Journal of Network Intelligence ©2024 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **9**, Number **4**, November **2024**

Research On the Evaluation System of Physical Rehabilitation Training Using Computer Vision

Lingfeng Huang

School of Electronic Information
Dongguan Polytechnic, Dongguan, 523808, P. R. China
rafehwang@gmail.com

Qian Du*

School of Information Science and Engineering
Linyi University, Linyi, 27600, P. R. China
duqian@lyu.edu.cn

Yeh-Cheng Chen

Department of Computer Science
University of California, Davis, 94555, USA
ycch@ucdavis.edu

Ruey-Shun Chen*

Department of Information Management
National Yang Ming Chiao Tung University, Hsinchu, 30050, Taiwan
rschen@nycu.edu.tw

Lunan Liu

School of Electronic Information
Dongguan Polytechnic, Dongguan, 523808, P. R. China
lln200412@outlook.com

Shengwen Huang

School of Electronic Information
Dongguan Polytechnic, Dongguan, 523808, P. R. China
vinegar_vip@163.com

*Corresponding author: Qian Du, Ruey-Shun Chen
Received December 21, 2023, revised March 21, 2024, accepted June 25, 2024.

ABSTRACT. *In recent years, relevant policies of rehabilitation medical care have been released successively, and the development of the medical industry and the construction of rehabilitation medical service system have been vigorously promoted from the macro policy level. However, there is still a large shortage of rehabilitation therapists at this stage, and there is obviously a shortage of supply in the rehabilitation medical industry. This research focuses on the current social problem of insufficient professional guidance in the rehabilitation medical industry, mainly for the lack of professional guidance in home rehabilitation training, using computer vision recognition technology, with the help of home computer equipment, to develop a set of suitable for community rehabilitation services and home rehabilitation training, operation easy recovery system. At the same time, the system design achieves the three goals of "low cost, high efficiency, and easy operation". While solving the social problems of difficult home rehabilitation training, it also reduces the treatment cost and space constraints of rehabilitation patients, so that home rehabilitation patients can receive professional and accurate rehabilitation training guidance at any time.*

Keywords: healthcare; computer vision; physical rehabilitation; open pose; home rehabilitation.

---

1. **Introduction.** Due to the increase in the elderly population, the number of patients with chronic diseases is gradually increasing. According to incomplete statistics, there are 260 million chronic disease patients, and the elderly are a high-risk group of chronic diseases. The incidence rate of the elderly over 65 years old is 64.5%, of which the urban incidence rate is 85.2%, and the rural incidence rate is 52.4%. Chronic diseases are protracted and have a high rate of death and disability, seriously affecting the health of the elderly. In addition to chronic diseases, many older adults face serious challenges with impaired physical functions such as muscle strength, balance, and mobility due to organ damage or decline in physical function [1, 2]. All these negative changes make it difficult for the elderly to maintain independence in daily life, which further leads to their anxiety, low self-esteem and reduced quality of life.

Traditional treatment often involves physical therapy and rehabilitation programs for older adults, which are critical to the recovery of older adults. Relevant studies have shown that participating in physical therapy and rehabilitation programs is an effective way for the elderly to recover. Appropriate physical therapy and physical rehabilitation exercises can improve the physical activity level and daily life exercise ability of the elderly [3, 4], thereby improving their quality of life. However, the rehabilitation training of the elderly usually requires multiple rehabilitation trainings in one day. If each rehabilitation training is supervised by a doctor or relevant medical staff for the elderly, it is obviously not feasible and economical. Even if family members are allowed time to supervise the rehabilitation of the elderly, they cannot make professional judgments on the rehabilitation content and movement norms [5, 6]. At the same time, since the outbreak of the novel coronavirus in 2020, my country's overall medical budget and medical staff gaps are very large, and the healthcare system is facing tremendous pressure. In the current epidemic stage, for non-emergency medical projects such as movement rehabilitation training for the elderly, the general rehabilitation plan is to allow patients to carry out under the direct supervision of clinicians. After a period of time, patients are asked to return to the corresponding community hospital or their own residence. Get some prescribed exercise. During this phase, patients are tasked with recording their daily rehabilitation exercises and regularly visiting the hospital for functional assessments [7].

However, in practice, many patients have low adherence to exercise regimens during home rehabilitation, hindering the healing process. Studies have shown that in home rehabilitation exercises, the non-adherence rate can reach 76% [5]. Failure to adhere to rehabilitation results in significant healthcare burdens and risks, such as personal health costs, unnecessary medical visits, prescription drug use, increased risk of complications, and decreased quality of life [6]. In some cases, incorrect rehabilitation exercises can even aggravate the injury in the recovering elderly. In addition, some foreign studies have shown that the effect of unsupervised home exercise is not good for patients [8]. Because patients receive little immediate feedback as they continue exercising at home. Even in hospitals, doctors cannot continuously track each patient during their workouts. Because doctors often alternate between several patients, or have other tasks, this situation can lead to inadequate, inaccurate and often subjective feedback. Therefore, how to reduce the difficulty of training auxiliary equipment, enable patients to effectively persist in training, and provide patients with real-time and accurate estimation feedback has become a key issue that needs to be solved urgently in home rehabilitation training for elderly patients [9, 10].

In this study, a rehabilitation training estimation system based on ordinary cameras is proposed to solve the problems of lagging feedback from doctors and complex use of rehabilitation equipment in home rehabilitation training for elderly patients with chronic diseases. Seniors can imitate their movements by following the standard movements in the software. The attitude estimation algorithm will calculate the key points of the elderly in real time, and display the training score in the software. Elderly patients can see their own training movements in real time through the software, which is convenient for immediate correction. After the elderly patient completes a series of actions, the system will score each action of the elderly and give action guidance suggestions. The exercise data of the elderly will be automatically stored in the computer, and doctors can check the exercise of the elderly at any time and propose targeted improvement measures. Compared with rehabilitation training robots and rehabilitation wearable devices on the market, the rehabilitation training evaluation system based on computer vision is more suitable for home rehabilitation training for the elderly with chronic diseases or limb injuries because of its "no need to wear and simple operation" [11,12].

1.1. **Related Work.** At present, various tools and devices for home physical rehabilitation have been developed both domestically and internationally, such as robot assisted systems, wearable devices, and Kinect based games. The E2 Robot developed by the University of Valladolid in Spain. The principle is based on the X and Y axes of Cartesians and adopts a cross axis structure. Robots have two rehabilitation modes: passive guidance and active assistance, both of which are based on graphical environments and utilize different game therapies. Similar commercialized rehabilitation robots are also emerging in China.

There are also centrally driven rehabilitation robots with seats developed domestically. Power is transmitted to the shoulder and elbow joints through synchronous belts and bevel gear transmission mechanisms. The robotic arm can perform active and passive rehabilitation training, and can also be trained according to the set trajectory. Robotics technology provides some interesting advantages, such as the ability to automate and personalize treatments to reduce fatigue caused by repetitive and monotonous exercises, or the ability to integrate sensors for quantitative estimation of recovery. Robot assisted systems typically have force feedback systems that can effectively assist patient movement, but due to their mechanical limitations, the degree of freedom of movement is often limited and can only mimic limited human movements. Moreover, the price of robot assisted systems, which often cost tens of thousands of yuan, is too expensive and currently only suitable for use in large hospitals. It is not very suitable for individuals to use as home rehabilitation training systems.

1.2. **Motivation and contribution.** In order to solve and estimate based on vision, the main research contents of this research are as follows: The purpose of this subject research is to use computer vision equipment and deep learning algorithms to construct a family rehabilitation training evaluation system, and apply it to the evaluation of family rehabilitation training for chronic diseases, impaired physical functions and other diseases of the elderly. The research uses algorithms such as pose estimation, data filtering, feature extraction, and machine learning to achieve objective and accurate estimation of motor function and predict scoring. Compared with the traditional home rehabilitation training system, the computer vision-based method greatly improves the ease of use of training equipment, reduces the cost of home rehabilitation training, and provides accurate and real-time evaluation feedback of movement rehabilitation training for elderly patients.

The main innovations and contributions of this work include:

(1) Using Mobile Net V3, shared convolution and residual hole convolution structure to improve and optimize the pose estimation network. Open Pose, realize real-time pose estimation and action scoring at about 30 frames per second on the home sensors-based computer platform.

(2) A special normalization algorithm is proposed, combined with sensors-based the angle information of the elbow joint and knee joint, to solve the error problem of motion information caused by differences in human body size or distance from the camera.

(3) The dynamic time warping algorithm is used to calculate the movement similarity between the old man and the standard action, which solves the problem of inaccurate action evaluation caused by the old man's movement hysteresis.

2. Literature review.

2.1. **Sensors.** Sensor technology [1] is the frontier technology of modern science and the important technical foundation of the new technology revolution and information society. In modern life and scientific research, a large number of reliable and accurate information provided by various types of sensors can not only replace the functions of human's five senses, but also detect the information that human's five senses cannot feel, so that human beings can better understand and transform the world. At present, sensor technology is widely used in the fields of cutting-edge technology such as aviation and aerospace, and in many aspects of human daily life such as industry and agriculture. The application penetration rate of sensors in the industrial sector has been regarded by the international community as an important indicator of a country's intelligence, digitalization and networking. Therefore, as a new subject closely related to modern science, sensor technology is developing rapidly and is being used more and more widely in many fields [13, 14].

Sensor refers to the general name of components or devices that have the function of sensing (or responding) and detecting certain information of the measured object, and convert it into corresponding output signals according to certain rules. The sensor is generally considered to be composed of three parts: sensitive element, conversion element and measuring circuit, and sometimes auxiliary power supply is required. The sensor can contact the measured object directly or not. There are usually many technical requirements for sensors, some of which are applicable to all types of sensors, and some of which are only applicable to specific types of sensors. The basic requirements for the working principle and structure of sensors in different occasions are: high sensitivity, anti-interference stability, linearity, easy adjustment, high precision, high reliability, no hysteresis, long working life, repeatability, anti-aging, high response rate, resistance to environmental impact, interchangeability, low cost, wide measurement range, small size, light weight, high strength, wide working range [15-18].

2.2. **Artificial Intelligence Internet of Things.** Artificial Intelligence Internet of Things integrates AI technology and IoT technology, generates and collects massive data from different dimensions and stores them in the cloud and edge through the Internet of Things [19-23], and then realizes the digitalization and intelligent connection of everything through big data analysis and higher forms of artificial intelligence. The ultimate goal of the integration of Internet of Things technology and artificial intelligence is to form an intelligent ecosystem. In this system, the integration of different intelligent terminal devices, different system platforms, and different application scenarios is achieved. In addition to the need for continuous innovation in technology, the research and development of AIoT-related technical standards and test standards, the implementation of relevant technologies and the promotion and scale application of typical cases are also important issues that need to be broken through in the field of Internet of Things and AI at this stage. With the maturity of the Internet of Things and artificial intelligence technology, more and more enterprises have listed AIoT as their main development direction. Since 2017, the word "AIoT" has appeared frequently and become a hot word in the Internet of Things industry. "AIoT", namely "AI+IoT", refers to the integration of AI technology and the Internet of Things in practical applications. More and more industrial applications combine AI with IoT. For example, manufacturers such as Xiaomi, Skyworth and Hisense have launched their own AIoT TV in succession. The purpose is to use TV as the main control center and realize the control of intelligent equipment such as air conditioners, refrigerators and washing machines through full-scene intelligence [35, 36]. AIoT has become the best channel for intelligent upgrading of major traditional industries, and will also become the inevitable trend of the development of the Internet of Things. The huge and complex data generated by the Internet of Things needs to be analyzed and processed, and AI technology is the best choice for effective information processing [37-41], which can make intelligent products better understand user intentions. Only IoT can provide AI data continuously [24-28]. The massive data provided by IoT can enable AI to acquire knowledge quickly. The integration with AI technology can bring broader market prospects for the Internet of Things, thus changing the existing industrial ecology and economic pattern, and even allowing us to enter the life scene like science fiction movies in advance [24-27].

2.3. **Smart Education.** Education has evolved into a lifelong process that extends far beyond classrooms and whiteboards. Learners can easily connect to digital libraries, take online classes, and submit homework and assignments electronically from virtually any kind of mobile device. That's why many education institutes provide mobile services that allow students to attend class whenever and wherever they like. Smart education is not only a complex construction process of comprehensively and deeply applying modern information technology to promote education reform and development in the field of education, such as education management, education and teaching, education and scientific research, but also a basic feature of digitalization, grid [29], intelligence and multimedia. At the same time, smart education will continue to introduce the Internet of Things technology, cloud computing, wireless communication and other new generation of information technology to create an interconnected, intelligent, perceptive, modern and personalized new education form and mode. Smart education is not only limited to the information of education, but also needs the perspective of technological innovation to apply emerging technologies to school management, education and teaching, and other aspects of real education equity, improve education quality and teaching effect, and create a new ecological model of smart education [30-33].

2.4. **Physical rehabilitation.** At present, a variety of tools and equipment for home physical rehabilitation have been developed at home and abroad, such as robot assistance systems, wearable devices, and games based on Kinect [12]. The E2 Robot developed by the University of Valladolid in Spain. Its principle is based on Cartesian X and Y axes, adopting cross axis structure. The robot has two rehabilitation modes, "passive guidance" and "active assistance", both of which are based on a graphical environment utilizing different play therapies. Domestic similar commercialized rehabilitation robots are also emerging in an endless stream [34].

The power is transmitted to the shoulder and elbow joints through the synchronous belt and bevel gear transmission mechanism. The robotic arm can perform active and passive rehabilitation training, and can also perform training according to the set trajectory [9]. Robotics offers some interesting advantages, such as the possibility of automating and personalizing therapy, which reduces fatigue from repetitive and monotonous exercises, or the integration of sensors for quantitative estimates of recovery. Robot-assisted systems usually have a force feedback system that can well assist patients in their movements. However, due to their mechanical structure limitations, the degree of freedom of movement is often limited to a certain extent, and can only imitate limited human movements. Moreover, the price of robot-assisted system, which costs tens of thousands of yuan, is too expensive. Currently, it is only suitable for use in large hospitals, not suitable for individuals to use as a home rehabilitation training system [34-36]. Based on the Kinect application, the developed intelligent motion analysis and training system is shown in Figure 1.

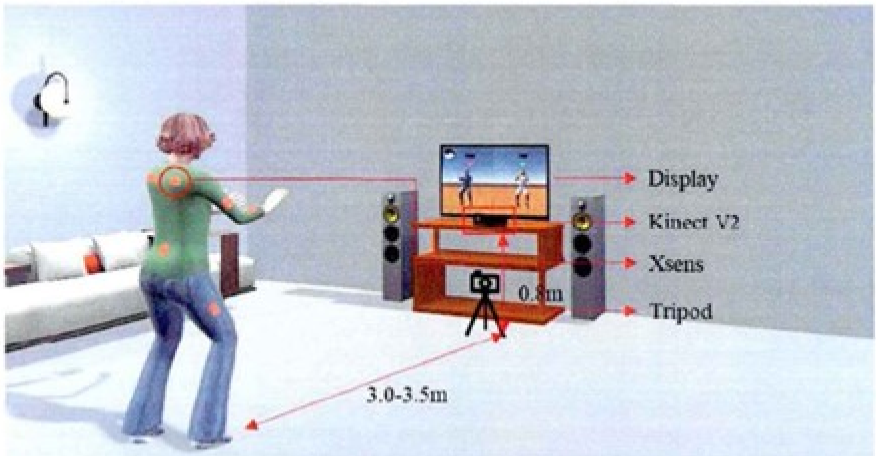

Figure 1. Kinect-based rehabilitation system

The system collects the patient's motion information through the Kinect camera, and displays the patient's motion on the screen as a mirror image, which is convenient for the patient to find errors in the motion and correct the motion in real time. Using Kinect for rehabilitation treatment, patients do not need to wear any sensors, and their motion information is obtained through the Kinect device, which greatly reduces the complexity of rehabilitation training for the elderly. Kinect provides a ready-made motion information capture algorithm, which facilitates the development of rehabilitation training estimation algorithms. However, the Kinect device itself is expensive, and most families need to purchase additional installations, and Microsoft has announced the discontinuation of Kinect earlier, so the rehabilitation application based on this device will be difficult to continue.

3. Research method.

This research aims to use computer vision equipment and attitude estimation algorithms to construct a family rehabilitation training estimation system, and apply it to the research on family rehabilitation training estimation of the elderly with chronic diseases, limb dysfunction and other diseases. Through the research on algorithms such as pose estimation, data filtering, feature extraction, and machine learning, the objective and accurate evaluation of motor function can be realized, and the score can be predicted. Compared with traditional home rehabilitation training systems, the computer vision-based method greatly improves the availability of training equipment, reduces the cost of home rehabilitation training, and provides patients with accurate and instant feedback on action estimation.

3.1. Improvement of attitude estimation algorithm.

Attitude estimation, referred to as HPE (Human, Pose, Estimation), its purpose is to obtain the posture of the human body from a given sensor input. Traditional attitude estimation is generally divided into two categories, one is to directly solve the attitude estimation problem into a classification or regression problem through global features [10]. The other category is graph-based models of structures and deformable components. This system uses a bottom-up multi-stage plane pose estimation algorithm Open Pose. The specific convolutional layer structure of the Open Pose network as Figure 2, after the image is passed through VGG-19, a set of feature maps F are generated and input to the initial stage. The initial stage consists of two branches, Branch 1 partly predicts the confidence map representing the location of body key points, and Branch 2 (Branch2) partly predicts the part affinity domain representing the degree of association between body key points. The enhancement stage also consists of two branches, but in order to maintain a large receptive field, three 7×7 planar convolutions are used. The reinforcement phase can be repeated multiple times to improve model accuracy.

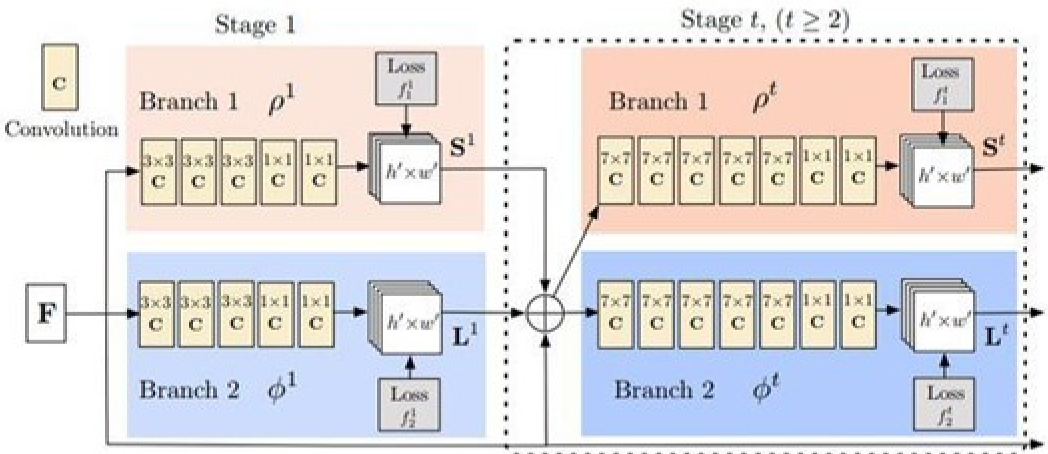

Figure 2. Open Pose network convolutional layer structure

The Open Pose convolutional neural network can predict the X and Y coordinate information of 25 key points such as human eyes, ears, nose, shoulders, elbows, hip joints, knee joints, and tread joints. The specific posture model is shown in FIGURE 3.

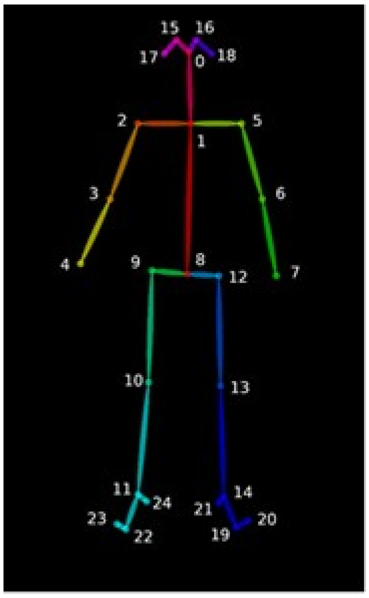

Figure 3. 25-point attitude model

However, after the image passes through the VGG-19 convolutional network, a set of features F is generated and input to the initialization stage. From the initialization phase, the two branches of the network will generate a set of confidence maps $S1$ and a set of partial affinity domains $L1$, respectively. After this strengthening phase, each repetition will produce a set of S_T and a set of L_T. Formula (1) and (2) are as follows:

$$L^t = \phi^t(F, L^{t-1}), \forall 2 \leq t \leq T_P \tag{1}$$

$$S^t = \rho^t(F, L^t, S^{t-1}), \forall T_P \leq t \leq T_P + T_C \tag{2}$$

where ϕ and ρ denote consecutive convolution operations for predicting partial affinity domains and confidence maps, respectively. T_P represents the maximum number of repetitions of the branch of the partial affinity domain, and T_C represents the maximum number of repetitions of the branch of the confidence map. To improve the accuracy of the model, the system adds intermediate supervision in each successive t stage, $t\phi^t\rho T_P T_C \in \{1 \dots T\}$. The loss function uses L2 loss, as in Formula (3) and (4):

$$\int_L^t = \sum_{c=1}^{C} \sum_P W(P) \cdot \left\| L_C^t(P) - L_C^*(P) \right\|^2 \tag{3}$$

$$\int_S^t = \sum_{j=1}^{J} \sum_P W(P) \cdot \left\| S_j^t(P) - S_j^*(P) \right\|^2 \tag{4}$$

where t_i refers to the i-th stage of the predicted partial affinity domain, and t_k refers to the k-th stage of the predicted confidence map. $L_C^*(P)$ is the true value of the partial affinity domain at pixel P, and similarly, $S_j^*(P)$ is the true value of the confidence map at pixel P. $W_L^C(P) S_j^t(P)$ $W(P)$ is a binary mask at pixel P, if pixel P is not marked, then $W(P) = 0$, otherwise $W(P) = 1$. The total loss function is the addition of the loss of the partial affinity domain and the loss of the confidence map, with the same weight. Formula (5) is as follows:

$$\text{Loss} = \sum_{t=1}^{T_P} f_L^t + \sum_{t=T_P+1}^{T_P+T_C} f_S^t \tag{5}$$

The confidence map represents the probability that a joint point of the body is located at each pixel in the image. The Open Pose network can predict 25 joint points of the human body, so the output of the model should have 18 confidence maps. The probability values in each confidence map are distributed in the form of a two-dimensional Gaussian function. In theory, the closer the pixel is to the real joint point, the greater the corresponding probability value in the confidence map, and vice versa. Typically, if there is only one person in the image, and every joint point of the person is visible in the image, there is only one distribution of Gaussian functions in each confidence map. If there are many people in the image, and there are K identical joint points visible, then there are K peak values of Gaussian functions in the confidence map corresponding to the joint point. The calculation Formula (6) for the value of each element in the real confidence map is as follows:

$$S_{j,k}^*(P) = \exp\left(-\frac{\|P - x_{j,k}\|^2}{2\sigma^2}\right) \tag{6}$$

where k means that there are k people in the picture, and j means the visible joints of a certain person. Represents the specific pixel position of the visible joint points of the k-th person, because the input of the image is two-dimensional, so the same for pixels. It can be seen that the confidence map is a two-dimensional Gaussian function, which controls the width of the peak. The accuracy of the plane human pose model established by the open pose convolutional neural network is about 60%.

However, due to its huge network and data volume, the reasoning speed on the home computer platform is slow, and the video frame rate is only about 13 FPS, which cannot Normal use in home rehabilitation. For dynamic interactive rehabilitation training, the video stream needs to be at least 24 FPS or higher, so that the human eye will not feel obvious lag. Moreover, if the number of frames of the video stream is higher, more key information will be extracted, which is conducive to real-time feedback and action scoring for rehabilitation training. To this end, the research group improved and optimized Open Pose: changed the Back Bone of Open Pose to Mobile Net V3, a mobile network with faster inference speed and higher accuracy; the initial stage and the strengthening stage refer to the lightweight model proposed by Intel Corporation The optimization method used in Open Pose [11]. In the initial stage, all layers except the last two layers are shared, which greatly reduces the number of parameters of the model. The specific method is shown in Figure 4; in the strengthening stage, each convolution block with the same receptive field is replaced by a convolution with a kernel size of 7×7 to capture long-range spatial correlations, three consecutive 1×1, 3×3 and 3×3 convolutions, with a dilation parameter equal to 2 in the last layer to preserve the initial perception.

It should be noted that after replacing the Back Bone part of Open Pose with Mobile Net V3, in order to achieve the matching of feature map size and channel, it is necessary to adjust the convolution step of some layers in Mobile Net V3, and delete the average value and The following fully connected layer uses 1×1 convolution to change the number of channels of its output. The network structure of Mobile Net V3 is shown in Figure 5. Mobile Net V3 inherits the depth wise separable convolution and the inverted residual structure with a linear bottleneck, redesigns the time-consuming layer, introduces the

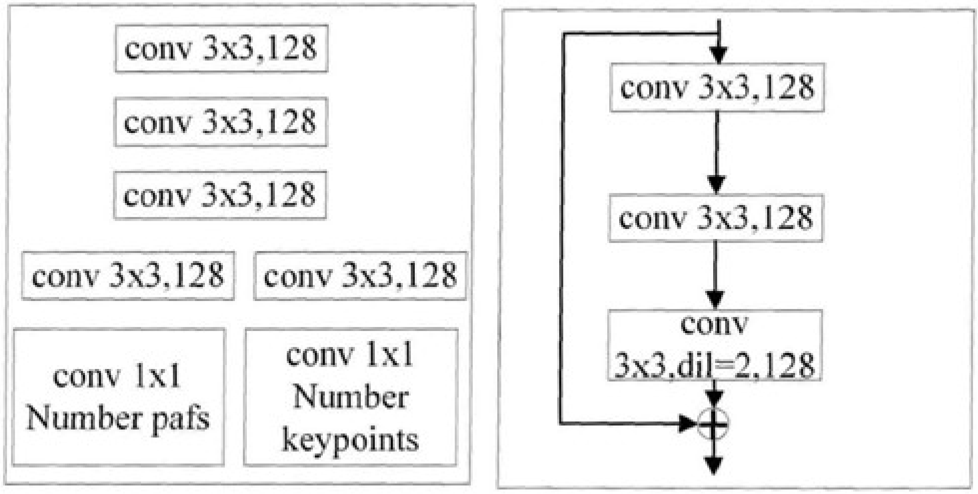

Figure 4. Open Pose improvement structure diagram

lightweight attention model "SENet" and adds channel fitting capabilities. In addition, it also uses the h-swish (see Formula (1)) activation function to improve the speed of the model. Thanks to the above techniques, Mobile Net V3 achieved an accuracy of 75.2% on the image classification dataset ImageNet.

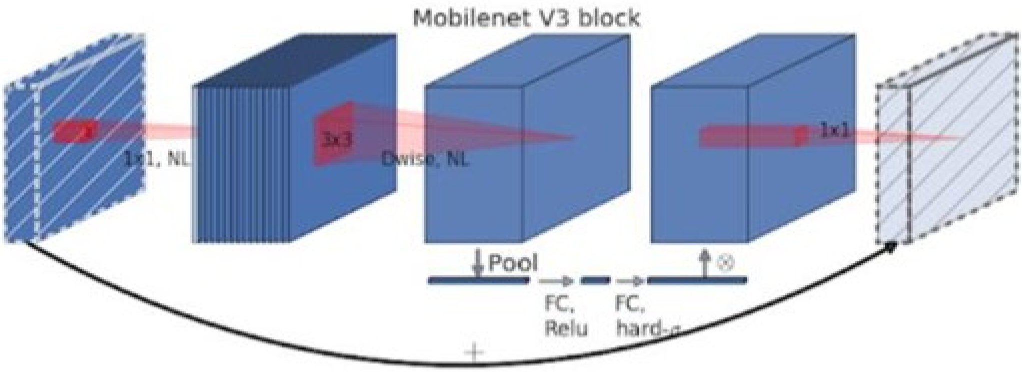

Figure 5. Mobile Net V3 network structure diagram

3.2. **Research on Scoring Algorithm for Action Rehabilitation Training.** The action evaluation algorithm mainly includes three main parts: human pose estimation, data filtering, and feature extraction. The specific algorithm flow is shown in Figure 6. After the information of the video stream is filtered by the pose estimation algorithm and data, the key point information of the human skeleton is obtained. After preliminary feature extraction, some useless key point information is deleted, and the angle information of the elbow joint and knee joint is calculated as a shallow feature. Further, deep feature extraction is performed to extract the dynamic time-warped similarity between the motion data sequence and the standard action sequence, as well as the AMP, MEAN, RMS, and JERK features of the key points to describe the amplitude, direction, dynamic energy, and smoothness of the motion and randomness. Finally, according to all the extracted features, the mapping model with clinical rehabilitation estimation is established through the support vector SVM algorithm.

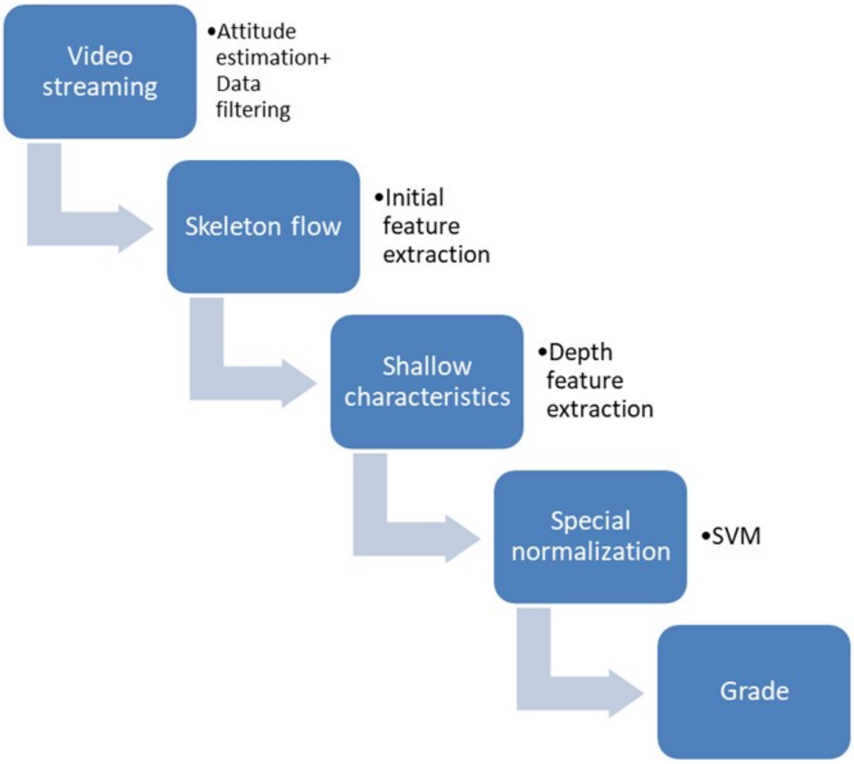

Figure 6. Action evaluation algorithm flowchart

3.2.1. *Data filtering.* The improved attitude estimation model has greatly improved the recognition speed, but its average accuracy has dropped from 60% to about 39%, so there is a certain noise signal in the detected key point information. In addition, when the user bends down, some joint points of the body may be blocked. At this time, the position of the joint points in Open Pose is represented by (-1, -1), which will cause jitter or deformity mutations in the bone joint point data. In order to reduce the impact of noise signal and joint point occlusion on the final estimation, the system performs limit filtering and median filtering on the bone joint point data.

The principle of clipping filtering is to compare the sampling values of two consecutive adjacent moments, and if the deviation of the two sampling values exceeds the threshold, the latter sampling value is considered to be abnormal. The threshold is usually an empirical value, which can be adjusted according to the actual experimental effect. There are many ways to deal with outliers, and there are two commonly used ones: the first is that the system considers the next sampling to be an outlier, delete it directly, and replace it with the sampling value at the previous moment. This method is simple and efficient; the second is the average value of the two sampling values is used to replace the second sampling value, and this method retains the sampling information as much as possible. In this system, the loss of joint points will directly cause the detected position to become (-1, -1), which is usually far from the sampled value at the previous moment. If the second method is used, the data will still be large fluctuation anomalies. Therefore, this system adopts the first method, which is directly replaced by the data at the previous moment, until the normal data is collected next time. The clipping filter method can simply and efficiently remove the influence caused by the sudden change in the position of a joint point when a certain joint point is suddenly occluded.

Median filtering has a good filtering effect on noise. Moreover, it can also protect the edge information of the signal and prevent the signal from being blurred when filtering out the noise. This is an excellent characteristic that the linear filtering method does not have, and this characteristic is also very suitable for the data processing of this system. In addition, the implementation of the median filter algorithm is simple, and when the window size is selected properly, the calculation speed is fast, which is conducive to the real-time data processing of this system. It is verified by experiments that the threshold of the clipping filter is set to 30, and the window size of the median filter is set to 5, and the effect is the best.

Figure 7. A picture of the experimenter's actions

Figure 7 takes the experimenter's action 1 as an example, Figure 8 showing the comparison before and after filtering of the real-time transformation curve with the key point being the X and Y coordinates of the left wrist. It can be seen that the filtering algorithm only deletes all abnormal points in the data, and does less smoothing on the data. The data does not lag or distort due to filtering, and the authenticity of the data is preserved to the greatest extent.

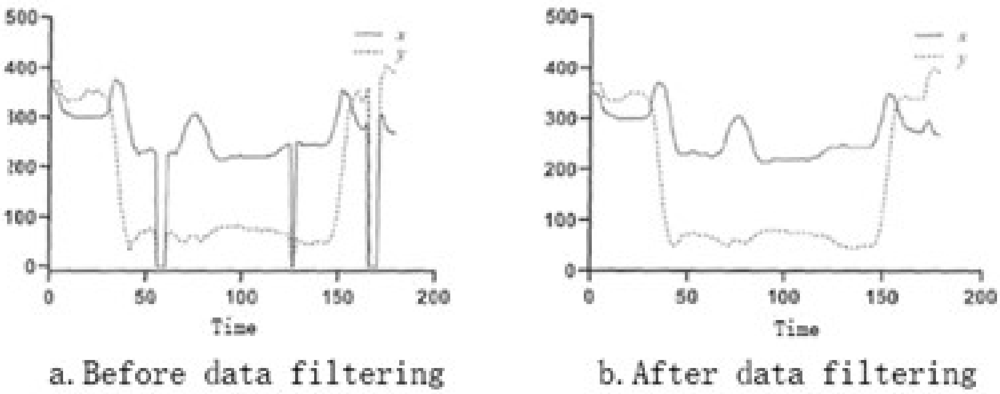

Figure 8. Comparison of data filtering algorithm before and after

3.2.2. *Preliminary feature extraction.* After data filtering, if the coordinate sequence of 25 key points of the human skeleton captured by the pose estimation model is directly used for motion similarity estimation, the estimation method is unscientific, and the final estimation result is also unreliable. Because an important problem in action scoring is that there are great differences in the appearance of the human body, such as height, fat and thin. Different people perform the same action. The point locations vary greatly.

Figure 9. The second picture of the experimenter's actions

For example, as Figure 9, when the experimenter raises his hand in action 2, when people of different heights raise their hand to the top of their head, the coordinates of the key points of the hand must be different. And even if the same person is at a different distance from the camera, the collected location information will be different. In order to solve this problem, the key points of the shoulder, elbow, wrist, waist, knee, and ankle joints are calculated by the Formula (7) to calculate the real-time transformation curves of the angles of the left and right elbow joints and left and right knee joints respectively.

$$\text{angle} = \frac{180}{\pi} \arccos\left(\frac{A^2 + B^2 - C^2}{2AB}\right) \tag{7}$$

For the upper limbs, A represents the distance from the shoulder to the elbow, B represents the distance from the wrist to the elbow, C represents the distance from the shoulder to the wrist, and the result angle represents the angle of the elbow joint. The distance can be calculated using the Formula (7) for the distance between two points. In the same way, the knee joint angle can be calculated by replacing the lower limbs with the corresponding key points.

In addition, the research team found in actual tests that the system often cannot accurately detect the position information of human eyes and ears due to different background light or people's hairstyles and decorations, such as strong light, wearing masks or hats. Therefore, the system deletes the coordinate information of the left eye, right eye, left ear, and right ear, and only retains the information of the nose to represent the position information of the head. The movement trends of these key points are almost identical, and the deletion will not affect the system's description of human movement information, and can also reduce the number of parameters and improve the operating speed of the system.

3.2.3. *Special Normalization Algorithm.* The change curves of the human elbow and knee joint angles will not be different due to the appearance of the human body and the distance from the camera, so it is very suitable as an effective feature for motion scoring. However, only the angle information is not enough. For example, if the experimenter raises his hand, if his arm is straightened but not placed on the top of the head, but only placed on the side of the body, the angle of the elbow joint is also 180°, but the movement is not standard. At this time, the position of the key points of the hand is still used for judgment, which shows that the position information of the key points is not completely useless. In order to fully show the trajectory of the human body without being affected by the size of the human body, the article proposes a special normalization algorithm.

$$x_{1,t}^{\text{new}} = \frac{x_{1,t} - x_{2,t}}{\|P_1 - P_2\|} \tag{8}$$

Take the upper limb as an example, where f represents the initial position of the wrist, g represents the initial position of the shoulder, which means the length of the human arm, A represents the X-axis position of the wrist at time t, and $X2t$ represents the X-axis position of the shoulder at time t. In the same way, the Y-axis coordinate information is also normalized. After normalization processing, it is equivalent to establishing a coordinate system with the shoulder as the center, the position of the wrist as the coordinate value, and the coordinate range is limited to the $(-1, +1)$ interval.

Similarly, for the key points of foot stepping, the hip joint can be used as $P2$ for normalization to represent the trajectory of the ankle. For the nose key, you can use the neck as $P2$ to represent the lift or rotation of the head. Note that the angle curves are not normalized, since there are no body size differences in the angle information. The normalized curve can well represent the movement trajectory of the wrist, ankle, and head of the joint points, and no longer contains specific position information, and the error caused by the difference in human body shape is removed.

3.2.4. *Dynamic Time Warping Similarity.* Due to the decline of motor function, the elderly usually cannot follow the standard video movement quickly, so there is always a problem of lag. But the lag of the movement does not mean that the old man's movement must not meet the standard. In this case, traditional methods such as Euclidean distance cannot be used to calculate the similarity between the old man's action and the standard action. Instead, the dynamic time warping algorithm (Dynamic Time Warping, referred to as DTW) should be used to calculate the similarity between the elderly's movement and the standard action. This algorithm can make up for the hysteresis between the elderly's standard actions. Because DTW can output the best alignment value between two curves, which is the minimum matching cost or cumulative distance. The closer the two-time series are, the better the matching performance and the smaller the DTW output value.

The specific calculation process of the dynamic time warping algorithm is as follows: Assume that the change sequence of the wrist coordinates in the standard action is $S = (S1, S2, S3...Sn)$, and the change sequence of the wrist coordinates in the elderly training action is $T = (t1, t2, t3...tm)$, where s and t both represent the 2D coordinate information of the wrist at a certain moment. In order to compare the similarity between sequences S and T, DTW needs to first construct a cost matrix $Cp(S, T)$ with n rows and m columns, whose definition is shown in Formula (9). Here it represents the distance between the sequence S at the i-th moment and the sequence T at the j-th moment. There may be multiple distortion paths P in the cost matrix.

$$Cp(S,T) = \sum_{k=1}^{s} C(S_{ik}, t_{jk}) \tag{9}$$

The goal of DTW is to find the distorted path with the smallest cumulative distance from all paths. The specific method is given by the recurrence equation Formula (10):

$$D(i,j) = \min\{D(i-1,j-1), D(i-1,j), D(i,j-1)\} + C(S_i, t_j) \tag{10}$$

where $1 < i < n, 1 < j < m$. $D(i,j)$ represents the matching cost between the standard action sequence S and the elderly motion data sequence T from time $(1,1)$ to time (i,j).

In addition to the action similarity, more features are needed to describe the user's action, such as the degree of completion, smoothness, randomness and other aspects of the action for feature extraction. The extracted features and their corresponding physical meanings are listed in Table 1.

Table 1. Extracted features and their physical meaning

No.	**Features**	**Definition**	**Physical meaning**
1	DTW	$D(S,T)$	Describe action similarity
2	AMP	$AMP = \max(x) - \min(x)$	Describe the amplitude of motion
3	MEAN	$MEAN = \frac{1}{N}\sum_{i=1}^{n} x_i$	Describe the direction of movement
4	RMS	$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$	Dynamic energy describing motion
5	JERK	$JERK = \sqrt{\frac{1}{N-1}\sum_{i=2}^{n}(x_i - x_{i-1})^2}$	Describe the smoothness of motion

3.2.5. *Key point angle calculation.* According to the description of rehabilitation training in the "Guidelines for Postoperative Rehabilitation in Orthopedics", patients must achieve the recommended range of physical activity within the specified rehabilitation cycle. Among them, the angle of limb movement is a very important indicator in rehabilitation training. Taking ankle joint and foot rehabilitation exercises after internal fixation of foot, ankle, and calf fractures as an example, according to the requirements of the rehabilitation data: "Patients' muscles in the affected limbs are completely relaxed, and they can control the progress of rehabilitation by themselves according to the degree of pain, the ankle joint reaches 120° dorsiflexion", the specific rehabilitation guidance requirements are shown in Table 2:

Table 2. Rehabilitation guidelines after ankle joint surgery

Stage	**Initiative**	**Passive**	**Effect**
Early stage (within 1 week)	Use the trunk, pelvis and the opposite limb to do the downward active movement, and perform the passive exercise of the foot and ankle joints of the affected limb	The whole affected limb is in a relaxed state, and the passive extension exercise of the ankle joint of the group can be achieved by using the role of the skeleton	For patients who delay the recovery time, take a standing position, pad their front feet high, use their weight, and perform passive back extension exercise of the ankle joint
Medium term (2-6 weeks)	The trunk, pelvis and lower limbs can be actively exercised	For patients with low pain area, foot pedal method is also required for passive exercise	The function of ankle joint returns to normal
Later stage (after 6 weeks)	ditto	For patients who delay the recovery time, take a standing position, pad their front feet high, use their weight, and perform passive back extension exercise of the ankle joint	The function of ankle joint can basically return to the functional position

Take the patient's ankle joint rehabilitation action in the Figure 10 as an example. To calculate the joint angles of key points, the law of cosines can be used. Included angles are represented by uppercase A, B, C, and side lengths are represented by lowercase a, b, c:

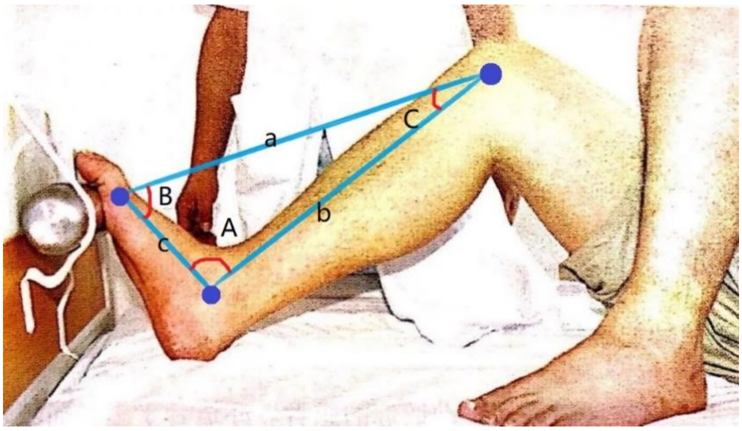

Figure 10. Ankle joint angle

The mathematical formula for the law of cosines is as follows: Therefore, it can be concluded that the calculation Formula (11)(12)(13) of each angle is:

$$a^2 = b^2 + c^2 - 2bc \cdot \cos A \tag{11}$$

$$b^2 = a^2 + c^2 - 2ac \cdot \cos B \tag{12}$$

$$c^2 = a^2 + b^2 - 2ab \cdot \cos C \tag{13}$$

```
Function Cos Angle (a, b, c: double)
    var a, b, c: double
    var p0, p1, p2: double
Begin
    a = (p1[0] - p0[0]) ** 2 + (p1[1] - p0[1]) ** 2
    b = (p1[0] - p2[0]) ** 2 + (p1[1] - p2[1]) ** 2
    c = (p2[0] - p0[0]) ** 2 + (p2[1] - p0[1]) ** 2
    If <a * b == 0>
      Then returns -1.0;
    Return (((a + b - c) / square(4 * a * b)) * 180 / 3.1415)
End
```

According to the calculation results of the program, the angle of the patient's joint can be obtained.

4. **Results and Analysis.** The Open Pose model has two computing versions, CPU and GPU. The CPU version was successfully compiled and run in the virtual machine through the downloaded Caffe model. However, in subsequent testing, it was found that the system had a high load on the CPU, so this option was discarded. The development version of the CPU computation is still reserved in the system options, providing a solution for some home computers with better performance.

Compilation of the GPU version requires the support of CUDA and cuDNN and depends on the computing power of the graphics card. During the compilation process, the default settings of Open Pose cannot automatically recognize the computing power of the corresponding machine, which may result in mismatched computing power after compilation and prevent the program from running. When setting up the system, the computing power needs to be explicitly set to match the current machine for stable performance.

4.1. **Bow Squat.** The angle between the torso and the ground: for a healthy person, this angle should be $90° \pm 10°$, with a result error within $\pm 2°$. Although the general trend is visible, it is difficult to select a single peak as a basis for judgment due to large fluctuations. It is more suitable to use the overall trend as a reference for assessment. The red color in Figure 11 indicates the trend of the curve, which closely resembles the condition of the subjects in the recorded video.

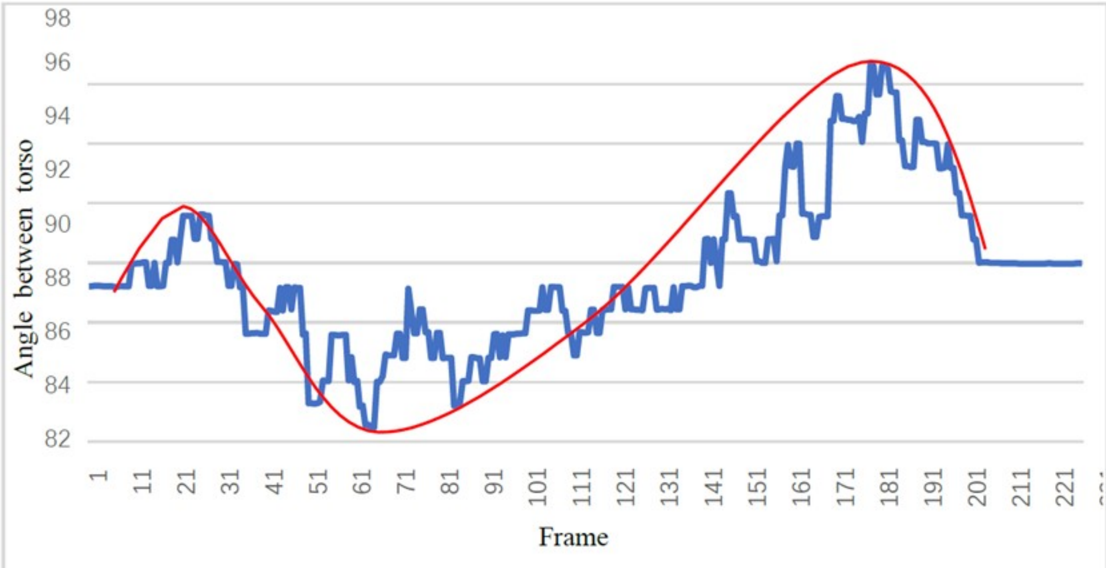

Figure 11. Bow squat - the angle between the torso and the ground Detection results

4.2. **Full Squat.** Knee angle: The lower the minimum knee angle value, the better. For a healthy person, the knee angle should be approximately $20° \sim 35°$. From Figure 12, it can be observed that the detection stability of the foot is higher than that of the torso. Side-view video footage is used for detection in this assessment. The knee angle value does not fluctuate excessively, and the error margin between the measured result and the actual value is within $\pm 2°$, allowing the results to be directly referenced with confidence.

The angle between the arm and the calf bone as Figure 13: take the minimum value, the smaller the angle, the worse the condition, the running result is consistent with the result, the healthy range should be greater than -20o, the error is ±3o.

4.3. **Knee Spacing.** Figure 14: Rehabilitation guidelines after ankle joint surgery use the value curve to deviate from the initial value and reach the maximum value. The maximum value and curve are consistent with the observations. When standing up and squatting, healthy individuals should keep their knees and shoulders at the same width, with the ratio of knee width to shoulder width between $0.7 \sim 1.3$. The curve should be as smooth as possible. In the figure below, fluctuations can be observed within the red circle, and the comparison video confirms that these fluctuations were due to vibrations caused by insufficient knee stability during the test.

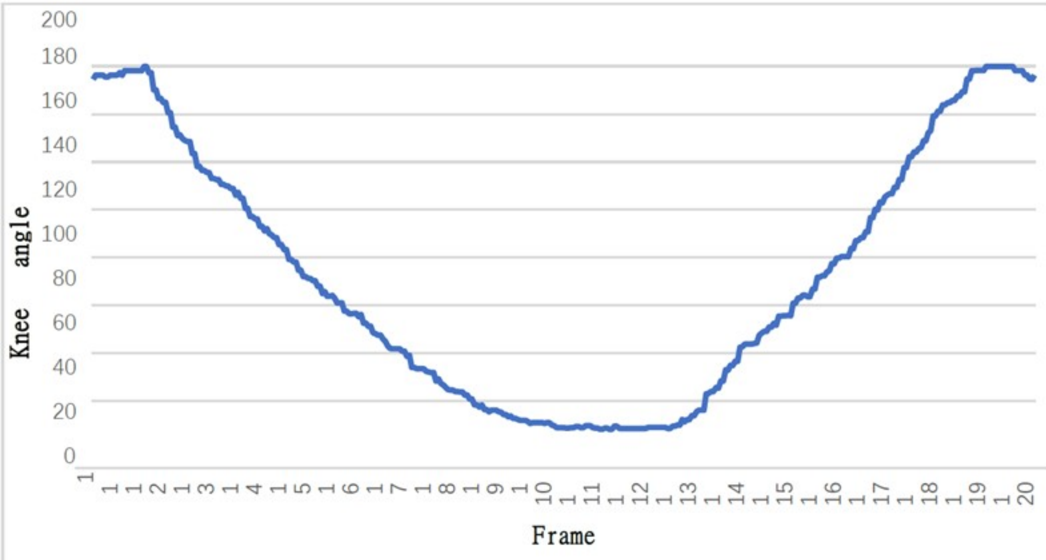

Figure 12. Full squat-knee angle detection results

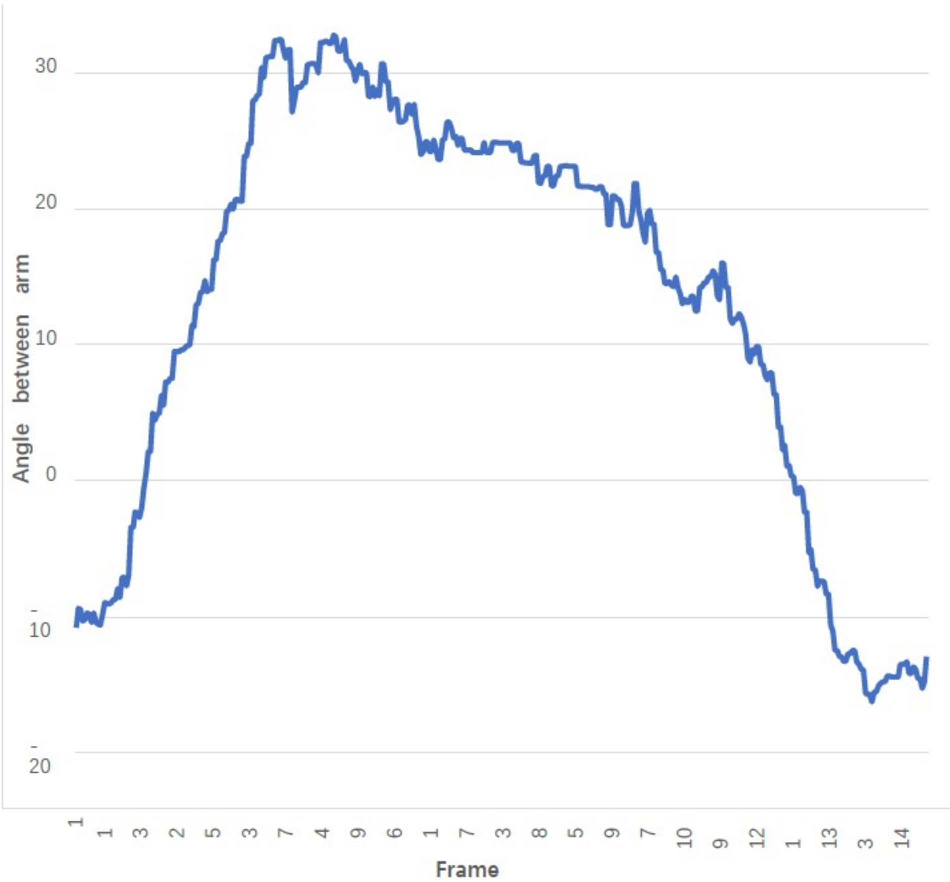

Figure 13. Full squat - the angle between the arm and the calf bone Detection results

4.4. **Touch the back with hands from bottom to top.** Back touch position as Figure 15. Rehabilitation guidelines after ankle joint surgery: Take the maximum value, the larger the value, the higher the back touch point and the better the range of motion of the joint. The result is consistent with the observed value.

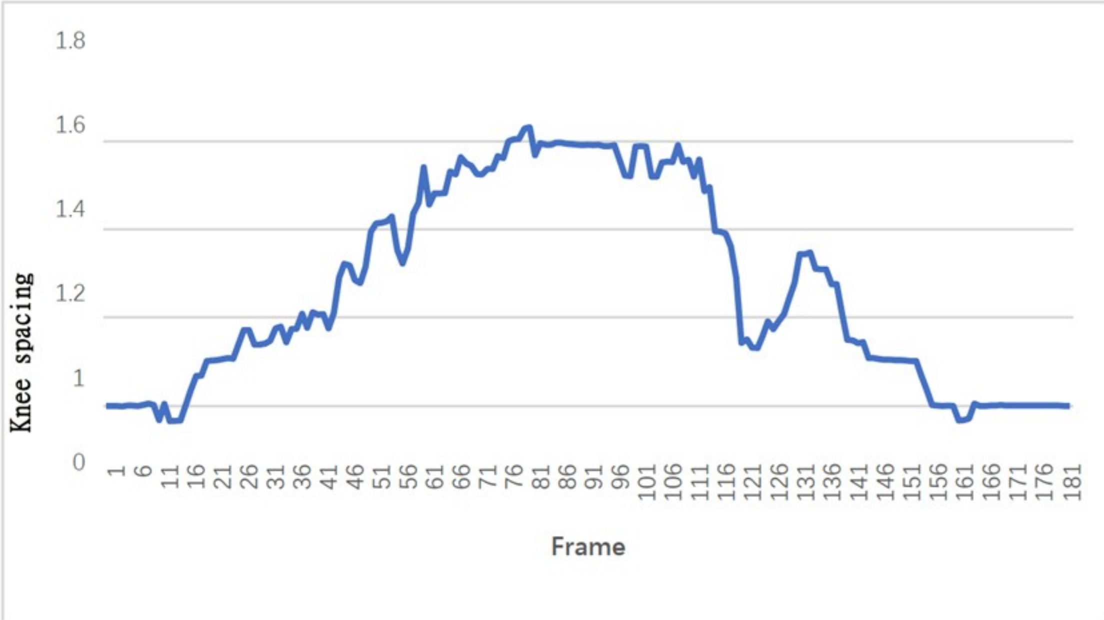

Figure 14. Full squat-knee distance detection results

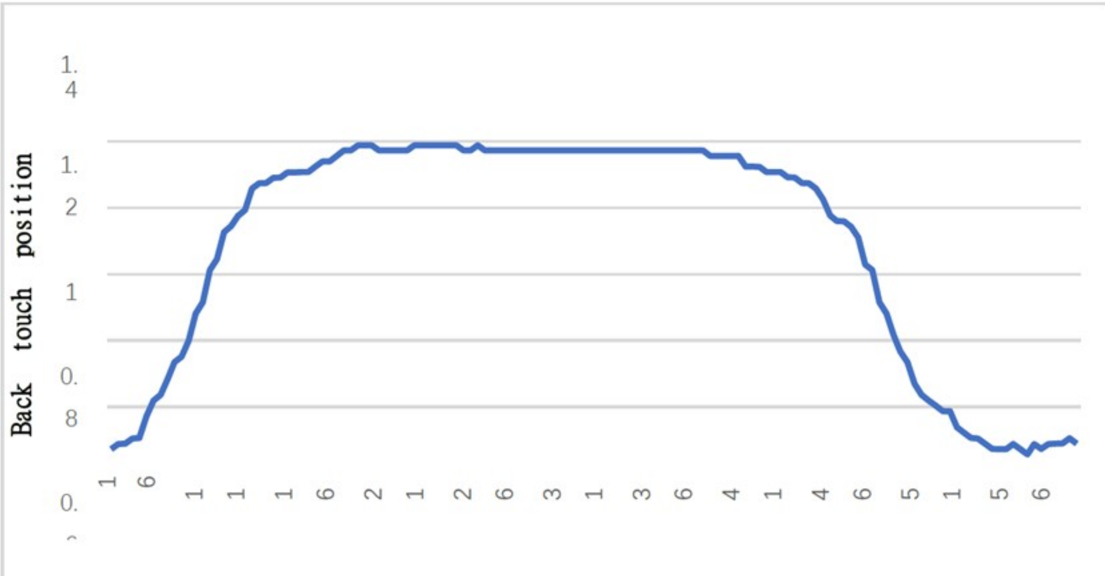

Figure 15. Touch the back of the hand from bottom to top - back touching position Detection results

5. **Discussion and application social benefits.** With the rapid development of society, population aging is rapidly becoming a global phenomenon, the incidence of chronic diseases is increasing, and more and more elderly people are facing serious challenges of impaired physical functions such as muscle strength, balance and mobility. According to incomplete statistics, there are 260 million chronic disease patients in my country, and the elderly are a group with a high incidence of chronic diseases. At present, the rehabilitation programs for the elderly in the health care systems of various countries are still in the initial stage. Routine rehabilitation is carried out under the direct supervision of hospital doctors. After discharge, they return to the community or family for further rehabilitation exercises to maintain and strengthen the effect of rehabilitation. However,

many elderly people have low adherence to the exercise program during home rehabilitation, which leads to prolonged treatment time and poor rehabilitation effect. Therefore, how to carry out effective exercise rehabilitation exercises in the home environment has become an urgent problem and research hotspot in home rehabilitation. The most difficult thing is how to use home rehabilitation equipment and how to accurately evaluate rehabilitation training.

This project proposes a computer vision-based family rehabilitation training evaluation method, and provides a system implementation. Using ordinary cameras to collect video stream information of patients' rehabilitation training, patients do not need to wear any equipment, which greatly reduces the complexity of traditional rehabilitation training equipment. The research results of this project have broad application prospects. It is planned to be promoted in the field of public medical care and community health care. It can effectively reduce the workload of physiotherapy rehabilitation medical staff and improve the persistence and effectiveness of family rehabilitation for the elderly.

6. **Conclusion.** Aiming at the difficulties and social problems of rehabilitation patients in home rehabilitation training, this paper proposes a home rehabilitation motion detection system based on the Open Pose open-source algorithm. Rehabilitation patients do not need to wear any equipment, and use home computers and ordinary camera equipment to systematically analyze and process video information of patients' rehabilitation training to obtain effective detection and guidance for rehabilitation actions. The system can solve difficult problems such as body recognition, motion detection and motion scoring at low cost and high efficiency. While solving the social problems of home-based rehabilitation training, it reduces the treatment cost and space constraints of rehabilitation patients, effectively reduces the workload of physiotherapy rehabilitation medical staff, and improves the persistence and effectiveness of home rehabilitation for the elderly. Allow home rehabilitators to receive professional and accurate rehabilitation training guidance at any time. The research results of this paper have broad application prospects are planned to be promoted in the field of public medical care and community healthcare.

Acknowledgement. This work is supported by the 2023 Horizontal Project Research Fund of the Scientific Research Office of Dongguan Polytechnic (ZXD202305) and supported by the 2023 Entrepreneurship Practice Project of Dongguan Polytechnic (2023CB02) and supported by the 2023 College Student Entrepreneurship Training Program of Dongguan Polytechnic (DC202301) and supported by the 2024 Entrepreneurship Practice Project of Dongguan Polytechnic (2024CB03) and Key Fields Special Project of Guangdong Provincial Department of Education in 2022 (2022ZDZX1074).

REFERENCES

[1] W. Gui, "Application of sensor and detection technology in Mechatronics System," *IOP Conference Series, Materials Science and Engineering*, vol. 750, pp. 27–29, 2019.

[2] Y. Cheng, H. Jiang, and F. Wang, "Using High-Bandwidth Networks Efficiently for Fast Graph Computation," *IEEE Transactions on Parallel and Distributed Systems*, vol. 30, pp. 1170–1183, 2019.

[3] Y. Dai, S. Wang, and X. Chen, "Generative adversarial networks based on Wasserstein distance for Knowledge graph embedding's," *Knowledge-Based Systems*, vol. 190, pp. 105165, 2020.

[4] L. Fu, D. Zhang, and Q. Ye, "Recurrent Thrifty Attention Network for Remote Sensing Scene Recognition," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, pp. 8257–8268, 2021.

[5] Q. Ye, P. Huang, Z. Zhang, Y. Zheng, L. Fu, and W. Yang, "Multi-view Learning with Robust Double-sided Twin SVM with Applications to Image Recognition," *IEEE Transactions on Cybernetics*, vol. 52, no. 12, pp. 12745–12758, 2022.

[6] L. Fu, Z. Li, Q. Ye, H. Yin, Q. Liu, X. Chen, X. Fan, W. Yang, and G. Yang, "Learning Robust Discriminant Subspace Based on Joint $L_{2,p}$- and $L_{2,s}$-Norm Distance Metrics," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 3, no. 1, pp. 130–144, 2022.

[7] X. Chen, M. Li, H. Zhong, Y. Ma, and C. Hsu, "DNN Off: Offloading DNN-based Intelligent IoT Applications in Mobile Edge Computing," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 4, pp. 2820–2829, 2022.

[8] X. Chen, J. Zhang, B. Lin, Z. Chen, K. Wolter, and G. Min, "Energy-Efficient Offloading for DNN-based Smart IoT Systems in Cloud-Edge Environments," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 3, pp. 683–697, 2022.

[9] X. Chen, J. Hu, Z. Chen, B. Lin, N. Xiong, and G. Min, "A Reinforcement Learning Empowered Feedback Control System for Industrial Internet of Things," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 4, pp. 2724–2733, 2022.

[10] X. Li, Y. Lin, and X. Liu, "Completely Independent Spanning Trees on BCCC Data Center Networks with an Application to Fault-Tolerant Routing," *IEEE Transactions on Parallel and Distributed Systems*, vol. 33, no. 8, pp. 1939–1952, 2022.

[11] G. Huang, C. Luo, K. Wu, Y. Ma, Y. Zhang, and X. Liu, "Software-Defined Infrastructure for Decentralized Data Lifecycle Governance: Principled Design and Open Challenges," *IEEE International Conference on Distributed Computing Systems*, pp. 1674–1683, 2019.

[12] Y. Chou, T. Kröger, and C. Pu, "Models of long-term care use among older people with disabilities in Taiwan: institutional care, community care, live-in migrant care and family care," *European Journal of Ageing*, vol. 12, no. 2, pp. 95–104, 2015.

[13] N. Tomita, K. Yoshimura, and N. Ikegami, "Impact of home and community-based services on hospitalization and institutionalization among individuals eligible for long-term care insurance in Japan," *BMC Health Services Research*, vol. 10, pp. 345–358, 2010.

[14] C. A. Garske, M. Dyson, S. Dupan, G. Morgan, and K. Nazarpour, "Serious Games Are Not Serious Enough for Myoelectric Prosthetics," *JMIR Serious Games*, vol. 9, no. 4, pp. 28079, 2021.

[15] P. F. Felzenszwalb, and D. P. Huttenlocher, "Pictorial Structures for Object Recognition," *International Journal of Computer Vision*, vol. 61, pp. 55–79, 2005.

[16] Z. Ambari-Nogini, R. Bidaki, and F. Maddizade, "The relationship between Zinc status and depression and anxiety in the elderly population," *Clinical Nutrition ESPEN*, vol. 37, pp. 51–59, 2020.

[17] D. Konstantinidis, V. Argyriou, T. Stathaki, and N. Grammalidis, "A modular CNN-based building detector for remote sensing images," *Computer Networks*, vol. 168, pp. 107034, 2020.

[18] G. Liu, X. Chen, and R. Zhou, "Social learning discrete Particle Swarm Optimization based two-stage X-routing for IC design under Intelligent Edge Computing architecture," *Applied Soft Computing*, vol. 10, pp. 107215, 2021.

[19] H. Xie, W. Lin, and S. Lin, "A multi-dimensional relation model for dimensional sentiment analysis," *Information Sciences*, vol. 579, pp. 832–844, 2021.

[20] J. Hargraves, J. Warren, and A. Celeste, "Community Health Workers Assisting Patients with Diabetes in Self-management," *Journal of Ambulatory Care Management*, vol. 35, no. 1, pp. 15–26, January/March 2012.

[21] G. Bamparopoulos, F. Konstantinidis, C. Bratsas, and P. D. Bamidis, "Towards exergaming commons. Composing the exergame ontology for publishing open game data," *Journal of Biomedical Semantics*, vol. 7, 2016.

[22] Y. Ling, L. P. TerMeer, Z. Yumak, and R. C. Veltkamp, "Usability Test of Exercise Games Designed for Rehabilitation of Elderly Patients after Hip Replacement Surgery," *JMIR Serious Games*, vol. 5, no. 4, pp. 7, 2017.

[23] R. Baluz, A. Teles, J. E. Fontenele, R. Moreira, R. Fialho, P. Azevedo, D. Sousa, F. Santos, V. H. Bastos, and S. Teixeira, "Motor Rehabilitation of Upper Limbs Using a Gesture-Based Serious Game," *Games for Health*, vol. 3, pp. 177–185, 2022.

[24] O. Mubin, F. Alnajjar, A. Al Mahmud, N. Jishtu, and B. Alsinglawi, "Exploring serious games for stroke rehabilitation. A scoping review," *Disability and Rehabilitation: Assistive Technology*, vol. 17, no. 2, pp. 159–165, 2022.

[25] C. H. Chen, and D. Ramanan, "3D Human Pose Estimation = 2D Pose Estimation + Matching," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp. 5759–5767, 2022.

[26] P. Verma, and R. Srivastava, "Reconsideration of multi-stage deep network for human pose estimation," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 9, no. 6, pp. 600, 2021.

[27] W. Chen, H. Wang, Y. Li, H. Su, and Z. Wan, "Synthesizing Training Images for Boosting Human 3D Pose Estimation," in *2016 Fourth International Conference on 3D Vision (3DV)*, Stanford, CA, USA, pp. 479–488, 2016.

[28] D. Mehta, H. Rhodin, D. Casas, and P. Fua, "Monocular 3D Human Pose Estimation in the Wild Using Improved CNN Supervision," in *2017 International Conference on 3D Vision (3DV)*, pp. 506–516, 2017.

[29] A. Sav, A. Salehi, F. S. Mair, and S. S. McMillan, "Measuring the burden of treatment for chronic disease: implications of a scoping review of the literature," *BMC Medical Research Methodology*, vol. 17, pp. 140, 2017.

[30] Y. C. Chou, and T. Kröger, "Community care in Taiwan: mere talk, no policy," *Social Work in Mental Health*, vol. 2, pp. 139–156, 2004.

[31] L. Dou, Z. Mao, Q. Fu, G. Chen, and S. Li, "Health-Related Quality of Life and Its Influencing Factors in Patients with Coronary Heart Disease in China," *Patient Preference and Adherence*, vol. 16, pp. 781–795, 2022.

[32] J. C. Campbell, N. Ikegami, and S. Kwon, "Policy learning and cross-national diffusion in social long-term care insurance: Germany, Japan, and the Republic of Korea," *International Social Security Review*, vol. 4, pp. 63–80, 2009.

[33] F. Le Bihan, and C. Martin, "Diversification of policy measures in the elderly care sector; towards greater flexibility for carers," *European Journal of Ageing*, vol. 9, pp. 2, 2012.

[34] C. Ionescu, J. Carreira, and C. Sminchisescu, "Iterated Second-Order Label Sensitive Pooling for 3D Human Pose Estimation," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, pp. 1661–1668, 2014.

[35] F. Zhang, T. Y. Wu, J. S. Pan, G. Ding, and Z. Li, "Human Motion Recognition Based on SVM in VR Art Media Interaction Environment," *Human-centric Computing and Information Sciences*, vol. 9, pp. 40, 2019.

[36] F. Zhang, T. Y. Wu, and G. Zheng, "Video salient region detection model based on wavelet transform and feature comparison," *EURASIP Journal on Image and Video Processing*, pp. 58, 2019.

[37] C. M. Chen, S. Liu, S. A. Chaudhry, and Y. C. Chen, "A Lightweight and Robust User Authentication Protocol with User Anonymity for IoT-Based Healthcare," *CMES-Computer Modeling in Engineering & Sciences*, vol. 131, 2022.

[38] G. Ke, R. S. Chen, Y. C. Chen, S. Wang, and X. Zhang, "Using ant colony optimization for improving the execution of material requirements planning for smart manufacturing," *Enterprise Information Systems*, vol. 16, pp. 379–401, 2022.

[39] L. Kang, W. Cao, R. S. Chen, and Y. C. Chen, "An adaptively reversed diffusion dual-drive evolutionary algorithm in dynamic environments for intelligence prediction," *Applied Soft Computing*, vol. 110, pp. 107761, 2021.

[40] B. Gan, C. Zhang, Y. Chen, and Y. C. Chen, "Research on role modeling and behavior control of virtual reality animation interactive system in Internet of Things," *Journal of Real-Time Image Processing*, vol. 18, pp. 1069–1083, 2021.

[41] T. Zhang, J. Yuan, Y. C. Chen, and W. Jia, "Self-learning soft computing algorithms for prediction machines of estimating crowd density," *Applied Soft Computing*, vol. 105, 2021.