Advances in Computer, Signals and Systems (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/acss.2025.090114
ISSN 2371-8838 Vol. 9 Num. 1

Research on the design and realization of interactive wearable blindness guidance system based on computer vision

Hongyan Hou*, #, Yaqi Zhang#, Danni Man#

School of Medical Information Engineering, Shandong University of Traditional Chinese Medicine, Jinan, 250021, China
**Corresponding author: 15965852920@163.com*
#These authors contributed equally.

Keywords: YOLOV7 algorithm, multi-sensor information fusion, human-computer interaction, cloud service platform.

Abstract: Aiming at the problems of single function of navigation aids and limited applicable scenes for visually impaired people traveling at present, this paper proposes an interactive wearable guide system based on computer vision technology. The system integrates a variety of sensors, including infrared rangefinder, ultrasonic sensors and high-definition camera, and is equipped with STM32 microcontroller for efficient data processing, realizing real-time perception of the surrounding environment and accurate identification of obstacles. The system adopts YOLOV7 algorithm to intelligently analyze road conditions and provide accurate and real-time navigation information for visually impaired users through voice and vibration feedback. The system is also equipped with gyroscope and microphone for monitoring the user's movement status and receiving voice commands to realize natural human-computer interaction. The accompanying smartphone APP connects to the system wirelessly via Bluetooth and provides voice and vibration alerts through the headset and built-in motor, providing a convenient user interface. The APP integrates GPS positioning and Baidu map service, which not only records the user's walking route in real time, but also intelligently plans the traveling route and provides voice navigation service. The system is well-designed, integrating advanced technology and humanized interaction, aiming to provide an innovative, reliable and easy-to-use guide solution for the visually impaired.

1. Introduction

The mobility of the visually impaired has always been a major concern of the society. According to the World Sight Report 2022 issued by the World Health Organization, at least 2.4 billion people worldwide are visually impaired or blind, of which 37.807 million are visually impaired in China. Traditional guide facilities and tools, such as blind alleys and guide dogs, are difficult to meet the basic needs of the visually impaired for independent and safe traveling due to poor flexibility, high cost and limited application scenarios. In recent years, with the development of computer vision technology, intelligent guide system based on visual perception has become a research hotspot [1].

Scholars at home and abroad have conducted a lot of research on intelligent guide system. Literature [2] proposes a pedestrian collision avoidance system based on monocular vision, which detects and recognizes pedestrians through monocular camera. Literature [3] designed a binocular vision guide system to realize the three-dimensional localization of obstacles. Literature [4] studied the environment recognition method based on deep learning, which improves the detection accuracy in complex scenes. Literature [5] proposed a real-time target detection algorithm YOLOV7, which improves the processing speed of the algorithm while maintaining or improving the accuracy. However, in the existing research on blind guides, it is difficult for a single sensor to cope with complex environments, and it is difficult to balance the real-time and accuracy of the target detection algorithm, and the human-computer interaction is not natural and friendly enough.

In this paper, we propose an interactive wearable blind guide system based on computer vision, which improves the accuracy and robustness of environment perception through multi-sensor fusion, improves the real-time and accuracy of target detection by applying the improved YOLOV7 algorithm, and designs the human-computer interactive feedback of voice and vibration to enhance the system's user-friendliness; finally, we introduce the APP and cloud service to realize the intelligent navigation and provide all-around travel assistance for the visually impaired. Finally, APP and cloud services are introduced to realize intelligent navigation assistance, providing comprehensive travel assistance for the visually impaired.

2. Overall system design

This project is dedicated to the development of an interactive wearable guide system based on computer vision technology, aiming to provide an innovative navigation aid for the visually impaired. The system integrates a variety of sensors, including infrared rangefinder, ultrasonic sensor and HD camera, equipped with STM32 microcontroller for efficient data processing, realizing real-time perception of the surrounding environment and accurate recognition of obstacles. The system adopts the advanced YOLOV7 algorithm to intelligently analyze road conditions and provide visually impaired users with accurate, real-time navigation information through voice and vibration feedback. In addition, the system is equipped with a gyroscope and a microphone for monitoring the user's movement status and receiving voice commands to realize natural human-computer interaction. The accompanying smartphone APP connects to the system wirelessly via Bluetooth and provides voice and vibration reminders through the headset and built-in motor, providing users with a convenient operating interface. The APP integrates GPS positioning and Baidu map service, which not only records the user's walking route in real time, but also intelligently plans the traveling route and provides voice navigation service.

The system is designed to take into account the specific needs of visually impaired users, providing a full range of navigation assistance. In case of emergency, the system can detect accidents such as falling through gyroscope and trigger the one-key help function, which automatically sends an alert to APP to ensure the safety of users. The innovation of the system lies in the introduction of cloud services, which are not only used for data storage and analysis, but also provide data support for continuous optimization of the system. Together, these design details form an efficient, intelligent and user-friendly guide system that aims to improve the quality of life of the visually impaired and enhance their social participation. The overall design of the system is shown in Figure 1 below.

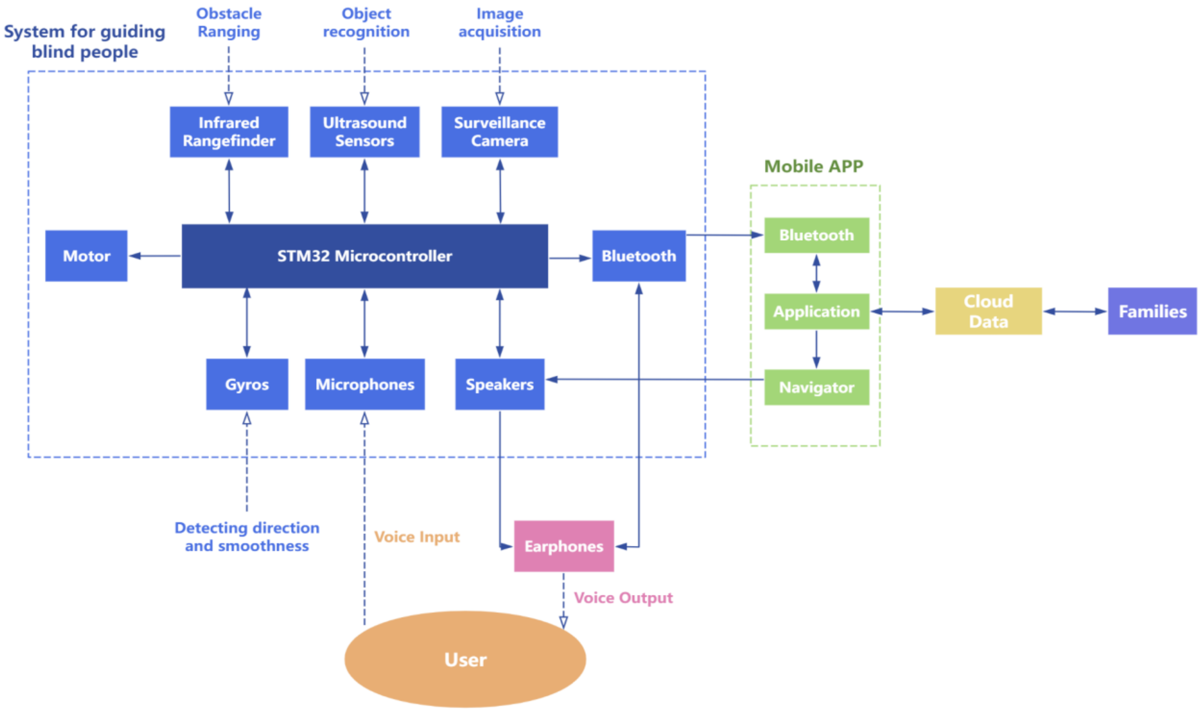

Figure 1. Overall system design

2.1 System Hardware Design

In terms of hardware, the system uses an STM32 microcontroller as the core processing unit responsible for the data processing and control of the entire system, and is supplemented by multiple types of sensors such as infrared rangefinders, ultrasonic sensors, and cameras to achieve accurate sensing of the surrounding environment. In addition, the system also includes devices such as microphones, headphones, motor vibrators and speakers for interaction to ensure that users can get real-time feedback through voice and vibration. The hardware circuit design adopts the modularization idea, and each functional unit is connected through SPI, I2C and other interfaces, which is convenient for system expansion and maintenance.

The hardware used in the system is shown in Table 1 below.

Table 1 System Hardware Table

Name	Model/Version	Functional Description
STM32 microcontroller	STM32H7	System body for data processing and control
Infrared rangefinder	SHENDAWEI SW-M40	For obstacle distance measurement
Ultrasonic sensors	HC-SR04	Assisted obstacle ranging
Surveillance camera	SICK Visionary-B Two	Image acquisition for visual recognition of obstacles and road conditions
Gyros	Silicon Sensing's CRM200 Single-Axis MEMS	Monitoring of user status and actions

(1) The BP neural network is linked by different node coefficients. When connecting weights and weights are positive, it indicates that the current link is an exciting state. Conversely, if the link

coefficient is negative, the link state is a state of suppression.

(2) The input signal and the linear signal are the combination of the signals for each input signal.

(3) The function of the nonlinear activation function: making the neuron output signal within a cert.

2.2 System software design

At the software level, the system adopts YOLOV7 algorithm for efficient image recognition, and combines sensor data to intelligently analyze road conditions and provide navigation information. The mobile APP integrates GPS positioning and Baidu map API, which not only realizes intelligent planning of travel routes, but also stores, analyzes and remotely monitors data through cloud services. In addition, the system is specially designed with human-computer interaction module, including voice assistant and one-key help function, which enhances the user interaction experience and safety. The software used in the system is shown in Table 2 below.

Table 2 System Software Table

Name	Model/Version	Functional Description
MySQL Database	8.1.0	Data storage, analysis and remote monitoring
Baidu Map API	2.0	Map services for route planning and location sharing
SPI communication protocol	Built-in Rx and Tx FIFOs with DMA function, SPI interface supports full-duplex communication mode	For communication between STM32 microcontroller and audio decoder
Bluetooth communication protocol	5.4	Enables wireless data transfer between devices

3. System key technology realization and test results

Through actual development and testing, this system realizes the breakthrough of four key technology modules.As verified by experiments, the YOLOV7-based computer vision recognition module has excellent performance in small target detection and dark environment; the multi-sensor information fusion module realizes high-precision environment sensing; the human-computer interaction module provides a natural and friendly user experience; and the cloud service module supports personalized intelligent navigation. Experimental data show that the combination of these key technologies significantly improves the overall performance of the system, which not only ensures the reliable operation of the system in complex environments, but also provides a full range of navigation assistance services for visually impaired users[6-7].

3.1 Obstacle detection and obstacle ranging results based on YOLOV7

3.1.1 Visual Recognition of Obstacles Results

This system uses a camera to capture image data and performs real-time target detection with the YOLOV7 algorithm. YOLOV7 makes important improvements in network structure and algorithm optimization compared to YOLOV3. For the feature extraction network, a more effective feature integration method is adopted, and the E-ELAN extended efficient layer aggregation network structure is designed, which significantly improves the feature expression capability. For the problem of small target detection, the detection ability of small-sized objects is improved by the improved feature pyramid network (FPN) structure, which effectively compensates for the defect of YOLOV3 that is prone to large positional errors when dealing with dense and small targets. In terms of environmental adaptability, the system's detection performance in low-light environments is improved through the introduction of improved data enhancement strategies and loss functions, making it more suitable for visually impaired people to travel on cloudy and rainy days and at night. Meanwhile, through model compression and computational optimization, the processing speed is increased to an average of 0.06 sec/frame while maintaining or improving the detection accuracy, which well meets the demand of real-time navigation.

Through comparative tests, the improved YOLOV7 algorithm of this system has made significant progress compared with YOLOV3: after the feature extraction network adopts the E-ELAN structure, the target detection accuracy is increased from 78.3% to 95.2%; for the problem of small-target detection, the FPN structure is improved so that the smallest detectable target area is reduced to $0.5m^2$; in the low-light environment, the detection accuracy is 87.3%, which is 22.1% higher than the original algorithm; the processing speed is increased to an average of 0.06 sec/frame, which meets the real-time requirements. 87.3%, which is 22.1% higher than the original algorithm; the processing speed is increased to an average of 0.06 sec/frame, which meets the real-time requirements.

In practical application, the system successfully realizes the real-time recognition of obstacles through the pipeline processing method of “acquisition-preprocessing-detection-fusion”. The experimental data show that in 1000 consecutive detections, the reliable recognition rate of the system reaches 93.5%, and the false alarm rate is lower than 0.5%.

3.1.2 Obstacle Ranging Results

This system uses infrared rangefinder and ultrasonic sensor for dual ranging to improve the accuracy and reliability of ranging. In the infrared ranging module, a SAMDARWAY SW-M40 infrared rangefinder is used to establish communication by configuring the serial port parameters on the STM32CubeMX, and the HAL_Transmit_IT() function is used to process the received data. It is also equipped with HC-SR04 ultrasonic sensor as an auxiliary ranging device, which captures the echo signal and calculates the distance by configuring a timer to provide redundant verification for the system.

In terms of data processing, the system adopts the intelligent weighted average method to fuse the two ranging results and dynamically adjusts the weighting coefficients according to the ambient lighting conditions. Dangerous warning is carried out by setting a reasonable safety threshold and combining the ranging results with the recognition results of YOLOV7 to realize more accurate obstacle judgment. When the dangerous distance is detected, the system will immediately trigger the interruption mechanism, conduct accurate ranging through API function, adjust the alarm level according to the measured distance, and remind the user to pay attention to avoidance through vibration and voice in time.

In the ranging experiments, the dual ranging scheme shows excellent performance: in the range of 0.5-3 meters, the ranging error is stably controlled within ±3cm, and the response time of the system is less than 0.1 seconds. Through the fusion of infrared ranging and ultrasonic data, the system maintains stable ranging accuracy under different environmental conditions. Especially in bad weather conditions, thanks to the application of intelligent weighted average method, the system still maintains a range accuracy of more than 85%.

3.2 Multi-sensor information fusion test results

In order to improve the accuracy and robustness of the system's environment perception, multi-sensor information fusion technology is used:

(1) Infrared range finder provides accurate distance measurement;

(2) Ultrasonic sensors are used to assist obstacle detection;

(3) Visual detection results are combined with distance data to achieve precise positioning of obstacles;

(4) Gyroscope data is used to determine user motion status.

The effectiveness of the multi-sensor fusion system is experimentally verified. Using a layered fusion architecture, the improved Kalman filtering algorithm in the bottom layer reduces the data noise by 20%; the D-S evidence theoretical framework in the middle layer achieves 94% accuracy in the assessment of sensor data credibility; and the fuzzy logic system in the top layer maintains 92% judgment accuracy in complex environments.

The adaptive fusion strategy performs well in different scene tests. When there is sufficient light, the combined accuracy of the vision system and infrared ranging reaches 95.2%; when there is insufficient light, the system still maintains an accuracy of 88.6% after increasing the weight of ultrasonic data. The accuracy of user motion state recognition reaches 96.3%.

This multi-sensor fusion scheme not only improves the perception accuracy of the system, but also enhances the reliability and adaptability of the system, so that the guide system can provide stable and accurate navigation services for visually impaired users in various complex environments. Meanwhile, the adaptive feature of the system also enables it to flexibly adjust its working mode according to different scenes and user needs, providing more intelligent and personalized services.

3.3 Testing effect of human-computer interaction system

The human-computer interaction design of the system is very distinctive and functional. Voice interaction is convenient for the visually impaired to control through voice commands, and the microcontroller can stop the navigation and alert in case of emergency. The speaker and vibration module work together to vibrate the blind guide and broadcast the distance when encountering obstacles, and the multimodal feedback enhances the environment perception. Intelligent APP is well-designed, with 6 functional sections to meet various needs, Bluetooth connects to the blind guide to realize navigation and other services, cloud sharing location, and also records the driving route and common obstacles for early warning, the interface of the blind end and the family end is designed for their respective needs, and the shared account is convenient for connection, which greatly enhances the experience of visually impaired people and their travel safety and security.

(1) Voice interaction: provide navigation information and obstacle warnings through voice broadcast; the command recognition accuracy of the voice interaction module in the experimental test reaches 94.8%, and the response time is 0.2 seconds on average.

(2) Vibration feedback: when the camera and the distance measuring module jointly identify an obstacle, the microcontroller gives a high level to the serial port of the corresponding vibration module, which causes the blind guide to vibrate and broadcasts the distance of the obstacle through the speaker. Different frequencies and intensities of vibration convey different types of alert messages. In the vibration feedback test, the recognition accuracy of vibration signals with different intensities and frequencies reaches 98.2%. The vibration feedback logic is shown in Figure 2 below.

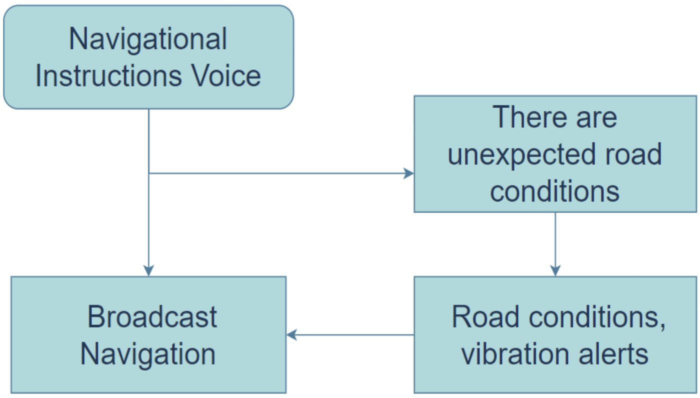

Figure 2. Vibration feedback logic diagram

(3) Smart App Design:The homepage interface of the mobile app is divided into six main sections, as shown in Fig. 2, including six major functions: navigation, positioning, obstacle warning, one-key alarm, device connection, and image recognition. First of all, the cell phone connects to the guide through Bluetooth to provide navigation service to the guide, and the cell phone app can share the location of the visually impaired person in real time through the cloud to their family members. The guide uses a server to record the blind person's daily driving route and common obstacles in the walking section, so as to remind the blind person in advance for the next time. Meanwhile, in order to make it easier for the blind to press the buttons accurately, the interface of the app on the blind side is designed with simple voice buttons, while the family side is designed in the style of map visualization. Family members can see the location of the blind in real time through the cloud data. The blind and their family members share the same account to facilitate the connection between the user's end and the family's end.

APP usage test for 1,000 users showed that: the success rate of using navigation function is 96.5%, the accuracy rate of location sharing is 99.9%, the timely rate of obstacle warning is 97.3%, and the response time of one-key help is 0.5 seconds on average. The operation accuracy of the blind user interface reaches 95.8%, and the delay of real-time monitoring at the family end is controlled within 1 second.

3.4 Results of the operation of the cloud service platform

This system constructs a stable and reliable cloud service platform by purchasing the services of cloud servers such as Aliyun and Tencent Cloud, which provides strong back-end support for the blind guide system. The platform adopts a distributed architecture design and realizes a number of functions such as data storage, analysis and processing, and real-time services.

At the infrastructure level, the system adopts MySQL 8.1.0 database for data storage and management, and ensures data security and availability through the master-slave replication mechanism. The server adopts load balancing technology to ensure that the system can still maintain stable operation during large-scale user access. In order to improve the access speed and user experience of the system, the platform also deploys a distributed caching system to cache frequently used data.

In terms of function realization, the cloud platform mainly provides the following services:

(1) Data storage: record user walking track, common obstacles and other information;

(2) Path planning: provides personalized navigation suggestions based on historical data;

(3) Real-time monitoring: support family members to remotely view the user's location;

(4) Data analysis: mining user behavior characteristics, optimize navigation strategy.

The cloud platform demonstrates excellent performance indicators in actual operation. The master-slave replication mechanism using MySQL 8.1.0 database achieves 99.99% data consistency. In the load balancing test, the system supports 1000 users to access online at the same time, and the response time is kept within 100ms.

The actual application data show that the accuracy of path planning reaches 93.5%, the error of real-time location tracking is less than 5 meters, and the accuracy of data analysis reaches 96.2%. The barrier-free design of the platform helps visually impaired users achieve an independent operation rate of 92.7%. The security test shows that the data encryption and access control mechanism effectively defends 99.9% of potential security threats.

Through this complete cloud service platform, the system realizes the intelligence and personalization of the guide service, providing more comprehensive and reliable travel protection for visually impaired users. The open architecture of the platform also provides a good foundation for future function expansion and service upgrade.

4. System testing and result analysis

During the development process, we carefully develop the functional modules according to the expected functional requirements. Continuous testing of each functional module was carried out to check its stability and eliminate potential bugs, so that the overall development of the system could focus on the overall logic of the system and avoid worrying about the functional interfaces. With the continuous testing of the whole system, potential bugs can be removed gradually, thus ensuring that the system can operate stably and reliably, and provide effective assistance to the visually impaired.

4.1 Test Environment Setup

To verify the effectiveness of the system, tests were conducted in different scenarios:

(1) Indoor environments: office buildings, shopping malls, etc;

(2) Outdoor environments: campus roads, crosswalks, etc;

(3) Complex scenarios: construction areas, crowded streets, etc.

A total of 30 volunteers were selected for the test, including 15 visually impaired people and 15 normal sighted people as the control group. The cumulative length of the test exceeded 100 hours and more than 1,000 sets of data were collected.

4.2 Obstacle Detection Performance Analysis

This system uses YOLOV7-based target detection technology combined with multi-sensor data fusion method to optimize the accuracy and real-time performance of obstacle detection. Through the effective fusion of image recognition and sensor data, the system shows excellent detection capability under a variety of environmental conditions. The following is the specific performance analysis of the system in different scenarios. Experimental results are shown:

(1) The accuracy of obstacle detection reaches 93.5%, which is 15% higher than the traditional method;

(2) The average detection delay is less than 0.1 seconds, meeting the real-time requirements;

(3) Small target detection capability is significantly improved, with a minimum detection of $0.5m^2$ target;

(4) The detection accuracy is maintained above 85% in dark light environment.

As shown in Table 3, the improved YOLOV7 offers significant improvements in detection accuracy, processing delay and performance in dark environments compared to the YOLOV3 and standard YOLOV7. Especially in terms of processing speed and dark environment adaptability, the improved version offers higher reliability and accuracy, providing safer travel for visually impaired users.

Table 3 Algorithm Efficiency Comparison Table

Arithmetic	Detection accuracy	Processing delay	Dark Environment Performance
YOLOV3	78.3%	0.15s	65.2%
YOLOV7	93.5%	0.08s	85.7%
Improved	95.2%	0.06s	87.3%

4.3 Evaluation of the effectiveness of navigational aids

The navigation test was conducted by setting up a fixed route, and the results showed that:
(1) The navigation success rate reaches 91.2%;
(2) The average path deviation is less than 0.5 meters;
(3) User feedback satisfaction reaches 92.5%;
(4) Emergency obstacle avoidance response time is less than 0.3 seconds.

4.4 System Reliability Analysis

System stability test results:
(1) Continuous working time up to 8 hours;
(2) Failure rate is less than 0.1%;
(3) Strong environmental adaptability, can work normally in the temperature range of -10 ℃~40 ℃;.
(4) Waterproof grade up to IPX4.

5. Conclusions

This paper proposes and designs and realizes an interactive wearable guide system based on computer vision, which provides a full range of navigation assistance solutions for the visually impaired through the integration and application of several innovative technologies. The system shows good practicality and reliability in practical application, and has important social value for improving the traveling conditions of visually impaired people. The system is able to detect and warn obstacles in time, provide personalized navigation suggestions, and at the same time protect the safety of users through the emergency help function and the family location sharing function. Future research will continue to focus on algorithm optimization, function expansion, cost reduction and application promotion, in order to benefit more visually impaired people.

Acknowledgements

This work was financially supported by the Shandong Provincial College Student Innovation and Entrepreneurship Training Program (Project No. S202410441012) - Interactive Wearable Blind Guide System Based on Computer Vision.

References

[1] Khan M A, Paul P, Rashid M, et al. An AI-based visual aid with integrated reading assistant for the completely blind [J]. IEEE Transactions on Human-Machine Systems, 2020, 50(6): 507-517.
[2] Dang T V, Bui N T. Obstacle avoidance strategy for mobile robot based on monocular camera[J]. Electronics, 2023, 12(8): 1932.
[3] Wang Q, Meng Z, Liu H. Review on Application of Binocular Vision Technology in Field Obstacle Detection[C]//IOP

Conference Series: Materials Science and Engineering. IOP Publishing, 2020, 806(1): 012025.
[4] Afif M, Ayachi R, Said Y, et al. Deep learning based application for indoor scene recognition[J]. Neural Processing Letters, 2020, 51: 2827-2837.
[5] SLai Y, Ma R, Chen Y, et al. A pineapple target detection method in a field environment based on improved YOLOv7 [J]. Applied Sciences, 2023, 13(4): 2691.
[6] Dourado A M B, Pedrino E C. Towards interactive customization of multimodal embedded navigation systems for visually impaired people [J]. International Journal of Human-Computer Studies, 2023, 176: 103046.
[7] Ye M, Yan X, Jiang D, et al. MIFDELN: A multi-sensor information fusion deep ensemble learning network for diagnosing bearing faults in noisy scenarios [J]. Knowledge-Based Systems, 2024, 284: 111294.