Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Data Glove-Based Hand Gesture Recognition for Improved Human-Computer Interaction

Dr. T. B. Sivakumar[1], S. Rahini Sudha[2], Lakshminarayanan S[3], Dr. M. Anand[4], A Fayaz Ahamed[5]

[1]Associate Professor, Department of CSE, School of Computing, Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, Chennai - 600 062, Tamilnadu, India, drsivakumartb@veltech.edu.in
[2]Assistant Professor, Department of CSBS, Panimalar Engineering College, Chennai, India, rahinisudha0699@gmail.com
[3]Assistant Professor, Department of CSE, Sri Sai Ram Engineering College, Chennai, India, lakshminarayanan.cj@sairam.edu.in
[4]Professor, Department of ECE, DR M.G.R Educational and Research Institute, India, anand.ece@drmgrdu.ac.in
[5]Assistant Professor, R.M.K Engineering College, Kavaraipettai, Tamilnadu, India, afd.eee@rmkec.ac.in

The term "human-machine interaction" (HMI) describes how a human and a machine communicate and interact through a user interface. An emotion identification system based on hand gestures and facial expressions is described in this paper. Natural user interfaces, like gestures, are becoming more and more popular these days since they let people use machines with instinctive and spontaneous behaviors. The ability to identify significant motions made by an individual using their hands, arms, face, head, and body is known as gesture recognition. The present techniques for recognizing gestures are data-glove-based, depending on how hands are input. To address some of the issues that the data glove currently has. The proposal is for a data glove system that uses a computer vision and pattern recognition algorithm and basic Electromyography (EMG) sensors to accurately recognize gestures. The hand data glove for gesture recognition is proposed in real-time human-computer Interaction (HCI). When creating a clever and effective human–computer interaction, hand gestures are more significant. Gesture recognition has a wide range of uses, from virtual reality to medical rehabilitation to sign languages. The range of recognition accuracy is large, ranging from 70% to 98%, with an average of 89.6 %. The restrictions taken into account include various textual interpretations, movements, and intricate non-rigid hand features. This work differs from other recent research in that it covers every kind of gesture recognition method.

Keywords: Data-glove-based, Hand Gesture Recognition, Lightweight Machine-Learning Algorithm, (MWCNT) Sensors.

1. Introduction

A computer system is now an extremely powerful device that is meant to make jobs for humans' easier thanks to the tremendous intake and development of technologies. The human-computer interface, or HCI, has grown significantly in our daily lives. As a species, we are now unable to ignore the effects of the swift growth and advancement of computer-mediated contact, which has taken precedence over all other considerations. We use technology for everything from shopping and employment to stimulation and communication. It is so ingrained in how we live. Numerous programs, such as Microsoft Office, Windows Picture Manager, and Media Player, need an intuitive and natural user interface. The majority of users these days communicate with computers via keyboard, mouse, pens, joysticks, and other devices, which are insufficient for them. These current technologies for computing, communication, and display will soon become a bottleneck, and further development in these areas will be necessary to create a system that feels as natural as feasible.

Moving towards "natural" means that people communicate with one another is one of the long-term objectives of HCI. For a long time, research on human speech detection conducted with this objective in mind. It has advanced significantly in its field. Nonetheless, there is a greater effort lately to integrate other human-to-human means of communication in HCI. Therefore, hand gestures made by humans can offer a sensory and organic substitute for some incompatible technology. Similar to how we rely on our fingertips to connect with and interact with computers regularly, we can to do similarly with computers. We use our hands to move, alter, and attest at individuals and things as well as to convey and transfer information. [1] We use our hands to move, alter, and transform objects as well as to point at people or things. We also utilize them to convey and transmit information. In a similar vein, we can express thoughts through gestures with our hands while we talk. To enable computer interpretation, a method for investigating the use of gestures in HCI must be provided. A machine is required to measure the static and dynamic shape of a human arm, hand, and possibly some other body part gestures for HCI to remain described. The design of particular specialized input devices has proven that they're highly significant in enabling and achieving enhanced communication between humans and computers.

The human hand, which is well-suited for gestures seen to be the most natural body part for human-to-human connection, is also the most appropriate for communication between a person and a computer. Virtual gaming controllers, spoken language recognition, direction indication by pointing, enabling younger kids to interact with computers, human-computer communication, robot control, lay detection, and other common applications are among the many common uses for hand gesture recognition. Researchers have conducted a great deal of research on this topic due to growing interest in it, as seen by the numerous surveys provided. These surveys have something to do with hand gesture recognition either directly or indirectly.

It is difficult to define the movements of the hands outside of the HCI context. If definitions at all exist, they are especially focused on the way that human hands and bodies move in communication. Gaits are described as "the use of movements of the limbs or body as an instrument of communication; an action usually of the human body or limb that expresses or emphasizes an idea, emotions, or attitude" in Webster's Dictionary, for instance. This wide term is typically narrowed by psychology and social sciences, which further tie it to human expression and social interaction. But gestures have an additional significance in the field of HCI. [2] In a computer-controlled setting, the goal is to employ the hand to do activities that resemble both the hand's natural function as a manipulator and its use in human-machine communication, which involves gesturing to control computer and machine functions. Conversely, traditional definitions of gestures never, if ever, address the previously stated utilization of the human hand (sometimes referred to as practical gestures).

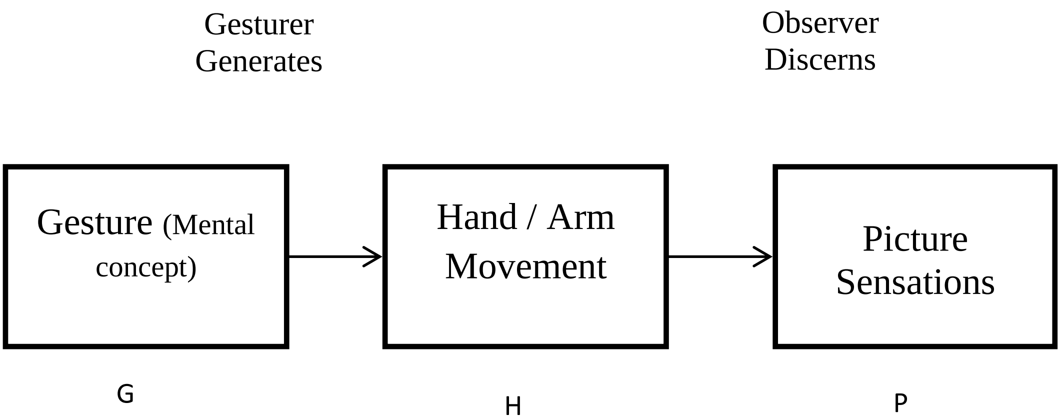


Fig. 1.1. Generating and Interpreting Gestures

Similar to language spoken, hand gestures can be used to communicate. Thus, a model frequently used in the area of spoken language identification can be used to represent the creation and interpretation of gestures. Fig. 1 shows this model's interpretation as it relates to gestures. The paradigm postulates that gestures begin as the thought of the gesture, sometimes in tandem with speech. They are conveyed through hand and arm movements, much like how speech is created through a human voice tract's circulatory system modulation. Additionally, observers view gestures as sequences of visual images that they then interpret based on their prior understanding of those motions.

The field of human-machine interaction is occasionally moving closer to simple and natural interfaces for users. Because humans can hold and manipulate objects with their hands well, keyboards and mice are more commonly used interfaces. Nowadays, the hand is utilized as well in a few HCI interfaces, such as those that use both dynamic and static recognition of gestures. In addition to its use in signaling language analysis and training, gesture-based data gloves are additionally being used in robots to operate the robot arms that are donning the glove.

This work primarily focuses on an accurate and successful grasping of the action by using the real-time input and output of data using the data glove. Hand Data Gloves is an electrical gadget with sensors that detect each finger and hand movement independently and provide an analog or digital signal to a computer. [3] These days, hand data gloves are utilized in a wide range of research areas, including virtual reality, playing games, robots, text verification and identification, and commerce applications. Gestures are even used in mobile video gaming platforms like the iPod, iPhone, and iPad. The data glove is available for use in motion animation and allows for the creation of numerous finger degrees of freedom. This is where hand data gloves come in handy for creating artwork and air-writing figures in a simpler, more actual time setting.

Glove-based methods offer dependable motion data and are simple to use. Numerous varieties of sensory gloves, including both commercialized and experimental models, are currently being developed. Since commercial products typically use resistive-bend sensors and pricey motion-sensing fibers, they are too expensive for the average consumer market. Thus, to reduce the price of such equipment, prototypes of data protection gloves are under development. The data glove has built-in flex or bending sensors. However, by placing the detector across the point of interest, the aforementioned sensors only measure the distance between the positions of flexible segments. This calls for precise sensor alignment with a specific joint. Furthermore, to reduce estimation errors caused by sensor changes, adjustment during operation is required. The absence of user customization for certain subject hands and the hand's palmar surface obstructing the sense of touch are two common drawbacks of data gloves. The attachment location needed to implant the sensors in garments is often associated with this. Magnet or inertial senses are induced to address the drawbacks.

Three metrics are used to evaluate the effectiveness of a hand gesture identification system: processing time, recognition accuracy, and classification accuracy. In this paper, the ideas of recognition and classification are distinguished. A sample is classified to determine its associated class. The prediction label and the actual label in the EMG data are simply compared in the classification evaluation process. The results of categorization are typically displayed as confusion matrices, where the gesture serves as a summary of sensitivity, precision, and accuracy. Classification is only one aspect of recognition; another is labeling a sample; recognition also needs time for the gesture's execution. [4]Therefore, the assessment of recognition accuracy makes a comparison between the basis for truth corresponding to a specific EMG sample and the vector of forecasts of an HGR system. The ground truth, which was previously acquired by an in-depth segmentation process, is a Boolean matrix set over points having muscle activity. This knowledge is included in each of the samples of the data set. According to this approach, a sample is deemed appropriately identified when the amount of overlap factor between the two vectors is more than an upper limit of 25%. Only a valid vector of predictions is used for this comparison. When there is just one continuous prediction segment that differs from the relaxing location and has the same label, the vector of guesses is considered legitimate.

The essay's remaining sections are organized as follows. Section 2 presents the research on the pertinent prior work. Section 3 describes the features of the proposed system, including the proposed system architecture, implementation model, characteristics of the graph-based technique, and data analysis. The implementation environment is described and the system's effectiveness is rated in Section 4. Section 5 provides the resolution.

2. Related Works

Using cameras with single, multiple, and depth perception lenses, a variety of computer vision-based research projects on human-machine interfaces were developed. Cameras equipped with various types of lenses can capture various types of information. [5] When more precise information is recovered, the recognition accuracy increases. According to the study, many techniques utilized to identify human gestures, including artificial neural networks, finite state machines, Markov models, principal component analysis (PCA), and specific filtering. In this brief paper, the author conducted research on hand gesture recognition and discusses the viability of these studies for hand gesture control of devices. The left arm of a human is detected using the Hough transform and the radon transform. In this case, the Hough and Radon transforms are used to perform background subtraction and straight-line detection. This method's inability to recognize characterized hand motions instinctively is one of its drawbacks.

The human face exhibits a great degree of irregularity and is a dynamic item. Previous projections have used a variety of methodologies. [6] Two categories of face detection algorithms exist: feature-based approaches and image-based approaches. The methods in the first category utilize the apparent characteristics of the face, including motion, skin tone, and facial geometry. While feature-based techniques can also detect faces quickly, they suffer from inconsistent results in low light. The image-based method for the second category makes use of recent developments in model recognition theory. The majority of image-based approaches use a window scanning technique, which is quite computationally intensive, to recognize faces.

A collection of permutations produced by hand and arm movements might be referred to as human hand gestures. These gestures could be as basic as pointing with the finger to more intricate gestures that are employed in interpersonal communication. In the course of human-computer interaction, the adoption of the hand and especially the palm and fingers as a means of input devices hence lowers the technical barrier in the interface between indifferent people and computers. [7] As we begin to use our hands as input devices, this offers a natural method to get rid of technological boundaries. This needs to be able to recognize patterns in people without the need for contact sensors.

A variation on this technique was put out to get around these issues and guarantee that only the user's hand would be identified. Instead of using skin tone to identify hands, this method uses glove color. Wearing a colored glove is required; after that, the program uses specified values for the color of the glove to detect the hand. [8] One needs just modify the application to match the range of colors utilized, thus any color glove can be used. With this strategy, the gloves must not match the color of the wearer or any items facing the camera. If not, there will also be a mistake in the hand detection outcome. Its extreme sensitivity to changes in glove or skin color presents yet another major issue with hand detection.

Researchers in a variety of domains have shown a great deal of interest in automatic affect analysis. There has been relatively little work done in multimodal affect analysis, and the majority of current computer-based methods for analyzing human affective states are unimodal in the input. The computer system processes are restricted to either voice signals or face images. Only four reports of bi-modal affect recognition can be located, according to the most recent summary of automatic affect recognition. [9] As such, although multimodal analysis of human affective states is becoming tractable due to recent research and technological advancements, this avenue of study is still in its early stages.

Through comparative analysis, we may determine that a single detection methodology is insufficient because several approaches can address various issues during the detection and recognition process. [10] For training classifiers, several machine learning techniques are accessible, including AdaBoost, support vector machine methodology, hidden Markov model, and principal component analysis. Additionally, there can be variations in the convex hull and contour detection of the hand region's border.

Human motion analysis has seen the use of broader vision techniques by numerous writers. Unlike digit eyes, these methods examine a portion of the entire hand movement, like a collection of motions or the inflexible movement of the palm. [11]A method for picking up and identifying dynamic hand gestures is described by Darrell and Pentland. Segen's related study approaches the detection of two-dimensional hand motions using artificial neural networks. These two methods operate in real-time on unmarked hand photographs, but they are not capable of producing 3D motion estimates, and their application to issues such as the 3D mouse interaction in Subsect would prove challenging.

3. Methods and Materials

3.1. Architecture of Human-Computer Interaction Systems:

Based on the quantity and variety of inputs and outcomes, an HCI system design may be roughly divided into two categories: unimodal HCI systems and multimodal HCI systems.

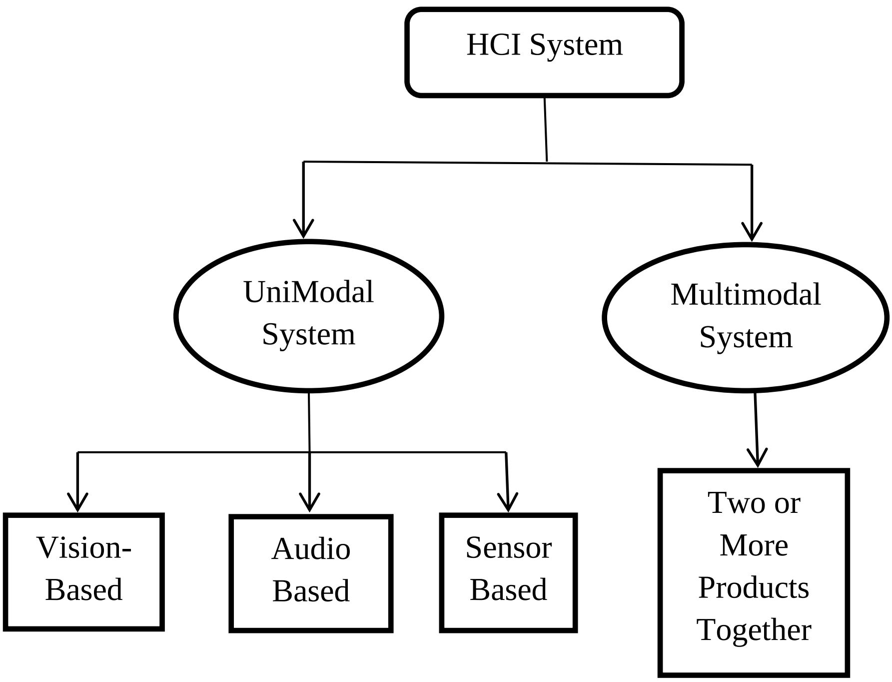


Fig. 1.2. HCI Systems

- **Unified Human-Computer Interfaces:** Unimodal systems may be audio or vision-based, depending on various kinds of sensors. Examples of the former include body motion recording, recognition of gestures, facial expression identification, and gaze detection.
- **Multimodal HCI systems:**HCI systems with several modes of operation Generally speaking, when people communicate with each other, they use several modalities. [12] This means that HCI frameworks should also include information gathered from several modalities to properly survey a user's expectations or behavior. The multimodal interface can be set up with combinations of data sources, such as gesture and speech, facial position and speech, and so on. E-commerce, intelligent gaming, smart video conferencing, driver monitoring, smart homes and offices, and help for individuals with impairments are a few of the main uses for multimodal systems.

3.2. Data Glove:

Fig. 1.3. Data Glove

Data gloves offer a unique method of communicating with computers that goes beyond the usage of a keyboard and mouse by utilizing sensors that provide crucial data regarding the motion or movements of the user. Users can communicate with computers without wearing anything tangible by using vision-based techniques, although this is impacted by occlusions, postures, and other ambient conditions. Conversely, data gloves are simple to use and offer accurate motion data. The advantages of producing data gloves are aided by the advancement of micro inertial devices, which offer benefits like compact size, large dynamic range, low power consumption, etc.

Magnetic detectors are used in conjunction with inertial sensors to achieve drift-free orientation and increased precision. [13] Without the use of external cameras, the inertial and magnetic measuring unit (IMMU) offers a superior method for precise body segment estimation.

Fig. 1.4. Flex band Resistor

Flex band sensors combined with microcontrollers & a transceiver module are used to create a basic data glove. By providing the values of a finger's bend, flex resistors are utilized to detect motion. These trigger resistors are positioned on the gloves' fingers. The user-pressed key values are virtually provided by the flex band resistors. A flex band resistor's resistance increases from 11k in its unbent condition to 150k at a bending degree of 140 degrees.

Fig. 1.5.Open Source Data Glove

These signals are sent to the transceiver unit after being processed by a microcontroller. Then, specific procedures or actions are carried out using these values. For interaction between the gloves and the system, a transceiver system which might be wired or wireless is employed. The advantages of data gloves over camera-based methods for virtual reality applications, emphasize the latter's inefficiency in terms of cost and other environmental considerations. Different neural network models' performances were compared.

3.3. The Electromyography [EMG]:

Electrical sensing techniques, such as electromyography (EMG) or electromagnetic impedance tomography (EIT), can measure muscle contraction by measuring the response of the muscles to an outside given current of electricity (within a safe range). The EMG and EIT sensors detect upper-limb movements, such as flexion, extension, and twist, brought on by hand and wrist movements, by detecting the electric signals. Specifically, numerous types of research on gesture detection have made use of a surface EMG (EMG) sensor. For example, in Figure 1.6, joint motions and muscle activity were recorded using the multifunctional and stretchy sEMG electrodes.

However, a variety of disruptions in real-world settings can easily affect sEMG signals, which could have an impact on feature extraction and, ultimately, classification or regression of hand gestures.

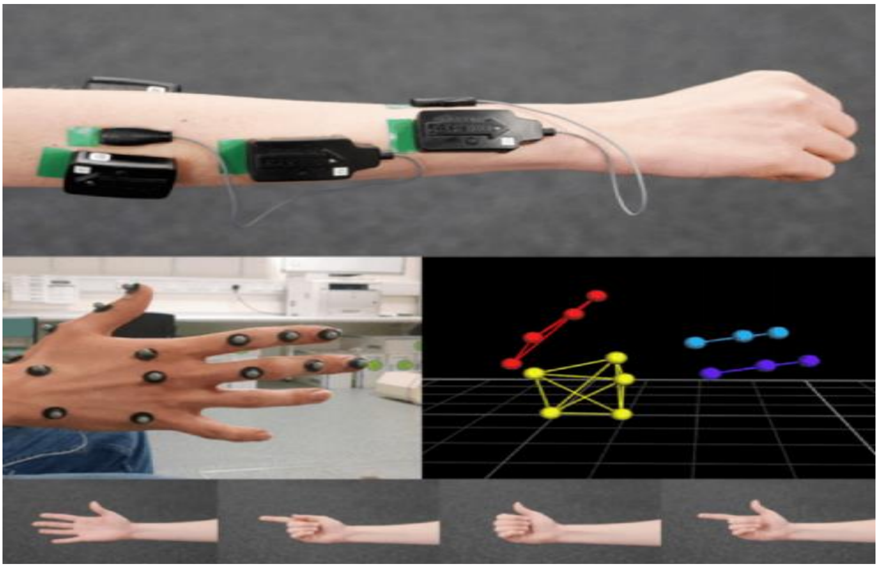


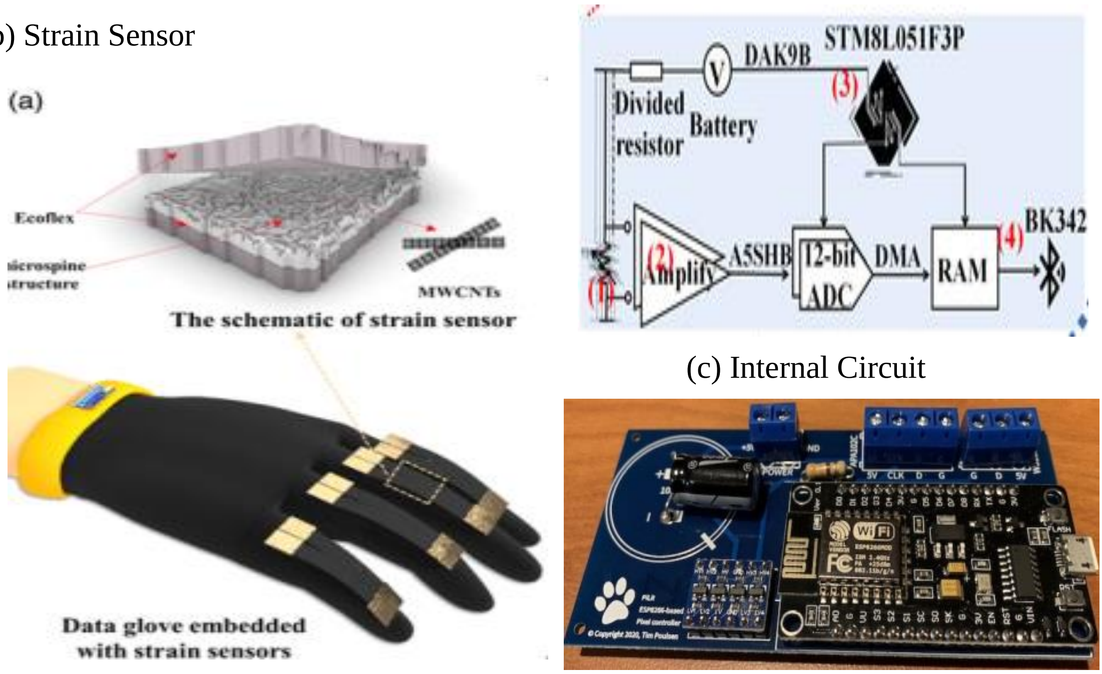
Fig. 1.6.EMG Electrodes location.

EMG signals can be confused by several factors, such as electrode moving, limb position, the strength of a muscle contraction, idea drift, etc. [14] The electromyography control method, sometimes referred to as myoelectric control mechanisms (MCSs), is typically used to classify surface EMG signals for control purposes. Powered upper limb prostheses and electric-powered wheelchairs are two of the primary potential uses for MCSs. To record the user's myoelectric activations while performing particular hand motions, EMG and fiducially marker-based monitoring are used. The apparatus displayed the ability to operate a dexterous robotic arm and hand system and interpret gestures to determine the preferred grip style for the robot hands. When determining a tissue's inner impedance dispersion and changes in conductivity, EIT is a non-invasive imaging technique that can be used. Through the use of high-frequency, low-amplitude electrical currents among pairs of sensors on the body, EIT records the potential among all other pairs of conductors in addition to measuring the internal resistance of the tissue. Though the electrodes need to make a thin contact with the outer layer of skin to function properly, this method has a high degree of precision. Compared to sEMG, EIT provides more freedom of design in terms of electrode count and measurement pattern, which is why it is commonly utilized in HGR.

Design of Data Glove System:

Using the MWCNT stress sensors, we created an internet-connected data glove system that can recognize gestures in real-time. Five sensors were fastened to the joint position, as seen in Figure 1.7a, to identify the releasing and stretching caused by the movements of several fingers. Figure 1.7b illustrates the strain sensor's unique structural construction; the red arrows point in the direction of the current flow. The cropped sensor displays an even greater shift in signal for the identical finger bending circumstance (see Figure S1, Additional Information), and Note S1, [15] Additional Information provides an analysis and justification of the unique structural design. Based on integrated circuit components, the customized wireless printed circuit board, or PCB, integrates several functionalities, including wireless transmission, signal amplification, and multichannel signal acquisition (Figure 1.7c). The integrated circuit's component locations that correlate to the digits in the PCB's internal circuit are shown in Figure 1.7d by the red dashed boxes. The analog-to-digital conversion (ADC) in the circuit amplifies the received signals and transforms them into digital signals. A low-energy Bluetooth connection then sends the digital signals to the computer. Less than a tenth of the bulk of an average pair of commercial data gloves, the systems overall mass is as low as 36.8 g.

(b) Strain Sensor



(c) Internal Circuit

(a) Data Glove

(d) Custom PCB

Fig. 1.7. MWCNT Strain Sensors are Incorporated with the Data Glove System.

3.4. Hand gestures for Human-Computer Interaction (HCI):

Open

Close

Cut

Paste

Maximize

Minimize

Fig. 1.8. Human-computer interaction (HCI) hand gestures

A crucial component of creating a suitable gesture language for Human-Computer Interaction (HCI) is choosing the right-hand motions. It is necessary to take into account the differences and anatomical potential of hand motions. One goal of human-computer interaction is to give a computer a set of hand gestures that serve as commands.Six instructions for static hand movements are created in our facial recognition system. Human-computer interaction hand gestures for static management are displayed in Figure 1.8.

We must pre-process the hand gesture pictures given by the digital camera to be able to gather the components required for the recognition of gestures process' categorization step. [16] Isolating and improving the hand region is the primary objective. Separation, filtration, and identifying edges are the steps in the pre-processing chain. Compared to the hand area, the rest of the image is thought to be nearly consistent and darker. Our attention is drawn to the execution of static hand gesture recognition using this constraint. A linear median filter is used to perform the enhancing process. The handle edge might be found using the Sobel operator.

3.5. Algorithmic methods for identifying hand gestures and positions:

If any gestures or postures have been identified, it is necessary to analyze the raw data obtained through a vision- or glove-based data-gathering device.

Template Alignment:

The most basic technique for hand posture recognition is template matching. One way to determine if an individual data post can be categorized as an element of a group of recorded data records is to use template matching. There are two components to employing template matching to recognize hand postures. The initial step involves gathering data values for every posture in the posture set to generate the templates. The second step involves contrasting the current sensor measurements with the provided set to determine which posture template most closely matches the current data record.

Analysis of Feature Extraction:

The process of extracting and analyzing low-level information from raw data to generate higher-level semantic information that is utilized to identify gestures and postures is known as the extraction of features and analysis. With over 97% accuracy, the system identified these motions. It has a strong recognition system for hand movements and postures. It is capable of identifying intricate hand motions and postures in addition to basic ones.

Model of Active Shapes:

Active shape models, sometimes known as "smart snakes," are a method for identifying a feature inside a still image. The shape of the feature to be traced is approximated by its contour on the image. The Contour is manipulated by repeatedly dragging it towards adjacent edges, which distorts the contour to fit the characteristic. Every frame has an active shape model applied to it, which uses the feature's location there as a starting point for the subsequent frame's approximation.

The analysis of principal components:

Principal Component Analysis is a statistical method that keeps the dataset's variety while lowering the dimensionality of a collection of data with numerous connected variables. Compression of the data set is the process of converting the old data into a new set of factors that are arranged so that the majority of the variation found in the original variables is included in the first few variables. [17] The initial data set is converted by calculating the eigenvalues and eigenvectors of the covariance matrix. The position, orientation, and sizing of the hand in the image all have a significant impact on the data while working with images.

Models of Linear Fingertips:

This model assumes that very few finger movements include rotation and that the majority of finger movements are linear. The approach allows for a model that depicts each fingertip journey across space as a simple vector and solely uses fingertips as input data. Motion correspondence is used to compute the fingertips' trajectories when they are detected. By storing the motion code, the name of the gesture, and the magnitude and direction vectors for every fingertip, the postures are generated from a short training set. If all of the direction and intensity vectors within a certain threshold coincide with a gesture trace in the training set, the postures are considered recognized.

Casual Evaluation:

Causal analysis is a vision-based recognition method derived from scene analysis research. Using a high-level understanding of the events in the scene and their relationships to each other and the surrounding environment, the approach retrieves data from a video stream. To detect gestures, the gesture filters normalize, aggregate, and apply causal knowledge about how people interact with real-world items. The system records the locations of the elbow, wrist, and shoulder joints in the picture plane. The system derives a feature set consisting of wrist velocity and acceleration, work done versus gravity, gesture size, area between arms, tilt between armpits, nearness to torso, and verticality from these locations. To identify movements like opening, pulling patting, pushing, pausing, and gripping, gesture filters normalize, aggregate, and apply causal information about how people interact with physical items. The accuracy of this procedure remains unclear. The fact that this technology does not utilize finger data is another drawback. If this method is reliable enough to be applied in any nontrivial applications, more investigation is required.

The Algorithm for Hand Tracking and Segmentation (HTS):

The goal of this algorithm is to remove complex backgrounds and detect skin colors reliably. It also attempted to detect and segment hands, tracked hands using the mean shift algorithm, and processed odd frames for speed. To improve performance, the user's skin color samples were input and an HSV histogram was created. The OpenCV library's Cam Shift function is used for monitoring and detection. The use of the edge traversal algorithm would obtain the fine contour of the hand shape. Since dynamic background was taken into consideration, it was possible to detect unwanted edges from the background while recording the user's signal after edge detection. Therefore, a traversal algorithm is developed only to identify the boundary of the user's hand edge.

- Obtain the camera's picture frames.
- Handle strange frames and use the CamShift feature to track the hand by giving it a flesh color.
- The run time samples.
- The experimented threshold value is supplied to the CamShift function for tracking the necessary hand portion, and an HSV histogram is formed.
- Divide the necessary hand area of the image.
- Use canny edge detection to locate the edges.
- Extend the picture.
- Remove the picture.
- Apply the edge traversal algorithm to get the finalContour.

3.6. The Implementation & Decision Structure:

The experimental design of our suggested system will be covered in this section. As previously mentioned, a single rechargeable battery between 2.5 and 6 volts powers the DG5 V Hand 3.0 data glove used in this experiment. In the temperature range of -43°F to 121°F, the Bi-Flex bend sensors detect the presser as well. This sensor is capable of measuring 1024 distinct positions for each finger in addition to three degrees of integrated tracking (x, y, and z axes) for roll, pitch, and yaw. An array with eight indexes serves as the data structure in this case. As seen in Table 1.1, the thumb, middle finger, index finger, ring finger, and tiny finger are the remaining five indexes. The initial three indexes are roll, pitch, and yaw. The axis values are easily determined using the following formulas, which take the numerical value of fingers and range from 0 to 1023:

$$x_b = \frac{2(x_{bl} + x_{bh} + 128)\sin x_{bh}}{32767}$$

$$y_b = \frac{2(y_{bl} + y_{bh} + 128)\sin y_{bh}}{32767}$$

$$z_b = \frac{2(z_{bl} + z_{bh} + 128)\sin z_{bh}}{32767}$$

(x_{bl}, x_{bh}) is the smaller and greater value associated with the x-axis at that particular moment for that gesture activity. The acceleration values range from -32676 to 32676.

Table 1.1. Feature Vector

FEATURE VECTOR							
x_b	y_b	z_b	Thumb f1	Index f2	Middle f3	Ring f4	Little f5

The next step after carrying out gesture operations is activity detection. In essence, activity detection is the use of machine learning algorithms to classify different events based on certain fundamentals. This study uses a basic decision tree to identify activity.

For decision-making problems, Decision Tree Đ (Δ) is potent and well-liked machine learning approach. It is applied to categorization issues as well. Several real-world applications, including fraud detection, weather forecasting, picture segmentation and processing, medical diagnosis, gesture recognition, and many more, have made use of decision trees. [18] It is straightforward to use and put into practice. Even though the decision tree is merely a collection of if-then-else rules, it is quite effective in identifying data that share characteristics. Consequently, a decision tree is a graphical representation of data divided into several classes based on their qualities. Furthermore, decision trees offer several other benefits, like the little need for data preprocessing and ease of comprehension and interpretation through the use of a white-box methodology. Its ability to function admirably with massive amounts of data quickly is maybe its most crucial feature. Figure 1.9 below depicts the decision tree, Đ (Δ) that was used in this paper to map the sub-gestures with the VLC interface. We have limited the number of gesture actions for VLC interaction to just 10 in this experiment to simplify it: Play, Pause, and Full Screen, pause, Mute, ahead, Reversed, after that, before, and Void gestures.

The above Đ (Δ) clearly illustrates that to perform the Next operation, the z-axis must be in a positive direction, the y-axis must remain constant, and the last requirement is that the shift in the x-axis has to be negative, meaning the user must wave his or her hand in the negative x-axis, as shown in the equation below:

$(x_{current} - x_{initial}) \leq 0 | Z \geq 0 \;\&\; Y = constant$

The value for the previous procedure needs to be positive, as indicated below:

$(x_{current} - x_{initial}) \geq 0 | Z \geq 0 \;\&\; Y = constant$

Where $x_{initial}$the original desire wasassociated with the x-axis and $x_{current}$is its present passion. Other gestures are computed and employed appropriately in a similar manner.

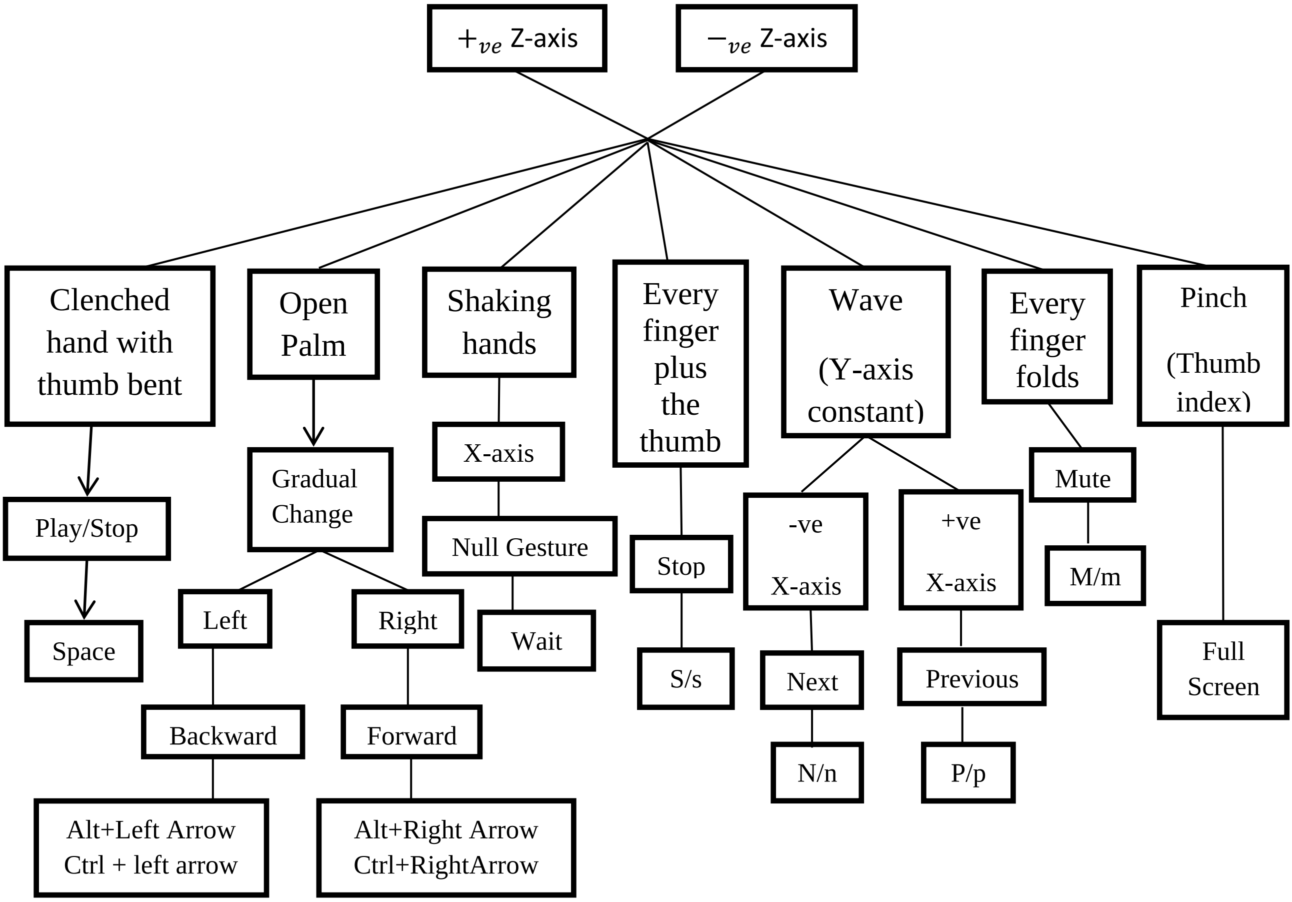


Fig. 1.9.Decision tree, Đ (Δ) for VLC gesture interface.

4. Result Analysis and Discussion

Outcomes of the HGR system:

The EMG-IMU-based HGR system's findings for the specific user's model training are shown in this subsection. As part of the validation process, we first analyze the data from 16 users to determine the optimal hyper-parameters. Next, we showcase the ultimate testing outcomes for sixteen distinct users, utilizing the optimal hyper-parameters discovered throughout the verification process. The following is an explanation of the testing and validation outcomes. Utilizing the test set consisting of sixteen distinct users, the best hyper-parameters discovered during the validation process were employed in this instance to assess the user-specific models and provide the testing outcomes. This helps analyze over-fitting by evaluating the best-found model using several sets of data. For 16 participants of the assessment set, the test accuracy scores for recognition and classification were 98.56 ±1.08 and 87.4 ± 2.6, respectively.

Even if the results are from different users, it is evident that there is not much of an impact between the testing accuracy and validation findings. [19] Therefore, for the suggested data set, we can conclude that our hypotheses are resistant to overfitting. Lastly, we also provide the confusion matrix in Table 1.2, which illustrates the test set categorization results. This matrix enables us to see the specific outcomes for every hand gesture. It is noteworthy that each window observation takes an average of 23 ms to process.

Table 1.2. Confusion Matrix

	Open	Close	Cut	Paste	Maximize	Minimize	
Open	227 15.4%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0%
Close	0 0.0%	234% 15.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0%
Cut	1 0.1%	2 0.1%	234% 15.2%	1 0.1%	4 0.3%	3 0.2%	97.3% 2.7%
Paste	0 0.0%	2 0.1%	1 0.1%	227% 15.7%	1 0.1%	0 0.0%	96.5% 3.5%
Maximize	0 0.0%	2 0.1%	1 0.1%	0 0.0%	222% 15.1%	3 0.2%	92.2% 7.8%
Minimize	2 0.1%	1 0.1%	3 0.2%	1 0.1%	2 0.1%	235% 15.3%	91.6% 8.4%
	97.7% 2.3%	96.8% 3.2%	96.8% 3.2%	94.0% 6.0%	97.0% 3.0%	97.2% 2.8%	96.4% 3.6%

Enhancement of Data:

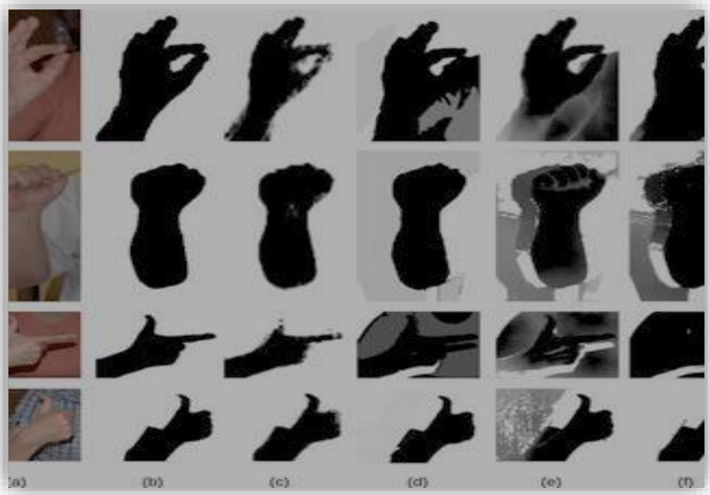


Fig. 1.10. Segmentations examples on the OUHANDS test set

Table 1.3. Segmentation Results on OUHANDS Hand Gestures

Method	Score	Time	Parameter	Size
FCN-8s	0.945	60	134M	537 MB
PSPNet	0.934	56	68.9M	456 MB
DeepLabv3	0.967	48	78.2M	326 MB
HGR-Net (no ASPP)	0.956	34	58.6M	534 MB
HGR-Net	0.925	21	1.45M	14MB

To prevent overfitting and enhance the model's generalization, we employ both online and offline data augmentation techniques to address the issue of dataset bias. Our offline data augmentation entails two changes: focusing (20-25%) and vertical or horizontal translations (20-25%), which together increase the training, set by four times. When using online data enhancement, we perform vertical/horizontal translations, rotation, shearing, and zooming of the data while it is being trained. Within the small batches that are introduced into the model, each of these processes is applied at random. We use lower modification values for these phases because we train our appearance and shape stream network from scratch, which speeds up the models' convergence.

Hand division trials:

This hand segmentation model was assessed using OUHANDS datasets. The F-score, or equally weighted mean of the recall and precision, is used to quantitatively compare the accuracy of the suggested approach. First, we compare each segmentation stage of HGR-Net (Stage 1) performance regardless of the ASPP module. To assess how well our segmentation model performs on this particular task, we also use the initial architectures of PSPNet, DeepLabv3, and FCN-8s. Table 1.3 reports the outcomes. [20] In combination with the ASPP component, which is roughly as precise as DeepLab3 and PSPNet but significantly less big and 3× quicker at run-time for a 320 × 320 RGB image, we find that our model performs better. For all DeepLabv3 and PSPNet architectures, we employ ResNet-40 as a feature extractor. In addition, our model is 2× faster and significantly smaller than FCN-7s, with a little performance advantage. Using comparison segmentation findings from the OUHANDS test set, Figure 1.10 provides a qualitative demonstration of our model's effectiveness. Images such as these demonstrate how well our segmentation model handles typical hand segmentation difficulties including erratic backdrops and lighting.

Table 1.4. Assessment of HGR-Net using various data augmentation methods for HGR tasks

Data Intensification	**Train**	**F-Score**
None	0.975	0.856
Offline	0.785	0.746
Online	0.978	0.587
Offline + Online	0.674	0.568

The HGR-Net is then assessed for the HGR task. Following various methods of data augmentation, Table 1.4 displays the classification scores for the OUHANDS test set. Table 1.4 clearly shows that test accuracy is higher when augmented online than when tested offline. However, training HGR-Net with online as well as offline methods yielded the best results. The aforementioned outcomes validate that implementing this data augmentation technique can effectively decrease overfitting and enhance the model's universality. Thus, it can be helpful for strong HGR in actual environments.

5. Conclusion

In this we will develop a hand gesture recognition model able to recognize dozens of gestures of the hand, with recognition accuracy greater than the real-time models proposed in the scientific literature and the number of gestures training data must be in the order of units, avoiding overfitting. To achieve these goals, we use Machine Learning algorithms for the mathematical model of the EMG signal is unknown. In the proposed hand gesture recognition model, we analyze preprocessing techniques, feature extraction techniques, classifiers, and post-processing techniques. The test accuracy scores for recognition and classification were 98.56 ±1.08 and 87.4 ± 2.6, respectively.

Electromyography (EMG) can identify muscle activity even when there is no visible movement. Identifying the highest energy pathway for a collection of synchronization EMG data of the wave-out gesture is the basis of the orientation correction algorithm. Positive results were found for both recognition and categorization. Even though the Myo bracelet detector is turned during training and testing, the suggested orientation correcting algorithm can enhance the hand gesture identification system's classification and recognition performance. This work primarily aims to enhance and streamline the conventional approach for hands-on gesture detection. Using depth vision allows for high grouping accuracy, which makes the EMG training process easier. According to standard finger behavior, the average percentage inaccuracy of the Data Glove for the angle of bending was found to be 25.33% & 17.42%. However, when comparable data is entered into the finite elements-based model, the modeled Easy Finger's response indicates an average percentage error of 35.21% for the Data Glove.

References

1. Sharma, R. P., & Verma, G. K. (2015). Human computer interaction using hand gesture. *Procedia Computer Science*, *54*, 721-727.
2. Pavlovic, V. I., Sharma, R., & Huang, T. S. (1997). Visual interpretation of hand gestures for human-computer interaction: A review. *IEEE Transactions on pattern analysis and machine intelligence*, *19*(7), 677-695.
3. Kumar, P., Rautaray, S. S., & Agrawal, A. (2012, March). Hand data glove: A new generation real-time mouse for human-computer interaction. In *2012 1st international conference on Recent Advances in Information Technology (RAIT)* (pp. 750-755). IEEE.
4. López, L. I. B., Caraguay, Á. L. V., Vimos, V. H., Zea, J. A., Vásconez, J. P., Álvarez, M., & Benalcázar, M. E. (2020). An energy-based method for orientation correction of EMG bracelet sensors in hand gesture recognition systems. *Sensors*, *20*(21), 6327.
5. Vasantrao, S. S., Prakash, C. M., Prakash, B. A., & Anant, G. A. (2014). Improved HCI using face detection and speech recognition. *International Journal*, *3*(2).
6. Rautaray, S. S., & Agrawal, A. (2012). Real time multiple hand gesture recognition system for human computer interaction. *International Journal of Intelligent Systems and Applications*, *4*(5), 56-64.
7. El Sibai, R., Abou Jaoude, C., & Demerjian, J. (2017, September). A new robust approach for real-time hand detection and gesture recognition. In *2017 International Conference on Computer and Applications (ICCA)* (pp. 18-25). IEEE.
8. Zeng, Z., Tu, J., Liu, M., Zhang, T., Rizzolo, N., Zhang, Z., ... & Levinson, S. (2004, October). Bimodal HCI-related affect recognition. In *Proceedings of the 6th international conference on Multimodal interfaces* (pp. 137-143).
9. Gurav, R. M., & Kadbe, P. K. (2015, May). Real time finger tracking and contour detection for gesture recognition using OpenCV. In *2015 International Conference on Industrial Instrumentation and Control (ICIC)* (pp. 974-977). IEEE.
10. Rehg, J. M., & Kanade, T. (1994, November). Digiteyes: Vision-based hand tracking for human-computer interaction. In *Proceedings of 1994 IEEE workshop on motion of non-rigid and articulated objects* (pp. 16-22). IEEE.

11. Dardas, N. H., & Alhaj, M. (2011). Hand gesture interaction with a 3D virtual environment. *The Research Bulletin of Jordan ACM*, *2*(3), 86-94.
12. Sarma, D., & Bhuyan, M. K. (2021). Methods, databases and recent advancement of vision-based hand gesture recognition for hci systems: A review. *SN Computer Science*, *2*(6), 436.
13. Aditya, K., Chacko, P., Kumari, D., Kumari, D., & Bilgaiyan, S. (2018, July). Recent trends in HCI: A survey on data glove, LEAP motion and microsoft kinect. In *2018 IEEE International Conference on System, Computation, Automation and Networking (ICSCA)* (pp. 1-5). IEEE.
14. Tchantchane, R., Zhou, H., Zhang, S., & Alici, G. (2023). A review of hand gesture recognition systems based on noninvasive wearable sensors. Advanced Intelligent Systems, 5(10), 2300207.
15. Li, Y., Yang, L., He, Z., Liu, Y., Wang, H., Zhang, W., ... & Song, G. (2022). Low‐Cost Data Glove based on deep‐learning‐enhanced flexible multiwalled carbon nanotube sensors for real‐time gesture recognition. *Advanced Intelligent Systems*, *4*(11), 2200128.
16. Stephan, J. J., & Sana'a Khudayer. (2010). Gesture Recognition for Human-Computer Interaction (HCI). *Int. J. Adv. Comp. Techn.*, *2*(4), 30-35.
17. Pradipa, R., & Kavitha, S. (2014). Hand gesture recognition–analysis of various techniques, methods and their algorithms. *International Journal of Innovative Research in Science, Engineering and Technology*, *3*(3), 2003-2010.
18. Uvarajan, K. P., and K. Usha. "Implement A System For Crop Selection And Yield Prediction Using Random Forest Algorithm." International Journal of communication and computer Technologies 12.1 (2024): 21-26.
19. Prasad, S., Kumar, P., & Sinha, K. P. (2014, August). A wireless dynamic gesture user interface for HCI using hand data glove. In *2014 Seventh International Conference on Contemporary Computing (IC3)* (pp. 62-67). IEEE.
20. Cruz, P. J., Vásconez, J. P., Romero, R., Chico, A., Benalcázar, M. E., Álvarez, R., ... & Valdivieso Caraguay, Á. L. (2023). A Deep Q-Network based hand gesture recognition system for control of robotic platforms. *Scientific Reports*, *13*(1), 7956.
21. Dadashzadeh, A., Targhi, A. T., Tahmasbi, M., & Mirmehdi, M. (2019). HGR‐Net: a fusion network for hand gesture segmentation and recognition. *IET Computer Vision*, *13*(8), 700-707.