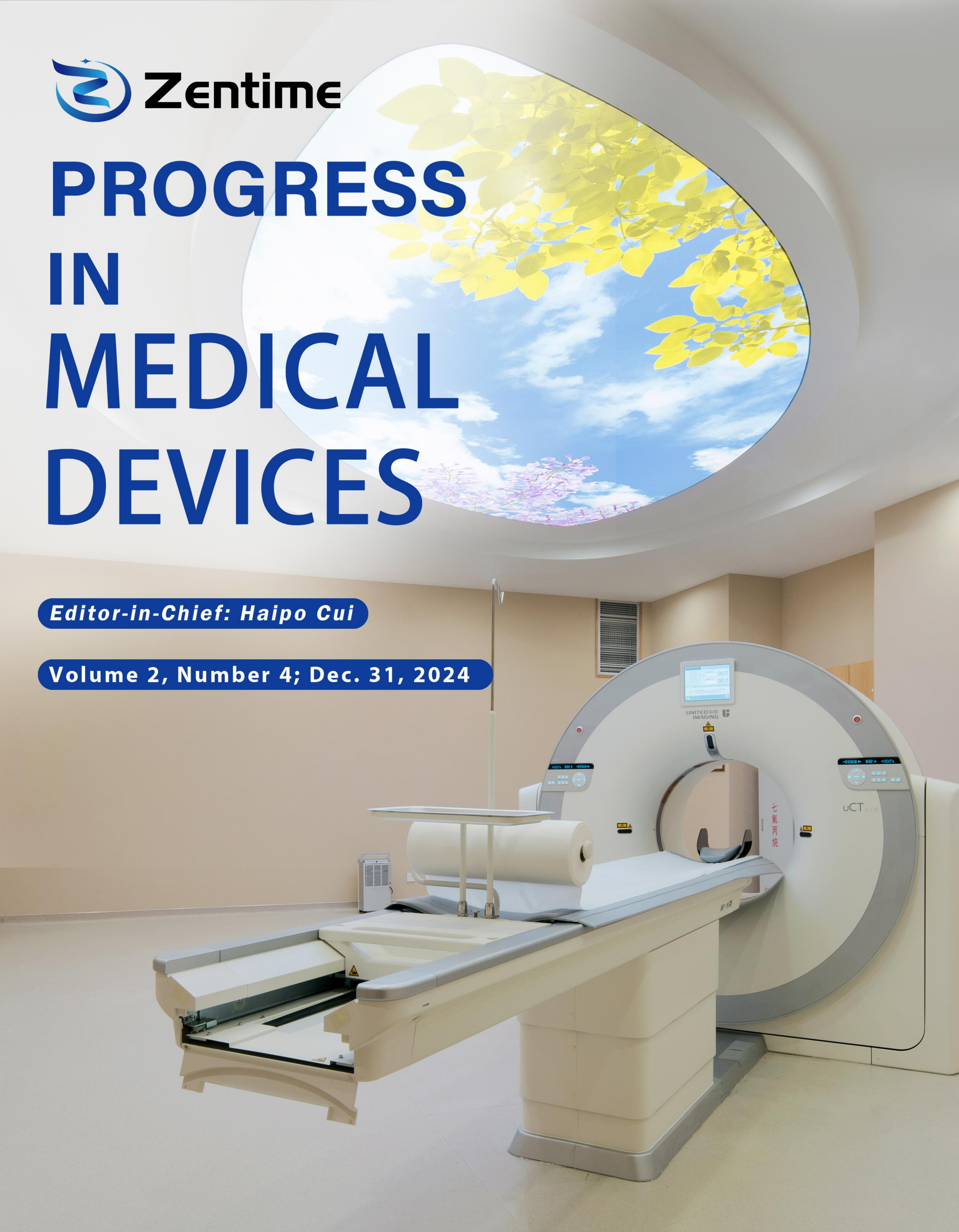


Zentime

PROGRESS IN MEDICAL DEVICES

Editor-in-Chief: Haipo Cui

Volume 2, Number 4; Dec. 31, 2024

ISSN: 2957-5478

Published by ZENTIME PUBLISHING CORPORATION

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Research Article https://doi.org/10.61189/434505lacrsk

An investigation of upper extremity impedance modeling and sensory thresholds in envelope wave electrical stimulation

Renling Zou[1], Yuhao Liu[1], Yicai Wu[1], Liang Zhao[1], Jigao Dai[1], Xiufang Hu[1], Xuezhi Yin[2]

[1]School of Health Sciences and Engineering, University of Shanghai for Science and Technology, Shanghai, 200000, China. [2]Shanghai Berry Electronic Technology Co., Ltd., Shanghai, 200000, China.

Corresponding authors: Renling Zou.

Acknowledgement: This work was supported by Science and Technology Commission of Shanghai Municipality (21S31906000), the National Natural Science Foundation of China (NSFC) Grant (61803265), and Medical-industrial cross-project of USST Grant (1022308524).

Declaration of conflict of interest: None.

Received April 24, 2024; Accepted June 25, 2024; Published December 31, 2024

Highlights

- This study introduces a novel impedance model for the human upper limb, providing a highly accurate fit between frequency and impedance values.
- The newly proposed Voltage Perception Threshold (VPT) method offers a more reliable measure of electrical stimulus sensation, independent of current magnitude and output frequency, compared to the traditional Current Perception Threshold (CPT).

Abstract

Objectives: To investigate how impedance values and sensory thresholds at various human upper limb sites affect the parameter settings of electrical stimulation equipment in low and medium frequency envelope electrical stimulation therapy. **Methods:** The study involved testing different upper limb sites on 22 healthy subjects (ages 21-25, 11 males and 11 females) by adjusting the modulation wave frequency, carrier frequency, and current intensity of the output. Five types of electrodes of various sizes were used in the tests. **Results:** The impedance test results for the human upper limb showed a wide range of impedance values across electrodes of different sizes. A new impedance model of the human upper limb was proposed, which accurately fits the relationship between frequency and impedance values. In electrical stimulus sensory experiments, the voltage perception threshold (VPT) introduced in this study was identified as a novel metric for electrical stimulus sensation. Unlike the current perception threshold, VPT does not consider the effects of current magnitude and output frequency. The range of sensory thresholds was 6-8 V, while the suprathreshold was 9-11 V. Neither experiment showed gender differences. **Conclusions:** Determining the value of the power supply and the output intensity of device power amplification circuitry based on the VPT can provide a more precise therapeutic dose for electrical stimulation therapy.

Keywords: Electrical stimulation, human impedance, voltage perception threshold, envelope

Introduction

As a rehabilitation technique for neurological injury diseases, electrical stimulation therapy has garnered extensive attention and research in recent years due to its ability to help patients rebuild or restore some motor functions [1, 2]. Currently, various forms of electrical stimulation instruments are available on the market. Most of these devices adjust treatment intensity based on subjective pain perception of patients, lacking capability to record the physiological information, such as bioimpedance, surface electromyography [3]. This kind of physiological information is essential for guiding role in the treatment of patients [4-7].

Address correspondence to: Renling Zou, School of Health Sciences and Engineering, University of Shanghai for Science and Technology, No.516 Jungong Road, Shanghai 200000, China. E-mail: zourenling@usst.edu.cn.

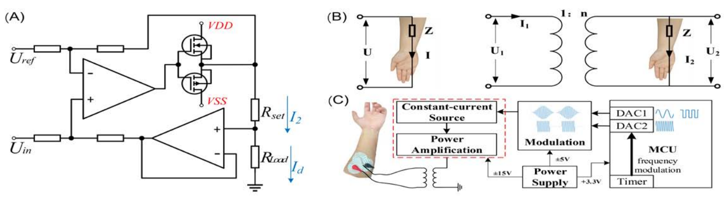


Figure 1. Electrical stimulation circuit block diagram. (A) Constant-current source model of the electrical stimulation device; (B) Functional block diagram of electrical stimulation module; (C) The relationship between the impedance of human body and the power supply of power amplifier circuit. VDD,voltage drain drain; VSS,voltage source and sink; DAC,digital-to-analog converter; MCU,microcontroller unit.

The electrical model of the human body is highly complex, and the biological passive impedance model contains complex domains, necessitating the consideration of both direct current and alternating current in waveform selection [8-12]. YinQimin noted an issue when comparing square waves with differential waves in experiments: square waves induce the human body to produce differential waves, making it impossible to evaluate the effect of the two waveforms by comparing two different differential waves [13-15]. One key point we propose is that a patient's sensation to a stimulus does not necessarily indicate that the waveform is undistorted. Many studies lack the measurement of relevant physiological information before setting electrical stimulation parameters [15]. Zhou et al. did not account for measuring human body impedance before designing the boost voltage source [16]. Boosting the power supply to a certain value without determining the impedance value can lead to excessive circuit power consumption or cause output waveform distortion in electrical stimulation equipment.

Pain and tolerance threshold can also be considered as a form of physiological information, providing valuable clinical insights for evaluating neurological defects, selecting appropriate treatments, and monitoring sensory recovery after injury [17]. Electrical stimulation is a simple method for measuring pain sensation [17-20]. Previous studies have predominantly used current perception threshold (CPT) as a criterion [21-25]. Hiremath SV et al. showed that the receptive intensity increases with the higher pulse amplitude and frequency [26]. However, there is limited research on the reasons for the differences in CPT across different genders and ages. This is because the adjustment of electrical stimulation intensity is usually based on subjective verbal feedback, and information about the stimulation waveform, electrode size and other parameters is often incomplete or inconsistent. Thus, relying on a patient's subjective experience to determine stimulation dose is unreliable, highlighting the need for methods to objectively quantify pain. Seno SI et al. measured CPT with five electrode sizes, finding that the CPT values increased with electrode size [18]. After adjusting for body fat percentage or body water content, the logarithmic conversion values showed no gender differences between electrode pairs.

The impedance value of the human upper limb directly affects the supply voltage required by stimulation equipment, while the sensory threshold affects the parameter setting of the electrical stimulation and the therapeutic effect. This study was conducted to investigate the influence of human upper limb impedance and electrical stimulation sensory threshold on the parameter setting of electrical stimulation equipment using a controlled variable method to obtain relevant physiological information. The variables examined included five different sizes of hydrogel electrodes, modulating waveforms ranging from 2-150 Hz, carrier waveforms with output frequencies from 1-10 kHz, and output currents ranging from 0-20 mA.

Methods and materials

Theory background

Figure 1A illustrates a model of a constant current source circuit used in electrical stimulation equipment. The current magnitude is given by (1):

$$I_d = I_2 = \frac{U_{in}}{R_{set}} \tag{1}$$

where I_d represents the current flowing through the human body (with R_{Load} as the impedance of the human upper limb), U_{in} is the input voltage signal, and VDD (voltage drain drain) is the power supply for the power amplifier circuit in the device. U_{ref} is referenced to ground, the output current is solely determined by the input voltage U_{in}, and is independent of the load impedance R_{Load}.

The relationship between the power supply of the power amplifier module and the impedance amplitude of the human body can be simplified to the model shown in **Figure 1B**. Previous studies have shown that the intensity of electrical stimulation is typically determined by the current magnitude [27]. When the output current I is fixed, the power supply voltage U should be set according to the maximum impedance Z. If $U < Z * I$, the output waveform will be distorted; if U is too large, it will increase the power consumption of the entire circuit. Therefore, before setting the power supply value, it is crucial to determine the range of the human impedance values. To prevent the output waveform distortion or excessive U values during the measurement, this study first measured the human impedance using the circuit shown in **Figure 1B**. A transformer was added between the constant-current source power amplifier circuit and the human body load. The transformer isolates the human body from the circuit board and acts as an impedance transformer. By converting the high impedance on the load side to a lower impedance on the output side, a lower supply voltage can be utilized without increasing circuit power consumption. In this study, the power supply $U1$ was set to 15 V, and the number of turns on the transformer's secondary coil was chosen to be 20, allowing for a maximum undistorted voltage of 300 V to be measured on the load side of the human body. The power supply voltage can be adjusted at a later stage according to the measured impedance values.

Instrument design

This device utilizes a constant-current source for the electrical stimulation output, ensuring that the output current remains unaffected by changes in the impedance of the human upper limb. **Figure 1C** presents a functional block diagram of the electrical stimulation module, including MCU (microcontroller unit), modulation module, constant-current source output and power amplification modules. The DAC1 (digital-to-analog converter) output of the MCU provides the modulating waveform, which can be selected as either a sinusoidal waveform or square waveform. Given the complex impedance characteristics of the human body, a square wave output will result in a differential wave, so DAC2 is configured to output only alternating current signals. The waveform frequency can be modulated by adjusting the timer settings. To enhance bandwidth capacity, the outputs from the MCU's two DACs are followed by a voltage follower. It is necessary to convert the DAC1 output to 0-1 V and the DAC2 output to -1 -1 V. These outputs are then fed into the AD835 multiplier, where they are modulated to produce the two waveforms depicted in the modulation module. Finally, the electrical stimulation is delivered to the human body through electrode pads via a constant current source and power amplification circuit.

Before conducting the experiment, it is essential to evaluate the output capability of the constant current source. This evaluation was performed using a 100 Hz sinusoidal modulating waveform and a 2 kHz sinusoidal carrier waveform while outputting a constant current. The voltage across the resistive loads, with varying resistance values, was measured. The output current strength was set in increments ranging from 5 to 25 mA, divided into 5 steps. The nominal resistance value of the loads were between 1 and 5 KΩ (with a power rating of 50W and an accuracy of 1%), and each load was measured individually. **Table 1** presents the measurement results, including the average value and standard deviations. The maximum standard deviation observed was 0.17 mA, which basically meets the requirements for a constant current source.

Results

The experiment used five hydrogel electrode pads in different sizes, as shown in **Figure 2**, to perform the following two tests: 1) A fixed current was output, while the modulating wave frequency and carrier frequency were adjusted. The impedance value of different upper limb sites was then determined using Ohm's law by measuring the voltage at various points on the upper limb. 2) Currents of varying sizes and frequencies were output to quantify the pain perception threshold. The study involved 11 healthy male subjects (age: 23±2 years, height: 173.2±5.1 cm, weight: 70.8±13.2 Kg) and 11 healthy female subjects (age: 22±2 years, height: 163±4.7 cm, weight: 52±8.6 Kg). All subjects participated voluntarily and were informed of the purpose of the experiment. They were also made aware that they could withdraw from the experiment at any time.

Human Upper Limb Impedance Measurement

The measurement of human upper limb impedance considered several factors: different parts of the human upper limb, modulating wave frequency, carrier frequency and electrode size. As shown in **Figure 3**, the positive electrode of the differential electrode pair was fixed on the left brachioradialis muscle, and the negative elec-

Table 1. Output current corresponding to different resistance loads and their average value and SD

		Current intensity / mA				
		5	10	15	20	25
Load/Ω	1K	5.10	10.8	15.20	20.40	25.40
	2K	5.15	10.2	15.00	20.20	25.20
	3K	5.00	9.98	14.96	19.96	24.88
	4K	4.10	10.10	15.10	20.10	25.12
	5K	5.08	10.20	15.20	20.28	25.26
Statistics	mv	5.08	10.18	15.09	20.19	25.17
	sd	0.05	0.14	0.10	0.15	0.17

Note: MV, mean value; SD, standard deviation.

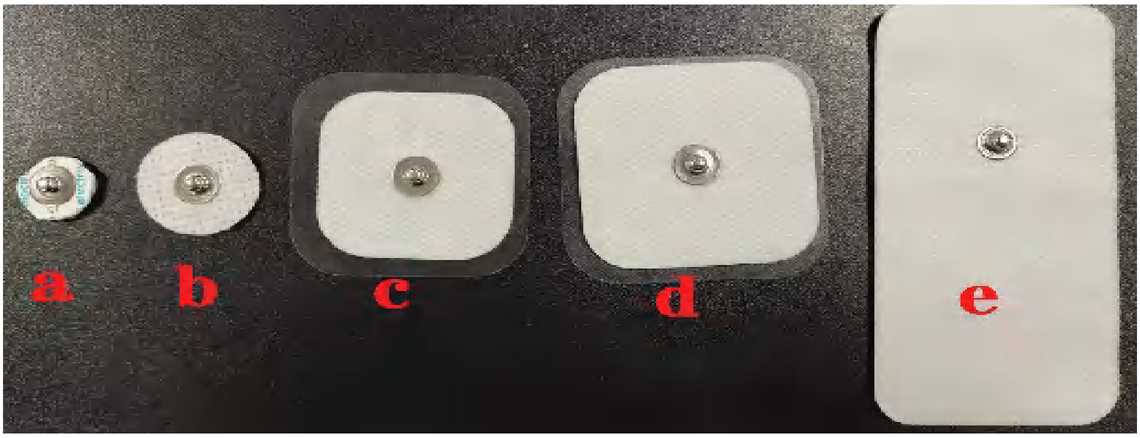


Figure 2. Hydrogel electrode size. The size of each electrode sheet: $3.53cm^2$ (a), $9.82cm^2$ (b), $16\ cm^2$ (c), $25\ cm^2$ (d), $50\ cm^2$ (e).

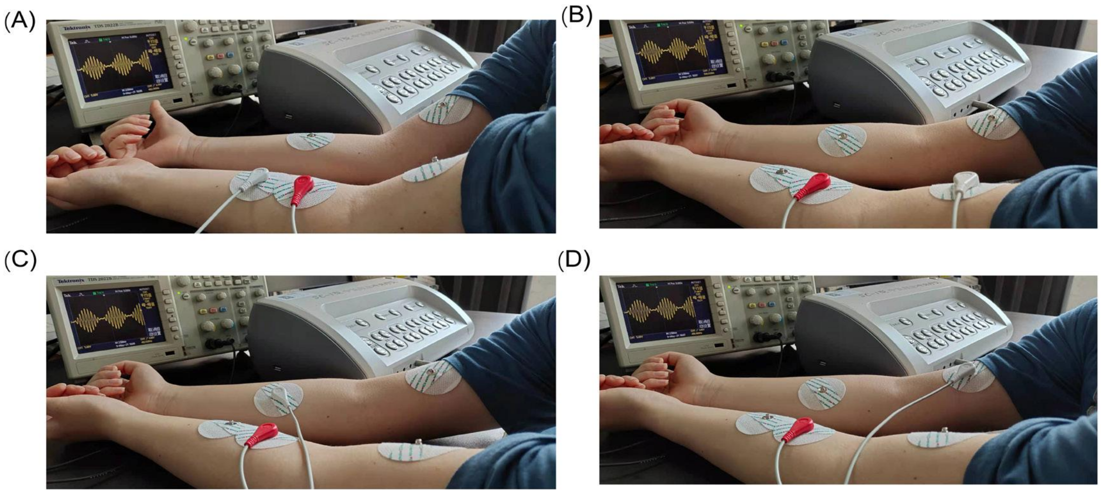


Figure 3. Impedance measurements were performed on different positions of the upper limb. (A) left brachioradialis-left brachioradialis; (B) left brachioradialis-left biceps brachii; (C) left brachioradialis-right brachioradialis; (D) left brachioradialis-right biceps brachii.

trode was placed on the following four different sites: (A) left brachioradialis muscle, (B) left biceps brachii muscle, (C) right brachioradialis muscle, and (D) right biceps brachii muscle. The electrodes used were Ag/AgCl disposable cardiac electrode sheets. The modulating waveform was a 10 Hz sine wave, and the carrier waveform was a 1000 Hz sine wave. During the experiment, the modulating wave frequency was varied in 1 Hz increments within the range of 2-150 Hz, with results consistently around 19.2 V. This consistency indicates that varying the electrode position on the upper limb or changing the modulating wave frequency had negligible effects on the measured impedance values. Consequently, this experiment primarily focused on examining the relationship between carrier frequency, electrode size, and the impedance of the upper limb.

Measurement result

Table 2. The impedance minima, maxima, and peak frequencies corresponding to the impedance maxima measured for the two test groups in the carrier frequency range of 100-10000 Hz (standard deviation in parentheses)

Electrode (cm^2)	Male (n=11)					Female (n=11)				
	3.53	9.82	16	25	50	3.53	9.82	16	25	50
Min/KΩ	0.51 (0.05)	0.45 (0.04)	0.39 (0.04)	0.34 (0.05)	0.27 (0.02)	0.51 (0.04)	0.44 (0.03)	0.42 (0.03)	0.38 (0.04)	0.31 (0.02)
Max/KΩ	21.80 (3.52)	10.98 (3.53)	8.06 (0.89)	6.18 (0.56)	3.79 (0.32)	19.80 (3.18)	11.68 (2.82)	8.16 (0.79)	5.56 (0.57)	3.29 (0.29)
Peak frequency/Hz	2000	1500	700	500	400	2000	1500	700	500	400

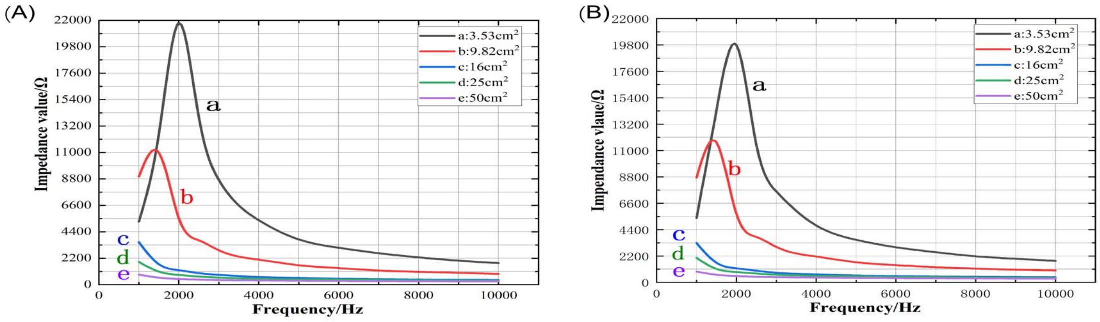


Figure 4. Carrier frequency (100-10000 Hz) corresponded impedance value of five electrode sizes. (A) Male subjects; (B) Female subjects.Increments of 100 Hz within the 100-1000 Hz range, 500 Hz within 1-3 kHz range, and 1000 Hz for adjustments within 3-10 kHz range.

The carrier frequency was adjusted in increments of 100 Hz within the 100-1000 Hz range, 500 Hz within 1-3 kHz range, and 1000 Hz for adjustments within 3 -10 kHz range. Measurements were made sequentially using five electrode pads. Tests were conducted separately on 11 males and 11 female subjects. **Figure 4A** shows the test results for the male group, and **Figure 4B** shows the test results for the female group. The data indicate that the impedance values varied significantly with different electrode sizes, although the overall trend remained consistent. As the carrier frequency increased, the impedance initially peaked before decreasing as the frequency continued to rise. Additionally, larger electrode sizes corresponded to lower carrier frequencies at which peak impedance occurred. When the carrier frequency exceeded 5 kHz, the impedance plateaued. No significant differences were observed between genders. **Table 2** lists the three impedances for both test groups: the impedance minima, maxima, and peak frequencies corresponding to the impedance maxima measured for each electrode size. The peak frequencies were identical, and the impedance values did not differ significantly by gender.

Human Upper Limb Impedance Model

Analysis of the experimental data reveals a maximum impedance peak at a specific frequency [28]. The three-element equivalent circuit model proposed by Cole is widely used in analyzing human impedance. In an RC (Resistance-capacitance) circuit model, impedance peaks are not possible; therefore, the traditional human impedance model is not suitable for electrical stimulation involving the envelope waveform formed by modulating and carrier waveforms. For low- and medium- frequency envelope waveform electric stimulation, a suitable passive circuit model can be used to fit the data, and the values of the parameters related to the impedance of the upper limb of the human body can be further analyzed. Based on the impedance variation pattern of the human upper limb shown in **Figure 5**, it is suggested that the upper limb exhibits not only capacitive characteristics but also possibly inductive characteristics. The combination of inductance and capacitance may result in a resonance frequency where the impedance peaks. The expression for the impedance value Z is:

$$Z = R_1 + \frac{1}{j\omega C} // (R_2 + j\omega L) = R_1 + \frac{\frac{1}{j\omega C}(R_2 + j\omega L)}{j\omega L + R_2 + \frac{1}{j\omega C}} \tag{2}$$

To achieve a peak frequency in this model, assuming , the expression of Z is:

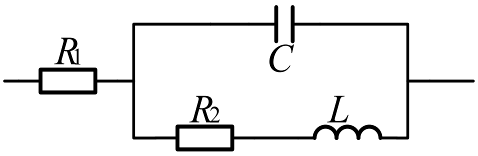


Figure 5. Passive circuit model of human impedance.

$$Z \approx R_1 + \frac{\frac{L}{C}}{R_2 + j(\omega L - \frac{1}{\omega C})} \quad (3)$$

When $\omega L = \frac{1}{\omega C}$, the impedance value Z reached its maximum, and the maximum value $z_m \approx L/(R_2C)$. The corresponding peak frequency $z_m \approx L/(R_2C)$. **Table 3** presents the relevant data measured in the experiment.

Figure 6 shows the fitting result of formula (3) to the impedance value and its corresponding frequency. Twenty-one discrete points were obtained from each electrode, with the solid line representing the fitted curve. **Table 4** shows the fitted values of the components in the circuit model and the theoretical peak frequencies and maxima impedance calculated from the fitted values. The analysis reveals that the inductance L remains nearly constant, while the fitted capacitance C increases with electrode size, and the fitted resistance R_2 decreases with electrode size. A linear relationship was observed between electrode area (Area), capacitance C, and resistance R_2. The relationship between C and electrode size (Area) is defined by equation (4) and the relationship between R_2 and C is defined by equation (5). Based on the fitted values, the peak frequency calculated by the formula $f_{peak}=1/(2\pi\sqrt{LC})$ closely matches the original measured peak frequency. Similarly, the theoretical impedance value calculated by $z_m \approx L/(R_2C)$ is consistent with the measured value. In summary, the circuit model accurately fits to the measured impedance of the human upper limb, providing a robust model for obtaining various physiological information in medical research involving low and medium frequency envelope electrical stimulation.

$$C = 5.577Area - 18.15 \quad (4)$$

$$R_2 = \frac{10000}{\sqrt{C}} \quad (5)$$

Pain perception threshold measurement

The relationship between pain perception, current intensity, and carrier frequency was investigated by adjusting these variables during experiments conducted with five electrodes of different sizes. Pain perception was assessed based on patient feedback, where the current intensity and carrier frequency were varied, and the patient's reactions were observed. When electrical stimulation was present, adjustments were made until the sensation disappeared, though current output remained—this point was recorded as the subsensory threshold. The sensory threshold was recorded when electrical stimulation was felt, and the supra-sensory threshold was recorded when there was noticeable electrical stimulation and numbness [29].

According to these three sensory criteria, the carrier frequency was adjusted within the 6-10 kHz, where the impedance changes slowly. The current intensity was varied from 0-20 mA, with a increment about 1 mA.

During measurement, as the carrier frequency increased, the voltage amplitude decreased, resulting in reduced electrical stimulation experienced by the subjects. Increasing the current at this point raised the voltage amplitude and enhanced the electrical stimulation. **Table 5** shows the test results for one subject. For each type of electrode, the standard deviation of the voltage measured at the three sensation levels was less than 1V across different frequencies. The results indicated consistent voltage thresholds for electrical stimulation among subjects. In summary, the voltage at both ends of the stimulation site can replace CPT as a measure of pain perception, namely voltage perception threshold (VPT).

Figure 7 shows the VPT values obtained from each electrode, separated by sex. **Figure 7a** shows the perceptual threshold corresponding to the five electrodes. The measured sensory threshold values for males were 11.04, 9.20, 6.97, 7.17 and 6.92 V, while the values for females were 12.75, 9.25, 6.68, 7.13 and 6.15 V, respectively. There were significant differences between sexes when using 3.53 cm^2 ($P < 0.001$) and 50 cm^2 ($P < 0.01$) electrodes, while the VPT values were similar for other three electrodes. Figure 7b shows the measurement results above the sensory threshold. The measured supra-sensory threshold values for males were 16.54, 11.36, 8.49, 8.89, and 8.51 V, whereas for females, the values were 18.58, 12.74, 8.61, 9.12, and 8.17 V. Significant differences between males and females were found when using the 3.53 cm^2 ($P < 0.001$) and 9.82 cm^2 ($P < 0.001$) electrodes, while VPT values remained consistent for the other three electrodes. For both sensory and supra-sensory

Table 3. Human body impedance obtained from five electrodes

Electrode	Z_{min}/KΩ	Z_{max}/KΩ	f_{peak}/Hz
EL 1	0.51	20	2000
EL 2	0.45	11	1500
EL 3	0.4	8	700
EL 4	0.39	6	500
EL 5	0.29	3.5	400

Note: Z_{max}: the maximum impedance value measured for each electrode; f_{peak}: the peak frequency corresponding to Z_{max}.

Table 4. Fitted values for each electrode in the circuit model

Electrode/ c^{m}2	L/H	C/nF	R_2/Ω	f_{peak}/Hz	Z_{max}/KΩ
EL 1(3.53)	0.6817	10.41	3178	1890	20.60
EL 2(9.82)	0.6839	22.27	2238	1290	13.72
EL 3(16)	0.7719	75.75	1204	659	8.46
EL 4(25)	0.7306	121.50	934.4	534	6.43
EL 5(50)	0.6777	261.30	659.7	378	3.93

Table 5. Pain perception value of subjects by adjusting electrode sizes and different carrier frequencies

Electrode	Perception		Carrier frequency/KHz					MV	SD
			6	7	8	9	10		
EL 1	Readout (Vpp)	subthreshold	8.4	8	9.4	8.4	8	8.44	0.51
		threshold	11	10.4	11	11.2	10.2	10.76	0.39
		suprathreshold	20	20.8	21.6	22	20	20.88	0.82
EL 2	Readout (Vpp)	subthreshold	9.2	9.8	9.2	10	9.2	9.48	0.35
		threshold	11	10	10.6	10	10.5	10.42	0.38
		suprathreshold	15.8	15.6	14.4	13.4	16	15.04	0.99
EL 3	Readout (Vpp)	subthreshold	6.8	6.8	7.68	7.6	6.72	7.12	0.42
		threshold	7.44	7.6	8	8.12	7.12	7.66	0.37
		suprathreshold	9.68	9.68	10.2	9.68	9.84	9.82	0.20
EL 4	Readout (Vpp)	subthreshold	7.12	6.4	7.2	8.12	6.72	7.11	0.58
		threshold	7.92	7.12	7.92	8.6	7.44	7.80	0.50
		suprathreshold	9.68	9.6	10.6	10.7	9.6	10.04	0.50
EL 5	Readout (Vpp)	subthreshold	6.16	5.84	6.4	6.72	7.44	6.51	0.55
		threshold	7.2	6.4	7.12	7.44	7.92	7.22	0.49
		suprathreshold	9.68	10.2	9.52	10.1	9.52	9.8	0.29

Note: Vpp, voltage peak-peak; MV, mean value; SD, standard deviation.

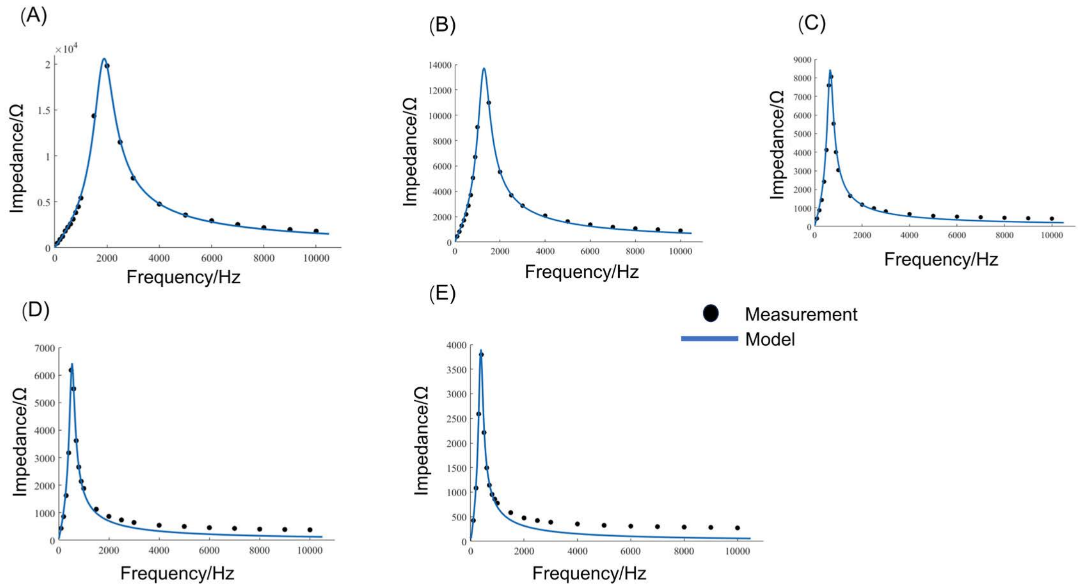


Figure 6. Fitting result of equation (3) for impedance value and its corresponding frequency. The solid dot is the measured data, and the solid line is the fitting curve. (A) Electrode 1; (B) Electrode 2; (C) Electrode 3; (D) Electrode 4; (E) Electrode 5.

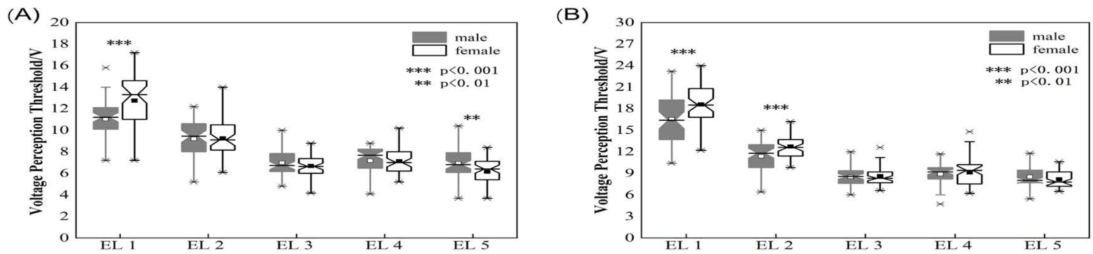


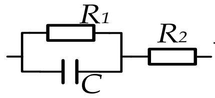
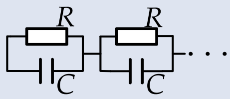
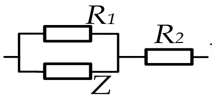
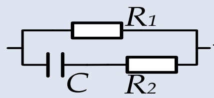
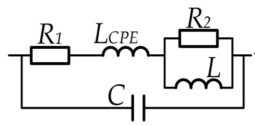
Figure 7. Comparison of the voltage perception threshold obtained from male (gray) and female (white) subjects using five electrodes of different sizes. (A) sensory threshold; (B) supra-sensory threshold. The box plot frames represent the 25th and 75th percentiles, while the whiskers indicate outliers. The small box within the plot represents the mean, the notch indicates the median, and * denotes significant differences between sexes ($^{***}p<0.001$, $^{**}p<0.01$).

threshold measurements, there was no significant difference in VPT values between the 16 cm^2 and 25 cm^2 electrodes. This indicates that the sensation of pain from electrical stimulation is determined by the voltage applied across the electrodes. In traditional electrical stimulation, a constant current source circuit is sued to achieve therapeutic effect by controlling parameters such as current intensity and frequency. However, individuals have different thresholds for current perception, meaning a unified current intensity standard may have different therapeutic effects for different individuals. As a new quantitative standard, VPT value can reflect the pain of electrical stimulation more simply and accurately, thereby achieving a consistent therapeutic effect across different individuals.

Discussion

This study aimed to investigate the maximum impedance of the human body during low-frequency envelope electrical stimulation and determine the appropriate voltage for the power amplifier circuit of electrical stimulation equipment. The measured impedance of the human body depends on the electrode size and the carrier frequency of the output. When using electrode EL1 for electrical stimulation, the impedance range is about 0.4 to 21.6 KΩ. According to **Figure 2a**, achieving a peak output current of 1mA requires a power supply of 216 V, while an output current of 2 mA would require 432 V, and so on. Such high voltage levels are impractical to achieve. The results of the electrical stimulation measurement show that the voltage perceived by the human body is much lower than the above-mentioned power supply value. As shown in **Figure 7**, when the

Table 6. Five common impedance models for human body

Definition	Model	References
Montague	 false	[31]
Tregear	 false	[31]
CPE_based	 false	[31]
Three-element	 false	[32]
FOIM	 false	[33]

Note: CPE, constant phase element; FOIM, fractional-order impedance models.

VPT exceeds about 10 V, the human body has an obvious sensation of electrical stimulation. The effect of varying electrode sizes is minimal; the sensory threshold ranges from 6-8 V for electrodes of 16 cm2, 25 cm^2 and 50 cm^2, while the supra-sensory threshold ranges from 9 to 11 V.

In addition to differentiating the test subjects by gender, we analyzed the factors affecting the impedance values of the human body using four parameters: different parts of the upper limbs, modulating wave frequency, carrier frequency, and electrode size. Only the electrode size and the output carrier frequency significantly affected the measured value. Notably, the impedance ranges measured by different sizes of electrodes varied significantly. The human impedance circuit model shown in **Figure 6** accurately represents the relationship between carrier frequency and human impedance. Skin impedance model was first developed by Cole, and since then, various impedance models have been proposed [30]. **Table 6** summarizes five common impedance models for the human body [31-33]. The differences in the studies may come from two main factors: 1) Most previous models used single waveform measurement, while this study focuses on the effect of human impedance on electrical stimulation parameters without comparing modulated waveforms to single waveform. 2) This study utilized five different electrode sizes for the measurement of human upper limb impedance, whereas most of the previous studies used a single electrode with unspecified dimensions and inconsistent electrode material. 3) In the previous research, when the frequency range used in the experiment was close to the peak frequency, the relationship between impedance and frequency was typically monotonous - either increasing steadily in the low-frequency range or decreasing steadily in the high-frequency range.

For the measurement of electrical stimulation perception, previous studies have suggested that the factors affecting CPT include current intensity, frequency and electrode size [26]. It is difficult to quantify so many variables. Compared with CPT, the VPT proposed in this study provides a better quantification of pain perception. As shown in **Figure 7**, the sensory threshold and supra-threshold values obtained from electrodes EL3 (electrode), EL4, and EL5 are comparable. The larger differences observed with EL1 and EL2 are due to their sensitivity to the measured values. As shown in **Figure 6**, within the 6-10 kHz range, the impedance values for EL1 and EL2 differ by approximately 1.5 KΩ and 500 Ω, respectively, despite the slower rate of change. This sensitivity is evident in that small currents and frequency changes cause significant voltage changes. While this study does not specifically explore the relationship between human impedance and electrical stimulation perception, the results suggest a positive correlation between these two physiological parameters.

One of the major limitations of this study is that the output intensity of electrical stimulation was not fully measured. Clinically, the output dose of electrical stimulation therapy is typically determined based on the patient's sensory or muscle response [29]. To avoid causing discomfort or pain to the subjects during the experiment, we did not use muscle response as a basis for measuring VPT. In addition, the measurement of human body impedance does not decisively determine the power required for the electrical stimulation power amplifier circuit. The final results show that the output current is not the primary factor affecting the perception of electrical stimulation. Instead, perception is solely related to the voltage across the stimulation site. Regardless of the current output, when the VPT exceeds 10 V, the human body experiences obvious electrical stimulation, which is far less than the product of the current and body impedance. Therefore, this finding suggests a potential shift towards using a constant voltage source circuit instead of constant current source circuit to achieve the desired stimulation effect in future electrical stimulation device. Additionally, we considered whether the inductive characteristic of human upper limb impedance is related to the material of hydrogel electrodes. To verify this, we conducted preliminary experiments with sponge electrodes and observed the same trends, thereby excluding the hydrogel material's influence on impedance values.

In summary, the purpose of this study was to measure the range of human impedance and electrical stimulation sensation to determine the appropriate power supply and output intensity for electrical stimulation devices. The results show that these two parameters should be set based solely on electrical stimulation sensation. A new impedance model proposed in this study accurately fits the measured values. For measuring electrical stimulus perception, this study proposes a new method using the VPT. focuses exclusively on the voltages across the stimulation site, excluding factors such as current intensity and frequency, which significantly simplifies the measurement criteria. Compared to CPT, VPT offers a more precise method for electrical stimulation therapy.

In the next step of the research, the device developed in this study will be refined by adding the voltage detection function to monitor the voltage across the patient's stimulation site in real time. Additionally, the method of measuring the upper limb impedance in this study relies on Ohm's law and collects the voltage at both ends of the electrode, including a part of electrode contact impedance. Future improvements in the measurement method will aim to exclude the influence of electrode contact impedance to provide more accurate results.

Author contributions: Renling Zou and Yuhao Liu are responsible for the overall conception and design of the paper, and put forward the core problems and experimental scheme of the research. Leading the data analysis and the interpretation of experimental results, and writing the main part of the paper, including the introduction and discussion. Yicaiwu, Liang Zhao and Ji Gaodai are responsible for the implementation of the experiment and data collection, participated in the experiment design, and provided suggestions and modifications for some chapters of the paper. Xuezhi Yin assists in the selection and maintenance of equipment and technical platforms used in the research process. Xiufang Hu is responsible for providing materials and financial support for the experiment. This passage helps me translate the following.

References

[1] Gill ML, Grahn PJ, Calvert JS, et al. Neuromodulation of lumbosacral spinal networks enables independent stepping after complete paraplegia. Nat Med 2018;24(11):1677-1682.

[2] Angeli CA, Boakye M, Morton RA, et al. Recovery of Over-Ground Walking after Chronic Motor Complete Spinal Cord Injury. N Engl J Med 2018;379(13):1244-1250.

[3] Rodrigues JC, Avila MA, Driusso P. Transcutaneous electrical nerve stimulation for women with primary dysmenorrhea: Study protocol for a randomized controlled clinical trial with economic evaluation. PLoS One 2021;16(5):e0250111.

[4] Zhou YX, Wang HP, Bao XL, et al. A frequency and pulse-width co-modulation strategy for transcutaneous neuromuscular electrical stimulation based on sEMG time-domain features. J Neural Eng 2016;13(1):016004.

[5] Sanchez B, Pacheck A, Rutkove SB. Guidelines to electrode positioning for human and animal electrical impedance myography research. Sci Rep 2016;6:32615.

[6] Son C, Kim S, Kim S-j, et al. Detection of muscle activation through multi-electrode sensing using electrical stimulation. Sens. Actuators, A 2018;275:19-28.

[7] Sun B, Baidillah MR, Darma PN, et al. Evaluation of the effectiveness of electrical muscle stimulation on human calf muscles via

frequency difference electrical impedance tomography. Physiol Meas 2021;42(3).

[8] Sanchez B, Aroul ALP, Bartolome E, et al. Propagation of Measurement Errors Through Body Composition Equations for Body Impedance Analysis. IEEE Trans Instrum Meas 2014;63(6):1535-1544.

[9] Chinen K, Kinjo I, Zamami A, et al. New equivalent-electrical circuit model and a practical measurement method for human body impedance. Biomed Mater Eng 2015;26 Suppl 1:S779-S786.

[10] Freeborn TJ, Elwakil AS, Maundy B. Compact Wide Frequency Range Fractional-Order Models of Human Body Impedance against Contact Currents. Math Probl Eng 2016;2016(1):4967937.

[11] Demir U, Kocaoğlu S, Akdoğan E. Human impedance parameter estimation using artificial neural network for modelling physiotherapist motion. Biocybern Biomed Eng 2016;36(2):318-326.

[12] Shiffman C. Scaling and the frequency dependence of Nyquist plot maxima of the electrical impedance of the human thigh. Physiol Meas 2017;38(12):2203-2221.

[13] Qiu S, Feng J, Xu R, et al. A Stimulus Artifact Removal Technique for SEMG Signal Processing During Functional Electrical Stimulation. IEEE Trans Biomed Eng 2015;62(8):1959-1968.

[14] Ding JH, Chen ZH, Li YR. Development of functional electrical stimulator with real-time adjustable stimulation parameters. J Fuzhou Univ (Nat Sci Ed) 2019;47(06):759-764.

[15] Yin Q, Li X, Liu Q. [Design of functional array electrode stimulation system with surface electromyography feedback]. Sheng Wu Yi Xue Gong Cheng Xue Za Zhi 2020;37(6):1045-1055.

[16] Zhou Y, Fang Y, Gui K, et al. sEMG Bias-Driven Functional Electrical Stimulation System for Upper-Limb Stroke Rehabilitation. IEEE Sens J 2018;18(16):6812-6821.

[17] van der Lely S, Liechti MD, Bachmann LM, et al. Quantitative electrical pain threshold assessment in the lower urinary tract. Neurourol Urodyn 2020;39(1):420-431.

[18] Seno SI, Shimazu H, Kogure E, et al. Factors Affecting and Adjustments for Sex Differences in Current Perception Threshold With Transcutaneous Electrical Stimulation in Healthy Subjects. Neuromodulation 2019;22(5):573-579.

[19] Oh KJ, Kim SH, Lee YH, et al. Pain-related evoked potential in healthy adults. Ann Rehabil Med 2015;39(1):108-115.

[20] Ji Ri Mu Tu Y, Yin Y, Zhu T. [Determination of electrical stimulation pain perception and analysis of the related factors]. Sheng Wu Yi Xue Gong Cheng Xue Za Zhi 2013;30(6):1200-1203, 1208.

[21] Chang W, Xu W, Hu R, et al. Current Perception Threshold Testing in Pharyngeal Paresthesia Patients with Depression or Anxiety. Neuropsychiatr Dis Treat 2020;16:1023-1029.

[22] Suzuki Y, Muramatsu K, Maruo K, et al. Pain thresholds are unaffected by age in a Japanese population. Muscle Nerve 2020;61(5):653-656.

[23] Solomons CD, Slovak M, Heller B, et al. Reducing the sensation of electrical stimulation with dry electrodes by using an array of constant current sources. Med Eng Phys 2018;51:91-95.

[24] Chan SCC, Peng J, Chan CCH. Reliability of measurements for sub-painful and painful perception on artificial electrical stimulations. Int J Psychophysiol 2018;123:35-41.

[25] Furuse N, Kimoto S, Nakashima Y, et al. Verification of the reliability of current perception threshold and pain threshold testing by application of an electrical current stimulus to mandibular mucosa in young adults. J Oral Rehabil 2019;46(6):556-562.

[26] Hiremath SV, Tyler-Kabara EC, Wheeler JJ, et al. Human perception of electrical stimulation on the surface of somatosensory cortex. PLoS One 2017;12(5):e0176020.

[27] Nitsche MA, Paulus W. Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. J Physiol 2000;527 Pt 3(Pt 3):633-639.

[28] Reilly JP. Applied bioelectricity: from electrical stimulation to electropathology. USA. Springer New York, NY, 2012.

[29] Shen Y and Zhang ZQ. Clinical work guidelines for rehabilitation therapists - physical factor therapy techniques. Beijing. People’s Health Publishing House, 2019. p. 171-172.

[30] Cole KS and Curtis HJ. Electric impedance of nerve and muscle. Cold Spring Harb Symp Quant Biol 1936;4:73-89.

[31] Bora DJ, Dasgupta R. Estimation of skin impedance models with experimental data and a proposed model for human skin impedance. IET Syst Biol 2020;14(5):230-240.

[32] Rutkove SB. Electrical impedance myography: Background, current state, and future directions. Muscle Nerve 2009;40(6):936-946.

[33] Ghita M, Neckebroek M, Juchem J, et al. Bioimpedance Sensor and Methodology for Acute Pain Monitoring. Sensors (Basel) 2020;20(23):6765.

Progress in Medical Devices
ISSN: 2957-5478

Research Article https://doi.org/10.61189/961030gznunx

Online recognition method for walking patterns of intelligent knee prostheses based on CNN-LSTM algorithm

Yibin Zhang[1], Yan Wang[1], Hongliu Yu[2]

[1]School of Medical Devices, Shanghai University of Medicine and Health Sciences, Shanghai 201318, China. [2]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

Corresponding author: Hongliu Yu.

Acknowledgement: This work was supported by National Natural Science Foundation of China (No.62073224).

Declaration of conflict of interest: None.

Received June 21, 2024; Accepted November 20, 2024; Published December 31, 2024

Highlights

- In prosthetics, using AI algorithms to identify the fused sensor data as known walking patterns has extremely strong expandability. Moreover, as the learning data continues to expand, the robustness of the model itself also increases accordingly.
- There are numerous AI algorithms currently available. The effective utilization of algorithm combination techniques to learn from each other's strengths can significantly improve the accuracy of identification. The combined model of convolutional neural networks (CNN) and bidirectional long short term memory (LSTM) attempted in this paper has witnessed a significant improvement in its comprehensive recognition rate.
- In the practical application of prosthetics, the real-time performance during the mode switching transition period is particularly important as it can reflect the flexibility of the prosthetics. In this paper, the algorithm optimized by the AI model has controlled the delay rate within one gait cycle, greatly enhancing the safety and reliability of prosthetics in actual use.

Abstract

To enhance the adaptive learning, self-organization, and fault tolerance capabilities of gait pattern recognition in intelligent knee prostheses, an online walking pattern recognition method based on the convolutional neural networks (CNN)-long short term memory (LSTM) model is proposed. Five test subjects wore the intelligent knee prostheses and performed four walking modes: level walking, uphill walking, downhill walking, and stair descent. The preprocessed gait data were fed into four neural network models: CNN, LSTM, CNN-LSTM, and CNN-bidirectional LSTM. Through hyperparameter tuning, the recognition accuracy of these models was compared. Real-time indicator, gait recognition delay, was also measured. Experimental results showed each model had its strengths and weaknesses. Overall, the CNN-LSTM model achieved the best recognition performance with accuracy rates of: level walking 89%±2.5%, uphill 72.8%±3.2%, downhill 71%±3.2%, and stair descent 96%±2.5%. When switching from level walking to downhill, gait recognition delay was 51.7%±15.6%, and vice versa it was 75.8%±11.5%; when switching from level walking to stair descent, gait recognition delay was 47.1%±17.1%, and vice versa it was 38.6%±10.5%. In summary, the application of the CNN-LSTM model for walking pattern recognition in unilateral intelligent knee prostheses is feasible, with accuracy and real-time performance meeting the control requirements of the prostheses.

Keywords: Neural network, walking pattern, gait recognition, recognition delay rate

Introduction

Traditional electronically controlled knee prostheses require the user to pre-select the current walking mode, such as level walking, uphill, downhill, or stairs. Once the walking mode

Address correspondence to: Hongliu Yu, School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai city Jungong road 516, Shanghai 200093, China. E-mail: yhl98@hotmail.com.

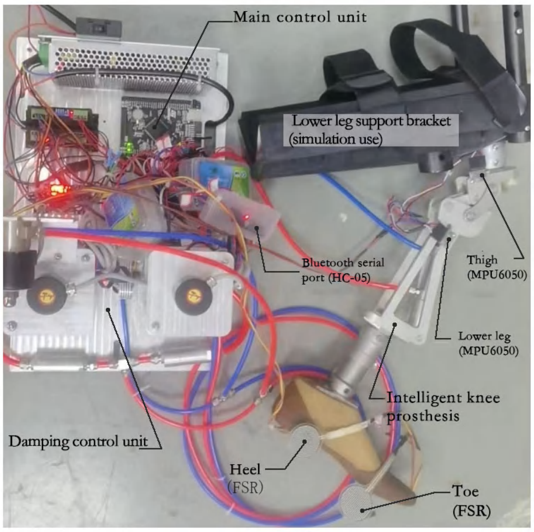


Figure 1. Intelligent above-knee prosthetic prototype.

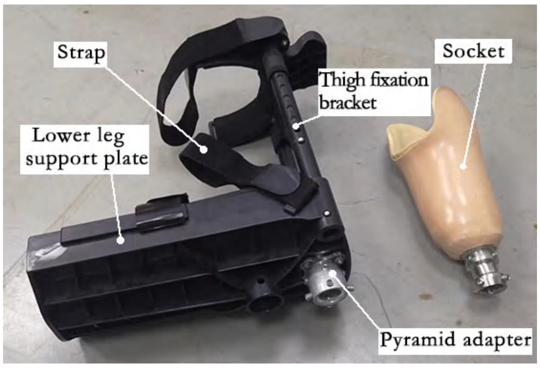


Figure 2. Alternative bracket for prosthetic socket.

is set, the prosthesis follows this predefined mode, and switching modes mid-walk can lead to control failure and an increased risk of falls. Currently, some advanced prostheses worldwide are transitioning from basic electronic control to intelligent systems. For example, the Genium X3 by Ottobock in Germany can intelligently and dynamically respond to various walking modes such as slow walking, fast walking, stair climbing, slope walking, and even obstacle crossing [1]. This indicates that intelligence is the future direction of knee prosthesis development.

Walking modes are classified based on daily human activities, which can be identified by collecting motion information using advanced sensors and extracting multiple features for classification. With the development of artificial intelligence and big data analysis, these technologies will undoubtedly excel in areas such as human-computer interaction, behavior analysis, rehabilitation, exoskeletons, and prosthetics. Currently, there are two mainstream methods for human activity recognition: image-based and sensor-based [2]. The former relies on fixed cameras to track target subjects and uses frame sequences to determine walking modes. However, this method has several limitations, such as being susceptible to environmental factors (e.g., poor lighting, background noise), complex image analysis algorithms, difficulty in long-term cross-environment monitoring, and the large, non-wearable nature of the equipment. In contrast, sensor-based methods benefit from advancements in microelectronics, with smaller and more affordable micro inertial measurement unit (MIMU) sensors entering the field of human activity recognition [3]. Due to their compact size and wearability, these sensors have garnered significant attention from researchers and have been widely applied in various fields [4-6].

This paper aims to use a multi-sensor fusion framework (combining MIMU sensors and thin-film pressure sensors) to collect multi-feature variables during the walking process of knee prostheses. Various neural network algorithms are utilized to classify and recognize the walking modes (level walking, uphill, downhill, and stairs) of the intelligent knee prostheses online. The goal is to achieve higher accuracy and stronger real-time performance, enabling intelligent multi-scenario applications for knee prostheses.

Intelligent knee prosthesis prototype

The intelligent knee prosthesis prototype shown in **Figure 1** was developed by the Rehabilitation Engineering and Technology Research Institute of the University of Shanghai for Science and Technology. The prototype mainly consists of a knee prosthesis and a damping control unit. It uses multiple sets of inertial sensors and foot pressure sensors to detect the prosthesis's motion state, enabling intelligent recognition of various walking modes (e.g., level walking, uphill, downhill, stairs). The knee prosthesis achieves swing phase speed-following control during level walking through a particle swarm-optimized backpropagation algorithm, as detailed in recent publications [7].

This prototype is still in a development and debugging stage. For safety considerations, the prototype intelligent knee prostheses were tested in healthy individuals simulating above-knee amputees. Since healthy individuals have intact knees and lower legs, a special support bracket is designed to bind the lower leg and knee, eliminating their respective movement func-

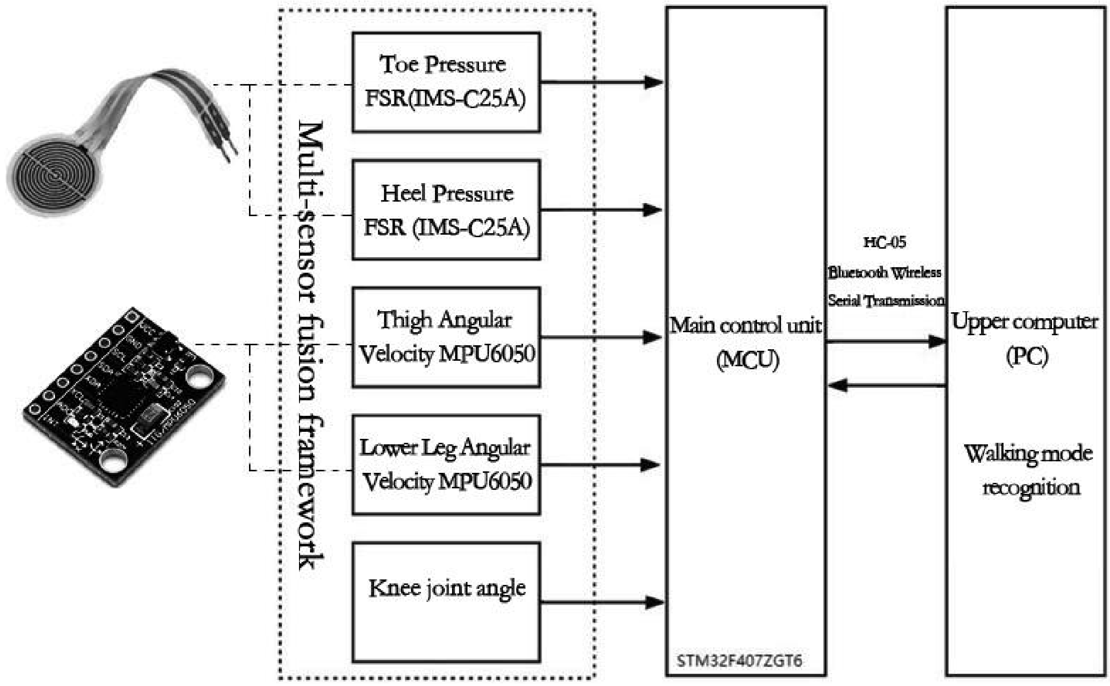


Figure 3. Main control unit framework.

Table 1. Information of test subjects

No.	Gender	Age	Height (m)	Weight (Kg)	Thigh length (m)	Lower leg length (m)	Stride length (m)
1	Male	53	1.70	70	0.42	0.48	0.63
2	Male	43	1.65	61	0.40	0.47	0.61
3	Male	29	1.78	75	0.44	0.51	0.66
4	Female	41	1.52	55	0.37	0.43	0.56
5	Female	55	1.60	57	0.39	0.46	0.59
Mean		44.2	1.65	63.6	0.404	0.47	0.61
Standard Deviation		10.45	0.098	8.59	0.027	0.029	0.038

tions. The structure of the lower leg support bracket is shown in **Figure 2**. The lower end of the bracket is equipped with a pyramidal interface, which connects to the knee prosthesis. The front and rear of the bracket have straps that secure it to the thigh, forming a unified structure that acts as a thigh prosthetic socket.

As shown in **Figure 3**, based on the characteristics of human walking and different road conditions, the selected lower limb kinematic parameters include thigh swing angular velocity, lower leg swing angular velocity, knee joint angle, heel pressure, and toe pressure. The thigh and lower leg angular velocities were collected using MIMU sensors (model MPU6050). The knee joint angle was obtained by integrating the differences between the angular velocity sensors of the thigh and lower leg. The foot pressure information was collected by thin-film pressure sensors Force Sensing Resistor [8, 9]. The main control unit operates using an STM32F407 embedded microcontroller [10, 11], and communication between the main control unit and the upper computer is achieved via a wireless Bluetooth serial port (model HC-05). The upper computer processes the data using a neural network model algorithm and returns the final recognition results. The hardware configuration of the upper computer includes a CPU (Intel i3-4160®), GPU (GeForce GT430®), 8GB RAM, and runs on a 64-bit Ubuntu 18.04 operating system with a Python-Anaconda 3.6, TensorFlow 1.7.0 neural network learning environment.

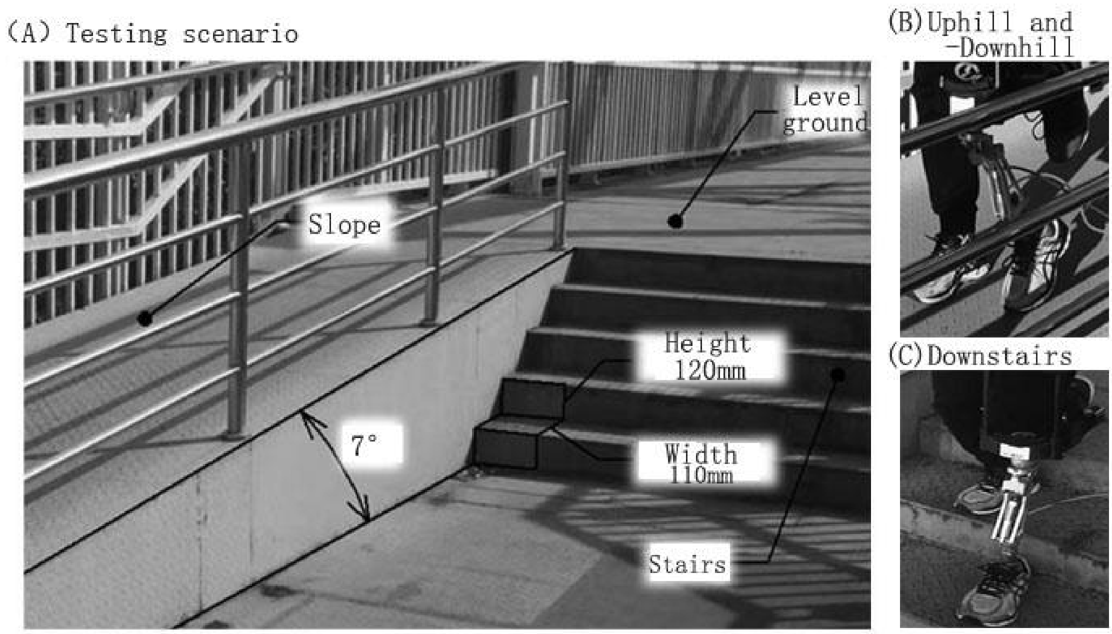


Figure 4. Gait testing in different walking modes. (A) Testing scenario; (B) Uphill and Downhill; (C) Downstairs.

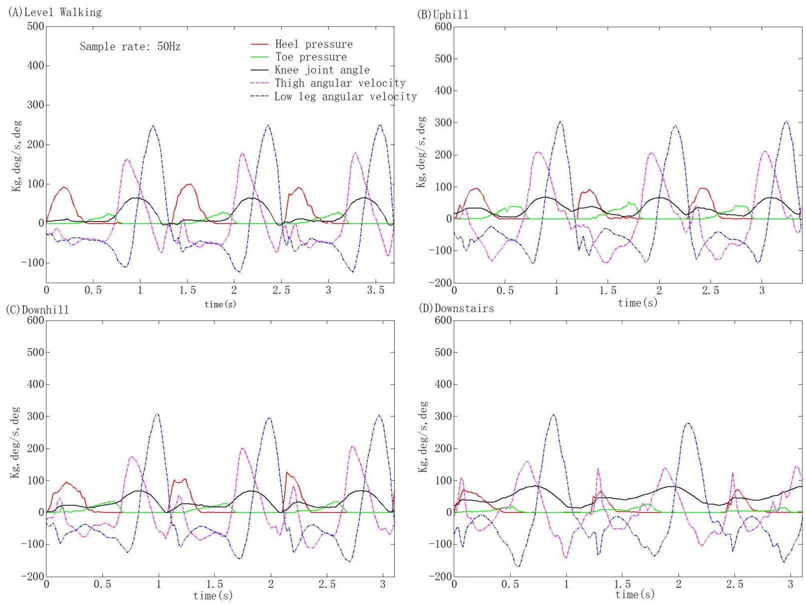


Figure 5. Gait data of test subjects. (A) Level Walking; (B) Uphill; (C) Downhill; (D) Downstairs.

Experimental methods

Experimental subjects

The gait data collection involved 5 healthy volunteers whose specific physical parameters are shown in **Table 1**. All participants were informed about the details of the experiment and the intended use of the collected data, and they signed informed consent forms.

Experimental conditions and procedure

1. The length of the lower leg section of the knee prosthesis was adjusted according to the height and leg length of the test subjects to ensure that both lower limbs were of equal height

2. Since the test subjects were healthy individuals (not above-knee amputees), they underwent a period of adaptive training to get used to wearing the prosthesis. However, during actual testing, subjects were allowed to hold onto handrails to prevent accidental falls.

3. The walking sequence for the test subjects

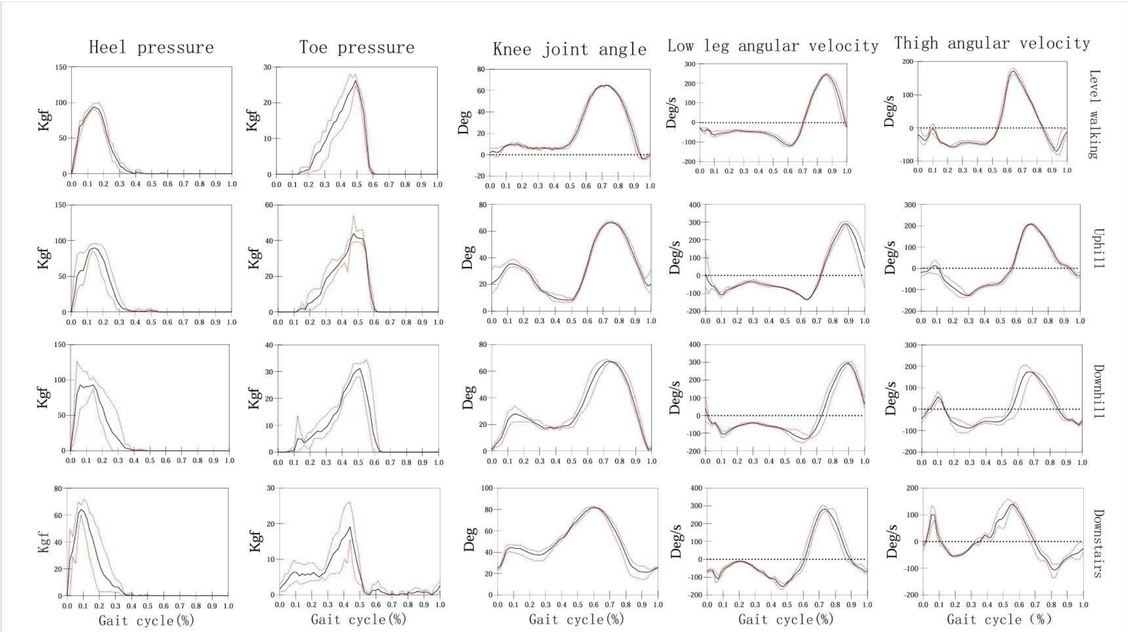


Figure 6. Statistical analysis of gait data.

Table 2 Gait Data Samples of Test Subject (#1)

Samples (N)	Toe force (kgf)	Heel force (kgf)	Knee angle (deg)	Thigh velocity (deg/s)	Leg velocity (deg/s)	Label
1	0	0	6	-25	-26	WE
596	0	0	18	3	-42	UR
929	0	0	1	-22	-32	DR
1,232	0	0	68	0	-29	DS

Note: WE, walking on even ground; UR, uphill walking; DR, downhill walking; DS, stair descent.

was level walking, uphill and downhill walking (**Figure 4B**), and stair descent (**Figure 4C**). To ensure an adequate sample size, each walking mode required at least 30 steps.

4. There were no specific requirements for the walking speed; subjects could walk at their preferred pace.

5. The testing environment is shown in **Figure 4A**, with a slope angle of 7°. The stairs had a width of 110 mm and a height of 120 mm, with handrails on one side.

6. There were no restrictions on which leg the prosthesis was worn on, so either the left or right leg could be used.

Data collection and analysis

Figure 5 shows the gait data from the prosthetic side of the test subjects, collected at a frequency of 50 Hz. It can be seen that different walking modes exhibit distinct gait characteristics, including the swing angular velocity of the thigh and lower leg, foot pressure, and the flexion-extension angle of the knee joint [14, 15]. All gait data from the test subjects were first normalized and then imported into GraphPad Prism 8.4 for statistical analysis. The mean and variance distributions are shown in **Figure 6**.

Neural network models

In this study, we utilized various neural network models to classify and recognize the four different walking modes: convolutional neural networks (CNN), long short-term memory (LSTM), CNN-LSTM, and Convolutional Neural Network - Bidirectional Long Short-Term Memory (CNN-BiLSTM) hybrid [16-19]. The classification performance of these models was compared through iterative learning.

Training set and test set

During the learning process of the neural network model, 80% of the data from each test subject in every walking pattern is used as the training set, while 20% is reserved as the test set. See **Table 2**.

Table 3. CNN-LSTM related parameters

Category	Subcategory	Parameter	Optimized value
Model	LSTM	Layers	2
	Conv1D	Units	64
		Kernel size	1×3
		Filters	64
	MaxPooling1D	Pool size	1
	Hyper	Batch size	25
	parameter	Epochs	500
		Dropout	0.2
		Learning rates	0.001

Note: CNN, convolutional neural networks; LSTM, long short-term memory.

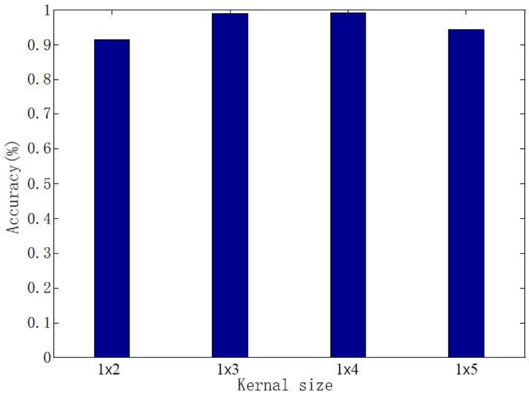


Figure 7. The Impact of CNN Convolution Kernel Size on Gait Pattern Recognition. CNN, convolutional neural networks

Parameter settings

Hyperparameter optimization included the number of LSTM layers and nodes, the size of CNN convolution kernels, dropout rates, and learning rates [20]. The walking modes were recognized using four neural network models: CNN, LSTM, CNN-LSTM, and CNN-BiLSTM, with the Mini SGD optimizer [21]. To balance the real-time performance and accuracy of recognition, a CNN convolution kernel size of 1×3 was chosen, and the test results are shown in **Figure 7**. **Table 3** presents the final optimized values for the convolution kernel size, the number of LSTM layers, and the learning rates for the CNN-LSTM model.

Model evaluation results

To evaluate the performance of the walking mode recognition algorithm, accuracy was used as the performance metric [22]. The formula is as follows:

$$Acc = \frac{TP+TN}{TP+TN+FP+FN}$$

where true positive is the correctly identified positive samples, true negative is the correctly identified negative samples, false positive is the incorrectly identified positive samples, and false negative is the incorrectly identified negative samples. Using this formula, the confusion matrices for the four neural network algorithms in recognizing the four different walking modes were plotted, as shown in **Figure 8** [23].

The recognition accuracies for the four different walking modes of all test subjects were imported into GraphPad Prism 8.4 for statistical analysis. The final statistical results are shown in **Figure 9**.

Testing the four neural network models revealed that the CNN-BiLSTM model had a recognition rate of 94.8%±1.6% for level walking and 98.6%±1.3% for stair descent, but only 37%±4.2% for downhill walking, leading to its exclusion. The remaining three neural network models have their own advantages and disadvantages:

- The CNN model achieved a recognition accuracy of 99%±1.4% for stair descent but only 78.6%±3.6% for level walking and 47%±6.8% for downhill walking.

- The LSTM model improved the downhill recognition rate to 60.2%±1.9% compared to the CNN model.

- The CNN-LSTM hybrid model further increased the downhill recognition rate to 71%±3.2%, although the recognition rates for other walking modes decreased slightly.

Overall, the CNN-LSTM model was determined to be the most advantageous compared to the other models.

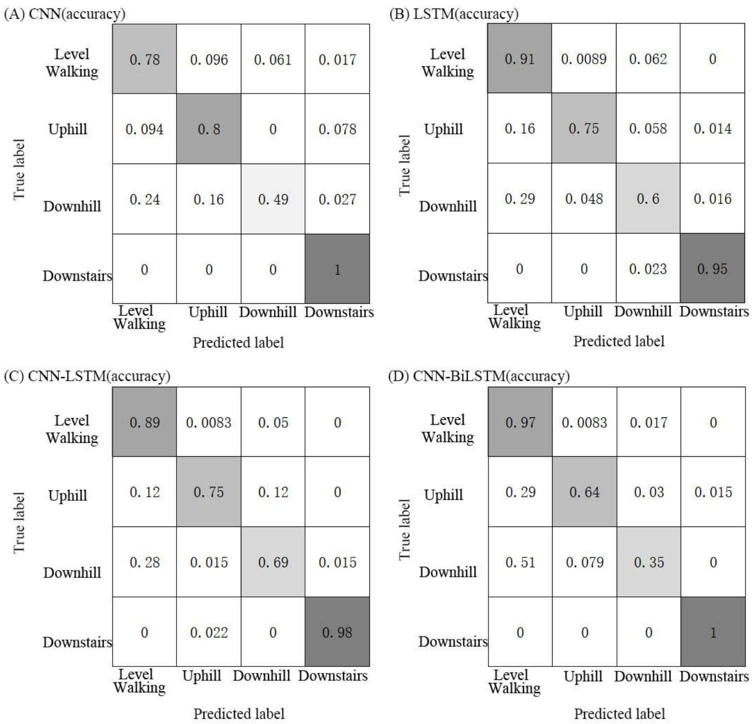


Figure 8. Accuracy evaluation of neural network models. A. CNN (accuracy); B. LSTM (accuracy); C. CNN-LSTM (accuracy); D. CNN-BiLSTM (accuracy). CNN: convolutional neural networks; LSTM: long short term memory; CNN-LSTM: convolutional neural networks-long short term memory; CNN-BiLSTM: convolutional neural networks-bidirectional long short term memory.

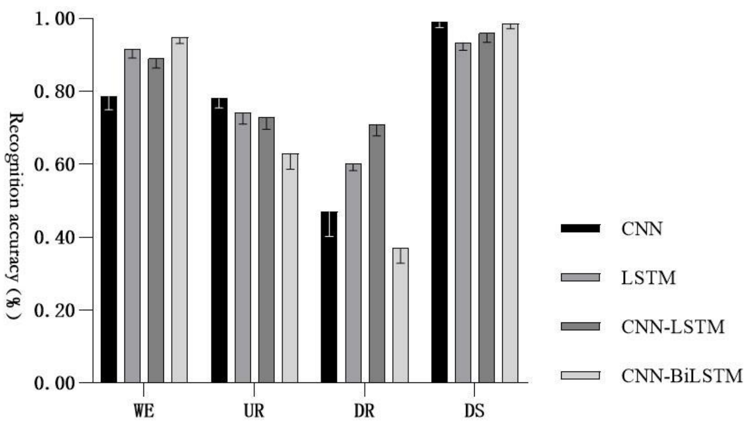


Figure 9. Statistical results of recognition accuracy. WE, walking on even ground; UR, uphill walking; DR, downhill walking; DS, stair descent.

Table 4. Delay Rate Test Results

Walking mode transition state	GRD
WEDR	51.7%15.6%
DRWE	75.8%11.5%
WEDS	47.1%17.1%
DSWE	38.6%10.5%

Note: WE, walking on even ground; DR, downhill walking; DS, stair descent; GRD, gait recognition delay.

Calculation of online recognition delay rate

The CNN-LSTM model, trained offline, has preliminary walking mode recognition capabilities. When one walking mode transitions to another, the online gait recognition system can promptly send the recognition results to the main control unit, identifying the current walking mode in real-time.

The formula for calculating the recognition delay rate is as follows:

$$GRD = \frac{T_s}{GC} \times 100\%$$

where GRD represents the recognition delay rate, GC represents the gait cycle, and Ts represents the system's recognition time, i.e., the time from when the gait data are sent to when the recognition result is returned [24].

There are 12 possible combinations for the transitions between the four walking modes: WE-UR, WE-DR, WE-DS, UR-DR, UR-DS, DR-DS. Considering the daily usage scenarios of the knee prosthesis prototype, this experiment only measured the delay rates for transitions between level walking (WE) and downhill (DR), and between level walking (WE) and stair descent (DS). The results, shown in **Table 4**, indicate that the recognition delay meets the usage requirements, allowing mode transitions to be recognized within one gait cycle. Among the four transitions, the delay is greatest when switching from downhill to level walking. The system can only recognize the mode transition near the end of the downhill step, further indicating that these two walking modes have less distinct features and require more classification time. In contrast, the delay is smallest when switching from stair descent to level walking, suggesting that these two walking modes have more distinct features and are easier to classify.

Conclusion

This paper proposes an exploratory solution to the problem of real-time walking mode recognition faced in fields such as prosthetics, exoskeletons, rehabilitation engineering, and behavior analysis [25]. By utilizing a unilateral knee prosthesis walking mode recognition system constructed with MIMU and Force Sensing Resistor sensors, we successfully collected gait data for level walking, uphill, downhill, and stair descent. The gait data for these four walking modes were classified using CNN, LSTM, CNN-LSTM, and CNN-BiLSTM models. Comprehensive evaluation confirmed that the CNN-LSTM model provided the best recognition results, with the highest relative accuracy. To verify the real-time performance of online recognition, we calculated the recognition delay rates for transitions between level walking and stair descent, and between level walking and downhill. The results showed that the delay time was less than one gait cycle, meeting the real-time requirements for knee prosthesis control.

Currently, the walking mode recognition system still uses a master-slave architecture, where unilateral lower limb gait data are sent via a wireless Bluetooth serial port to a more powerful PC for processing by neural network models. The processed classification results are then returned to the main control unit via the serial port. Future research will involve porting the code to the main control unit, eliminating the master-slave architecture, and facilitating the independent application of the prosthesis [26]. Additionally, the experiment did not include stair ascent testing because the current knee prosthesis prototype is still a passive type (unable to provide additional assistive power), making it difficult to achieve such a walking mode [27].

Author contributions: Yibin Zhang was responsible for the hardware design of the prosthetic prototype and the construction of the experimental scenario. Yan Wang undertook the hardware design of the data acquisition system and also carried out the gait data collection and organization of the prosthetic during the testing process. Meanwhile, Hongliu Yu focused on the algorithm optimization of the road condition recognition model.

References

[1] Sun JY, Yu HL, Wang XM, et al. Design and Implementation of Control System for Intelligent Hydraulic Prosthetic Knee Joint. Software Guide 2020;019(007):89-93.

[2] Chen DL, Liu XH. Human Activity Recognition Based on Adaptive Sample Entropy with Wearable Sensors. J Wuhan Univ Technol 2022;44(08):91-96.

[3] Lu YL, Chen YW, Li Y, et al. Adaptive Gait Detection Algorithm Based on MIMU. Piezoelectr Acoustoopt 2018;40(5):768-771.

[4] Palla S, Sahu G, Parida P. A contemporary Survey on Human Gait Recognition. 2020;15(3):94-106.

[5] Nweke HF, Teh YW, Al-garadi MA, et al. Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges. Expert Systems with Applications 2018;105:233-261.

[6] Gupta H, Anil A, Gupta R. On the combined use of Electromyogram and Accelerometer in Lower Limb Motion Recognition. 2018 IEEE 8th International Advance Computing Conference (IACC) 2018;240-245.

[7] Zhang YB, Lv J, Yu HL. Research on an Intelligent Prosthetic Knee Joint and Its Control Algorithm. Mod Instrum Med Treat 2022;28(6):19-27.

[8] Zhu L, Xing X, Meng ZX, et al. Wood Planarity Measurement Method Based on Flexible Film Sensors. Sens Microsyst 2021;40(11):143-145+149.

[9] Zhao YQ, Zhao CL, Wang K. Personalized Customization Method for Orthotic Insoles Based on Film Pressure Sensors. Comput Meas Control 2020;28(11):265-269.

[10] Zhang KZ, Hu M. Design of a Biomimetic Hexapod Obstacle-Crossing Robot System Based on STM32. Mod. Manuf Technol Equip 2022;58(07):90-94.

[11] Li XY, Han FR, Liu ZX, et al. Design of Ankle Joint Gait Collection and Processing System Based on STM32. Exp Sci Technol 2020;18(06):40-45.

[12] Wang C, Fang Z, Yu T, et al. A Method for Predicting Walking Frequency on Flat Ground Based on Kalman Filter Model. J Harbin Eng Univ 2015;36(08):1109-1113.

[13] Huang CZ, Chen X, Zhou YM. Design of a Tactile Feedback Intelligent Glove Based on Kalman Filter. Guangdong Sci Technol 2022;31(04):79-81.

[14] Zhang LX, Wang LJ, Wang FL, et al. A Method for Measuring Human Gait Trajectories. Meas Control Technol 2009;28(02):24-27.

[15] Zhang RH, Jin DW, Zhang JC, et al. Study on Normal Gait Characteristics on Different Road Conditions. J Tsinghua Univ, Sci Technol 2000;(08):77-80.

[16] Zhang x, Zhao D, Tao SH. A Method for Lower Limb Action Pattern Recognition Based on Convolutional Neural Networks. J Hebei Univ Sci Technol 2022;43(04):347-354.

[17] Yu JJ, Li C, Zuo GY, et al. A Method for Generating Biped Robot Cyclic Gait Based on

LSTM Neural Network. J Beijing Univ Technol 2020;46(12):1335-1344.

[18] Qi YJ, Kong YP, Wang JJ, et al. A Gait Recognition Method Combining LSTM and CNN. J Xidian Univ 2021;48(05):78-85.

[19] Sun YX, Chen JB, Wu DH. Human Activity Recognition Method Based on Convolutional Neural Network and Bi-Directional Long Short-Term Memory Network. Sci Technol Eng 2022;22(04):1517-1525.

[20] Huang Y, Gu CG, Yang HJ. Strategy for Deleting Garbage Neurons in Neural Network Hyperparameter Optimization. Acta Phys Sin 2022;71(16):137-145.

[21] Li M, Lai GH, Chang YM, et al. Performance Analysis of Different Optimizers in Deep Learning Algorithms. Inf. Technol Informatization 2022;(3):206-209.

[22] Wu XF, Duan R, Yang LW, et al. Analysis of Evaluation Indicators of Scientific Journals Based on Keras Neural Network Model. Tianjin Sci Technol 2022;49(08):102-106.

[23] Yu Y, Yang TT, Yang BY. Confusion Matrix Classification Performance Evaluation and Python Implementation. Mod Comput 2021;(20):70-73+79.

[24] Wang KJ, Liu LL, Ding XN, et al. Gait Cycle Detection Method Based on Convolutional Neural Network. J Harbin Eng Univer 2021;42(05):656-663.

[25] Sheng M, Liu SQ, Wang J, et al. Real-Time Recognition of Motion Intention of Intelligent Lower Limb Prosthesis Based on Improved Template Matching. Control Decis 2020;35(09):2153-2161.

[26] Zhang PF, Ye ZJ, Yang JL, et al. Design of CNN-Based Water Meter Pointer Reading Recognition and STM32 Implementation Scheme. Electron Meas Technol 2021;44(23):61-67.

[27] Zhang N, Li J. Research Progress on Powered Intelligent Prosthetic Knee Joint. Beijing Biomed Eng 2018;37(04):427-432.

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Review Article https://doi.org/10.61189/521889ygxkmc

Overview of the current development in Visual-Inertial Systems

Mingxia Wei, Qingyun Meng

School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

Corresponding author: Qingyun Meng.

Declaration of conflict of interest: None.

Received May 15, 2024; Accepted June 25, 2024; Published December 31, 2024

Highlights

- A comprehensive overview of Visual-Inertial Navigation Systems.
- Exploration of key technologies, including image processing methods for visual odometry.

Abstract

The Visual-inertial navigation system (VINS) integrates visual and inertial sensors, providing advanced navigation capabilities. Recent improvements in these sensors have made VINS widely applicable in areas such as robotic navigation and autonomous driving, due to its complementary functionality and decreasing sensor costs. This paper reviews the latest developments in VINS, introducing visual simultaneous localization and mapping and its role in VINS. Key technologies, including direct and indirect methods for image processing in visual odometry and Inertial Measurement Unit preintegration for robot motion estimation, are discussed. Both filter-based and optimization-based state estimation methods in VINS are examined. Filter-based methods, like the Extended Kalman Filter, offer real-time state updates for attitude tracking and map construction, while optimization-based methods focus on minimizing reprojection or other error metrics to improve localization accuracy and robustness. The application of dynamic simultaneous localization and mapping, which addresses dynamic objects in complex environments, is also explored. This paper summarizes current research challenges and proposes future directions in the field of dynamic simultaneous localization and mapping.

Keywords: Visual-Inertial Navigation System, visual simultaneous localization and mapping, extended Kalman Filter, dynamic simultaneous localization and mapping, semantic methods

Introduction

Simultaneous localization and mapping (SLAM) is a computational process of simultaneously constructing or updating maps of unknown environments while tracking an agent's position within them. It is a key concept in robotics and autonomous driving, with a development history spanning several decades. Early research focused on fundamental SLAM, such as determining a robot's position and mapping the environment without prior knowledge, primarily relying on probabilistic methods and estimation theory [1].

As SLAM research evolved towards more practical solutions—particularly in handling large-scale environments and managing the increasing computational complexity of accumulating data—sensor fusion approaches gained prominence, with Visual-inertial navigation system (VINS) attracting significant attention [2]. VINS employs cameras and inertial measurement units (IMUs) as sensors to measure displacement and attitude changes. Visual SLAM (VSLAM), an integrated technology combining localization and mapping, relies on extracting feature information or pixel grayscale from image sequences captured by visual sensors to infer the camera's pose. VSLAM provides rich environmental information and avoids drift when static, delivering strong localization and mapping performance in static environments. In contrast, IMUs, independent of external sig-

Address correspondence to: Qingyun Meng, School of Health Science and Engineering, University of Shanghai for Science and Technology, Yangpu District, Shanghai 200093, China. Tel: +86-13761813609. E-mail: mengqy@sumhs.edu.cn.

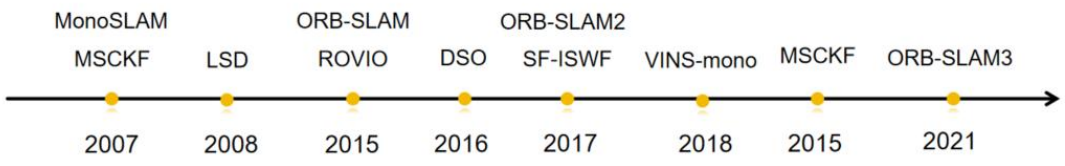


Figure 1. Classic filtering based and optimization based VINS algorithms. VINS, Visual-inertial navigation system.

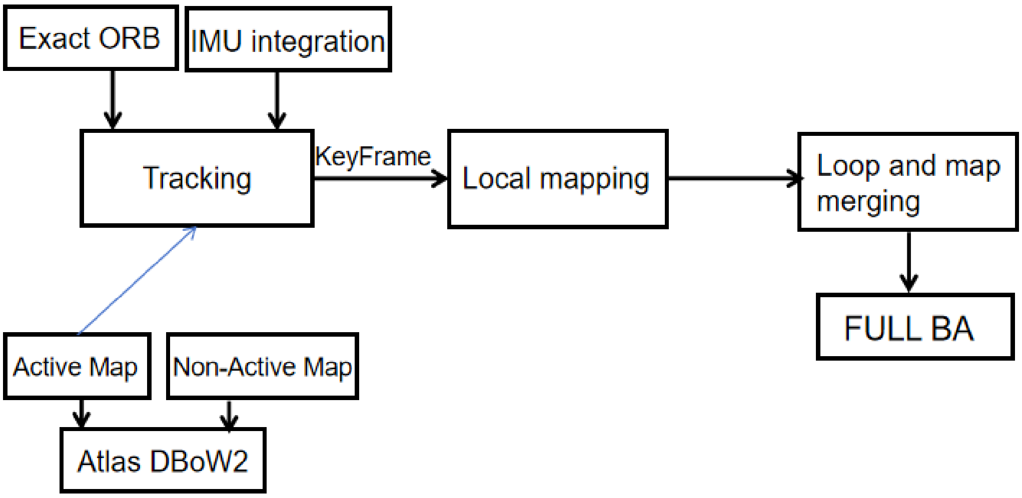


Figure 2. ORB-SLAM3 Algorithm Framework. ORB-SLAM3, Oriented FAST and Rotated BRIEF-Simultaneous Localization and Mapping 3.

nals and with a high sampling rate, capture rapid motion and determine the carrier's position and orientation through double integration.

VINS leverages the complementary strengths of visual sensors, which provide detailed environmental information, and IMUs, which measure motion. VSLAM helps correct IMU drift and performs loop closure detection, while IMUs assist VSLAM in addressing localization challenges in scenes with sparse textures or rapid motion, thereby enhancing the accuracy of visual-inertial systems in diverse settings [3]. VINS has broad applications in fields such as autonomous driving, drone navigation, augmented reality, and robotics. With advances in hardware and algorithm optimization, VINS continues to evolve, with ongoing research focusing on improving localization accuracy, enhancing map quality, and achieving more effective data fusion, particularly in dynamic environments. The following sections will elaborate on the current state of VINS development.

Key Technologies

Visual Sensors

With the advancement of computer technology, visual sensors have seen significant improvements in resolution, pixel density, and focus. Based on their operational mode, they can be categorized into monocular cameras, stereo cameras, and RGB-D cameras [4]. Monocular cameras are known for their simple design, low cost, and high computational efficiency, but they lack depth perception. Stereo cameras, on the other hand, derive depth information from the baseline distance between two monocular cameras, enabling depth estimation in both indoor and outdoor settings. However, stereo cameras require more complex configuration and calibration, as well as significant computational resources. RGB-D cameras, which became popular around 2010, capture both color and depth information directly, making them suitable for various routine visual tasks [5]. **Figure 1** shows the timeline of classic VINS algorithms based on filtering and optimization.

Visual odometry

Nister introduced the concept of visual odometry (VO), which involves analyzing image sequences to estimate the pose (position and orientation) of a moving robot in real-time using computer vision techniques [6]. This method overcomes the limitations of traditional odometry methods, providing more accurate localization. There are three main types of algorithms for processing images captured by visual sensors: feature-based methods, optical flow methods, and direct methods. According to the residual model definition, feature-based and optical flow methods are collectively referred to as indirect methods. **Figure 2** is ORB-SLAM3 Algorithm's Framework.

Feature-based methods extract sparse features from images through detection or general-purpose operations. These features are tracked between consecutive frames by matching feature points and adding descriptors to enhance robustness. Prominent feature-based methods include MonoSLAM, proposed by Davison et al. in 2007, which was the first real-time monocular visual SLAM system; ORB-SLAM, ORB-SLAM2, and ORB-SLAM3 [7-10]. Features in images may have specific structures, such as corners, edges, or block objects. It is generally easier to match corners or edges between images than block objects. To improve robustness, line feature extraction has gained attention in recent research. In 2016 , Meier et al. proposed an algorithm using Bezier curves as static landmark primitives instead of point features, combining curve landmarks for mapping to generate continuous curve-based maps [11]. Experiments demonstrated that line features improved computational efficiency compared to point features.

Gomez-Ojeda et al. proposed the PL-SLAM, which combines point and line features for robust performance, especially in scenarios where point features are sparse or unevenly distributed [12]. PL-SLAM does not assume line endpoint positions, enabling its tracking frontend to handle partially occluded lines. He et al. introduced the PL-VIO monocular visual-inertial odometry system, which tightly couples point and line features [13]. It optimizes system state by minimizing IMU preintegration constraints and point and line reprojection errors in a sliding window. Fu et al. improved the LSD algorithm from OpenCV by using Plucker coordinates to represent spatial lines as infinite lines and modeling line reprojection residuals based on point-to-line distances [14-16]. Integrated with VINS-Mono, the improved LSD algorithm significantly increased speed and introduced the first complete VINS framework for point-line feature fusion [17]. Xu et al. developed the EPLF-VINS algorithm, which enhances traditional line feature detection by fusing short lines, ensuring uniform line feature distribution, and using adaptive threshold extraction methods [18]. Teng et al. presented a tightly coupled stereo visual-inertial SLAM system with point-line features, improving line feature detection by introducing new criteria for curvature detection and minimum line segment length [19]. Zheng et al. proposed the LRPL-VIO visual-inertial odometry algorithm, which assumes invariant luminance values between endpoints and midpoints of consecutive frames for fast line matching and to address localization challenges in complex environments [20].

Direct methods minimize photometric errors, while indirect methods minimize geometric displacements. Direct methods, with smaller attraction regions, require good initial guesses, whereas indirect methods consume additional computational resources for feature extraction and matching. Indirect methods are more widely used in practical applications due to their robustness and maturity. However, direct methods, based on photometric consistency, require good initial guesses and high frame rates for accurate image alignment and are more suitable for dense mapping, showing potential in textureless scenes. SVO, introduced by C. Forster et al. in 2014, is a semi-direct monocular VO algorithm [21]. It uses semi-sparse direct methods for pixel-level matching and features for stable tracking and employs image pyramids for scale-space feature extraction and tracking. Engel et al. introduced direct sparse odometry, which minimizes photometric errors while jointly optimizing all model parameters, making it one of the few systems to compute VO purely through direct methods [22]. Zhao et al. enhanced monocular direct VO by developing deep direct sparse odometry, a new architecture that improves scale consistency in the trajectory network TrajNet using geometric constraints [23].

Direct methods minimize photometric errors, while indirect methods minimize geometric displacements. Direct methods require good initial guesses due to their small attraction regions, while indirect methods require additional computational resources for feature extraction and matching. Indirect methods are more robust and widely used, but direct methods are more adaptable for dense mapping and show promise in textureless environments [17].

IMU preintegration

IMU preintegration is a crucial step in VINS for estimating a device's displacement, velocity, and attitude over time by integrating linear acceleration and angular velocity measurements from the IMU. This process addresses challenges such as repeated integration of IMU measurements, which occur at a higher frequency compared to visual sensor data, and error accumulation during prolonged IMU motion [24-27]. IMU preintegration typically involves preprocessing raw IMU data to remove noise, performing coordinate transformations, estimating attitude, velocity, and position using filters like the Kalman or Extended Kalman filters, and predicting motion using physical models. The predicted values are then integrated using a mathematical model to reduce error accumulation.

IMU preintegration is essential for improving pose estimation accuracy in visual-inertial SLAM systems. By accounting for the correlation between IMU measurements and errors, it provides more accurate pose information, thus enhancing the overall performance of SLAM systems.

State Estimation Optimization

State estimation in VINS is an optimization problem where camera pose variations are computed through iterative image alignment [3]. Solutions are primarily divided into filter-based methods and optimization-based methods. Below is an introduction to these two types of approaches.

Filter-based methods

Filter-based methods are a class of state estimation techniques that utilize sensor data to predict and update a system's state. Based on Bayesian filtering theory, these methods combine sensor data with the system's dynamic model to perform recursive state estimation. Common filter-based approaches include the Kalman filter, extended kalman filter (EKF), unscented kalman filter, and particle filter. These methods typically estimate the state by recursively predicting and updating the probability distribution of the system's state.

Previously, filter-based methods were the primary approach for the back-end of V-SLAM, with the core objective of iteratively updating the state using Bayesian filtering. In VINS, EKF and unscented kalman filter, extensions of the Kalman filter, are commonly used to handle nonlinear systems. EKF achieves state estimation by linearizing the state transition and observation models, followed by applying standard Kalman filter methods. Although simple to implement, EKF may introduce errors when linearizing nonlinear functions and requires computing Jacobian matrices, reducing convergence and accuracy in highly nonlinear systems.

On the other hand, unscented kalman filter offers higher accuracy and stability, as it does not require Jacobian matrices and avoids linearization errors, making it more efficient in highly nonlinear systems. In 2007, Mourikis et al. proposed the Multi-State Constraint Kalman Filter, an extension of EKF, which improves robustness and accuracy by introducing multiple state constraints [1]. Multi-State Constraint Kalman Filter handles multiple state variables and observation constraints, fusing them through nonlinear optimization, making it suitable for complex VINS applications.

In 2015, Bloesch et al. introduced the ROVIO algorithm, a VIO based on the direct EKF method [28]. It enhances the state estimation of visual and inertial measurements by incorporating outlier suppression and preintegration techniques, achieving strong performance in dynamic environments. SR-ISWF is an extension of Multi-State Constraint Kalman Filter that uses a single-precision representation in square root form, avoiding numerical instability [29]. It employs an inverse filter for iterative linearization, similar to optimization-based algorithms, batch graph optimization, or bundle adjustment techniques, to optimize all measurement data for the best state estimation.

Optimization-based methods

Optimization-based methods adjust camera poses or 3D point clouds by minimizing reprojection error, photometric error, or other cost functions. The primary goal is to improve pose estimation accuracy, with reprojection error being a common metric. Reprojection error measures the Euclidean distance between predicted and observed image feature point positions. Optimization techniques like the Newton-Gauss method or the Levenberg-Marquardt algorithm are often employed. Levenberg-Marquardt, an algorithm for nonlinear least squares problems, combines the strengths of gradient descent and Newton's method, offering good convergence and robustness.

Another widely used optimization approach is pose graph optimization, which minimizes the error between consecutive poses and constraints from loop closures or sensor measurements. This method is frequently applied in real-time visual SLAM systems, where new sensor data is continuously processed, and incremental optimization is required.

In 2016, Engel et al. introduced direct sparse odometry, a direct VO method that estimates camera motion by minimizing brightness error between consecutive frames, without relying on feature extraction and matching [22]. It employs sliding window optimization, marginalizing out old camera poses and points that leave the field of view. In 2018, Qin et al. proposed VINS-mono, which features a robust initialization process [17]. By optimizing only a fixed number of recent IMU poses and velocities, VINS-mono reduces computational complexity. It also introduces a tightly coupled relocalization module that minimizes drift by establishing feature-level connections between loop closure candidates and the current frame, integrating them into the monocular VIO module.

In 2021, Campos et al. presented ORB-SLAM3, a tightly integrated visual-inertial SLAM system that builds upon its predecessors, ORB-SLAM and ORB-SLAM2 [10]. ORB-SLAM3 fully utilizes maximum a posteriori estimation during IMU initialization, resulting in significantly higher accuracy in both indoor and outdoor environments, with improvements up to tenfold over previous versions. Additionally, ORB-SLAM3 introduced the first multi-map system, leveraging short-term, medium-term, and long-term data associations to reuse information from high-parallax co-visible keyframes across all stages of the algorithm. This system performs loop closure and relocalization using the DBoW2 vocabulary library.

Chen et al. analyzed the application of the first estimate jacobian in VINS [30]. When implemented correctly in a nonlinear optimization framework, first estimate jacobian addresses information error caused by marginalization during linearization of fixed points, ensuring estimation consistency and improving system accuracy.

Overall, optimization-based methods provide a flexible and powerful framework for state estimation, allowing for the incorporation of complex sensor models and system dynamics. These methods are particularly well-suited for scenarios requiring accurate estimates and where computational resources are available for solving large-scale optimization problems.

Dynamic SLAM

Most visual SLAM algorithms assume a static scene, leading to poor performance in dynamic environments. Dynamic objects can introduce outliers that significantly affect the accuracy and stability of these algorithms. Dynamic SLAM methods are generally categorized into geometric-based and semantic-based approaches. Below is an overview of the principles and recent research achievements in these two categories.

Geometric-based dynamic SLAM

Geometric-based dynamic SLAM methods rely on matching key feature points extracted from consecutive image frames to provide geometric constraints for SLAM. These methods remove outliers from dynamic objects by analyzing motion models or applying geometric constraints. In 2017, Wang et al. proposed an end-to-end monocular VO framework based on deep recursive convolutional neural networks [31]. Recursive convolutional neural networks combine convolutional neural networks to extract spatial features from images with recurrent neural networks to capture temporal dynamics, enabling simultaneous feature extraction and sequential VO modeling. The model estimates new poses by inputting consecutive image frames.

In 2018, DeTone et al. introduced SuperPoint, a method unrelated to visual-inertial navigation but relevant for automatically detecting interest points in images and describing them using deep learning techniques [32]. SuperPoint can be used for keypoint detection and description in VINS. Bian et al. published a study on unsupervised learning of depth and camera motion from monocular videos [33]. They proposed geometric consistency losses for scale-consistent prediction and introduced self-discovery masks to handle moving objects and occlusions, addressing the issue of frame scale ambiguity in geometric image reconstruction. In 2022, Liu et al. presented lightweight feature tracking based on RGB-D sensor depth images and IMU inertial data. They divided images into grids, padded grid boundaries to prevent edge features from being ignored, and performed feature detection only on grids with insufficient feature matching [34]. By integrating geometric and depth information, they achieved dynamic feature recognition comparable to semantic segmentation.

Geometric-based methods offer good stability, providing reliable localization and mapping with high precision in various environments. They are computationally efficient and can achieve real-time localization. However, these methods perform poorly in feature-scarce environments and may cause instability in localization and mapping during scene changes. Additionally, their robustness in dynamic scenes is lower than that of semantic-based methods, as they lack semantic information.

Semantic-based dynamic SLAM

Semantic segmentation is a crucial task in computer vision that uses deep learning models to classify each pixel in an image into different semantic categories. This process generates a semantically rich segmentation map, where each pixel is labeled with a specific category, such as road, car, pedestrian, or building. Semantic segmentation is typically modeled using deep convolutional neural networks, with popular architectures including fully convolutional networks, U-Net, DeepLab, and SegNet. These networks perform convolution and deconvolution operations on images, progressively upsampling feature maps to generate segmentation maps of the same size as the original image, assigning each pixel to its most probable semantic category.

Sünderhauf et al. proposed a method to integrate semantic information into vision-based SLAM, utilizing feature-associated semantic data within a factor graph framework to improve outlier/inlier computation, based on a pre-trained deep learning network [35]. Niko et al. introduced a semantic mapping method by combining ORB-SLAM2 with a deep learning-based object detector, performing three-dimensional unsupervised segmentation using planar information from the object detector [36]. Bavle et al. extended this approach to develop a more complete SLAM framework, using landmark covariance to improve Mahalanobis distance calculations, optimizing the overall process [37].

Parkhiya et al. proposed using deep networks to build class-specific models for monocular object SLAM, estimating the relative 3D pose of semantic objects and integrating these with the robot's VO poses into a graph optimization framework [38]. Mccormac et al. introduced Fusion++, an online object-level SLAM system that constructs accurate 3D models of objects, using Mask- recursive convolutional neural network instance segmentation for initializing compact volumetric signed distance fields and incrementally improving them via depth fusion for tracking, relocalization, and loop detection, while optimizing pose graphs [39].

Some research combines geometric and semantic methods to enhance system understanding and environmental processing. Yan et al. proposed DGS-SLAM, which combines geometric and semantic information through a motion segmentation detection module based on a multi-residual model and semantic segmentation techniques, designed to identify potential moving targets in dynamic scenes [40]. Cheng et al. introduced SG-SLAM, a real-time semantic visualization system built on the ORB-SLAM2 framework, which adds two parallel threads for target detection and semantic segmentation [41]. These threads provide dynamic object information for feature rejection strategies and integrate 2D keyframe information and 3D point clouds into a semantic object database, generating an intuitive semantic metric map in the ROS system. Sun et al. proposed a method for classifying dynamic objects based on their motion states, using IMU data and geometric constraints to identify and verify temporarily static features [42]. Additionally, a new adaptive weight loss function based on BA dynamic factors is used in the backend to refine camera poses and spatial point positions.

Semantic methods accurately classify each pixel into specific semantic categories, providing high-quality information for subsequent image analysis and addressing issues with object detection bounding boxes. However, semantic methods heavily depend on neural network quality, making it difficult to balance speed and accuracy simultaneously. Combining geometric and semantic methods will be essential for enhancing the robustness of VINS in dynamic environments.

Conclusion

VINS remains a prominent research topic in fields like robot navigation and autonomous driving. This paper provides a review of the current state of visual-inertial systems, highlighting the continuous advancements in filtering/optimization-based methods and geometric/semantic-based methods in dynamic SLAM. While significant progress has been made over the past decade, several challenges remain:

Data Association: Many methods based on IMU and VO rely on prior assumptions, such as normal distribution and stationary noise, or the accuracy of sensors.

Sensor Fusion: Integrating additional sensors, such as LiDAR, can provide richer information for visual SLAM algorithms.

Deep Learning: Incorporating deep learning methods offers a learnable strategy for SLAM, enabling the use of past SLAM experiences to improve performance in unknown environments.

Author Contributions: Mingxia Wei contributed to the conception of the study and performed the discussion and wrote the manuscript. Qingyun Meng contributed to check the manuscript.

References

[1] Mourikis AI, Roumeliotis SI. A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation. Proceedings 2007 IEEE International Conference on Robotics and Automation 2007;3565-3572.

[2] He M, Zhu C, Huang Q, et al. A review of monocular visual odometry. Visual Comput 2019;36(5):1053-1065.

[3] Gui J, Gu D, Wang S, et al. A review of visual inertial odometry from filtering and optimisation perspectives. Adv Rob 2015;29(20):1289-1301.

[4] Huang G. Visual-Inertial Navigation: A Concise Review. 2019 International Conference on Robotics and Automation (ICRA) 2019;9572-9582.

[5] Kerl C, Sturm J, Cremers D. Dense visual SLAM for RGB-D cameras. 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems 2013;2100-2106.

[6] Nister D, Naroditsky O, Bergen J. Visual odometry. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. 2004;I-I.

[7] Davison AJ, Reid ID, Molton ND, et al. MonoSLAM: real-time single camera SLAM. IEEE Trans Pattern Anal Mach Intell 2007;29(6):1052-1067.

[8] Mur-Artal R, Montiel JMM, Tardos JD. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. IEEE Trans Rob 2015;31(5):1147-1163.

[9] Mur-Artal R, Tardos JD. ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras. IEEE Trans Rob 2017;33(5):1255-1262.

[10] Campos C, Elvira R, Rodriguez JJG, et al. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual–Inertial, and Multimap SLAM. IEEE Trans Rob 2021;37(6):1874-1890.

[11] Meier K, Chung SJ, Hutchinson S. Visual-inertial curve SLAM. 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2016;1238-1245.

[12] Gomez-Ojeda R, Moreno F-A, Zuniga-Noel D, et al. PL-SLAM: A Stereo SLAM System Through the Combination of Points and Line Segments. IEEE Trans Rob 2019;35(3):734-746.

[13] He Y, Zhao J, Guo Y, et al. PL-VIO: Tightly-Coupled Monocular Visual-Inertial Odometry Using Point and Line Features. Sensors (Basel). 2018:10;18(4):1159.

[14] Fu Q, Wang J, Yu H, et al. PL-VINS: Real-Time Monocular Visual-Inertial SLAM with Point and Line Features. arXiv:2009.07462 2022.

[15] Grompone von Gioi R, Jakubowicz J, Morel JM, et al. LSD: a fast line segment detector with a false detection control. IEEE Trans Pattern Anal Mach Intell 2010;32(4):722-732.

[16] Kaehler A, Bradski G. Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library. O'Reilly Media, Inc., 2016.

[17] Qin T, Li P, Shen S. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator. IEEE Trans Rob 2018;34(4):1004-1020.

[18] Xu L, Yin H, Shi T, et al. EPLF-VINS: Real-Time Monocular Visual-Inertial SLAM With Efficient Point-Line Flow Features. IEEE Rob Autom Lett 2023;8(2):752-759.

[19] Teng Z, Han B, Cao J, et al. PLI-SLAM: A Tightly-Coupled Stereo Visual-Inertial SLAM System with Point and Line Features. Remote Sens 2023;15(19):4678.

[20] Zheng F, Zhou L, Lin W, et al. LRPL-VIO: A Lightweight and Robust Visual-Inertial Odometry with Point and Line Features. Sensors 2024;24(4):1322.

[21] Forster C, Pizzoli M, Scaramuzza D. SVO: Fast semi-direct monocular visual odometry. 2014 IEEE International Conference on Robotics and Automation (ICRA) 2014;15-22.

[22] Engel J, Koltun V, Cremers D. Direct Sparse Odometry. IEEE Trans Pattern Anal Mach Intell 2018;40(3):611-625.

[23] Zhao C, Tang Y, Sun Q, et al. Deep Direct Visual Odometry. IEEE trans Intell Transp Syst 2022;23(7):7733-7742.

[24] Liu Z, Shi DX, Yang SW, et al. Review of visual-inertial navigation system initialization method. J Natl Univ Def Technol, China 2023;45(02):15-26.

[25] Lupton T, Sukkarieh S. Visual-Inertial-Aided Navigation for High-Dynamic Motion in Built Environments Without Initial Conditions. IEEE Trans Rob 2012;28(1):61-76.

[26] Christian F, Luca C, Frank D, et al. IMU Preintegration on Manifold for Efficient Visual-Inertial Maximum-a-Posteriori Estimation. Robotics: Sci Syst 2015.

[27] Forster C, Carlone L, Dellaert F, et al. On-Manifold Preintegration for Real-Time Visual--Inertial Odometry. IEEE Trans Rob 2017;33(1):1-21.

[28] Bloesch M, Omari S, Hutter M, et al. Robust visual inertial odometry using a direct EKF-based approach. 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2015;298-304.

[29] Paul MK, Wu K, Hesch JA, et al. A comparative

analysis of tightly-coupled monocular, binocular, and stereo VINS. 2017 IEEE International Conference on Robotics and Automation (ICRA) 2017;165-172.
[30] Chen C, Geneva P, Peng Y, et al. Optimization-Based VINS: Consistency, Marginalization, and FEJ. 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2023;1517-1524.
[31] Wang S, Clark R, Wen H, et al. DeepVO: Towards end-to-end visual odometry with deep Recurrent Convolutional Neural Networks. 2017 IEEE International Conference on Robotics and Automation (ICRA) 2017;2043-2050.
[32] DeTone D, Tomasz Malisiewicz, Andrew Rabinovich. SuperPoint: Self-Supervised Interest Point Detection and Description. arXiv:1712.07629 2018.
[33] Bian JW, Li Z, Wang N, et al. Unsupervised scale-consistent depth and ego-motion learning from monocular video. In: editors. Proceedings of the 33rd International Conference on Neural Information Processing Systems. Curran Associates Inc., 2019. p. Article 4.
[34] Liu J, Li X, Liu Y, et al. RGB-D Inertial Odometry for a Resource-Restricted Robot in Dynamic Environments. IEEE Rob Autom Lett 2022;7(4):9573-9580.
[35] Murali V, Chiu HP, Samarasekera S, et al. Utilizing semantic visual landmarks for precise vehicle navigation. 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC) 2017;1-8.
[36] Sünderhauf N, Pham TT, Latif Y, et al. Meaningful maps with object-oriented semantic mapping. 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2017;5079-5085.
[37] Bavle H, Puente PDL, How JP, et al. VPS-SLAM: Visual Planar Semantic SLAM for Aerial Robotic Systems. IEEE Access 2020;8:60704-60718.
[38] Parkhiya P, Khawad R, Murthy JK, et al. Constructing Category-Specific Models for Monocular Object-SLAM. 2018 IEEE International Conference on Robotics and Automation (ICRA) 2018;4517-4524.
[39] Mccormac J, Clark R, Bloesch M, et al. Fusion++: Volumetric Object-Level SLAM. 2018 International Conference on 3D Vision (3DV) 2018;32-41.
[40] Yan L, Hu X, Zhao L, et al. DGS-SLAM: A Fast and Robust RGBD SLAM in Dynamic Environments Combined by Geometric and Semantic Information. Remote Sens 2022;14(3):795.
[41] Cheng S, Sun C, Zhang S, et al. SG-SLAM: A Real-Time RGB-D Visual SLAM Toward Dynamic Scenes With Semantic and Geometric Information. IEEE Trans Instrum Meas 2023;72:1-12.
[42] Sun Y, Wang Q, Yan C, et al. D-VINS: Dynamic Adaptive Visual–Inertial SLAM with IMU Prior and Semantic Constraints in Dynamic Scenes. Remote Sens 2023;15:3881.

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Review Article https://doi.org/10.61189/673672yizrwd

Review of gait prediction of lower extremity exoskeleton robot

Haonan Geng[1], Xudong Guo[1], Haibo Lin[1], Youguo Hao[2], Guojie Zhang[3]

[1]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China. [2]Shanghai Putuo District People's Hospital, Shanghai 200060, China. [3]LingYuan Iron and Steel CO., LTD, Lingyuan 122500, Liaoning Province, China.

Corresponding authors: Xudong Guo and Youguo Hao.

Declaration of conflict of interest: None.

Received September 8, 2024; Accepted November 6, 2024; Published December 31,2024

Highlights

- Gait prediction relies on multimodal sensor data, and the acquisition of multimodal information, such as physical sensors and bioelectrical signal sensors, is introduced in order to monitor and analyze the lower limb movement in real time, and provide a data basis for prediction.
- The application of machine learning algorithms in gait prediction technology, such as Support Vector Machine, Random Forest, and Back Propagation Neural Network, is reviewed to construct an optimized gait prediction model, which provides effective support for the intelligent control of exoskeleton.
- Compared with machine learning algorithms, the article summarizes the researchers' efforts to extract and un derstand the hidden patterns in gait data by constructing neural network models related to different deep learn ing algorithms, which are used to improve the accuracy and robustness of gait prediction.

Abstract

In recent years, gait prediction has gradually become a cutting-edge research direction in the fields of biomechanics and artificial intelligence. Gait prediction technology, which analyzes an individual's walking patterns to predict future changes, is crucial for the precision of rehabilitation and exoskeleton robot control. This paper reviews the recent research progress in the field of gait prediction, focusing on the multimodal information acquisition methods based on physical sensors and bioelectric signals, as well as the application of machine learning and deep learning algorithms in gait prediction. By analyzing different sensor data fusion strategies, the importance of multimodal information fusion for improving the accuracy of gait prediction is emphasized. Furthermore, this paper introduces the performance of traditional machine learning algorithms such as Support Vector Machine, Random Forest, and Back Propagation Neural Network, as well as deep learning models such as Long Short-Term Memory, Convolutional Neural Network, and Transformer in gait prediction, highlighting the advantages of deep learning in feature extraction and adaptability to complex scenarios. Finally, this paper explores future directions for the development of gait prediction technology, emphasizing improvements in timeliness, accuracy, and personalization to advance exoskeleton robotics and related fields.

Keywords: Gait prediction, multimodal information acquisition, machine learning, deep learning

Introduction

Amidst the intensifying trend of global population aging, the incidence of neurological disorders precipitated by factors such as traumatic brain injury, spinal cord injury, stroke, and Freezing of Gait has been on the rise [1-4]. Exoskeleton robot is a kind of intelligent mechanical structure worn on the outside of the user 's body. By combining sensing, control, information, fusion, mobile computing and other technologies, the human body's feeling, thinking, movement and other organs are combined with the machine's perception system, intelligent processing center and control execution system, so as to improve and strengthen the human body 's physical function or for medical rehabilitation [5]. It is a new type of electromechanical integration device, which has the functions of support, movement and protection.

Address correspondence to: Xudong Guo, School of Health Science and Engineering, University of Shanghai for Science and Technology, No.516 Jungong Road, Yangpu District, Shanghai 200093, China. Email: guoxd@usst.edu.cn; **Youguo Hao**, Shanghai Putuo District Central Hospital, No.1291 Jiangning Road, Putuo District, Shanghai, 200060, China. Email: youguohao6@163.com.

Combined with key technologies such as mechanical structure, control, driving mode and human-computer interaction, it enables the wearer to perform tasks that would otherwise be impossible [6].

Gait prediction is a frontier research direction in the fields of biomechanics and artificial intelligence, involving the analysis and prediction of human walking patterns [7]. It has numerous applications in lower limb exoskeleton robots, with real-time gait monitoring and adjustment being the primary application. This technology integrates various sensors (accelerometers, gyroscopes, pressure sensors, etc.) and computer vision to capture the lower limb movement data of users, and analyzes and identifies the trends of gait changes. Based on these data, predictive models trained by machine learning and deep learning algorithms can monitor the gait patterns of users in real-time and predict changes in gait. Exoskeleton robots can adjust the amount of assistance, motion synchronization, and movement patterns according to the predicted results, helping users walk more naturally. This is particularly important for patients with neuromuscular diseases, those with sports injuries, or those who have been bedridden for a long time, as it provides personalized gait assistance to help them regain the ability to walk.

Furthermore, gait prediction technology can significantly enhance the motion stability and fall prevention capabilities of exoskeleton robots. By analyzing the gait characteristics and load distribution of users, predictive models can identify unstable states in advance, allowing the exoskeleton system to react preemptively, such as increasing support force when the user is about to lose balance. This is especially suitable for the elderly, patients with balance disorders, or users in the rehabilitation phase, effectively reducing the secondary injuries caused by falls.

In the field of rehabilitation, gait prediction offers the possibility of personalized rehabilitation training programs for exoskeleton robots. By collecting and analyzing the gait data of patients over the long term, predictive models can assess the progress of rehabilitation, identify problems in gait, and adjust the rehabilitation plan. For example, assistance can be adjusted according to the patient's recovery, and help the patient to gradually regain the ability to walk independently [8]. This not only accelerates the rehabilitation process but also improves the effectiveness of rehabilitation.

Gait prediction technology can also help healthy individuals enhance their athletic performance. For professions that require long periods of walking or walking with loads, such as soldiers and firefighters, exoskeleton robots can optimize movement through gait prediction models, adjust energy output, and allow users to complete longer walking tasks with less physical exertion. This feature can also be applied to outdoor activities such as hiking and mountaineering, helping users maintain stable walking on complex terrains and reducing the risk of injury [9].

Currently, gait prediction, as a key technology in lower limb exoskeleton robots, integrates sensors, computer vision, machine learning, and deep learning to achieve high-precision capture and analysis of gait [10]. By learning from a vast amount of gait data, predictive models can identify specific patterns in individual gait and forecast potential trends of change, providing a scientific basis and decision support for related fields, thereby further advancing the development of lower limb exoskeleton robots.

Acquisition of multi-modal information for gait prediction

In the process of gait prediction, the acquisition of multimodal information is necessary, and methods based on sensor data play a crucial role in gait prediction research. Different types of sensors, such as Inertial Measurement Units (IMUs), surface electromyography (sEMG), and plantar pressure sensors, are widely used for the collection, analysis, and prediction of gait data [11-14]. These sensor data allow for real-time monitoring and analysis of the state of human lower limb movement, providing a solid foundation for the prediction of gait patterns. The commonly used multi-modal information acquisition methods for gait prediction are shown in **Figure 1**.

Signal acquisition based on physical sensors

During the acquisition of sensor data, IMUs are widely used for capturing gait information due to their high precision and portability. An IMU is a device that integrates accelerometers, gyroscopes, and magnetometers, capable of measuring and reporting the specific forces, angular velocities, and magnetic fields surrounding an object's attitude. In gait analysis, IMUs are typically attached to the lower limbs or feet to capture data such as angular velocities and accelerations of various human joints during walking. By combining these data with deep learning models, such as long short-term memory (LSTM) networks, high-accuracy gait

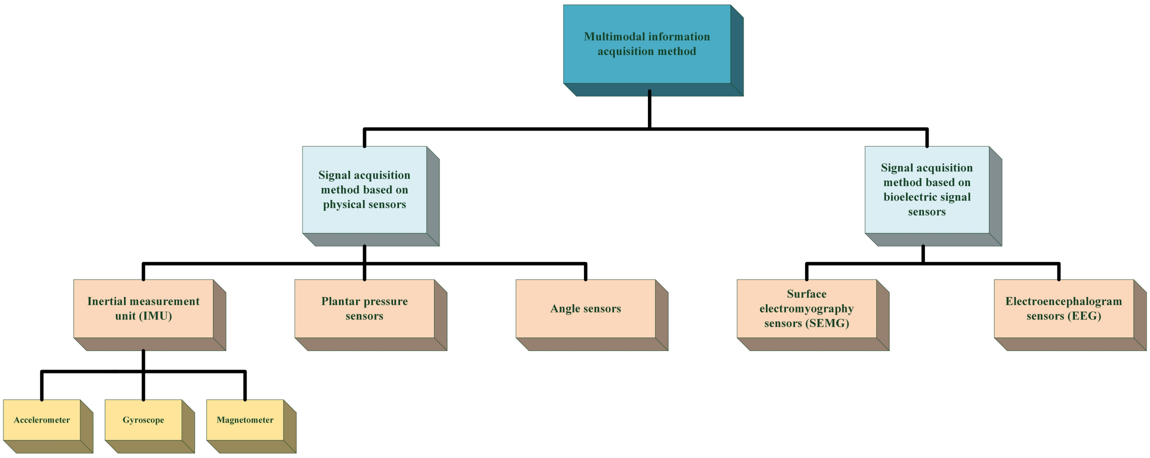


Figure 1. Multi-modal information acquisition methods for gait prediction.

classification and prediction can be achieved [15]. Zhang and colleagues have mentioned that fusing IMU sensor data with data from other sensors can significantly enhance the effectiveness of gait feature extraction, thereby improving the accuracy of gait recognition [16]. Additionally, there is literature introducing the embedding of IMU sensors within waist control components to obtain real-time joint angle information, combined with an LSTM model augmented with an attention mechanism for gait prediction [17].

Plantar pressure sensors provide detailed gait information by measuring the force and distribution of pressure as the foot contacts the ground during walking [8]. These sensors are usually embedded in the insole or dedicated pressure measurement board, which can capture the pressure distribution of the foot in the support phase, the movement path of the pressure center, and the pressure peak. By analyzing the pressure distribution across different regions of the foot, subtle changes in gait can be inferred. A study has introduced a method for analyzing human motion by collecting pressure data from four specific locations on the plantar surface, based on the pressure variation patterns observed at different foot positions during movement [18]. Four sets of sensors were placed at different locations on the left and right feet to detect signal variations in the corresponding regions. This setup assists the exoskeleton robot's perception system in detecting changes in foot gait signals during walking at a constant speed of 1.5 m/s. Through these signals, the sample data of the human body during exercise can be obtained to facilitate subsequent research. In addition, the plantar pressure sensor can also be used to detect gait events such as standing phase and swing phase, providing important input data for gait prediction algorithms [19].

Signal acquisition based on bioelectric signal sensors

Bioelectrical signals from surface sEMG and electroencephalogram (EEG) can reflect the state of muscle activity and brain neuron activity in a very short time. sEMG and EEG commonly utilize sensors that are affixed to the muscular and cerebral surfaces, thereby providing significant portability. The integration of sEMG signals with data from additional sensors enables the effective identification of gait patterns and the prediction of lower limb movement intentions [20]. Liu used a special double-sided tape to attach the sEMG sensor to three muscles closely related to joint movement, namely rectus femoris, vastus lateralis and biceps femoris, combined with an angle sensor to record the coordinated motion data of the subjects. Their study established an attention-based Convolutional Neural Network (CNN)-LSTM model for generating personalized gait trajectories of affected limbs [21]. Some researchers combined electromyography (EMG) signals with EEG signals [22]. The feature-level fusion method was used to fuse EEG and EMG for lower limb motion intention recognition, with accuracy rate reached 90.23%. Compared with the single EEG feature, the feature fusion method improved the recognition accuracy by 2.73%.

Multi-sensor data fusion

Multiple studies have shown that relying on a

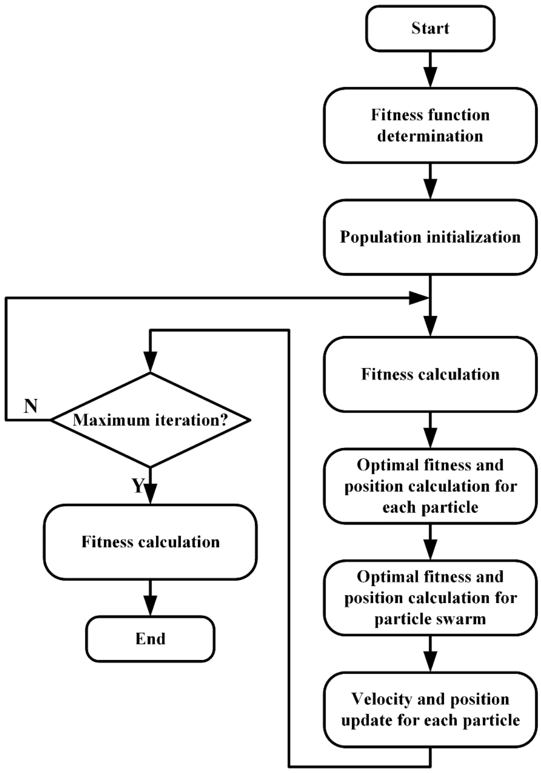


Figure 2. Flowchart of PSO-SVM algorithm. This figure was cited from [30]. PSO, Particle Swarm Optimization; SVM, Support Vector Machines.

single sensor often fails to achieve sufficient accuracy [23-25]. The fusion of data from multiple sensors is a key strategy for improving the accuracy of gait prediction. By integrating data from various sensors, such as accelerometers, gyroscopes, and pressure sensors, more comprehensive and precise gait characteristic information can be obtained. A multimodal neural network model, which combines sEMG signals with inertial sensor data, can effectively predict gait patterns. These methods leverage the temporal continuity and spatial complementarity of multi-sensor data, significantly enhancing the stability and accuracy of gait prediction [26].

In a study, researchers combined plantar pressure sensors and attitude angle sensors to identify and predict gait [23]. The plantar pressure sensors were primarily distributed across three key load-bearing areas: the heel, the forefoot, and the toes, allowing simultaneous collection of pressure values from these regions during movement. Attitude angle sensors were affixed to the thigh, calf, and foot to obtain real-time posture angles. The study further combined least mean squares and recursive least squares algorithms, resulting in minimal error between the original and predicted data, demonstrating effective prediction performance. Multi-sensor data fusionprovides more accurate, more comprehensive, and more robust recognition capabilities in gait prediction, which helps to improve performance and user experience in various application scenarios [27].

Classification of gait prediction algorithms

Machine learning algorithms

In gait prediction for intelligent lower limb exoskeleton robots, machine learning methods have shown significant potential, particularly in processing sensor data to recognize and predict gait phases and trajectories [28]. By integrating and analyzing data collected from various sensors, optimized gait prediction models can be developed, providing effective support for the intelligent control of exoskeletons.

Traditional machine learning algorithms, such as Support Vector Machines (SVM), Random Forest (RF), Backpropagation (BP) Neural Networks, and Linear Regression models, play a crucial role in the field of gait prediction.

SVM

SVM is a supervised learning method that offers significant advantages, particularly in handling high-dimensional data. It possesses strong generalization capabilities and robust pattern recognition abilities, making it well-suited for nonlinear problems. Due to these unique optimization objectives and classification abilities, SVM holds potential for important applications in the field of gait analysis and prediction [29]. The research team led by Wu employed a Particle Swarm Optimization (PSO)-SVM approach, which uses PSO to optimize the SVM parameters C and G, to develop a predictor for gait phase prediction based on spatial features. The specific algorithm flow is shown in **Figure 2**.

In terms of input, the exoskeleton robot is equipped with four angular sensors that capture the hip and knee joint angles during walking, forming a sensor network. Finally, according to the plantar pressure distribution, four gait phases of heel-contact, foot-flat, heel-off and toe-high are determined. The experimental results show that the PSO-SVM method can predict the gait phase well [30]. In another study, Luo and colleagues proposed using SVM to train the mapping relationship between EMG signals and motion, aiming to classify lower limb gait patterns and predict the continuous movements of the hip, knee, and ankle joints in the sagittal plane [31]. When performing regression predictions on joint angles, the

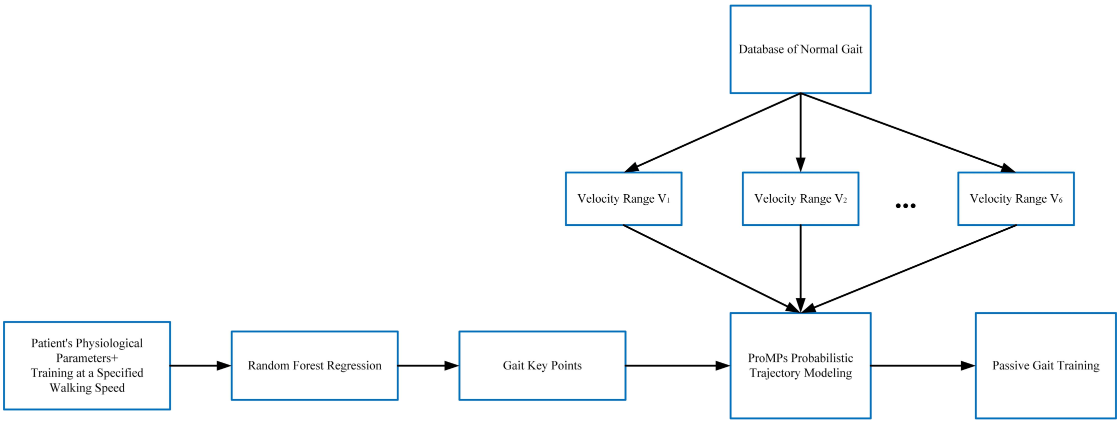


Figure 3. Structure of the passive training gait prediction algorithm combined with ProMPs. This figure was cited from [35]. ProMPs, Probabilistic Movement Primitives.

best-performing subject exhibited a minimum prediction error of 7.02° for the knee joint, 7.97° for the hip joint, and 5.11° for the ankle joint. For all subjects, during walking at a comfortable speed, the root mean square errors (RMSE) for the predicted hip, knee, and ankle joint angles were 9.35°, 10.82°, and 6.87°, respectively, with percentage errors of 20.63%, 15.95%, and 17.11%.

To predict continuous lower limb movements, Duan et al. also employed SVMs to develop a model for each gait phase [32]. In their study, an SVM classifier was first used to identify the corresponding phase model, and then SVM was employed to predict the joint angles. They also compared SVM with the traditional joint angle prediction method using the whole segment motion data for modeling, and found that the effect of joint angle prediction was significantly improved, especially the correlation coefficient of the ankle joint with smaller amplitude and more complex motion increased from about 0.9 to more than 0.99, and the average RMSE decreased from 1.5°- 3.0° to 0.2°- 1.0°. In addition, Qin et al. proposed a prediction algorithm based on least squares support vector regression (LS-SVR) using human lower limb gait sample data to predict the next human lower limb movement, with an accuracy rate of 99.6% [19].

RF

RF can be applied for gait prediction to enhance both accuracy and generalization capabilities [33]. As a powerful ensemble learning method, RF constructs multiple decision trees and combines their predictions to effectively handle nonlinear problems and reduce the risk of model overfitting. It is well-suited for directly processing high-dimensional datasets and can significantly shorten training time through parallel processing. Ren et al. proposed a gait pattern prediction method based on anthropometric features, using the RF algorithm to design personalized gait trajectories for different patients [34]. Finally, the prediction accuracy of the RF model was evaluated by 5-fold cross-validation, and the RMSE and Pearson correlation coefficient were used as evaluation criteria. The study also compared the prediction results with those of the Multilayer Perceptron Neural Network (MLPNN) model, finding that the RMSE of MLPNN exhibited greater volatility. Due to the need to apply gait prediction to some hemiplegic rehabilitation patients in daily life, Lu and others proposed a gait prediction algorithm based on RF and Probabilistic Movement Primitives (ProMPs) to solve the problem of single gait and poor adjustability in passive training of hemiplegia rehabilitation, and the specific process is shown in **Figure 3** [35]. The algorithm takes normal human gait data and the specified gait speed V as inputs, trains the RF-ProMPs, and outputs the predicted angles y_1^*... y_m^* for m knee gait keypoints, and then updates the trajectory model according to the predicted angle values y_1^* of the gait key points to predict the trajectory. Using the mean absolute error (MAE) as well as the correlation coefficient ρ, the errors were found to be at a low level. For active training of hemiplegic patients, the study suggests that it was feasible to use only hip and knee angles for gait phase prediction. After testing, the Pearson-RF model was used, combining the variable importance of RF and the Pearson's correlation coefficients for the input feature filtering. Compared with the SVM using all the features, the average precision rate of the RF with feature filtering increased by 10.04%, and the average recall rate increased by 10.24%.

BP neural network

In addition, BP neural networks have been widely used in gait prediction, and have shown great potential in gait prediction and analysis with their powerful nonlinear mapping ability. By correctly selecting the network structure, activation function, loss function, and reasonable optimization algorithm, the accuracy and efficiency of gait prediction can be effectively improved. Niu's research team used plantar pressure data to predict lower limb gait cycle characteristics (hip and knee angle curves), and obtained gait trajectories by querying the database. The database is constructed by BP neural network used to predict the gait data of the subjects [36]. The experimental results showed that the average error of the hip angle curve was 0.247° and the average error of the knee angle curve was 0.435°, which illustrates the effectiveness of the gait prediction method. In order to predict the gait to be able to offset the problems such as sluggishness of the exoskeleton robot that may be caused by the delay of the system, so that the exoskeleton robot can follow the human behavior in real time and control, Ma et al. designed two algorithms, interpolation and BP neural network, to predict the gait data, and found that the BP neural network exhibited better prediction [37].

Other machine learning algorithms

In addition to the common machine models mentioned above, there are other machine modeling algorithms that have fetched good results. Tanghe et al. proposed a real-time method for predicting future joint angles, velocities, accelerations, and acuteness during walking with a wearable exoskeleton device, as well as the timing of gait events such as the initial contact, foot flat, heel off and toe off [38]. The core of the approach is to utilize gait patterns learned from normal walking using probabilistic principal component analysis, which are then used in the online phase for state estimation and prediction.

Then, cross validation was used to prove the effectiveness of this method for trajectory and event prediction. Another team of researchers proposed a gait prediction method based on Deep Gaussian Processes (DGP) to estimate the operator's gait intention in real time and to compensate for the measurement delay of the IMU. In order to deal with the problem of gait prediction accuracy under small-scale training data, the study again used the Dual Stochastic Variational Inference (DSVI)-DGP model [39]. DSVI-DGP improves the prediction performance of the model by simplifying the intra-layer correlation while maintaining the inter-layer correlation. The method is effective in predicting the operator's gait and has a high prediction accuracy that keeps the error within acceptable limits.

In order to develop a personalized gait pattern prediction model for gait planning of lower limb exoskeleton robots, Shen et al. proposed a gait parameter extraction based on Fourier transform and a gait parameter prediction model based on Multilayer Perceptron Neural Network (LASSO) regression [40]. Fourier transform is used to convert the sagittal angle curves of ankle, knee and hip joints at different speeds into discrete gait parameters, which are used as input and trained in the regression algorithm [40]. LASSO regression can select variables during model training by adding L1 norm penalty term, so as to simplify the model and reduce the risk of over-fitting. Comparing this model with the SVR model and Gaussian Process Regression model constructed by scikit-learn packaged in python, the LASSO regression model showed better prediction performance with higher correlation coefficients and lower mean absolute errors. **Table 1** presents gait prediction studies related to machine learning algorithms.

Deep learning algorithms

Deep learning is a type of machine learning that uses data for learning and prediction, but puts more emphasis on modeling and predicting complex data through deep feature extraction and representation learning [41]. Deep learning plays a central role in gait prediction, which significantly improves the accuracy and robustness of gait prediction by constructing complex neural network models to extract and understand hidden patterns in gait data. Specifically, deep learning can automatically learn complex features from large amounts of data, reducing the need for manual feature engineering and improving the model's ability to generalize to complex scenes.

The classical deep learning algorithms commonly applied to gait prediction are LSTM Network, CNN, Transformer and so on [8]. The application of these deep learning algorithms in gait prediction not only improves the accuracy of recognition, but also enhances the model's ability to adapt to complex scenes.

LSTM

LSTM networks excel in processing time-series data and capturing long-term dependencies. The hierarchical features obtained from training on large-scale datasets have the ability to inte-

Table 1. Related studies based on machine learning algorithms

Researcher	Prediction Method	Predictive Accuracy
Wu et al.[30]	PSO-SVM	Average simulated gait pattern pattern success rate: 87.19%
Luo et al.[31]	SVM	Hip angle prediction: 20.63%； Ankle angle prediction: 17.11%; Knee angle prediction: 15.95%;
Duan et al.[32]	SVM	Correlation coefficient between predicted and measured values: > 0.99 Average Root Mean Squared Error (RMSE): 0.2°–1.0°
Qin et al.[19]	LS-SVR	Accuracy: 99.6%
Ren et al.[34]	RF	Compared with MLPNN, the RMSE volatility of MLPNN is relatively larger.
Lu et al.[35]	Hemiplegic passive rehabilitation training：RF-ProMPs Hemiplegia active rehabilitation training：Pearson-RF	Hemiplegic passive rehabilitation training: The adjustment error of gait trajectory is less than 0.3 °. Hemiplegia active rehabilitation training: Compared with the SVM using all features, the average precision rate of the random forest after feature selection is increased by 10.04 %, and the average recall rate is increased by 10.24 %.
Niu et al.[36]	BP Neural Network	Average error for hip joint angle curve: 0.247° Average error for knee joint angle curve: 0.435°
Ma et al.[37]	BP Neural Network	Compared with the interpolation algorithm, the prediction effect is better.
Tanghe et al.[38]	Probabilistic principal component analysis	Using cross-validation, it is proved that this method is effective for trajectory and event prediction.
Chen et al.[39]	DSVI-DGP	Comparison with K-Nearest neighbors, support vector, back-propagation neural network, gaussian process, bayesian neural network, and DGP: The DSVI-DGP model has the smallest RMSE, indicating the best prediction performance.
Shen et al.[40]	LASSO	Comparison with SVR and Gaussian Process Regression: The LASSO regression model has a higher correlation coefficient and lower MAE.

Note: PSO, Particle Swarm Optimization; SVM, Support Vector Machines; LS-SVR: least squares support vector regression; RF, Random Forest; BP, Backpropagation; SVM, Support Vector Machines; ProMPs, Probabilistic Movement Primitives; MLPNN: Multilayer Perceptron Neural Network; RMSE, root mean square errors; DSVI, Dual Stochastic Variational Inference; DGP, Deep Gaussian Processes; LASSO, Multilayer Perceptron Neural Network; MAE, mean absolute error.

grate data from different sensors, strong generalization ability, and the robustness to factors such as occlusion, change of view angle, and change of clothing in real application scenarios. Su et al. used the LSTM network to predict the future trajectory (angular velocity) and five gait stages (loading response, mid-stance, terminal stance, preswing and swing) of the thighs, shanks, and feet through the collected IMU data [42]. All predicted trajectories exhibit a strong correlation with the measured trajectories, with correlation coefficients exceeding 0.98. Additionally, the prediction accuracy for gait phases reaches 95%. Sun et al. used inertial assessment platform sensors to collect gait information, which was then preprocessed to form input features [43]. These features, along with the corresponding assistive force cycles, were fed into an LSTM model to predict gait patterns, including uniform walking, variable walking, upstairs, and uphill. The final mean squared error (MSE) values were all below 0.1, and the proposed method achieved an average prediction accuracy of 97.23%. This demonstrates the high accuracy and reliability of the method in predicting assistive force cycles. Moosavian et al. proposed using LSTM networks to predict future trajectories and classify walking phases for RoboWalk (a weight support assist device) [44]. The accuracy of the prediction is about 98.5%, and the overall error of all gait patterns is less than 5 degrees. In addition, Lu et al. used the deep neural network of LSTM architecture to accurately predict gait phase through thigh and calf inertia data. During the test, the gait data of two healthy participants were calibrated. Through the classification and prediction of the LSTM model, a stable gait phase prediction was achieved using the time history data of 40 sample points. During walking experiences in structural and over-ground environments, the algorithm achieved an average accuracy of 85.6% and 73.3%, respectively. The lowest RMSE was 0.143, achieved during the identification of 40 consecutive gait phase estimates for the current gait phase [45].

CNN

The majority of researchers tend not to use one model alone, but build gait prediction models combing models such as CNN, Deep Neural Network (DNN), Conformer and others. CNN are outstanding in feature extraction and image processing. Ren et al. constructed a gait trajectory prediction model based on CNN, LSTM and attention mechanism gait trajectory prediction [46]. The model achieved low MSE and MAE values for the tasks of predicting right knee, left knee, right hip, and left hip angles, indicating that the model was able to accurately predict gait trajectories. In addition, Zhu et al. also used a CNN-LSTM model for the prediction of joint angles [47]. In order to obtain the eigenvalues reflecting the gait characteristics, they introduced an improved principal component analysis (PCA) method during data preprocessing for dimensionality reduction, effectively removing redundant information. Finally, by comparing with traditional machine learning methods such as BP neural network, RF, SVR, etc., it is found that the improved PCA algorithm outperforms the traditional BP neural network and SVM in terms of convergence time and prediction accuracy, with better final prediction performance. Liu et al. also used the collaborative gait prediction model of CNN-LSTM network based on attention mechanism to predict the gait trajectory of the affected limb according to the sEMG signal of the healthy lower limb and the knee joint angle [21]. Coordinated gait trajectories with low MAE and high Pearson correlation coefficient were generated.

Transformer

Using traditional Recurrent Neural Networks or LSTM networks, it is difficult to capture long-term dependencies in time-series data through the mechanism of self-attention, whereas this can be effectively realized in the Transformer model. Through its parallel processing capabilities and positional coding, Transformer not only improves processing efficiency, but also ensures chronological accuracy. Its flexible modeling structure allows for adaptation to data of varying complexity, while the encoder-decoder architecture is particularly well suited for predicting future points in time. Therefore, some researchers have constructed a transformer-based network to predict the angle of the hip and knee joints. They adjusted the Transformer architecture while retaining the 'sequence-to-sequence' structure, and proposed a gait prediction model (Transformer-Based Neural Network: TFSformer) to integrate information in plantar pressure data and predict joint angles [48]. The experiments compared TFSformer with CNN, Transformer and CNN Transformer. In the single-step prediction task, the average MSE of TFSformer was reduced by 10.83%, 15.04% and 4.74% compared to CNN, Transformer and CNN Transformer, respectively. The average MAE was reduced by 20.40%, 29.09% and 10.43%, respectively. In the multi-step prediction task, TFSformer achieved the lowest MAE. TFSfomer showed strong robustness and stability in this study. Regardless of the individual gait diversity, it can predict the gait trajectory of different people with a small error. Zhao et al. proposed a novel gait prediction method that combines LSTM networks with Conformer models [49]. Among them, the Conformer model is a Seq2Seq (sequence-to-sequence) model that combines the CNN and Transformer models, utilizing the CNN to enhance the ability of extracting local features and retaining the Transformer model's adeptness at capturing content-based global interaction capabilities. The temporal characteristics of the gait data are initially extracted using LSTM, and the depth temporal and spatial characteristics in the lower limb gait data are further mined using Conformer. The results also show that the accuracy of the model proposed in this paper is greatly improved, with an average accuracy of 0.9489, which is 3.17% higher than the average accuracy of LSTM, and 4.56% higher than the average accuracy of Recurrent Neural Networks.

Multi-model fusion strategy

In the field of gait prediction, multi-model fusion strategy can significantly improve the accuracy and robustness of prediction. Since gait data usually contain complex biomechanical characteristics and variable walking patterns, a single model may be difficult to capture the full picture. Multi-model fusion can understand and predict gait patterns more comprehensively by combining the advantages of different algorithms. In addition, the fusion strategy can reduce the risk of overfitting a specific data subset and improve the generalization ability of the model in different populations and under different conditions. Guo et al. proposed a new method of fusing multimodal signals-transferable multi-modal fusion (TMMF), which achieves continuous prediction of knee joint angle and corresponding gait phase by fusing multimodal signals [50]. The architecture is shown in **Figure 4**. In the process of collecting data, virtual reality tracker, EMG signal acquisition equipment, and gyroscope are fixed on the left leg of the subject to obtain virtual reality data, EMG data, and gyroscope data. The data

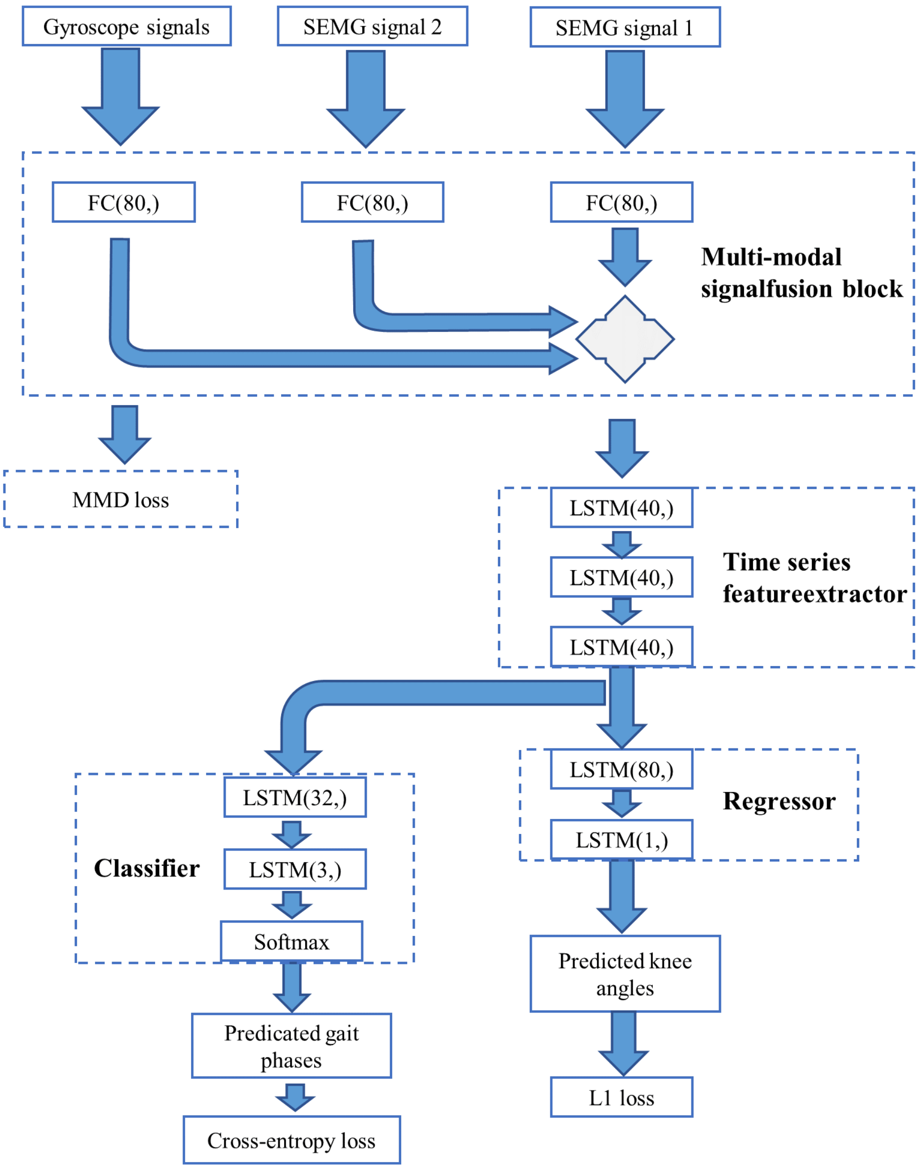


Figure 4. The visualization of the proposed TMMF architecture configured with LSTMs and FC layers. This figure was cited from [50]. LSTM, long short-term memory; FC, fully connected; TMMF, transferable multi-modal fusion.

were preprocessed to generate three-channel data inputs into the TMMF to predict knee angles and gait phase profiles. In particular, the TMMF consists of a multimodal signal fusion module (three FC layers (fully connected layers) and a cascade operation to extract distributional features), a time-series feature extractor (consisting of three LSTM layers, each containing 40 hidden units), a regressor (consisting of two FC layers), and a classifier (consisting of two FC layers and a softmax layer). The TMMF achieved a RMSE of 0.090 ± 0.022 for knee angle prediction, indicating that the predictions were very close to the actual values. In terms of gait phase prediction, TMMF achieved an accuracy of 83.7% ± 7.7%, showing high classification accuracy. Finally, the classifier was used to predict the corresponding gait phase.

some researchers have applied generative adversarial networks (GANs) to the field of gait prediction. Wang et al. enhanced human gait data through GANs and attention mechanisms, and input the collected real gait data and synthetic gait data into the two-stage attention model (LSTM based baseline prediction model and attention model) that was constructed for gait trajectory prediction [51]. The two-stage attention model performed better, with the attention mechanism showing a higher ability to learn the dependencies between historical gait data, which help to accurately predict the current values of hip angle and knee angle in the gait trajectory. The predicted gait trajectories based on historical gait data can be further used in gait trajectory tracking strategies. A study has described a method called GaIn to

Table 2. Related research based on deep learning algorithms

Researcher	Prediction Method	Predictive Accuracy
Su et al.[42]	LSTM	Correlation coefficient between predicted and measured trajectories: >0.98 Prediction accuracy for gait phases: 95%
Sun et al.[43]	LSTM	MSE for gait pattern prediction: <0.1 Average prediction accuracy: 97.23%
Moosavian et al.[44]	LSTM	Prediction accuracy for trajectories: ~98.5%
Lu et al.[45]	LSTM	Average accuracy in structured and over-ground environments: 85.6% and 73.3%
Ren et al.[46]	CNN-LSTM	Prediction of right knee, left knee, right hip, and left hip joint angles: Achieved low MSE and MAE (Mean Absolute Error) values
Zhu et al.[47]	CNN-LSTM	Prediction error: 1.34 ± 0.25 degrees Comparison with traditional machine learning methods (BP Neural Network, RF, SVR): The improved PCA algorithm outperformed traditional BP Neural Networks and Support Vector Machines in both convergence time and prediction accuracy.
Liu et al.[21]	CNN-LSTM Based on Attention Mechanism	Convergence time and prediction accuracy: Outperformed traditional BP Neural Networks and Support Vector Machines Generated gait trajectories: Exhibited lower MAE and higher Pearson Correlation Coefficient
Chereshnev et al.[52]	GaIn	Average prediction error: 4.55
Wang et al.[51]	GANs+LSTM+ Attention Mechanism	Hip joint angle prediction: RMSE: 0.0031–0.0129 MAE: 0.0027–0.0112 Knee joint angle prediction: RMSE: 0.0069–0.0280 MAE: 0.0051–0.0236
Guo et al.[50]	TMMF	Knee joint angle prediction: RMSE: 0.090 ± 0.022 Gait phase prediction: Accuracy: 83.7% ± 7.7%
Ren et al.[48]	TFSformer	Single-step prediction task: Average MSE: TFSformer reduced MSE by 10.83%, 15.04%, and 4.74% compared to CNN, Transformer, and CNN Transformer, respectively. Average MAE: TFSformer reduced MAE by 20.40%, 29.09%, and 10.43% compared to CNN, Transformer, and CNN Transformer, respectively. Multi-step prediction task: MAE: TFSformer achieved the lowest MAE.
Zhao et al.[49]	LSTM-Conformer	Average Accuracy: 94.89%

Note: LSTM, long short-term memory; CNN, Convolutional Neural Network; MSE, mean squared error; MAE, mean absolute error; RF, Random Forest; BP, Backpropagation; SVR: least squares support vector regression; PCA, principal component analysis; RMSE, root mean square errors; GANs, generative adversarial networks; TMMF, transferable multi-modal fusion; TFSformer, Transformer-Based Neural Network.

control non-invasive robotic prostheses and to predict leisure walking-related activities (such as walking, taking stairs, and running) through GaIn's thigh-based exercise [52]. The user's activity intention can be recognized by sEMG sensors, while in the gait inference component recurrent neural networks Recurrent Neural Networks and LSTMs are used. Based on the angular and angular velocity data of the thighs, the patterns learned by the network are used to infer the motion of the calves and feet. The system has high prediction accuracy, with an average prediction error for calf motion of 4.55. **Table 2** presents gait prediction studies related to deep learning algorithms.

Conclusion and future prospects

Summary

The wearable lower extremity exoskeleton robot shows great potential in the field of rehabilitation treatment and assisted walking. This article introduces some of the current cutting-edge robots (BLEEX, BaiHong, ReWalk, HAL, etc.), focusing on their assistance and rehabilitation functions in lower extremity exoskeleton robotics. Additionally, the acquisition methods of multi-modal data are introduced, including acquisition based on physical sensors and physiological electrical signal sensors. The fusion of multi-modal data provides a more comprehensive and accurate understanding of gait characteristics. Gait prediction has attracted much attention as a key technology for improving the control accuracy and adaptability of exoskeleton robots. Researchers have constructed biomechanical models by combining multimodal sensor data such as IMUs, plantar pressure sensors, and accelerometers, and have applied methods such as SVMs, RFs, BP neural networks, LSTMs, and CNNs, to achieve gait phase recognition and prediction. Future research directions include combining multimodal sensor data and advanced algorithms to improve the timeliness, accuracy, and robustness of prediction, while also developing personalized gait prediction methods. Gait prediction technology holds great potential in the fields of rehabilitation therapy, assisted walking, and sports training, driving the evolution of exoskeleton robots towards intelligence and personalization.

Challenges of gait prediction technology

The challenges of gait prediction techniques include the following:

1. Complexity of data acquisition and processing: Gait prediction needs to utilize multimodal sensor data, such as IMU, force sensitive resistor, and acceleration sensors, but the acquisition and processing of these data are more complex. The fusion, synchronization, and calibration of different sensor data are technical challenges, and the collected raw data often have noise and uncertainty, and requires effective signal processing and filtering to accurately reflect the user's gait information.

2. Selection and optimization of machine learning algorithms: Gait prediction involves various machine learning algorithms, such as SVM, RF, BP, probabilistic principal component analysis, and DGP. In practical applications, it is essential to select the appropriate algorithm based on the characteristics of the specific task and the nature of the data, followed by optimizing the algorithm's parameters to enhance both prediction accuracy and efficiency.

3. Handling of long-term dependency relationships: Gait, as a complex motor process, involves time series data with long-term dependency relationships. Traditional machine learning algorithms have limitations when dealing with such extended temporal sequences. Therefore, more specialized deep learning methods, such as LSTM networks and DNN, are required to better capture time-series information and improve prediction accuracy and robustness.

4. Implementation of personalized gait prediction: There are significant individual differences among users, including variations in body structure and gait habits. Accurately predicting personalized gait and adjusting prediction models based on individual differences are critical challenges that need to be addressed in future gait prediction technologies.

Future gait prediction technologies should focus on timeliness, accuracy, and personalization. By integrating multimodal sensor data with advanced machine learning algorithms, these technologies can further enhance the control precision and adaptability of lower limb exoskeleton robotics, providing greater support in fields such as rehabilitation, assisted walking, and movement training. The continuous development and refinement of gait prediction technology will open up broader prospects and opportunities for research and applications in the field of biomedical engineering.

Future research directions

Several key development directions are anticipated for future research in gait prediction for lower limb exoskeleton robots. First, integrating multimodal sensor data with advanced machine learning algorithms will significantly enhance the timeliness, accuracy, and robustness of gait prediction, thereby improving the control and adaptability of exoskeleton robots. Second, the study of personalized gait prediction methods is expected to become a focal point. By analyzing individual differences such as physiological characteristics and movement habits, more precise and adaptive models can be developed, leading to customized rehabilitation outcomes and optimized athletic performance. Finally, the application of gait prediction holds great promise in fields such as rehabilitation therapy, assistive walking, and athletic training, offering more effective treatment and support solutions for both rehabilitative and healthy populations, ultimately improving human health and quality of life.

Author Contributions:Haonan Geng was re-

sponsible for the conception, drafting, revision, and summarization of the entire manuscript. Xudong Guo was in charge of the conception and revision of the manuscript, providing guiding support, and notifying the status of the manuscript. Haibo Lin was tasked with the acquisition and integration of literature related to "multimodal information acquisition." Youguo Hao provided guidance on the manuscript and reviewed the content of the manuscript. Guojie Zhang offered relevant suggestions for the revision of the manuscript and reviewed the content of the manuscript.

References

[1] Nizamis K, Stienen AHA, Kamper DG, et al. Transferrable Expertise From Bionic Arms to Robotic Exoskeletons: Perspectives for Stroke and Duchenne Muscular Dystrophy. IEEE Trans Med Robot Bionics 2019;1(2):88-96.

[2] Gassert R, Dietz V. Rehabilitation robots for the treatment of sensorimotor deficits: a neurophysiological perspective. J Neuroeng Rehabil 2018;15:46.

[3] Kim YK, Lee DG, Ieee. Robot Aided Rehabilitation System for Gait Training of Stroke Patients. 9th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI) 2012;505-506.

[4] Pan LX. Early Prediction of Freezing of Gait in Parkinson's Disease Based on Deep Learning. School of Life Science and Technology 2023.

[5] Han JJ, Wang YH, Wan DQ. Research progress and development trend of lower extremity exoskeleton rehabilitation robot. J Shanghai Jiao Tong Univer (Med Sci). 2022;42(02):241-246.

[6] Zhang WR. Research on Wearable Lower Limb Exoskeleton Rehabilitation Robot. North China University of Science and Technology 2022.

[7] Fukuchi CA, Duarte M. A prediction method of speed-dependent walking patterns for healthy individuals. Gait & Posture 2019;68:280-284.

[8] Li YN, Huang GW, Xie LH. Gait trajectory prediction of lower limb rehabilitation exoskeleton based on deep learning. Rob Tech Appl 2024;(04):28-32.

[9] Su BT, Li JL, Xu HH, et al. Scientific training aid: Flexible wearable sensor motion monitoring applications. Scientia Sinica (Informationis) 2022;52(01):54-74.

[10] Zhou YK, ZengY, Li Y, et al. Design of exoskeleton robot based on advanced sensing and control technology. Technol Innovation Appl 2024;14(25):9-12.

[11] He CY, Zhang B, Zhou K, et al. Gait feature estimation method based on inertial sensor. Acta Armamentarii 2024;45(S1):200-208.

[12] Du YC, Yang LD, Wang XM, et al. Lower limb motion intention recognition based on FMG and IMU (Advance online publication). Inf Control 2024;1-12.

[13] Chang Y, Wang LAN, Li MIN, et al. RESEARCH ON GAIT RECOGNITION OF SURFACE EMG SIGNAL BASED ON MPSO-LSTM ALGORITHM. J Mech Med Biol 2023;23(08):2340066.

[14] Zeng DZ. Research on lower limb assisted exoskeleton based on multi-source signal gait recognition. South China University of Technology 2021.

[15] Sun XQ, Yu XD. A review of gait recognition systems for lower limb exoskeleton robots. AI-View 2024;(01):66-80.

[16] Zhang XD, Chen W, Sun YC, et al. Human gait recognition algorithm based on Tsfresh-RF feature extraction. Comput Syst Appl 2021;30(06):168-175.

[17] Yuan M, Wang X, Zhao Q, et al. Method for predicting human body gait track based on attention mechanism, involves using pre-processed original gait data to train prediction model to obtain prediction model, and performing human body track prediction based on prediction model. China patent CN116807458-A. 2023/10/30.

[18] Huang ZL, Fang CH, Ouyang XP, et al. Research on sensing system of lower limb exoskeleton robot based on multi-information fusion. Chin J Eng Des 2018;25(02):159-166.

[19] Qin T, Yang Y, Wen B, et al. Research on human gait prediction and recognition algorithm of lower limb-assisted exoskeleton robot. Intell Serv Rob 2021;14(3):445-457.

[20] Wei PN, Zhang JH, Tian FF, et al. A comparison of neural networks algorithms for EEG and sEMG features based gait phases recognition. Biomed Signal Process 2021;68:102587.

[21] Liu HJ, Zhu C, Zhou ZD, et al. Synergetic gait prediction and compliant control of SEA-driven knee exoskeleton for gait rehabilitation. Front Bioeng Biotechnol 2024;12:1358022.

[22] Zheng CK. Research on lower limb movement recognition based on eeg and emg signals. Southeast University 2021.

[23] Yao J. Gait recognition and prediction based on plantar pressure measurement. Southwest Jiaotong University 2017.

[24] Zhu MY. Research on gait prediction and recognition algorithm based on deep learning. Guizhou University 2024.

[25] Zhan HR, Kou JE, Cao YC, et al. Human Gait phases recognition based on multi-source data fusion and BILSTM attention neural network. Measurement 2024;238:115396.

[26] Wang L, Guo HX. Exploring Key Technologies of Multi-Sensor Data Fusion. 4th International Conference on Machinery, Materials and Information Technology Applications (ICMMITA) 2016;64-67.

[27] Jiao WC, Han S, Cui PZ, et al. The Research of Multi-sensor Data Fusion Technology. International Conference on Artificial Intelligence and Computer Science (AICS) 2016;294-299.

[28] Yan JJ, Xu YJ, Lin Y, et al. Gait phase prediction of exoskeleton robot enhanced by DGNN based on channel attention mechanism (Advance online publication). J East China Univ Sci Technol 2024;1-12.

[29] Miao X, Wu J. Gait recognition method based on deep learning integrates support vector machine (SVM), involves completing identity recognition by using SVM classifier, after feature vector is constructed by extracting required features from gait sequence. China patent CN114373091-A. 2022/06/05.

[30] Wu G, Wang C, Wu X, et al. Gait phase prediction for lower limb exoskeleton robots. 2016 IEEE International Conference on Information and Automation (ICIA) 2016;19-24.

[31] Luo WY. Human lower limb posture and motion state prediction. Huazhong University of Science and Technology 2017.

[32] Duan YK, Chen XG, Gui J, et al. Prediction of continuous lower limb motion based on phase division. J Zhejiang Univ (Eng Sci) 2021;55(01):89-95.

[33] Lu K, He ZQ. Lower limb exoskeletons gait recognition based on optimized random forest. Intell Comput Appl 2022;12(12):100-103.

[34] Ren SX, Wang WQ, Hou ZG, et al. Anthropometric Features Based Gait Pattern Prediction Using Random Forest for Patient-Specific Gait Training. 25th International Conference on Neural Information Processing (ICONIP) 2018;15-26.

[35] Lu K. Gait prediction for lower limb exoskeleton robots oriented to hemiplegic rehabilitation. Guizhou University 2023.

[36] Niu ZY, Liu H, Cheng HS, et al. Gait Prediction of Swing Phase Based on Plantar Pressure. IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM) 2020;91-96.

[37] Ma XM. Research on Gait Recognition and Prediction of Exoskeleton Robot Based on DTW Template Matching. Wuhan University of Technology 2019.

[38] Tanghe K, De Groote F, Lefeber D, et al. Gait Trajectory and Event Prediction from State Estimation for Exoskeletons During Gait. IEEE Trans Neural Syst Rehabil Eng 2020;28(1):211-220.

[39] Chen ZL, Guo Q, Li TS, et al. Gait Prediction and Variable Admittance Control for Lower Limb Exoskeleton With Measurement Delay and Extended-State-Observer. IEEE Trans Neural Netw Learn Syst 2023;34(11):8693-8706.

[40] Shen F. Research on human gait parameter prediction algorithm based on LASSO regression. Shenzhen University 2020.

[41] Zhou Q. Human Gait Prediction and Recognition Based on Deep Learning and Inertial Data. Anhui University of Technology 2021.

[42] Su BB, Gutierrez-Farewik EM. Gait Trajectory and Gait Phase Prediction Based on an LSTM Network. Sensors 2020;20(24):7127.

[43] Sun L, Wang H, Lu RD. Assistive Force Cycle Prediction Method for Lower Limb Exoskeleton Based on Independent Attitude Processing Sensor. IEEE Sens J 2024;24(7):9562-9572.

[44] Moosavian SAA, Kiani A, Akbari V, et al. RoboWalk Trajectory Planning Based on the Human Gait Prediction Using LSTM. 2021 9th RSI International Conference on Robotics and Mechatronics (ICRoM) 2021;433-438.

[45] Lu LJ, Liu SY, Zhou ZY, et al. Continuous Gait Phase Recognition and Prediction Using LSTM Network Across Walking Speeds and Inclinations. 7th IEEE International Conference on Advanced Robotics and Mechatronics 2022;101-106.

[46] Ren J, Wang A, Ma Z, et al. Gait Prediction for Rehabilitation Robots Based on Deep Learning. 2022 International Conference on Advanced Mechatronic Systems (ICAMechS) 2022;86-90.

[47] Zhu M, Guan XR, Li Z, et al. sEMG-Based Lower Limb Motion Prediction Using CNN-LSTM with Improved PCA Optimization Algorithm. Journal of Bionic Eng 2023;20(2):612-627.

[48] Ren JL, Wang AH, Li HY, et al. A Transformer-Based Neural Network for Gait Prediction in Lower Limb Exoskeleton Robots Using Plantar Force. Sensors 2023;23(14):6547.

[49] Zhan ZJ, Wang T, Bei TX, et al. Gait prediction method of lower limb exoskeleton based on LSTM and Conformer neural network. Comput ERA 2022;(08):1-5.

[50] Guo ZP, Zheng HX, Wu HR, et al. Transferable multi-modal fusion in knee angles and gait phases for their continuous prediction. J Neural Eng 2023;20(3).

[51] Wang Y, Li ZK, Wang X, et al. Human Gait Data Augmentation and Trajectory Prediction for Lower-Limb Rehabilitation Robot Control Using GANs and Attention Mechanism. Machines 2021;9(12):367.

[52] Chereshnev R, Kertész-Farkas A. GaIn: Human Gait Inference for Lower Limbic Prostheses for Patients Suffering from Double Trans-Femoral Amputation. Sensors 2018;18(12):4146.

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Review Article https://doi.org/10.61189/403384jfzmyx

Application and progress of functionalized magnetic bead-based biosensors for protein detection

Haoyuan Su[1], Yuehua Liao[2], Shu Wu[1], Jun Ji[1], Shuya An[1], Dongdong Zeng[2]

[1]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China. [2]School of Medical device, Shanghai University of Medicine & Health Sciences, Shanghai 201318, China.

Corresponding author: Dongdong Zeng.

Declaration of conflict of interest: None.

Received September 25, 2024; Accepted November 6, 2024; Published December 31, 2024

Highlights

- In the field of bioanalysis, a new biosensor technology based on functionalized magnetic beads is leading a new direction in protein detection. With its excellent separation efficiency and sensitivity, it provides a powerful tool for early disease diagnosis and biomarker monitoring.
- This article explores the latest advancements in this technology, including innovative magnetic bead designs, diverse detection strategies, and the technical challenges and future development directions. It reveals the poten tial and application prospects of biosensor technology in biomarker detection.

Abstract

In the field of bioanalysis, the integration of magnetic beads and biosensors provides a protein detection platform with high separation efficiency and sensitivity. The superparamagnetism of magnetic beads, combined with surface functional modifications, forms the basis for selectively capturing and effectively separating target proteins. Additionally, the high sensitivity and specificity of biosensors ensure precise quantitative analysis of captured proteins. This article systematically reviews the synthesis strategies of functionalized magnetic beads, detection methods for proteins and nucleic acids, as well as the current technical challenges and future development directions.

Keywords: Protein, magnetic beads, biosensor

Introduction

Protein biomarkers play a critical role in medical research and clinical diagnostics. They are indicators that can reflect biological states or conditions, capable of detecting biological activities and processes [1-4]. The significance of protein biomarkers is evident in areas such as disease diagnosis, monitoring, and personalized medicine. For example, the abnormal expression of Programmed death ligand-1 (PD-L1) protein is associated with the prognosis and clinical pathological features of non-small cell lung cancer; abnormal deposition of amyloid beta protein, forming amyloid beta plaques, is one of the main pathological features of Alzheimer's disease [5, 6]. Such protein biomarkers can be detected through traditional methods like Enzyme-Linked Immunosorbent Assay (ELISA) and Polymerase Chain Reaction (PCR). However, in early disease diagnosis, pathological changes in protein biomarkers may not be significant enough, which can lead to undetectable increases in protein expression. These traditional methods have long detection cycles, complex operations, and limited sensitivity and specificity, which are insufficient for rapid and precise detection [7]. Biosensors, with high sensitivity, selectivity, and the ability for quick detection, can better address these issues.

A biosensor is an analytical tool consisting of a biological recognition element (such as enzymes, antibodies, and nucleic acids) and a signal transducer, capable of converting the recognition of biological molecules into measurable

Address correspondence to: Dongdong Zeng, School of Medical device, Shanghai University of Medicine & Health Sciences, No. 268 Zhouzhu Highway, Pudong, Shanghai 201318, China. E-mail: zengdd@sumhs.edu.cn.

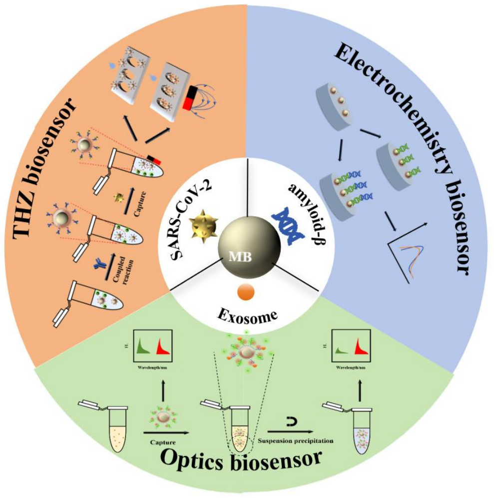


Figure 1. Detection of target proteins using functional magnetic beads combined with various biosensors.

signals. Therefore, biosensors can specifically identify target molecules, improving detection specificity, and, due to the presence of transducers and signal amplifiers, they can detect extremely low concentrations of biomarkers, enhancing sensitivity. Specifically, the working principle of a biosensor is based on the recognition and binding of a specific target molecule by the biometric element. This binding alters the physical or chemical properties of the biosensor, such as conductivity, light absorption, or mass change. These changes are converted by the signal transducer into measurable electrical or optical signals, which are then used to detect the presence and quantity of the target based on signal changes and their intensity [8-11]. Common biosensors include electrochemical biosensors, optical biosensors, terahertz metamaterial biosensors, and giant magnetoresistance biosensors, etc.

Magnetic beads are superparamagnetic microparticles with particle size distribution ranging from micrometers to nanometers. They can play multiple roles and perform various key functions in biosensor systems. For example, due to their excellent molecular dynamics and the inherent magnetism of the magnetic bead material, they can achieve precise magnetic separation, targeted positioning, and delivery through the application of an external magnetic field, allowing for the rapid and efficient separation of biomolecules from other components in a solution and signal modulation for signal transduction via their magnetic properties [12]. The core of the magnetic beads is mainly composed of magnetite (Fe_3O_4), while the outer material may include silica, agarose, or polymers, which minimizes their biological toxicity, making them suitable for binding various biomolecules [13, 14]. Additionally, the small size and large specific surface area of the magnetic beads allow for the binding of more biomolecules, enhancing the sensitivity of the biosensor. This makes them suitable for detecting trace amounts of samples, reducing reagent consumption, and promoting the development of portable and automated detection systems. Surface modification techniques are used to introduce various functional groups, such as carboxyl, amino, biotin, and avidin, to the outer layer of the beads, serving as solid-phase carriers to provide stable interfaces for biomolecules [15, 16].

This article explores the latest advancements in magnetic bead-based biosensor detection of proteins, including innovative magnetic bead designs, diverse detection strategies, and current challenges and future directions. It reveals the potential and application prospects of biosensor technology in biomarker detection (**Figure 1**).

Functionalized magnetic beads

Functionalized magnetic beads offer unique structures, controllable preparation methods, and diverse functional modifications. By selecting appropriate magnetic core materials, coating layers, and functionalization techniques, efficient capture and detection of target molecules can be achieved, providing a powerful tool for bioanalysis.

Structure

Magnetic beads can be classified into three types based on their physical composition: core-shell structure, hollow structure, and composite structure. Among them, the core-shell structure is the most common and widely used, which uses magnetic nanoparticles as the core, with outer layers typically made of materials such as silica, polymers, or gold. This core-shell structure significantly enhances the chemical stability of the beads and provides a larger specific surface area for further functional modifications. By directly modifying functional groups or biomolecules on the shell, a monolayer surface modification can be achieved. This simple monolayer structure is suitable for the rapid capture and separation of biomolecules. For example, Mao et al. used a supramolecular layer-by-layer self-assembly strategy to construct immunomagnetic beads, which significantly increased the number of functional sites on the beads, improving both capture efficiency and specificity [17].

Hollow structure magnetic beads have high specific surface areas, light weight, good magnetic responsiveness, and strong drug loading capacity. Their internal cavities can be used to load drugs or other active substances, with adjustable pore sizes. Ding et al. developed a hollow cellulose/carbon nanotube composite microbead with a patterned porous structure [18]. As the concentration of carbon nanotubes and cellulose increases, the diameter of the composite magnetic beads grow, effectively adsorbing methylene blue. Composite structure magnetic beads combine the advantages of different materials, offering multifunctionality and high efficiency. Their design enables the integration of various functions such as magnetism, fluorescence, and catalysis, providing broader application possibilities. The composite structure enhances the stability and biocompatibility of magnetic beads, improving their targeting and specific capture ability. Chen et al. assembled signal probes based on gold nanoparticles onto the surface of magnetic beads by the specific binding between the target and the aptamer, creating functional composite structure magnetic beads for colorimetric detection [19].

In target molecule capture, efficient separation methods are crucial for improving capture efficiency. To achieve separation between the target and biological samples, several magnetic separation methods have been developed based on the inherent magnetism of magnetic beads, all of which have shown promising results. These methods can be categorized into dynamic magnetic separation, static magnetic separation, gradient magnetic separation, automatic magnetic separation, magnetic fluidized bed, and magnetic microsphere aggregation, depending on the application of the magnetic field and the separation mechanism. These methods are selected based on sample size and experimental design, and they have been proven to enhance detection efficiency and accuracy in nucleic acid and protein detection [20-22].

Preparation methods

In recent years, core-shell structure magnetic beads have been most used in biological detection processes, such as nucleic acid and protein assays [23]. Magnetite (Fe_3O_4) is often chosen as the core material due to its simple synthesis, superparamagnetism, controllable magnetization, and low biological toxicity. In protein and nucleic acid detection, the stability and dispersion of magnetic beads are crucial, necessitating the modification and protection of magnetic core material. This can be achieved by using polymer stabilizers or coatings of inert metals. Additionally, functionalized magnetic beads should be equipped with target-specific ligands or functional groups on their surface to ensure stronger binding forces during target capture, improving the success rate and speed of the binding process [24].

The preparation methods for superparamagnetic beads mainly include physical, chemical, and microbial methods. Physical methods often require harsh conditions and are difficult to control the size and shape of the beads [25]. Microbial methods are hindered by long preparation times and low productivity, making them less suitable as high-priority methods for magnetic bead preparation. However, chemical methods can overcome these problems and produce more uniform magnetic beads with consistent size and shape due to their high controllability, making it more suitable for large-scale production [26]. The most common chemical synthesis methods include co-precipitation, microemulsion, and hydrothermal methods.

Co-precipitation is a commonly used method for preparing micron to nanometer-sized magnetic beads. This method involves adding a precipitant to a solution containing soluble salts, forming insoluble hydroxides or oxides, which are then decomposed by heating to produce nanomaterials. This method is simple, cost-effective, and allows for control over the morphology and size of magnetic nanoparticles by adjusting factors such as the solution pH and reaction time. While it is possible to control particle size by regulating conditions, achieving precise control may be more difficult compared to other methods, such as the microemulsion method. Additionally, without proper stabilizers, the particles produced may agglomerate, negatively affecting their performance in biomedical applications [27, 28].

Therefore, the preparation of magnetic beads using the microemulsion method has been a focus in the field of biomaterials. Zhi et al. proposed a new method for the in-situ preparation of magnetic chitosan/Fe_3O_4 composite nanoparticles in microaqueous pools of oil-in-water microemulsions by adding NaOH as an alkaline precipitant and using the oil-in-water microemulsion containing chitosan and ferrous salts as a microreactor [29]. Marija et al. used a microemulsion system with cetyltrimethylammonium bromide as the surfactant, butanol as the auxiliary surfactant, hexanol as the oil phase, and an aqueous solution as the water phase to prepare magnetic beads [30]. The microemulsion method allows for precise control over the size and shape of magnetic beads, as the size and morphology of the microemulsion droplets directly affect the properties of the resulting nanoparticles. Therefore, magnetic beads prepared using the microemulsion method typically exhibit high monodispersity, and this method allows for direct surface modification during particle formation, which is beneficial for improving the biocompatibility and functionality of the magnetic beads. However, the process is more complex and requires precise control of reaction conditions. Additionally, beads prepared by the microemulsion method may require extra purification steps to remove residual surfactants and other organic solvents, which increases the overall cost. Nonetheless, microemulsion method is a versatile and effective technique for synthesizing monodisperse magnetic beads with controlled size, microstructure, and properties [29].

The hydrothermal method is a more classical method for synthesizing magnetic nanoparticles. Typically carried out in a high-pressure autoclave, this method uses water or alcohol as solvents, transforming iron precursors into magnetic nanoparticles under elevated temperature and pressure. This method is simple to operate, easy for mass production, and produces magnetic nanoparticles with good dispersion, uniform particle size, and complete crystal structure. Hemauer et al. synthesized Fe_3O_4 nanoparticles using the hydrothermal method, employing various crystal growth control agents to control the morphology of iron oxide nanoparticles [31].

Functional modification

Surface functionalization and modification of magnetic beads is a key aspect in various biomedical applications. Magnetic beads have gained attention for their diagnostic and therapeutic capabilities, allowing for simultaneous diagnosis and treatment of diseases [32]. Functionalized magnetic beads typically consist of a magnetic core and a surface functional layer. The magnetic core is primarily made of magnetic materials such as iron oxide, while the surface functional layer is modified through chemical or biological methods to impart specific biological recognition abilities to the beads. For example, antibodies, nucleic acids, enzymes, and other biomolecules can be immobilized on the surface of magnetic beads by covalent bonding, physical adsorption, or bio-affinity. Under the influence of an external magnetic field, the functionalized beads can rapidly bind to and separate target molecules. By employing appropriate functional groups and surface modification techniques, magnetic beads can be customized with various biometric elements or functional groups tailored to specific detection needs, significantly enhancing their capabilities [28]. There are various types of functionalized magnetic beads, commonly categorized based on their surface modifications. For example, carboxyl-functionalized magnetic beads are suitable for chemically cross-linking and immobilizing proteins or other biomolecules; amino-functionalized magnetic beads can undergo various chemical reactions for further functionalization; streptavidin-functionalized magnetic beads specifically bind biotin-labeled molecules; and antibody-functionalized magnetic beads are used to capture corresponding antigens [33-36]. In practical applications, such as in Chen et al.'s research, multifunctional carboxylated magnetic beads were prepared, allowing bovine serum albumin to be immobilized on the particle surface [37]. Overall, the surface functionalization of magnetic beads plays a pivotal role in enhancing their performance in various biomedical applications, ranging from drug delivery systems to diagnostic tools. With the fast and specific binding of streptavidin and biotin, coupled with their strong binding affinity, researchers have conducted extensive studies. For example, Zhao et al. used streptavidin-modified magnetic beads to specifically capture biotin-containing peptides [38]. In the field of antibody magnetic beads, Wang et al. developed a high-sensitivity sandwich assay method using antibody magnetic beads and aptamers to detect C-reactive protein (CRP) [39]. In conclusion, these functionalized magnetic beads, due to their unique surface chemistry and biocompatibility, provide diverse tools and methods for protein detection in biomedical research, with immense potential to drive technological advancements in the field.

Protein detection strategies based on magnetic beads and biosensors

In the ever-evolving field of biochemical diagnostics, magnetic bead-based detection technology has become a promising frontier. These technologies offer a range of advantages that redefine the possibilities for biomarker detection. At the core of their effectiveness are features such as a high surface-to-volume ratio, rapid response, and high throughput, positioning these technologies at the forefront of cancer research. The ability to achieve simultaneous multi-channel detection, automate the processing of large sample volumes, and maintain high sensitivity and specificity has made magnetic bead-based biosensors indispensable tools in both research and clinical settings. In protein detection, biosensors are of particular interest due to their superior accuracy and sensitivity. In recent years, with the rapid development of nanotechnology, magnetic beads, as excellent biomaterials, have been combined with biosensors to give rise to various efficient protein detection strategies. In the following section, some of the popular methods are introduced.

Electrochemical biosensors

The principle of this method involves transferring magnetic beads that are bound to the target protein onto the surface of an electrochemical sensor. Electrochemical techniques, such as cyclic voltammetry and differential pulse voltammetry, are used to detect the presence and concentration of the target protein. The intensity of the electrochemical signal is proportional to the concentration of the target protein, allowing for quantitative analysis through calibration curves. Compared to traditional methods, electrochemical sensor-based methods offer advantages in terms of detection speed, sensitivity, specificity, and ease of operation. Specifically, methods based on magnetic beads and electrochemical biosensors enable rapid detection, often completing the process within minutes to hours. This is because magnetic beads can quickly capture the target protein, and electrochemical biosensors can monitor the reaction signals in real-time. In contrast, traditional detection methods such as ELISA and Western blot typically require longer times for sample preparation, reaction, and detection, often taking hours to days.

In terms of sensitivity and specificity, the method based on magnetic beads and electrochemical biosensors offers high sensitivity and specificity, allowing for the detection of low concentrations of target proteins. The high specific surface area of the functionalized magnetic beads enhances the binding efficiency to the target protein, while electrochemical biosensors can detect even the slightest changes in electrochemical signals, enabling the detection of trace proteins. On the other hand, traditional methods often have lower sensitivity and specificity, particularly when detecting low-concentration proteins, which may require more complex sample processing and signal amplification steps, thus increasing detection difficulty.

Methods based on magnetic beads and electrochemical biosensors generally have a higher degree of automation, are easy to operate, and require fewer sample processing steps, with no need for complex equipment. Due to the magnetic properties of the beads, the separation and enrichment of the target protein can be controlled by a magnetic field, reducing the need for time-consuming procedures like centrifugation or filtration. In contrast, traditional methods often involve more complicated steps, especially when multiple sample processing and detection stages are involved, requiring skilled technicians and more reagents and consumables.

Based on these characteristics, Zhang et al. developed an electrochemical biosensor for the sensitive detection of human epidermal growth factor receptor 2 (HER2) protein by utilizing DNA aptamers and peptide nucleic acids as recognition probes, combined with magnetic FeO/α-FeO@Au nanocomposites [40]. The aptamers captured large HER2 protein molecules and released single-stranded DNA, resulting in sensitive changes in the electrochemical signal. Peptide nucleic acids then captured the released single-stranded DNA chain, converting the electrochemical signal changes induced by HER2 into the changes driven by the quantity of short-chain single-stranded DNA, thereby expanding the detection range.

Roy et al. developed a compact, yet highly sensitive electrochemical sensor based on micropores [41]. This sensor features a densely packed microelectrode array, which utilizes the structure to precisely control the position and quantity of magnetic beads on the working and counter electrodes, resulting in different bead distributions on the working electrode. This configuration reduced variations in the biosensor signal, enabling stable and highly sensitive detection of amyloid proteins.

Gutierrez-Galvez focused on the study of exhaled gas electrochemical sensing and developed an electrochemical immunosensor for detecting the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) (**Figure 2**) [42]. This figure was cited from [42].The sensor consists of two parts: an exhaled gas collection module and a detection module. This immunosensor demonstrated high selectivity for both the SARS-CoV-2 spike protein and the SARS-CoV-2 nucleocapsid protein.

Optical biosensors

Optical biosensors achieve their detection functionality by interacting an optical field with biological recognition components. They are one of the most widely used biosensors due to their minimal sample requirements, high sensitivity,

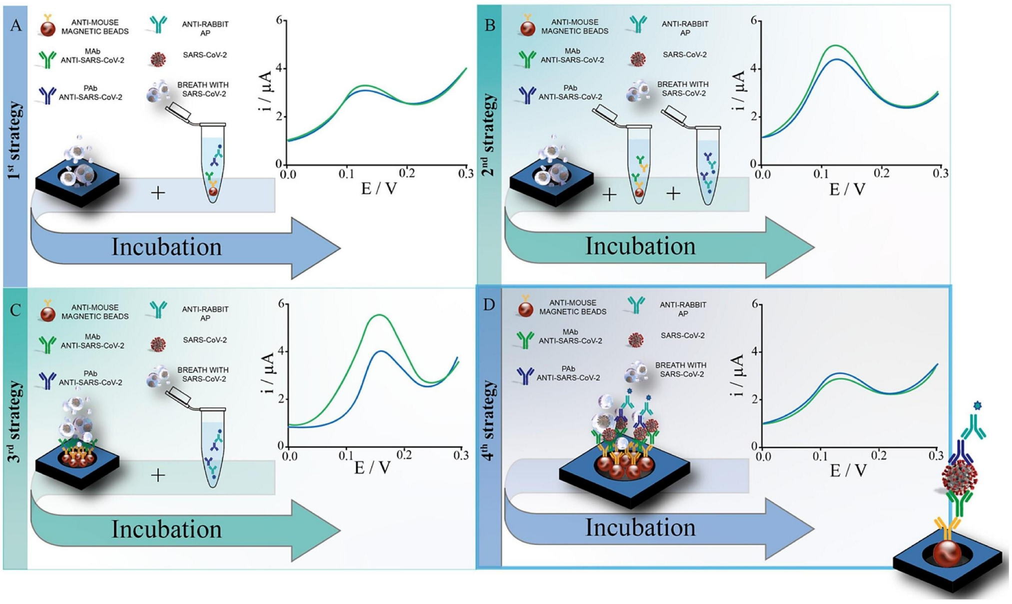


Figure 2. An electrochemical biosensor based on magnetic beads for the detection of SARS-Cov-2. This figure was cited from [42].

and specificity. Generally, optical biosensors can be divided into two categories: label-free and labeled. Label-free sensing means that the signal is directly generated by the interaction between the sample and the sensor. In contrast, labeled methods requires the target analyte to be tagged with a reporter molecule to enable detection via fluorescence, chemiluminescence, or colorimetric signals [43].

Luo et al. developed an integrated magnetic-fluorescent exosome nanosensor for the rapid and sensitive detection of cancer-derived proteins [44]. First, magnetic beads were coated with DNA tetrahedron lipid probes to effectively capture exosomes. A well-designed bifunctional aptamer specifically recognized the exosome PD-L1 protein, initiating catalytic hairpin assembly, which converted the protein signal into a double-strand signal. Finally, the double-strand body activated CRISPR-associated protein 12a, which cleaved substrate reporter molecules, amplifying the fluorescence signal. By integrating magnetic separation and fluorescence signal amplification, this nanosensor exhibited excellent specificity and sensitivity for PD-L1-positive exosomes.

Fluorescence-based optical biosensor detection is another highly efficient strategy. Wu et al. proposed a ratio-fluorescence method for one-step, label-free protein detection [45]. The method uses a fixed protein-reactive dye, Cy5-labeled Cluster of Differentiation 63 hairpin aptamer (Cy5-Apt) on the surface, along with a dual-color streptavidin magnetic bead to recognize the protein. The Cy5-Apt stem is embedded with the green SYBR Green I to detect the protein. After successful capture, the Cy5-Apt undergoes a conformational change and releases the encapsulated SYBR Green I, allowing the exosome to be measured based on the fluorescence ratio of Cy5 and SYBR Green I.

Additionally, surface plasmon resonance (SPR) is an optical phenomenon occurring on a metal surface. When light hits the metal surface, free electrons in the metal interact with photons, generating surface plasmon waves, which are electromagnetic waves propagating along the metal-medium interface. SPR occurs when the propagation speed of the surface plasmon wave matches the speed of the incident light. This phenomenon can also be combined with magnetic beads to construct optical biosensors for protein detection. For example, Sun et al. established an SPR biosensor system for immunoassays, where magnetic microbead conjugates coupled with antibodies were captured firmly on the Au film by magnetic forces [46]. They used magnetic microbeads as solid supports for Heat Shock Protein 70 (Hsp 70) antibodies, with antibody-functionalized magnetic microbeads replacing single antibody assays to detect Hsp 70.

Wang et al. developed a grating-coupled surface plasmon resonance sensor by combining

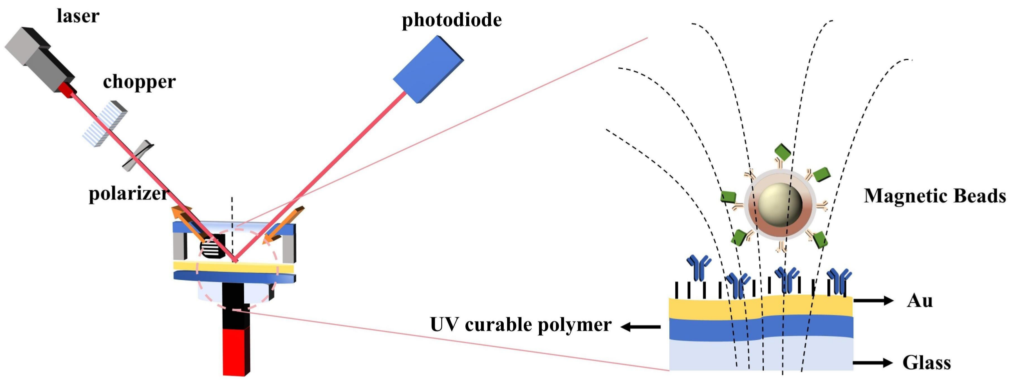


Figure 3. Plasmon resonance biosensor based on magnetic bead. This figure was cited from [47].

SPR biosensor technology with magnetic bead assays (**Figure 3**) [47]. This figure was cited from [47]. This system utilizes a metal diffraction grating sensor chip, to recognize target analytes after functionalization with specific antibodies. Magnetic beads, coupled with specific antibodies, not only serve as signal amplifying markers to enhance refractive index changes caused by target molecule capture but also act as carriers to rapidly deliver analytes to the sensor surface. The efficient collection of magnetic beads was achieved by applying a magnetic field gradient to the grating-coupled surface plasmon resonance sensor chip. The role of the movement speed of magnetic beads in the presence of the magnetic field in enhancing the sensitivity of grating-coupled surface plasmon resonance biosensors was investigated. Additionally, the validity of the technique was confirmed in an immunoassay modeled on β-Human Chorionic Gonadotropin.

Terahertz metamaterial biosensors

Terahertz (THz) metamaterial biosensors combine the detection principles of biological magnetic beads with the strong localized field enhancement effects of THz metamaterials. By integrating the high specific surface area and magnetic separation characteristics of biological magnetic beads, these biosensors enable high-sensitivity detection of biomolecules. In the detection process, biological magnetic beads, acting as solid-phase carriers, can capture and enrich target biomolecules through functional modifications on their surface, such as the attachment of antibodies or specific molecules. When these magnetic beads are combined with THz metamaterials, the localized field enhancement effect of the THz waves significantly boosts the sensitivity for detecting changes in the biomolecules on the magnetic bead surface [48]. Furthermore, under the influence of an external magnetic field, the magnetic beads can quickly separate from the solution, simplifying the operation process and shortening the reaction time.

For instance, Bi et al. proposed a THz metamaterial biosensor for detecting the SARS-CoV-2 spike protein antigen [49]. By utilizing the positioning of magnetic beads in the sensor (resonant metamaterial unit cell), the target molecules induced response signals within the domain. The sensor successfully detected trace concentrations of the SARS-CoV-2 spike protein antigen, with a detection range of 0.005 ng/mL to 1000 ng/mL and a detection limit of 0.002 ng/mL. However, challenges remain, including technical complexity, sample preparation and processing issues, biocompatibility and stability concerns, as well as data standardization and analysis.

To address these challenges, researchers are working on optimizing design and manufacturing processes to reduce costs and enhance the accessibility of THz metamaterials. They are also applying advanced signal processing algorithms and machine learning techniques to improve the analytical capabilities and accuracy of the signals. Additionally, efforts are focused on standardizing sample processing protocols to reduce variability during sample preparation and improve reproducibility, conducting more biocompatibility studies to ensure the safety of THz metamaterials and magnetic beads in biomedical applications, and developing validated data analysis workflows to improve the accuracy and reliability of the detection results [50].

Giant magnetoresistance biosensors

Giant magnetoresistance (GMR) biosensors combine the detection principles of biological magnetic beads with the giant magnetoresis-

Table 1. Detection performance of various magnetic bead-based biosensors

Sensor type	Magnetic bead type	Test object	LOD
Zhang's Electrochemistry [40]	Gold nanocomposite magnetic beads	HER2	4.1 fg/mL
Roy's Electrochemistry [41]	Antibody magnetic bead	Amyloid proteinβ	<10 fg/mL
Gutierrez-Galvez's Electrochemistry [42]	Antibody magnetic bead	SARS-CoV-2 Spike protein	1 ng/mL
Luo's Optics [44]	DNA tetrahedral functionalized magnetic beads	PD-L1	1.71×10^3particle/μL
Wu's Optics [45]	Fluorescent magnetic bead	Exosome	4.0×10^4 particle/μL
Sun's Optics [46]	Antibody magnetic bead	Hsp 70	0.30 μg/mL
Wang's Optics [47]	Antibody magnetic bead	βHuman chorionic gonadotropin	0.45 pM
Bi's Terahertz [49]	Gold nano antibody composite magnetic beads	SARS-CoV-2	0.002 ng/mL
Gupta's GMR [53]	Antibody magnetic bead	ESAT-6	12 pg/mL
Ding's Electrochemistry [54]	DNA functionalized magnetic beads	protamine	80 pM

Note: LOD, limit of detection; HER2, Human Epidermal growth factor Receptor 2; PD-L1, programmed death-ligand 1; Hsp 70, Heat Shock Protein 70; SARS-CoV-2, Severe Acute Respiratory Syndrome Coronavirus 2; ESAT-6, Early Secretory Antigenic Target 6 kDa.

tance effect, where the electrical resistance of a material changes significantly as its magnetization state varies. These sensors consist of immune magnetic beads, a highly magnetic-sensitive GMR sensor, and associated readout circuitry.

In the detection process, a biological probe specific to the target molecule is first equipped on the surface of the GMR sensor. The sample containing the target molecules is then passed over the sensor surface, where the target molecules are captured by the probe. Immune magnetic microspheres are added, which interact with the target molecules to complete the labeling process. An external gradient magnetic field is applied to separate the unmarked excess immune magnetic microspheres, reducing background noise and improving detection accuracy. Finally, an alternating magnetic field is applied to magnetize the magnetic labels. The resulting alternating magnetic field generated by the magnetic labels causes changes in the sensor's magnetoresistance, which are then read out to determine the presence and concentration of the target molecules in the sample [51, 52].

For example, Gupta et al. proposed a highly sensitive and specific technique based on the principle of GMR sensors for early tuberculosis diagnosis [53]. This GMR biosensor assay used monoclonal antibodies specific to the early secreted antigenic target-6 (ESAT-6) secreted by M. tuberculosis antigen, with magnetic beads serving as the labels. In the presence of ESAT-6, the magnetic beads bind to the GMR sensor, and the binding is proportional to the concentration of the ESAT-6 protein. This causes a change in the overall resistance of the GMR sensor, and the resistance variation corresponds to the ESAT-6 concentration, achieving the detection goal.

Trends and challenges in protein detection based on magnetic beads and biosensors

As discussed above, the combination of magnetic beads and biosensors for protein detection is an efficient analytical method, which leverages the rapid separation capabilities of magnetic beads and the high-sensitivity detection characteristics of biosensors (**Table 1**). This technology is advancing toward higher sensitivity, faster detection speeds, and improved user-friendliness. Below are some key trends and challenges in this field.

Improving sensitivity and selectivity

Background noise is one of the main factors that affect the sensitivity of protein detection. For instance, when fluorescent probes are used to label target analytes, low concentrations of the target molecule often result in very weak signal, which can be easily obscured by background noise. In fluorescence-based analyses, common sources of background noise include Raman scattering from the solvent, residual fluorescence from unbound fluorescent molecules, and autofluorescence from the capture surface. Although washing steps can help reduce background noise, they add complexity and time to the detection process [55].

Additionally, some magnetic bead and biosensor systems may have limited dynamic detection ranges, which restricts their ability to detect proteins over a wide concentration range. Developing sensors with a broader dynamic range is crucial for enhancing their clinical applicability.

In terms of selectivity, due to the large specific surface area and numerous active sites of magnetic beads, they tend to have some degree of adsorption to biomolecules. The non-specific binding of magnetic beads to biomolecules can lead to false positive signals, affecting the accuracy of detection [56]. To reduce non-specific binding, the surface of the magnetic beads can be appropriately functionalized to enhance the specificity of binding between the bead surface and the target protein, thereby minimizing the occurrence of false positive signals.

Additionally, in practical operations, most researchers use blocking strategies for the active sites to reduce the generation of false positive signals. This is typically done by adding excess target protein aptamers or sufficient bovine serum albumin to block the non-specific binding sites on the magnetic beads.

Multi-target detection capability

In the field of protein detection using magnetic beads and biosensors, despite rapid technological advancements, there are still challenges and limitations in achieving multi-target detection. These challenges are particularly evident in areas such as the selection of biological recognition elements, signal amplification and detection technologies, as well as sample separation and processing techniques [57].

First, the complexity of multi-target detection is a key issue. Interactions between different proteins can lead to cross-reactivity, which may compromise the accuracy of the detection results. For example, the concentration of certain disease biomarkers may be much lower than that of more common proteins, requiring extremely high sensitivity and a broad dynamic range of the sensors. Furthermore, the selectivity of the biological recognition elements is another important consideration. It's essential to choose an appropriate recognition element for each target protein, especially when the target proteins are highly similar, which remains challenging in practical applications [58]. To address these challenges, one approach is to develop and screen highly specific antibodies or aptamers to ensure accurate recognition of the target proteins. Additionally, bioinformatics methods can be employed to predict and design biological recognition elements that specifically bind to the target proteins, enhancing the precision of multi-target detection [49].

Furthermore, sample processing and separation techniques are necessary steps before conducting multi-target detection to reduce interference from sample matrices. Magnetic beads are widely used in sample preparation due to their high separation efficiency, but in multi-target detection, effectively separating and enriching different targets remains a challenge. Additionally, data interpretation and analysis also present difficulties in multi-target detection. Detection generates large amounts of data, which requires sophisticated data processing and analysis methods to distinguish and interpret signals from different targets [59]. To address these challenges, researchers have implemented strategies such as using nanomaterials to enhance signals and developing new signal transduction methods. Specifically, the optical or electrochemical properties of nanomaterials, such as gold nanoparticles and quantum dots, are utilized to amplify signals, while enzyme-catalyzed reactions or the plasmonic resonance effects of nanostructures are employed to further enhance signals [49].

In conclusion, while magnetic beads and biosensor technology show great potential in protein detection, achieving effective multi-target detection still requires overcoming the challenges outlined above. Future research needs to focus on improving the selectivity of biological recognition elements, developing new signal amplification strategies, refining sample processing and separation techniques, and advancing data processing methods to enable more efficient and accurate multi-target detection.

Portability and real-time monitoring

With the development of wearable devices and mobile health technologies, portable and real-time protein detection techniques have garnered increasing attention. This requires magnetic bead and biosensor technologies to not only offer high sensitivity but also be portable and easy to operate, for better meeting the demands of various monitoring environments. For example, wearable devices can monitor proteins in sweat, saliva, or tears in real time, providing dynamic data for disease diagnosis and health management [60]. Furthermore, the ability to monitor protein biomarkers in real time will significantly enhance the level of precision medicine. Current technologies, such as ELISA and Lateral Flow Immunoassays, can only provide single-time measurements with time delays, limiting timely responses to rapidly changing life-threatening conditions [61].

However, achieving multi-target detection is still challenging. For example, interactions between different proteins may lead to cross-reactivity, affecting the accuracy of the detection results. Additionally, sample preparation and separation are necessary steps before multi-target detection to reduce interference from sample matrices. Magnetic beads are widely used in sample pre-processing due to their high separation efficiency, but in multi-target detection, effectively separating and enriching different targets remains a challenge [62].

Future research needs to focus on improving the selectivity of biosensors, developing new signal amplification strategies, refining sample processing and separation techniques, and advancing data processing methods. With these advancements, magnetic beads and biosensor technologies are expected to exert more profound influences in wearable devices and mobile health, providing more accurate and convenient protein detection solutions for individuals and clinical healthcare.

Conclusion

This review summarizes the research progress and applications of functionalized magnetic beads in protein detection. Functionalized magnetic beads, as a novel tool for bio-separation and detection, exhibit significant advantages in biomolecule capture, separation, and detection due to their unique physical structure and surface modification capabilities. The article first introduces the three main physical structures of magnetic beads: core-shell, hollow, and composite structures, and discusses in detail the characteristics of each structure and their specific applications in protein detection. Core-shell magnetic beads, with their high stability and large specific surface area, are widely used in various bio-detection platforms, while hollow and composite magnetic beads, by providing more functional sites and enhanced capture ability, further improve the sensitivity and specificity of detection.

In practical applications, functionalized magnetic beads, when combined with electrochemical biosensors, demonstrate numerous advantages over traditional detection methods. Through surface functionalization of the magnetic beads, target proteins can be rapidly and efficiently captured, and electrochemical biosensors can be used for real-time detection of reaction signals. This method not only accelerates the detection speed but also enhances sensitivity and specificity, simplifies the operation process, reduces sample processing time, and is well-suited for multiplex detection and on-site applications. Furthermore, the cost-effectiveness of functionalized magnetic beads in high-throughput detection makes them an indispensable part of modern bio-detection technologies.

Overall, this article thoroughly discusses the key technological advancements of functionalized magnetic beads in protein detection, highlighting their broad application prospects in fields such as biomedicine, environmental monitoring, and food safety. Future research will continue to optimize surface functionalization techniques for magnetic beads, further improving their detection performance and promoting their wider practical application.

In future bio-detection, the synergistic application of magnetic beads and biosensors is expected to make significant progress in several areas, particularly in technological innovations, the development of multimodal detection platforms, and the creation of portable detection devices.

Future research is likely to explore novel magnetic materials and surface modification techniques to enhance the magnetic properties and biocompatibility of magnetic beads. For instance, by introducing rare earth elements or transition metals, magnetic nanoparticles with higher magnetic saturation and superior magnetic responsiveness can be synthesized. Additionally, advancements in surface modification techniques, such as dopamine polymerization or silanization reactions, will further improve the stability and functionality of magnetic beads. These innovations will make magnetic beads more efficient in biomolecule capture and separation while reducing non-specific binding and improving the specificity and sensitivity of detection.

Integrated biosensors combining multiple detection modes will be developed for multiparameter analysis of proteins. This multimodal platform will combine electrochemical, optical, and magnetic resonance technologies to provide a more comprehensive analysis of protein characteristics. For example, electrochemical biosensors will allow monitoring of protein redox states, while optical biosensors will detect interactions between proteins and fluorescent markers. The application of magnetic resonance technology will provide detailed information about protein structure and dynamics. This integrated detection strategy will significantly enrich our understanding of protein functions and behaviors.

To meet the demand for rapid on-site detection, the development of portable and on-site detection devices will become an important direction. These devices will integrate miniaturization technologies, wireless communication, and user-friendly interfaces, making protein detection more convenient and accessible. With the advancement of microfluidic technologies and nanomanufacturing, future detection devices will be capable of performing complex biochemical analyses on smaller platforms. Additionally, by integrating with smartphones or other mobile devices, these systems will provide instant detection results and data analysis, greatly expanding the application scope of bio-detection technologies.

In conclusion, future research will continue to drive the development of magnetic beads and biosensor technologies to enable faster, more accurate, and cost-effective protein detection. These technological advancements will not only provide powerful tools for biomedical research but also bring revolutionary changes to fields such as clinical diagnostics and environmental monitoring.

Author Contributions: Haoyuan Su organized papers and wrote the manuscript; Yuehua Liao, Shu Wu, Jun Ji, Shuya An have made important contributions to the analysis and manuscript preparation, Dongdong Zeng gave instruction and revision of manuscript writing.

References

[1] Das S, Lyon CJ, Hu T. A Panorama of Extracellular Vesicle Applications: From Biomarker Detection to Therapeutics. ACS Nano 2024;18(14):9784-9797.

[2] Qi L, Hu Q, Kang Q, et al. Detection of Biomarkers in Blood Using Liquid Crystals Assisted with Aptamer-Target Recognition Triggered in Situ Rolling Circle Amplification on Magnetic Beads. Anal Chem 2019;91(18):11653-11660.

[3] Ouzounis CA, Coulson RM, Enright AJ, et al. Classification schemes for protein structure and function. Nat Rev Genet 2003;4(7):508-19.

[4] Cianci M. INTRODUCTION TO PROTEINS, Introduction to Proteins: Structure, Function, and Motion, 2nd edition. Crystallogr Rev 2021;27:47-50.

[5] Chen YY, Wang LB, Zhu HL, et al. Relationship Between Programmed Death-ligand 1 and Clinicopathological Characteristics in Non-small Cell Lung Cancer Patients. Chin Med Sci J 2013;28(03):147-151.

[6] Mukherjee S, Poudyal M, Dave K, et al. Protein misfolding and amyloid nucleation through liquid–liquid phase separation. Chem Soc Rev 2024;53(10):4976-5013.

[7] Van Thanh Nguyen N, Smadja C, Taverna M, et al. Immuno-capture of biomolecules on functionalized magnetic beads: From characterization to fluorescent detection of a biomarker for Alzheimer's disease toward high performance biosensing. Talanta Open 2023;8:100264.

[8] Fahmy HM, Abu Serea ES, Salah-Eldin RE, et al. Recent Progress in Graphene- and Related Carbon-Nanomaterial-based Electrochemical Biosensors for Early Disease Detection. ACS Biomater Sci Eng 2022;8(3):964-1000.

[9] Sankar K, Kuzmanovic U, Schaus SE, et al. Strategy, Design, and Fabrication of Electrochemical Biosensors: A Tutorial. ACS Sens 2024;9(5):2254-2274.

[10] Saldanha DJ, Cai A, Dorval Courchesne NM. The Evolving Role of Proteins in Wearable Sweat Biosensors. ACS Biomater Sci Eng 2023;9(5):2020-2047.

[11] Menard-Moyon C, Bianco A, Kalantar-Zadeh K. Two-Dimensional Material-Based Biosensors for Virus Detection. ACS Sens 2020;5(12):3739-3769.

[12] Carinelli S, Luis-Sunga M, González-Mora JL, et al. Synthesis and Modification of Magnetic Nanoparticles for Biosensing and Bioassay Applications: A Review. Chemosensors 2023;11(10):533.

[13] Loichen J, Hartmann U. Cell-surface interactions involving immobilized magnetite nanoparticles on flat magnetic substrates. Eur Biophys J 2009;38(7):891-901.

[14] Yoon TJ, Lee H, Shao H, et al. Multicore assemblies potentiate magnetic properties of biomagnetic nanoparticles. Adv Mater 2011;23(41):4793-4797.

[15] Zhao W, Zhang WP, Zhang ZL, et al. Robust and highly sensitive fluorescence approach for point-of-care virus detection based on immunomagnetic separation. Anal Chem 2012;84(5):2358-2365.

[16] Berg Luecke L, Gundry RL. Assessment of Streptavidin Bead Binding Capacity to Improve Quality of Streptavidin-based Enrichment Studies. J Proteome Res 2021;20(2):1153-1164.

[17] Mao Y, Zhang Y, Yu Y, et al. Self-assembled supramolecular immunomagnetic nanoparticles through pi-pi stacking strategy for the enrichment of circulating tumor cells. Regen Biomater 2023;10:rbad016.

[18] Ding F, Ren P, Wang G, et al. Hollow cellulose-carbon nanotubes composite beads with aligned porous structure for fast methylene blue adsorption. Int J Biol Macromol 2021;182:750-759.
[19] Chen S, Chu LT, Yeung PP, et al. Enzyme-Free Amplification by Nano Sticky Balls for Visual Detection of ssDNA/RNA Oligonucleotides. ACS Appl Mater Interfaces 2015;7:22821-22830.
[20] Spicer V, Ezzati P, Neustaeter H, et al. 3D HPLC-MS with Reversed-Phase Separation Functionality in All Three Dimensions for Large-Scale Bottom-Up Proteomics and Peptide Retention Data Collection. Anal Chem 2016;88(5):2847-2855.
[21] Huang Z, Gao N, Lin Y, et al. Exploring the migration and transformation of lattice oxygen during chemical looping with NiFe2O4 oxygen carrier. Chem Eng J 2022;429:132064.
[22] Cao Q, Li Z, Wang Z, et al. Disaggregation and separation dynamics of magnetic particles in a microfluidic flow under an alternating gradient magnetic field. J Phys D:Appl Phys 2018;51(19):195002.
[23] Pahlow S, Richard-Lacroix M, Hornung F, et al. Simple, Fast and Convenient Magnetic Bead-Based Sample Preparation for Detecting Viruses via Raman-Spectroscopy. Biosensors (Basel) 2023;13(6):594.
[24] Zhou X, Zhang Y, Kang K, et al. Controllable Environment Protein Corona-Disguised Immunomagnetic Beads for High-Performance Circulating Tumor Cell Enrichment. Anal Chem 2022;94(11):4650-4657.
[25] Zhu N, Ji H, Yu P, et al. Surface Modification of Magnetic Iron Oxide Nanoparticles. Nanomaterials (Basel) 2018;8(10):810.
[26] Huber DL. Synthesis, properties, and applications of iron nanoparticles. Small 2005;1(5):482-501.
[27] López Pérez JA, López Quintela MA, Mira J, et al. Advances in the Preparation of Magnetic Nanoparticles by the Microemulsion Method. J Phys Chem B 1997;101(41):8045-8047.
[28] Shahbazi R, Behbahani FK. Synthesis, modifications, and applications of iron-based nanoparticles. Mol Divers 2024;28(6):4515-4552.
[29] Zhi J, Wang Y, Lu Y, et al. In situ preparation of magnetic chitosan/Fe3O4 composite nanoparticles in tiny pools of water-in-oil microemulsion. React Funct Polym 2006;66(12):1552-1558.
[30] Šuljagić M, Vulić P, Jeremić D, et al. The influence of the starch coating on the magnetic properties of nanosized cobalt ferrites obtained by different synthetic methods. Mater Res Bull 2021;134:111117.
[31] Hemauer MA, Lee S, Bakshi MS. Nanorod morphology control of iron oxide nanoparticles induced by ionic surfactants and their solid – liquid interfacial adsorption. J Cryst Growth 2019;508:34-41.
[32] Araujo EV, Carneiro SV, Neto DMA, et al. Advances in surface design and biomedical applications of magnetic nanoparticles. Adv Colloid Interface Sci 2024;328:103166.
[33] Yao J, Adnan A, Wang HB. Separation mixed semen of two individuals using magnetic beads coupled ABH blood group antibody. Electrophoresis 2023;44(19-20):1539-1547.
[34] Martinkova P, Konecna H, Gintar P, et al. Benchmarking of Two Peptide Clean-Up Protocols: SP2 and Ethyl Acetate Extraction for Sodium Dodecyl Sulfate or Polyethylene Glycol Removal from Plant Samples before LC-MS/MS. Int J Mol Sci 2023;24(24):17347.
[35] Kou HS, Lin KH, Sebuyoya R, et al. Dual-probe ligation without PCR for fluorescent sandwich assay of EGFR nucleotide variants in magnetic gene capture platform. Mikrochim Acta 2023;190(9):375.
[36] Xu S, Chen J, Yang F, et al. A FEN 1-driven DNA walker-like reaction coupling with magnetic bead-based separation for specific SNP detection. Front Bioeng Biotechnol 2023;11:1279473.
[37] Chen L, Gao T, Wu X, et al. Polycarboxylate functionalized magnetic nanoparticles Fe3O4@SiO2@CS-COOH: Preparation, characterization, and immobilization of bovine serum albumin. Int J Biol Macromol 2024;260(Pt 2):129617.
[38] Zhao J, Zhang S, Dong J, et al. Screening and identification of peptidyl arginine deiminase 4 inhibitors from herbal plants extracts and purified natural products by a trypsin assisted sensitive immunoassay based on streptavidin magnetic beads. Talanta 2024;279:126611.
[39] Wang W, Geng L, Zhang Y, et al. Development of antibody-aptamer sandwich-like immunosensor based on RCA and Nicked-PAM CRISPR/Cas12a system for the ultra-sensitive detection of a biomarker. Anal Chim Acta 2023;1283:341849.
[40] Zhang Y, Wang J, Liu M, et al. Magnetically induced self-assembly electrochemical biosensor with ultra-low detection limit and extended measuring range for sensitive detection of HER2 protein. Bioelectrochemistry 2024;155:108592.
[41] Roy S, Kang S, Choi KY, et al. Implementation of an ultra-sensitive microwell-based electrochemical sensor for the detection of Alzheimer's disease. Biosens Bioelectron 2024;247:115898.
[42] Gutierrez-Galvez L, Seddaoui N, Fiore L, et al. Functionalized N95 Face Mask with a Chemical-Free Paper-Based Collector for Exhaled

Breath Analysis: SARS-CoV-2 Detection with a Printed Immunosensor as a Case Study. ACS Sens 2024;9(8):4047-4057.
[43] Yasamineh S, Nikben N, Hamed Ahmed M, et al. Increasing the sensitivity and accuracy of detecting exosomes as biomarkers for cancer monitoring using optical nanobiosensors. Cancer Cell Int 2024;24(1):189.
[44] Luo S, Wu Y, Pan W, et al. An integrated magneto-fluorescent nanosensor for rapid and sensitive detection of tumor-derived exosomes. Sens Actuators, B 2023;374:132792.
[45] Wu Y, Gao Z, Chai Y, et al. One-step and label-free ratiometric fluorescence assay for the detection of plasma exosome towards cancer diagnosis. Talanta 2024;271:125700.
[46] Sun Y, Bai Y, Song D, et al. Design and performances of immunoassay based on SPR biosensor with magnetic microbeads. Biosens Bioelectron 2007;23(4):473-478.
[47] Wang Y, Dostalek J, Knoll W. Magnetic nanoparticle-enhanced biosensor based on grating-coupled surface plasmon resonance. Anal Chem 2011;83(16):6202-6207.
[48] Wang Y, Cui ZJ, Zhang XJ, et al. Research progress of metamaterials powered advanced terahertz biochemical sensing detection techniques. Acta Phys Sin 2021;70(24):247820.
[49] Bi H, You R, Bian X, et al. A magnetic control enrichment technique combined with terahertz metamaterial biosensor for detecting SARS-CoV-2 spike protein. Biosens Bioelectron 2024;243:115763.
[50] Zhu J, Li C, Fang B, et al. Advancements in biosensing detection based on terahertz metasurfaces. Opt Laser Technol 2024;176:110978.
[51] Zurnansyah, Jayanti PD, Mahardhika LJ, et al. Real-time biomolecule detection using GMR chip-based sensor with green-synthesized Fe3O4/rGO nanocomposites as magnetic labels. Sens Actuators, A 2024;375:115493.
[52] Wu K, Tonini D, Liang S, et al. Giant Magnetoresistance Biosensors in Biomedical Applications. ACS Appl Mater Interfaces 2022;14(8):9945-9969.
[53] Gupta S, Bhatter P, Kakkar V. Point-of-care detection of tuberculosis using magnetoresistive biosensing chip. Tuberculosis (Edinb) 2021;127:102055.
[54] Ding J, Gu Y, Li F, et al. DNA Nanostructure-Based Magnetic Beads for Potentiometric Aptasensing. Anal Chem 2015;87(13):6465-6469.
[55] Roth S, Margulis M, Danielli A. Recent Advances in Rapid and Highly Sensitive Detection of Proteins and Specific DNA Sequences Using a Magnetic Modulation Biosensing System. Sensors (Basel) 2022;22(12):4497.
[56] Vanova V, Mitrevska K, Milosavljevic V, et al. Peptide-based electrochemical biosensors utilized for protein detection. Biosens Bioelectron 2021;180:113087.
[57] Klebes A, Ates HC, Verboket RD, et al. Emerging multianalyte biosensors for the simultaneous detection of protein and nucleic acid biomarkers. Biosens Bioelectron 2024;244:115800.
[58] Wu J, Liu H, Chen W, et al. Device integration of electrochemical biosensors. Nat Rev Bioeng 2023;1(5):346-360.
[59] Cheng HP, Yang TH, Wang JC, et al. Recent Trends and Innovations in Bead-Based Biosensors for Cancer Detection. Sensors (Basel) 2024;24(9):2904.
[60] Lu L, Zhang J, Xie Y, et al. Wearable Health Devices in Health Care: Narrative Systematic Review. JMIR Mhealth Uhealth 2020;8(11):e18907.
[61] Parolo C, Idili A, Ortega G, et al. Real-Time Monitoring of a Protein Biomarker. ACS Sens 2020;5(7):1877-1881.
[62] Chadha U, Bhardwaj P, Agarwal R, et al. Recent progress and growth in biosensors technology: A critical review. J Ind Eng Chem 2022;109:21-51.

Progress in Medical Devices
ISSN: 2957-5478

Zentime

Review Article https://doi.org/10.61189/974215qcjfzk

Advancements in finite element analysis for prosthodontics

Yan Wang, Liwen Chen

School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

Corresponding author: Liwen Chen.

Acknowledgement: This work was partly supported by Shanghai Sailing Program (22YF1430500).

Declaration of conflict of interest: None.

Received November 11, 2024; Accepted December 3, 2024; Published December 31, 2024

Highlights

- This paper presents a comprehensive review of the advancements in finite element analysis (FEA) within the field of prosthodontics over the past five years.
- It examines the role of FEA in aiding the selection of restorative materials, enhancing prosthetic designs, and investigating the dynamic interactions between prostheses and natural dentition.
- Integrating FEA findings with clinical practice enhances treatment outcomes and patient satisfaction.

Abstract

Finite element analysis (FEA) is a computer-aided tool widely employed in the field of prosthodontics, offering a comprehensive understanding of biomechanical behavior and assisting in the design and evaluation of dental prostheses. By dividing a model into finite elements, FEA enables accurate predictions of stress, strain, and displacement of structures. This review summarizes recent research developments in the application of FEA across various aspects of prosthodontics, including dental implant, removable partial denture, fixed partial denture and their combinations. FEA plays a significant role in selecting restoration materials, optimizing prosthetic designs, and examining the dynamic interactions between prostheses and natural teeth. Its computational efficiency and accuracy have expanded its application potentials for preoperative planning in custom-made prosthodontics. Upon the physician's assessment of the repair requirements tailored to the individual patient's condition, FEA can be employed to evaluate the stress distribution, displacement, and other relevant outcomes associated with the proposed restoration. When integrated with clinical expertise, it facilitates assessing design feasibility, identifying necessary adjustments, and optimizing prosthetic solutions to mitigate the risk of failure. Additionally, FEA helps identify potential complications arising from long-term prosthetics use, allowing for the implementation of preventive strategies. Presenting FEA results to patients enhances their understanding of the scientific basis and rationale behind the design, thereby bolstering patient confidence in the proposed intervention. Despite its ongoing limitations, FEA underscores the importance of integrating computational findings with clinical judgment and supplementary diagnostic tools. This review emphasizes the growing role of FEA in advancing prosthodontics by offering computational analysis and design optimization, ultimately improving treatment outcomes and patient satisfaction.

Keywords: Finite element analysis, dental implant, removable partial denture, fixed denture, fixed partial denture

Introduction

Dental caries, a prevalent oral disease, can develop at any stage of life, potentially leading to tooth damage and eventual tooth loss [1]. Clinically, dental caries progression can be categorized into three stages: tooth defect, dentition defect, and edentulousness. Tooth defect refers to damaged or abnormal tooth morphology. As the condition worsens, multiple tooth extractions may become necessary, leading to a dentition defect. In severe cases, edentulousness - complete loss of teeth in the upper and (or) lower jaw, occurs. These conditions significantly impact dental aesthetics and functionality, potentially causing anxiety, diminished

Address correspondence to: Liwen Chen, School of Health Science and Engineering, University of Shanghai for Science and Technology, No.516 Jungong Road, Yangpu District, Shanghai 200093, China. E-mail: chenlw@usst.edu.cn.

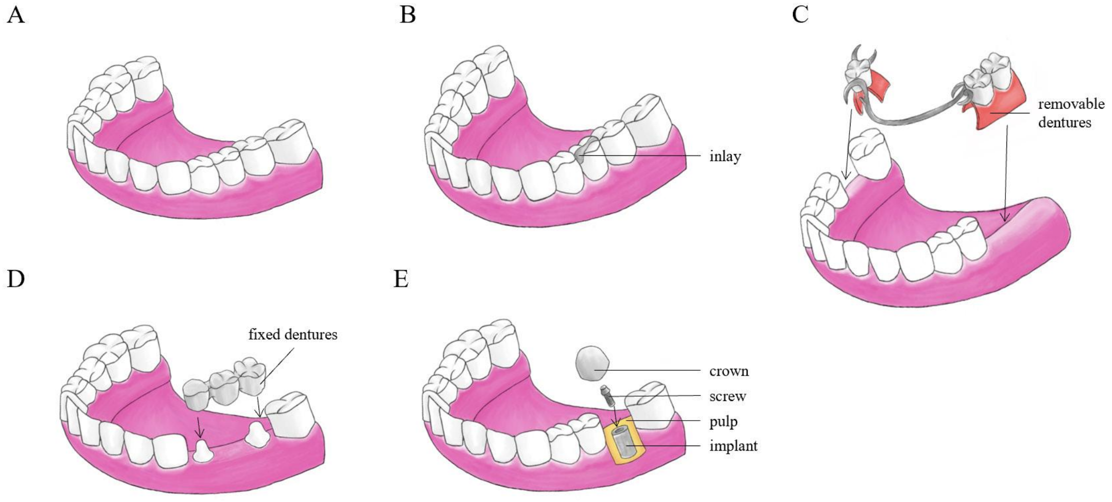


Figure 1. Healthy teeth and common treatment modalities. Schemes of (A) healthy teeth and various aspects of treatment modalities within the field of prosthodontics. Specifically, the schemes encompass: (B) inlays, with the gray area denoting the inlay component; (C) removable dentures, depicted as situated above the natural teeth; (D) fixed dentures, which are also represented in gray; and (E) dental implants, where the pulp is highlighted in yellow, alongside the incorporation of implants and screws within the dental architecture.

self-esteem, and other negative emotions that affect patients' physical and mental well-being.

Prosthodontics offers a range of treatment options tailored to varying tooth defects, including inlays, implants, removable dentures, fixed dentures, and combination treatments, as shown in **Figure 1**. Compared to onlays that sit partially above the tooth surface, inlays are completely embedded within the crown, which restore tooth shape and function while preventing further loss of tooth substance. As for dentition defects, removable partial dentures (RPDs) are commonly employed due to their wide range of applications and cost-effectiveness. However, RPDs rely on a dual support system that may compromise retention and stability, particularly with multiple missing teeth or posterior tooth loss, where rotation during chewing can affect chewing function [2, 3].

In contrast, fixed dentures, also known as fixed bridges, are securely attached to neighboring teeth or implants, providing greater stability. However, their application often requires removing more healthy tooth material, and they are more challenging to clean thoroughly, potentially leading to infection. For patients with edentulousness, traditional complete dentures, removable appliance, share similar drawbacks to RPDs [4]. Implant dentures, on the other hand, have rapidly evolved, demonstrating high success rates (89-100%) across various indications and therapies, making them a viable and increasingly popular solution [5, 6]. Furthermore, the integration of implants with other restoration methods has emerged as a new direction for prosthodontics.

Despite the availability of various treatment options, determining the most appropriate repair methods tailored to the specific requirements of each individual case remains a significant challenge. In vitro testing is generally accepted in dentistry, notwithstanding its inherent limitations concerning precision and intricacy [7]. The integration of digital technologies, such as computer-aided design (CAD) and computer-aided manufacturing, virtual reality, additive manufacturing (3D printing) and finite element analysis (FEA), is transforming the field [8-23]. Such technologies hold promises in enhancing diagnostic precision, tailoring treatments to individual needs, streamlining operational efficiency, promoting patient experiences, and fostering clinical excellence in dentistry.

Simulating the complex human biomechanical system such as bone and soft tissue is a challenging task in the medical field, and FEA emerges as an effective solution to address these complexities [24, 25]. It is a widely used numerical method that uses CAD to construct and analyze mathematical models of engineering or physical systems [25, 26]. Introduced by Courant in 1943, FEA found application in orthopedics in the 1970s, where it was used to evaluate stress and deformations in human bones under functional loading [27-29]. Since its introduction to dentistry in 1973, FEA has been extensively used to simplify complex oral systems into manageable data models [26, 30, 31]. FEA is used in dentistry for various purposes, including analyzing soft tissue changes be-

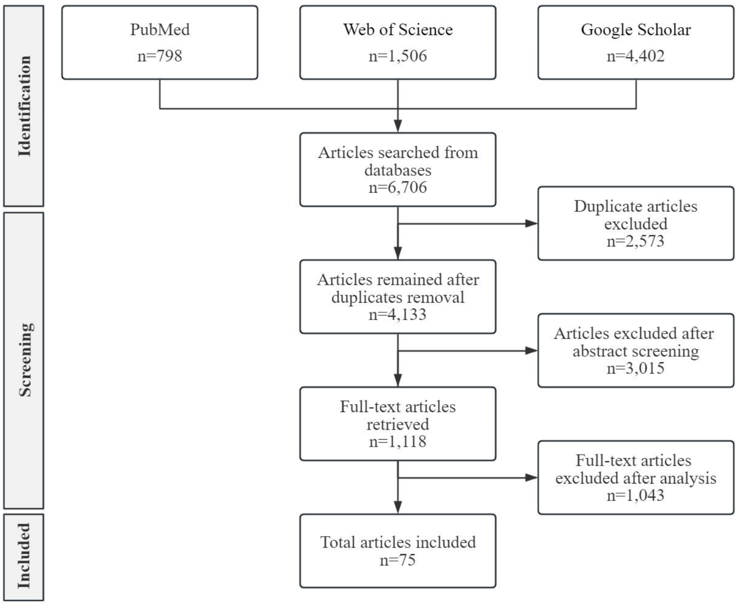


Figure 2. Review flow chart. The methodology involves utilizing specific databases and search terms to conduct an initial screening based on titles and abstracts. This is followed by a more comprehensive evaluation of the full-text content, guided by established inclusion and exclusion criteria.

fore and after operations, simulating craniofacial bone anatomy, customizing optimal surgical strategies by evaluating stress distribution and displacement in reconstruction components, and assessing bone conditions and load-carrying capacity around implants [32-36].

The aim of this review is to summarize the recent advancements in the application FEA in prosthodontics. The paper is structured as follows: Chapter 2 introduces the literature search method, establishing a foundation for the subsequent research. Chapter 3 describes the fundamental steps of FEA, offering theoretical and technical framework. Chapters 4 to 6 constitute the core of this paper, covering the application of FEA to the analysis of restorative materials, the design of restorative structures, and the analysis of the interaction between the restoration and the teeth. Finally, Chapter 7 concludes the review by summarizing current trends and proposing potential future directions in this field.

Literature search

This review systematically searched the existing literature in databases including PubMed, Web of Science, and Google Scholar, from 2019 to 2024. The search terms included “dental implant,” “removable partial denture,” “fixed denture,” “fixed partial denture,” and “finite element,” used alone or combined with Boolean operators (e.g. AND, OR). References from the identified articles were also examined to ensure a comprehensive collection. **Figure 2** shows the process of article screening. The retrieved literature was imported into EndNote to remove duplicates. Initially, a preliminary screening based on article titles was performed, selecting those explicitly referencing the specified keywords. Articles unrelated to FEA, such as purely biological research or descriptions of clinical procedures, were excluded. Subsequently, the abstracts of the selected articles were examined to ascertain the clarity of the study’s purpose and a concise summary of the main findings. Abstracts incorporating keywords pertinent to prosthodontics and presenting novel methods, models, or insights were prioritized. Following this initial screening, full-text articles were thoroughly evaluated. The assessment focused on the validity of the study design, including the development of finite element models, the appropriate application of boundary conditions and loads. Inclusion criteria stipulate that studies should be published in English, specifically address FEA in the context of prosthodontics, and provide comprehensive descriptions of the models and analyses of the results. Articles lacking adequate model validation, employing outdated methodologies, or unrelated to prosthodontics were excluded. Finally, the obtained literature was analyzed in detail.

General procedure of finite element analysis

As a computer-based approach, FEA employs mathematical models to analyze practical problems in various fields. The primary objective of FEA is not to achieve an exact result, but to offer approximate solutions that include some fundamental information, such as displacement, force, temperature. **Figure 3** presents the general analysis process of FEA. The first step is to construct an overall geometric model. This model is then divided into a finite number of small elements for generating meshes, which is termed discretization. FEA focuses on the analysis of these individual elements. Mechanical properties and material characteristics of individual elements are analyzed independently, constructing the corresponding stiffness matrix. Following this step, the mechanical properties of the entire model are calculated, allowing for the estimation of stress, strain, displacement

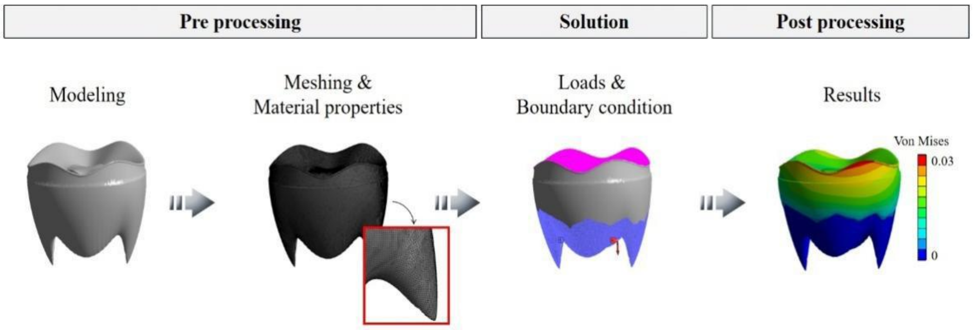


Figure 3. FEA procedure. The pre-processing phase encompasses several critical steps, including the development of a model, the creation of a mesh, and the specification of material parameters. Subsequently, appropriate load and boundary conditions are established, followed by the analysis of results utilizing finite element calculations. FEA, Finite element analysis procedure.

and other response outcomes [25, 29]. The key steps in the FEA procedure are briefly described below. For interested readers, more detailed tutorials may be found in these reviews [37-39].

Construction of geometric model

The construction of a geometric model is the first step in FEA for prosthodontics, including multiple components, such as prosthesis, bone, and tooth. The prosthesis geometry is usually based on the design data provided by manufacturers.

The bone models are categorized into simple and complex structures [25]. For simple structures, the bone is represented through an array of 2D or 3D shapes, while both cortical and trabecular components are retained. Complex structures are typically obtained using medical imaging techniques, such as intraoral scanners and cone beam computer tomography [10, 40]. FEA can be performed directly on the geometric model obtained from medical images [37]. However, the bone model to be analyzed must first be segmented, typically through image processing software like Mimics or Amira. The segmented images are then converted into finite element models using CAD tools (e.g. Geomagic, SolidWorks).

Tooth models typically consist of enamel, dentin, and pulp. The modeling process is similar to that of bone models (i.e. images are obtained through medical imaging technology, processed and modeled using CAD). In some cases, ex vivo test models are used, where a physical tooth model is created from extracted teeth [41, 42]. Alternatively, finite element models may be directly generated from existing image data [43].

Discretization and assignment of material properties

Meshing, or discretization, is a crucial step in finite element analysis. This process involves dividing the geometric model into smaller computational elements. To ensure accuracy and efficiency of the analysis, the mesh size must be sufficiently fine so that the results are independent of the mesh size. Additionally, it is important to strike a balance between geometric complexity and computational efficiency, as the choice of element type significantly influences the analysis. In 2D finite element modeling, structures are commonly divided into triangular or quadrilateral elements, while tetrahedral or hexahedral elements are typically employed in 3D modeling. Tetrahedral elements are more frequently used compared to hexahedral ones, not only due to more convenient meshing automation, but also better fitting to complex geometries [44]. Researchers often use pre-processing software, such as Hypermesh, HyperWorks, and VR Mesh Studio, to improve mesh quality and analysis productivity [42, 45-49].

To meet specific application requirements, it is necessary to assign appropriate material properties for all model components, as these properties significantly influence the accuracy of FEA results. The mechanical properties of bone, teeth, and restorative materials, such as elastic modulus, shear modulus, and Poisson's ratio, can be obtained experimentally or referenced from previous reports.

FEA simplifies the model's complexity by adjusting the quantity and nature of elements and specifying material properties to ensure precision. However, clinical validation remains essential to confirm the accuracy of the approximate outcomes [50]. Despite the inherent complexity of biological structures, models are commonly simplified with material properties often assumed to be homogeneous, isotropic and linear elastic for reducing computation time. This does not mean that important details should be left out or omitted. Oversimplification may lead to the exclusion of essential characteristics, resulting in inaccurate representation. Hence, it is imperative to achieve a delicate balance between computational efficiency and model accuracy.

To better simulate real-life scenarios, some researchers adjust material properties to better match actual conditions. For instance, Chakraborty modeled the mandible as a heterogeneous material, considering different bone conditions to evaluate the implant's biomechanical potential for promoting osseointegration at the peri-implant bone [51]. This method can simulate the implant process based on the patient's bone condition, which aids in selecting suitable implants and creating personalized treatment plans. Meanwhile, Ortiz-Puigpelat developed a methodology utilizing a nonlinear geometric model and material properties specific to the periodontal ligament to predict its mechanical response during occlusion [52]. This approach closely replicate actual clinical conditions, offering enhanced precision and reliability in FEA results, which can inform clinical assessments and therapeutic interventions.

Boundary conditions

After discretization, boundary conditions and loads should be appropriately assigned based on different experimental purposes. The boundary conditions can be fixed, free, periodic, or other types. Most studies assume that the connections between the prosthesis, abutments, and bone interfaces are tightly coupled, with nodes constrained in all directions to prevent translation or rotation. Previous studies have shown that fixed boundary conditions are commonly used to simulate tooth roots fixed in the bone when the periodontal ligament model is neglected [41, 53, 54]. This involves fixing the tooth from the enamel-cementum junction to the apical region. Additionally, masticatory forces are typically simulated by applying vertical and/or oblique loads. To simulate the actual chewing force, some studies incorporate food clumps or spherical objects on the tooth contact surface to represent food and apply loads accordingly [48, 55, 56].

Solution and analysis

The accuracy of FEA depends on the above steps. With the help of software specialized FEA, a large amount of data can be obtained for the solution of the finite element model which has been set up. After obtaining the solution, the results may be analyzed and displayed by post-processing, which involves calculating additional physical quantities (e.g. stress, strain, displacement) and visualizing the results. In the following chapters, FEA results are analyzed from three application perspectives.

Analysis of restoration materials

The selection of a restoration material significant influences the magnitude and distribution of stress in teeth [57]. Various materials are available to dentists, including zirconia, lithium disilicates, ceramics, amalgam, and resin-based composites. **Figure 4** provides a comparative illustration of clinically used restorative materials based on relevant literature [58-61]. However, selecting the most suitable material for a given situation can be challenging [49]. FEA enables the creation of restoration models using different materials, simulate bite forces with varied directions and magnitudes, and evaluates the resulting stress and strain under maximum load conditions (**Table 1**) [43]. The primary focus of material analysis is to evaluate the modulus of elasticity (MOE).

FEA studies demonstrate that significant discrepancies between the MOE of dental restorative materials and dental tissues can lead to extremely large stress values. Therefore, a comparable MOE between materials and dental tissues reduces stress transmitted to the dental tissues [43]. Kim et al. utilized FEA to analyze optimal materials for mitigating residual crack propagation, finding that materials with lower MOE exhibited deformation due to stress absorption, while those with higher MOE exerted increased stress and stress concentration in the teeth, potentially enhancing aesthetic quality and fracture resistance [48]. Darwich et al. came to the similar conclusion when evaluating the material of the inner crown [40]. Therefore, it is recommended to use hard materials with high MOE to reduce the displacement of the tooth structure.

FEA can also be applied to study the fatigue properties of restoration materials. As high stress or stress concentration can potentially lead to cracks, Kim et al. evaluated crack propagation by assessing residual stress beyond the lower edge of the crack surface after treatment, identifying gold crown combined with resin filling as the most effective treatment method [48]. Demirel et al. compared the fatigue properties of various ceramic materials by simulating crack generation and propagation in repair materials through FEA, and found that polymer-infiltrated ceramic networks exhibited greater brittleness due to their lower MOE com-

Table 1. Material types and material properties analyzed in the different studies

Applications	Materials	Properties		Load direction & magnitude	Conclusion	Ref.
		MOE (MPa)	Poisson's ratio			
Endocrown	Translucent zirconia	210,000	0.307	Constant load: 600 N Direction: -Axial load: perpendicular to the occlusal surface of the crown -Oblique load: tilted by 45° with respect to the tooth axis	Translucent zirconia is considered the optimal material for internal crowns to preserve the dental restoration complex. Both lithium disilicate glass ceramic and zirconia-reinforced lithium silicate ceramic are appropriate materials to produce internal crowns, but resin nanoceramics may not be the most suitable option.	[40]
	Zirconia-reinforced lithium silicate ceramic	102,900	0.208			
	Lithium disilicate glass ceramic	83,000	0.210			
	Polymer-infiltrated network ceramic	30,000	0.230			
	Resin nanoceramic	12,700	0.470			
Inlay, Onlay, Overlay	Lithium disilicates (LS)	98,000	0.230	Constant load: 600 N Direction: perpendicular to the occlusal surface of the crown	LS exhibits superior properties relative to PICN materials.	[41]
	Polymer-infiltrated ceramic network (PICN)	35,000	0.230			
Inlay	Direct composite resin (Filtek Z350)	9,800	0.250	Constant load: 1,000 N Sliding-type contact: the contact points of a 6 mm diameter sphere* placed in the fovea	A gold crown filled internally with resin is an advantageous method to prevent further crack propagation in cracked teeth.	[48]
	Indirect composite resin (Tescera ATL)	14,200	0.250			
	Ceramic (Emax)	94,100	0.250			
	Gold	85,000	0.250			
RPD	Zirconia	210,000	0.300	Constant load: 100 N Direction: perpendicular to the bottom surface of the small connector	Zirconia occlusal rests, derived from conventional occlusal rest designs, are regarded as occlusal supports capable of enduring repetitive masticatory forces. The limited fatigue resistance of PEEK may render it unsuitable for extended clinical use as a dental scaffold material.	[97]
	Polyether-ether-ketone (PEEK)	4,100	0.400			
RPD	Zirconia (3Y-TZP)	200,000	0.300	Constant load: 600 N perpendicular to the occlusal surface of the crown	PEEK demonstrates enhanced biomechanical characteristics when utilized as the primary crown within a double-crown attachment system, but its effectiveness still needs additional validation.	[98]
	Titanium (Ti-6Al-4V)	110,000	0.300			
	CoCr	220,000	0.300			
	PEEK	5,100	0.400			
FPD	Zirconia	220,000	0.300	Constant load: 300 N Direction: perpendicular to the occlusal surface of the crown	Zirconia material can be utilized to fabricate implants-assisted FPD.	[50]
	Titanium	110,000	0.300			
FPD	Zirconia	210,000	0.260	Constant load: 1,000 N Direction: -Axial load: three vertical force vectors -Oblique load: tilted by 60° with respect to the tooth axis	The first principal stress at the inlay fixator and the connector is affected differently by the stiffness variation of the support structure.	[99]

Note: *A 6 mm diameter sphere was used to simulate food clump.

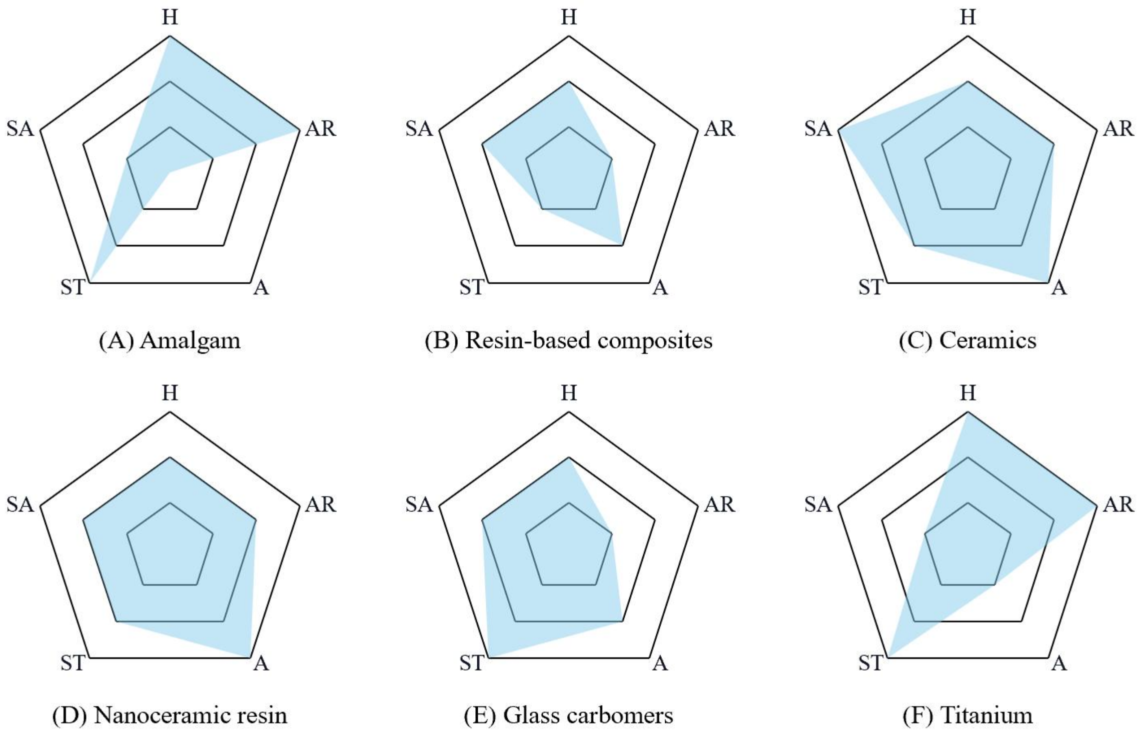


Figure 4. Comparison of performance of restoration materials. Six commonly used repair materials, i.e. (A) amalgam, (B) resin-based composites, (C) ceramics, (D) nanoceramic resin, (E) glass carbomers and (F) titanium, are evaluated relative to hardness (H), abrasion resistance (AR), aesthetics (A), service time (ST) and scope of application (SA).

pared to lithium disilicate glass ceramic [41].

Zirconia fixed partial dentures (FPDs) are known for their exceptional fracture resistance and superior fatigue properties compared to other dental ceramics, as demonstrated by recent FEA studies [50]. The fatigue resistance of zirconia FPDs is significantly higher than that of lithium disilicates FPDs, making them particularly suitable for cases involving molar replacement [62]. It is worth noting that suitable core materials can help reduce the fracture risk of zirconia FPDs. Canipaner et al. compared sound dentin, resin composite cores, and metallic cores, finding that while resin composite cores are less strong than metal cores, their flexibility helps disperse stress, thereby reducing fracture risk [63]. This underscores the importance of prosthesis design when selecting zirconia materials. However, Ceddia et al. recently compared crowns made of ceramic and zirconia with natural enamel to assess their mechanical properties under different loading conditions [64]. Vertical, oblique (45°), and horizontal loads were applied to simulate chewing. While ceramic crowns exhibited inferior mechanical properties and were more prone to fracture or fragmentation, zirconia crowns displayed higher stress levels under vertical and horizontal loads. This suggests that zirconia crowns may not be ideal for patients with oral dysfunction. Therefore, clinicians should choose appropriate materials according to the patient's specific dental condition.

In addition to repetitive functional loading, human teeth are subject to thermal stress from lifelong temperature fluctuations [65]. FEA simulations have shown that both hot and cold loads result in elevated stress levels in the inlays.

Structural designs of the dental prosthesis

Figure 5 shows the general procedure of dental restoration, from the beginning of clinical diagnosis to the final prosthesis installation, including material selection and shape. The effectiveness of restorative treatments depends not only on the properties of the restorative material but also on the design of the prosthesis structures. FEA provides a virtual platform to simulate various design iterations, enabling the identification of potential issues without the need for physical prototype.

Proper cavity designs are essential for inlays, as they can make use of residual tooth tissue while ensuring adequate fracture resistance and retention force, thereby reducing the risk of fracture. Cheng et al. used EFA to demonstrate that a concave design significantly improves the fracture resistance of ceramic inlays compared to flat designs [45]. Similarly, Babaei

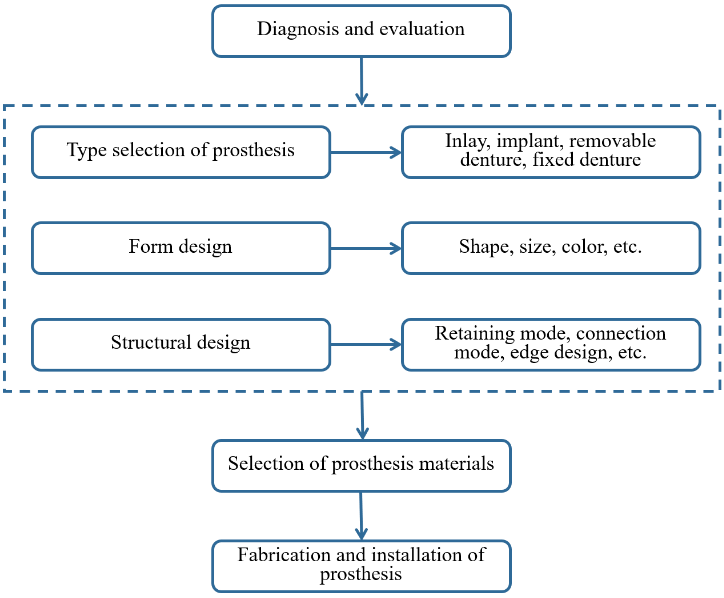


Figure 5. General procedures for dental restoration. Following the dentist's diagnosis, the selection of the prosthesis is determined based on the patient's specific condition. Subsequently, the design of the shape and structure is undertaken, after which the most suitable material for the fabrication and installation of the prosthesis is chosen.

et al. analyzed the effects of occlusal cavity depth and internal cavity angle (α) on the biomechanical stability of treated molars [7]. They demonstrated that as the occlusal cavity depth increases, the pressure level of the restoration also increases, resulting in reduced fracture resistance. Conversely, increasing α and reducing occlusal cavity depth lowers the pressure levels within the restoration, thereby improving stability and fracture resistance. This conclusion is consistent with the results reported by She et al. [66].

The design of the pier abutment and connector plays a crucial role in determining repair effectiveness and service life of FPD [67, 68]. Recent FEA studies have shown that the connector volume is a key factor affecting the maximum principal stress of dentures [10]. Smaller connector volumes result in higher maximum equivalent stress. Additionally, FEA has demonstrated that the area surrounding the connector experiences the highest stress levels, regardless of the design or material used [11, 12]. To compensate for differences in movement between implants and natural teeth, both rigid and non-rigid connectors can be utilized depending on the specific cases [69]. FEA results indicate that the non-rigid connectors reduce stress concentration in the bone surrounding the implant but increase stress levels in the alveolar bone [70]. In the design of a 5-element FPD, non-rigid connector was found to be more suitable [67, 68]. However, Naguib et al. observed through FEA that rigid connectors generate lower stress values than non-rigid connectors [69]. Therefore, when designing a three-element FPD, it is recommended to consider the optimal implant diameter and length and choose a rigid connection design to minimize stress on both the prosthesis and supporting bone.

Numerous studies have investigated the application of FEA in implant dentures. As for implants, Falcinelli et al. provided a comprehensive review of FEA application in implant dentures over the past 20 years [25]. Most studies focused on evaluating stress distribution in implants and surrounding bone, demonstrating that factors such as implant geometry, connection design, implant material, and properties of the bone-implant interface significantly influence the stress field. Key factors affecting load transfer and stress distribution include diameter of the implant, the length of the bone-implant interface, and the spacing, shape, and depth of the threads. The dimensions of an implant, specifically its diameter and length, directly impacts the stress distribution in the surrounding bone [71]. Increasing implant diameter reduces bone stress near the implant [72]. By optimizing the geometric configuration of the implant, it is possible to mitigate bone resorption around the implant, thereby enhancing both implant stability and the likelihood of long-term success. Furthermore, design of the implant-abutment connection also plays a critical role in stress transfer mechanisms, influencing stress distribution at the interface between the bone and the implant. An appropriately designed connection can alleviate stress concentration, lowering the risk of implant loosening and failure [73, 74]. Additionally, implant material selection, as discussed in Chapter 4, is vital for achieving optimal biocompatibility and mechanical compatibility with adjacent bone tissue. Selecting materials with an appropriate MOE minimizes bone resorption and mitigates implant failure risks. The nature of the implant connection exerts a more significant influence on the stress experienced by the implant and abutment than the diameter of the components. Tissue-level connections demonstrate a more advantageous stress distribution compared to bone-level connections. Lee recommends tissue-level connections for patients with low-density cancellous bone [75]. Additionally, bone strain is influenced by bone mineral density, with higher bone mineral density levels correlate with reduced stress on both the bone and the implants [76].

Implant-supported dentures are a reliable and predictable treatment option with high survival rates and long-term stability, owing to the advancements and widespread adoption of implant technology [77]. Studies using FEA have demonstrated consistent results for implant-supported dentures. For instance, Güzelce et al. compared the stress effects of various implant types for overdentures on supporting bone, implants, and frame materials under masticatory forces [78]. The study revealed that overdentures supported by conventional implants are more stable than those supported by micro-implants.

It is worth noting that traditional implant-supported fixed complete dentures typically require 6-8 implants and involve a long recovery time. The concept of All-on-4®, a clinical plan for implant-supported and immediate restoration of edentulous jaws proposed by Maló, has been extensively studied in clinical practice [79-84]. Similar to findings in implant dentures and implant-supported FPDs, FEA analyses of the All-on-4® system indicate that modifying the implant design in complete fixed denture supported by four implants has minimal impact on bone stress and strain. Therefore, altering the implant design may have little effect on implant bone remodeling.

Analysis of the interaction between the restoration and the tooth

Despite the material selection and structural design, the long-term stability and success of dental restorations also rely on the interaction between the prosthesis and the tooth. FEA can accurately simulate this interaction, delivering objective and comprehensive evaluations of prosthesis performance that are invaluable for clinical decision-making. By simulating forces under various conditions, FEA examines how the prosthesis influences stress distribution, displacement, and other parameters in the tooth. This analytical approach aids in uncovering the inherent principles governing the prosthesis - tooth interaction and elucidating the reciprocal mechanisms.

The integrity of the adhesive interface is important for the success of inlay repairs [85]. Previous studies have shown that the thickness of the adhesive layer affects the stress distribution and wear, though the ideal thickness remains debated [56]. Restorative materials also play a crucial role in the stress distribution and damage degree at the enamel-inlay bonding interface. He et al. conducted FEA to analyze the stress distribution of the inner crown and cement layers using different restoration materials and resin cements [55]. Their findings revealed that higher material hardness reduced stress on the cement layer because high-MOE materials concentrated the stress within themselves, thereby lowering peak stress on the cement layer. Therefore, high-MOE material is beneficial for bonding. Ceramic inlays bear the stress primarily on the ceramic structure, minimizing damage to the enamel. In contrast, resin composite inlays release shrinkage stress at the interface with the tooth structure, promoting crack propagation towards the inner enamel. Therefore, ceramic inlays offer superior protection residual dental tissue compared to resin composite inlays, which is consistent with the findings in Chapter 4.

Implant-assisted RPDs (IARPDs) enhance retention force and stability through both direct effects of the implant and indirect effects on the bone. Additionally, they also help maintain bone levels around the implant, particularly in cases of posterior tooth loss [86]. In a review of FEA, Ichikawa et al. analyzed the strain distribution across the metal frame, tooth root, and other components, focusing on the impact of implant positioning on the load distribution of IARPDs supporting elements [87]. Their results indicate that the position of implants in IARPDs can affect the load distribution of abutment teeth and implants. Implant support and/or retention in IARPDs can prevent wobble motion along the rotation axis of the prosthesis, thereby reducing the risk of abutment teeth loss. Comparative studies between in vitro experiments and FEA highlight that FEA can simulate not only vertical loads but also tilted or horizontal loads, offering more accurate experimental results that closely mimic real bite force and masticatory force.

For fixed bridges, the distribution of stress is highly dependent on the number and placement of implants. FEA studies has shown that optimal bone stress distribution and prevention of bone resorption caused by stress concentration can be achieved by using at least three implants to support a four-unit implant-supported FPD [47, 77, 80]. FEA has shown that the highest stress concentration occurs in the contact area between the abutment and implant, regardless of the abutment type [88]. Naguib et al. found through FEA that using implants with a larger width is more conducive to stress propagation, aligning with prior FEA research on implant dentures [71, 89]. Furthermore, FEA also identified a proportional relationship between implant length and stress: increasing the length of the distal implant while maintaining a fixed diameter may result in higher torque and stress [71].

Researchers have increasingly focused on tilted implants in complete denture restoration. FEA was used to evaluate the maximum von-Mises stress experienced by the implant and the adjacent cortical bone. Studies have shown that stress and strain exerted on the surrounding bone tissue are significantly influenced by the positioning of the load, whereas the implant's design has minimal impact [35]. Elevated bone stresses and strains were observed when the load was positioned near the tilted posterior implant. Similarly, Ibrahim et al. utilized FEA in their investigation and confirmed that stress levels increased with the angle of posterior implant tilt, suggesting that optimizing the implant's tilt angle can mitigate deformation and enhance performance [90].

While most FEA studies focus on mechanical factors, they may not fully account for biological responses, such as bone remodeling and tissue healing, which are critical to clinical outcomes. FEA provides valuable insights into the influence of prosthetic design on the healing of adjacent tissues. Implant designs that induce localized stress concentrations may disrupt the normal healing process, potentially delaying or impairing recovery. The adaptation of bone tissue in response to mechanical loading, known as bone remodeling, is a critical determinant of the long-term stability of implants. FEA models can incorporate bone remodeling processes to forecast alterations in bone density surrounding the implant and assess the long-term effects of prosthetic design. For instance, Poovarodom developed a bone remodeling algorithm based on FEA to simulate changes in bone density in response to mechanical stimulation, such as bite force [74]. Strain induced by mechanical loading influences bone density changes over time. By implementing a 12-month bone remodeling algorithm, the study optimized key parameters, including implant placement depth, abutment taper, and gingival height of the titanium abutment. Findings indicated that both implant placement depth and taper significantly affect bone remodeling, with the optimized model improving bone density uniformity. Rajaeirad employed the same theoretical framework of bone remodeling algorithms to forecast the impact of functionally graded materials on the process of bone remodeling [91]. FEA results indicated that the bone density and mechanical properties of titanium-hydroxyapatite (Ti-HAP) implants were influenced by the volume fraction of hydroxyapatite within functionally graded materials, consequently enhancing osseointegration and remodeling.

Current trends and future directions

FEA continues to play a pivotal role in advancing biomechanical research in prosthodontics and offers substantial potential for clinical application. While FEA is currently employed in prosthetics under static and ideal conditions, it has significant advantages in scientific research, including short experimental duration, good repeatability, and the ability to test a wide range of complex working conditions without causing damage. These attributes make it an invaluable predictive tool for clinical performance [85].

In clinical practice, the selection of an appropriate repair method should consider the patient's specific needs, oral condition, anticipated outcomes, and financial resources. FEA aids clinicians by providing critical insights into the properties of various materials and techniques, thereby facilitating more informed clinical decision-making. Inlays are frequently employed for small to medium-sized defects in posterior teeth, particularly when the remaining tooth structure is sufficiently robust to support the restoration [43, 85]. RPDs are indicated in scenarios involving multiple missing teeth, particularly when fixed prostheses or implants are not financially viable [92]. RPDs are relatively easy to clean and maintain and effectively restore functionality [2]. Nevertheless, their stability and comfort are generally inferior to other options and they may exert stress on adjacent natural teeth and surrounding soft tissues [2, 93]. FPDs are ideal for the restoration of one or more adjacent missing teeth, particularly in cases where the neighboring teeth are in good health [77]. They offer excellent stability, comfort, and functional and aesthetic outcomes resembling natural dentition. However, the preparation often involves the removal of a substantial amount of dental tissue from supporting teeth [94]. Implants are ideal for the restoration of one or more missing teeth, particularly for patients seeking a restoration that closely resembles natural dentition. They offer optimal stability and support, while also preserving alveolar bone structure without relying on adjacent teeth. However, implants are cost-intensive and may require an extended treatment timeline [90].

In terms of biomechanical properties, dental implants generally outperform other restorative methods, with fixed dentures ranking second and removable dentures exhibiting the least favorable characteristics. Implants exhibit the highest rates of survival and success, followed by fixed dentures, whereas removable dentures may compromise long-term outcomes due to various factors, including inadequate stability and challenges in maintenance. While removable dentures are the most economical option, fixed dentures and inlays follow, with implants representing the most expensive but delivering superior functional and aesthetic results.

Continuous technological innovation have enabled the combination of various restorative methods. For example, inlay-retained FPD provides a viable alternative for patients who decline implant treatment or traditional FPD due to aesthetic or other reasons [95]. The design of the bridge and connector is critical in such cases. FEA focuses on examining the stress levels and fracture resistance of bridges. For single missing tooth, a two-element resin-bonded FPD offers a feasible and minimally invasive treatment option [42]. Dental implant FPD is also a feasible and predictable treatment [69]. Additionally, for patients who have undergone root canal treatment, the use of internal crowns can help retain more tissue and reduce the risk of FPD fracture or debonding failure [9]. A hybrid prosthesis known as a fixed- removable denture combines a fixed component attached to abutment teeth with a removable component. Retention between the two is achieved through mechanical mechanisms, including friction, spring force, locking force, or magnetic retainers. Future research in FEA can focus on combined restoration technology. Simulation analysis can provide valuable insights into optimizing clinical strategies to meet aesthetic requirements while minimizing patient discomfort.

Nonetheless, there are notable limitations in the application of FEA to clinical practices. The accuracy of predictive FEA depends on various factors, including the geometric details of the object, boundary conditions, loads, and material properties [29]. FEA is frequently employed to model and simplify intricate biological systems, including the mechanical behavior of bone structures at both microscopic and macroscopic levels, the nonlinear characteristics of soft tissues, and the interactions between prosthetic materials and host tissues. This modeling process abstracts the actual biological environment, aiming to capture the primary mechanical features and response behaviors of the system through numerical calculations. When studying oral structures, it is also crucial to consider their deformation dynamics. Existing computer simulations often fall short of fully capturing the complexity of human biological behaviors. Most FEA studies assume that the properties of bone, teeth, and restoration materials are homogeneous, isotropic, and linearly elastic [40, 41, 88]. As stated in Section 3.2, while these assumptions enable reasonably accurate FEA results, clinical validation remains essential to ensure their applicability in real-world scenarios.

When developing FEA models, simplifying material properties is an essential step. Given the variety and complexity of biomaterials—such as the heterogeneity of bone, the viscoelastic nature of soft tissue, and the varying mechanical properties of prosthetic materials—models typically rely on average or idealized values for these properties. While this approach simplifies calculations, it may not accurately represent the true variability of material properties in clinical settings, particularly when accounting for individual differences, disease conditions, or age-related changes. Additionally, accurately simulating the biological characteristics of the occlusion process also presents challenges [96]. Loads settings in FEA are generally simplified to a single force, with only vertical or oblique load directions, fixed boundary conditions, and without considering the influence of multi-point forces. These limitations also do not invalidate the value of FEA as a predictive tool [50]. These limitations emphasize the necessity of thoughtful evaluation when using FEA in clinical scenarios, underscoring the significance of integrating FEA findings with clinical expertise and additional diagnostic methods. Further laboratory and clinical studies, including in vitro studies and long-term follow-ups, are necessary to validate the outcomes suggested by FEA [96]. For example, Ye et al. conducted a long-term follow-up study to assess the individual cases analyzed through FEA [47]. The findings showed no significant reduction in bone level or bone resorption around the implants, indicating that the implant design scheme was satisfactory in terms of function and denture stability.

The purpose of FEA is to provide dentists with an efficient tool for mechanical analysis and a reliable theoretical basis for studying the mechanical properties of restorations. In the future, developing new methods for the automatic generation and analysis of finite element models, potentially leveraging machine learning, will be essential for expanding its practical applications. In clinical practice, such advancements must focus on speed, high automation, and the ability to deliver accurate reference indices tailored to individual patient needs.

Author Contributions: Yan Wang and Liwen Chen contributed to the conception of the study; Liwen Chen contributed significantly to analysis and manuscript preparation; Yan Wang collected the data and analyzed it and wrote the manuscript; Yan Wang and Liwen Chen performed the analysis with constructive discussions.

References

[1] Chabouis HF, Faugeron VS, Attal JP. Clinical efficacy of composite versus ceramic inlays and onlays: A systematic review. Dent Mater 2013;29(12):1209-1218.

[2] Davenport JC, Basker RM, Heath JR, et al. The removable partial denture equation. Br Dent J 2000;189(8):414-424.

[3] Preshaw PM, Walls AWG, Jakubovics NS, et al. Association of removable partial denture use with oral and systemic health. J Dent 2011;39(11):711-719.

[4] Solberg K, Heinemann F, Pellikaan P, et al. Finite element analysis of different loading conditions for implant-supported overdentures supported by conventional or mini implants. Comput Methods Biomech Biomed Engin 2017;20(7):770-782.

[5] Qin SQ, Gao Z. Comparative evaluation of short or standard implants with different prosthetic designs in the posterior mandibular region: A three-dimensional finite element analysis study. Comput Methods Biomech Biomed Engin 2023;26(12):1499-1509.

[6] Oliveira H, Velasco AB, Ríos-Santos JV, et al. Effect of different implant designs on strain and stress distribution under non-axial loading: A three-dimensional finite element analysis. Int J Environ Res Public Health 2020;17(13):4738.

[7] Babaei B, Cella S, Farrar P, et al. The influence of dental restoration depth, internal cavity angle, and material properties on biomechanical resistance of a treated molar tooth. J Mech Behav Biomed Mater 2022;133:105305.

[8] Peskersoy C, Sahan HM. Finite element analysis and nanomechanical properties of composite and ceramic dental onlays. Comput Methods Biomech Biomed Engin 2022;25(14):1649-1661.

[9] Tribst JPM, Dal Piva AMdO, Muris J, et al. Endocrown fixed partial denture: Is it possible? J Prosthet Dent 2023;131(6):1118-1125.

[10] Osman MLM, Lim TW, Chang HC, et al. Structural integrity of anterior ceramic resin-bonded fixed partial denture: A finite element analysis study. J Funct Biomater 2023;14(2):108.

[11] Campaner LM, Silveira MPM, de Andrade GS, et al. Influence of polymeric restorative materials on the stress distribution in posterior fixed partial dentures: 3D finite element analysis. Polymers 2021;13(5):758.

[12] Kihara T, Shigeta Y, Ikawa T, et al. Designing anterior cantilever resin-bonded fixed dental prostheses based on finite element analysis. J Prosthodont Res 2023;67(3):418-423.

[13] Eijlers R, Utens E, Staals LM, et al. Systematic review and meta-analysis of virtual reality in pediatrics: Effects on pain and anxiety. Anesth Analg 2019;129(5):1344-1353.

[14] Ayoub A, Pulijala Y. The application of virtual reality and augmented reality in Oral & Maxillofacial Surgery. BMC Oral Health 2019;19(1):1-8.

[15] Monaghesh E, Negahdari R, Samad-Soltani T. Application of virtual reality in dental implants: A systematic review. BMC Oral Health 2023;23(1):603.

[16] Joda T, Gallucci GO, Wismeijer D, et al. Augmented and virtual reality in dental medicine: A systematic review. Comput Biol Med 2019;108:93-100.

[17] Chiari A, Mantovani S, Berzaghi A, et al. Load bearing capability of three-units 4Y-TZP monolithic fixed dental prostheses: An innovative model for reliable testing. Mater Des 2023;227:111751.

[18] Dawood A, Marti BM, Sauret-Jackson V, et al. 3D printing in dentistry. Br Dent J 2015;219(11):521-529.

[19] Tahayeri A, Morgan M, Fugolin AP, et al. 3D printed versus conventionally cured provisional crown and bridge dental materials. Dent Mater 2018;34(2):192-200.

[20] Della Bona A, Cantelli V, Britto VT, et al. 3D printing restorative materials using a stereolithographic technique: a systematic review. Dent Mater 2021;37(2):336-350.

[21] Stansbury JW, Idacavage MJ. 3D printing with polymers: Challenges among expanding options and opportunities. Dent Mater 2016;32(1):54-64.

[22] Han X, Yang D, Yang C, et al. Carbon fiber reinforced PEEK composites based on 3D-printing technology for orthopedic and dental applications. J Clin Med 2019;8(2):240.

[23] Schweiger J, Edelhoff D, Güth JF. 3D printing in digital prosthetic dentistry: An overview of recent developments in additive manufacturing. J Clin Med 2021;10(9):2010.

[24] Rigotti RLD, Tardelli JDC, Dos Reis AC, et al. Influence of dental implant/mini-implant design on stress distribution in overdentures: A systematic review. Oral Maxillofac Surg

2023;1-13.
[25] Falcinelli C, Valente F, Vasta M, et al. Finite element analysis in implant dentistry: State of the art and future directions. Dent Mater 2023;39(6):539-556.
[26] Lisiak-Myszke M, Marciniak D, Bieliński M, et al. Application of finite element analysis in oral and maxillofacial surgery-a literature review. Materials 2020;13(14):3063.
[27] Courant R. Variational methods for the solution of problems of equilibrium and vibrations. Bull Am Math Soc 1943;1-23.
[28] Brekelmans W, Poort H, Slooff T. A new method to analyse the mechanical behaviour of skeletal parts. Acta Orthop Scand 1972;43(5):301-317.
[29] Choi AH, Conway RC, Ben-Nissan B. Finite-element modeling and analysis in nanomedicine and dentistry. Nanomedicine 2014;9(11):1681-1695.
[30] Thresher RW, Saito GE. The stress analysis of human teeth. J Biomech 1973;6(5):443-449.
[31] Farah J, Craig RG, Sikarskie DL. Photoelastic and finite element stress analysis of a restored axisymmetric first molar. J Biomech 1973;6(5):511-514, IN519, 515-520.
[32] Chabanas M, Luboz V, Payan Y. Patient specific finite element model of the face soft tissues for computer-assisted maxillofacial surgery. Med Image Anal 2003;7(2):131-151.
[33] Marcin M, Katarzyna C. Stress and displacement patterns during orthodontic intervention in the maxilla of patients with cleft palate analyzed by finite element analysis: A systematic review. BMC Oral Health 2023;23(1):93.
[34] Ferguson BM, Entezari A, Fang J, et al. Optimal placement of fixation system for scaffold-based mandibular reconstruction. J Mech Behav Biomed Mater 2022;126:104855.
[35] Wu AYJ, Hsu JT, Fuh LJ, et al. Biomechanical effect of implant design on four implants supporting mandibular full-arch fixed dentures: In vitro test and finite element analysis. J Formos Med Assoc 2020;119(10):1514-1523.
[36] Lee H, Park S, Noh G. Biomechanical analysis of 4 types of short dental implants in a resorbed mandible. J Prosthet Dent 2019;121(4):659-670.
[37] Schillinger D, Ruess M. The finite cell method: A review in the context of higher-order structural analysis of CAD and image-based geometric models. Arch Comput Methods Eng 2015;22(3):391-455.
[38] Müzel SD, Bonhin EP, Guimarães NM, et al. Application of the finite element method in the analysis of composite materials: A review. Polymers 2020;12(4):818.
[39] Liu WK, Li SF, Park HS. Eighty years of the finite element method: Birth, evolution, and future. Arch Comput Methods Eng 2022;29(6):4431-4453.
[40] Darwich MA, Aljareh A, Alhouri N, et al. Biomechanical assessment of endodontically treated molars restored by endocrowns made from different CAD/CAM materials. Materials 2023;16(2):764.
[41] Demirel MG, Mohammadi R, Keçeci M. Crack propagation and fatigue performance of partial posterior indirect restorations: An extended finite element method study. J Funct Biomater 2023;14(9):484.
[42] Chen YC, Fok A. Shape optimization of a 2-unit cantilevered posterior resin-bonded fixed dental prosthesis. J Prosthet Dent 2023;129(1):181-190.
[43] Guler MS. Low and high viscosity bulk-fill composite resins stress distribution in primary molar tooth inlay cavity. Comput Methods Biomech Biomed Engin 2024;27(3):411-418.
[44] Burkhart TA, Andrews DM, Dunning CE. Finite element modeling mesh quality, energy balance and validation methods: A review with recommendations associated with the modeling of bone tissue. J Biomech 2013;46(9):1477-1488.
[45] Cheng CW, Chen WP, Chien YT, et al. Biomechanical behavior of cavity design on teeth restored using ceramic inlays: An approach based on three-dimensional finite element analysis and ultrahigh-speed camera. Acta Biomater 2019;89:382-390.
[46] Zhu J, Luo D, Rong Q, et al. Effect of biomimetic material on stress distribution in mandibular molars restored with inlays: A three-dimensional finite element analysis. PeerJ 2019;7:E7694.
[47] Ye Z, Ye H, Jiang Z, et al. Influence of different implant designs on replacement of four teeth of the posterior free-end edentulism: Three-dimensional finite element analysis and clinic case validation. Ann Anat 2023;249:152111.
[48] Kim S-Y, Kim B-S, Kim H, et al. Occlusal stress distribution and remaining crack propagation of a cracked tooth treated with different materials and designs: 3D finite element analysis. Dent Mater 2021;37(4):731-740.
[49] Alp Ş, Gulec Alagoz L, Ulusoy N. Effect of direct and indirect materials on stress distribution in class II MOD restorations: A 3D-finite element analysis study. Biomed Res Int 2020; 2020:7435054.
[50] Melo de Matos JD, Scalzer Lopes GdR, Nomura Nakano LJ, et al. Biomechanical evaluation of 3-unit fixed partial dentures on monotype and two-piece zirconia dental implants. Comput Methods Biomech Biomed Engin 2022;25(3):239-246.
[51] Chakraborty A, Datta P, Majumder S, et al.

Finite element and experimental analysis to select patient's bone condition specific porous dental implant, fabricated using additive manufacturing. Comput Biol Med 2020;124:103839.

[52] Ortiz-Puigpelat O, Lázaro-Abdulkarim A, de Medrano-Reñé JM, et al. Influence of implant position in implant-assisted removable partial denture: A three-dimensional finite element analysis.J Prosthodont 2019;28(2):E675-E681.

[53] Gönder HY, Demirel MG, Mohammadi R, et al. The effects of using cements of different thicknesses and amalgam restorations with different young's modulus values on stress on dental tissue: An investigation using finite element analysis. Coatings 2023;13(1):6.

[54] Gönder HY, Fidancioğlu YD, Fidan M, et al. Comparison of resin cement's different thicknesses and poisson's ratios on the stress distribution of class II amalgam restoration using finite element analysis. Appl Sci 2023;13(7):4125.

[55] He J, Zheng Z, Wu M, et al. Influence of restorative material and cement on the stress distribution of endocrowns: 3D finite element analysis. BMC Oral Health 2021;21:1-9.

[56] Xiong Y, Xu Z, Yu P, et al. Influence of adhesive layer thickness on wear damage of the bonded interface between enamel and composite inlays. Tribol Lett 2021;69:1-15.

[57] Dejak B, Młotkowski A. A comparison of mvM stress of inlays, onlays and endocrowns made from various materials and their bonding with molars in a computer simulation of mastication - FEA. Dent Mater 2020;36(7):854-864.

[58] Paolone G, Mandurino M, De Palma F, et al. Color stability of polymer-based composite CAD/CAM blocks: A systematic review. Polymers (Basel) 2023;15(2):464.

[59] Awada A, Nathanson D. Mechanical properties of resin-ceramic CAD/CAM restorative materials. J Prosthet Dent 2015;114(4):587-593.

[60] Zarone F, Di Mauro MI, Ausiello P, et al. Current status on lithium disilicate and zirconia: a narrative review. Bmc Oral Health 2019;19(1):1-14.

[61] Moxon R, Xu ZG, Tettey F, et al. Dental metal matrix composites: The effects of the addition of titanium nanoparticle particles on dental amalgam. Materials 2024;17(7):1662.

[62] Heintze S, Monreal D, Reinhardt M, et al. Fatigue resistance of all-ceramic fixed partial dentures–Fatigue tests and finite element analysis. Dent Mater 2018;34(3):494-507.

[63] Canipaner LM, Pinto Metro ADO, Mendes Tribst JP, et al. Loading stress distribution in posterior teeth restored by different core materials under fixed zirconia partial denture: A 3D-FEA study. Am J Dent 2021;34(3):157-162.

[64] Ceddia M, Lamberti L, Trentadue B. FEA comparison of the mechanical behavior of three dental crown materials: Enamel, ceramic, and zirconia. Materials 2024;17(3):673.

[65] Han W, Kim J-H, Kwon H-B, et al. Effect of cyclic thermal stress on the fatigue life of teeth restored with gold inlay. Dent Mater J 2022;41(4):567-572.

[66] She Y-H, Zhang Y-Y, Liu Y-X, et al. Three-dimensional finite element analysis of cuspal-coverage thickness influence on the stress distribution of all-ceramic onlay-restored premolars. Hua xi kou qiang yi xue za zhi 2019;37(6):636-641.

[67] Ebadian B, Fathi A, Tabatabaei S. Stress distribution in 5-unit fixed partial dentures with a pier abutment and rigid and nonrigid connectors with two different occlusal schemes: A three-dimensional finite element analysis. Int J Dent 2023;2023(1):3347197.

[68] Kumar SM, Jei JB, Krishnan M. An investigation of effect of rigid and nonrigid connector designs on implant as pier abutment in implant tooth-supported fixed dental prosthesis with three-dimensional finite element analysis: An in vitro study. Adv Hum Biol 2021;11(4):S45-S50.

[69] Naguib GHH, Hashem ABH, Abougazia A, et al. Effect of non-rigid connector on the stress distribution of tooth-implant supported fixed prostheses using different implant length and diameter: A comparative 3D finite element study. J Prosthodont 2023;32(6):E129-E138.

[70] Huang LS, Huang YC, Yuan CD, et al. Biomechanical evaluation of bridge span with three implant abutment designs and two connectors for tooth-implant supported prosthesis: A finite element analysis. J Dent Sci 2023;18(1):248-263.

[71] Naguib GH, Hashem ABH, Natto ZS, et al. The effect of implant length and diameter on stress distribution of tooth-implant and implant supported fixed prostheses: An in vitro finite element analysis study. J Oral Implantol 2023;49(1):46-54.

[72] Eazhil R, Swaminathan SV, Gunaseelan M, et al. Impact of implant diameter and length on stress distribution in osseointegrated implants: A 3D FEA study. J Int Soc Prev Community Dent 2016;6(6):590-596.

[73] Shash M, Nazha H, Abbas W. Influence of different abutment designs on the biomechanical behavior of one-piece zirconia dental implants and their surrounding bone: a 3D-FEA. Irbm 2019;40(6):313-319.

[74] Poovarodom P, Rungsiyakull C, Suriyawanakul

J, et al. Multi-objective optimization of custom implant abutment design for enhanced bone remodeling in single-crown implants using 3D finite element analysis. Sci Rep 2024;14(1):15867.
[75] Lee H, Jo M, Sailer I, et al. Effects of implant diameter, implant-abutment connection type, and bone density on the biomechanical stability of implant components and bone: A finite element analysis study. J Prosthet Dent 2022;128(4):716-728.
[76] Leblebicioğlu Kurtuluş I, Kilic K, Bal B, et al. Finite element analysis of the stress distribution associated with different implant designs for different bone densities. J Prosthodont 2022;31(7):614-622.
[77] Cheng KC, Liu PH, Chen HS, et al. Stress distribution of four-unit implant-supported fixed partial prosthesis with different numbers and positions of fixtures in maxilla anterior region-3D FEA. J Med Biol Eng 2022;42(4):526-533.
[78] Güzelce E. Biomechanical comparison of diffirent framework materials in mandibular overdenture prosthesis supported with implants of different sizes: A finite element analysis. BMC Oral Health 2023;23(1):450.
[79] Huang HL, Tsai HL, Wu YL, et al. Biomechanical evaluation of bone atrophy and implant length in four implants supporting mandibular full-arch-fixed dentures. Materials 2022;15(9):3295.
[80] Skiedraitė I, Diliūnas S, Varinauskas V. A study case of short dental implants loading in the all-on-4® system on fixed dental prostheses: A finite element analysis. Mechanika 2023;29(2):104-109.
[81] Kucukkurt S, Konuklu R, Çelikkol Ş. Is it beneficial to support interforaminal implant placement techniques with ultra-short implants in the posterior region? A 3D finite element analysis. J Osseointegration 2023;15(2):95-104.
[82] Sharaf MA, Jiang J, Wang S, et al. Clinical and patient-centered outcomes following rehabilitation of atrophic edentulous maxilla using six implants placed simultaneously with bilateral maxillary sinus augmentation: A retrospective case series. J Stomatol Oral Maxillofac Surg 2023;124(6S):101480.
[83] Sun X, Cheng KJ, Liu YF, et al. Biomechanical comparison of all-on-4 and all-on-5 implant-supported prostheses with alteration of anterior-posterior spread: a three-dimensional finite element analysis. Front Bioeng Biotechnol 2023;11:1187504.
[84] Wang PS, Tsai MH, Wu YL, et al. Biomechanical analysis of titanium dental implants in the all-on-4 treatment with different implant-abutment connections: A three-dimensional finite element study. J Funct Biomater 2023;14(10):515.
[85] Yu P, Xiong YH, Zhao P, et al. On the wear behavior and damage mechanism of bonded interface: Ceramic vs resin composite inlays. J Mech Behav Biomed Mater 2020;101:103430.
[86] Campos CH, Gonçalves TMSV, Garcia RCMR. Implant-supported removable partial denture improves the quality of life of patients with extreme tooth loss. Braz Dent J 2015;26(5):463-467.
[87] Ichikawa H, Yoda N, Ogawa T, et al. Impact of implant location on load distribution of implant-assisted removable partial dentures: a review of in vitro model and finite-element analysis studies. Int J Implant Dent 2023;9(1):31.
[88] Scalzer Lopes GdR, Melo de Matos JD, Queiroz DA, et al. Influence of abutment design on biomechanical behavior to support a screw-retained 3-unit fixed partial denture. Materials 2022;15(18):6235.
[89] Lee H, Jo M, Noh G. Biomechanical effects of dental implant diameter, connection type, and bone density on microgap formation and fatigue failure: A finite element analysis. Comput Methods Programs Biomed 2021;200:105863.
[90] Ibrahim CRM, Sameh A, Askar O. A finite element analysis study on different angle correction designs for inclined implants in all-on-four protocol. BMC Oral Health 2024;24(1):331.
[91] Rajaeirad M, Fakharifar A, Posti MHZ, et al. Evaluating the effect of functionally graded materials on bone remodeling around dental implants. Dent Mater 2024;40(5):858-868.
[92] Silva LS, Bento VAA, Brunetto JL, et al. Polyetheretherketone materials for removable partial denture frameworks: an integrative review. Gen Dent 2023;71(4):58-62.
[93] Rodrigues MT, Gowda BHH, Alva B. Stress distribution in tooth supported removable partial denture fabricated using two different materials: A 3-dimensional finite element analysis. Mater Today: Proc 2021;46:7643-7650.
[94] Kasem ATT, Elsherbiny AA, Abo-Madina M, et al. Biomechanical behavior of posterior metal-free cantilever fixed dental prostheses: effect of material and retainer design. Clin Oral Investig 2023;27(5):2109-2123.
[95] Tatarciuc M, Maftei GA, Vitalariu A, et al. Inlay-retained dental bridges-A finite element analysis. Appl Sci 2021;11(9):3770.
[96] Yu WQ, Chen SY, Ma L, et al. Biomechanical analysis of different framework design, framework material and bone density in the edentulous mandible with fixed implant-

supported prostheses: A three-dimensional finite element study. J Prosthodont 2023;32(4):309-317.
[97] Lyu H, Murakami N, Yamazaki T, et al. Evaluation of PEEK and zirconia occlusal rest designs for removable partial dentures based on finite element analysis. J Prosthodont Res 2023;67(2):196-205.
[98] Tribst JPM, Dal Piva AMD, Syed AU, et al. Comparative stress analysis of polyetherketoneketone (PEKK) telescopic crowns supported by different primary crown materials. Appl Sci 2022;12(7):3446.
[99] Waldecker M, Rues S, Rammelsberg P, et al. Validation of in-vitro tests of zirconia-ceramic inlay-retained fixed partial dentures: A finite element analysis. Dent Mater 2019;35(3):E53-E62.