Hand Gesture Recognition with Focus on Leap Motion: An Overview, Real World Challenges and Future Directions

Nahla MAJDOUB BHIRI[a], Safa AMEUR[a,*], Ihsen Alouani[b], Mohamed Ali MAHJOUB[a], Anouar BEN KHALIFA[a,c]

[a] *Université de Sousse, Ecole Nationale d'Ingénieurs de Sousse, LATIS- Laboratory of Advanced Technology and Intelligent Systems, 4023, Sousse, Tunisie;*
[b] *Center for Secure Information technologies, Queen's University Belfast, UK;*
[c] *Université de Jendouba, Institut National des Technologies et des Sciences du Kef, 7100, Le Kef, Tunisie;*
nahla.majdoubbhiri@eniso.u-sousse.tn, safa.ameur@eniso.u-sousse.tn, i.alouani@qub.ac.uk, mohamedali.mahjoub@eniso.rnu.tn, anouar.benkhalifa @eniso.u-sousse.tn

---

Abstract

In the recent years, a steady growth of Hand Gesture Recognition (HGR) based applications has been observed. Thus, significant progress has been made in the field of hand detection and tracking. However, hand poses remain particularly challenging and can deeply affect the results due to several factors especially congenital hand problems. Besides, HGR applications performance is very dependent on the sensor used for gesture acquisition. Owing to its high performance, the Leap Motion Controller (LMC), has recently become a by default sensor for HGR systems and provide high performance in terms of recognition accuracy. This work proposes a comprehensive literature review of HGR based on different modalities with a focus on the LMC based systems. This survey retrieves the main challenges faced in literature with an investigation of the state of the art solutions to address them. Moreover, we discuss numerous existing methods for static and dynamic HGR from different modalities. We reveal remaining challenges in the literature and highlight promising future research directions.

*Corresponding author at ENISo, 4023 Sousse, Tunisia
Email adress: safa.ameur@eniso.u-sousse.tn (Safa AMEUR)

Keywords: HGR, Leap Motion Controller, Modalities, Challenges, future research directions.

1. Introduction

The development of Artificial Intelligence (AI) based on Human Action Recognition (HAR) lays the foundation of a wide range of applications. HAR aims to recognize human behavior and classify each action. The automatic recognition of a human action is a requirement in many application domains: industrial, medical, surveillance, personal assistance, domestic, robotics, education, etc. Gestures are considered as a natural expression of the human body. Moreover, a gesture could be classed into : body gestures, head or facial gestures (including shaking head, winking, eye gaze, nodding..), and the third class arm or hand gestures (recognition of the hand position). Automatically recognizing hand gestures, is among the most interesting research topics in the quest of efficient Human Computer Interaction (HCI). A recent study [1] mentions that 21 % of people use the hand for communication with a machines, peoples or devices. The recent release of sensors, such as the Leap Motion Controller (LMC), enables the acquisition of a very informative and accurate description of the hand pose and motion that can be used for precise gesture recognition. According to the manufacturer, this instrument is tiny in size (only 80 x 30 x 12.7 mm). It has three infrared LEDs, which are utilized for scene brightness and two stereo cameras that record pictures at frame rates ranging from 50 to 200 frames per second. The detection precision of each fingertip position is around 0.01mm [2, 3, 4, 5]. The LMC provides information in the device's field of view covering an inverted pyramid. The sensors have a field of vision of roughly 150 degrees and are pointed along the y-axis. The LMC takes use of optical sensors as well as infrared light. Moreover, HGR applications with LMC touches various interesting area including sign language [6, 7, 8], also it has been used in robotics [9, 10], HAR [11, 12], physical exercise monitoring [13] , medical actions [13, 2] , security [14], gaming [15], car automation and controlling smart home/assisted living applications [16, 17]. Therefore, establishing a harmonious and natural HCI environment that conforms to human communication habits is a vital research topic.

Meanwhile, numerous research studies related to HGR have been conducted. We recorded more than 241.456 scientific papers which have been published between 2015 and 2022 on different scientific databases. There are several survey papers on HGR, but a few of them dealt with HGR from different modalities and discuss its impact on both static and dynamic HGR. Hence, this survey aim to perform a systematic literature review of the existing studies in HGR fields, its challenges, limits, and approaches. In the process of studying the state-of-the-art reviews, their had been no review which examined the extent of research made towards the development of possible future directions and axes that touch HGR area. In this aim, the current survey fulfills that need by evaluating current and previous work to investigate the achievements of vision-based hand gesture recognition systems produced so far. Thus, our contributions are summarized as follows:

- We conduct a detailed scientific review that covers the LMC based systems with the focus on applications.
- We extensively examine the challenges faced in the literature for HGR systems, as well as the techniques proposed to overcome these issues.
- We give an overview of multi-modal approaches used for static and dynamic HGR frameworks.
- We discuss various future directions for HGR research advances.

The rest of paper is organized as follows:

The second section present a meta-review of HGR. Section 3 present the most important fields of LMC. Section 4 focus on all of the challenges and limits faced so far on HGR research field. In section 5, we discuss approaches and techniques used for dynamic and static gesture recognition from different modalities. Section 6 gives an overview on promessing future directions for HGR field. Finally, we conclude in section 7. All of the mentioned sections are illustrated in Figure1.

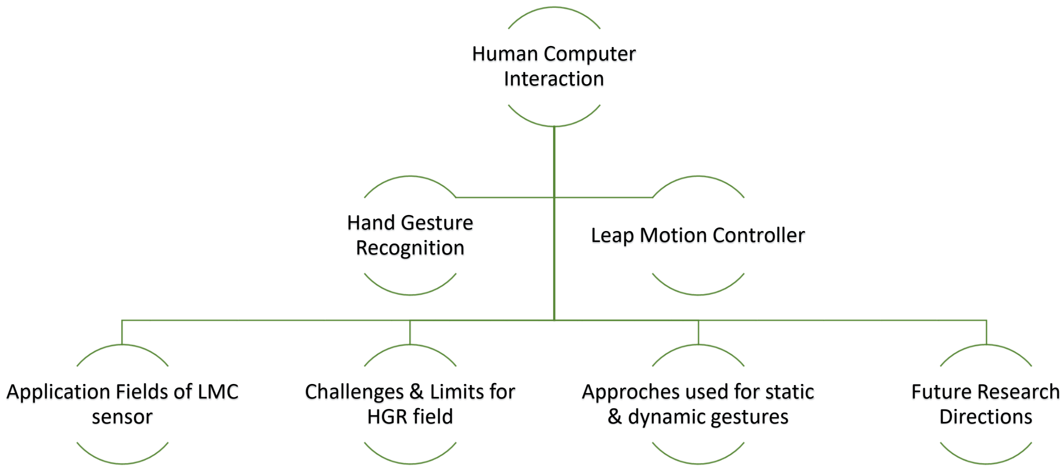

Figure 1: Hierarchy used for this survey.

2. Review of reviews

Several reviews have already been published. Each of the existing reviews highlights a particular step in gesture recognition: gathering dataset, preprocessing, training and testing techniques, and reviewing the results for each strategy. In this work, we have collected a total of 90 survey papers from many well known databases such as: IEEE Xplore, Science Direct, Springer, Wiley and MDPI for recent works on HGR. The results of this research are depicted in Figure 2. In this setting, none of the evaluations had gone into detail about the challenges or resolved all of the issues. Our study provides a broad conclusion for other researchers' reviews and informs us whether others have previously focused on the limits encountered throughout each step of the research. Furthermore, we will discuss the various difficulties that might arise within the framework of HGR, including the static and dynamic gestures with LMC. The challenges are not just connected to the method employed for preprocessing or classification, but also to the sensors or the dataset itself. Several works [18, 19, 20, 21] showed

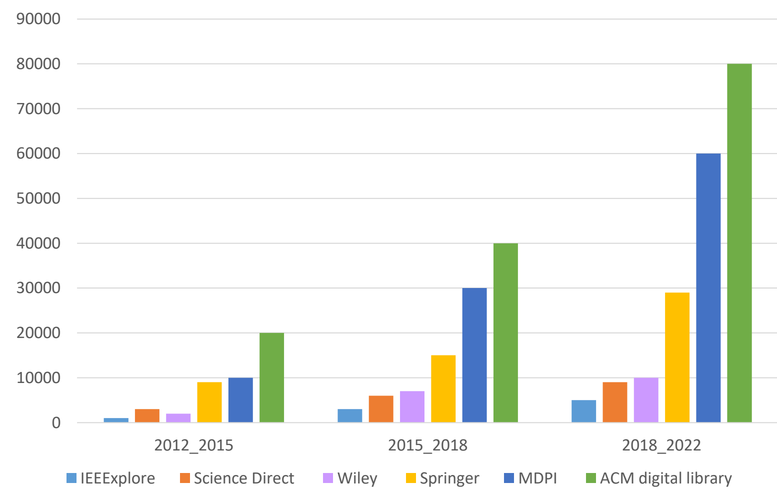

Figure 2: Recent statistic of published paper in the most popular scientific databases from 2015 to 2022 for HGR field

study cases for HGR and payed close attention to various issues encountered, whether or not they were overcame.

One of reviews have focused on HGR problem's progress using infrared data [22]. It presents a systematic literature review that retrieves information about the models' architectures, the implemented techniques in each module, also the category of learning used (supervised, unsupervised..), and recognition accuracy classification, and the processing time. However, [22] does not focused on the main challenges faced in literature in HGR field, since, it only focused on different methods and technique for several input modalities. Table1 summarize the recent review papers with the focus on their main contributions.

Ref and Year	***Domain***	***Scientific Databases***	***Contributions***
Rastgoo et al. [19] (2021)	Sign Language	Science Direct	−Review of recent works for sign language recognition. −Discussion of features, modalities, evaluation metrics, applications and datasets.
Yuanyuan et al. [23] (2021)	Dynamic Gesture recognition	IEEE Xplore	−Latest research and most advanced deep learning-based methods. −The challenges of gesture recognition from videos, and the limitations of existing methods.
Singh et al. [24] (2021)	Human Action Recognition	Springer	−An introduction to the various methods and approaches for detecting motion cues. −Accounts for the most commonly used datasets for action recognition and accuracy.
Ahmed et al. [18] (2021)	HGR + Radar senor	MDPI	−Overview of much available radar technology. −Analysis of the available literature for HGR through radar sensors.
Nogales et al. [22] (2021)	HGR + Infrared data	Springer	−HGR problem's progress using infrared information.
Colombini et al. [25] (2021)	LMC + HGR	MDPI	−An overview of existing applications of LEAP Motion for different psychological domains
Bachmann et al. [26] (2018)	HCI + LMC	MDPI	− An overview of methods for gesture design and recognition, Interaction design
Neiva et al. [27] (2018)	HGR + Mobile	Science Direct	−The authors cover static and dynamic gestures, simple and complex backgrounds facial and the use of special hardware particularly in a mobile context.
Li et al. [28] (2019)	Human Hand Motion	IEEE Xplore	−A discussion of the realistic challenges of mapping human hand motion
Wang et al. [29] (2019)	HGR	Google scholar	−Summarizing of the current technology and research results in HGR field

Table 1: Most recent survey papers in the HGR field and their main contribution extracted from most popular scientific databases.

We notice that the majority of recent survey described the domain of HGR from a specific point of view, such as field of application, new approaches and techniques, different input modalities used etc.. In this survey, we cover the main aspect of HGR system missed in literature including challenges faced and solutions proposed.

3. Scope of LMC applications

The LMC has numerous applications in a variety of domains. To allow contactless device control while reducing the need for attentive button pressing and averting eye attention. Furthermore, there has been a rapid advancement in many fields with the use of LMC including: robotics [30, 3, 31], medical tasks [2], Physical rehabilitation restore physical exercises [13, 14], house control [17, 32], education [11], sign language [6, 7, 16, 33, 34], gaming [4], and in-vehicle menu control to eliminate the need of looking for controls while driving [4] .

3.1. Leap motion in Robotic

The development of the LMC currently drives considerable innovation in robotics research. Developers analyze the hand movements and fingers and use this data to generate robotized technologies. In [9], a smart mobile robotic arm based on gesture control was proposed, consisting of a remote control smart automobile mounted with a robotic arm. The LMC have been used also a processing API function have been employed for the extraction of information of both hands in 3D space and translate it into other control data. The left-hand gesture signal was used to control the car's movement, while the right-hand gesture signal was processed into 5 driving equipment Pulse Width Modulation (PWM) signals to control the robot arm's elbow and wrist motion and grab. Li *et al.* [35] worked on gesture recognition based on mapping methods, the strategy is to develop a clear connection between human and robotic hand positions for teleportation. They attempt to make a connection between vision-based hand movements motion estimation and teleportation, which have been achieved with skeleton device CyberGlove for the caption of hand motion. Two distinct evaluation metrics were used. The first metric describes the entire estimation error across all joints, whereas the second metric represents the robustness of the outlier joints. However, the accuracy of the predicted human hand motions cannot be rigorously guaranteed. Thus, the authors proposed the use of infrared device capture the LMC. Since, the theoretical accuracy of LMC in fingertip detection is approximately 0.01 mm. Similarly, Boyali *et al.* [36] used Leap Motion to drive a robotic wheel chair. The sensor acquired five gestures of the left hand and then translated them into commands for controlling the power wheelchair. The device was successfully tested the device in a hall with tables and chairs.

3.2. Leap motion for house control

Home automation aims at centralizing the command of systems and subsystems at home. Moreover, home automation systems have gained popularity in recent years as a means of making people's lives more easy and more comfortable. About 15% of the world's population lives with some form of disability [32] and daily tasks may be challenging for them. Zaiti *et al.* [37] collected people's gesture patterns for communicating with TV by capturing 3-D finger movements and hand attitudes using the LMC as a tracking device. Furthermore, they present a set of five gesture measurements that characterize the spatial and kinematic aspects of leap gestures, including volume gesture, length gesture, and finger-to-palm distance. A set of design guidelines for gesture interfaces for the TV set was made and a dataset collection was made with LMC sensor composed of 378 distinct samples collected from 18 participants. Sant *et al.* [38] created an efficient system for controlling electrical home appliances in real time using a leap motion sensor. The LMC by default uses a euclidean distance algorithm to differentiate the different gestures of the human hand. The action depends on the gesture and the matching character for that particular gesture. Many other researchers [19, 21, 21] worked on smart house control based on gesture recognition. A gesture is defined as a mental concept of an idea associated with an action, the interaction between the user and home devices is literally a new concept. Hence, many researchers want to enlarge the search and the application of it. Home automation systems were experiencing, however have various challenges. Each person, aged people or paralysed individual, has their unique comfort zone and hand motion with specific frame rate when making a hand gesture. Hatwar *et al.* [39] that have developed a speech and gesture recognition system to help aged people controlling there houses. A sophisticated hardware system has been designed for speech recognition then gesture recognition. Different gestures were converted into signal features that can be used to create a command to control home appliances. Dinh *et al.* [40] proposed a new HGR and HCI system that recognizes each hand component in a hand depth silhouette and generates control commands for smart home devices. They gathered there own dataset using depth sensor then they train a random forest classifier. An average recognition rate of 98.50% from five different subjects was finally obtained.

3.3. Leap motion for sign language

Nonverbal communication is crucial in our lives since it conveys around 65% of messages, compared to verbal communication, which contributes just 35% of our interactions [41]. Other study [42] has shown that nearly 2.5 billion people are projected to have some degree of hearing loss and at least 700 million will require hearing rehabilitation. Thus, there are many people who suffer from communication challenges. The most widely and common communication form is sign language. It consists of a sequence of movements to express vocabulary that forms the overall notion of the sentence to be expressed. LMC is able to capture gestures used for sign language that could be static or dynamic. Enikeev *et al.* [33] conducted a study in which the LMC was employed to create an American sign language recognition system. The device identifies the user's hands and converts them into a three-dimensional representation, including the location and posture of the user's hands and thumbs. The authors faced some difficulties in sign language recognition system which is hand segment extraction because it depends on background color's that present uncontrolled challenge for HGR systems. However, the proposed algorithm have been successfully classified different hand gestures images with a 95% of precision based on a convolutional neural network. Kumar *et al.* [43] presented a new multi-sensor fusion that combines the Microsoft Kinect and LMC for HGR in the sign language field. When the Microsoft Kinect and LMC are coupled, the total recognition performance improves. Moreover, the LMC has been kept below the hand, while the Kinect is positioned in front of the signer to capture the horizontal and vertical movement of fingers during sign gestures. Moreover, when data from both sensors is fused, accuracy can be enhanced compared to single sensor-based recognition. Furthermore, a deep survey [19] have been published recently that focused on the importance of sign language domain with the use of LMC with the focus on various techniques. The authors summarized the vision based sign language recognition models corresponding to the achieved results with different datasets gathered generally from infrared sensors like the LMC. There are other several type of sign languages used all over the world as HGR field including: Myanmar sign language [6], Croatian sign language [16], Australian sign language, Arabic sign language [7], British sign language, etc.

3.4. Leap motion for health care

Healthcare is one of the fields where contact-free interaction is critical since it reduces the danger of infection especially with diseases and viruses like covid-19. Applications that could benefit from LMC in this domain include those used for rehabilitation, surgery, telemedicine, physiotherapy, medical visualization. Therefore, many studies make use of LMC to enhance appropriate interventions in healthcare settings. In [2], the authors took advantage of the LMC and worked extensively on commanding medical imagery with hand gestures while surgery. This method reduce and save time while its application also it prevent contamination during the surgery. They used hand movements to change the images or to zoom in to focus on tumour location. The dataset gathered with the LMC was inspired by actions used widely in the surgery room. To date, several studies have been focused on the effectiveness of LMC as a touchless device for medical application in various disease including Central Nervous System (CNS) disease that present a critical brain dis-function that lead to movement disorder such as Parkinson's disease. An LMC was utilized in [44] for the recovery of patients from CNS. Results showed that the LMC was a beneficial and effective tactile VR device for enhancing numerous areas of upper extremity motor function in CNS patients. In addition, a neurological disorders like Cerebral Palsy which obstructs physical and mental development of children are very tough to cure. Using LMC gaming helps many categories of disabled people to interact fluently and correctly. Furthermore, an LMC can be used in physical rehabilitation such as measuring flexibility of fingers and wrist, physical exercises targeted at rehabilitating the fingers and wrist range of motion were recognized by the LMC. Kavian *et al.* [13] developed an application which records joints location from a LMC. These data were used to generate two sets of features. The first category has been used to recognize hand postures, while the second was used to recognize exercises.

A leap motion could be used for predicting neurological diseases including autism spectrum disorder with the hand movement since autistic people have some special movements like hyperactivity motions or gesture that helps diagnosing such disorders [25].

3.5. Leap motion for Virtual Reality

Virtual reality (VR) is an artificial environment that is created with software and presented to the user in such a way that the user suspends belief and accepts it as a real environment. VR refers to the ability to communicate with the digital world in order to generate an experience that is not achievable in regular physical reality. Thus, with a new hand-tracking system, LMC is aiming to make VR more intuitive than ever before. LMC allows more natural interactions with digital contents, via an optical mode tracking of hand and finger movements. The data coming from the LMC allow users to interact within a virtual environment in a touchless way, by using natural hand gestures as input commands. LMC has been integrated into gamed virtual worlds to engage users in performing a particular task, which is specifically intended to evaluate or strengthen compromised functions. Also we found LMC application for virtual chemistry laboratory [34]. LMC contributed to some projects in Augmented Reality (AR), and Mixed Reality (MR) technology, as new interactive personal computers. Rakib *et al.* [45] make a fusion of various fields based on VR which are physical rehabilitation for Post-Stork patient and gaming with the use of LMC. They proposed an innovative approach based on interactive games called VR rehabilitation that has more advantages and reported to have efficient effect on the upper limb function of post stroke patients. The rehabilitation system comprised of leap motion sensor integrated with Unity software, the LMC is needed to engage the subjects with the virtual environments and the Unity software is used to build an interesting VR environment. Three games that focus on hand and finger motion have been created. The games are called as space game, piano game and apple game. Those games enable patient to practice move, open and close their finger as they need. The hardware was chosen because it is appropriate for detecting hand and finger motion in post-stroke patients and has a shorter distance range. Both the LMC and the software were successfully integrated to provide a completely functional VR interface. Many other researchers employed the LMC sensor for VR and AR systems and achieved significant results [46, 47, 48, 49].

3.6. Leap motion for Entertainment

This section highlights entertainment ideas with the LMC, which provides sensor-assisted interaction interfaces for entertainment. Many authors explore the suitability of this input device for interactive entertainment with a focus on usability, user engagement, and personal motion control sensitivity, and compare it with traditional keyboard and mouse controls. In the same setting, Miao *et al.* [11] presented a successful and low-cost method for preschool teaching. They demonstrate a computer-assisted toy block constructing method. The device, which is equipped with a leap motion controller, allows users to handle blocks in the virtual scene in a free-hand manner, giving them a natural experience. The LMC provides a novel gaming experience. Simply, one must assign actions to various movements and make them conform to what the favorite video games demand. Furthermore, many researchers investigate how to enhance playing music from such a device LMC. In [50] the leap motion's introduction of a hand class SDK is used to obtain the position of the hand, and the dominant hand may be employed to adjust the pitch, which can be identified automatically. In the Musical Instrument Digital Interface standard, the pitch is defined as the number of musical notes numbered from zero to 88, with 12 tones for each octave. Thus they proposed an integration of leap motion optical hand monitoring to play music.

4. Challenges related to HGR

Any issues or difficulties related to HGR especially during data acquisition it will have harmful effects on users experience. A recent work on the challenges of HGR with LMC was presented by Ameur *et al.* [51]. They summarized the most critical difficulties as well as some proposed solutions faced in literature. Confusions in gesture recognition derive not only from the problem of defining the motion of hand, but also from a plethora of different real-world challenges [13, 52, 40]. Thus, in this survey, we will develop a comprehensive discussion that covers challenges faced in HGR with LMC as well as new issues related to the LMC sensor for hand gesture data acquisition.

4.1. Hand challenges

For HGR systems the hand is the main item that any problem related to it can lead to crucial issues while collecting data with the hand tracking device LMC.

4.1.1. Size

HGR remains challenging due to its extraordinary complexity especially while gathering dataset. Moreover, hand size proved an important prior for robust HGR system, much in the way that it has been shown for whole body tracking. Hand size varies depending on age, genders, and finger flexibility which make the recognition process more complicated [2, 1]. As a consequence, researchers should enhance existing approaches for tracking and controlling these changes using hand models. Several research have addressed this issue, Ameur *et al.* [51] showed a case of study for different hand models provided from LMC SDK. They noticed that hand shape present a major effect in hand modeling then hand classification. Furthermore, Khamis *et al.* [53] developed a model based on realistic human hand shapes and poses with LMC. They have extracted skeleton and depth data with LMC. The model was built from a set of noisy depth images of a diverse set of subjects performing different poses with their hands with a variety of shapes. Result have demonstrated the effectiveness of the model proposed. Moreover, Colombini *et al.* [25] indicated negative feedback addressing Leap Motion's lack of sensitivity with super young children's hands that were too small to be correctly detected. As a consequence, they required to adjust its abilities to fit with child's characteristics with the introduction of other depth sensor like Kinect. An other solution have been proposed in [54, 55] which was the normalization of the features values by describing the enlargement of hand shape. Lu *et al.* [54] extract five Euclidean distances between the fingertips and the hand center as features. They were then scaled to normalize them. Likewise, Marin *et al.* [55] estimated the scale factor in the initial condition. It indicates the distance in between palm center and the middle fingertip, and it was used to standardize all feature values in the range [0,1]. In the same setting, Zhang *et al.* [56] introduced a new distance metric called the finger-earth mover's distance to distinguish hand gestures. It recently produced remarkable results by encoding 3D hand shape information from a depth map using his proposed histogram of 3D facets.

4.1.2. Hand Deformity

Some persons have significant hand contraction and deformities as a result of osteoarthritis and arthritis. However, this group is very rare, but it persists, and they are unhappily afraid to do anything yet they need to rely on others to help them. Moreover, muscle or neurological illness produce paralysis of the upper limb in an injured or elderly person. Deformity owing to any rheumatic disease, and musculoskeletal problem affecting the hand, wrist or elbow. Hence, their hand movement will differ, the duration of the action changes, agitated actions persists, in-completed movement will frequently appear, that lead to ironed data [31].

Arman *et al.* [57] showed a case of study for gesture recognition for healthy participants then a comparison of the proposed model with disabled ones. The authors used LMC-based Fizyosoft HandROM System for hand tracking. It was shown that measuring active hand and wrist ROM was possible and effective. Hand measurements were recorded by the LMC software system, hand information (right or left) were determined by the Fizyosoft system. Furthermore, the proposed method based on goniometer measurements better able to give an effective classification range for HGR with LMC device. Indeed, this technique must be done with controlled situation if they were effected by bone deformity. Yet, the challenge appears with people suffering from bones deformity. This methods doesn't overcomes this issues, so the HandROM system should have a high precision measurement. Nevertheless, LMC present the best solution for physical rehabilitation, since it is able to recognize all hand and fingers part [13, 58, 59]. Kavian *et al.* [13] presented a method for measuring the flexibility of wrist and fingers of damaged hand using leap motion camera. Leap motion was incorporated to acquire the 3D position of hand joints. From the acquired joints, using spatial-temporal features of hand joints, physical exercises targeted at rehabilitating the harmed hand fingers and wrist range of motion were recognized. The accuracy of leap motion sensor for wrist and fingers range of motion was verified against standard goniometry. Results showed the effectiveness of the LMC as bone coordinate tracking model and help the patient to make the appropriate movement for hand rehabilitation. Additionally, chronic disease that affects young people is a significant concern. Tarakci *et al.* [58] presented a study of potential efficacy of 8-week LMC program set as an upper extremity rehabilitation program by comparing conventional rehabilitation program in children and adolescents with physical disabilities such as Juvenile idiopathic arthritis that cause hand deformity. Therefore for this type of person, we should pick simple gestures for them and simplify the controlling orders. In addition, hand shaking of aged people or people with neurological disorder, hyperactive persons like autistic ones make hand movement looks differently. Thereby, many researches have been used the LMC to implement therapies for neurodevelopmental and neurocognitive disorders. Following the work of Colombini *et al.* [25] that have extensively reviewed all the neurological problems that lead to different hand gestures such as: mild cognitive impairment, autism spectrum disorder, dementia, attention-deficit and hyperactivity disorder. Those problems make a huge challenge for data acquisition step.

4.1.3. Skin color challenge

The appropriate segmentation of skin-colored objects (e.g., hands, face) against a complicated static background is a major problem in gesture recognition. The precision of skin segmentation algorithms is restricted by background objects that are similar in color to humans. Yang *et al.* [60] the authors applied a kinect sensor for gesture recognition also they used the skin color detection based on threshold technique. According to different areas the threshold value applied, the segmentation method can be divided into two groups, one is the threshold value only used in one region and the other is the threshold value used in the whole image. In this setting the threshold value is used to classify the points which have similar features on the image into the same class. This approach based on skin detection usually led to inflexibility of hand movements. However, it has much higher recognition accuracy. Hein *et al.* [6] had an input from an Red, Green, Blue (RGB) camera, they made a preprocessing stage to enhance RGB color space and transform RGB to Luminance; Chroma: blue; Chroma: red (YCbCr). Afterwards, they detected component by using component detection method. After detecting component, they found the localization of head and hand by using our proposed localization method and finally, they trained the skin component of hand feature by using Machine Learning (ML). The main challenge of this work was the used device which can't work properly in the presence of skin-colored objects in the background. Even with skin color-based enhancement method and color-based segmentation method for detecting skin color of hands and the system based on ML they still have some difficulties

to detect hand color variation. This issue can also result in erroneous results in a number of cases. Thus, if the background is constructed from any skin portion the model will have difficulties for recognizing and understanding the source of the gesture, resulting in the misclassification of the motion,to overcome this limitation, colorful markers or data gloves have been used to improve visual tasks[6, 61]. In addition, to distinguish between the background and the hands, gloves are occasionally employed to avoid skin detection for HGR, especially when the background contains a person with hand movement. However, the gloves must be flexible enough to allow for simple hand motions, which is not always the case, so the person who must move his fingers in a flexible manner has a hurdle [1].
All challenges related to the user's hand are depicted in Table.2 with some suggested solutions faced in literature.

Main challenge	Sub challenges	Ref and Year	Solutions
Hand challenges	Hand size	Hein *et al.* [6] (2021) Ameur *et al.* [51] (2020) Khamis *et al.* [53] (2015) Colombini *et al.* [25] (2021) Lu *et al.* [54] (2016) Zhang *et al.* [56] (2016)	−Feature normalization −Adding Kinect
	Hand deformity	Tarakci *et al.* [58] (2019) Kavian *et al.* [13] (2020) Rubio *et al.*[59] (2022)	−Intel realsense −Artificial arm
	Skin color	Hein *et al.* [6] (2020) Yang *et al.* [60] (2013) Cheng *et al.* [61] (2016)	−Segmentation −Multi LMC −Data gloves

Table 2: Hand challenges faced in literature with suggested solutions to each problem's requirement.

4.2. LMC sensor challenges

Several challenges could be caused by the acquisition device itself [13, 15]. We usually look for the right sensor that fulfill all the users needs a perfect cost and an easy way to work with. For HGR, many devices are used and achieved high accuracy. Since the LMC is the chosen device to work on for HGR area we should enumerate it limitations.

The LMC is a low-cost commercial marker-less optical sensor that really can track a human hand's motion by recording various parameters. It presents a new technology specifically designed for hand gesture recognition but it still have some restrictions. The major challenges related to the LMC sensor are depicted in Figure 3. which show the LMC coordinate system, field of view, touch zone.

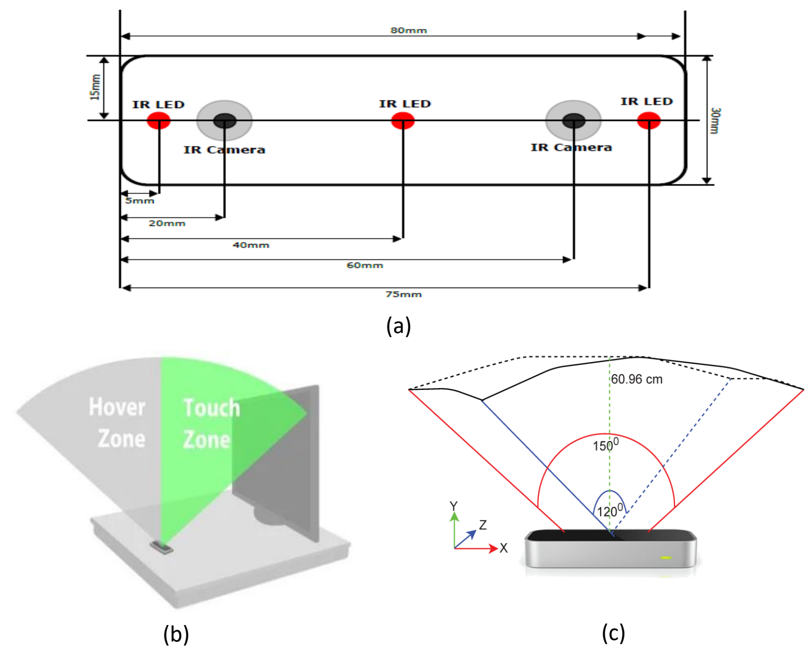

Figure 3: Hardware system major challenges: (a) Schematic view, (b) touche zone , (c) LMC coordinate system and field of view [4]

4.2.1. Frame Rate

A frame rate is the rate at which the Leap Motion system provides data frames per second. Thus, the rate at which the same hand gesture is executed varies with each repetition, based on the user's speed as well as the computer's power and settings. Each person performs the move in his or her own unique manner. Furthermore, there is no guarantee that a person will always repeat the gesture at the same time. Besides, by switching the USB, from USB0.2 port to USB0.3 port the frame rate could be multiplied by one and a half [62]. Thus, the frame rate is affected by computer settings and USB versions. In general, the frame rate ranges from 20 to 200 frames per second [63], depending on the system configuration. This fluctuation in frame rate should be considered in a gesture recognition system [54].

4.2.2. Position

All contactless technologies and devices had a calculated field of vision, which was also known as the interaction zone. With the focus on the LMC sensor, this device scans a zone shaped like an inverted pyramid and stretching upwards from the sensor's center [31]. Because of the LMC field of view it should have a specified direction to identify the gesture, we notice that not all positions and orientations are adequate for data acquisition [2]. An attractive case of study in [51] showed that the position of LMC played an efficient role for hand tracking. The authors tried five positions of LMC : placed up right, on it side, rotated 90 degree on y-axe, placed on an angle and in a head mounted. Results showed that the best position of LMC was the first one since it help to cover all the hand. Belanova *et al.* [49] showed a study of the robustness of LMC to track hand correctly. They concluded that the best position of leap motion was the head mounted which was useful for for patient in physical rehabilitation system. The leap motion is intended to be placed on a desk, but what if the user wants to change the orientation. Many studies in the literature discuss the LMC interaction area while introducing sensor technology, but only a few of them investigate its impact on recognition rate. The authors in [2] have discussed the problem with the LMC position, as this will change the device's perspective of the hands. As a result, they conduct some testing to identify the best direction and positioning for optimizing the LMC response. Therefore, in every application, it is critical to transfer the coordinate values received from the LMC to the right coordinate system.

4.2.3. Field of view

The LMC has a maximum field of view of 150°, which covers an inverted pyramid with the vertex placed in the middle of the device: the effective detection range roughly extends from 0.25 to 60 cm above the device [64] . As declared by the manufacturer, the controller has a maximum frame rate of 200 fps and an accuracy of 0.01 mm in the fingertip position detection. Thus, many orientations may be examined to determine the optimal one for obtaining an accurate pose for hand or the best orientation for acquiring hand motions. Because of the limited range of view, we must select an appropriate orientation for the hand depending on the field of view of the device. This is where the issue arises. When motion originates from diverse angles, certain techniques fail in some circumstances. Furthermore, it has been discovered to be weaker when the hand is positioned perpendicular to the device [65, 19, 64, 66]. The LMC sensor creates a set of hand representations, and each hand loss occurs when participants move their hands outside the LMC visible zone that we cannot really see it.

All challenges related to the LMC sensor are summarized in Table 3.

4.3. Scene challenges

4.3.1. Light condition

The Leap hardware includes two infrared cameras to obtain a stereoscopic picture in order to detect movement and analyze distance. There are also three infrared LEDs strategically positioned to make a great and wide field of view. Normally, the only objects apparent are those that have been directly lighted by the LMC's infrared LEDs. Yet, the quality of hand gesture modeling can be greatly influenced by the light condition. Consequently, hand gesture may appear differently in different lighting situations for the same person. This has a major impact on the HGR system's performance. Particularly, sunlight

Main challenge	Sub challenges	Ref and Year	Solutions
LMC challenges	Frame rate	Lu *et al.* [54] (2016) Guzsvinecz *et al.* [62] (2019) Placidi *et al.* [63] (2022)	−Virtual glove −Two orthogonal LMC −Better hardware equipment
	Position	Belanova *et al.* [49] (2020) Ameur *et al.* [2, 51] (2020) Naydanova *et al.* [31] (2020)	−Head mounted −Desktop in front of computer screen
	Field of view	Ameur *et al.* [2, 51] (2020) Wang *et al.* [65] (2021) Islam *et al.* [64] (2020) Rastgoo*et al.* [19] (2021)	−Particular position −Adding camera

Table 3: This table resume the main challenges for the LMC sensor and the sub challenges related to it with some proposed solutions.

poses a significant obstacle for hand gesture detection and picture processing, because of its uncontrolled state, the acquisition component is often problematic [4, 67]. Thus, because of a variable lighting conditions, hand segmentation may be noisy. In [68], an SVM was used to classify 24 alphabet letters because it provide effective classification models while being trained with a small collection of data. Room lighting combines yellow and white colored lamps, which are commonly used as home lighting and have 50Lx and 25Lx power. Data was collected using LMC with a hand and sensor distances ranging from 5 to 50 cm. From a single input data, 17 frames with 17 hand vector coordinate values are generated. The results indicated that the maximum accuracy is obtained by utilizing a white light 50Lx. Attarde *et al.* [69] used the LMC to control a drone, the LMC was used in different light conditions. The five gestures were: moving right, left, moving forward backward and centre to control the flying drone. Results showed that LMC worked adequately under the less light condition.

4.3.2. Background

The contexts in which hand gestures are recorded is critical for accurately recognizing actions. Most methods presume that HGR algorithms perform well in an indoor environment with a uniform and static background. However, in an outdoor environment, this performance decreases dramatically.

- Cluttered background

Cluttered background is very challenging for reliable hand pose estimation. Mueller *et al.* [70] proposed SynthHands and EgoDexter datasets that contains annotated sequences of challenging cluttered scenes which both present new dataset that combines real captured hand motion, natural backgrounds, and virtual elements. SynthHands dataset was captured with an LMC. It records natural hand motions such as posture, skin color, form, texture, background clutter and hand-object interactions. The most common approach for an ordinary camera to overcome the cluttered background challenges, is to separate the background from the subject. As a result, many researchers have suggested color-based and region-based segmentation approaches that depends on a fixed background for front tracking and segmentation [65, 71].

- Dynamic background

It present the second background challenge for action and hand detection systems. Many solution have been proposed to overcome this challenge including segmentation which is critical in most vision-based gesture recognition systems and try to separate the hand from the background also the prefiltering technique. In [18] the author utilised the prefiltering as a prepossessing step followed by a data formatting which depends on the next step for hand gesture classification. Mapping the background is another approach that seems to be more robust than the other techniques that involve the creation of a graphical model of the background. Due to the uniform background, using a static camera makes this work easier. A simple reduction of the cluttered static background yields simply by moving object. Zhang *et al.* [66], presented a two-stage gesture detection approach based on robust hand posture estimation to deal with the problem of complex background. Furthermore, a hand pose classifier based on Fuzzy Gaussian Mixture Models is proposed to classify the gesture which performs well in rejecting the non gestures with limited numbers of non gesture training samples. The results indicate that the proposed algorithm was effective, robust to complicated backdrops, and meets real-time requirements. Thus, a number of experts and researchers suggested several solutions to the complex background problem for HGR area such as working in fixed background for gathering dataset, threshold technique to eliminate all useless data. Besides, when a dynamic gesture is presented against a dynamic background, the gesture detection suffers, and the results may be smoothed out or incorrectly classified, especially if the background contains a skin component or hand motions [72].

4.3.3. Occlusion

The LMC offers a fluid and accurate tracking mode with sub-millimeter precision. Nevertheless, when occlusion appears, it cannot ensure a consistent and accurate tracking mode. In the same context, occlusion is described as the temporary absence of human body parts caused by being behind a larger apparent diameter item or person. Thus, occlusion can be classified into two main types:

- Self-occlusion:

Self-occlusion occurs when one portion of an item is covered by another part from a certain viewpoint. Three cases could include self-occlusion. In the first case, one finger extends and one part of the finger is covered by the remaining part of the finger. The second instance is when one finger or hand is occluded by another. Finally, when the finger is blocked by the palm. For HGR self-occlusion present a critical challenge that can crucially lead to miss recognition problems in what follows including the classification process. The number of fingers employed in the hand gesture is the primary cause of self-occlusion. However, the number of hands can also be a source of occlusion, when fingers are self occluded essential data tracking will greatly reduced. Following the work of Kiselev *et al.* [73], where they used 3 LMCs for hand tracking to overcome the problem of self-occlusion. The 3 sensors were positioned in a specific area that make the participants free to move their hands in multiple views and reduce the probability of self-occlusion. Results showed that the use of multiple LMC increase the recognition rate comparing to one or two sensors, since training a classifier using features extracted from multiple view point helps the model to classify accurately. Shen *et al.* [74] proposed a multi-sensor data fusion method, by tracking precision of assembly motion in the range of target visual field. This technique based on 3 LMC sensor showed an improvement of precision while hand tracking and avoid self-occlusion issue. Kavian *et al.* [13] examined the ability of Leap Motion to handle occlusion in a case study. They evaluate the LMC in various occlusion settings. They demonstrated the stability of Leap Motion front of hand self-occlusion or two-hand occlusion. However, it was rendered ineffective when long periods of continuous occlusion occurred, which plainly has an impact on gesture recognition. Even in virtual worlds, self occlusion is a difficult problem in which elements of the hand overlap with one another. In the same sense, Ameur *et al.* [51] showed an impressive case of study for different type of occlusion including the case when two hand overlapped on each other. The LMC in this case suffer and give erroneous information about the hands and number of fingers.

Furthermore, Qingchao *et al.* [48] evaluated the LMC performance with astronaut training program, as input device. In addition, a virtual hand was created in a virtual environment using hand pose data from Leap Motion. A virtual assembly simulation platform was developed to carry on test. The experiment results showed robustness of Leap Motion, no matter hand self-occlusion or two hands occlusion.

- Occlusion by an object:

This type occurs when some bodily parts are obscured by an object from one viewpoint. Recognizing hand activity "thump up" for example and putting a physical object like a glass of water from the front view, is a difficult task to detect. When a person reaches across their body to reach the sensor, an object, or their sleeve, arm, or jewelry, can block the controller from seeing their hands clearly. This can impair tracking accuracy or completely prevent it. Furthermore, certain hand gestures need the use of both hands, which increase the difficulty. Chen *et al.* [75] proposed a vision-based two hand tracking algorithm that aims to achieve the robustness and real-time efficiency while considering occlusion, disturbance, and appearance variation of the similar actions. They achieve a success rate of 96.3% of the proposed approach. Furthermore, Deriche *et al.* [76] used two synchronised LMC in the same time to identify large vocabulary of static and dynamic gestures for Arabic sign language under challenging scenarios and results showed that the proposed technique addressed the issue of missed or partial data as a result of occlusion. Similarly, Kumar *et al.* [77] used a combination of Leap Motion and Kinect to capture data because they have similar properties. The authors combined the data from both sensors. Their proposed approach appeared to be applicable to the occlusion challenge, as overall recognition performance has been improved beyond individual sensor-based performances. In addition, Ponraj *et al.* [78] coupled the LMC with a Flex sensor to deal with occlusion difficulties. An improvement of accuracy have been achieved in occluded scenarios. The following Table 4. resume all the challenges related to the scene worked on with some suggested solutions.

4.4. Datasets challenges

We discovered difficulties on datasets in multiple studies starting with small datasets to imbalanced ones. As a result, it makes certain subsequent stages harder or even produces low results due to restricted subject number or repetition, resulting in a tiny dataset that is not stable and good enough for future work [16, 40, 80]. Some of the available datasets, including DHG, only had extremely brief samples. In addition, some datasets had noisy and incomplete motions. Furthermore, some gesture classes, such as LMDHG, are executed with only one hand [81]. To solve data insufficiency problem, Simon *et al.* used

Main challenge	Sub challenges	Ref and Year	Solutions
Scene challenges	Light condition	Yang *et al.* [67] (2020)	−Controlled light condition
		Insani *et al.*[68](2019)	
		Vadia *et al.* [69] (2019)	
	Background	Zhang *et al.* [66] (2020)	−Segmentation
		Mueller *et al.* [70] (2021)	−Gloves
		Kavian *et al.* [13] (2020)	−Thresholding
			−Subtraction
			−Colored fingers
	Occlusion	Ponraj *et al.* [78] (2018)	−Virtual gloves
		Kumar *et al.* [77] (2017)	−Multi LMC
		Deriche *et al.* [76] (2019)	−Kinect
		Chen *et al.* [75] (2013)	−Webcam
		Kiselev *et al.* [73] (2019)	−Wearable sensors
		Placidi *et al.* [63] (2022)	−New SDK version
		Avola *et al.* [79] (2019)	

Table 4: Brief summarizing of scene challenges faced in literature with the proposition of several solutions.

a multi-camera approach to estimate hand pose. They obtained 2D pose estimations of all camera views by applying a weak hand pose estimator trained by a labeled dataset of hands. Knowing the relative physical positions of cameras, the 2D estimations were converted into 3D estimations by triangulation.

Even when executing the same gestures, each individual behaves differently. Some motions appear to be the same, but in reality, tiny variations or orientations lead them to be completely different. For example, it's very rare to have the identical gesture from young or aged people this came from the previous challenge which is the hand size problem for HGR systems. Thus, for long and complex gesture, it is difficult to develop one simple model of the same motion. Moreover, the user's hand fingers are often trembling and cannot remain constant during the interaction process that lead to crucial problem for HGR. Several efforts have been made to obtain discriminative features in order to minimize inter-class similarity and intra-class variability. A robust and effective HGR strategy should account for class differences and similarities. Because of the substantial overlap between classes, this task will be more challenging for a large number of activity classes. In [82], the authors tackled from low inter-class and high intra-class variability, a multilevel temporal sampling method was proposed that is based on the motion energy of keyframes of depth sequences. As a result, long, middle, and short sequences were generated that contain the relevant gesture information. This technique enhanced intra-class similarity while reducing inter-class dissimilarity. Probabilistic approaches, including Hidden Markov Models which are statistical models that predict a set of unknown or hidden variables based on a set of data, is among the most popular and widely used solutions to minimize inter-class similarity [83]. Li *et al.* [84] worked on gesture recognition method using LMC and based on a fused dataset. They emphasised on gesture similarity as a critical difficulty and proposed spatial fuzzy matching algorithm for dynamic hand gesture recognition. Results showed that the system was better able to recognize gestures which can greatly enhance the usability of LMC. Moreover, HGR systems based on dynamic Bayesian network framework effected positively in recognizing various and similar gestures [85]. For most instances, datasets contain both inter-class similarity and intra-class variability. Many similar gestures have been demonstrated, and each gesture is performed numerous times by various subjects. Table 5. summarises the most popular public datasets with brief description and the challenges faced.

Year	Datasets	S; G;R	Challenges	Modality
2015	VIVA hand gesture	8S; 19G	Frequent occlusion for studying natural human activities in real-world driving settings	Multimodal
2017	LMDHG	21S; 13G; 2R;	Had noisy and incomplete gestures some gesture classes performed using one hand	Skeleton

2017	DHG	14G; 10R	Very short clips of hand motion	Depth
2017	SHREK 17 Track	28S; 14G	Heavy data burden	Depth-Skeleton
2017	MSRA	6S; 20G	All subject are hand righted	Depth
2017	Synth-Hands	6S	All of images were captured with an object in the hand	RGB-Depth
2018	SHSL	40S; 25G; 2R	Small dataset	RGB-Depth
2019	ASL	5S; 26G	Some gestures are considered as the same	RGB-Depth
2019	HANDS 2019	15S	Problem with hand-object interaction system	RGB-Depth
2019	Sebastian dataset	10S	Different pictures are subjected to different illumination and scale conditions. Only static gestures were considered	PGM
2020	Leap-gesturesDB	120S; 11G; 5R	Limited gestures with one input modality	skeleton
2020	FPHA	45G	Inter-subject and Intra-subject variations on style, speed, scale, and viewpoint	RGB-Depth
2020	HGM-4	26G	Only static gestures, Unbalanced dataset	RGB
2021	Myanmar sign language	10S; 35G; 20R	Skin color detection problem	RGB
2021	DHG 14-28	20S; 14G; 5R	Small dataset	RGB

Table 5: Public Datasets in HGR fields with the focus on modalities and challenges faced .

A statistical study has been carried out to focus on the most common problems faced in the literature and depicted in Figure 4.

5. Hand gesture recognition approaches

HGR has increasingly played a significant part in the field of HCI. Despite the fact that many studies have been conducted to enhance HGR methods, it still presents numerous challenges in terms of calculation time and detection

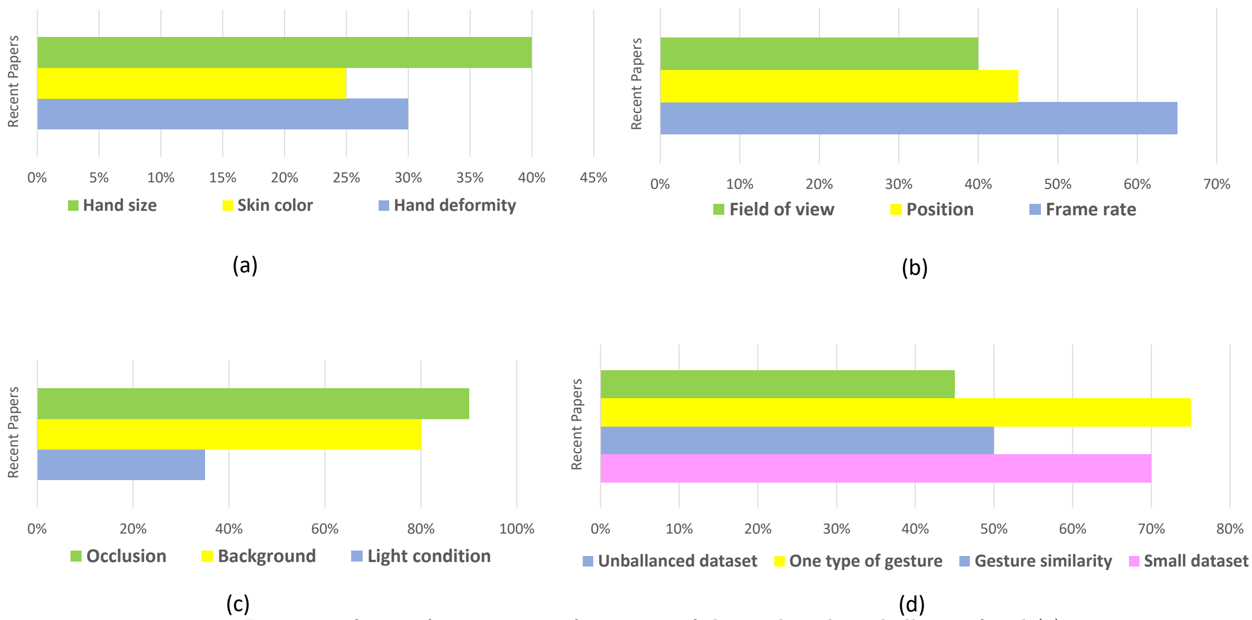

Figure 4: Partition of 139 reference papers from state of-the-art based on challenges faced (a) Challenges related to hand (b) Challenges related to LMC device (c) Challenge of the scene (d) Dataset challenges

accuracy. In general, every type of information can be thought of as a modality. An image could be classified as RGB, depth, skeleton, or optical flow data based on the type of sensor that captured it. Thereby, each of them can be thought of as a modality. Even when fundamental distinguishing elements are shown consistently across modalities, there are typically variations in the expression of details. Thus, HGR has been a highly important research area, with a diversity of methodologies presented. Each modality has an appropriate methodology. Adding to that, gestures can be divided into two categories: unimodal gestures and multimodal gestures, and these two groups can be further subdivided [7, 16]. The aim of this section is to highlight numerous approaches discovered in the literature for various gesture modalities.

5.1. RGB-based HGR approaches

The RGB modality presents the most classical and useful model for hand or object detection since it give a clear information also it facilitate prediction task. Hein *et al.* [6] worked on RGB modality for hand and head detection they transformed the RGB color space to YCbCr in the preprocessing step. The authors used color based segmentation to detect skin color and also to avoid wearing gloves or using other sensors. This technique was based on the threshold value of pixel. Furthermore, they applied component detection method for feature extraction step. A supervised ML algorithm was applied for gesture classification which is the Support Vector Machine (SVM). Jayaweera *et al.* [32] proposed watershed algorithm and use it for complex images for segmentation process. In addition a Convolutional Neural Network (CNN) was applied for gesture prediction to command house holding for bedridden people. Simon *et al.* [19] presented a real-time convolutional-based technique for hand detection from RGB group of images using a multi-view camera system. Yan *et al.* [86] developed a multi-scale CNN for unrestricted hand detection based on RGB modality. Furthermore, they applied a generic region proposal technique for HGR followed by multi-scale data fusion from the fully convolutional model. The authors combined information from different levels of a CNN model to create a multi-scale representation of hand items. In [16], a great performance was reported using RGB data and Histogram Of Gradient (HOG) method. Authors in [87] worked on RGB data captured by a web-camera for HGR purpose. For preprocessing step they transform RGB to HSV space and used segmentation as well as background subtraction technique. A high performance was reported by fusion of two 3DCNN networks for RGB and depth data in [16]. Khan *et al.* [88] used the Mask-Recurrent Neural Network (RCNN) for higher segmentation of RGB data. Further, the integration of grass hopper optimization provided a robust gesture recognition system. Mueller *et al.* [89] addressed the problem of real time 3D hand tracking based on monocular RGB modality. Moreover, they proposed a tracking method that combine CNN and Kinematic 3D hand model to have an effective system which deal with unseen data. Zimmermann *et al.* [90] proposed a 3D hand pose with the PosePrior network from colored images which increase the gesture recognition accuracy.

5.2. Skeleton-based HGR approaches

HGR based on hand skeleton tracking became the default interaction technique for the next generation of VR applications. Moreover, the skeleton data is an effective representation of human and hand motion. It is not only resistant to background influence, but it also has computational advantages because each skeleton has a minimal number of joints. As a result, much recent research has relied on 3D skeleton data in HGR field. Devineau *et al.* [91] developed a deep CNN model for HGR employing just skeletal data from the hand. They employed parallel convolutions to train sequences of hand-skeletal joint positions at various temporal resolutions. Their model proved that parallel sequence processing with CNNs can compete with neural networks that use cells particularly intended for sequence processing, such as GRU and LSTM cells. Hu *et al.* [92] used the skeleton data generated from LMC; They applied Hidden Markov Model (HMM) based approach, Fuzzy logic and CNN for dynamic gesture recognition and classification. Zhao *et al.* [93] worked on a skeleton-based dynamic HGR approach, in which the skeleton structure of the hand captured by leap motion depth sensor is firstly exploited and the spatiotemporal multi-fused features that concatenate four skeleton hand shape features and one hand direction feature are extracted. Moreover, a Fisher Vector was obtained from a Gaussian Mixture Model (GMM) that was fed into a linear SVM classifier. Smdet *et al.* [94] used an Intel RealSense camera to extract an effective descriptor from hand skeleton connected joints. They proposed a method based in Gaussian Mixture Model (GMM) used the DHG-14/28 dataset and outperformed depth based approaches. Miao *et al.* [11] showed an outperforming method based on skeleton data. Additionally, they proposed a new graph convolutional operator called central difference graph convolution for skeleton based gesture recognition. Experiments on two popular large-scale datasets NTU RGB+D 60 and 120 have demonstrated the efficacy of the proposed approach. Caputo *et al.* [95] worked with the unidirectional Gated Recurrent Units uDeepGRU model which is based on RNN. Although, RNN-based models have demonstrated great potential in sequence modeling and HGR tasks.

5.3. Depth-based HGR approaches

Many studies on HGR have been performed since the emergence of depth cameras such as LMC and kinect. Hence, a depth image is a rich data that contains a wealth of information about hand gestures [34, 4]. Several approaches have been proposed in the literature for depth-based HGR task. However, as pointed in [67], Yang *et al.* presented a new gesture recognition system based on Deep Neural Network (DNN) with depth images generated from LMC sensor. Moreover, by testing 30000 gesture frames results showed that the suggested system's recognition accuracy can reach 98%, demonstrating that the scheme can achieve the specified gesture recognition with a high average recognition rate. Sokhib *et al.* [96] worked on a combined method of Skin-and Depth-based HGR with a Kinect device as a depth sensor. The authors processed on coordinate matching and candidate hand region detection. They showed the robustness of the proposed method by training the deep model. Experiments showed the effectiveness of this modality for hand identification. Banerjee *et al.* [34] gathered data from leap motion and used as a depth image. Moreover, they preporcessed data and used the CNN architecture followed by a deep model training to have an accurate results. Adding the work of Alberto *et al.* [4], which worked on physical rehabilitation using the LMC sensor. Depth images were used to evaluate the proposed approach with the fusion of several deep recognition models. Ahn *et al.* [97] applied the depth camera to generate depth dataset and suggested a method for HGR in a smart device environment. Finger detection was made by the application of Convex Hull algorithm. An SVM was used to compare a newly identified hand motion with previously learned data. An actual smart device system was employed for experiments to evaluate the proposed method that gave an important recognition rate.

5.4. Graph-based HGR approaches

Many applications in HCI relies on data generated as graphs. Subsequently there has been a demand for considering graph presentations as a new modality for HGR field. Several studies were applied for this type of data. We encountered numerous algorithms for graph hand gesture identification in the literature. Hou *et al.* [98] suggested an hybrid design that combines a CNN with a Long Short-Term Memory (LSTM) recurrent network for graph data constructed from LMC images. Therefor, the CNN is used to extract spatial characteristics, and the LSTM recurrent network is utilized to capture patterns related to time evolution. Mirehi *et al.* [99] proposed a new hand gesture recognition approach called GNG-IEMD, Growing Neural Gas (GNG) graph to model the image and Improved Earth Mover's Distance (IEMD) metric used to mesure dissimilarity between hand gestures. They compared the performance of the proposed approach to state-of-the-art approaches on challenging public datasets such as NTU Hand Digits, HKU, HKU multi-angle, and ASL. The experimental results showed the effectiveness of the proposed method. The authors proposed a Dynamic Graph-Based Spatial-Temporal Attention (DG-STA) method for hand gesture recognition. Chen *et al.* [100] used for HGR a Dynamic Graph-Based Spatial-Temporal Attention (DG-STA) technique. The main concept was to create a fully-connected graph from a hand skeleton, and then automatically learn the node features and links using a self-attention method that works in both spatial and temporal domains. They have also proposed the use of spatio-temporal characteristics of joint locations to ensure robust recognition in difficult conditions. Furthermore, a unique spatial-temporal mask was used to dramatically reduce the computational cost by 99%. They conducted a comprehensive trials on benchmarks (DHG-14/28 and SHREC'17) and demonstrated that the proposed method outperformed state-of-the-art methods. Besides, Li *et al.* [101] proposed a weighted scheme for Elastic Graph Matching (EGM) that present a technique used for hand and object recognition, where they are represented by a labeled bunch graph. The bunch graph consists of a connected graph where the most salient features on the image are represented as series of nodes HGR. This EGM approach was used to classify ten hand gestures, and it was found that the poses were recognized with a recognition accuracy of 97.08% on average. Further, for modeling static gestures, Avraam *et al.* [102] suggested a fusion of graph-based attributes and appearance-based descriptors such as identified edges.

5.5. Multi-modal-based HGR approaches

With the diversity of modality, there is also a combination of methods used to recognize and categorize the gesture. Considering the work of Cui *et al.* [103] that created a continuous sign language recognition system employing CNN and Bi-LSTM. To acquire the representative features from CNN, they used an iterative optimization technique. The recognition model's training and tuning processes are used to increase model performance iteratively. Furthermore, this model gains from the multi-modal fusion of RGB pictures and depth data. Experiment outcomes on two public benchmarked datasets, RWTH-PHOENIXWeather 2014 and SIGNUM, demonstrate the outperforming of state-of-the-art by more than 15%. Cardenas and Chavez [104] suggested an hybrid approach for HGR that combines CNN with the Histogram of Cumulative Magnitudes (HCM). They utilized three different input modalities: RGB, Depth, and Skeleton. To include a fixed number of keyframes from the input video, a skeleton estimate method and a sampling method are used. The retrieved spatiotemporal features were fused and put into a linear SVM classifier for final recognition. While this model performed similarly to state-of-the-art approaches on isoGD and UFOP-LIBRAS datasets, it improved state-of-the-art on UTD-MHAD with a 0.16% improvement. Huang *et al.* [105] suggested a deep sign recognition model based on multi-modal inputs employing 3D CNN. To improve recognition accuracy, three input modalities, including RGB, depth, and skeletal data, were employed as multi-channel video streams. They evaluated the model using their own dataset and reported that the model was effective, with a recognition accuracy of 94.2%. D'Eusanio *et al.* [106] suggested framework based on a multimodal combination of Convolutional Neural Networks using depth and infrared pictures as input, achieving a high level of light invariance, which is critical in vision-based in-car systems. The system was tested on a recent multimodal dataset collected in a realistic automotive setting as well as a different well-known public datasets, created for the interaction between the driver and the car. The proposed technique's efficacy and real-time performance are demonstrated by experimental results on both datasets. Moreover, Jadooki *et al.* [107] have fused features mining for depth-based hand gesture recognition to classify blind human communication. They combined depth data with the color information for more reliable recognition. The proposed approach explained how the hand can be separated from the scene by depth data, then in the purpose of extracting significant features they introduced a method for combination features that gave relevant image features. Furthermore, Mahmud *et al.* [108] have extensively worked on deep learning-based multimodal data for dynamic HGR system. They proposed many approaches to deal with the high variance problem in existing multimodal fusion CRNN systems. The authors tested the strategy by using two datasets: the DHG-14/28 dataset and the SHREC'17 track dataset. This approach outperformed previous similar multimodal methods in terms of accuracy and parameter efficiency, giving results that are comparable to the state-of-the-art.

6. Discussion

HGR and feature extraction are a hard challenges in each type of discussed modalities. However, several approaches dealt with HGR could be divided according to hand gesture types. Hence, hand gesture could be gathered into two main categories: static and dynamic gestures[6, 109, 5, 110, 71]. Static gesture recognition employs a gesture image acquired at a specific point in time, with the recognition result based on the location, shape, and texture [20]. However, dynamic gestures are referred to the variation of hand movement in a period of time [23, 81, 111]. Thus, for those two main types of gestures there have been numerous methods and approaches used so far. For static hand gestures several studies showed it's effectiveness in HGR field, Jalab *et al.* [112] proposed a novel algorithm to recognize a set of static hand gestures for the HCI based on hand segmentation using both wavelet network for images feature extraction, and supervised feed-forward neural network with back propagation training algorithm for recognition. Furthermore, Pinto *et al.* [113] worked on two static datasets: ASL and there own dataset for HGR purpose. The authors used segmentation techniques for preprocessing and CNNs for classification. Thus, with the proposed methodology, they demonstrated with simple architectures of CNN an achievement of interesting results for static gesture classification. Moreover, a static HGR method deploying CNN was proposed in [114]. Consequently, Islam *et al.* proved the greatness of CNN for image representation and classification with static gestures. Further, Avram *et al.* [102] proposed a novel combination of features, classic appearance-based characteristics and graph descriptors with the use of static gestures. The authors concluded that these features are sufficient to describe in a discriminative way a gesture. whereas, other studies [10, 16, 20, 81, 98] have focused on dynamic gesture since it works in both the spatial and temporal domains and convey more interesting information about the gesture, and facilitate HGR systems. In this setting, Mahmud *et al.* [108] focused on improving the multimodal fusion approach to dynamic HGR. They extracted all the hand region of interests from the depth image frames, from both Benchmark datasets DHG14/28 and SHREC-2017. A comparison between approaches including Res-TCN, DG-STA, and pointLSTM, and their own technique was shown. The authors concluded that with a fusion of modalities and approaches named gVar-FL-Fusion contribute to great accuracy. Further, Smedt *et al.* [10] have presented an approach to recognize dynamic hand gesture as time series of 3D hand skeleton returned by the Intel RealSense depth camera. As an input the authors took a set of relevant joints inferred from 3D hand skeleton and proposed a compact representation using Fisher Vector kernel and on multi-level encoding the temporal nature of gestures. Experimental results showed the high performance of the proposed approach, with the use of the Benchmark dataset DHG14/28. Moreover, Kraljević *et al.* [16] worked in sign language field to help deaf people controlling their houses, also a multi-modal fusion of RGB and depth modality was used. They proposed sign recognition module based on the Conv3D network that have demonstrate its effectiveness in Croatian sign language dataset. We note that the use of dynamic gestures have improve HGR systems performance since it convey an informative movement.

Nevertheless, gesture classification and recognition tasks have been studied in literature as a major step in HGR system [16, 2]. Hence, for each type of modality and gestures there have been several methods used :traditional machine learning methods based on hand crafted features as well as deep learning methods based on raw input data. An important advantage of the handcrafted methods was discussed in [19] since it provide high precision. However, deep learning-based models attracted more attention in the research community in comparison with the traditional machine learning techniques. Several HGR models have been proposed in recent years using deep learning approaches, especially CNN and RNN. Using the 3D information of depth maps has led to significant improvement in this area. This is not to argue that classic computer vision techniques are no longer useful. Some issues may benefit from a trade-off between deep learning's strong capabilities (particularly when dealing with massive amounts of data) and the drawback design of handcrafted features. For that purpose, some researchers have recently suggested using the skills of both approaches. In hybrid models, a combination of a deep-based and unsupervised feature extractor, such as CNN, and a traditional classifier approach, such as SVM, has been frequently used [16, 33, 109]. Furthermore, combining certain hand-crafted characteristics with CNN-based features is a current method in hybrid models. Thus, combining prior knowledge with deep-based features can facilitate the creation of systems that are less complex and possibly more accurate in some particular areas.

We notice that, the majority of the proposed models in the HGR domain are based on depth and skeleton inputs, with only a few models suggested for RGB inputs. Using depth cameras such as LMC has enabled researchers to create large-scale datasets with automatic annotations of keypoint. While these types of sensors provide correct annotations for depth and skeleton inputs, they are inefficient for RGB inputs because they deform the representation of the hand in the RGB inputs.

Despite the fact that HGR has been intensively investigated in recent years and several models have been presented, there are still many issues that must be addressed. Because of extensive occlusions in hand key points, even accurate manual feature point annotations are impossible. As a result, providing an accurate model to learn the hand features is difficult in this domain. While hand detection is frequently the first stage in many tasks, such as action recognition and sign language recognition, robotics, it is a challenging task due to the wide range of hand shapes and gestures. Some of the most critical problems in this field are heavy occlusion, limited resolution, different lighting conditions, hand deformity, diverse hand gestures, and complicated interactions between hands and objects or other hands. Another significant challenge is the fact that a hand could grasp things, appear at different scales with closed or open palms, or hold other hands. We believe that combining features from other hand parts, with various types of input data, such as image, skeleton, graph, flow information, as well as using another deep learning model fused with traditional approaches, could help to improve HGR accuracy. Furthermore, by combining new hardware capabilities and advancements with efficient model implementations, actual hand detection in real-world applications can be reached [70, 89]. Figure 5. illustrates a statistics of recent papers based on the used technique in preprocessing, feature extraction and classification steps.

7. Future research directions for HGR:

Hand gestures are a sophisticated type of communication with numerous applications in the field of HCI. Furthermore, HGR has been employed in a variety of critical applications ranging from scientific to commercial to household. However, in order to fully exploit it, other regions can be researched in the coming future. Thus, future research can be discovered based on a comprehensive review of earlier research in HGR area. Over and above, the upcoming researches should consider the following directions for HGR domain.

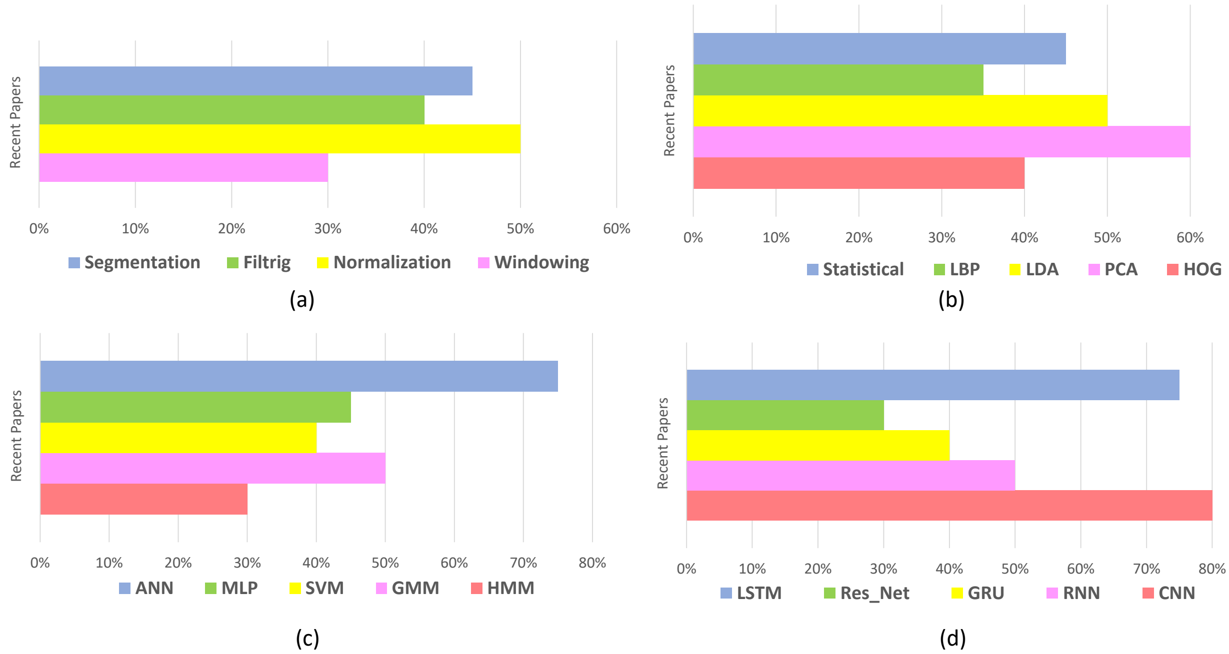

Figure 5: Partition of 139 reference papers from state of-the-art papers based on methods used in each step of a HGR approach (a) Methods used in preprocessing step (b) Feature extraction methods (c) Most used classifiers in literature (d) Deep learning approaches

7.1. Generalizability

Generalizability is the ability of HGR models to become invariant to unexpected contexts. The original objective of generalization is to improve the HGR system's performance on unseen data. Moreover, transfer learning, which includes training the system on some tasks in order to increase its performance on others, may fail in the real-world conditions due to the underspecification of these models. Underspecification is a significant barrier to the generalizability of machine learning models and refers to the pipeline's inability to ensure that the model has encoded the system's inner logic. This phenomenon is distinct from and less well understood than over-fitting in artificial intelligence, but it can be analyzed and resolved with the use of testing process [115, 116].

We notice that in a variety of HGR systems, underspecification affected how models struggle from hand data acquisition. For example, depending on the frame rate, a pipeline could produce a model that depends more or less on correlations involving gender, hand size, subject ages or hand deformity when making predictions. We believe that studies on HGR approaches have to emphasise further on the generalizabilty of the proposed frameworks.

7.2. Reproducibility

Reproducibility is an increasing concern in HGR frameworks. The HGR field evolves with the availability of public datasets and codes implementations; therefore, researchers should work more on making their knowledge reproducible [117]. Being able to reproduce a HGR models is a crucial task, as it is closely tied to various steps like pre-processing, training, testing, validating and evaluating. However, HGR models are challenging to be reproduced due to issues like randomness in the software such as the case of deep learning algorithms, and non-determinism in the hardware when using Graphical Processing Units for example. An interesting study for the reproducibility of deep learning models [78] shows a systematic approach to reproduce these models.

According to HGR domain, making a reproducible and intelligent model depends on three main elements: the code, the dataset and the environment. For the code and to guarantee reproducibility, changes in code and algorithms must be tracked and recorded during experimentation. Furthermore, adding new datasets, adjusting data distribution, and increasing sample size will all impact the outcome of a model. To achieve reproducibility, dataset configuration and change tracking must be recorded. Adding to that, the environment in which a project was created must be captured in order to be reproducible. Framework dependencies, versions, hardware, and any other environmental elements must be documented and reproducible. The environment should satisfy the several requirements such as: using all available computational power, setting randomization parameters, employing identical versions on multiple machines, using the latest library and document version, take use of the most recent library and document versions and returning to the previous state without losing the setup. Thereby, evaluating the performance and reproducibility of HGR models on real world data would definitely be interesting for future work.

7.3. Fairness

Fairness is a prevalent term in the HGR field. More precisely, the fairness in data collection, and algorithms development is critical to build safe and responsible systems. A detailed survey on bias and fairness in machine learning models [118] investigated that fairness is the process of understanding bias introduced by data, and ensuring that models provides equitable predictions across all demographic groups.

To trust HGR systems, we expect them to be fair and not discriminatory especially against classes that are legally protected, such as by race, gender, disability and more. Furthermore, we argue that it is unethical to rely merely on useful models with low error rates when the safety of a human being is at stake. Given an example of HGR RGB-based model, a highly accurate prediction will be supposed when classifying different hand gestures. Here, fairness can come into question if this RGB data actually included different hand skin colors. This fairness issue could be caused by collecting bias introduced by RGB data. Hence, to address this issue, it is essential to assure that predictions are scaled for each hand skin color group. In general, we should apply continuously fairness analysis throughout the entire development steps by identifying imbalanced datasets, treating all classes fairly, predicting thresholds, and testing model behaviour on real data.

By the perspective on fairness for HGR systems, we hope to expand the horizons of the researchers to think deeply while contributing new HGR approach to ensure that it has a low likelihood of causing potential bias or harm toward a specific class.

7.4. Federated Learning (FL)

With the development of big data technologies, the amount of data is no longer the primary focus. The most critical concern that must be addressed is rather data privacy and security which is fulfilled with federated learning. In the same setting, federated learning is a computer vision strategy that allows machine learning models to extract knowledge from different datasets without having access to them. It allows global model collaborative learning without sharing users' raw data. FL enables stronger models, lower latency, and lower power usage while maintaining privacy. In fact, each system uses local information for local training, then transfers the model update to the server for aggregation, and finally, the server delivers the model updated to the participants [28]. This setting also allows the training data decentralized to ensure the data privacy of each device. FL for HGR systems present a distributed training method, allows users in different regions to interact with other users to create machine learning models based on HGR, and all private details, including sensitive personal information like hand bone deformity, chronic brain diseases that effect hand movement, can be stored on the device. On this account, FL is a trending approach to resolve privacy issues in learning problems for HGR field. On the other hand, existing federated learning techniques, produce insufficient HGR system performance because they fail to dynamically combine models based on the statistical diversity of users' data.

7.5. Affordability

The affordability of machine learning systems refers to the model's low cost in hardware and software tools. Achieving affordability is a challenging task, from both a theoretical and practical point of view. Following the work of [119] that proposed a model-based pricing framework, which instead of pricing the data, directly prices machine learning model instances. Moreover, making HGR systems more affordable and less time-consuming with customized hardware, FPGAs can make DNN more accessible with lesser technical expertise. The affordable and accurate prediction of hand gesture plays a key role in the analysis of increasingly complex experiments of the field. Towards low cost machine learning system present a crucial research topic that researchers doesn't deeply emphasise on it. In addition, in HGR domain little work studies investigate how to reduce the cost of data acquisition, to guarantee affordability. A HGR system could be affordable with the use of low-budget data gloves that provide high performance for gesture recognition as mentioned in [120].

7.6. Causality

Causality gives us tools to address the question," *Why does something happen?* " This advances beyond standard statistical or machine learning approaches that are concerned with future associations instead of predicting outcomes. Thus, causal models also allow us to respond to situations we haven't seen before. Recent developments at the intersection of causality and HGR systems are making the discovery of such causal relationships easier [121]. Causality is also about calculating the effect of actions, and allows to transfer knowledge to new, unfamiliar situations.

7.7. Explainability

Bueff *et al.* [122] defined explainability as the ability to elucidate the reasons behind a machine learning model's prediction and how to trust the result obtained by the a given model. One could fine-tune and optimize the model if they understand why and how it works. The growing complexity of machine learning algorithms limits the ability to understand what the model has learned or why a given prediction was made. Similarly, for HGR systems, explainability must maintain in order to explain the reasons behind the model's prediction. We found two major types of explainability: global explainability that refers to understand the general behavior of a machine learning model as a whole, and local explainability that refers to understand why the machine learning model made a specific prediction. For HGR domain, explainability allows the user to understand and interpret the model's behavior. Also to select which features does the model consider important in HGR systems.

8. Conclusion

Several sensors have been employed for HGR systems for data acquisition. Researchers focus on the device accuracy, precision and cost while designing a new application. A competitive study of sensors proved that the LMC offers the best price-performance ratio for HGR systems. This led to the widespread adoption of LMC in several areas. For instance, it is mostly utilized in sign language, robotic, education, home assisting, virtual reality and medicine. However, HGR challenges still persist and affect the performance of the models.

This work presents an extensive systematic literature review for HGR domain with an introduction to the main applications of the hand tracking device (LMC), a summary of the most crucial challenges faced in literature with some suggested solutions, as well as numerous possible future directions of this impressive research field. This survey examines the HGR system from a different angle, highlighting a crucial and efficient aspect that should be considered: the main challenges that could affect the application fiability, the model performance, and the settings of the used approaches.

We enumerated the major challenges that affect HGR systems, especially challenges related to hands, sensors, scene and datasets. For hand challenge we focused on the user's hand deformity and neurological diseases which present a natural limitations that harmfully disrupt the user's interaction experience. In addition, we enumerate a set of LMC challenges, scene problems including cluttered and dynamic background, uncontrolled light condition that could be solved in some cases with the several proposed solutions faced in literature. Moreover, deficiency in data acquisition protocols, limited number of samples and few repetitions of the gestures acquired with LMC lead to unbalanced datasets. Moreover, we present recent approaches and techniques employed for HGR from preprocessing to classification of different input modalities: RGB, skeleton, depth and graph. In the last section of our survey, we give insights for unconsidered directions in HGR domain. We believe that in this survey is comprehensive of the HGR problem and hope it helps the community investigate its challenges further.

References

M. Oudah, A. Al-Naji, J. Chahl, Hand gesture recognition based on computer vision: a review of techniques, journal of Imaging 6 (8) (2020) 73.

S. Ameur, A. Ben Khalifa, M. S. Bouhlel, Hand-gesture-based touchless exploration of medical images with leap motion controller, in: 2020 17th

International Multi-Conference on Systems, Signals Devices (SSD), 2020, pp. 6–11. doi:10.1109/SSD49366.2020.9364244.

K. Nosirov, S. Begmatov, M. Arabboev, Analog sensing and leap motion integrated remote controller for search and rescue robot system, in: 2020 International Conference on Information Science and Communications Technologies (ICISCT), 2020, pp. 1–5. doi:10.1109/ICISCT50599.2020.9351425.

C. A. Aguilar-Lazcano, E. J. Rechy-Ramirez, Performance analysis of leap motion controller for finger rehabilitation using serious games in two lighting environments, Measurement 157 (2020) 107677. doi:https://doi.org/10.1016/j.measurement.2020.107677.
URL https://www.sciencedirect.com/science/article/pii/S0263224120302153

S. Ameur, A. B. Khalifa, M. S. Bouhlel, A comprehensive leap motion database for hand gesture recognition, in: 2016 7th International Conference on Sciences of Electronics, Technologies of Information and Telecommunications (SETIT), 2016, pp. 514–519. doi:10.1109/SETIT.2016.7939924.

Z. Hein, T. P. Htoo, B. Aye, S. M. Htet, K. Z. Ye, Leap motion based myanmar sign language recognition using machine learning, in: 2021 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), 2021, pp. 2304–2310. doi:10.1109/ElConRus51938.2021.9396496.

C.-C. Lee, Z. Gao, Sign language recognition using two-stream convolutional neural networks with wi-fi signals, Applied Sciences 10 (24). doi:10.3390/app10249005.
URL https://www.mdpi.com/2076-3417/10/24/9005

H. Fasihuddin, S. Alsolami, S. Alzahrani, R. Alasiri, A. Sahloli, Smart tutoring system for arabic sign language using leap motion controller, in: 2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE), 2018, pp. 1–5. doi:10.1109/ICSCEE.2018.8538425.

H. He, Y. Dan, The research and design of smart mobile robotic arm based on gesture controlled, in: 2020 International Conference on Advanced Mechatronic Systems (ICAMechS), 2020, pp. 308–312. doi:10.1109/ICAMechS49982.2020.9310156.

Q. De Smedt, H. Wannous, J.-P. Vandeborre, Skeleton-based dynamic hand gesture recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2016, pp. 1206–1214. doi:10.1109/CVPRW.2016.153.

S. Miao, Y. Hou, Z. Gao, M. Xu, W. Li, A central difference graph convolutional operator for skeleton-based action recognition.

X. Jiang, k. xu, T. Sun, Action recognition scheme based on skeleton representation with ds-lstm network, IEEE Transactions on Circuits and Systems for Video Technology PP (2019) 1–1. doi:10.1109/TCSVT.2019.2914137.

M. Kavian, A. Nadian-Ghomsheh, Monitoring wrist and fingers range of motion using leap motion camera for physical rehabilitation, in: 2020 International Conference on Machine Vision and Image Processing (MVIP), 2020, pp. 1–6. doi:10.1109/MVIP49855.2020.9116876.

M. Luimula, J. Ranta, M. Al-Adawi, Hand tracking in fire safety - electric cabin fire simulation, in: 2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), 2020, pp. 000221–000222. doi:10.1109/CogInfoCom50765.2020.9237818.

A. Dzikri, D. E. Kurniawan, Hand gesture recognition for game 3d object using the leap motion controller with backpropagation method, in: 2018 International Conference on Applied Engineering (ICAE), 2018, pp. 1–5. doi:10.1109/INCAE.2018.8579400.

L. Kraljević, M. Russo, M. Pauković, M. Šarić, A dynamic gesture recognition interface for smart home control based on croatian sign language, Applied

Sciences 10 (7). doi:10.3390/app10072300.
URL https://www.mdpi.com/2076-3417/10/7/2300

P.-J. Gonzalo, A. Holgado-Terriza Juan, Control of home devices based on hand gestures, in: 2015 IEEE 5th International Conference on Consumer Electronics - Berlin (ICCE-Berlin), 2015, pp. 510–514. doi:10.1109/ICCE-Berlin.2015.7391325.

S. Ahmed, K. D. Kallu, S. Ahmed, S. H. Cho, Hand gestures recognition using radar sensors for human-computer-interaction: A review, Remote Sensing 13 (3). doi:10.3390/rs13030527.
URL https://www.mdpi.com/2072-4292/13/3/527

R. Rastgoo, K. Kiani, S. Escalera, Sign language recognition: A deep survey, Expert Systems with Applications 164 (2021) 113794.

S. Yuanyuan, L. Yunan, F. Xiaolong, M. Kaibin, M. Qiguang, Review of dynamic gesture recognition, Virtual Reality & Intelligent Hardware 3 (3) (2021) 183–206.

P. K. Singh, S. Kundu, T. Adhikary, R. Sarkar, D. Bhattacharjee, Progress of human action recognition research in the last ten years: A comprehensive survey, Archives of Computational Methods in Engineering (2021) 1–41.

R. E. Nogales, M. E. Benalcázar, Hand gesture recognition using machine learning and infrared information: a systematic literature review, International Journal of Machine Learning and Cybernetics 12 (10) (2021) 2859–2886. doi:10.1007/s13042-021-01372-y.
URL https://doi.org/10.1007%2Fs13042-021-01372-y

Y. SHI, Y. LI, X. FU, M. Kaibin, M. Qiguang, Review of dynamic gesture recognition, Virtual Reality Intelligent Hardware 3 (3) (2021) 183–206. doi:https://doi.org/10.1016/j.vrih.2021.05.001.
URL https://www.sciencedirect.com/science/article/pii/S2096579621000279

P. Singh, S. Kundu, T. Adhikary, R. Sarkar, D. Bhattacharjee, Progress of human action recognition research in the last ten years: A comprehensive survey, Archives of Computational Methods in Engineeringdoi:10.1007/s11831-021-09681-9.

G. Colombini, M. Duradoni, F. Carpi, L. Vagnoli, A. Guazzini, Leap motion technology and psychology: A mini-review on hand movements sensing for neurodevelopmental and neurocognitive disorders, International Journal of Environmental Research and Public Health 18 (8). doi:10.3390/ijerph18084006.
URL https://www.mdpi.com/1660-4601/18/8/4006

D. Bachmann, F. Weichert, G. Rinkenauer, Review of three-dimensional human-computer interaction with focus on the leap motion controller, Sensors 18 (7). doi:10.3390/s18072194.
URL https://www.mdpi.com/1424-8220/18/7/2194

D. H. Neiva, C. Zanchettin, Gesture recognition: A review focusing on sign language in a mobile context, Expert Systems with Applications 103 (2018) 159–183.

Q. Li, Z. Wen, Z. Wu, S. Hu, N. Wang, X. Liu, B. He, A survey on federated learning systems: Vision, hype and reality for data privacy and protection, CoRR abs/1907.09693. arXiv:1907.09693.
URL http://arxiv.org/abs/1907.09693

T. Wang, Y. Li, J. Hu, A. Khan, L. Liu, C. Li, A. Hashmi, M. Ran, A survey on vision-based hand gesture recognition, in: ICSM, 2018.

A. Koenig, F. R. y Baena, R. Secoli, Gesture-based teleoperated grasping for educational robotics, in: 2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN), IEEE, 2021, pp. 222–228.

E. Naydanova, M. J. Kim, B. Y. Hwang, K. A. Mills, W. S. Anderson, Y. Salimpour, Objective evaluation of motor symptoms in parkinson's disease via a dual system of leap motion controllers, in: 2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE), IEEE, 2020, pp. 826–829.

N. Jayaweera, B. Gamage, M. Samaraweera, S. Liyanage, S. Lokuliyana, T. Kuruppu, Gesture driven smart home solution for bedridden people, in: 2020 35th IEEE/ACM International Conference on Automated Software Engineering Workshops (ASEW), 2020, pp. 152–158. `doi:10.1145/3417113.3422998`.

D. Enikeev, S. Mustafina, Recognition of sign language using leap motion controller data, in: 2020 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA), 2020, pp. 393–397. `doi:10.1109/SUMMA50634.2020.9280795`.

T. Banerjee, K. P. Srikar, S. A. Reddy, K. S. Biradar, R. R. Koripally, G. Varshith, Hand sign recognition using infrared imagery provided by leap motion controller and computer vision, in: 2021 International Conference on Innovative Practices in Technology and Management (ICIPTM), 2021, pp. 20–25. `doi:10.1109/ICIPTM52218.2021.9388334`.

R. Li, H. Wang, Z. Liu, Survey on mapping human hand motion to robotic hands for teleoperation, IEEE Transactions on Circuits and Systems for Video Technology 32 (5) (2022) 2647–2665. `doi:10.1109/TCSVT.2021.3057992`.

A. Boyali, N. Hashimoto, Hand posture control of a robotic wheelchair using a leap motion sensor and block sparse representative classification method.

I.-A. Zaiti, S. G. Pentiuc, R.-D. Vatavu, On free-hand tv control: experimental results on user-elicited gestures with leap motion, Personal and Ubiquitous Computing 19 (2015) 821–838.

S. K. S. C. Santhanakrishnan, Thejeshwar ReddyYerasi, Hand gesture based home automation using leap motion sensor, International Journal of Advanced Science and Technology 29 (6s) (2020) 1926 – 1931. URL `http://sersc.org/journals/index.php/IJAST/article/view/9360`

P. D. Hatwar, N. A. Wahile, I. M. Padiya, Home automation system based on gesture recognition system, International Journal of Emerging Technologies in Engineering Research (IJETER) 5 (3).

D.-L. Dinh, J. Kim, T.-S. Kim, Hand gesture recognition and interface via a depth imaging sensor for smart home appliances, Energy Procedia 62. `doi:10.1016/j.egypro.2014.12.419`.

N. Mohamed, M. B. Mustafa, N. Jomhari, A review of the hand gesture recognition system: Current progress and future directions, IEEE Access 9 (2021) 157422–157436. `doi:10.1109/ACCESS.2021.3129650`.

D. Hirafuji Neiva, C. Zanchettin, Gesture recognition: A review focusing on sign language in a mobile context, Expert Systems with Applications 103 (2018) 159–183. `doi:https://doi.org/10.1016/j.eswa.2018.01.051`. URL `https://www.sciencedirect.com/science/article/pii/S0957417418300654`

P. Kumar, H. Gauba, P. Pratim Roy, D. Prosad Dogra, A multimodal framework for sensor based sign language recognition, Neurocomputing 259 (2017) 21–38, multimodal Media Data Understanding and Analytics. `doi:https://doi.org/10.1016/j.neucom.2016.08.132`. URL `https://www.sciencedirect.com/science/article/pii/S092523121730262X`

I. Cortés-Pérez, N. Zagalaz-Anula, D. Montoro-Cárdenas, R. Lomas-Vega, E. Obrero-Gaitán, M. C. Osuna-Pérez, Leap motion controller video game-based therapy for upper extremity motor recovery in patients with central

nervous system diseases. a systematic review with meta-analysis, Sensors 21 (6). doi:10.3390/s21062065.
URL https://www.mdpi.com/1424-8220/21/6/2065

N. F. Rakib, N. H. Mahmood, N. Ramli, N. A. Zakaria, M. A. A. Razak, Preliminary results of hand rehabilitation for post stroke patient using leap motion-based virtual reality, in: 2020 IEEE Student Conference on Research and Development (SCOReD), 2020, pp. 259–262. doi:10.1109/SCOReD50371.2020.9250985.

N. F. Rakib, N. H. Mahmood, N. Ramli, N. A. Zakaria, M. A. A. Razak, Preliminary results of hand rehabilitation for post stroke patient using leap motion-based virtual reality, in: 2020 IEEE Student Conference on Research and Development (SCOReD), 2020, pp. 259–262. doi:10.1109/SCOReD50371.2020.9250985.

F. Wijaya, Y.-C. Tseng, W.-L. Tsai, T.-Y. Pan, M.-C. Hu, Vr piano learning platform with leap motion and pressure sensors, in: 2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), 2020, pp. 584–585. doi:10.1109/VRW50115.2020.00143.

X. Qingchao, C. Jiangang, The application of leap motion in astronaut virtual training, IOP Conference Series: Materials Science and Engineering 187 (2017) 012015. doi:10.1088/1757-899x/187/1/012015.
URL https://doi.org/10.1088/1757-899x/187/1/012015

D. BELANOVÁ, K. YOSHIDA, Hand position tracking correction of leap motion controller attached to the virtual reality headset, International Journal of Biomedical Soft Computing and Human Sciences: the official journal of the Biomedical Fuzzy Systems Association 25 (1) (2020) 29–37. doi:10.24466/ijbschs.25.1_29.

A. Sourin, Music in the air with leap motion controller, in: 2019 International Conference on Cyberworlds (CW), 2019, pp. 57–60. doi:10.1109/CW.2019.00018.

S. Ameur, A. B. Khalifa, M. S. Bouhlel, Chronological pattern indexing: An efficient feature extraction method for hand gesture recognition with leap motion, Journal of Visual Communication and Image Representation 70 (2020) 102842.

B. Yu, Z. Luo, H. Wu, S. Li, Hand gesture recognition based on attentive feature fusion, Concurrency and Computation: Practice and Experience 32 (22) (2020) e5910.

S. Khamis, J. Taylor, J. Shotton, C. Keskin, S. Izadi, A. Fitzgibbon, Learning an efficient model of hand shape variation from depth images, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 2540–2548.

W. Lu, Z. Tong, J. Chu, Dynamic hand gesture recognition with leap motion controller, IEEE Signal Processing Letters 23 (9) (2016) 1188–1192. doi:10.1109/LSP.2016.2590470.

G. Marin, F. Dominio, P. Zanuttigh, Hand gesture recognition with leap motion and kinect devices, in: 2014 IEEE International Conference on Image Processing (ICIP), 2014, pp. 1565–1569. doi:10.1109/ICIP.2014.7025313.

C. Zhang, Y. Tian, Histogram of 3d facets: A depth descriptor for human action and hand gesture recognition, Computer Vision and Image Understanding 139 (2015) 29–39. doi:https://doi.org/10.1016/j.cviu.2015.05.010. URL https://www.sciencedirect.com/science/article/pii/S1077314215001216

N. Arman, A. B. Oktay, D. Tarakci, E. Tarakci, Y. S. Akgul, The validity of an objective measurement method using the leap motion controller for fingers wrist, and forearm ranges of motion, Hand Surgery and Rehabilitation 40 (4) (2021) 394–399. doi:https://doi.org/10.1016/j.hansur.2021.03.007. URL https://www.sciencedirect.com/science/article/pii/S2468122921000955

E. Tarakci, S. Arman (BAYDOGAN), D. Tarakci, O. Kasapcopur, Leap motion controller–based training for upper extremity rehabilitation in children and adolescents with physical disabilities: A randomized controlled trial, Journal of Hand Therapy 33. doi:10.1016/j.jht.2019.03.012.

Á. Aguilera-Rubio, I. M. Alguacil-Diego, A. Mallo-López, A. Cuesta-Gómez, Use of the leap motion controller® system in the rehabilitation of the upper limb in stroke. a systematic review, Journal of Stroke and Cerebrovascular Diseases 31 (1) (2022) 106174.

S. Yang, P. Premaratne, P. Vial, Hand gesture recognition: An overview, in: 2013 5th IEEE International Conference on Broadband Network Multimedia Technology, 2013, pp. 63–69. doi:10.1109/ICBNMT.2013.6823916.

H. Cheng, L. Yang, Z. Liu, Survey on 3d hand gesture recognition, IEEE Transactions on Circuits and Systems for Video Technology 26 (9) (2016) 1659–1673. doi:10.1109/TCSVT.2015.2469551.

T. Guzsvinecz, V. Szucs, C. Sik-Lanyi, Suitability of the kinect sensor and leap motion controller—a literature review, Sensors 19 (5). doi:10.3390/s19051072.
URL https://www.mdpi.com/1424-8220/19/5/1072

G. Placidi, A. Di Matteo, F. Mignosi, M. Polsinelli, M. Spezialetti, Compact, accurate and low-cost hand tracking system based on leap motion controllers and raspberry pi., in: ICPRAM, 2022, pp. 652–659.

J. Islam, A. Ghosh, M. I. Iqbal, S. Meem, N. Ahmad, Integration of home assistance with a gesture controlled robotic arm, in: 2020 IEEE Region 10 Symposium (TENSYMP), 2020, pp. 266–270. doi:10.1109/TENSYMP50017.2020.9230893.

Y. Wang, Y. Wu, S. Jung, S. Hoermann, S. Yao, R. W. Lindeman, Enlarging the usable hand tracking area by using multiple leap motion controllers in vr, IEEE Sensors Journal 21 (16) (2021) 17947–17961. doi:10.1109/JSEN.2021.3082988.

T. Zhang, H. Lin, Z. Ju, C. Yang, Hand gesture recognition in complex background based on convolutional pose machine and fuzzy gaussian mixture models, International Journal of Fuzzy Systems 22 (4) (2020) 1330–1341.

Q. Yang, W. Ding, X. Zhou, D. Zhao, S. Yan, Leap motion hand gesture recognition based on deep neural network, in: 2020 Chinese Control And Decision Conference (CCDC), 2020, pp. 2089–2093. doi:10.1109/CCDC49329.2020.9164723.

C. N. Insani, I. Nurtanio, A. A. Ilham, The effect of light on leap motion controller in the classification of sign language translator system, in: 2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), IEEE, 2019, pp. 296–300.

G. Attarde, O. Vaidya, G. Phade, Leap motion estimation for controlling and tracking of drone 6 (2019) 12016 –12019.

F. Mueller, D. Mehta, O. Sotnychenko, S. Sridhar, D. Casas, C. Theobalt, Real-time hand tracking under occlusion from an egocentric rgb-d sensor, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 1154–1163.

A. Haria, A. Subramanian, N. Asokkumar, S. Poddar, J. S. Nayak, Hand gestur51e recognition for human computer interaction, Procedia Computer Science 115 (2017) 367–374, 7th International Conference on Advances in Computing Communications, ICACC-2017, 22-24 August 2017, Cochin, India. doi:https://doi.org/10.1016/j.procs.2017.09.092.
URL https://www.sciencedirect.com/science/article/pii/S1877050917319130

T. Kapuściński, D. Warchoł, Hand posture recognition using skeletal data and distance descriptor, Applied Sciences 10 (6) (2020) 2132.

V. Kiselev, M. Khlamov, K. Chuvilin, Hand gesture recognition with multiple leap motion devices, in: 2019 24th Conference of Open Innovations Association (FRUCT), 2019, pp. 163–169. doi:10.23919/FRUCT.2019.8711887.

H. Shen, X. Yang, H. Hu, Q. Mou, Y. Lou, Hand trajectory extraction of human assembly based on multi-leap motions, in: 2019 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM), 2019, pp. 193–198. doi:10.1109/AIM.2019.8868827.

B.-J. [2], C.-M. Huang, A.-S. Liu, T.-E. Tseng, L.-C. Fu, Hands tracking with self-occlusion handling in cluttered environment, in: 2013 9th Asian Control Conference (ASCC), 2013, pp. 1–8. doi:10.1109/ASCC.2013.6606126.

M. Deriche, S. O. Aliyu, M. Mohandes, An intelligent arabic sign language recognition system using a pair of lmcs with gmm based classification, IEEE Sensors Journal 19 (18) (2019) 8067–8078. doi:10.1109/JSEN.2019.2917525.

P. Kumar, H. Gauba, P. P. Roy, D. P. Dogra, Coupled hmm-based multi-sensor data fusion for sign language recognition, Pattern Recognition Letters 86 (2017) 1–8. doi:https://doi.org/10.1016/j.patrec.2016.12.004. URL https://www.sciencedirect.com/science/article/pii/S0167865516303518

G. Ponraj, H. Ren, Sensor fusion of leap motion controller and flex sensors using kalman filter for human finger tracking, IEEE Sensors Journal 18 (5) (2018) 2042–2049. doi:10.1109/JSEN.2018.2790801.

D. Avola, M. Bernardi, L. Cinque, G. L. Foresti, C. Massaroni, Exploiting recurrent neural networks and leap motion controller for the recognition of sign language and semaphoric hand gestures, IEEE Transactions on Multimedia 21 (1) (2019) 234–245. doi:10.1109/tmm.2018.2856094. URL https://doi.org/10.1109%2Ftmm.2018.2856094

P. Chophuk, S. Chumpen, S. Tungjitkusolmun, P. Phasukkit, Hand postures for evaluating trigger finger using leap motion controller, in: 2015 8th Biomedical Engineering International Conference (BMEiCON), IEEE, 2015, pp. 1–4.

K. Lupinetti, A. Ranieri, F. Giannini, M. Monti, 3d dynamic hand gestures recognition using the leap motion sensor and convolutional neural networks.

R. Azad, M. Asadi-Aghbolaghi, S. Kasaei, S. Escalera, Dynamic 3d hand gesture recognition by learning weighted depth motion maps, IEEE Transactions on Circuits and Systems for Video Technology 29 (6) (2019) 1729–1740. doi:10.1109/TCSVT.2018.2855416.

D. Enikeev, S. Mustafina, Recognition of sign language using leap motion controller data, in: 2020 2nd International Conference on Control Systems, Mathematical Modeling, Automation and Energy Efficiency (SUMMA), 2020, pp. 393–397. doi:10.1109/SUMMA50634.2020.9280795.

H. Li, L. Wu, H. Wang, C. Han, W. Quan, J. Zhao, Hand gesture recognition enhancement based on spatial fuzzy matching in leap motion, IEEE Transactions on Industrial Informatics 16 (3) (2020) 1885–1894. doi:10.1109/TII.2019.2931140.

H.-I. Suk, B.-K. Sin, S.-W. Lee, Hand gesture recognition based on dynamic bayesian network framework, Pattern Recognition 43 (2010) 3059–3072. doi:10.1016/j.patcog.2010.03.016.

S. Yan, Y. Xia, J. S. Smith, W. Lu, B. Zhang, Multiscale convolutional neural networks for hand detection, Applied Computational Intelligence and Soft Computing 2017.

M. A. Ur Rahman, M. S. U. Miah, M. A. Fahad, D. Karmaker, Shimpg: Simple human interaction with machine using physical gesture, in: 2014 13th International Conference on Control Automation Robotics Vision (ICARCV), 2014, pp. 301–305. doi:10.1109/ICARCV.2014.7064322.

F. Khan, M. N. Haji Mohd, D. Soomro, S. Bagchi, M. D. Khan, 3d hand gestures segmentation and optimized classification using deep learning, IEEE Access PP (2021) 1–1. doi:10.1109/ACCESS.2021.3114871.

F. Mueller, F. Bernard, O. Sotnychenko, D. Mehta, S. Sridhar, D. Casas, C. Theobalt, Ganerated hands for real-time 3d hand tracking from monocular rgb, in: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 49–59. doi:10.1109/CVPR.2018.00013.

C. Zimmermann, T. Brox, Learning to estimate 3d hand pose from single RGB images, CoRR abs/1705.01389. arXiv:1705.01389.
URL http://arxiv.org/abs/1705.01389

G. Devineau, F. Moutarde, W. Xi, J. Yang, Deep learning for hand gesture recognition on skeletal data, in: 2018 13th IEEE International Conference on Automatic Face Gesture Recognition (FG 2018), 2018, pp. 106–113. doi:10.1109/FG.2018.00025.

B. Hu, J. Wang, Deep learning based hand gesture recognition and uav flight controls, in: 2018 24th International Conference on Automation and Computing (ICAC), 2018, pp. 1–6. doi:10.23919/IConAC.2018.8748953.

D. Zhao, Y. Liu, G. Li, Skeleton-based dynamic hand gesture recognition using 3d depth data, Electronic Imaging 2018 (18) (2018) 461–1.

Q. De Smedt, H. Wannous, J.-P. Vandeborre, Skeleton-based dynamic hand gesture recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2016, pp. 1–9.

A. Caputo, A. Giachetti, S. Soso, D. Pintani, A. D'Eusanio, S. Pini, G. Borghi, A. Simoni, R. Vezzani, R. Cucchiara, A. Ranieri, F. Giannini, K. Lupinetti, M. Monti, M. Maghoumi, J. J. LaViola, M.-Q. Le, H.-D. Nguyen, M.-T. Tran, Shrec 2021: Track on skeleton-based hand gesture recognition in the wild (2021). doi:10.48550/ARXIV.2106.10980.
URL https://arxiv.org/abs/2106.10980

S. Tukhtaev, T. K. Whangbo, A combined method of skin-and depth-based hand gesture recognition, The International Arab Journal of Information Technology 17 (2019) 137–145. doi:10.34028/iajit/17/1/16.

Y. Ahn, Y.-C. Park, The hand gesture recognition system using depth camera, in: ACHI 2017, 2017.

J. Hou, G. Wang, X. Chen, J.-H. Xue, R. Zhu, H. Yang, Spatial-temporal attention res-tcn for skeleton-based dynamic hand gesture recognition, in: Proceedings of the European Conference on Computer Vision (ECCV) Workshops, 2018, pp. 0–0.

N. Mirehi, M. Tahmasbi, A new gng graph-based hand gesture recognition approach (2019). doi:10.48550/ARXIV.1909.03534.
URL https://arxiv.org/abs/1909.03534

Y. Chen, L. Zhao, X. Peng, J. Yuan, D. N. Metaxas, Construct Dynamic Graphs for Hand Gesture Recognition via Spatial-Temporal Attention, arXiv e-prints (2019) arXiv:1907.08871arXiv:1907.08871.

Y.-T. Li, J. P. Wachs, Recognizing hand gestures using the weighted elastic graph matching (wegm) method, Image and Vision Computing 31 (9) (2013) 649–657. doi:https://doi.org/10.1016/j.imavis.2013.06.008.
URL https://www.sciencedirect.com/science/article/pii/S0262885613001030

M. Avraam, Static gesture recognition combining graph and appearance features, International Journal of Advanced Research in Artificial Intelligence 3 (2). doi:10.14569/IJARAI.2014.030201.
URL http://dx.doi.org/10.14569/IJARAI.2014.030201

R. Cui, H. Liu, C. Zhang, A deep neural framework for continuous sign language recognition by iterative training, IEEE Transactions on Multimedia 21 (7) (2019) 1880–1891. doi:10.1109/TMM.2018.2889563.

E. Escobedo-Cardenas, G. Camara-Chavez, A robust gesture recognition using hand local data and skeleton trajectory, in: 2015 IEEE International Conference on Image Processing (ICIP), 2015, pp. 1240–1244. doi:10.1109/ICIP.2015.7350998.

J. Huang, W. Zhou, H. Li, W. Li, Sign language recognition using 3d convolutional neural networks, in: 2015 IEEE International Conference on Multimedia and Expo (ICME), 2015, pp. 1–6. doi:10.1109/ICME.2015.7177428.

A. D'Eusanio, A. Simoni, S. Pini, G. Borghi, R. Vezzani, R. Cucchiara, Multimodal hand gesture classification for the human–car interaction, Informatics 7 (3). doi:10.3390/informatics7030031.
URL https://www.mdpi.com/2227-9709/7/3/31

S. Jadooki, D. B. Mohamad, T. Saba, A. S. Almazyad, A. Rehman, Fused features mining for depth-based hand gesture recognition to classify blind human communication, Neural Computing and Applications 28 (2016) 3285–3294.

H. Mahmud, M. M. Morshed, M. K. Hasan, A deep learning-based multimodal depth-aware dynamic hand gesture recognition system (2021). doi:10.48550/ARXIV.2107.02543.
URL https://arxiv.org/abs/2107.02543

A. Ikram, Y. Liu, Real time hand gesture recognition using leap motion controller based on cnn-svm architechture, in: 2021 IEEE 7th International Conference on Virtual Reality (ICVR), 2021, pp. 5–9. doi:10.1109/ICVR51878.2021.9483844.

S. Y. Boulahia, E. Anquetil, F. Multon, R. Kulpa, Dynamic hand gesture recognition based on 3d pattern assembled trajectories, in: 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA), 2017, pp. 1–6. doi:10.1109/IPTA.2017.8310146.

Q. De Smedt, H. Wannous, J.-P. Vandeborre, Skeleton-based dynamic hand gesture recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2016, pp. 1206–1214. doi:10.1109/CVPRW.2016.153.

H. A. Jalab, Static hand gesture recognition for human computer interaction, Information Technology Journal 11 (9) (2012) 1265.

R. F. Pinto, C. D. Borges, A. Almeida, I. C. Paula, Static hand gesture recognition based on convolutional neural networks, Journal of Electrical and Computer Engineering 2019.

M. Z. Islam, M. S. Hossain, R. ul Islam, K. Andersson, Static hand gesture recognition using convolutional neural network with data augmentation, in: 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), IEEE, 2019, pp. 324–329.

B. K. Olorisade, P. Brereton, P. András, Reproducibility in machine learning-based studies: An example of text mining, 2017.

A. D'Amour, K. A. Heller, D. Moldovan, B. Adlam, B. Alipanahi, A. Beutel, C. Chen, J. Deaton, J. Eisenstein, M. D. Hoffman, F. Hormozdiari, N. Houlsby, S. Hou, G. Jerfel, A. Karthikesalingam, M. Lucic, Y. Ma, C. Y. McLean, D. Mincu, A. Mitani, A. Montanari, Z. Nado, V. Natarajan, C. Nielson, T. F. Osborne, R. Raman, K. Ramasamy, R. Sayres, J. Schrouff, M. Seneviratne, S. Sequeira, H. Suresh, V. Veitch, M. Vladymyrov, X. Wang, K. Webster, S. Yadlowsky, T. Yun, X. Zhai, D. Sculley, Underspecification presents challenges for credibility in modern machine learning, CoRR abs/2011.03395. arXiv:2011.03395.
URL https://arxiv.org/abs/2011.03395

B. Chen, M. Wen, Y. Shi, D. Lin, G. K. Rajbahadur, Z. Jiang, Towards training reproducible deep learning models.

N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman, A. Galstyan, A survey on bias and fairness in machine learning.

L. Chen, P. Koutris, A. Kumar, Towards model-based pricing for machine learning in a data marketplace, in: Proceedings of the 2019 International Conference on Management of Data, 2019, pp. 1535–1552.

O. Luzanin, M. Plancak, Hand gesture recognition using low-budget data glove and cluster-trained probabilistic neural network, Assembly Automation 34. doi:10.1108/AA-03-2013-020.

G. Xu, T. D. Duong, Q. Li, S. Liu, X. Wang, Causality learning: A new perspective for interpretable machine learning (2021). arXiv:2006.16789.

A. Bueff, I. Papantonis, A. Simkute, V. Belle, Explainability in machine learning: a pedagogical perspective (2022). doi:10.48550/ARXIV.2202.10335. URL https://arxiv.org/abs/2202.10335