Systematic Review

Research Progress of Human–Computer Interaction Technology Based on Gesture Recognition

Hongyu Zhou [1,2], Dongying Wang [2], Yang Yu [2] and Zhenrong Zhang [1,*]

[1] Key Laboratory of Multimedia Communication and Network Technology in Guangxi, School of Computer, Electronics and Information, Guangxi University, Nanning 530004, China

[2] College of Liberal Arts and Sciences, National University of Defense Technology, Changsha 410073, China

[*] Correspondence: zzr76@gxu.edu.cn

Abstract: Gesture recognition, as a core technology of human–computer interaction, has broad application prospects and brings new technical possibilities for smart homes, medical care, sports training, and other fields. Compared with the traditional human–computer interaction models based on PC use with keyboards and mice, gesture recognition−based human–computer interaction modes can transmit information more naturally, flexibly, and intuitively, which has become a research hotspot in the field of human–computer interaction in recent years. This paper described the current status of gesture recognition technology, summarized the principles and development history of electromagnetic wave sensor recognition, stress sensor recognition, electromyographic sensor recognition, and visual sensor recognition, and summarized the improvement of this technology by researchers in recent years through the direction of sensor structure, selection of characteristic signals, the algorithm of signal processing, etc. By sorting out and comparing the typical cases of the four implementations, the advantages and disadvantages of each implementation and the application scenarios were discussed from the two aspects of dataset size and accuracy. Based on the abovementioned discussion, the problems and challenges of current gesture recognition technology were discussed in terms of the biocompatibility of sensor structures, wearability and adaptability, stability, robustness, and crossover of signal acquisition and analysis algorithms, and the future development directions in this field were proposed.

Keywords: gesture recognition; human–computer interaction; sensor technology; electromagnetic wave sensing recognition; mechanical sensing recognition; electromyo–graphic sensing recognition; visual sensing recognition

1. Introduction

During the early stages of the Second Industrial Revolution, humans controlled and regulated the speed and direction of trains, as well as the rotational speed and power output of steam engines, by manipulating buttons and levers, marking the initial form of human–machine interaction. With the advent of computers, this mode of information transmission gradually gave way to keyboard input and mouse control. In recent years, the rapid development of signal acquisition technology and machine learning has led to increased information transmission speed and improved accuracy of information recognition. Complex tasks can now be precisely accomplished by simply driving the computer with simple signals. As technology has advanced, researchers have also developed various human–machine interaction technologies, including voice control [1–3], brain-computer interfaces [4–7], facial expression control [8–11], and gesture recognition, among others, further enhancing the freedom, applicability, and efficiency of human–machine interaction.

Among the numerous human–machine interaction technologies, gesture recognition holds a significant position due to the frequent use of hand movements by humans to convey and receive information. Studies have shown that gestures account for 55% of

the importance in information transmission, while sound and text each account for the remaining 45% [12]. This highlights the crucial role that body language plays in expressing emotions and teaching, making gesture recognition a core technology in human–machine interaction, offering advantages such as simplicity, flexibility, and rich connotations [13]. This article primarily discussed gesture recognition technologies related to the palm and its surrounding area. Depending on the method of implementation, these technologies can be broadly classified into four categories: electromagnetic wave sensing recognition, mechanical sensing recognition, electromyography sensing recognition, and visual sensing recognition. Due to the unique advantages associated with each gesture recognition approach, extensive research has been conducted worldwide to explore diverse implementation methods.

After this introduction, the article is structured as follows: In the second section, the article provides an in-depth introduction to the principles and implementation methods of various gesture recognition technologies. It categorizes and exemplifies the advancements made by researchers and scholars in recent years, focusing on areas such as feature extraction methods, artificial intelligence algorithms, and sensor material structure characteristics. The third section compiles and compares typical cases of four implementation methods, comprehensively discussing the advantages, disadvantages, and applicable scenarios of each approach from the perspectives of dataset size and accuracy. In the fourth section, the article delves into the applications and development of gesture recognition technology in modern production and daily life, encompassing areas such as improving traditional control methods, medical applications, and sports training. The fifth and sixth sections explore the existing problems and challenges of current gesture recognition technology, considering factors such as the biocompatibility and wearability of sensor structures, as well as the adaptability, stability, robustness, and cross-functionality of signal acquisition and analysis algorithms. Finally, these sections provide a summary and future outlook on the development directions within this field.

This paper provides a systematic summary and analysis of the current state and developmental trajectory of gesture recognition-based human–computer interaction technology. It identifies prevailing issues and proposes future directions for development. The findings are expected to facilitate the advancement and practical application of gesture recognition technology while aiding researchers and scholars in selecting implementation methods that align with their research objectives and application requirements. Moreover, this work serves as a foundation for enhancing and innovating this technology.

2. Research Methods and Current Situation

2.1. Electromagnetic Wave Sensing Recognition

The principle underlying gesture recognition using electromagnetic wave sensing is based on the physical phenomena of reflection, refraction, and scattering that occur when electromagnetic waves encounter obstacles, specifically human hands, in their path. These phenomena lead to changes in the intrinsic parameters of the original electromagnetic waves. By analyzing the variations in the transmitted and received signals and utilizing demodulation techniques, gesture poses can be identified. Currently, electromagnetic wave sensing for gesture recognition can be categorized into two main types: Wi-Fi-based recognition and radar-based recognition.

In a static propagation model, as electromagnetic waves propagate, they experience not only a direct path but also reflection, refraction, and scattering due to the presence of human hands. Consequently, the receiving end captures multiple signals from different paths, resulting in the occurrence of multipath effects [14]. For the direct path, the Friis free-space propagation equation [15] can be employed to determine the received signal strength:

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 d^2} \tag{1}$$

where P_t represents the transmission power, $P_r(d)$ represents the received power, G_t and G_r are the transmission and reception gains, respectively, λ represents the wavelength, and d represents the propagation distance. When there is a human hand in the propagation path, Equation (1) becomes

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 (d + 4h + \Delta)^2} \qquad (2)$$

where Δ represents an approximation of the path length change caused by the disturbance of the human hand, while h represents the distance between other reflection points and the direct path, excluding the human hand [16]. It can be observed that the received power is inversely proportional to the length of the propagation path. When a human hand obstructs the propagation path, it introduces disturbances and creates a new propagation path.

The dynamic propagation model primarily relies on the Doppler effect, which describes the change in the wavelength of radiation emitted by an object due to the relative motion between the source and the observer. Assume the original wavelength of the source is λ, the wave speed is u, and the velocity of the observer is v. When the observer approaches the source, the observed frequency $f1$ of the source can be calculated as follows:

$$f1 = \frac{u^2}{\lambda(u - v)} \qquad (3)$$

Otherwise, the observed wave source frequency $f2$ is

$$f2 = \frac{u^2}{\lambda(u + v)} \qquad (4)$$

In the presence of a moving wave source, the wave is compressed, resulting in a shorter wavelength and a higher frequency. Conversely, when the wave source is behind in motion, the opposite effect occurs, causing the wavelength to lengthen and the frequency to decrease [17].

By considering these two effects, Wi-Fi signals at the Medium Access Control (MAC) layer can be represented by the Received Signal Strength Indication (RSSI), which accounts for the accumulation of propagation delay, amplitude attenuation, and phase shift along different propagation paths [18–20]. However, this representation has limitations, such as low-ranging accuracy and deviation in recognition due to the influence of static multipath propagation on RSSI fluctuations. With the continuous improvement of Wi-Fi protocols, the development of Orthogonal Frequency Division Multiplexing (OFDM) technology has enabled the utilization of Channel State Information (CSI) at the physical layer to reflect the signal state during propagation [21]. Due to the implementation of OFDM technology, the channel between the transmitter and receiver is partitioned into several subcarriers, as depicted in Figure 1. These subcarriers are employed to capture characteristics such as signal scattering and multipath attenuation. The CSI (Channel State Information) signals exhibit consistency for the same gesture, while variations among CSI signals corresponding to different gestures enable differentiation between various types of gestures. Compared to the RSSI representation at the MAC layer, the CSI representation at the physical layer is less susceptible to multipath interference and provides a more precise characterization at each subcarrier.

Building upon this representation, researchers have attempted to overcome the inherent limitations of Wi-Fi signal sensing, such as low time resolution, vulnerability to interference, and narrow bandwidth. Qirong Bu et al. [22] extracted gesture segments based on changes in CSI amplitude and transformed the problem of Wi-Fi-based gesture recognition into an image classification task by representing CSI streams as image matrices and inputting them into deep learning networks for recognition. Zhanjun Hao et al. [23] established a correlation mapping between the amplitude and phase difference information of subcarrier levels in wireless signals and sign language gestures. They combined an

effective denoising method to filter environmental interference and efficiently selected optimal subcarriers, thereby reducing system computation costs. Li Tao et al. [24], utilizing the nexmon firmware, obtained 256 CSI subcarriers from the underlying layer of smartphones operating in IEEE 802.11 ac mode with an 80 MHz bandwidth. They then fused the extracted CSI features in the time and frequency domains using a cross-correlation method and ultimately recognized gestures using an improved DTW algorithm, breaking the limitation that gesture recognition can only be performed within relatively fixed regions along the transmission link.
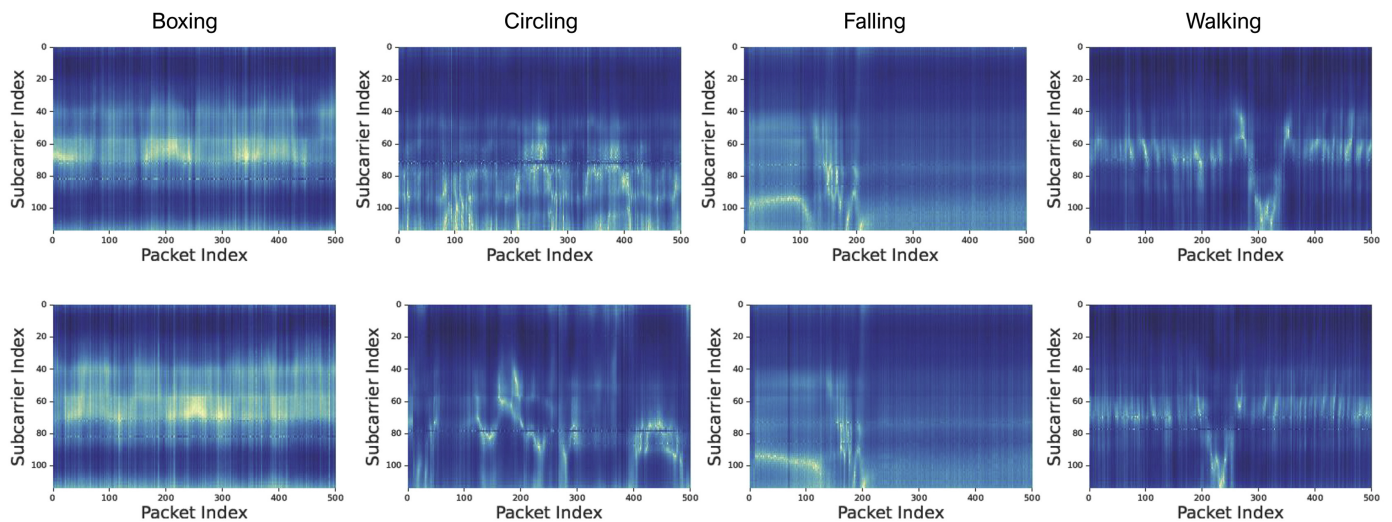
Figure 1. CSI signal samples during two human activities detected by Wi-Fi [25].

Compared to Wi-Fi gesture recognition, continuous wave radar gesture recognition technology, which utilizes Doppler information, demonstrates superior performance in dynamic recognition [26]. The general process is illustrated in Figure 2. Radar signals reflected in the backward direction contain abundant gesture-related information, including distance-Doppler information, angle-Doppler information, distance-angle information, and time-frequency information, enabling the differentiation of various types of gestures. Skaria Sruthy et al. [27] employed a low-cost dual-antenna Doppler radar to generate co-frequency and orthogonal components of the gesture signal, mapping them to three input channels of a deep convolutional neural network (DCNN). This approach yielded two spectrograms and an arrival angle matrix for recognition, achieving an accuracy rate exceeding 95%. Wang Yong et al. [28] developed a gesture recognition platform based on frequency-modulated continuous wave (FMCW) radar, which extracted features from the obtained range-Doppler map (RDM) and range-angle map (RAM). The accuracy rate for simultaneous recognition of gestures involving both hands surpassed 93.12%. Yan Baiju et al. [29] harnessed the secrecy and non-contact advantages of millimeter-wave radar to develop a gesture recognition system. They divided the collected data, including the distance-Doppler image (RDI), distance-angle image (RAI), Doppler-angle image (DAI), and micro-Doppler spectrum, into training and testing sets. Utilizing a semi-supervised learning (SSL) model, they achieved outstanding performance, with average accuracy rates of 98.59% for new users, 96.72% for new locations, and 97.79% for new environments.
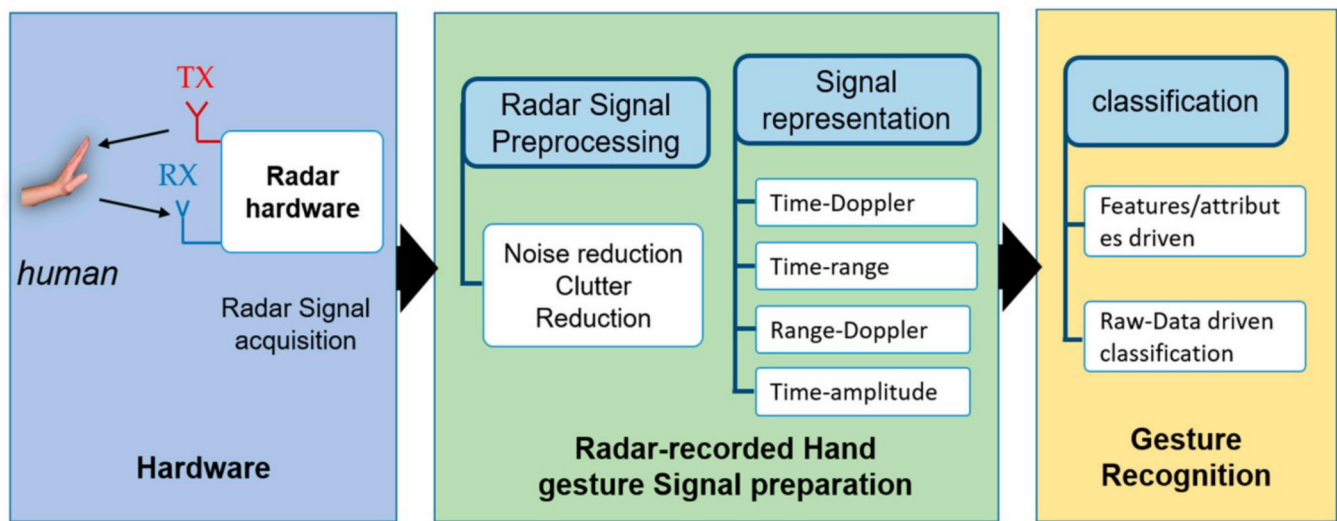
Figure 2. Radar gesture recognition process [30].

2.2. Strain-Sensing Recognition

Strain-sensing recognition technology is based on percolation theory and tunneling theory. It converts the changes in stress on the sensing element into corresponding electrical signals, which are then used for gesture recognition through machine learning algorithms. In the early stages, strain-sensing recognition devices often utilized rigid substrates, which resulted in mechanical incompatibility with the flexible human skin. Additionally, these devices were typically bulky and inconvenient to carry. However, with technological advancements, these devices have undergone miniaturization and are now widely integrated into data gloves and smart wristbands.

There are three main types of strain-sensing devices based on the method of signal extraction: piezoelectric [31,32], capacitive, and resistive. Piezoelectric sensors are self-generating electromechanical transducers. When subjected to stress, the material generates charges on its surface, which are then converted into electrical signals and amplified. The strength of the electrical signal is generally proportional to the applied pressure, enabling strain-sensing recognition. Capacitive sensing technology commonly employs parallel plate capacitors. When the excitation electrode and the sensing electrode are placed in parallel opposition, the electric field lines between the electrodes expand uniformly between the plates as they slowly separate. As shown in Figure 3, when the two electrodes reach a coplanar state, the edge electric field between them becomes dominant. Capacitive sensors utilize this effect to recognize dynamic gestures within the sensitive area [33].

Figure 3. Parallel plate capacitance to planar capacitance.

Resistive sensors transform non-electrical variations into changes in resistance by modifying the shape of a stress-sensitive elastic conductor. The typical structure of a resistive sensor, illustrated in Figure 4, includes a flexible contact layer for the first and fifth layers, a first electrode layer and a second electrode layer for the second and fourth layers, respectively, and a sensing layer typically positioned in the middle as the third layer. Compared to other sensor types, resistive sensors have garnered significant attention

due to their high sensitivity, wide measurement range, excellent reusability, and simple structure, making them a primary focus of research.
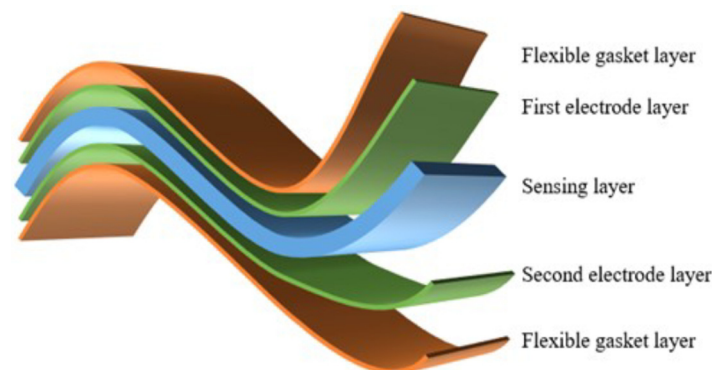
Figure 4. General structure diagram of the strain-sensing element.

In recent years, researchers have pursued investigations into approaches that involve coupling resistive sensors with other hierarchical or microstructures. The aim is to enhance biocompatibility and sensing accuracy. These structures introduce an intermediate functional layer, which acts as a composite material combining dielectric materials with specific components. In comparison to pure dielectric materials, this intermediate layer possesses a lower effective Young's modulus, making it more susceptible to deformation. When external pressure is applied, the microstructure and air–composite functional layer expel air, resulting in an increased proportion of components with high dielectric constant. This leads to a more pronounced change in capacitance compared to conventional hierarchical structures. Moreover, it offers a higher signal-to-noise ratio and improved stability. Zhu Penghua et al. [34] proposed a stretchable resistive thread sensor based on a composite material comprising silver-coated glass microspheres (Ag@GMs) and solid rubber (SR). Wang Shuai et al. [35] introduced a gradient porous/pyramid hybrid structure (GPPHS) conductive composite film with gradient compression characteristics and superior structural compressibility. This innovation simultaneously enhances the detection accuracy and range of resistive sensors. Liu Caixia et al. [36] developed a flexible resistive sensor with a crack structure and high strain coefficient. It is composed of a biodegradable and stretchable gelatin composite material integrated with a fabric substrate. Additionally, specific materials can be plated onto the contact surface between the conductive composite structure and the conductive layer. This allows the conductivity of the device to be influenced by changes in the distance between the composite material and the contact area with the conductive layer. As a result, the sensor exhibits an expanded measurement range and improved measurement accuracy [37].

For traditional resistive sensors, the most common method of signal classification is the fixed threshold judgment approach. This method is simple and fast, but it has significant drawbacks. It exhibits poor resistance to interference and lacks accuracy. Therefore, introducing efficient machine learning algorithms has become a key focus of research in this field. Fan Tianyi et al. [38] used binary neural networks and convolutional neural networks to process the collected signals. For a dataset consisting of 10×200 gesture samples, they achieved a recognition accuracy of 98.5%. Liu Caixia et al. [36] employed support vector machines to process data from nine types of gestures, greatly improving the recognition accuracy to an average of 99.5%. Wang Ziyi et al. [39] invented a gesture recognition system based on a planar capacitive array. They conducted experiments using a dataset of 6×25 gesture samples and achieved a recognition rate exceeding 95% by employing hidden Markov models. These results demonstrate the significant application potential of strain-sensing technology with the intervention of artificial intelligence algorithms.

2.3. EMG Sensing Recognition

When skeletal muscles are in a state of resting relaxation, the membrane potential of each muscle cell, also known as muscle fiber, is approximately -80 mV [40]. During muscle contraction, there is a potential difference between the muscle cells and the motor neurons that form the motor unit, leading to the conduction of electrical signals. Additionally, research suggests that the sum of action potentials within all muscle cells of a motor unit is referred to as the motor unit action potential. When a skeletal muscle contracts, the electromyographic (EMG) signal is the linear sum of the related motor units. The strength of the EMG signal is strongly correlated with the state and quantity of muscle contraction, thus providing possibilities for gesture recognition technology.

Based on the location of EMG signal generation, there are two types: surface EMG signals (sEMG) and intramuscular EMG (iEMG) signals, each with its own characteristics as shown in Table 1. Surface EMG signal sensing technology collects signals by placing non-invasive electrodes on the surface of muscles, providing information about muscle movement. It is widely used in medical rehabilitation and human–computer interaction fields due to its ability to accurately record actual action potentials by ignoring factors such as lighting conditions and occlusions. Furthermore, studies have shown that this method can capture relevant electrical signals approximately 200 ms before physical movement [40], which means that this technology also has the potential for action prediction. On the other hand, intramuscular EMG signal acquisition often requires inserting electrodes into specific points within muscle tissue to capture muscle activity signals. This invasive collection method unavoidably carries a slight risk of tissue damage, infection, and discomfort for the individual. Additionally, it has drawbacks such as high equipment costs and the need for frequent maintenance, making it unsuitable for flexible and dynamic human–computer interaction environments.

Table 1. Comparison of surface electromyographic signals and intramuscular electromyographic signals.

Characteristic	sEMG	iEMG
Signal source	Muscle Contraction	Muscle Contraction
Measurement position	Muscle surface	Intramuscular
Damaged tissue	No	Yes
Measurement difficulty	Relatively simple	Relatively difficult
Measurement accuracy	Relatively low	Relatively high
Application area	Gesture recognition, intelligent wear, rehabilitation therapy, and other fields	Muscle evaluation, biomedical engineering, and other fields

The application of sEMG signals can be traced back to the 1990s when researchers used muscle signals to control robot movements [41] for human–computer interaction. Over time, the collection technology for sEMG signals has been continuously improved and expanded, finding applications in gesture recognition technology. Similar to other gesture recognition approaches, sEMG-based gesture recognition primarily relies on improved feature extraction methods and machine learning algorithms to enhance recognition accuracy. Due to the nonlinear, stochastic, and highly variable nature of EMG signals, processing and analyzing the entire signal waveform become particularly challenging. Lv Zhongming et al. [42] proposed a feature selection and classification recognition algorithm based on a Self-Organizing Map (SOM) and Radial Basis Function (RBF), followed by Principal Component Analysis (PCA) to reduce the size of feature vectors. Anastasiev Alexey et al. [43] used a novel four-component multi-domain feature set and feature vector weight addition for signal segmentation and feature extraction, enabling the more accurate investigation of features and patterns during muscle contraction. Mahmoud Tavakoli et al. [44], Vimal Shanmuganathan et al. [45], and Guo Weiyu [46] respectively employed Support Vector Machines, R-CNN, and Long Exposure Neural Networks for processing the collected EMG signals, achieving recognition accuracies exceeding 95%.

Furthermore, EMG-based gesture recognition is also applicable to individuals with disabilities, as paralysis or amputation of the elbow or palm does not affect the generation of muscle contractions and EMG signal generation within the arm's neural and muscular system. By capturing and analyzing EMG signals, it becomes possible to recognize and control gestures and movements for people with disabilities. Gu Guoying et al. [47] invented a neuroprosthetic hand, as shown in Figure 5, which utilizes non-invasive electrodes placed on the amputee's elbow to capture EMG signals and control the neuroprosthetic hand. Furthermore, the pressure sensors on the fingertips of the prosthetic hand generate electrical stimulation to simulate tactile feedback. Such technology not only helps individuals with disabilities achieve normal hand movements but also assists doctors in better understanding the muscular condition of patients, facilitating their treatment.
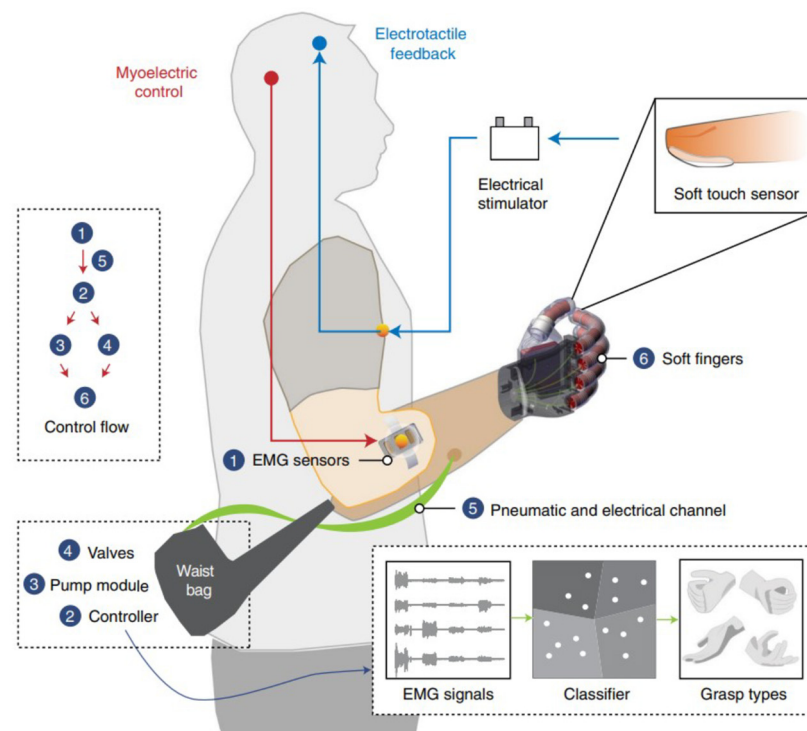
Figure 5. Structure diagram of the disabled artificial limb based on EMG gesture recognition [47].

### 2.4. Visual Sensing Recognition

Gesture recognition systems based on visual sensing have reached a relatively advanced stage of development. They recognize gestures by converting images containing human hands into color channel data and keypoint position information. Early research primarily focused on sign language translation and gesture control technology. With the continuous improvement and expansion of computer hardware and image processing techniques, visual sensing and recognition technology has become more precise and standardized.

During the image acquisition and preprocessing stage of visual sensing and recognition technology, human hand images are captured using image capture devices. Generally, higher-pixel cameras result in higher accuracy for visual gesture recognition, but they also increase the computational load. Therefore, preprocessing of the collected hand image is essential, as illustrated in Figure 6. The image undergoes processing techniques such as filtering, grayscale conversion, and binarization to extract the contour information and finger coordinates of the hand, while simplifying the data.

Given the variations in hand size, skin color, and shape among individuals, as well as the potential interference caused by scene lighting and hand motion speed, effectively segmenting the meaningful hand gesture from the complex background environment remains an urgent challenge. Clearly, relying solely on color segmentation is insufficient. In

2019, Manu Martin et al. proposed seven methods for extracting pixel or neighborhood-related features from color images. They evaluated and combined these methods using random forest [48]. Danilo Avola et al. [49] normalized the extracted images and employed a multi-task semantic feature extractor to derive 2D heatmaps and hand contours from RGB images. Finally, a viewpoint encoder was used to predict hand parameters. With the advancement of structured light technology and depth image capture, gesture recognition based on depth images has gained considerable attention. Mature commercial products such as Kinect and Leap Motion are currently available in this field. Depth images utilize grayscale values to represent the relative distances of pixels from the capturing system, reducing the impact of complex backgrounds on gesture segmentation to a certain extent. Moreover, the data structure of depth images is well-suited for input into artificial intelligence algorithms such as random forests and neural networks, effectively transforming complex feature extraction problems into simpler pixel classification tasks. Consequently, this approach is gradually emerging as the mainstream trend in gesture segmentation. Xu et al. [50] introduced traditional depth-based gesture segmentation algorithms, including RDF, R-CNN, YOLO, and SegNet. They enhanced the baseline SegNet algorithm by incorporating class weights, transpose convolutions, hybrid dilated convolution combinations, and concatenation and merging skip connections between the encoder and decoder. These enhancements resulted in F2-Score improvements of 7.6% and 5.9% for the left and right hand, respectively, compared to the baseline method.
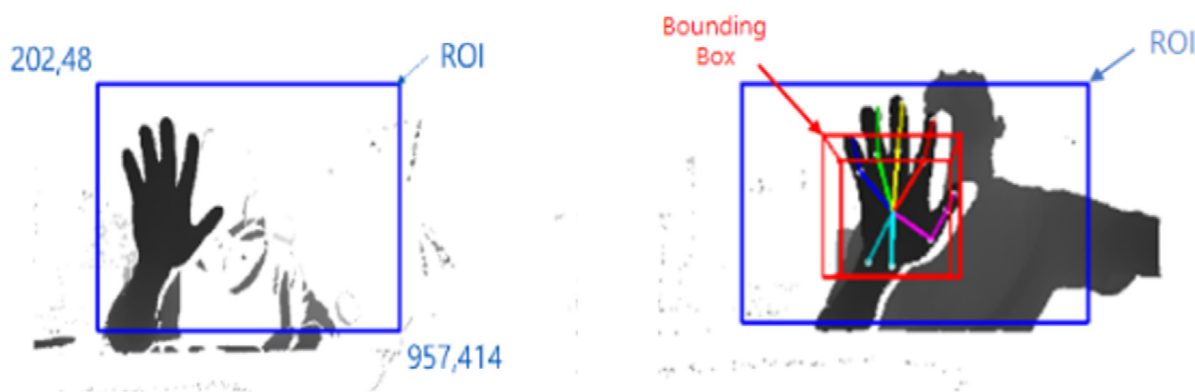
Figure 6. Hand boundary detection and hand key point detection [51].

After completing the aforementioned steps, the next phase involves extracting gesture-related signals from the preprocessed images, such as the number of fingers, finger length, finger opening angle, and more. These signals are then classified using machine learning algorithms, which are trained with a large number of samples to achieve accurate recognition. Khan Fawad Salam et al. [52] employed the MASK-RCNN combined with the Grass Hopper algorithm for classifying the obtained RGB images and hand keypoints. Jaya Prakash Sahoo et al. [53] developed an end-to-end fine-tuning method for pre-trained CNN models using fractional-level fusion technology. This approach demonstrates excellent gesture prediction performance even with a limited training dataset. Building upon this technology, the team designed and developed the real-time American Sign Language (ASL) recognition system [54].

3. Comparison and Analysis

Table 2 includes representative examples and relevant parameters of various detection methods. By analyzing these typical cases, we can infer the strengths and limitations of each detection method. Typically, larger datasets lead to higher accuracy and better performance of gesture recognition methods. However, it is crucial to consider the variations in experimental conditions, such as diverse environments and dataset formats used in different methods. As a result, the final results may exhibit certain biases. Therefore, this paper provides a comprehensive analysis of the pros and cons of each detection method.

Table 2. Comparison of cases of various testing methods.

Author	Time	Method	Algorithm	Data Set	Accuracy
Xiaorui Ma et al. [55]	1 January 2020	Wi-fi	Deep learning	14 volunteers × 26 datasets × 20 samples	Average over 90%
Jiang Dehao et al. [56]	23 August 2020	Wi-fi	SVM	Widar 3.0 and SignFi datasets	98% and 95.6%
Skaria Sruthy et al. [27]	15 April 2019	Dual antenna Doppler radar	DCNN	14 Gestures × 250 Samples	Over 95%
Wang Yong et al. [28]	14 May 2022	FMCW radar	3DCNN	Over 4500 data (3 Men + 2 women)	93.12%
Liu Caixia et al. [36]	6 February 2023	Resistive sensing element	SVM	9 × n Group Gestures	99.50%
Wang Ziyi et al. [39]	15 December 2022	Capacitive sensing element	HMM	6 × 25 sets of gestures	Average 95%
Yang Zhiwen et al. [57]	22 October 2021	sEMG	ResNet	Ninapro DB1	89.65%
Anastasiev Alexey et al. [43]	11 November 2022	sEMG	Four-component multidomain feature set and eigenvector weight addition	19 volunteers × 10 gestures	94.20–96.62% on one side of hemiplegia and 88.48–90.37% on the other side of non-hemiplegia.
Khan Fawad Salam et al. [52]	1 January 2021	Visual recognition	MASK-RCNN and Grass Hopper	Microsoft Kinect and Leap Motion Common Datasets	99.05% and 99.29%
Xu Chi et al. [58]	2 February 2021	Visual recognition	Deep learning	LaRED Common Datasets	99.96%

After comparison, it is evident that gesture recognition technologies employing machine learning and artificial intelligence algorithms generally achieve higher accuracy. Among them, electromagnetic wave-based gesture recognition stands out with its relatively high accuracy, non-contact nature, and independence from line-of-sight and lighting conditions. It is suitable for recognizing gestures of different sizes and shapes. However, this approach demands high hardware requirements and comes with a higher cost, sometimes necessitating multiple signal transmission devices to achieve high-precision results. Strain-based gesture recognition also demonstrates a high level of accuracy and strong reliability, allowing for operation in various environments. However, its gesture recognition range is limited, requiring the deployment of sensor hardware devices on the hand, which can impact accuracy based on the specific deployment environment. On the other hand, surface electromyography (sEMG) gesture recognition exhibits relatively lower accuracy due to significant noise in the electromyographic signals. It requires advanced denoising algorithms and inconveniently involves the placement of electrodes on the skin. Moreover, specific muscle training is necessary, and accuracy is influenced by muscle fatigue. However, sEMG has the capability to detect subtle muscle changes and enables gesture recognition for individuals with disabilities, thus holding significant importance in the medical field. Visual sensing-based gesture recognition achieves the highest accuracy, thanks to the rapid development of color and depth imaging technologies. It offers advantages such as low cost, ease of implementation, and non-contact operation. However, it is susceptible to signal distortions caused by background, lighting, and occlusion effects, necessitating image preprocessing. Furthermore, the recognition accuracy for random samples not present in the database requires further investigation.

Existing recognition technologies also face common issues that urgently require solutions. Despite the tremendous potential of gesture recognition technology, there are significant differences between static and dynamic gesture recognition, as well as variations in the semantic-syntactic structures of different gestures. Insufficient analysis algorithms, datasets, visual corpora, and other factors hinder the ability to perform in-depth semantic analysis. Currently, there is no fully automated model or method that can be widely applicable to multiple static or dynamic gesture recognition systems. In fact, as demonstrated in Table 3, numerous unimodal or multimodal corpora are already accessible to researchers. The STF-LSTM [59] deep neural network model architecture effectively incorporates gesture and lip movement information features, achieving a gesture recognition rate of 98.56% on

the extensive multimodal Turkish Sign Language dataset, AUTSL [60]. Notably, the article introduces a model compression technique utilizing ONNX Runtime, which significantly reduces the computational and memory requirements of the model. Consequently, this advancement facilitates seamless and efficient operation on prevalent mobile devices such as the Samsung Galaxy S22. Another notable framework, the SAM-SLR [61] architecture proposed by Songyao Jiang et al., attained recognition rates of 98.42% and 98.53% on the RGB and RGB-D tracks of AUTSL, respectively. Nonetheless, existing corpora still confront certain challenges, including limited accessibility, insufficient data volume and diversity, and the inclusion of only one type of dynamic or static gesture. These constraints hinder progress in deep semantic analysis. Furthermore, a research gap exists in the domain of gesture recognition concerning cutting-edge technologies such as femtosecond laser recognition, fiber optic sensing recognition, and acoustic recognition. Considering practical application scenarios, there is room for improvement in the comfort and portability of certain wearable devices.

Table 3. Partial hand dataset.

Dataset	Publication Time	Content
NVGesture [62]	2016	Driving scene gesture
RHD [63]	2017	Depth image and segmentation mask
FreiHAND [64]	2019	3D hand posture
AUTSL [60]	2020	RGB, depth, and skeleton formats
WLASL [65]	2020	Sign language video
SLOVO [66]	2023	Sign language dataset

4. Technology Application

Hand gesture recognition technology is characterized by convenience, intuitiveness, and intelligence, and it holds tremendous significance and potential for human production and daily life. This section aims to provide a comprehensive overview of the applications of hand gesture recognition technology in various domains of modern production and daily life.

4.1. Improved Traditional Control Mode

Gesture recognition technology-based human–computer interaction control methods offer several advantages over traditional approaches. Firstly, they align more naturally with human habits, eliminating the need for additional learning of input devices. Secondly, gesture recognition technology allows for the design of diverse gestures to meet the specific requirements of different application scenarios. This enables more flexible and intuitive interaction in fields such as gaming, healthcare, and education. Extensive research and applications in various domains have demonstrated the significant potential of gesture recognition-based human–computer interaction control methods. For example, Strazdas Dominykas et al. [67] developed a non-contact multimodal human–computer interaction system called Rosa, which integrates gesture recognition, facial recognition, and speech technologies to efficiently and securely control mechanical systems. Su Mu Chun et al. [68] proposed a gesture recognition-based home appliance control system that achieved a recognition accuracy of 91% and allowed wireless control of household appliances using a small set of gestures. In the field of unmanned aerial vehicles (UAVs), gesture recognition technology has been employed as an alternative to traditional joystick controls. Lee JiWon et al. [69] implemented a UAV gesture control system based on IMU (inertial measurement unit) recognition components. During control, obstacle information on the UAV's heading can be conveyed to the user through vibration feedback, enhancing safety. Konstantoudakis Konstantinos et al. [70] developed an AR-based single-hand control system for UAVs, addressing the issue of visual fatigue caused by prolonged focus on joysticks and control screens. Moreover, for highly dynamic underwater UAV environments, Yu Jiang et al. [71] proposed a gesture interaction system for underwater autonomous underwater vehicles

(AUVs) that employed fuzzy control to overcome challenges such as fluctuation interference and light attenuation. These studies collectively highlight the efficiency, safety, and speed advantages offered by gesture recognition technology compared to traditional methods, positioning it as a core technology for emerging human–computer interaction control systems.

4.2. Medical Applications

Gesture recognition technology has found extensive applications in the medical field, playing a crucial role in improving the efficiency and quality of healthcare services, facilitating convenient patient care by medical professionals, and aiding patients in their rehabilitation and return to a normal life. A noteworthy example is the work of Korayem M.H et al. [72], who designed a high-precision remote laparoscopic surgery system based on the Leap Motion platform. This system enables skilled surgeons to perform laparoscopic procedures on patients regardless of geographical constraints. Xie Baao et al. [64] and Gu Guoying et al. [47] utilized electromyography (EMG) recognition technology to achieve gesture recognition based on the residual limbs of disabled individuals. By accurately controlling prosthetic hands through EMG signals, they empowered disabled patients to independently perform actions such as grasping and placing objects, significantly enhancing their quality of life (Figure 7b). Stroh Ashley et al. [73] addressed the challenges faced by individuals with conditions such as cerebral palsy or muscle malnutrition, who struggle with precise muscle control required for operating electric wheelchairs using traditional joysticks. They devised an electric wheelchair control system based on EMG gesture recognition, offering a solution that improves mobility for individuals with limited muscle control. Additionally, Nourelhoda M. Mahmoud et al. [74] developed a remote patient monitoring system that tracks patients' hand movements, detects pain levels, and monitors muscle function recovery. This system holds great significance in disease monitoring and the adjustment of subsequent treatment plans.

4.3. Physical Training

In sports training, hand gestures can effectively reflect the athletes' performance. By calibrating hand postures and analyzing movements, gesture recognition technology helps coaches gain a comprehensive understanding of athletes' skill levels and provides targeted training recommendations. Therefore, gesture recognition technology plays a crucial role in the field of sports. Li Guangjing et al. [75] analyzed static images and video sequences and proposed a method that utilizes multi-scale feature approximation to improve the speed of hand feature extraction. This advancement in athlete posture analysis provides a theoretical foundation for subsequent analysis of athlete gesture movements. Rong Ji [76] introduced an approach based on image feature extraction and machine learning for recognizing basketball shooting gestures. By classifying hand movements during shooting, this method provides a scientific basis for training basketball shooting techniques. Shuping Xu et al. [77] conducted similar work focused on table tennis, significantly enhancing the efficiency of analyzing match recordings of highly skilled table tennis players during training sessions. Furthermore, gesture recognition technology can be applied to the professional training of referees, enhancing their accuracy in interpreting rules and making real-time assessments of game situations, thereby ensuring the fairness of sports competitions. Tse-Yu Pan et al. [78] developed a referee training system where trainees wear MYO electromyography (EMG) gesture recognition armbands to watch pre-recorded match videos. The system facilitates training and corrective actions based on the consistency between the trainee's EMG signals and the official EMG signals from the recorded videos. Paulo Trigueiros et al. [79] created an application that can recognize the hand gestures of the main referee in real-time matches, assisting assistant referees and video assistant referees (VARs) in making real-time judgments on the game situation.

4.4. Other Areas

In addition to the three domains mentioned above, gesture recognition technology has found wide applications in various other technical fields. As shown in Figure 7d, Wang Xin et al. [80] developed a robotic system to address issues such as low productivity, low safety, and labor shortages in construction sites. By analyzing the hand gestures of workers through visual analysis, they laid the technological foundation for the subsequent development of mechanical hands for construction workers. Alexander Riedel et al. [81] used visual gesture recognition to analyze the hand movements of a large number of workshop workers, enabling them to predict industry-standard production times for assembly line design and product cost estimation. Through this quantitative analysis of hand gestures, accurate predictions of future data trends can be made.
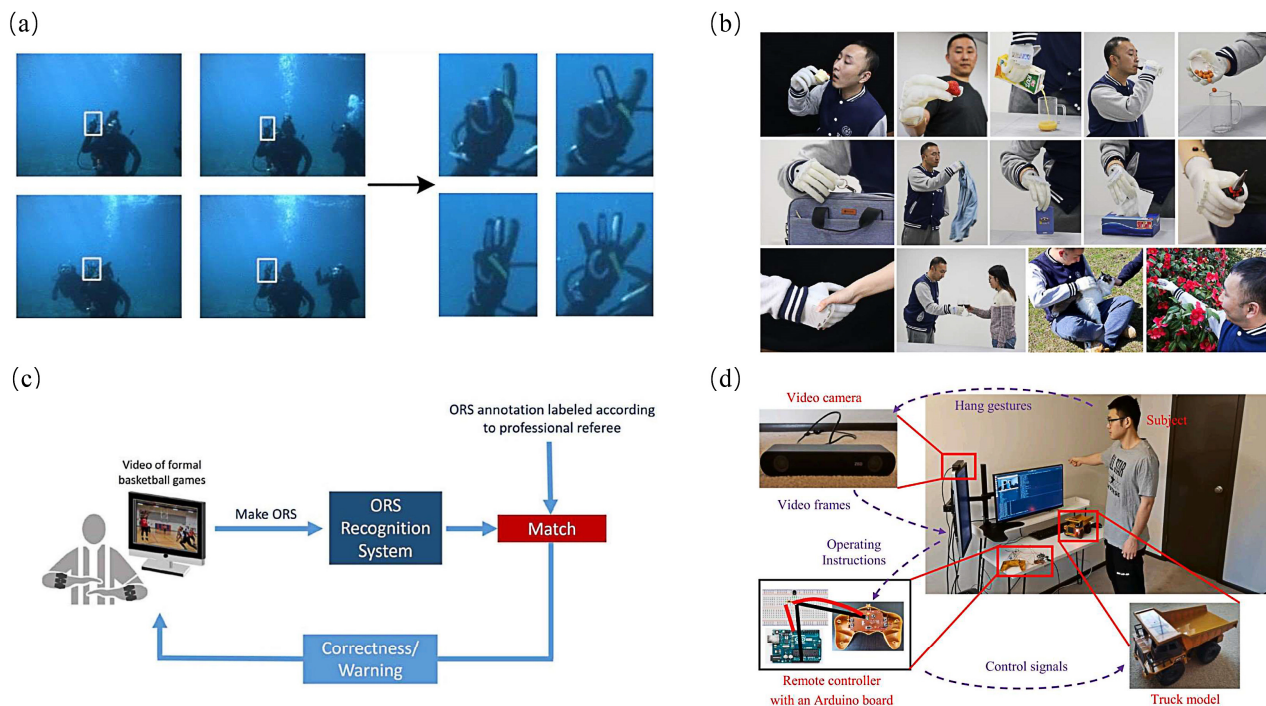
Figure 7. Application of gesture recognition technology in modern production and life (**a**) gesture recognition screen of autonomous underwater vehicle [71]. (**b**) Artificial hand based on EMG gesture recognition helps the disabled to live normally [47]. (**c**) Flow chart of a referee training system based on gesture recognition [79]. (**d**) Robot system based on gesture recognition of construction workers [80].

5. Future Outlook

Gesture recognition-based human–computer interaction technology is an emerging field with immense potential for development. In recent years, it has attracted considerable attention as a core technology in human–computer interaction. The performance of gesture recognition technology is expected to make significant advancements in the future. In this section, we will analyze its potential directions of development in four parts.

5.1. Biocompatibility and Wearability

The wearability of gesture recognition devices will become a key focus of future development. This is due to the current technological limitations that result in the mismatch between rigid materials and flexible skin, as well as the drawbacks of large volume and weight in existing sensing devices. The advancement of new material technologies brings new possibilities to gesture recognition technology. Some materials not only exhibit good biocompatibility but also enhance the sensing performance of sensors. Combining sens-

ing systems with lightweight materials is becoming a necessary trend in order to select appropriate materials based on the specific usage environment.

5.2. Adaptability

Virtually all gesture recognition technologies rely on machine learning or artificial intelligence algorithms to classify gestures. From numerous research findings, it is evident that the combination of different machine learning techniques and gesture recognition implementation methods leads to varying levels of accuracy. In future research, in addition to advancing more efficient algorithms, researchers should also investigate the compatibility between different technologies, algorithms, and application domains. This includes efforts to reduce usage and manufacturing costs, enhance product usability, and identify the optimal combinations.

5.3. Stability and Robustness

The current gesture recognition technology is significantly influenced by the surrounding environment. Gesture recognition in practical applications needs to take into account factors such as lighting, medium properties, and occlusions. It goes beyond the recognition of static gestures from a specific database. Improving the stability and reliability of gesture recognition technology requires mitigating the impact of variations in gesture speed and motion trajectories during movement. The next step in the research will involve optimizing both hardware and algorithms to address these challenges effectively.

5.4. Overlapping

The effectiveness of gesture recognition varies in different environments depending on the chosen implementation approach. Gesture recognition technology should not be confined to a single sensing method. By combining two or more methods, such as electromagnetic wave sensing, strain sensing, electromyography (EMG) sensing, and visual sensing, or by exploring emerging sensing methods such as fiber optic sensing and acoustic sensing, the reliability and robustness of the technology can be significantly improved. The integration of multiple signals in applications will greatly enhance the overall performance of the technology.

6. Conclusions

This article comprehensively examined the principal implementation methods employed in recent years for gesture recognition-based human–computer interaction technologies. Specifically, it focused on electromagnetic wave sensing technology, strain sensing technology, EMG sensing technology, and visual sensing technology. The article also highlighted the latest advancements in these implementation methods and their associated technologies. Our study presented, analyzed, and discussed the findings derived from an extensive review of 73 pertinent publications pertaining to this technology. Below are our findings in regard to the four implementation methods.

Electromagnetic wave sensing recognition: We primarily focused on two categories of implementation methods: Wi-Fi sensing technology and radar sensing technology. We provided an overview of the disturbance model and the Doppler effect in the electromagnetic wave transmission process. Building upon this foundation, we delved into the characteristics of commonly used RSSI channel description methods and CSI channel description methods in Wi-Fi sensing technology. Furthermore, we explored various channel description methods based on Doppler information utilized in radar sensing technology. In comparison to other recognition technologies, electromagnetic wave sensing gesture recognition demonstrates relatively higher accuracy, non-contact nature, and immunity to line-of-sight and lighting conditions. It is suitable for gestures of varying sizes and shapes. However, achieving high recognition accuracy heavily relies on the quality of collected channel information, thereby imposing stringent requirements on acquisition

devices and entailing higher costs. In certain cases, multiple signal-transmitting devices may be necessary to obtain highly precise results.

Strain sensing recognition: Drawing upon percolation theory and tunneling theory, the change in stress applied to sensing elements is effectively converted into variations in electrical signals to facilitate recognition. We elucidated the underlying principles of three stress-based gesture recognition techniques: piezoelectric, capacitive, and resistive methods. Furthermore, we delved into the enhancements introduced by researchers in terms of sensing layer materials, structural considerations, and pertinent algorithms for electrical signal processing. This technology demonstrates robust recognition reliability and can be effectively deployed in diverse operational environments. However, its gesture recognition capabilities are constrained, necessitating the deployment of sensor hardware on the hand, thereby influencing accuracy. Aspects such as sensor reusability, mitigation of hair growth interference, and resilience to electromagnetic interference warrant further exploration and discussion by future scholars.

EMG sensing recognition: Implemented primarily through the utilization of generated potential differences during muscle movement, EMG technology can be classified into two categories: sEMG and iEMG. A comparative analysis of their respective features reveals that the non-invasive nature of sEMG renders it more suitable for practical gesture recognition applications. Given the substantial noise inherent in EMG signals, demanding noise reduction algorithms are imperative. We explored researchers' endeavors in signal segmentation, highlighting that the integration of artificial intelligence algorithms significantly enhances the accuracy of EMG signal recognition. Nonetheless, successful implementation of this technology still necessitates specific muscle training, with precision exhibiting a certain degree of correlation with muscle fatigue. Furthermore, we discussed relevant efforts in employing EMG signal-based gesture recognition for individuals with disabilities, emphasizing the potential benefits not only in facilitating natural hand movements but also in providing healthcare practitioners with invaluable insights into patients' muscle conditions, thereby aiding in treatment in and holding significant implications for the medical field.

Visual sensing recognition: By converting images containing human hands into color data or depth data, we explored the advancements made by researchers in various aspects of gesture recognition, including hand segmentation methods and coordinate representation techniques. With the rapid development of depth imaging technology and the introduction of devices such as Kinect, this technology has exhibited additional features such as cost-effectiveness, ease of implementation, and non-contact operation. However, it is worth noting that the literature predominantly focuses on discussing the recognition accuracy of samples from existing databases, while further investigation is needed to assess the recognition accuracy of random samples not included in the database.

Technological applications and future work: We outlined some of the transformative effects that gesture recognition technology has brought to modern life and industrial production. Furthermore, we explored the potential future directions of this technology. In the coming years, there will be an increasing demand for gesture recognition technology, accompanied by higher performance expectations. To delve deeper, our next steps will involve thorough investigations into both materials and algorithms. The development of novel materials technology has given rise to high-quality structures that are lightweight, possess high electrical conductivity, and have low Young's modulus. The maturation of 3D printing has made it possible to fabricate sensing layers with microstructured surfaces, enabling sensors to achieve broader detection ranges and heightened sensitivity. The advent of unsupervised learning methods, such as deep learning and cluster learning, allows for the formation of models even when target data lack labels. This provides a means of data fitting for emerging gesture recognition technologies, including fiber optic sensing recognition and acoustic sensing recognition, which deal with complex signals. Based on these insights, we can anticipate the emergence of gesture recognition systems that are highly wearable, exhibit exceptional stability, and demonstrate robustness.

Abbreviations

The following abbreviations are used in this manuscript:

MAC	Medium Access Control
RSSI	Received Signal Strength Indication
OFDM	Orthogonal Frequency Division Multiplexing
CSI	Channel State Information
DCNN	Deep Convolutional Neural Network
FMCW	Frequency-modulated Continuous Wave
RDM	Range-Doppler Map
RAM	Range-angle Map
RDI	Distance-Doppler Image
RAI	Distance-angle Image
DAI	Doppler-angle Image
SSL	Semi-supervised learning
Ag@GMs	Silver-coated Glass Microspheres
SR	Solid Rubber
GPPHS	Gradient Porous/pyramid Hybrid Structure
EMG	Electromyographic
sEMG	Surface Electromyographic
iEMG	Intramuscular Electromyographic
SOM	Self-Organizing Map
RBF	Radial Basis Function
PCA	Principal Component Analysis
RDF	Random Forest
R-CNN	Region with Convolutional Neural Networks Features
YOLO	You Only Look Once
SVM	Support Vector Machine
HMM	Hidden Markov Model
UAVs	Unmanned Aerial Vehicles
AUVs	Autonomous Underwater Vehicles

References

1. Lin, Z.W.; Zhang, G.Q.; Xiao, X.; Christian, A.; Zhou, Y.H.; Sun, C.C.; Zhou, Z.H.; Yan, R.; Fan, E.; Si, S.B.; et al. A Personalized Acoustic Interface for Wearable Human–Machine Interaction. *Adv. Funct. Mater.* **2021**, *32*, 2109430. [CrossRef]
2. Yang, B.; Li, L.; Jerry, C.; Liu, J.; Chen, Y.Y.; Yu, J.D. Acoustic-based sensing and applications: A survey. *Comput. Netw.* **2020**, *181*, 107447.
3. Peng, Z.Q.; Wen, H.Q.; Jian, J.N.; Andrei, G.; Wang, M.H.; Huang, S.; Liu, H.; Mao, Z.H.; Chen, K.P. Identifications and classifications of human locomotion using Rayleigh-enhanced distributed fiber acoustic sensors with deep neural networks. *Sci. Rep.* **2020**, *10*, 21014. [CrossRef]
4. Li, J.C.; Wang, F.; Huang, H.Y.; Qi, F.F.; Pan, J.H. A novel semi-supervised meta learning method for subject-transfer brain–computer interface. *Neural Netw.* **2023**, *163*, 195–204. [CrossRef]
5. Liao, W.Z.; Li, J.H.; Zhang, X.S.; Chen, L. Motor imagery brain–computer interface rehabilitation system enhances upper limb performance and improves brain activity in stroke patients: A clinical study. *Front. Hum. Neurosci.* **2023**, *17*, 1117670. [CrossRef] [PubMed]
6. Greenwell, D.; Vanderkolff, S.; Feigh, J. Understanding De Novo Learning for Brain Machine Interfaces. *J. Neurophysiol.* **2023**, *129*, 749–750. [CrossRef]

7. Li, Q.; Sun, M.Q.; Song, Y.; Zhao, T.J.; Zhang, Z.L.; Wu, J.L. Mixed reality-based brain computer interface system using an adaptive bandpass filter: Application to remote control of mobile manipulator. *Biomed. Signal Process. Control* **2023**, *83*, 104646. [CrossRef]

8. Liu, Y.Y.; Wang, W.B.; Feng, C.X.; Zhang, H.Y.; Chen, Z.; Zhan, Y.B. Expression snippet transformer for robust video-based facial expression recognition. *Pattern Recognit.* **2023**, *138*, 109368. [CrossRef]

9. Zhang, R.Y.; Ma, R.L. Facial expression recognition method based on PSA—YOLO network. *Front. Neurorobot.* **2023**, *16*, 1057983.

10. Castillo, T.G.; RoigMaimó, M.F.; MascaróOliver, M.; Esperança, A.; Ramon, M. Understanding How CNNs Recognize Facial Expressions: A Case Study with LIME and CEM. *Sensors* **2022**, *23*, 131.

11. Sun, Z.; Zhang, H.H.; Bai, J.T.; Liu, M.Y.; Hu, Z.P. A discriminatively deep fusion approach with improved conditional GAN (im-cGAN) for facial expression recognition. *Pattern Recognit.* **2023**, *135*, 109157. [CrossRef]

12. Ma, X.L.; Zong, Y.P. The application of body language in the teaching of "Morality and Rule of Law" in junior middle school. *Knowl. Econ.* **2019**, *511*, 169–170.

13. Hu, Y.H.; Chen, Y.D.; Zhang, T. A review of sensor-based wrist motion detection and gesture recognition methods. *Sens. Microsyst.* **2022**, *41*, 1–3.

14. Lu, Y.; Lu, S.H.; Wang, X.D. Review on human behavior perception technology based on WiFi signal. *J. Comput. Sci.* **2019**, *42*, 1–21.

15. Rappaport, T.S. *Wireless Communications: Principles and Practice*; Prentice Hall: Hoboken, NJ, USA, 1996.

16. Wang, Y.; Wu, K.; Ni, L.M. WiFall: Device-free fall detection by wireless networks. *IEEE Trans. Mob. Comput.* **2017**, *16*, 581–594. [CrossRef]

17. Alex, M.; Ann, B. *The Story of Light Goes from Atom to Galaxy*; Fu, Z.X.; Lin, B.X., Translators; China University of Science and Technology Press: Hefei, China, 2015; Volume 8, p. 147.

18. Bianchi, V.; Ciampolini, P.; Munari, I.D. RSSI-Based Indoor Localization and Identification for ZigBee Wireless Sensor Networks in Smart Homes. *IEEE Trans. Instrum. Meas.* **2019**, *68*, 566–575. [CrossRef]

19. Carpi, F.; Davoli, L.; Martalo, M. RSSI-based Me-thods for LOS/NLOS Channel Identification Indoor Scenarios. In Proceedings of the 2019 16th International Symposium on Wireless communication Systems (ISWCS), Oulu, Finland, 27–30 August 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 171–175.

20. Han, K.; Xing, H.; Deng, Z.; Du, Y.C. A RSSIPDR-based probabilistic position selection algorithm with NLOS identification for indoor localization. *ISPRS Int. J. Geo-Inf.* **2018**, *7*, 232. [CrossRef]

21. Li, J.A. Cross Domain Gesture Recognition Based on WiFi Channel State Information. Master's thesis, Hefei University of Technology, Hefei, China, 2022. [CrossRef]

22. Bu, Q.; Yang, G.; Ming, X.; Zhang, T.; Feng, J.; Zhang, J. Deep transfer learning for gesture recognition with WiFi signals. *Pers. Ubiquitous Comput.* **2020**, *26*, 543–554. [CrossRef]

23. Hao, Z.; Duan, Y.; Dang, X.; Liu, Y.; Zhang, D. Wi-SL: Contactless Fine-Grained Gesture Recognition Uses Channel State Information. *Sensors* **2020**, *20*, 4025. [CrossRef]

24. Li, T.; Shi, C.; Li, P.; Chen, P. A Novel Gesture Recognition System Based on CSI Extracted from a Smartphone with Nexmon Firmware. *Sensors* **2020**, *21*, 222. [CrossRef]

25. Yang, J.; Chen, X.; Zou, H.; Lu, C.X.; Wang, D.; Sun, S.; Xie, L. SenseFi: A library and benchmark on deep-learning-empowered WiFi human sensing. *Patterns* **2023**, *4*, 100703. [CrossRef] [PubMed]

26. Zhang, Z.Y.; Tian, Z.S.; Zhou, M. Latern: Dynamic Continuous Hand Gesture Recognition Using FMCW Radar Sensor. *IEEE Sens. J.* **2018**, *18*, 3278–3289. [CrossRef]

27. Skaria, S.; Hourani, A.; Lech, M.; Evans, R.B. Hand-Gesture Recognition Using Two-Antenna Doppler Radar with Deep Convolutional Neural Networks. *IEEE Sens. J.* **2019**, *19*, 3041–3048. [CrossRef]

28. Wang, Y.; Wang, D.; Fu, Y.H.; Yao, D.K.; Xie, L.B.; Zhou, M. Multi-Hand Gesture Recognition Using Automotive FMCW Radar Sensor. *Remote Sens.* **2022**, *14*, 2374. [CrossRef]

29. Yan, B.J.; Wang, P.; Du, L.D.; Chen, X.X.; Fang, Z.; Wu, Y.R. mmGesture: Semi-supervised gesture recognition system using mmWave radar. *Expert Syst. Appl.* **2023**, *213*, 119042. [CrossRef]

30. Ahmed, S.; Kallu, K.D.; Ahmed, S.; Cho, S.H. Hand Gestures Recognition Using Radar Sensors for Human-Computer-Interaction: A Review. *Remote Sens.* **2021**, *13*, 527. [CrossRef]

31. Szumilas, M.; Władziński, M.; Wildner, K. A Coupled Piezoelectric Sensor for MMG-Based Human-Machine Interfaces. *Sensors* **2021**, *21*, 8380. [CrossRef]

32. Cha, Y.S.; Seo, J.; Kim, J.S.; Park, J.M. Human–computer interface glove using flexible piezoelectric sensors. *Smart Mater. Struct.* **2017**, *26*, 057002. [CrossRef]

33. Ma, M.; Guo, P.F.; Dong, Y.Z. Defect de-tection of composite materal based on dual-plane ca-pacitance sensor. *Comput. Appl. Soft-Ware* **2021**, *38*, 82–86.

34. Zhu, P.H.; Zhu, J.; Xue, X.F.; Song, Y.T. Stretchable Filler/Solid Rubber Piezoresistive Thread Sensor for Gesture Recognition. *Micromachines* **2021**, *13*, 7. [CrossRef]

35. Wang, S.; Zhang, Z.X.; Yang, B.; Zhang, X.; Shang, H.M.; Jiang, L.; Liu, H.; Zhang, J.; Hu, P.A. High sensitivity tactile sensors with ultrabroad linear range based on gradient hybrid structure for gesture recognition and precise grasping. *Chem. Eng. J.* **2023**, *457*, 141136. [CrossRef]

36. Liu, C.X.; Sun, Y.F.; Liu, P.; Ma, F.; Wu, S.G.; Li, S.; Hu, R.H.; Wang, Z.H.; Wang, Y.B.; Liu, G.Q.; et al. Fabrication and characterization of highly sensitive flexible strain sensor based on biodegradable gelatin nanocomposites and double strain layered structures with crack for gesture recognition. *Int. J. Biol. Macromol.* **2023**, *231*, 123568. [CrossRef] [PubMed]

37. Yan, L.X.; Mi, Y.J.; Lu, Y.; Qin, Q.H.; Wang, X.Q.; Meng, J.J.; Liu, F.; Wang, N.; Cao, B. Weaved piezoresistive triboelectric nanogenerator for human motion monitoring and gesture recognition. *Nano Energy* **2022**, *96*, 107135. [CrossRef]

38. Fan, T.Y.; Liu, Z.Y.; Luo, Z.W.; Li, J.D.; Tian, X.Y.; Chen, Y.S.; Feng, Y.; Wang, C.L.; Bi, H.C.; Li, X.M.; et al. Analog Sensing and Computing Systems with Low Power Consumption for Gesture Recognition. *Adv. Intell. Syst.* **2020**, *3*, 2000184. [CrossRef]

39. Wang, Z.Y.; Shen, S.M.; She, S.J. Dynamic Gesture Recognition Technology Based on Planar Capacitive Sensor Array. *J. Test. Technol.* **2023**, *37*, 54–59.

40. Xu, Z.; Lv, J.; Pan, W.J. Analysis of upper extremity fatigue based on surface EMG signal and motion capture. *J. Biomed. Eng.* **2022**, *39*, 92–102.

41. Hudgins, B.; Parker, P.; Scott, R.N. A new strategy for multifunction myoelectric control. *IEEE Trans. Bio-Med. Eng.* **1993**, *40*, 82–94. [CrossRef]

42. Lv, Z.M.; Xiao, F.Y.; Wu, Z.; Liu, Z.S.; Wang, Y. Hand gestures recognition from surface electromyogram signal based on self-organizing mapping and radial basis function network. *Biomed. Signal Process. Control* **2021**, *68*, 102629. [CrossRef]

43. Anastasiev, A.; Kadone, H.; Marushima, A.; Watanabe, H.; Zaboronok, A.; Watanabe, S.; Matsumur, A.; Suzuki, K.; Matsumaru, Y.; Ishikawa, E. Supervised Myoelectrical Hand Gesture Recognition in Post-Acute Stroke Patients with Upper Limb Paresis on Affected and Non-Affected Sides. *Sensors* **2022**, *22*, 8733. [CrossRef]

44. Mahmoud, T.; Carlo, B.; Pedro, A.L.; Osorio, L.B.; Almeda, A.T. Robust hand gesture recognition with a double channel surface EMG wearable armband and SVM classifier. *Biomed. Signal Process. Control* **2018**, *46*, 121–130.

45. Vimal, S.; Harold, R.Y.; Mohammad, S.K.; Manju, K.; Amir, H.G. Gandomi. R-CNN and wavelet feature extraction for hand gesture recognition with EMG signals. *Neural Comput. Appl.* **2020**, *32*, 16723–16736.

46. Guo, W.; Ma, C.; Wang, Z.; Zhang, H.; Farina, D.; Jiang, N.; Lin, C. Long exposure convolutional memory network for accurate estimation of finger kinematics from surface electromyographic signals. *J. Neural Eng.* **2021**, *18*, 026027. [CrossRef] [PubMed]

47. Gu, G.Y.; Zhang, N.B.; Xu, H.P.; Lin, S.T.; Yu, Y.; Chai, G.H.; Ge, L.; Yang, H.L.; Shao, Q.W.; Sheng, X.J.; et al. A soft neuroprosthetic hand providing simultaneous myoelectric control and tactile feedback. *Nat. Biomed. Eng.* **2021**, *7*, 589–598. [CrossRef] [PubMed]

48. Manu, M.; Thang, N.; Shahrouz, Y.; Li, B. Comprehensive features with randomized decision forests for hand segmentation from color images in uncontrolled indoor scenarios. *Multimed. Tools Appl.* **2019**, *78*, 20987–21020.

49. Avola, D.; Cinque, L.; Fagioli, A.; Foresti, G.L.; Fragomeni, A.; Pannone, D. 3D hand pose and shape estimation from RGB images for keypoint-based hand gesture recognition. *Pattern Recognit.* **2022**, *129*, 108762. [CrossRef]

50. Xu, Z.Z.; Zhang, W.J. An improved method for left and right hand synchronous segmentation based on depth images. *J. Shanghai Univ. (Nat. Sci. Ed.)* **2021**, *27*, 454–465.

51. Chanhwi, L.; Jaehan, K.; Seoungbae, C.; Jinwoong, K.; Jisang, Y.; Soonchul, K. Development of Real-Time Hand Gesture Recognition for Tabletop Holographic Display Interaction Using Azure Kinect. *Sensors* **2020**, *20*, 4566.

52. Khan, F.S.; Mohd, N.H.; Soomro, D.M.; Bagchi, S.; Khan, M.D. 3D Hand Gestures Segmentation and Optimized Classification Using Deep Learning. *IEEE Access* **2021**, *9*, 131614. [CrossRef]

53. Sahoo, J.P.; Prakash, A.J.; Pławiak, P.; Samantray, S. Real-Time Hand Gesture Recognition Using Fine-Tuned Convolutional Neural Network. *Sensors* **2022**, *22*, 706. [CrossRef]

54. Barczak, A.; Reyes, N.; Abastillas, M.; Piccio, A.; Susnjak, T. A new 2D static hand gesture colour image dataset for ASL gestures. *Res. Lett. Inf. Math. Sci.* **2011**, *15*, 12–20.

55. Ma, X.R.; Zhao, Y.N.; Zhang, L.; Gao, Q.H.; Pan, M.; Wang, J. Practical Device-Free Gesture Recognition Using WiFi Signals Based on Metalearning. *IEEE Trans. Ind. Inform.* **2020**, *16*, 228–237. [CrossRef]

56. Jiang, D.H.; Li, M.Q.; Xu, C.L. WiGAN: A WiFi Based Gesture Recognition System with GANs. *Sensors* **2020**, *20*, 4757. [CrossRef] [PubMed]

57. Yang, Z.W.; Jiang, D.; Sun, Y.; Tao, B.; Tong, X.L.; Jiang, G.Z.; Xu, M.M.; Yun, J.T.; Liu, Y.; Chen, B.J.; et al. Dynamic Gesture Recognition Using Surface EMG Signals Based on Multi-Stream Residual Network. *Front. Bioeng. Biotechnol.* **2021**, *9*, 969. [CrossRef]

58. Xu, C.; Jiang, Y.K.; Zhou, J.; Liu, Y. Semi-Supervised Joint Learning for Hand Gesture Recognition from a Single Color Image. *Sensors* **2021**, *21*, 1007. [CrossRef]

59. Ryumin, D.; Ivanko, D.; Ryumina, E. Audio-Visual Speech and Gesture Recognition by Sensors of Mobile Devices. *Sensors* **2023**, *23*, 2284. [CrossRef]

60. Zimmermann, C.; Ceylan, D.; Yang, J.; Russell, B.; Argus, M.; Brox, T. Freihand: A dataset for markerless capture of hand pose and shape from single rgb images. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Republic of Korea, 27 October–2 November 2019; pp. 813–822.

61. Jiang, S.; Sun, B.; Wang, L.; Bai, Y.; Li, K.P.; Fu, Y. Skeleton aware multi-modal sign language recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 3413–3423.

62. Yang, Y.; Newsam, S. Bag-of-visual-words and spatial extensions for land-use classification. In Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, San Jose, CA, USA, 3–5 November 2010.

63. Zimmermann, C.; Brox, T. Learning to estimate 3d hand pose from single rgb images. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 4903–4911.

64. Sincan, O.M.; Keles, H.Y. AUTSL: A Large Scale Multi-Modal Turkish Sign Language Dataset and Baseline Methods. *IEEE Access* **2020**, *8*, 181340–181355. [CrossRef]

65. Li, D.; Rodriguez, C.; Yu, X.; Li, H.D. Word-level deep sign language recognition from video: A new large-scale dataset and methods comparison. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Snowmass Village, CO, USA, 1–5 March 2020; pp. 1459–1469.

66. Kapitanov, A.; Kvanchiani, K.; Nagaev, A.; Petrova, E. Slovo: Russian Sign Language Dataset. *arXiv* **2023**, arXiv:2305.14527.

67. Strazdas, D.; Hintz, J.; Khalifa, A.; Abdelrahman, A.A.; Hempel, T.; Al-Hamadi, A. Robot System Assistant (RoSA): Towards Intuitive Multi-Modal and Multi-Device Human-Robot Interaction. *Sensors* **2022**, *22*, 923. [CrossRef] [PubMed]

68. Su, M.C.; Chen, J.H.; Arifai, A.M.; Tsai, S.Y.; Wei, H.H. Smart Living: An Interactive Control System for Household Appliances. *IEEE Access* **2021**, *9*, 14897–14904. [CrossRef]

69. Lee, J.W.; Yu, K.H. Wearable Drone Controller: Machine Learning-Based Hand Gesture Recognition and Vibrotactile Feedback. *Sensors* **2023**, *23*, 2666. [CrossRef]

70. Konstantoudakis, K.; Christaki, K.; Tsiakmakis, D.; Sainidis, D.; Albanis, G.; Dimou, A.; Daras, P. Drone Control in AR: An Intuitive System for Single-Handed Gesture Control, Drone Tracking, and Contextualized Camera Feed Visualization in Augmented Reality. *Drones* **2022**, *6*, 43. [CrossRef]

71. Yu, J.; Peng, X.L.; Xue, M.Z.; Wang, C.; Qi, H. An Underwater Human–Robot Interaction Using Hand Gestures for Fuzzy Control. *Int. J. Fuzzy Syst.* **2021**, *23*, 1879–1889.

72. Korayem, M.H.; Vosoughi, R.; Vahidifar, V. Design, manufacture, and control of a laparoscopic robot via Leap Motion sensors. *Measurement* **2022**, *205*, 112186. [CrossRef]

73. Stroh, A.; Desai, J. Hand Gesture-based Artificial Neural Network Trained Hybrid Human–machine Interface System to Navigate a Powered Wheelchair. *J. Bionic Eng.* **2021**, *18*, 1045–1058. [CrossRef]

74. Nourelhoda, M.; Hassan, F.; Ahmed, M. Smart healthcare solutions using the internet of medical things for hand gesture recognition system. *Complex Intell. Syst.* **2020**, *7*, 1253–1264.

75. Li, G.J.; Zhang, C.P. Research on static image recognition of sports based on machine learning. *J. Intell. Fuzzy Syst.* **2019**, *37*, 6205–6215.

76. Ji, R. Research on Basketball Shooting Action Based on Image Feature Extraction and Machine Learning. *IEEE Access* **2020**, *8*, 138743–138751. [CrossRef]

77. Xu, S.P.; Liang, L.X.; Ji, C.B. Gesture recognition for human–machine interaction in table tennis video based on deep semantic understanding. *Signal Process. Image Commun.* **2020**, *81*, 115688. [CrossRef]

78. Pan, T.Y.; Tsai, W.L.; Chang, C.Y.; Yeh, C.W.; Hu, M.C. A Hierarchical Hand Gesture Recognition Framework for Sports Referee Training-Based EMG and Accelerometer Sensors. *IEEE Trans. Cybern.* **2020**, *52*, 3172–3184. [CrossRef]

79. Paulo, T.; Fernando, R.; Luis, P.R. Generic System for Human-Computer Gesture Interaction: Applications on Sign Language Recognition and Robotic Soccer Refereeing. *J. Intell. Robot. Syst.* **2015**, *80*, 573–594.

80. Wang, X.; Zhu, Z.H. Vision–based framework for automatic interpretation of construction workers' hand gestures. *Autom. Constr.* **2021**, *130*, 103872. [CrossRef]

81. Riedel, A.; Brehm, N.; Pfeifroth, T. Hand Gesture Recognition of Methods-Time Measurement-1 Motions in Manual Assembly Tasks Using Graph Convolutional Networks. *Appl. Artif. Intell.* **2022**, *36*, 2014191. [CrossRef]