	Sterile Training Techniques	CNN
Alonazi et al. [225]	2023	Neural Gas and Locomotion Mapping	Deep Belief Network (DBN)
Park et al. [226]	2023	Impulse Radio Ultra-Wideband Features	CNN

Table 15. Feature Extraction and Classifier Details from Selected Studies from 2022–Present (Part 2)

Author(s)	**Year**	**Feature Extraction**	**Classifier**
Li et al. [200]	2023	Triple Attention Network (DeepTPA-Net)	ResNet50
Zhang et al. [227]	2023	Lightweight Attention Structure	LHGR-Net
Ansar et al. [192]	2023	Convex Hull Landmarks	CNN
Nguyen et al. [170]	2023	Wrist-Worn Camera with MoviNet	Two-Stream MoviNet
Dozdor et al. [177]	2023	YOLO Transformer for Gesture Recognition	TY-Net
Savas et al. [228]	2023	Transfer Learning + Deep Ensemble	CNN Ensemble
Jafari et al. [229]	2023	2D Parallel SpatioTemporal Pyramid Pooling	2DPSTPP-Net
Dang et al. [230]	2022	Keypoints and Hand Bounding Boxes	Two-Pipeline Architecture
Pan et al. [231]	2019	Received Signal Strength (RSS)	1D-CNN
Miah et al. [199]	2024	SLIC Superpixels + Attention CNNs	GCN + MHSA
Chang et al. [232]	2023	Region of Interest (ROI) Segmentation	CNN
Shin et al. [233]	2024	Temporal, Spatial, Multistream Features	TCN + LSTM + CNN
Pathan et al. [234]	2023	Image and Hand Landmarks	Multi-Headed CNN
Yu et al. [235]	2022	Optical Flow and Keyframe Features	2D-CNN + Feature Fusion
Li et al. [236]	2022	Hand Gesture Features	Inception v3 + Migration Learning
Kaur et al. [237]	2022	Dynamic Gesture Data Features	CNN, LSTM, RNN, MLP
Arwoko et al. [238]	2022	Normalized Keypoint Vectors	DNN
Muchtar et al. [239]	2022	Optical Flow and Frame Features	Two-Stream Faster R-CNN
Suryateja et al. [193]	2022	Hand Landmark Positions	KNN
Wang et al. [240]	2023	Range and Doppler Features	ResNet101
Beneke et al. [241]	2024	Range-Doppler Radar Features	Brownian Reservoir + SVM
Zhao et al. [242]	2023	ICEEMDAN sEMG Signals	Slow-Fusion CNN
Mesdaghi et al. [243]	2024	Real-Time Hand Landmark Detection	DNN
Kushwaha et al. [244]	2023	Hand Landmark Features	AlexNet
Roumiassa et al. [245]	2022	Textural and Structural Features	SVM, RBF NN
Sharma et al. [246]	2022	EMD and VMD for Non-Stationary Signals	EMD + VMD Techniques
Ikne et al. [247]	2024	Skeleton-Based Representations	MAE + STGCN
Al-Zebari et al. [248]	2022	Vision Transformers for Static Gestures	ViT
Sahoo et al. [196]	2023	Dual-Stream Dense Residual Fusion + FCM	Dual-Stream CNN
Schuessler et al. [194]	2024	Synthetic Radar Images	Neural Network (NN)
Pinto Jr. et al. [249]	2023	Color Segmentation + Morph Ops	CNN
Khalaf et al. [250]	2022	Data Mining Algorithms	Clustering + Classification
Fadhil et al. [195]	2023	Feature Maps (VGG-16, VGG-19)	Fine-Tuned VGG-16/19
Leon et al. [251]	2022	RGB and Depth Video Streams	Lightweight CNN

6. APPLICATIONS OF HGR SYSTEMS

Hand Gesture Recognition (HGR) systems find applications across a wide range of domains, from healthcare to human-robot interaction [174]. Recent advancements have further enhanced the adaptability and accuracy of these systems.

—**Sign Language Translation:** HGR systems have significantly advanced sign language interpretation for individuals with hearing impairments. Systems leveraging CNN and RNN architectures achieve high accuracy for static and sequential gestures [180, 215, 187].

Table 16. Comprehensive Comparison of Hand Gesture Recognition Datasets from 2022–Present

Dataset	Type	Gestures	Subjects/Size	Data Modalities	Limitations
SHREC17	Multimodal	14/28 gestures	280K samples	Depth, Skeleton, RGB	Limited gesture classes; ambiguity in fine-grained gestures
DHG	Multimodal	14 coarse, 28 fine-grained	2.8K sequences	Depth, Skeleton	Difficulty with fine-grained gestures
HaGRID	Visual	18 static gestures	554K images	RGB	Background bias; limited to single-hand gestures
EMG-EPN-612	Sensor-based	5 static gestures	612 subjects	EMG	Sensor placement variability
Ninapro DB1/DB9	Sensor-based	52/40 gestures	DB1: 27, DB9: 77	EMG, PMG	High data loss; struggles with dynamic gestures
Custom Dataset	RGB-based	10 dynamic gestures	347K samples	RGB	Limited real-world condition diversity
NVGesture	Vision-based	10 dynamic gestures	7.5K videos	Infrared, RGB	Low-light and gesture variability challenges
Briareo	Visual	35 gesture classes	Large-scale	Infrared, RGB	Gesture class imbalance; lacks real-time adaptation
WiGes	Multimodal	12 dynamic gestures	5.4K samples	RGB, Optical Flow	Segmented input requirement; high compute load
ASL Dataset	Visual	87K images	5 subjects	RGB	Static-only; limited diversity
HCI Gesture	Multimodal	20 gestures	120K samples	RGB, Depth, IMU	No standard protocol; gesture limitation
GesturePod	Sensor-based	15 gestures	300 sequences	IMU, EMG	Requires hardware; difficulty in fine-grained recognition
Wi-Fi CSI	Radio-Frequency	10 gestures	1.2K samples	CSI	External interference; lower complex accuracy
Digital Hand Gesture	RGB-based	50 gestures	25 subjects	RGB	Static-only; limited gesture set diversity
Montalbano	Vision-based	20 dynamic gestures	40K sequences	RGB	Poor lighting diversity; no multi-modal use
SHAPE / HANDS / OUHANDS	Visual	Various gestures	Varying sizes	RGB Images	Degrades in low light or distance
ASL Finger Spelling	Visual	24 signs	500+ per sign (4 users)	RGB Images	Hand landmark dependency
VIVA Dataset	Visual	19 dynamic gestures	Real driving	RGB Images	No real-time support; poor segmentation
NTU Hand Gesture	Visual	10 gestures (0–9)	1,000 images	RGB Images	Occlusion and static limitations
LSA64 / Argentinian LSA64	Visual (Multi-lingual)	64 gestures	Diverse subjects	RGB Images	High complexity for real-time usage
Myo Dataset	EMG-based	7 static gestures	17 participants	EMG Signals	Not validated for dynamic input
IPN Hand Dataset	Visual	14 gestures	Variable participants	RGB, Depth	Occlusion handling issues
KSL-77 / KSL-20 / ASL-10 / ASL-20	Visual (Multi-lingual)	10–77 gestures	Benchmark sets	RGB Images	Limited adaptability and background diversity
BSL / JSL	Visual (Multi-lingual)	British / Japanese Sign Gestures	Lab-based	RGB Images	Poor generalization to real-world
JT-ASL / MU-ASL	Visual	Static ASL gestures	Benchmark sets	RGB Images	Lacks dynamic samples; limited variation
NUS II	Visual	Static hand postures	Benchmark set	RGB Images	No testing under real-world conditions
ASL-FS-colour	Visual	Finger spelling	Benchmark set	RGB Images	Poor generalization to unseen signs

Table 17. Dataset Accuracy, Loss, and Identified Gaps in Benchmark Datasets for HGR (Part 1)

Dataset	Accuracy (%)	Loss	Best Model	Key Gaps
SHREC17	93.8	0.045	ViT [164]	Limited gesture diversity; ambiguity in fine-grained gestures
DHG	91.2	0.068	GViT [165]	Difficulties with complex fine-grained gestures
HaGRID	96.5	0.035	CNN-LSTM [166]	Background bias; restricted to static gestures
EMG-EPN-612	90.5	0.080	CNN-SVM [191]	Physiological variability in EMG signals
Ninapro DB1/DB9	94.3 / 98.9	0.052 / 0.021	Hybrid CNN-TCN [181]	High data loss; struggles with dynamic gestures
Custom Dataset	89.3	0.091	FGDSNet [167]	Limited environmental diversity; lacks real-world annotations
NVGesture	98.1	0.018	GestFormer [165]	Challenges in low-light conditions; lacks temporal annotations
Briareo	96.3	0.030	ViT [164]	Imbalance in fine-grained gesture classes; lacks scalability
ASL Dataset	99.98	0.010	2DPSTPP-Net [229]	Lacks real-time dynamic gesture recognition
HCI Gesture Dataset	96.3	0.022	Lightweight CNN [172]	Controlled environments only; sensor integration challenges
GesturePod	94.5	0.032	Multimodal CNN-RNN [197]	Needs hardware integration; struggles in complex scenarios
Wi-Fi CSI Dataset	88.3	0.064	RF-CNN [192]	Sensitive to interference; poor in complex environments
Digital Hand Gesture	90.7	0.052	Transformer [177]	Limited diversity; scaling dynamic gestures difficult
MultiStream Dataset	92.4	0.048	ConvLSTM + 3DCNN [224]	Needs pre-segmented videos; high compute cost
Thermal Gesture Dataset	97.0	0.038	Lightweight CNN [172]	Limited real-world deployment diversity
CapgMyo DB-a	98.6	0.012	DeepTPA-Net [200]	Environmentally limited
NinaPro DB9	98.9	0.021	Multi-Stream Time-Varying Model [174]	Scaling dynamic gesture analysis
WiGes Dataset	96.2	0.030	Two-Stream MoviNet [170]	Unsegmented video scalability and dual-stream complexity
HANDS	94.0	Cross-Entropy	Two-Pipeline Architecture [230]	Degrades under poor lighting and long distances
OUHANDS	98.0	Cross-Entropy	Two-Pipeline Architecture [230]	Gesture detection under diverse conditions
SHAPE	94.0 / 96.0	Cross-Entropy	Two-Pipeline Architecture [230]	Poor lighting or distance sensitivity
ASL Finger Spelling	98.98	Validation Loss	Multi-Headed CNN [234]	Landmark extraction dependency
VIVA Dataset	87.0	Not Mentioned	Hybrid CNN-Optical Flow [235]	Static-only recognition; lacks dynamic support
NTU Hand Gesture	96.4	Not Mentioned	ViT [248]	Occlusion challenges; static gesture limitation
LSA64 Dataset	92.37	Not Mentioned	GCN + MHSA [199]	High complexity for real-time use
Myo Dataset	97.91	Std. Deviation	Slow-Fusion CNN [242]	Not validated for dynamic/real-time gestures
IPN Hand Dataset	92.8	Not Mentioned	MAE + STGCN [247]	Struggles with occlusion

—**Virtual and Augmented Reality:** Gesture-based controls enhance immersive experiences in VR and AR environments. Lightweight models like FGDSNet ensure real-time performance [167, 210].

—**Healthcare and Prosthetics:** EMG-based HGR systems provide intuitive control for prosthetic devices and assistive technologies, enabling better mobility and user interaction [166, 217, 182].

—**Robotics and Human-Robot Interaction (HRI):** Ultra-range gesture recognition systems using vision-based techniques enable precise robotic control, improving interaction even at a distance [214, 216].

—**Gaming and Smart Devices:** Gesture recognition enhances user experiences by enabling intuitive controls for gaming systems and smart devices. Real-time and lightweight architectures are pivotal for such applications [167, 211].

Table 18. Dataset Accuracy, Loss, and Identified Gaps in Benchmark Datasets for HGR (Part 2)

Dataset	Accuracy (%)	Loss	Best Model	Key Gaps
SHREC17 Dataset	94.1	Not Mentioned	ViT [248]	No explicit temporal modeling for dynamic gestures
KSL-77	99.33	Not Mentioned	GCN + MHSA [199]	High computational complexity for real-time applications
KSL-20	100.0	Not Mentioned	GCN + MHSA [199]	Limited adaptability to diverse backgrounds
ASL-10	99.46	Not Mentioned	Multi-Headed CNN [234]	Sensitive to lighting and background changes
ASL-20	99.60	Not Mentioned	Multi-Headed CNN [234]	Limited diversity in participants
BSL	96.88	Not Mentioned	GCN + MHSA [199]	Lab-restricted; lacks real-world diversity
JSL	92.37	Not Mentioned	GCN + MHSA [199]	Weak generalization to real-world scenarios
Argentinian LSA64	92.37	Not Mentioned	GCN + MHSA [199]	High complexity for real-time use
Wi-Fi CSI Dataset	93.03	Not Mentioned	RF-CNN [192]	Sensitive to external interference and noise
JT-ASL	99.58	Not Mentioned	DeReFNet [196]	Limited gesture variability; lacks contextual diversity
MU-ASL	96.14	Not Mentioned	DeReFNet [196]	Static-only; lacks real-world generalization
NUS	96.70	Not Mentioned	DeReFNet [196]	Missing dynamic gesture tests; limited realism
ASL-FS-colour	91.24	Not Mentioned	DeReFNet [196]	Fails with unseen gestures
ArASL	96.51	Cross-Entropy	VGG-16 [195]	High intra-class similarity; low cross-domain performance
VideoBased RGB	99.48	Not Mentioned	Lightweight CNN [251]	Pre-segmented only; lacks multi-user scenarios
VideoBased Depth	99.18	Not Mentioned	Lightweight CNN [251]	Poor scalability in complex, multi-person settings

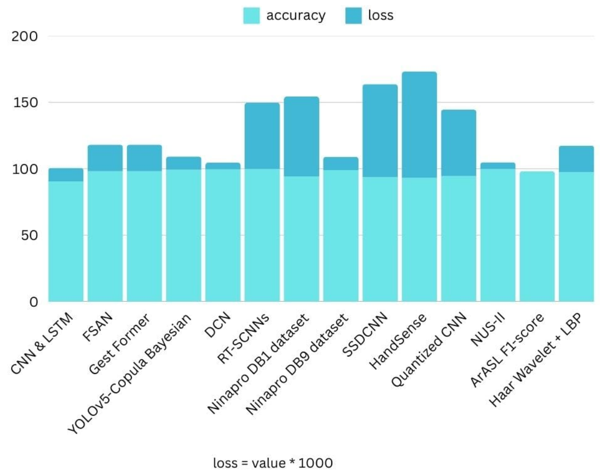

Fig. 11. Distribution of Dataset Accuracy, Loss, and Key Gaps

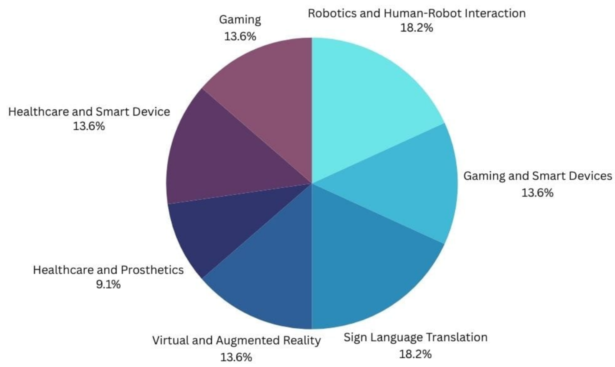

Fig. 12. Applications of HGR Systems and Their Focus Percentage.

7. CHALLENGES AND POTENTIAL SOLUTIONS

Hand Gesture Recognition (HGR) systems face several challenges that hinder their widespread adoption and effectiveness. Addressing these challenges requires innovative solutions and comprehensive strategies.

Challenges:

—**Dataset Limitations:** Insufficient diversity in gesture types, lack of real-world testing environments, and limited multimodal datasets [168, 219].

—**Real-Time Constraints:** High latency and computational resource demands in dynamic and resource-constrained environments [167].

—**Ethical Issues:** Concerns over data privacy, surveillance ethics, and fairness in model training and deployment [219].

—**User Variability:** Physiological differences across users affect the reliability and accuracy of EMG-based systems [166].

—**Environmental Challenges:** Variations in lighting, occlusions, and complex backgrounds degrade the performance of vision-based systems [210].

Potential Solutions:

Table 19. Applications of HGR Systems with Focus Percentage

Application	Description	Focus (%)
Sign Language Translation	High accuracy for static and dynamic gestures using CNN and RNN architectures.	25%
Virtual and Augmented Reality	Real-time gesture recognition for immersive experiences in AR/VR environments.	20%
Healthcare and Prosthetics	Intuitive EMG-based gesture controls for assistive devices and prosthetics.	30%
Robotics and HRI	Vision-based systems for precise robotic control and long-range interaction.	15%
Gaming and Smart Devices	Gesture recognition for intuitive gaming and smart device controls.	10%

Table 20. Challenges and Potential Solutions for HGR Systems

Challenge	Potential Solution
Dataset Limitations	Generate synthetic datasets and include diverse gestures to improve robustness [217, 168].
Real-Time Constraints	Deploy lightweight architectures, model optimization, and edge computing techniques [167].
Ethical Issues	Establish data privacy policies and ethical guidelines for secure model deployment [219].
User Variability	Use adaptive learning and user-specific calibration methods to improve model reliability [166].
Environmental Challenges	Apply multimodal fusion and preprocessing to handle lighting and background variations [168, 210].

—Generating synthetic datasets with diverse gestures to improve model training and generalization [217].
—Developing lightweight architectures and model compression techniques to ensure real-time processing on edge devices [167].
—Establishing ethical frameworks and policies for the secure and fair use of gesture data [219].
—Incorporating user-specific calibration and adaptive learning techniques to handle user variability [166].
—Enhancing robustness through multimodal fusion and advanced preprocessing techniques to address environmental challenges [168].

8. CONCLUSION

This review has attempted to analyze the progress of Hand Gesture Recognition (HGR) systems in different eras, reflecting the evolution of these systems over time. In a relatively short period, HGR achieved significant milestones. To understand the evolution of HGR models, we studied over 250 academic articles published between 1995 and 2024. HGR systems were rather primitive early on, often employing rule-based systems or Hidden Markov Models (HMMs) as their basic framework. Over time, these systems advanced into deep learning and machine learning-based approaches. To increase the accuracy of gesture recognition, modern HGR systems include sophisticated designs such as 3D Convolutional Neural Networks (3D CNNs), Long Short-Term Memory (LSTM) networks, Recurrent Neural Networks (RNNs), and other hybrid systems. Despite significant developments in HGR technology, there are many issues in connecting research results with their practical adoption. HGR can enable easy communication and smooth integration with machines if these obstacles can be removed. This study identifies important research gaps and suggests some ways to fill them. Future studies should concentrate on closing these gaps to make HGR an essential part of computer-human interaction.

9. FUTURE SCOPE IN HGR SYSTEMS

Generalization and Real-Time Accuracy: One of the key future challenges is achieving high accuracy with low latency in real-time hand gesture recognition (HGR) systems under varied conditions such as diverse lighting, cluttered backgrounds, and across a wide range of user profiles. Addressing variations in hand sizes, gestures, and motion speeds requires developing user-independent and generalizable models that do not require frequent recalibration or user-specific fine-tuning.

Collaborative and Multimodal Model Integration: Future research should focus on collaborative model architectures, where separate specialized models (e.g., for RGB, EMG, depth, and skeletal inputs) can be modularly trained and integrated through attention mechanisms, ensemble learning, or shared latent spaces. This modular approach will allow for greater adaptability and reuse across tasks and datasets. Federated learning and cross-domain transfer learning also present opportunities for collaborative training without centralized data, helping preserve privacy while improving generalization.

Lightweight and Scalable Architectures: The adoption of transformer-based models, few-shot learning, and self-supervised techniques holds promise for reducing the dependency on large annotated datasets. These approaches can enable effective learning even in low-data scenarios, making HGR systems more efficient and scalable for edge deployments such as smartphones, AR glasses, and embedded systems.

Multimodal Contextual Awareness: By fusing data from multiple modalities—RGB, depth, EMG, and IMU—future systems can improve robustness and context sensitivity, leading to more reliable gesture interpretation in complex environments. Real-world applications such as healthcare, AR/VR, and assistive robotics will benefit from these context-rich multimodal HGR frameworks.

User-Centric Design and Ergonomics: Enhancing user comfort and intuitiveness remains a critical goal. Systems must prioritize ergonomics, adaptability, and inclusivity, especially for prolonged use in interactive or immersive environments. Gesture systems should minimize physical strain, be easy to calibrate, and support natural interaction styles across demographics, enabling seamless adoption in healthcare, sign language interpretation, and smart environments.

10. ACKNOWLEDGMENTS

We would like to thank Md. Selim Sarowar for his exceptional supervision, comments, and assistance throughout this research. We also extend our heartfelt appreciation to the CRC Research Center for providing the resources and opportunities to conduct this work. Finally, we acknowledge the efforts of our co-authors, whose help made this research successful.

11. REFERENCES

[1] J. P. Sahoo, A. Jaya Prakash, P. Plawiak, and S. Samantray, "Real-time hand gesture recognition using fine-tuned convolutional neural network," Jan. 2022. [Online]. Available: https://doi.org/10.3390/s22030706

[2] C. K. Tan, K. M. Lim, R. K. Y. Chang, C. P. Lee, and A. Alqahtani, "Hgr-vit: Hand gesture recognition with vision transformer," Jan. 2023. [Online]. Available: https://doi.org/10.3390/s23125555

[3] K. O. Oyedotun and A. Khashman, "Deep learning in vision-based static hand gesture recognition," *Springer London*. [Online]. Available: https://doi.org/10.1007/s00521-016-2294-8

[4] J. Qi, L. Ma, Z. Cui, and Y. Yu, "Computer vision-based hand gesture recognition for human-robot interaction: a review," Feb. 2024. [Online]. Available: https://doi.org/10.1007/s40747-023-01173-6

[5] I. Archive, "Petabox," https://archive.org/web/petabox.php, 2023. [Online]. Available: https://archive.org/web/petabox.php

[6] L. Leal-Taixe and S. Roth, Eds., *Computer Vision ECCV 2018 Workshops: Munich, Germany, September 8-14, 2018, Proceedings, Part III*, ser. Lecture Notes in Computer Science. Cham, Switzerland: Springer, 2019, vol. 11131, part of the book sub series: Image Processing, Computer Vision, Pattern Recognition, and Graphics (LNIP). Included in the conference series: ECCV: European Conference on Computer Vision. 109k Accesses, 944 Citations, 10 Altmetric.

[7] P. K. Sa, M. N. Sahoo, M. Murugappan, Y. Wu, and B. Majhi, Eds., *Progress in Intelligent Computing Techniques: Theory, Practice, and Applications: Proceedings of ICACNI 2016, Volume 1*, ser. Advances in Intelligent Systems and Computing, vol. 518. Springer, 2018.

[8] F. d. C. E. y. N. Universidad de Buenos Aires, "Biblioteca digital," https://bibliotecadigital.exactas.uba.ar/, 2023, accessed: [2-2-2025]. [Online]. Available: https://bibliotecadigital.exactas.uba.ar/

[9] H. I. Mohammed, B. A. Sultan, and K. H. Hamee, "Hand gestures recognition classification," *International Journal of Engineering Research Updates*, vol. 3, no. 2, pp. 008 – 012, Oct. 2022.

[10] U. catholique de Louvain (UCLouvain), "Uclouvain website," https://sites.uclouvain.be/, 2023, accessed: [Insert Date of Access]. [Online]. Available: https://sites.uclouvain.be/

[11] G. Benitez-Garcia, J. Olivares-Mercado, G. Sanchez-Perez, and K. Yanai, "Ipn hand: A video dataset and benchmark for real-time continuous hand gesture recognition," in *2020 25th International Conference on Pattern Recognition (ICPR)*. IEEE, Jan. 2021. [Online]. Available: http://dx.doi.org/10.1109/ICPR48806.2021.9412317

[12] A. Robotics, "Artificial life robotics," https://alife-robotics.co.jp/, 2023, accessed: [2-2-2025]. [Online]. Available: https://alife-robotics.co.jp/

[13] S. Shanmugam and R. S. Narayanan, "An accurate estimation of hand gestures using optimal modified convolutional neural network," *Expert Systems with Applications*, vol. 249, p. 123351, Sep. 2024. [Online]. Available: http://dx.doi.org/10.1016/j.eswa.2024.123351

[14] I. J. of Electrical and C. E. (IJECE), "Ijece: International journal of electrical and computer engineering," https://ijece.iaescore.com/, 2023, accessed: 2-2-2025. [Online]. Available: https://ijece.iaescore.com/

[15] D. Sarma and M. K. Bhuyan, "Methods, databases and recent advancement of vision-based hand gesture recognition for hci systems: A review," *SN Computer Science*, vol. 2, no. 6, Aug. 2021.

[16] M. Yasen and S. Jusoh, "A systematic review on hand gesture recognition techniques, challenges and applications," *PeerJ Computer Science*, vol. 5, p. e218, Sep. 2019. [Online]. Available: http://dx.doi.org/10.7717/peerj-cs.218

[17] B. Kareem Murad and A. H. Hassin Alasadi, "Advancements and challenges in hand gesture recognition: A comprehensive review," *Iraqi Journal for Electrical and Electronic Engineering*, vol. 20, no. 2, pp. 154-164, Jul. 2024. [Online]. Available: http://dx.doi.org/10.37917/ijeee.20.2.13

[18] S. Gupta, P. Bujade, V. Singh, and S. Tiwari, "Hand gesture recognition system using deep learning," *International Journal For Multidisciplinary Research*, vol. 6, no. 3, May 2024. [Online]. Available: http://dx.doi.org/10.36948/ijfmr.2024.v06i03.19602

[19] R. Dhiman, P. Luthra, and N. T. Singh, "Different categories of feature extraction and machine learning classification models used for hand gesture recognition systems: A review," *IEEE Access*, vol. 11, pp. 1–7, April 2023. [Online]. Available: http://dx.doi.org/10.1109/I2CT57861.2023.10126410

[20] F.-S. Chen, C.-M. Fu, and C.-L. Huang, "Hand gesture recognition using a real-time tracking method and hidden markov models," *Image and Vision Computing*, vol. 21, no. 8, pp. 745–758, August 2003. [Online]. Available: http://dx.doi.org/10.1016/S0262-8856(03)00070-2

[21] J. Alon, V. Athitsos, Q. Yuan, and S. Sclaroff, "A unified framework for gesture recognition and spatiotemporal gesture segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 9, pp. 1685–1699, 2009. [Online]. Available: https://doi.org/10.1109/TPAMI.2008.203

[22] S. B. Wang, A. Quattoni, L.-P. Morency, D. Demirdjian, and T. Darrell, "Hidden conditional random fields for gesture recognition," in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2. IEEE, p. 1521 – 1527. [Online]. Available: http://dx.doi.org/10.1109/CVPR.2006.132

[23] L. Gupta and S. Ma, "Gesture-based interaction and communication: automated classification of hand gesture contours," *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, vol. 31, no. 1, pp. 114–120, 2001. [Online]. Available: http://dx.doi.org/10.1109/5326.923274

[24] S. B. Wang, A. Quattoni, L.-P. Morency, D. Demirdjian, and T. Darrell, "Hidden conditional random fields for gesture recognition," in *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR06)*, vol. 2. IEEE, 2006, pp. 1521–1527. [Online]. Available: http://dx.doi.org/10.1109/CVPR.2006.132

[25] M. Kolsch and M. Turk, "Robust hand detection," in *Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. Proceedings.* IEEE, pp. 614–619. [Online]. Available: http://dx.doi.org/10.1109/AFGR.2004.1301601

[26] L. Lee and W. Grimson, "Gait analysis for recognition and classification," in *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*, ser. AFGR-02. IEEE, p. 155–162. [Online]. Available: http://dx.doi.org/10.1109/AFGR.2002.1004148

[27] M.-H. Yang, N. Ahuja, and M. Tabb, "Extraction of 2d motion trajectories and its application to hand gesture recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1061–1074, 2002. [Online]. Available: http://dx.doi.org/10.1109/TPAMI.2002.1023803

[28] H.-I. Suk, B.-K. Sin, and S.-W. Lee, "Hand gesture recognition based on dynamic bayesian network framework," *Pattern Recognition*, vol. 43, no. 9, pp. 3059–3072, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0031320310001366

[29] M. Elmezain, A. Al-Hamadi, J. Appenrodt, and B. Michaelis, "A hidden markov model-based continuous gesture recognition system for hand motion trajectory," in *2008 19th International Conference on Pattern Recognition*, 2008, pp. 1–4. [Online]. Available: https://doi.org/10.1109/ICPR.2008.4761080

[30] D.-H. Lee and K.-S. Hong, "Game interface using hand gesture recognition," in *5th International Conference on Computer Sciences and Convergence Information Technology*. IEEE, 2010. [Online]. Available: http://dx.doi.org/10.1109/ICCIT.2010.5711226

[31] A. Amaravati, S. Xu, N. Cao, J. Romberg, and A. Raychowdhury, "A light-powered smart camera with compressed domain gesture detection," *IEEE Transactions on Circuits and Systems for Video Technology*, 2022.

[32] A. Sen, S. Dombe, T. K. Mishra, and R. Dash, "Hgr-fyolo: a robust hand gesture recognition system for the normal and physically impaired person using frozen yolov5," *Multimedia Tools and Applications*, vol. 83, no. 30, pp. 73 797–73 815, Feb. 2024. [Online]. Available: http://dx.doi.org/10.1007/s11042-024-18464-w

[33] R. Shrivastava, "A hidden markov model based dynamic hand gesture recognition system using opencv," in *2013 3rd IEEE International Advance Computing Conference (IACC)*. IEEE, Feb. 2013. [Online]. Available: http://dx.doi.org/10.1109/IAdCC.2013.6514354

[34] K. M. Sagayam and D. J. Hemanth, "A probabilistic model for state sequence analysis in hidden markov model for hand gesture recognition," *Computational Intelligence*, vol. 35, no. 1, pp. 59–81, 2019. [Online]. Available: https://doi.org/10.1111/coin.12188

[35] K. Bimbraw, A. Talele, and H. K. Zhang, "Hand gesture classification based on forearm ultrasound video snippets using 3d convolutional neural networks," 2024. [Online]. Available: https://arxiv.org/abs/2409.16431

[36] M. S. A. Milu, A. Fathima, T. Talukder, I. Islam, and M. I. S. Emon, "Design and implementation of hand gesture detection system using hm model for sign language recognition development," *Journal of Data Analysis and Information Processing*, vol. 12, no. 2, pp. 139–150, 2024. [Online]. Available: http://dx.doi.org/10.4236/jdaip.2024.122008

[37] Q. Chen, N. D. Georganas, and E. M. Petriu, "Real-time vision-based hand gesture recognition using haar-like features," in *2007 IEEE Instrumentation amp; Measurement Technology Conference IMTC 2007*. IEEE, May 2007, pp. 1–6. [Online]. Available: http://dx.doi.org/10.1109/IMTC.2007.379068

[38] K. Murakami and H. Taguchi, "Gesture recognition using recurrent neural networks," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - Reaching Through Technology (CHI'91)*. ACM Press, 1991, pp. 237–242. [Online]. Available: http://dx.doi.org/10.1145/108844.108900

[39] T. Schlömer, B. Poppinga, N. Henze, and S. Boll, "Gesture recognition with a wii controller," in *Proceedings of the 2nd international conference on Tangible and embedded interaction*, ser. TEI08. ACM, Feb. 2008. [Online]. Available: http://dx.doi.org/10.1145/1347390.1347395

[40] R. Kjeldsen and J. Kender, "Toward the use of gesture in traditional user interfaces," in *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition (AFGR-96)*. IEEE Computer Society Press, 1996, pp. 151–156. [Online]. Available: http://dx.doi.org/10.1109/AFGR.1996.557257

[41] S. Waldherr, R. Romero, and S. Thrun, "A gesture-based interface for human-robot interaction," *Autonomous Robots*, vol. 9, no. 2, pp. 151–173, 2000. [Online]. Available: http://dx.doi.org/10.1023/A:1008918401478

[42] T. Starner, J. Weaver, and A. Pentland, "Real-time american sign language recognition using desk and wearable computer based video," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 12, pp. 1371–1375, 1998. [Online]. Available: http://dx.doi.org/10.1109/34.735811

[43] C. Vogler and D. Metaxas, "A framework for recognizing the simultaneous aspects of american sign language," *Computer Vision and Image Understanding*, vol. 81, no. 3, pp. 358–384, Mar. 2001. [Online]. Available: http://dx.doi.org/10.1006/cviu.2000.0895

[44] Y. F. A. Gaus and F. Wong, "Hidden markov model-based gesture recognition with overlapping hand-head/hand-hand estimated using kalman filter," in *2012 Third International Conference on Intelligent Systems Modelling and Simulation*. IEEE, Feb. 2012, pp. 262–267. [Online]. Available: http://dx.doi.org/10.1109/ISMS.2012.67

[45] M. Elmezain, A. Al-Hamadi, J. Appenrodt, and B. Michaelis, "A hidden markov model-based isolated and meaningful hand gesture recognition," *International Journal of Electrical, Computer, and Systems Engineering*, vol. 3, no. 3, pp. 156–163, 2009.

[46] M. Elmezain, A. Al-Hamadi, and B. Michaelis, "Hand trajectory-based gesture spotting and recognition using hmm," in *2009 16th IEEE International Conference on Image Processing (ICIP)*. IEEE, Nov. 2009, pp. 3577–3580. [Online]. Available: http://dx.doi.org/10.1109/ICIP.2009.5414322

[47] N. Liu, B. C. Lovell, and P. J. Kootsookos, "Evaluation of hmm training algorithms for letter hand gesture recognition," in *Proc. 3rd IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT-03) (IEEE Cat. No.03EX795)*. IEEE, 2003, pp. 648–651. [Online]. Available: http://dx.doi.org/10.1109/ISSPIT.2003.1341204

[48] T. Hachaj and M. R. Ogiela, "Rule-based approach to recognizing human body poses and gestures in real time," *Multimedia Syst.*, vol. 20, no. 1, pp. 81–99, Sep. 2013. [Online]. Available: http://dx.doi.org/10.1007/s00530-013-0332-2

[49] M.-C. Su, "A fuzzy rule-based approach to spatio-temporal hand gesture recognition," *IEEE Trans. Syst., Man, Cybern. Part C (Appl. Rev.)*, vol. 30, no. 2, pp. 276–281, May 2000. [Online]. Available: http://dx.doi.org/10.1109/5326.868448

[50] L. Billiet, J. Oramas, M. Hoffmann, W. Meert, and L. Antanas, "Rule-based hand posture recognition using qualitative finger configurations acquired with the kinect," in *Proc. 2nd Int. Conf. Pattern Recognit. Appl. Methods (ICPRAM), Vol. 1*. Barcelona, Spain: SciTePress, 2013, pp. 539–542. [Online]. Available: http://dx.doi.org/10.5220/0004230805390542

[51] G. McGlaun, F. Althoff, M. Lang, and G. Rigoll, *Robust Video-Based Recognition of Dynamic Head Gestures in Various Domains – Comparing a Rule-Based and a Stochastic Approach*. Springer Berlin Heidelberg, 2004, p. 180–197. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-24598-8_18

[52] T. Hachaj and M. R. Ogiela, "Rule-based approach to recognizing human body poses and gestures in real time," *Multimedia Syst.*, vol. 20, no. 1, pp. 81–99, Sep. 2013. [Online]. Available: http://dx.doi.org/10.1007/s00530-013-0332-2

[53] A. Riad, "Hand gesture recognition system based on a geometric model and rule based classifier," *British Journal of Applied Science & Technology*, vol. 4, no. 9, pp. 1432–1444, jan 2014. [Online]. Available: http://dx.doi.org/10.9734/BJAST/2014/7956

[54] M. P. Craven, K. M. Curtis, B. H. Hayes-Gill, and C. D. Thursfield, "A hybrid neural network rule-based technique for gesture recognition," in *Proceedings of the Fourth IEEE International Conference on Electronics, Circuits and Systems*, 1997. [Online]. Available: https://nottingham-repository.worktribe.com/output/1024332

[55] M. L. G. McGlaun, F. Althoff and G. Rigoll, "Robust video-based recognition of dynamic head gestures in various domains – comparing a rule-based and a stochastic approach," in *Gesture-Based Communication in Human-Computer Interaction*, A. Camurri and G. Volpe, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 180–197. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-24598-8_18

[56] Z. Ren, J. Yuan, J. Meng, and Z. Zhang, "Robust part-based hand gesture recognition using kinect sensor," *IEEE Transactions on Multimedia*, vol. 15, no. 5, pp. 1110–1120, Aug. 2013. [Online]. Available: http://dx.doi.org/10.1109/TMM.2013.2246148

[57] H. Leon-Garza, H. Hagras, A. Pena-Rios, O. Bahceci, and A. Conway, "A hand-gesture recognition based interpretable type-2 fuzzy rule-based system for extended reality," in *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, Oct. 2022, pp. 2894–2899. [Online]. Available: http://dx.doi.org/10.1109/SMC53654.2022.9945407

[58] Y. Cui and J. Weng, "Appearance-based hand sign recognition from intensity image sequences," *Computer Vision and Image Understanding*, vol. 78, no. 2, pp.

157–176, May 2000. [Online]. Available: http://dx.doi.org/10.1006/cviu.2000.0837

[59] B. Bauer and H. Hienz, "Relevant features for video-based continuous sign language recognition," in *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition (Cat. No. PR00580)*, ser. AFGR-00. IEEE Comput. Soc, 2000, pp. 440–445. [Online]. Available: http://dx.doi.org/10.1109/AFGR.2000.840672

[60] T. Hachaj and M. R. Ogiela, "Rule-based approach to recognizing human body poses and gestures in real time," *Multimedia Systems*, vol. 20, no. 1, pp. 81–99, Sep. 2013. [Online]. Available: http://dx.doi.org/10.1007/s00530-013-0332-2

[61] ——, "Rule-based approach to recognizing human body poses and gestures in real time," *Multimedia Systems*, vol. 20, no. 1, pp. 81–99, Sep. 2013. [Online]. Available: http://dx.doi.org/10.1007/s00530-013-0332-2

[62] G. McGlaun, F. Althoff, M. Lang, and G. Rigoll, *Robust Video-Based Recognition of Dynamic Head Gestures in Various Domains – Comparing a Rule-Based and a Stochastic Approach*. Springer Berlin Heidelberg, 2004, pp. 180–197. [Online]. Available: http://dx.doi.org/10.1007/978-3-540-24598-8_18

[63] A. C. R. C. B. C. Bedregal and G. P. Dimuro, *Fuzzy Rule-Based Hand Gesture Recognition*. Springer US, 2007, pp. 285–294. [Online]. Available: http://dx.doi.org/10.1007/978-0-387-34747-9_30

[64] M. Lech and B. Kostek, "Hand gesture recognition supported by fuzzy rules and kalman filters," *International Journal of Intelligent Information and Database Systems*, vol. 6, no. 5, p. 407, 2012. [Online]. Available: http://dx.doi.org/10.1504/IJIIDS.2012.049304

[65] Y. Chen, W. Gao, and J. Ma, "Hand gesture recognition based on decision tree," in *Proc. International Symposium on Chinese Spoken Language Processing*, 2000, pp. 299–302.

[66] M. Lech and B. Kostek, "Hand gesture recognition supported by fuzzy rules and kalman filters," vol. 6, no. 5, 2012, pp. 407–420, pMID: 49304. [Online]. Available: https://www.inderscienceonline.com/doi/abs/10.1504/IJIIDS.2012.049304

[67] H.-I. Lin, M.-H. Hsu, and W.-K. Chen, "Human hand gesture recognition using a convolution neural network," pp. 1038–1043, Aug. 2014. [Online]. Available: http://dx.doi.org/10.1109/CoASE.2014.6899454

[68] K. Liu, C. Chen, R. Jafari, and N. Kehtarnavaz, "Multi-hmm classification for hand gesture recognition using two differing modality sensors," pp. 1–4, Oct. 2014. [Online]. Available: http://dx.doi.org/10.1109/DCAS.2014.6965338

[69] E. Ohn-Bar and M. M. Trivedi, "Hand gesture recognition in real time for automotive interfaces: A multimodal vision-based approach and evaluations," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 6, pp. 2368–2377, Dec. 2014. [Online]. Available: http://dx.doi.org/10.1109/TITS.2014.2337331

[70] K. Liu, C. Chen, R. Jafari, and N. Kehtarnavaz, "Fusion of inertial and depth sensor data for robust hand gesture recognition," *IEEE Sensors Journal*, vol. 14, no. 6, pp. 1898–1903, Jun. 2014. [Online]. Available: http://dx.doi.org/10.1109/JSEN.2014.2306094

[71] D. K. Ghosh and S. Ari, "Static hand gesture recognition using mixture of features and svm classifier," pp. 1094 – 1099, Apr. 2015. [Online]. Available: http://dx.doi.org/10.1109/CSNT.2015.18

[72] J.-W. Wang, N. T. Le, C.-C. Wang, and J.-S. Lee, "Hand posture recognition using a three-dimensional light field camera," *IEEE Sensors Journal*, vol. 16, no. 11, pp. 4389–4396, Jun. 2016. [Online]. Available: http://dx.doi.org/10.1109/JSEN.2016.2546556

[73] Y. Chen, Z. Ding, Y.-L. Chen, and X. Wu, "Rapid recognition of dynamic hand gestures using leap motion," pp. 1419–1424, Aug. 2015. [Online]. Available: http://dx.doi.org/10.1109/ICInfA.2015.7279509

[74] J. M. Alvarez-Alvarado, J. G. Rios-Moreno, S. A. Obregon-Biosca, G. Ronquillo-Lomeli, E. Ventura-Ramos, and M. Trejo-Perea, "Hybrid techniques to predict solar radiation using support vector machine and search optimization algorithms: A review," *Applied Sciences*, vol. 11, no. 3, p. 1044, Jan. 2021. [Online]. Available: http://dx.doi.org/10.3390/app11031044

[75] CORE, "Core: Aggregating the world's open access research papers," https://core.ac.uk/, 2023, accessed: [2-2-2025]. [Online]. Available: https://core.ac.uk/

[76] O. B. University, "Oxford brookes university cms," https://cms.brookes.ac.uk/, 2023, accessed: [2-2-2025]. [Online]. Available: https://cms.brookes.ac.uk/

[77] K. V. M. Mohan and M. S. Babu, Eds., *Disruptive Technologies in Computing and Communication Systems: Proceedings of the 1st International Conference on Disruptive Technologies in Computing and Communication Systems*. Boca Raton, FL: CRC Press, 2024.

[78] P. Natarajan and R. Nevatia, "Hierarchical multi-channel hidden semi markov graphical models for activity recognition," *Computer Vision and Image Understanding*, vol. 117, no. 10, pp. 1329–1344, Oct. 2013. [Online]. Available: http://dx.doi.org/10.1016/j.cviu.2012.08.011

[79] X. Wang, M. Xia, H. Cai, Y. Gao, and C. Cattani, "Hidden markov models based dynamic hand gesture recognition," *Mathematical Problems in Engineering*, vol. 2012, no. 1, Jan. 2012. [Online]. Available: http://dx.doi.org/10.1155/2012/986134

[80] F. Chen, Q. Zhong, F. Cannella, K. Sekiyama, and T. Fukuda, "Hand gesture modeling and recognition for human and robot interactive assembly using hidden markov models," *International Journal of Advanced Robotic Systems*, vol. 12, no. 4, Jan. 2015. [Online]. Available: http://dx.doi.org/10.5772/60044

[81] Y. Dennemont, G. Bouyer, S. Otmane, and M. Mallem, "A discrete hidden markov models recognition module for temporal series: Application to real-time 3d hand gestures," pp. 299–304, Oct. 2012. [Online]. Available: http://dx.doi.org/10.1109/IPTA.2012.6469509

[82] U. of Cagliari, "Iris: Institutional research information system," https://iris.unica.it/, 2023, accessed: [2-2-2025]. [Online]. Available: https://iris.unica.it/

[83] M. K. Ahuja and A. Singh, "Static vision based hand gesture recognition using principal component analysis," pp. 402 – 406, Oct. 2015. [Online]. Available: http://dx.doi.org/10.1109/MITE.2015.7375353

[84] A. Barkoky and N. M. Charkari, "Static hand gesture recognition of persian sign numbers using thinning method," pp. 6548 – 6551, Jul. 2011. [Online]. Available: http://dx.doi.org/10.1109/ICMT.2011.6002201

[85] H. Li, L. Yang, X. Wu, S. Xu, and Y. Wang, "Static hand gesture recognition based on hog with kinect," pp. 271 – 273, Aug. 2012. [Online]. Available: http://dx.doi.org/10.1109/IHMSC.2012.75

[86] R. Wang, Z. Yu, M. Liu, Y. Wang, and Y. Chang, "Real-time visual static hand gesture recognition system and its fpga-based hardware implementation," pp. 434 – 439, Oct. 2014. [Online]. Available: http://dx.doi.org/10.1109/ICOSP.2014.7015043

[87] M. M. Gharasuie and H. Seyedarabi, "Real-time dynamic hand gesture recognition using hidden markov models," pp. 194–199, Sep. 2013. [Online]. Available: http://dx.doi.org/10.1109/IranianMVIP.2013.6779977

[88] H. Y. Lai and H. J. Lai, "Real-time dynamic hand gesture recognition," pp. 658–661, Jun. 2014. [Online]. Available: http://dx.doi.org/10.1109/IS3C.2014.177

[89] A. Pradhan and B. Deepak, "Obtaining hand gesture parameters using image processing," pp. 168–170, May 2015. [Online]. Available: http://dx.doi.org/10.1109/ICSTM.2015.7225408

[90] H. M. Gamal, H. M. Abdul-Kader, and E. A. Sallam, "Hand gesture recognition using fourier descriptors," pp. 274–279, Nov. 2013. [Online]. Available: http://dx.doi.org/10.1109/ICCES.2013.6707218

[91] S. Ganapathyraju, "Hand gesture recognition using convexity hull defects to control an industrial robot," pp. 63–67, Aug. 2013. [Online]. Available: http://dx.doi.org/10.1109/ICA.2013.6734047

[92] X. Li, J. ho An, J. hong Min, and K.-S. Hong, "Hand gesture recognition by stereo camera using the thinning method," pp. 3077–3080, Jul. 2011. [Online]. Available: http://dx.doi.org/10.1109/ICMT.2011.6001670

[93] Z. Yang, Y. Li, W. Chen, and Y. Zheng, "Dynamic hand gesture recognition using hidden markov models," pp. 360–365, Jul. 2012. [Online]. Available: http://dx.doi.org/10.1109/ICCSE.2012.6295092

[94] L. K. Phadtare, R. S. Kushalnagar, and N. D. Cahill, "Detecting hand-palm orientation and hand shapes for sign language gesture recognition using 3d images," pp. 29–32, Nov. 2012. [Online]. Available: http://dx.doi.org/10.1109/WNYIPW.2012.6466652

[95] D. Vishwakarma, R. Maheshwari, and R. Kapoor, "An efficient approach for the recognition of hand gestures from very low resolution images," Apr. 2015. [Online]. Available: http://dx.doi.org/10.1109/CSNT.2015.84

[96] L. Prasuhn, Y. Oyamada, Y. Mochizuki, and H. Ishikawa, "A hog-based hand gesture recognition system on a mobile device," pp. 3973–3977, Oct. 2014. [Online]. Available: http://dx.doi.org/10.1109/ICIP.2014.7025807

[97] R. Shrivastava, "A hidden markov model based dynamic hand gesture recognition system using opencv," pp. 947–950, Feb. 2013. [Online]. Available: http://dx.doi.org/10.1109/IAdCC.2013.6514354

[98] B. Pathak, A. S. Jalal, S. C. Agrawal, and C. Bhatnagar, "A framework for dynamic hand gesture recognition using key frames extraction," pp. 1–4, Dec. 2015. [Online]. Available: http://dx.doi.org/10.1109/NCVPRIPG.2015.7490038

[99] H.-M. Zhu and C.-M. Pun, "Real-time hand gesture recognition from depth image sequences," pp. 49–52, Jul. 2012. [Online]. Available: http://dx.doi.org/10.1109/CGIV.2012.13

[100] T. Wan, Y. Wang, and J. Li, "Hand gesture recognition system using depth data," pp. 1063–1066, Apr. 2012. [Online]. Available: http://dx.doi.org/10.1109/CECNet.2012.6201837

[101] L. Li and S. Dai, "Bayesian neural network approach to hand gesture recognition system," pp. 2019–2023, Aug. 2014. [Online]. Available: http://dx.doi.org/10.1109/CGNCC.2014.7007487

[102] Y. Chen, B. Luo, Y.-L. Chen, G. Liang, and X. Wu, "A real-time dynamic hand gesture recognition system using kinect sensor," Dec. 2015. [Online]. Available: http://dx.doi.org/10.1109/ROBIO.2015.7419071

[103] R. Harshitha, I. A. Syed, and S. Srivasthava, "Hci using hand gesture recognition for digital sand model," pp. 453–457, Dec. 2013. [Online]. Available: http://dx.doi.org/10.1109/ICIIP.2013.6707633

[104] S. Shiravandi, M. Rahmati, and F. Mahmoudi, "Hand gestures recognition using dynamic bayesian networks," pp. 1–6, Apr. 2013. [Online]. Available: http://dx.doi.org/10.1109/RIOS.2013.6595318

[105] N. N. Bhat, Y. V. Venkatesh, U. Karn, and D. Vig, "Hand gesture recognition using self organizing map for human computer interaction," pp. 734–738, Aug. 2013. [Online]. Available: http://dx.doi.org/10.1109/ICACCI.2013.6637265

[106] X. Wu, C. Yang, Y. Wang, H. Li, and S. Xu, "An intelligent interactive system based on hand gesture recognition algorithm and kinect," pp. 294–298, Oct. 2012. [Online]. Available: http://dx.doi.org/10.1109/ISCID.2012.225

[107] A. R. Asif, A. Waris, S. O. Gilani, M. Jamil, H. Ashraf, M. Shafique, and I. K. Niazi, "Performance evaluation of convolutional neural network for hand gesture recognition using emg," *Sensors*, vol. 20, no. 6, p. 1642, Mar. 2020. [Online]. Available: http://dx.doi.org/10.3390/s20061642

[108] R. Antonius and H. Tjahyadi, "Electromyography gesture identification using cnn-rnn neural network for controlling quadcopters," *Journal of Physics: Conference Series*, vol. 1858, no. 1, p. 012075, Apr. 2021. [Online]. Available: http://dx.doi.org/10.1088/1742-6596/1858/1/012075

[109] V. Adithya and R. Rajesh, "Hand gestures for emergency situations: A video dataset based on words from indian sign language," *Data in Brief*, vol. 31, p. 106016, Aug. 2020. [Online]. Available: http://dx.doi.org/10.1016/j.dib.2020.106016

[110] G. Yuan, X. Liu, Q. Yan, S. Qiao, Z. Wang, and L. Yuan, "Hand gesture recognition using deep feature fusion network based on wearable sensors," *IEEE Sensors Journal*, p. 1, 2020. [Online]. Available: http://dx.doi.org/10.1109/JSEN.2020.3014276

[111] H. Zhang, Z.-H. Bo, J.-H. Yong, and F. Xu, "Interactionfusion: real-time reconstruction of hand poses and deformable objects in hand-object interactions," *ACM Transactions on Graphics*, vol. 38, no. 4, pp. 1-11, Jul. 2019. [Online]. Available: http://dx.doi.org/10.1145/3306346.3322998

[112] K. Lai and S. N. Yanushkevich, "Cnn+rnn depth and skeleton based dynamic hand gesture recognition," pp.

3451-3456, Aug. 2018. [Online]. Available: http://dx.doi.org/10.1109/ICPR.2018.8545718

[113] J. Wan, Y. Zhao, S. Zhou, I. Guyon, S. Escalera, and S. Z. Li, "Chalearn looking at people rgb-d isolated and continuous datasets for gesture recognition," June 2016.

[114] A. Amir, B. Taba, D. Berg, T. Melano, J. McKinstry, C. Di Nolfo, T. Nayak, A. Andreopoulos, G. Garreau, M. Mendoza, J. Kusnitz, M. Debole, S. Esser, T. Delbruck, M. Flickner, and D. Modha, "A low power, fully event-based gesture recognition system," July 2017.

[115] 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 network," June 2016.

[116] Z. Zhou, K.-S. Lui, V. W. Tam, and E. Y. Lam, "Applying (3+2+1)d residual neural network with frame selection for hong kong sign language recognition," pp. 4296-4302, Jan. 2021. [Online]. Available: http://dx.doi.org/10.1109/ICPR48806.2021.9412075

[117] U. von Agris, M. Knorr, and K.-F. Kraiss, "The significance of facial features for automatic sign language recognition," pp. 1-6, Sep. 2008. [Online]. Available: http://dx.doi.org/10.1109/AFGR.2008.4813472

[118] S. Albanie, G. Varol, L. Momeni, H. Bull, T. Afouras, H. Chowdhury, N. Fox, B. Woll, R. Cooper, A. McParland, and A. Zisserman, "Bbc-oxford british sign language dataset," 2021. [Online]. Available: https://arxiv.org/abs/2111.03635

[119] J. Fink, B. Frenay, L. Meurant, and A. Cleve, "Lsfb-cont and lsfb-isol: Two new datasets for vision-based sign language recognition," pp. 1-8, Jul. 2021. [Online]. Available: http://dx.doi.org/10.1109/IJCNN52387.2021.9534336

[120] C. C. de Amorim, D. Macêdo, and C. Zanchettin, "Spatial-temporal graph convolutional networks for sign language recognition," pp. 646-657, 2019. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-30493-5_59

[121] O. M. Sincan and H. Y. Keles, "Autsl: A large scale multi-modal turkish sign language dataset and baseline methods," *IEEE Access*, vol. 8, pp. 181 340-181 355, 2020. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2020.3028072

[122] J. P. Vasconez, L. I. Barona Lopez, L. Valdivieso Caraguay, and M. E. Benalcazar, "Hand gesture recognition using emg-imu signals and deep q-networks," *Sensors*, vol. 22, no. 24, p. 9613, Dec. 2022. [Online]. Available: http://dx.doi.org/10.3390/s22249613

[123] medRxiv, "medrxiv: The preprint server for health sciences," https://www.medrxiv.org/, 2023, accessed: [2-2-2025]. [Online]. Available: https://www.medrxiv.org/

[124] S. Shaik and S. Saraswathi, "Enhance the ai virtual system accuracy with novel hand gesture recognition algorithm comparing to convolutional neural network," vol. 491, p. 04022, 2024. [Online]. Available: https://doi.org/10.1051/e3sconf/202449104022

[125] X.-H. Zhang, J.-J. Wang, X. Wang, and X.-L. Ma, "Improvement of dynamic hand gesture recognition based on hmm algorithm," pp. 401-406, Jun. 2016. [Online]. Available: http://dx.doi.org/10.1109/ISAI.2016.0091

[126] Y. Li, X. Feng, Y. Xu, X. Dong, Z. Xu, J. Huang, and L. Lu, "A dynamic hand gesture recognition model based on the improved dynamic time warping algorithm," pp. 1-6, Sep. 2019. [Online]. Available: http://dx.doi.org/10.23919/IConAC.2019.8895002

[127] A. Mohanarathinam, K. Dharani, R. Sangeetha, G. Aravindh, and P. Sasikala, "Study on hand gesture recoginition by using machine learning," pp. 1498-1501, Nov. 2020. [Online]. Available: http://dx.doi.org/10.1109/ICECA49313.2020.9297513

[128] A. V. and R. R., "A deep convolutional neural network approach for static hand gesture recognition," *Procedia Computer Science*, vol. 171, pp. 2353-2361, 2020. [Online]. Available: http://dx.doi.org/10.1016/j.procs.2020.04.255

[129] O. Mazhar, S. Ramdani, and A. Cherubini, "A deep learning framework for recognizing both static and dynamic gestures," *Sensors*, vol. 21, no. 6, p. 2227, Mar. 2021. [Online]. Available: http://dx.doi.org/10.3390/s21062227

[130] S.-H. Yang, W.-R. Chen, W.-J. Huang, and Y.-P. Chen, "Ddanet: Dual-path depth-aware attention network for fingerspelling recognition using rgb-d images," *IEEE Access*, vol. 9, pp. 7306 – 7322, 2021. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2020.3046667

[131] P. Koch, M. Dreier, M. Maass, M. Bohme, H. Phan, and A. Mertins, "A recurrent neural network for hand gesture recognition based on accelerometer data," pp. 5088-5091, Jul. 2019. [Online]. Available: http://dx.doi.org/10.1109/EMBC.2019.8856844

[132] W. Zhang, J. Wang, and F. Lan, "Dynamic hand gesture recognition based on short-term sampling neural networks," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 1, pp. 110 – 120, Jan. 2021. [Online]. Available: http://dx.doi.org/10.1109/JAS.2020.1003465

[133] W. Yu and E. N. Sanchez, Eds., *Advances in Computational Intelligence: Proceedings of the IWACI'09 Workshop, 22nd-23rd June 2009, Mexico City*, ser. Advances in Intelligent and Soft Computing. Berlin, Heidelberg: Springer, 2009, vol. 61, presents latest research in Computational Intelligence.

[134] E. Tsironi, P. Barros, C. Weber, and S. Wermter, "An analysis of convolutional long short-term memory recurrent neural networks for gesture recognition," *Neurocomputing*, vol. 268, pp. 76-86, Dec. 2017. [Online]. Available: http://dx.doi.org/10.1016/j.neucom.2016.12.088

[135] W. Tao, M. C. Leu, and Z. Yin, "American sign language alphabet recognition using convolutional neural networks with multiview augmentation and inference fusion," *Engineering Applications of Artificial Intelligence*, vol. 76, pp. 202 – 213, Nov. 2018. [Online]. Available: http://dx.doi.org/10.1016/j.engappai.2018.09.006

[136] C. Bhuvaneshwari and A. Manjunathan, "Advanced gesture recognition system using long-term recurrent convolution network," *Materials Today: Proceedings*, vol. 21, pp. 731 – 733, 2020. [Online]. Available: http://dx.doi.org/10.1016/j.matpr.2019.06.748

[137] N. B. Ibrahim, M. M. Selim, and H. H. Zayed, "An automatic arabic sign language recognition system (arslrs)," *Journal of King Saud University - Computer and Information Sciences*, vol. 30, no. 4, pp. 470 – 477, Oct. 2018. [Online]. Available: http://dx.doi.org/10.1016/j.jksuci.2017.09.007

[138] M. Oliveira, H. Chatbri, Y. Ferstl, M. Farouk, S. Little, N. E. O'Connor, and A. Sutherland, "A dataset for irish sign language recognition," Aug. 2017.

[139] Y. SHI, Y. LI, X. FU, M. Kaibin, and M. Qiguang, "Review of dynamic gesture recognition," *Virtual Reality amp; Intelligent Hardware*, vol. 3, no. 3, pp. 183 – 206, Jun. 2021. [Online]. Available: http://dx.doi.org/10.1016/j.vrih.2021.05.001

[140] 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," *International Journal of Computer Vision*, vol. 126, no. 24, pp. 430 – 439, Oct. 2016. [Online]. Available: http://dx.doi.org/10.1007/s11263-016-0957-7

[141] D. Jiang, G. Li, Y. Sun, J. Kong, B. Tao, and D. Chen, "Grip strength forecast and rehabilitative guidance based on adaptive neural fuzzy inference system using semg," *Personal and Ubiquitous Computing*, vol. 26, no. 4, pp. 1215 – 1224, Jul. 2019. [Online]. Available: http://dx.doi.org/10.1007/s00779-019-01268-3

[142] P. Sharma and R. S. Anand, "Depth data and fusion of feature descriptors for static gesture recognition," *IET Image Processing*, vol. 14, no. 5, pp. 909 – 920, Mar. 2020. [Online]. Available: http://dx.doi.org/10.1049/iet-ipr.2019.0230

[143] 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," pp. 107 – 118, Aug. 2019. [Online]. Available: http://dx.doi.org/10.1007/978-3-030-29933-0_9

[144] Z. Gao, P. Wang, H. Wang, M. Xu, and W. Li, "A review of dynamic maps for 3d human motion recognition using convnets and its improvement," *Neural Processing Letters*, vol. 52, no. 2, pp. 1501 – 1515, Jul. 2020. [Online]. Available: http://dx.doi.org/10.1007/s11063-020-10320-w

[145] K. Yin and J. Read, "Better sign language translation with stmc-transformer," 2020. [Online]. Available: https://arxiv.org/abs/2004.00588

[146] Q. De Smedt, H. Wannous, and J.-P. Vandeborre, "3d hand gesture recognition by analysing set-of-joints trajectories," pp. 86 – 97, 2018. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-91863-1_7

[147] ——, "Skeleton-based dynamic hand gesture recognition," June 2016.

[148] S. Y. Boulahia, E. Anquetil, F. Multon, and R. Kulpa, "Dynamic hand gesture recognition based on 3d pattern assembled trajectories," pp. 1 – 6, Nov. 2017. [Online]. Available: http://dx.doi.org/10.1109/IPTA.2017.8310146

[149] J. Liu, A. Shahroudy, D. Xu, A. C. Kot, and G. Wang, "Skeleton-based action recognition using spatio-temporal lstm network with trust gates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 12, pp. 3007 – 3021, Dec. 2018. [Online]. Available: http://dx.doi.org/10.1109/TPAMI.2017.2771306

[150] D. Konstantinidis, K. Dimitropoulos, and P. Daras, "A deep learning approach for analyzing video and skeletal features in sign language recognition," pp. 1 – 6, Oct. 2018. [Online]. Available: http://dx.doi.org/10.1109/IST.2018.8577085

[151] J. C. Nunez, R. Cabido, J. J. Pantrigo, A. S. Montemayor, and J. F. Velez, "Convolutional neural networks and long short-term memory for skeleton-based human activity and hand gesture recognition," *Pattern Recognition*, vol. 76, pp. 80 – 94, Apr. 2018. [Online]. Available: http://dx.doi.org/10.1016/j.patcog.2017.10.033

[152] S. Yan, Y. Xiong, and D. Lin, "Spatial temporal graph convolutional networks for skeleton-based action recognition," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, Apr. 2018. [Online]. Available: http://dx.doi.org/10.1609/aaai.v32i1.12328

[153] C. Ma, A. Wang, G. Chen, and C. Xu, "Hand joints-based gesture recognition for noisy dataset using nested interval unscented kalman filter with lstm network," *The Visual Computer*, vol. 34, no. 6-8, pp. 1053 – 1063, May 2018. [Online]. Available: http://dx.doi.org/10.1007/s00371-018-1556-0

[154] W. Wei, Y. Wong, Y. Du, Y. Hu, M. Kankanhalli, and W. Geng, "A multi-stream convolutional neural network for semg-based gesture recognition in muscle-computer interface," *Pattern Recognition Letters*, vol. 119, pp. 131 – 138, Mar. 2019. [Online]. Available: http://dx.doi.org/10.1016/j.patrec.2017.12.005

[155] C. Si, W. Chen, W. Wang, L. Wang, and T. Tan, "An attention enhanced graph convolutional lstm network for skeleton-based action recognition," June 2019.

[156] S. Jiang, B. Sun, L. Wang, Y. Bai, K. Li, and Y. Fu, "Sign language recognition via skeleton-aware multi-model ensemble," 2021. [Online]. Available: https://arxiv.org/abs/2110.06161

[157] ——, "Skeleton aware multi-modal sign language recognition," pp. 3413 – 3423, June 2021.

[158] R. Rastgoo, K. Kiani, and S. Escalera, "Sign language recognition: A deep survey," *Expert Systems with Applications*, vol. 164, p. 113794, Feb. 2021. [Online]. Available: http://dx.doi.org/10.1016/j.eswa.2020.113794

[159] H. Zhang, D. Liu, and Z. Xiong, "Two-stream action recognition-oriented video super-resolution," October 2019.

[160] U. Cote-Allard, C. L. Fall, A. Drouin, A. Campeau-Lecours, C. Gosselin, K. Glette, F. Laviolette, and B. Gosselin, "Deep learning for electromyographic hand gesture signal classification using transfer learning," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 27, no. 4, pp. 760 – 771, Apr. 2019. [Online]. Available: http://dx.doi.org/10.1109/TNSRE.2019.2896269

[161] H. Wei, R. Jafari, and N. Kehtarnavaz, "Fusion of video and inertial sensing for deep learning–based human action recognition," *Sensors*, vol. 19, no. 17, p. 3680, Aug. 2019. [Online]. Available: http://dx.doi.org/10.3390/s19173680

[162] K. H. Lee, J. Y. Min, and S. Byun, "Electromyogram-based classification of hand and finger gestures using artificial neural networks," *Sensors*, vol. 22, no. 1, p. 225, Dec. 2021. [Online]. Available: http://dx.doi.org/10.3390/s22010225

[163] G. Dougherty, *Pattern Recognition and Classification: An Introduction*. New York, NY: Springer, 2013.

[164] B.-A. Awaluddin, C.-T. Chao, and J.-S. Chiou, "A hybrid image augmentation technique for user- and environment-independent hand gesture recognition based on deep learning," *Mathematics*, vol. 12, no. 9, p. 1393, May 2024. [Online]. Available: http://dx.doi.org/10.3390/math12091393

[165] M. Garg, D. Ghosh, and P. M. Pradhan, "Gestformer: Multiscale wavelet pooling transformer network for dynamic hand gesture recognition," 2024. [Online]. Available: https://arxiv.org/abs/2405.11180

[166] L. I. Barona Lopez, F. M. Ferri, J. Zea, L. Valdivieso Caraguay, and M. E. Benalcazar, "Cnn-lstm and post-processing for emg-based hand gesture recognition," *Intelligent Systems with Applications*, vol. 22, p. 200352, Jun. 2024. [Online]. Available: http://dx.doi.org/10.1016/j.iswa.2024.200352

[167] G. Zhou, Z. Cui, and J. Qi, "Fgdsnet: A lightweight hand gesture recognition network for human robot interaction," *IEEE Robotics and Automation Letters*, vol. 9, no. 4, pp. 3076 – 3083, Apr. 2024. [Online]. Available: http://dx.doi.org/10.1109/LRA.2024.3362144

[168] P. Balaji and M. Ranjan Prusty, "Multimodal fusion hierarchical self-attention network for dynamic hand gesture recognition," *Journal of Visual Communication and Image Representation*, vol. 98, p. 104019, Feb. 2024. [Online]. Available: http://dx.doi.org/10.1016/j.jvcir.2023.104019

[169] G. Bhaumik and M. C. Govil, "Spatnet: a spatial feature attention network for hand gesture recognition," *Multimedia Tools and Applications*, vol. 83, no. 14, pp. 41 805 – 41 822, Oct. 2023. [Online]. Available: http://dx.doi.org/10.1007/s11042-023-16988-1

[170] H.-Q. Nguyen, T.-H. Le, T.-K. Tran, H.-N. Tran, T.-H. Tran, T.-L. Le, H. Vu, C. Pham, T. P. Nguyen, and H. T. Nguyen, "Hand gesture recognition from wrist-worn camera for human-machine interaction," *IEEE Access*, vol. 11, pp. 53 262 – 53 274, 2023. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2023.3279845

[171] Y. Zhang, W. Kang, and W. Song, "Robust and accurate hand gesture authentication with cross-modality local-global behavior analysis," *IEEE Transactions on Information Forensics and Security*, vol. 19, pp. 8630 – 8643, 2024. [Online]. Available: http://dx.doi.org/10.1109/TIFS.2024.3451367

[172] S. Birkeland, L. J. Fjeldvik, N. Noori, S. R. Yeduri, and L. R. Cenkeramaddi, "Thermal video-based hand gestures recognition using lightweight cnn," *Journal of Ambient Intelligence and Humanized Computing*, vol. 15, no. 12, pp. 3849 – 3860, Sep. 2024. [Online]. Available: http://dx.doi.org/10.1007/s12652-024-04851-6

[173] M. Garg, D. Ghosh, and P. M. Pradhan, "Mvtn: A multiscale video transformer network for hand gesture recognition," 2024. [Online]. Available: https://arxiv.org/abs/2409.03890

[174] J. Shin, A. S. M. Miah, S. Konnai, I. Takahashi, and K. Hirooka, "Hand gesture recognition using semg signals with a multi-stream time-varying feature enhancement approach," *Scientific Reports*, vol. 14, no. 1, Sep. 2024. [Online]. Available: http://dx.doi.org/10.1038/s41598-024-72996-7

[175] A. S. M. Miah, N. Hassan, M. Maniruzzaman, N. Asai, and J. Shin, "Emg-based hand gesture recognition through diverse domain feature enhancement and machine learning-based approach," 2024. [Online]. Available: https://arxiv.org/abs/2408.13723

[176] H. Mahmud, M. M. Morshed, and M. K. Hasan, "Quantized depth image and skeleton-based multimodal dynamic hand gesture recognition," *The Visual Computer*, vol. 40, no. 1, pp. 11 – 25, Jan. 2023. [Online]. Available: http://dx.doi.org/10.1007/s00371-022-02762-1

[177] Z. Dozdor, Z. Kalafatic, Ban, and T. Hrkaz, "Ty-net: Transforming yolo for hand gesture recognition," *IEEE Access*, vol. 11, pp. 140 382 – 140 394, 2023. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2023.3341702

[178] Yaseen, O.-J. Kwon, J. Kim, S. Jamil, J. Lee, and F. Ullah, "Next-gen dynamic hand gesture recognition: Mediapipe, inception-v3 and lstm-based enhanced deep learning model," *Electronics*, vol. 13, no. 16, p. 3233, Aug. 2024. [Online]. Available: http://dx.doi.org/10.3390/electronics13163233

[179] J. W. Smith, S. Thiagarajan, R. Willis, Y. Makris, and M. Torlak, "Improved static hand gesture classification on deep convolutional neural networks using novel sterile training technique," *IEEE Access*, vol. 9, pp. 10 893 – 10 902, 2023. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2021.3051454

[180] D. Sarma, H. P. J. Dutta, K. S. Yadav, M. Bhuyan, and R. H. Laskar, "Attention-based hand semantic segmentation and gesture recognition using deep networks," *Evolving Systems*, vol. 15, no. 1, pp. 185 – 201, Jul. 2023. [Online]. Available: http://dx.doi.org/10.1007/s12530-023-09512-1

[181] Z. Zhang, Q. Shen, and Y. Wang, "Electromyographic hand gesture recognition using convolutional neural network with multi-attention," *Biomedical Signal Processing and Control*, vol. 91, p. 105935, May 2024.

[182] S. Zhang, H. Zhou, R. Tchantchane, and G. Alici, "A wearable human-machine-interface (hmi) system based on colocated emg-pfmg sensing for hand gesture recognition," *IEEE/ASME Transactions on Mechatronics*, pp. 1 – 12, 2024. [Online]. Available: http://dx.doi.org/10.1109/TMECH.2024.3386929

[183] N. Fadel and E. I. A. Kareem, "Real-time hand gesture recognition based on multi-connect architecture associative memory in human computer interaction," vol. 3061, p. 030003, 2024. [Online]. Available: http://dx.doi.org/10.1063/5.0203656

[184] A. Chandrabose, X. Fernando, and E. Mercier-Laurent, Eds., *Computer, Communication, and Signal Processing. Smart Solutions Towards SDG: 8th IFIP TC 12 International Conference, ICCCSP 2024, Chennai, India, March 20-22, 2024, Revised Selected Papers*, ser. IFIP Advances in Information and Communication Technology. Cham, Switzerland: Springer, 2024, vol. 723.

[185] Q. U. of Technology, "Qut eprints," https://eprints.qut.edu.au/, 2023, accessed: [2-2-2025]. [Online]. Available: https://eprints.qut.edu.au/

[186] A. Kumar and R. Saini, "A comparative study of machine learning techniques for hand gesture recognition to guide multimedia operation," *COMPUTER*, vol. 24, no. 8, pp. 145 – 160, 2024.

[187] T. Jawalkar, S. S. Khalate, S. A. Medhe, and K. S. Palaskar, "Hand gesture recognition using ai/ml," *International Journal of Advanced Engineering Application*, vol. 10, pp. 95–115, 2024.

[188] E. Andrei, T. Cornel, G. Culea, S. B. Constantin, and U. A. Gabriel, "Romanian sign language and mime-gesture recognition," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 8, 2024. [Online]. Available: http://dx.doi.org/10.14569/IJACSA.2024.0150888

[189] L. Zholshiyeva, Z. Manbetova, D. Kaibassova, A. Kassymova, Z. Tashenova, S. Baizhumanov, A. Yerzhanova, and K. Aikhynbay, "Human-machine interactions based on hand gesture recognition using deep learning methods," *International Journal of Electrical and Computer Engineering*

(IJECE), vol. 14, no. 1, p. 741, Feb. 2024. [Online]. Available: http://dx.doi.org/10.11591/ijece.v14i1.pp741-748

[190] X. Savarimuthu, S. Subramani, and A. N. J. Raj, Eds., *Artificial Intelligence for Multimedia Information Processing: Tools and Applications*. Boca Raton, FL: CRC Press, 2024, 69 B/W Illustrations.

[191] W. Han, M. Hao, Y. Yuan, and P. Liu, "Fusion enhancement of yolov5 and copula bayesian classifier for hand gesture recognition in smart sports venues," *IEEE Access*, vol. 12, pp. 67 005 – 67 012, 2024. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2024.3398142

[192] H. Ansar, N. A. Mudawi, S. S. Alotaibi, A. Alazeb, B. I. Alabdullah, M. Alonazi, and J. Park, "Hand gesture recognition for characters understanding using convex hull landmarks and geometric features," *IEEE Access*, vol. 11, pp. 82 065 – 82 078, 2023. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2023.3300712

[193] C. M. Suryateja, S. Boppu, L. R. Cenkeramaddi, and B. Ramkumar, "Hand gesture recognition system in the complex background for edge computing devices," pp. 13 – 18, Dec. 2022. [Online]. Available: http://dx.doi.org/10.1109/iSES54909.2022.00016

[194] C. Schuessler, W. Zhang, J. Bräunig, M. Hoffmann, M. Stelzig, and M. Vossiek, "Radar-based recognition of static hand gestures in american sign language," 2024. [Online]. Available: https://arxiv.org/abs/2402.12800

[195] O. Y. Fadhil, B. S. Mahdi, and A. R. Abbas, "Using vgg models with intermediate layer feature maps for static hand gesture recognition," *Baghdad Science Journal*, Feb. 2023. [Online]. Available: http://dx.doi.org/10.21123/bsj.2023.7364

[196] J. P. Sahoo, S. P. Sahoo, S. Ari, and S. K. Patra, "Derefnet: Dual-stream dense residual fusion network for static hand gesture recognition," *Displays*, vol. 77, p. 102388, Apr. 2023. [Online]. Available: http://dx.doi.org/10.1016/j.displa.2023.102388

[197] Y. Jiang, L. Song, J. Zhang, Y. Song, and M. Yan, "Multi-category gesture recognition modeling based on semg and imu signals," *Sensors*, vol. 22, no. 15, p. 5855, Aug. 2022. [Online]. Available: http://dx.doi.org/10.3390/s22155855

[198] M. Al-Hammadi, M. A. Bencherif, M. Alsulaiman, G. Muhammad, M. A. Mekhtiche, W. Abdul, Y. A. Alohali, T. S. Alrayes, H. Mathkour, M. Faisal, M. Algabri, H. Altaheri, T. Alfakih, and H. Ghaleb, "Spatial attention-based 3d graph convolutional neural network for sign language recognition," *Sensors*, vol. 22, no. 12, p. 4558, Jun. 2022. [Online]. Available: http://dx.doi.org/10.3390/s22124558

[199] 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 Journal of the Computer Society*, vol. 5, pp. 144 – 155, 2024. [Online]. Available: http://dx.doi.org/10.1109/OJCS.2024.3370971

[200] Q. Li, Z. Luo, R. Qi, and J. Zheng, "Deeptpa-net: A deep triple attention network for semg-based hand gesture recognition," *IEEE Access*, vol. 11, pp. 96 797 – 96 807, 2023. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2023.3312219

[201] U. Imran, A. Waris, S. O. Gilani, J. Iqbal, M. Jamil, G. E. Eldesoky, and S. Mushtaq, "Patient-specific movement regime: Investigating the potential of upper-extremity motions for intelligent myoelectric prosthetic control," *IEEE Access*, vol. 12, pp. 35 663–35 682, 2024. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2024.3365639

[202] arXiv, "arxiv.org e-print archive: Computer science," 2025, accessed: 2025-02-02. [Online]. Available: https://arxiv.org/search/cs

[203] V. Sharmila, S. Kannadhasan, A. R. Kannan, P. Sivakumar, and V. Vennila, Eds., *Challenges in Information, Communication, and Computing Technology: Proceedings of the 2nd International Conference on Challenges in Information, Communication, and Computing Technology (ICCICCT 2024), April 26th & 27th, 2024, Namakkal, Tamil Nadu, India*. Boca Raton, FL: CRC Press, 2024.

[204] V. Mottini, "Precision neural interfaces through intrinsically stretchable electronics," Ph.D. dissertation, Michigan State University, 2024.

[205] I. Society, "Insight society," https://insightsociety.org/, 2023, accessed: [2-2-2025]. [Online]. Available: https://insightsociety.org/

[206] S. Manoharan, A. Tugui, and Z. Baig, Eds., *Proceedings of the 4th International Conference on Artificial Intelligence and Smart Energy (ICAIS 2024), Volume 2*. Springer Nature, 2024.

[207] T. N. S. R. L. M. (TNSRLM), "Tamil nadu state rural livelihoods mission," https://www.tnsroindia.org.in/, 2023, accessed: [2-2-2025]. [Online]. Available: https://www.tnsroindia.org.in/

[208] J. Shin, A. S. M. Miah, S. Konnai, I. Takahashi, and K. Hirooka, "Hand gesture recognition using semg signals with a multi-stream time-varying feature enhancement approach," *Scientific Reports*, vol. 14, no. 1, Sep. 2024. [Online]. Available: http://dx.doi.org/10.1038/s41598-024-72996-7

[209] Sunanda, A. Balmik, and A. Nandy, "A novel feature fusion technique for robust hand gesture recognition," *Multimedia Tools and Applications*, vol. 83, no. 25, pp. 65 815 – 65 831, Jan. 2024. [Online]. Available: http://dx.doi.org/10.1007/s11042-024-18173-4

[210] F. A. Farid, N. Hashim, J. B. Abdullah, M. R. Bhuiyan, M. Kairanbay, Z. Yusoff, H. A. Karim, S. Mansor, M. T. Sarker, and G. Ramasamy, "Single shot detector cnn and deep dilated masks for vision-based hand gesture recognition from video sequences," *IEEE Access*, vol. 12, pp. 28 564 – 28 574, 2024. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2024.3360857

[211] M. Jaiswal, V. Sharma, A. Sharma, S. Saini, and R. Tomar, "Quantized cnn-based efficient hardware architecture for real-time hand gesture recognition," *Microelectronics Journal*, vol. 151, p. 106345, Sep. 2024. [Online]. Available: http://dx.doi.org/10.1016/j.mejo.2024.106345

[212] J. Tang, K. Gou, C. Wang, M. Wei, Q. Tan, and G. Weng, "Self-powered and 3d printable soft sensor for human health monitoring, object recognition, and contactless hand gesture recognition," *Advanced Functional Materials*, vol. 34, no. 52, Sep. 2024. [Online]. Available: http://dx.doi.org/10.1002/adfm.202411172

[213] Y. Wang, Y. Tian, J. Zhu, H. She, Y. Jiang, Z. Jiang, and H. Yokoi, "A hand gesture recognition strategy

based on virtual-dimension increase of emg," *Cyborg and Bionic Systems*, vol. 5, Jan. 2024. [Online]. Available: http://dx.doi.org/10.34133/cbsystems.0066

[214] A. Sen, S. Dombe, T. K. Mishra, and R. Dash, "Hgr-fyolo: a robust hand gesture recognition system for the normal and physically impaired person using frozen yolov5," *Multimedia Tools and Applications*, vol. 83, no. 30, pp. 73 797 – 73 815, Feb. 2024. [Online]. Available: http://dx.doi.org/10.1007/s11042-024-18464-w

[215] Z. Mohammadi, A. Akhavanpour, R. Rastgoo, and M. Sabokrou, "Diverse hand gesture recognition dataset," *Multimedia Tools and Applications*, vol. 83, no. 17, pp. 50 245 – 50 267, Nov. 2023. [Online]. Available: http://dx.doi.org/10.1007/s11042-023-17268-8

[216] E. Bamani, E. Nissinman, I. Meir, L. Koenigsberg, and A. Sintov, "Ultra-range gesture recognition using a web-camera in human-robot interaction," *Engineering Applications of Artificial Intelligence*, vol. 132, p. 108443, Jun. 2024. [Online]. Available: http://dx.doi.org/10.1016/j.engappai.2024.108443

[217] Sundaram and B. C. Sahana, "Multivariate emg signal based automated hand gestures recognition framework for elder care," *International Journal of Precision Engineering and Manufacturing*, Aug. 2024. [Online]. Available: http://dx.doi.org/10.1007/s12541-024-01116-2

[218] A. Shaaban, M. Strobel, W. Furtner, R. Weigel, and F. Lurz, "Rt-scnns: real-time spiking convolutional neural networks for a novel hand gesture recognition using time-domain mm-wave radar data," *International Journal of Microwave and Wireless Technologies*, vol. 16, no. 5, pp. 783 – 795, Jan. 2024. [Online]. Available: http://dx.doi.org/10.1017/S1759078723001575

[219] T.-H. Tsai, Y.-C. Ho, P.-T. Chi, and T.-J. Chen, "A deep neural network for hand gesture recognition from rgb image in complex background," *Signal, Image and Video Processing*, vol. 18, no. S1, pp. 861 – 872, May 2024. [Online]. Available: http://dx.doi.org/10.1007/s11760-024-03198-x

[220] S. Padmakala, S. O. Husain, E. Poornima, P. Dutta, and M. Soni, "Hyperparameter tuning of deep convolutional neural network for hand gesture recognition," p. 1–4, Aug. 2024. [Online]. Available: http://dx.doi.org/10.1109/NMITCON62075.2024.10698984

[221] S. N. Uke and A. Zade, "Optimal video processing and soft computing algorithms for human hand gesture recognition from real-time video," *Multimedia Tools and Applications*, vol. 83, no. 17, pp. 50 425 – 50 447, Nov. 2023. [Online]. Available: http://dx.doi.org/10.1007/s11042-023-17608-8

[222] K. Bimbraw, A. Talele, and H. K. Zhang, "Hand gesture classification based on forearm ultrasound video snippets using 3d convolutional neural networks," 2024. [Online]. Available: https://arxiv.org/abs/2409.16431

[223] S. Misal, "Hand gesture recognition using deep learning," *International Journal of Innovative Science and Research Technology (IJISRT)*, pp. 69 – 72, Aug. 2024. [Online]. Available: http://dx.doi.org/10.38124/ijisrt/IJISRT24AUG154

[224] Q. Gao, Y. Chen, Z. Ju, and Y. Liang, "Dynamic hand gesture recognition based on 3d hand pose estimation for human-robot interaction," *IEEE Sensors Journal*, vol. 22, no. 18, pp. 17 421 – 17 430, Sep. 2022. [Online]. Available: http://dx.doi.org/10.1109/JSEN.2021.3059685

[225] M. Alonazi, H. Ansar, N. A. Mudawi, S. S. Alotaibi, N. A. Almujally, A. Alazeb, A. Jalal, J. Kim, and M. Min, "Smart healthcare hand gesture recognition using cnn-based detector and deep belief network," *IEEE Access*, vol. 11, pp. 84 922 – 84 933, 2023. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2023.3289389

[226] G. Park, V. K. Chandrasegar, and J. Koh, "Accuracy enhancement of hand gesture recognition using cnn," *IEEE Access*, vol. 11, pp. 26 496 – 26 501, 2023. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2023.3254537

[227] M. Zhang, Z. Zhou, T. Wang, and W. Zhou, "A lightweight network deployed on arm devices for hand gesture recognition," *IEEE Access*, vol. 11, pp. 45 493 – 45 503, 2023. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2023.3273713

[228] S. Savas and A. Ergüzen, "Hand gesture recognition with two stage approach using transfer learning and deep ensemble learning," 2023. [Online]. Available: https://arxiv.org/abs/2309.11610

[229] F. Jafari and A. Basu, "Two-dimensional parallel spatio-temporal pyramid pooling for hand gesture recognition," *IEEE Access*, vol. 11, pp. 133 755 – 133 766, 2023. [Online]. Available: http://dx.doi.org/10.1109/ACCESS.2023.3336591

[230] T. L. Dang, S. D. Tran, T. H. Nguyen, S. Kim, and N. Monet, "An improved hand gesture recognition system using keypoints and hand bounding boxes," *ScienceDirect*, vol. 16, p. 100251, Dec. 2022. [Online]. Available: http://dx.doi.org/10.1016/j.array.2022.100251

[231] X. Pan, T. Jiang, X. Li, X. Ding, Y. Wang, and Y. Li, "Dynamic hand gesture detection and recognition with wifi signal based on 1d-cnn," pp. 1 – 6, May 2019. [Online]. Available: http://dx.doi.org/10.1109/ICCW.2019.8756690

[232] V. Chang, R. O. Eniola, L. Golightly, and Q. A. Xu, "An exploration into human-computer interaction: Hand gesture recognition management in a challenging environment," *SN Computer Science*, vol. 4, no. 5, Jun. 2023. [Online]. Available: http://dx.doi.org/10.1007/s42979-023-01751-y

[233] J. Shin, A. S. M. Miah, S. Konnai, I. Takahashi, and K. Hirooka, "Hand gesture recognition using semg signals with a multi-stream time-varying feature enhancement approach," *Scientific Reports*, vol. 14, no. 1, Sep. 2024. [Online]. Available: http://dx.doi.org/10.1038/s41598-024-72996-7

[234] R. K. Pathan, M. Biswas, S. Yasmin, M. U. Khandaker, M. Salman, and A. A. F. Youssef, "Retracted article: Sign language recognition using the fusion of image and hand landmarks through multi-headed convolutional neural network," *Scientific Reports*, vol. 13, no. 1, Oct. 2023. [Online]. Available: http://dx.doi.org/10.1038/s41598-023-43852-x

[235] J. Yu, M. Qin, and S. Zhou, "Dynamic gesture recognition based on 2d convolutional neural network and feature fusion," *Scientific Reports*, vol. 12, no. 1, Mar. 2022. [Online]. Available: http://dx.doi.org/10.1038/s41598-022-08133-z

[236] H. Li and H. Guo, "Design of bionic robotic hand gesture recognition system based on machine vision," pp. 960 – 964, May 2022. [Online]. Available: http://dx.doi.org/10.1109/CVIDLICCEA56201.2022.9825148

[237] A. Kaur and S. Bansal, “Deep learning for dynamic hand gesture recognition: Applications, challenges and future scope,” pp. 1–6, Nov. 2022. [Online]. Available: http://dx.doi.org/10.1109/IMPACT55510.2022.10029100

[238] H. Arwoko, E. M. Yuniarno, and M. H. Purnomo, “Hand gesture recognition based on keypoint vector,” pp. 530 – 533, Aug. 2022. [Online]. Available: http://dx.doi.org/10.1109/IES55876.2022.9888333

[239] R. A. Muchtar, R. Yuniarti, and A. Komarudin, “Hand gesture recognition for controlling game objects using two-stream faster region convolutional neural networks methods,” pp. 59 – 64, Nov. 2022. [Online]. Available: http://dx.doi.org/10.1109/ICITRI56423.2022.9970207

[240] Y. Wang, J. Zhang, and X. Zhao, “Research on hand gesture recognition based on millimeter wave radar,” pp. 205 – 209, Jan. 2023. [Online]. Available: http://dx.doi.org/10.1109/ICCECE58074.2023.10135494

[241] G. Beneke, T. B. Winkler, K. Raab, M. A. Brems, F. Kammerbauer, P. Gerhards, K. Knobloch, S. Krishnia, J. H. Mentink, and M. Kläui, “Gesture recognition with brownian reservoir computing using geometrically confined skyrmion dynamics,” *Nature Communications*, vol. 15, no. 1, Sep. 2024. [Online]. Available: http://dx.doi.org/10.1038/s41467-024-52345-y

[242] L. Zhao, Z. Yu, and J. Cai, “A novel method for semg-based hand gesture recognition,” pp. 1668 – 1672, Apr. 2023. [Online]. Available: http://dx.doi.org/10.1109/ICSP58490.2023.10248486

[243] S. Mesdaghi, R. P. Hasanzadeh, and F. Janabi-Sharifi, “Finger-hand rehabilitation using dnn-based gesture recognition of low-cost webcam images,” pp. 1 – 6, Mar. 2024. [Online]. Available: http://dx.doi.org/10.1109/MVIP62238.2024.10491167

[244] R. Kushwaha, G. Kaur, and M. Kumar, “Hand gesture based sign language recognition using deep learning,” pp. 293 – 297, May 2023. [Online]. Available: http://dx.doi.org/10.1109/ICSCCC58608.2023.10176912

[245] F. Roumiassa, S. E. Agab, and F. Z. Chelali, “Hand gesture recognition system based on textural features,” pp. 1 – 6, Oct. 2022. [Online]. Available: http://dx.doi.org/10.1109/ICAEE53772.2022.9962080

[246] B. Sharma and J. Panda, “Hand gesture recognition using emd and vmd techniques,” pp. 1 – 5, Dec. 2022. [Online]. Available: http://dx.doi.org/10.1109/IATMSI56455.2022.10119304

[247] O. Ikne, B. Allaert, and H. Wannous, “Skeleton-based self-supervised feature extraction for improved dynamic hand gesture recognition,” pp. 1 – 10, May 2024. [Online]. Available: http://dx.doi.org/10.1109/FG59268.2024.10581975

[248] A. Al-Zebari, N. Omar, and A. Sengur, “Vision transformers-based hand gesture classification,” pp. 1 – 3, Dec. 2022. [Online]. Available: http://dx.doi.org/10.1109/IISEC56263.2022.9998295

[249] F. Noble, M. Xu, and F. Alam, “Static hand gesture recognition using capacitive sensing and machine learning,” *Sensors*, vol. 23, no. 7, p. 3419, Mar. 2023. [Online]. Available: http://dx.doi.org/10.3390/s23073419

[250] L. I. Khalaf, S. A. Aswad, S. R. Ahmed, B. Makki, and M. R. Ahmed, “Survey on recognition hand gesture by using data mining algorithms,” pp. 1 – 4, Jun. 2022. [Online]. Available: http://dx.doi.org/10.1109/HORA55278.2022.9800090

[251] D. G. Leon, J. Groli, S. R. Yeduri, D. Rossier, R. Mosqueron, O. J. Pandey, and L. R. Cenkeramaddi, “Video hand gestures recognition using depth camera and lightweight cnn,” *IEEE Sensors Journal*, vol. 22, no. 14, pp. 14 610 – 14 619, Jul. 2022. [Online]. Available: http://dx.doi.org/10.1109/JSEN.2022.3181518

[252] H. Quintián, E. Corchado, A. T. Lora, H. P. Garcia, E. J. Perez, J. L. C. Rolle, F. J. M. de Pison, P. G. Bringas, F. M. Alvarez, A. Herrero, and P. Fosci, Eds., *Hybrid Artificial Intelligent Systems: 19th International Conference, HAIS 2024, Salamanca, Spain, October 9–11, 2024, Proceedings, Part I*, ser. Lecture Notes in Computer Science. Springer Nature Switzerland, 2024.