A Probabilistic Neural Network-Based Dynamic Perception Framework for Rehabilitation Robots Using Multi-Modal Sensor Fusion

Lin Gui*, Zhenlai Chen
School of Automotive and Aeronautical Engineering, Henan Polytechnic Institute, Nanyang, 473000, China
E-mail: glgl520520@163.com, 19836796353@163.com
*Corresponding author

Keywords: multi-modal, sensors, information fusion, robot, dynamic perception

Received: March 24, 2025

The current mechanical dynamic perception methods for rehabilitation-assisted robots have low accuracy and low efficiency. In response to this problem, a robot mechanical dynamic perception model based on multi-modal sensor fusion is proposed in the research. The interaction force sensor and weight sensor are used to collect the interaction force and weight loss values during the patients' motion process, with a sample size of 5,378 collected. Then, the Kalman filtering algorithm is used for data processing. Finally, the sensor information is fused using a weighted fusion method, and a probabilistic neural network is used to determine the patient's motion, thereby achieving intelligent dynamic perception in rehabilitation training. The probabilistic neural network architecture includes the input layer, the hidden layer, and the output layer. The experiment was simulated and analyzed using MATLAB, and five volunteers were selected to carry out practical experiments. The results showed that the mechanical dynamic perception model proposed in the study could accurately perceive the motion intention of the user through multi-modal sensors and make accurate judgments. Its judgment accuracy reached 95.52%, and the response time was only 0.85 seconds. Compared with traditional rehabilitation assisted robot mechanical dynamic perception methods, this model significantly improves accuracy and efficiency. The method proposed in the study can further enhance the dynamic perception ability of robots through multi-modal sensor fusion, thereby improving the intelligence of rehabilitation training.

Povzetek: Predlagan je model dinamične zaznave rehabilitacijskega robota, ki s fuzijo interakcijskih in težnostnih senzorjev ter probabilističnim nevronskim omrežjem izboljša natančnost prepoznavanja pacientovega gibanja.

1 Introduction

The key technology for robots to achieve intelligent operation and motion control involves comprehensive perception of the environment and their own state, providing accurate information support for robot decision-making [1]. Driven by intelligent technology, the application of sensor technology in the field of robotics is becoming increasingly widespread. However, the information acquisition of a single sensor often has limitations and cannot meet the perception needs in complex environments. Therefore, multi-modal sensor fusion technology has emerged. The information from different sensors is integrated to improve the robot's perception accuracy of the external environment [2]. Challa et al. found that the stability of bipedal robots during walking was difficult to maintain. Therefore, a gait trajectory generation method based on long short-term memory networks was built. This method adopted motion capture sensors to capture gait data of humans walking on a treadmill. Finally, the sensor data was applied to the gait trajectory generation model. This method could effectively reduce the swinging of bipedal robots during walking and improve walking stability [3]. Zan proposed a path planning method based on whale optimization algorithm and computer perception to improve the path planning performance of robots in complex environments. This method achieved path obstacle avoidance through visual sensors. The path efficiency and smoothness planned by this method were significantly improved [4]. Shi et al. combined ultrasonic sensors and flexible frictional electric sensors to assist robots in perceiving objects to improve their ability to recognize and judge captured objects. The method achieved a recognition accuracy of 92.48% for captured objects [5].

At present, robots with high requirements for mechanical dynamic perception are mainly used in high-precision manufacturing, medical surgery, aerospace, and other fields. In these applications, robots need to be able to perceive and operate accurately to complete tasks [6-8]. Rajendran et al. reviewed the current research challenges of the selective harvesting robot to enhance the dynamic perception capability. The improvement

effect of integrating artificial intelligence, soft robots, and data-driven methods on the perception ability was discussed. The results indicated that data-driven methods and integrated artificial intelligence could significantly enhance the dynamic perception ability of robots, especially when dealing with complex and changing environments [9]. Wang et al. found that currently most soft actuators only had a single driving element and lacked perception. A dielectric elastomer actuator with multi-degree of freedom driving and perception functions was proposed to address this issue. By stimulating the selective dynamic of electrodes, all-round braking and perception were achieved. After applying this method, the robot's grasping action was more flexible [10]. Chen et al. designed a multi-modal fusion method based on deep learning to explore the collaborative intention recognition of robots towards human colleagues. This method combined visual, auditory, and tactile sensors to predict the behavior and intention through deep learning models. This method had high accuracy in human colleague collaboration intention recognition and could effectively improve the collaboration efficiency between robots and human colleagues [11]. Yan et al. proposed a robot with dynamic perception to assist surgeons in surgical procedures. The robot system predicted and perceived the insertion status of the guide wire through force feedback perception. The robot system assisted physicians in judging surgical progress and making intelligent decisions based on prediction and perception results. The perception accuracy reached 93.62% [12]. The comparison and summary of the performance indicators of the existing related work are shown in Table 1.

Table 1: Comparison and summary of the existing related work

Document Number	Title of Literature	Response time (s)	Perception accuracy (%)
[3]	An optimized-LSTM and RGB-D sensor-based human gait trajectory generator for bipedal robot walking	0.73	85.46
[4]	Research on robot path perception and optimization technology based on whale optimization algorithm	0.68	89.77
[5]	Soft robotic perception system with ultrasonic auto-positioning and multimodal sensory intelligence	0.71	90.12
[6]	Tacto: A fast, flexible, and open-source simulator for high-resolution vision-based tactile sensors	0.65	90.35
[9]	Towards autonomous selective harvesting: A review of robot perception, robot design, motion planning and control	0.66	90.74
[10]	Electroactive polymer-based soft actuator with integrated functions of multi-degree-of-freedom motion and perception	0.69	87.11
[11]	Influence of mobile robots on human safety perception and system productivity in wholesale and retail trade environments: A pilot study	0.70	90.23
[12]	Machine learning-based surgical state perception and collaborative control for a vascular interventional robot	0.77	89.56

As shown in Table 1, the current mechanical dynamic perception model still has problems such as low accuracy and poor real-time performance in perceiving the operator's motion intention in the process of assisting rehabilitation medical treatment [13]. Existing perception models have problems such as single sensor mode, insufficient multi-modal fusion, and low efficiency in PNN applications, which limit their perception accuracy and real-time performance in complex environments, making it difficult to meet the requirements of efficient collaboration and precise operation. This is because the sensor data fusion algorithm is not optimized enough, fails to make full use of multi-source information, and the PNN has low computational efficiency when processing large-scale data, resulting in the response speed and accuracy of the model in a dynamic environment being difficult to meet the clinical application standards. Therefore, multi-modal sensing information is combined to further enhance the dynamic perception capability of machinery. A robot mechanical dynamic perception model based on multi-modal sensors is proposed. This model analyzes the gait of patients and uses multiple sensors to identify and judge the intentions, further achieving the intelligence of rehabilitation training. The innovation of the research lies in taking the interaction force sensor and the weight sensor as the main features in the dynamic perception process of the robot. The accuracy of dynamic perception is further improved through the feature fusion of multi-modal sensors. Furthermore, the research proposed a mechanical motion pattern judgment model based on probabilistic neural networks, further enhancing machine's power sensing capability and assisting it in determining the appropriate next action. The research aims to improve the response speed and accuracy of the model in complex environments through multi-modal fusion and probabilistic neural network optimization, meet the requirements of clinical applications, and ultimately achieve efficient and precise rehabilitation training assistance. This is to address the bottleneck of perception accuracy and real-time performance caused by insufficient fusion of single perception mode and multiple modalities in current research, ensuring that the model

can quickly and accurately identify the patient's intention in a dynamic environment, and improving the effectiveness and safety of rehabilitation training.

2 Methods and materials

To assist users in motion, a robot mechanical dynamic perception model based on multi-modal sensor fusion is proposed. This model combines interaction force sensors and weight sensors to perceive the patient's motion intention, thereby assisting users in exercise training.

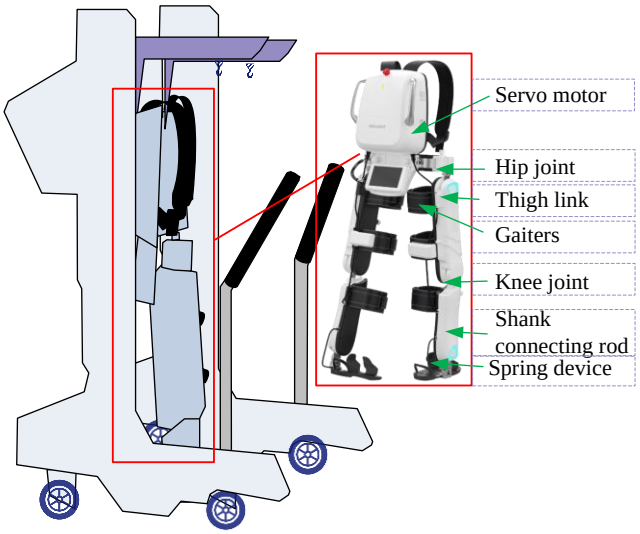
2.1 Perception data acquisition and signal processing based on multi-modal sensor fusion

Conventional image perception acquisition methods generally suffer from low acquisition accuracy. Therefore, the study first takes a gait analyzer to accurately divide the support phase and swing phase during walking [14-15]. In addition, interaction force sensors are used to obtain human-machine interaction force to evaluate the dynamic state of robot users during motion. To obtain more accurate dynamic perception data, the weight sensor is combined to further analyze the gait characteristics of users by monitoring the changes in their lower limb weight-bearing capacity. The placement position, sensor accuracy, sampling rate, and data range of each sensor are shown in Table 2.

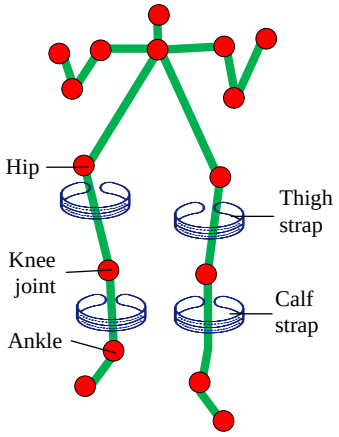
Table 2: Details of sensor placement

Serial number	Placement position	Type	Accuracy	Sampling rate (Hz)	Data range
1	Servo motor	Interaction force sensor	±0.5N	100	The full cycle of walking
2	Hip joint		±0.5N	100	The full cycle of walking
3	Thigh link		±0.5N	100	The full cycle of walking
4	Gaiters		±0.5N	100	The full cycle of walking
5	Knee joint		±0.5N	100	The full cycle of walking
6	Shank connecting rod		±0.5N	100	The full cycle of walking
7	Spring device		±0.5N	100	The full cycle of walking
8~9	At the left and right pull rods	Weight sensor	±0.1kg	50	The full cycle of walking

The dynamic perception robot proposed in the study and the fixed method acting on the patient's motion stage are shown in Figure 1.



(a) Lower limb rehabilitation robot structure



(b) The fixed pattern of the rehabilitation training phase

Figure 1: The proposed dynamic perception robot and the fixed mode acting on the patient's motion stage.

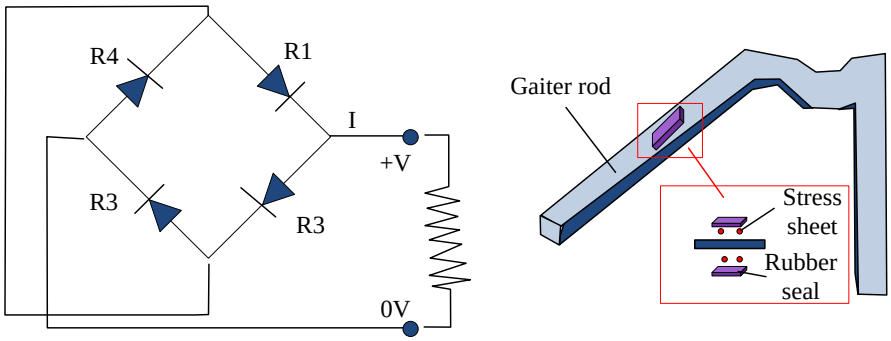
In Figure 1 (a), the robot in the study has an adjustable fixing device for adjusting between users of different heights and body types. Multiple sensor interfaces are also equipped in the structure to connect gait analyzers, interaction force sensors, and weight sensors, ensuring accurate data collection and transmission. In Figure 1 (b), the robot is bound to the user through leg warmers and straps. The interaction torque between the lower limbs and the machine is transmitted to the user's legs through the leg warmers. Therefore, the leg warmer is taken to design interaction force sensors that extract the interaction force during the motion. The dynamic equation of the motion perception robot is shown in equation (1).

$$\gamma = M(\alpha)\alpha'' + C(\alpha,\alpha')\alpha' + g(\alpha) - \psi \tag{1}$$

In equation (1), α, α', and α'' represent the swing angle of the rehabilitation robot and its first and second derivatives. $C(\alpha,\alpha')$ is the acceleration of the robotic arm. $g(\theta)$ is the gravity matrix. $M(\alpha)$ is the Coriolis acceleration matrix. γ is the joint torque of the rehabilitation robot arm. ψ is the interaction force between humans and machines. The interaction force between humans and machines is shown in equation (2).

$$\psi = \begin{bmatrix} S_3\cos\varepsilon & 0 \\ -S_3\cos\varepsilon & 0 \end{bmatrix}^T \phi\Delta x_3 + \begin{bmatrix} L_3\cos\varepsilon + S_4\cos(\eta+\varepsilon) & S_4\cos(\eta+\varepsilon) \\ -L_3\cos\varepsilon - S_4\cos(\eta+\varepsilon) & -S_4\cos(\eta+\varepsilon) \end{bmatrix}^T \phi\Delta x_4 \tag{2}$$

In equation (2), S_3 represents the distance from the patient's hip joint to the thigh binding rod. L_3 represents the length of the thigh and the thigh connecting rod. S_4 represents the distance from the knee joint to the calf connecting rod. ε represents the angle between the hip joint of the leg and the hip joint of the machine relative to the vertical line of the ground. η represents the angle between the knee joint of the human and the machine relative to the thigh connecting rod. φ represents the stiffness coefficient of the binding rod. It is set as a constant value in the research. The value is 1,000 N/m. Δx_3 and Δx_4 are respectively the displacements of the centroids of the thigh connecting rod and the calf connecting rod. Based on the interaction force calculation method and interaction force sensor data, the current motion status of rehabilitation training is perceived. Due to the high accuracy, wide measurement range, and long lifespan of resistance strain gauges, they are chosen for research to achieve interaction force perception. The circuit and installation location of the resistance strain gauge used in the study are shown in Figure 2.



(a) Strain gauge full bridge circuit (b) Stress sensor layout

Figure 2: Circuit of resistance strain gauge and installation position.

As shown in Figure 2 (a), a full bridge circuit is taken to connect resistance strain gauges to improve measurement accuracy and stability. This circuit can eliminate the influence of environmental factors such as temperature and ensure the accuracy of interaction force data. As shown in Figure 2 (b), four strain gauges are placed on each leg rod. These strain gauges can sense the slight deformation of the leg rod under force and convert it into a change in resistance value. During the user's motion, there are interactions with the machine, causing deformation of the strain gauge and resulting in a change in its resistance value. By measuring the changes in these resistance values, the interaction force between the user and the machine can be indirectly obtained. The specific process of resistance change is shown in equation (3).

$$K = \frac{\Delta R}{R} \tag{3}$$

In equation (3), R signifies the original resistance, measured in Ω. ΔR is the change in resistance, measured in Ω. K is the sensitivity coefficient of the resistance strain gauge, which is dimensionless. In addition to interaction force perception, the weight sensor is taken to detect changes in the user's body center of gravity and lower limb weight-bearing capacity. The weight perception proposed in the study is achieved through tension sensors arranged at different positions. The screw sensor is used to measure weight loss data during walking. After wearing the weight loss suit, the user connects the weight loss rope on the suit to the simple tension sensor on the weight reducing arm. By measuring the tension of the weight loss rope during the patient's walking process through a tension sensor, the weight-bearing capacity of the affected limb can be indirectly obtained. To avoid interference signals during the perception process, the Kalman filtering algorithm is used to preprocess sensor signals. The structure of the Kalman filter is shown in Figure 3.

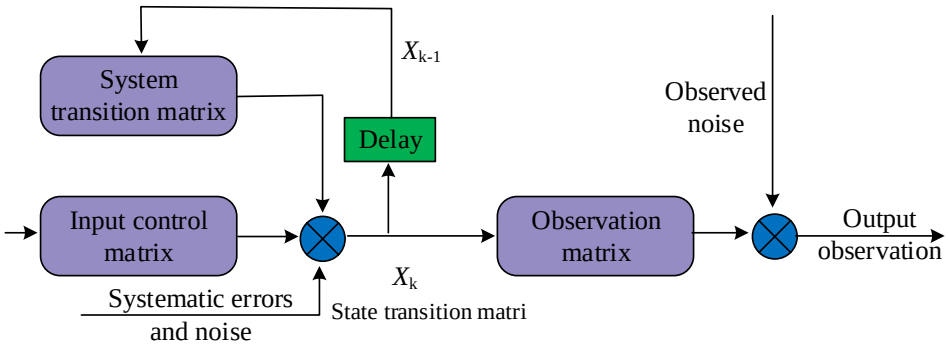


Figure 3: Kalman filter structure.

In Figure 3, the Kalman filter algorithm can effectively remove noise and interference, improve signal accuracy and reliability through prediction and update. In the prediction stage, the algorithm estimates the current state based on the state of the previous moment. In the update phase, the algorithm combines the current observation values to correct the prediction results and obtain the optimal state estimation. Based on this approach, the study ensures the stability and accuracy of sensor signals, providing a reliable data foundation for subsequent power perception models.

2.2 Mechanical dynamic perception model based on probabilistic neural network and multi-modal sensors

After setting up sensors and introducing filtering algorithms to process sensor signals, the probabilistic neural network is introduced to achieve mechanical

dynamic perception of robots. Through dynamic perception, robots assist users in motion. The user's motion mode includes two types: active and passive, which are classified according to the user's autonomous walking ability [16]. To enhance the applicability of the proposed robot, the information obtained from multi-modal sensors is utilized to achieve mechanical dynamic perception, and uses probabilistic neural networks to determine the user's motion pattern. The data during the training process is collected in real time through multi-modal sensors, including parameters such as force, displacement, and velocity. After preprocessing, it is input into the probabilistic neural network. The dataset is divided into the training set and the test set with a ratio of 7:3. The training process is subjected to ten-fold cross-validation to ensure the generalization ability of the model. The study takes high-performance computers for model training, equipped with Intel i9 processors, 64GB of memory, and NVIDIA RTX 3080 graphics cards. The operating system is Windows 11, the programming environment is Python 3.9, and the main dependency libraries include TensorFlow and Scikit learn. The sensor calibration program uses standard weights for static calibration to ensure measurement accuracy. The model parameters are optimized through grid search, and the final number of hidden layer nodes is determined to be 20 with a learning rate of 0.01. During the training process, the changes in the loss function are monitored in real-time to avoid over-fitting. To fuse multi-modal perception information, the weighted summation method is taken to integrate the feature information from different sensors. The specific process of the sensor data fusion method used in the study is shown in Figure 4.

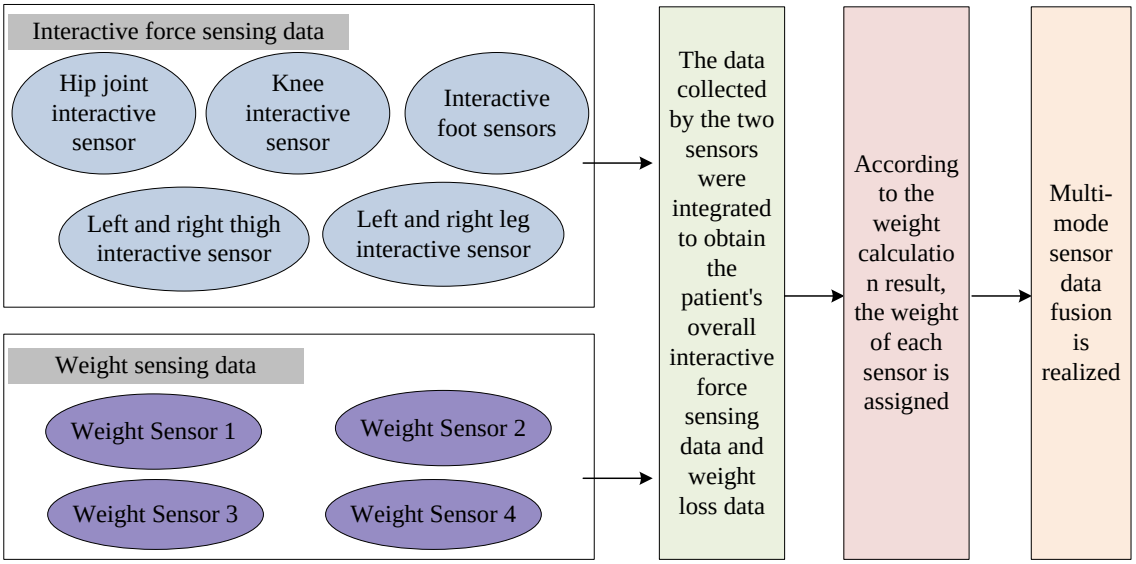


Figure 4: The specific process of sensor data fusion method.

In Figure 4, firstly, the data collected by multiple sensors is pieced together to obtain the overall motion data of the patient. Then, the weight assignment is performed based on the weight values of each sensor to further achieve feature fusion. The weight calculation method for sensors is shown in equation (4).

$$w_i = \left[\sigma_i^2 \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \right]^{-1} \tag{4}$$

In equation (4), σ_i is the measurement variance of the sensor. i is the sensor serial number. n is the number of sensors. w_i signifies the weight of the corresponding sensor. The measurement variance of the sensor is calculated based on multiple experimental data during the training process. The benchmark data adopted in the study is the kinematic model established by healthy subjects performing standard motions, which serves as a reference standard for auxiliary rehabilitation training. After obtaining the corresponding interaction force and weight loss value generated by the user during the gait cycle through data fusion, the probabilistic neural network is used to judge their motion patterns and further enhance the intelligence. Probabilistic neural network is a radial basis function network, which includes an input layer, a pattern layer, a summation layer, and an output layer [17]. Figure 5 displays the structure of the probabilistic neural network.

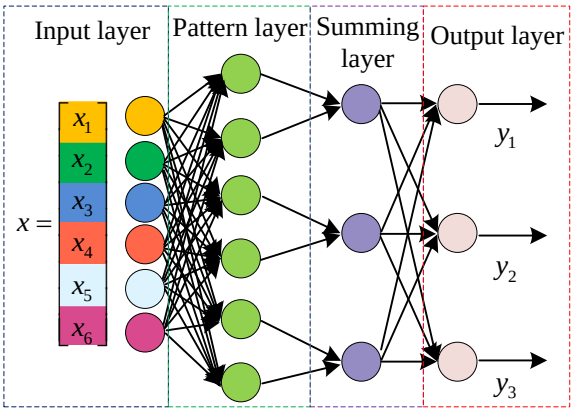


Figure 5: The structure of probabilistic neural network.

In Figure 5, in the probabilistic neural network, the input layer is responsible for receiving fused data from sensors. The pattern layer calculates the similarity between the input data and each training sample. The summation layer weights and sums the similarity to obtain probability estimates for each category. The output layer determines the final classification result based on the probability estimate. When using the probabilistic neural network for classification, the losses and risks in practical applications are considered. Probabilistic neural networks are sensitive to smoothing factors and data distribution, and need to be finely adjusted to optimize the classification effect. Therefore, the grid search method is taken to determine its smoothing parameters. The research conducts grid search, traverses different combinations of smooth parameters, evaluates the classification accuracy, and finally selects the optimal parameters. The input-output relationship determined in the hidden layer is shown in equation (5).

$$\xi_{kj}(x) = e^{-\frac{(x-x_{kj})(x-x_{kj})^T}{\sigma^2}} \cdot \left[\sqrt{2\pi}\sigma^d\right]^{-1} \tag{5}$$

In equation (5), $\xi_{ij}(x)$ is the output of the hidden layer. d is the dimension value of the sample space. x' signifies the sample value. k and j are the sample category and sample number, respectively. After outputting the hidden layer, the summation layer is used to weight the outputs of neurons belonging to the same type of hidden layer. The specific calculation process is shown in equation (6).

$$\zeta_k = \frac{\sum_{j=1}^{N} \varpi_{kj}}{N} \tag{6}$$

In equation (6), N signifies the number of samples in that category. ϖ_{kj} is a weighted value. ζ is the output result of the summation layer. The specific process of the proposed motion mode judgment method is shown in Figure 6.

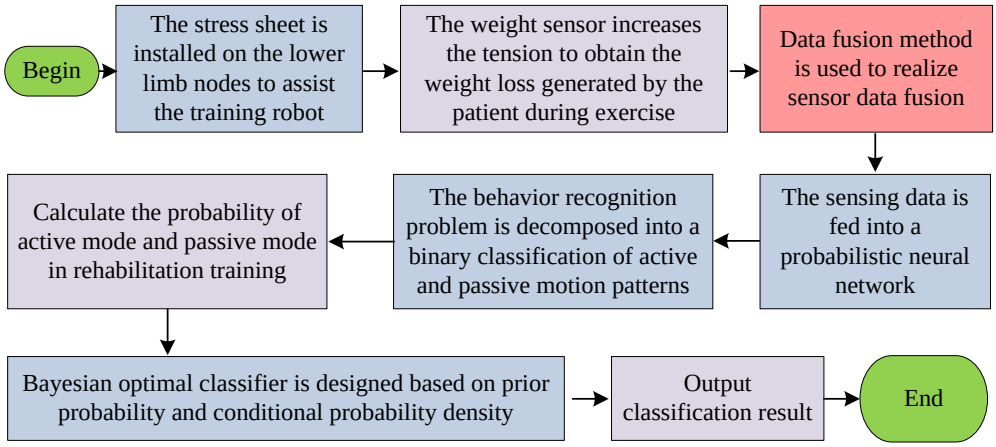


Figure 6: The specific process of the proposed motion mode judgment method.

As shown in Figure 6, the study inputs the interaction force data of each node and the data measured by weight sensors into the network model. The input variables also include the phase of each joint. These data are standardized after pre-processing steps to ensure that data of different dimensions can be compared at the same scale. These input variables can comprehensively reflect the dynamic state and motion characteristics. In the training phase, the research utilizes rich historical data to train the network to learn the mapping relationship between different motion patterns and sensor data. The probability density function for each motion mode is presented in equation (7).

$$f_{x,\gamma'} = \frac{(x-\gamma_i')^T (x-\gamma_i')}{2\varphi^2} \tag{7}$$

In equation (7), γ' signifies the connection weight. φ signifies the smoothing factor. $f_{x,\gamma'}$ is the probability density function for each motion mode. Based on the probability density calculation results, the motion mode

that best matches the current sensor data can be determined. Firstly, a large amount of rehabilitation training data is classified and labeled through the manual labeling method to ensure the accuracy and reliability of the data. The categories of motion intentions include standing, walking, sitting, standing, etc., and each category corresponds to a specific sensor data pattern. Subsequently, these annotated data are used for model training, gradually optimizing network parameters and improving recognition accuracy.

3 Results

To test the perception ability and application effect of the proposed mechanical dynamic perception model, a series of experiments are designed to analyze.

3.1 Analysis of data acquisition and processing effectiveness based on multi-modal perception and kalman filter

To test the perception ability and application effect of the method proposed in the research, simulation experiments and field operation experiments were designed for the research. The specific experimental setup process is shown in Figure 7.

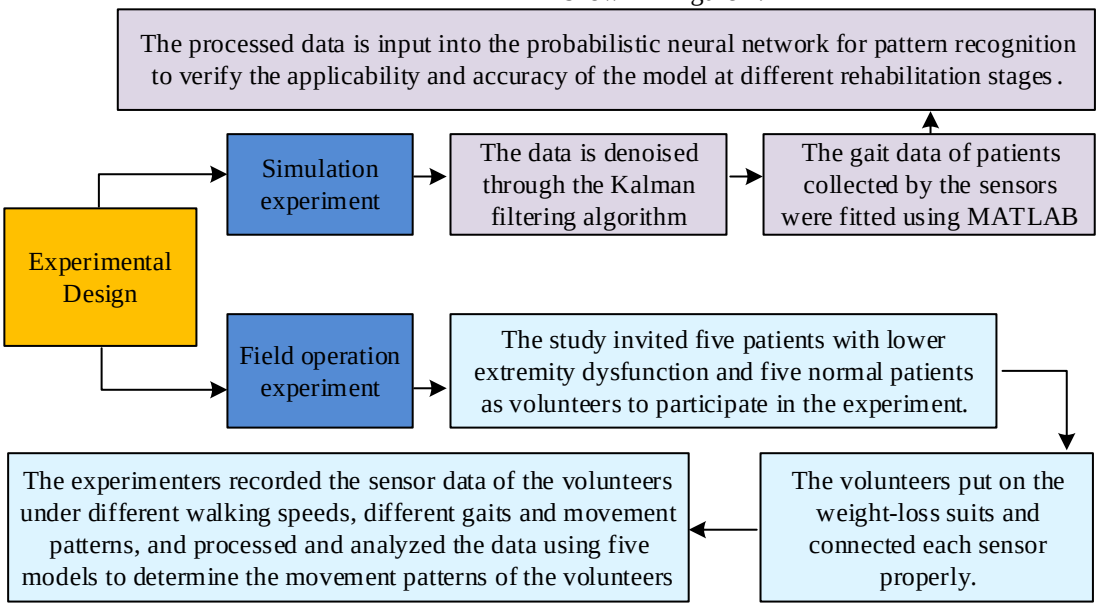
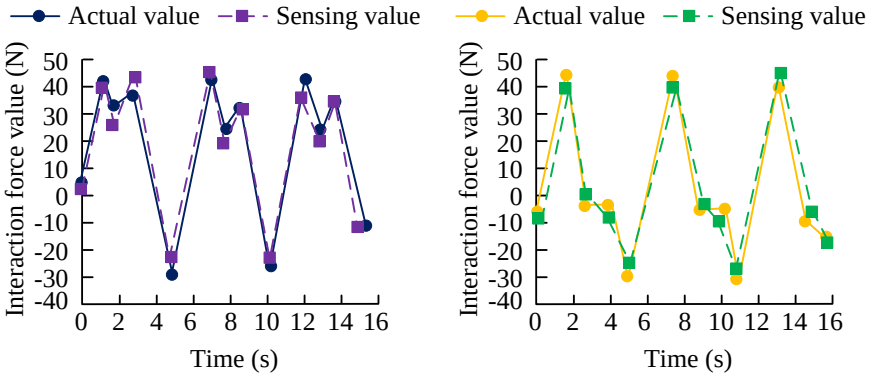


Figure7: The specific experimental setup process

As shown in Figure 7, the MATLAB was taken to fit the patient gait data collected by the sensor. Then, the Kalman filter algorithm was used to denoise data to ensure its accuracy and reliability. Subsequently, the processed data was input into the probabilistic neural network for pattern recognition to validate the applicability and accuracy of the model at different rehabilitation stages. In the field operation experiment, five patients with lower extremity dysfunction and five normal patients were invited as volunteers to participate in the experiment. During the experiment, volunteers put on weight-loss suits and correctly connected each sensor. Experimenters recorded sensor data from volunteers at different walking speeds, gaits, and motion modes, and used five models to process and analyze the data to determine the volunteers' motion mode. A total of 10 field operation experiments were conducted, each lasting for 30 minutes, to ensure the data diversity and repeatability. The patients were divided into two groups: normal patients and those with lower extremity dysfunction. They were grouped according to the severity of the patients' conditions and the rehabilitation stage to ensure the scientificity and comparability of the experimental results. The experimental results of the 10 patients did not differ much, and the results before and after the experiment were basically consistent.

To validate the accuracy of data collection from interaction force sensors, taking the rehabilitation training robot as an example, the patient gait data collected by the sensors is fitted using MATLAB. The simulation results of the exercise status of the same patient participating in rehabilitation training are compared with the actual measurement results collected by sensors. The results are shown in Figure 8.



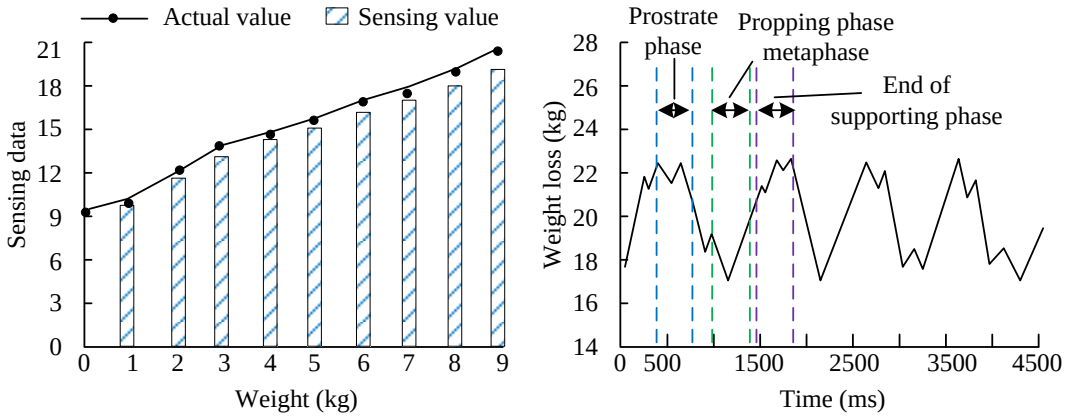
(a) Fitted curve of thigh interaction force value

(b) Leg interaction force fitting curve

Figure 8: Accuracy analysis of interaction force perception data.

In Figure 8 (a), the simulation results of thigh interaction force were basically consistent with the data curve collected by the sensor, reaching 0.95. As shown in Figure 8 (b), the simulation and measured data of leg interaction force also showed high consistency, further confirming the reliability of data collection.

Weight sensors were employed to determine the gait cycle of patients during the rehabilitation process, thereby assisting in dynamic perception. To validate the accuracy of the weight sensor installation method and the data reliability, a dummy model is used and weights of 1, 2, 3, 4, 5, 6, 7, 8, and 9kg are sequentially added to the dummy model to simulate patients of different weights. Then, the data from the weight sensor is collected at each weight. The original weight of the dummy is 10kg. The measurement results of four weight sensors and the changes in weight loss values of normal patients with walking status are shown in Figure 9.



(a) Four weight sensor measurements

(b) Change of weight loss value with walking state

Figure 9: Measurement curves of four weight sensors with different weights.

As shown in Figure 9 (a), with the increase of weights, the tension value detected by the sensor gradually increased, and the curve fitting between its measured value and the actual weight value of the dummy reached 0.97. As shown in Figure 9 (b), the measurement results of weight loss values also exhibited periodic changes over time. In the early stage of the supporting phase, the weight loss value gradually increased, which corresponded to the process of the patient gradually transferring their body center of gravity to the weight loss device. After entering the mid support phase, the weight loss value remained stable. At the end of the supporting phase, the weight loss value gradually decreased, as the patient begins to shift their body weight from the weight loss device back to their own legs. This indicates that the weight loss measurement of the weight sensor can further assist the robot in determining the current motion status of the patient.

3.2 Perception effect analysis based on multi-modal sensor fusion

To process the collected data and avoid interference noise affecting the accuracy of dynamic perception, the Kalman filter is used for data preprocessing. To validate the rationality of the proposed data processing method (Method 1), the study compares it with the currently popular data preprocessing methods. The comparative

methods include sliding average filtering (Method 2), median filtering (Method 3), and wavelet transform denoising (Method 4). The study first simulates sensor data containing different levels of noise, and then applies these four methods for preprocessing. The Mean Square Error (MSE) and R-squared (R value) of the interaction force perceived by the model after four processing methods are taken to evaluate the performance, as displayed in Table 3.

Table 3: Quality comparison of sensor data before and after noise reduction.

Project	SNR=10dB		SNR=20dB		SNR=30dB	
	MSE	R value	MSE	R value	MSE	R value
Method 1	0.05	0.93	0.03	0.95	0.01	0.97
Method 2	0.10	0.87	0.09	0.89	0.07	0.90
Method 3	0.09	0.88	0.08	0.90	0.06	0.91
Method 4	0.07	0.90	0.05	0.92	0.04	0.93
Method 5	0.12	0.85	0.11	0.87	0.08	0.88

As shown in Table 3, Method 1 outperformed other comparison methods on MSE and R-squared under SNR of 10dB, 20dB, and 30dB. Especially, at SNR=10dB, Method 1 had the lowest MSE of only 0.01 and the highest R value, reaching 0.97.

To further verify the rationality of the sensor fusion method, the study uses data before and after fusion for dynamic perception, judges the patient's motion mode, and compares the results. The accuracy and running time of the motion mode judgment of the research model before and after data fusion are shown in Figure 10.

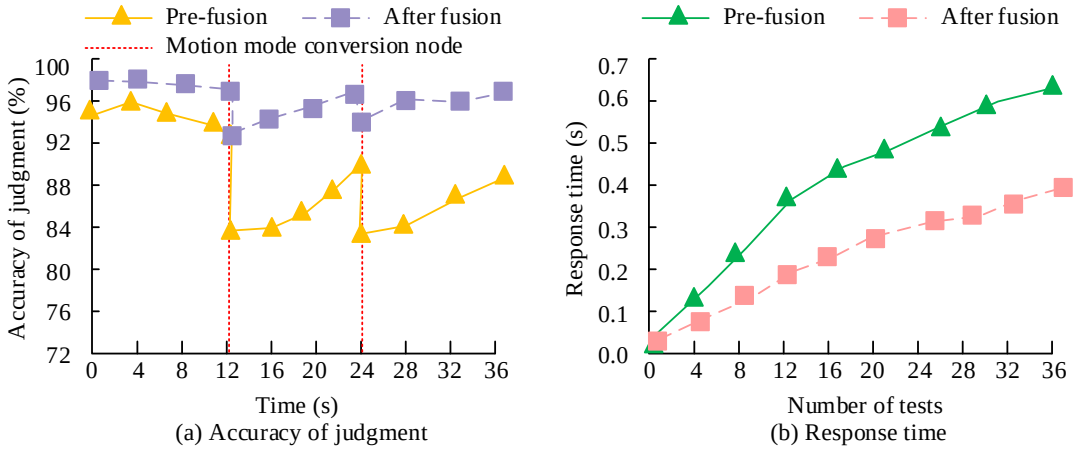


Figure 10: The accuracy and running time before and after the fusion data.

In Figure 10 (a), the accuracy of data judgment before fusion was relatively low. Especially at the patient motion mode transition, the model was prone to misjudgment, with an average accuracy of 85.47%. The fused data significantly improved the judgment accuracy, with an accuracy of 95.84%, which enabled more accurate identification of the patient's motion intention. As shown in Figure 10 (b), the model's running time was reduced by 0.28s after data fusion, indicating that the sensor fusion method not only improved the judgment accuracy, but also enhanced the running efficiency.

3.3 Analysis of the practical application effect of dynamic perception model

To test the performance of the proposed multi-modal perception fusion robot mechanical dynamic perception model (Model 1), a comparative experiment is conducted on Model 2 in reference [18], Model 3 in reference [19], Model 4 in reference [20], and Model 5 in reference [21]. Model 2 focuses on the adaptability of soft robot skin and enhances the interactive ability of collaborative robots through body tactile perception. Model 3 combines machine learning and artificial synapses to accurately respond to intelligent flexible sensing systems. Model 4 utilizes active visual tactile interaction to optimize object pose estimation. Model 5 is based on neural network assisted filtering to improve the robustness of multi-modal indirect sensing in soft robots. These five models are applied to the perception process of robot rehabilitation training. Five patients with lower limb dysfunction and five normal patients are recruited as volunteers to participate in the experiment. During the experiment, volunteers wear weight loss clothing and connect various sensors. The experimenters record the sensor data of volunteers at different walking speeds, gaits, and motion patterns, and use five models to process and analyze the data to determine the volunteers' motion patterns. The motion mode refers to active walking and passive walking for patients. Motion intention recognition

refers to accurately identifying the volunteer's motion intention, including actions such as starting, accelerating, decelerating, and stopping. The method for determining the ground conditions is to measure the ground reaction force through high-precision pressure sensors, capture the ground texture information in combination with visual sensors, and comprehensively judge parameters such as ground hardness and friction coefficient to ensure that the model can accurately identify the motion intention under different ground conditions. During the experiment, the accuracy of motion mode judgment for five models is recorded, and the results are shown in Figure 11.

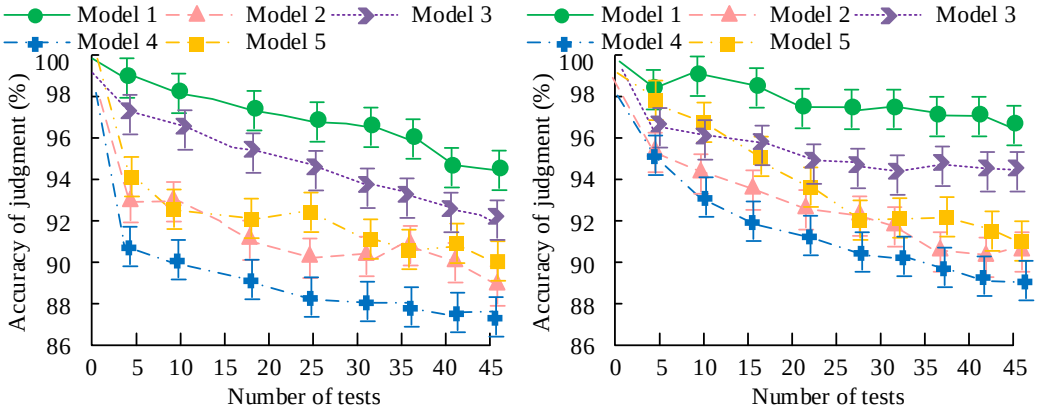


Figure 11: Motion mode judgment accuracy of five models.

In Figure 11 (a), the accuracy of each model in judging the motion mode of patients with lower limb dysfunction gradually decreased with the increase of testing times. However, the overall judgment performance of Model 1 was significantly higher than other methods, with a judgment accuracy rate consistently above 92%. As shown in Figure 11 (b), for normal patients, the accuracy of the five models was relatively higher, with Model 1 having the highest accuracy, at 96.88%.

To further test the practical application effect, volunteers were asked to judge their walking speed, gait characteristics, and motion modes during rehabilitation training in different scenarios, and the accuracy and stability were evaluated. The experimental scenarios include walking on flat ground, crossing obstacles, and going up and down stairs. Volunteers wear weight loss clothing and connect sensors to carry out rehabilitation training activities according to instructions. The accuracy and response time for each model in three scenarios are presented in Table 4.

Table 4: Motion intention recognition accuracy and response time of each model under three scenarios.

Project	Walking on flat ground		Crossing obstacles		Going up and down stairs	
	Intention recognition accuracy (%)	Response time (s)	Intention recognition accuracy (%)	Response time (s)	Intention recognition accuracy (%)	Response time (s)
Model 1	95.48±1.55*&%#	0.85±0.09*&%#	95.72±1.69*&%#	0.89±0.08*&%#	95.36±1.47*&%#	0.82±0.07*&%#
Model 2	85.69±2.43	1.62±0.12	85.71±1.40	1.68±0.12	85.65±1.68	1.77±0.14
Model 3	90.34±1.78	1.05±0.09	91.00±1.83	1.08±0.10	90.52±1.63	1.07±0.11
Model 4	83.54±1.49	1.81±0.10	84.00±1.74	1.82±0.12	84.11±1.72	1.80±0.13
Model 5	88.45±2.00	1.21±0.11	88.90±1.69	1.12±0.11	88.23±1.88	1.36±0.12

Note: In Table 4, * indicates that the comparison results between Model 1 and Model 2 are different (p<0.05). & Indicates that the comparison results between Model 1 and Model 3 are different (p<0.05). % Indicates that there is a difference in the comparison results between Model 1 and Model 4 (p<0.05). # Indicates that the comparison results between Model 1 and Model 5 are different (p<0.05).

According to Table 4, the intention recognition accuracy of Model 1 was significantly higher than the other four models in these three scenarios, reaching 95.48%, 95.72%, and 95.36%, respectively. Meanwhile,

the response time of Model 1 was relatively short, at 0.85s, 0.89s, and 0.82s respectively, indicating its high real-time performance. In contrast, the other models had lower accuracy in intent recognition and longer response time. Especially in complex scenarios such as crossing obstacles and going up and down stairs, Model 1 had more obvious advantages, with an intention recognition accuracy about 10% higher than other models and a response time shortened by nearly 50%. This indicates that Model 1 can quickly and accurately perceive the patient's motion intention, providing timely and effective auxiliary support for the robot. In addition, Model 1 has demonstrated good stability and adaptability, and can maintain high accuracy in different patients and rehabilitation stages.

To test the effectiveness of the proposed intent recognition model applying Kalman filtering and multi-sensor fusion in the research, ablation experiments are conducted on the model. The Kalman filter and sensor fusion modules are removed separately for testing. The comparison results are shown in Table 5.

Table 5: Ablation experiments of the intention recognition model

Project	Module			Accuracy (%)
	Kalman filter	Multi-sensor fusion	Basic model	
1	-	-	√	87.47
2	-	√	√	90.36
3	√	-	√	90.11
4	√	√	√	95.48

As shown in Table 5, when using the multi-sensor fusion or Kalman filter module alone, the model accuracy was improved, but did not reach the level of Model 1. The combination of the two can significantly improve the accuracy and real-time performance of intent recognition, verifying the effectiveness of the research method. Further analysis shows that Kalman filtering effectively reduces noise interference, while multi-sensor fusion enhances data integrity. The synergistic effect of the two significantly improves the performance of the model in complex scenarios.

4 Discussion and conclusion

Traditional rehabilitation training methods for lower limb dysfunction often rely on the therapist's experience and judgment, lacking objective and accurate evaluation methods. Therefore, a robot mechanical dynamic perception model based on multi-modal sensor fusion was proposed, providing a new solution for rehabilitation training of patients with lower limb dysfunction. A set of behavior recognition input variables was constructed by collecting patient interaction force perception data and weight loss value data. Afterwards, the probabilistic neural network was taken to recognize the patient's motion intention, assisting the robot to generate correct responses to the patient's motions. The fitting degree between the interaction force perception data and the measured data reached 0.95, and the data changes transmitted through the weight sensor could accurately determine the patient's motion phase. The R-squared of the data processed by Model 1 remained above 0.93 and the MSE did not exceed 0.05 at 10dB, 20dB, and 30dB. In practical applications, the perception model proposed in the research can accurately perceive the motion intentions of patients and thereby achieve accurate auxiliary training. Its average recognition accuracy reached 95.52%, and the average response time was only 0.85 seconds. The recognition accuracy of the other four models reached 95.48%, 95.72% and 95.36% respectively. The response time was 0.85s, 0.89s and 0.82s, respectively. Compared with these four models, the model proposed in the study has higher real-time performance and accuracy. The performance difference occurs because the research combines Kalman filtering and multi-modal data fusion technology, effectively reducing noise interference and improving the accuracy and efficiency of data processing. The Kalman filtering algorithm optimizes the sensor data in real time, ensuring the intention recognition accuracy. Meanwhile, multi-modal fusion enhances the adaptability of the model to different scenarios, enabling Model 1 to maintain excellent performance in complex environments. This indicates that the proposed model not only improves the accuracy of motion intention recognition, but also shortens response time, providing patients with more timely and effective auxiliary support. At present, the research has not considered eliminating signal interference through machine structures. In the future, mechanical structure optimization will be considered to further improve the accuracy and stability of data acquisition.

References

[1] Y. Lin, J. Lloyd, A. Church, and N. F. Lepora, "Tactile gym 2.0: Sim-to-real deep reinforcement learning for comparing low-cost high-resolution robot touch," IEEE Robotics and Automation Letters, vol. 7, no. 4, pp. 10754-10761, 2022. https://doi.org/10.1109/LRA.2022.3195195

[2] X. Liu, C. Huang, H. Zhu, Z. Wang, J. Li and A. Cangelosi, "State-of-the-Art Elderly Service Robot: Environmental Perception, Compliance Control, Intention Recognition, and Research Challenges," IEEE SYST MAN CYBERN, vol. 10, no. 1, pp. 2-16, 2024. https://doi.org/10.1109/MSMC.2023.3238855.

[3] S. K. Challa, A. Kumar, V. B. Semwal, and N. Dua, "An optimized-LSTM and RGB-D sensor-based human gait trajectory generator for bipedal robot

walking," IEEE Sensors Journal, vol. 22, no. 24, pp. 24352-24363, 2022. https://doi.org/10.1109/JSEN.2022.3222412

[4] J. Zan, "Research on robot path perception and optimization technology based on whale optimization algorithm," Journal of Computational and Cognitive Engineering, vol. 1, no. 4, pp. 201-208, 2022. https://doi.org/10.47852/bonviewJCCE597820205514

[5] Q. Shi, Z. Sun, X. Le, J. Xie, and C. Lee, "Soft robotic perception system with ultrasonic auto-positioning and multimodal sensory intelligence," ACS Nano, vol. 17, no. 5, pp. 4985-4998, 2023. https://doi.org/10.1021/acsnano.2c12592

[6] S. Wang, M. Lambeta, P. W. Chou, and R. Calandra, "Tacto: A fast, flexible, and open-source simulator for high-resolution vision-based tactile sensors," IEEE Robotics and Automation Letters, vol. 7, no. 2, pp. 3930-3937, 2022. https://doi.org/10.1109/LRA.2022.3146945

[7] F. Mumuni, and A. Mumuni, "Bayesian cue integration of structure from motion and CNN-based monocular depth estimation for autonomous robot navigation," International Journal of Intelligent Robotics and Applications, vol. 6, no. 2, pp. 191-206, 2022. https://doi.org/10.1007/s41315-022-00226-2

[8] R. J. Kirschner, H. Mayer, L. Burr, N. Mansfeld, S. Abdolshah, and S. Haddadin, "Expectable motion unit: Avoiding hazards from human involuntary motions in human-robot interaction," IEEE Robotics and Automation Letters, vol. 7, no. 2, pp. 2993-3000, 2022. https://doi.org/10.1109/LRA.2022.3144535

[9] V. Rajendran, B. Debnath, S. Mghames, W. Mandil, S. Parsa, S. Parsons, and E. A. Ghalamzan, "Towards autonomous selective harvesting: A review of robot perception, robot design, motion planning and control," Journal of Field Robotics, vol. 41, no. 7. pp. 2247-2279, 2024. https://doi.org/10.48550/arXiv.2304.09617

[10] R. Wang, C. Zhang, W. Tan, J. Yang, D. Lin, and L. Liu, "Electroactive polymer-based soft actuator with integrated functions of multi-degree-of-freedom motion and perception," Soft Robotics, vol. 10, no. 1, pp. 119-128, 2023. https://doi.org/10.1089/soro.2021.01

[11] Y. Chen, C. Yang, Y. Gu, and B. Hu, "Influence of mobile robots on human safety perception and system productivity in wholesale and retail trade environments: A pilot study," IEEE Transactions on Human-Machine Systems, vol. 52, no. 4, pp. 624-635, 2022. https://doi.org/10.1109/THMS.2021.3134553

[12] Y. Yan, H. Wang, H. Yu, F. Wang, J. Fang, J. Niu, and S. Guo, "Machine learning-based surgical state perception and collaborative control for a vascular interventional robot," IEEE Sensors Journal, vol. 22, no. 7, pp. 7106-7118, 2022. https://doi.org/10.1109/jsen.2022.3154921

[13] L. Zongxing, Z. Jie, Y. Ligang, C. Jinshui and L. Hongbin, "The Human–Machine Interaction Methods and Strategies for Upper and Lower Extremity Rehabilitation Robots: A Review," IEEE SENS J, vol. 24, no. 9, pp. 13773-13787, 2024. https://doi.org/10.1109/JSEN.2024.3374344.

[14] H. S. Nasution, A. Jayadi, and R. Rikendry, "Implementasi metode fuzzy logic untuk sistem pengereman robot mobile berdasarkan jarak dan kecepatan," Jurnal Teknik Dan Sistem Komputer, vol. 3, no. 1, pp. 15-24, 2022. https://doi.org/10.33365/jtikom.v3i1.1634

[15] A. B. Latupeirissa, C. Panariello, and R. Bresin, "Probing aesthetics strategies for robot sound: Complexity and materiality in movement sonification," ACM Transactions on Human-Robot Interaction, vol. 12, no. 4, pp. 1-22, 2023. https://doi.org/10.1145/3585277

[16] B. Lacevic, A. M. Zanchettin, and P. Rocco, "Safe human-robot collaboration via collision checking and explicit representation of danger zones," IEEE Transactions on Automation Science and Engineering, vol. 20, no. 2, pp. 846-861, 2022. https://doi.org/10.1109/TASE.2022.3167772

[17] D. Felix Brown and S. Quan Xie, "Effectiveness of Intelligent Control Strategies in Robot-Assisted Rehabilitation—A Systematic Review," IEEE T NEUR SYS REH, vol. 32, no. 3, pp. 1828-1840, 2024. https://doi.org/10.1109/TNSRE.2024.3396065.

[18] Z. Ye, G. Pang, K. Xu, Z. Hou, H. Lv, Y. Shen, and G. Yang, "Soft robot skin with conformal adaptability for on-body tactile perception of collaborative robots," IEEE Robotics and Automation Letters, vol. 7, no. 2, pp. 5127-5134, 2022. https://doi.org/10.1109/LRA.2022.3155225

[19] T. Sun, B. Feng, J. Huo, Y. Xiao, W. Wang, J. Peng, Z. Li, C. Du, W. Wang, G. Zou, and L. Liu, "Artificial intelligence meets flexible sensors: Emerging smart flexible sensing systems driven by machine learning and artificial synapses," Nano-Micro Letters, vol. 16, no. 1, pp. 14-21, 2024. https://doi.org/10.1007/s40820-023-01235-x

[20] P. K. Murali, A. Dutta, M. Gentner, E. Burdet, R. Dahiya, and M. Kaboli, "Active visuo-tactile interactive robotic perception for accurate object pose estimation in dense clutter," IEEE Robotics and Automation Letters, vol. 7, no. 2, pp. 4686-4693, 2022. https://doi.org/10.1109/LRA.2022.3150045

[21] J. Y. Loo, Z. Y. Ding, V. M. Baskaran, S. G. Nurzaman, C. P. Tan, "Robust multimodal indirect sensing for soft robots via neural network-aided filter-based estimation," Soft Robotics, vol. 9, no. 3, pp. 591-612, 2022. https://doi.org/10.1089/soro.2020.0024