Exploration on Intelligent Interaction Design and User Experience of VR Platform Based on Intelligent Perception and Deep Learning

Yadong Zhang[1*], Yujuan Zheng[2], Chen Yin[1], Chunju Shi[1]

{[*]Corresponding author: 153318291@qq.com}
{ 124725859@qq.com, 63089223@qq.com, 363141404@qq.com}

[1]School of Information Engineering, Shandong Vocational University of Foreign Affairs, Weihai, Shandong Province, China

[2]Information Center, Shandong Drug and Food Vocational College, Weihai, Shandong Province, China

Abstract: With the popularization of mobile terminals, the development and application fields of VR (virtual reality) technology are receiving increasing attention. Intelligent interactive systems play an important role in this field, and VR scenarios are a new way of digital information dissemination. This article took user experience as the starting point for research. Firstly, relevant literature at home and abroad was organized and summarized. Secondly, after analyzing the integration of existing mainstream platforms with deep learning (DL) theory, an intelligent interaction design platform based on DL model fusion was designed and developed, and experimental and evaluation results were compared and analyzed based on the algorithm. The test results showed that the overall flexibility of the system's interaction was over 70%; the overall delay rate was within 0.08; the tracking accuracy was above 89%. This indicated that through the optimization of DL algorithms, the intelligent interactive design VR platform can meet user needs.

Keywords: Intelligent Perception, Deep Learning, VR Platform, Intelligent Interaction

1. Introduction

With the rapid development of digital media technology, the VR application market has also expanded. In the information age, with the widespread application of digital technology and networks, humans have entered a virtual environment based on multimedia and characterized by communication. At present, the design of VR platform systems mainly achieves interaction by establishing models. These models utilize animation, 3D image videos, and other methods to transform virtual reality interactions into mutual contrast effects between real environment images and simulated devices, thus forming a complete visual experience. At the same time, by capturing the position and posture of objects in the scene with multiple cameras, a

ICSETPSD 2023, November 17-18, Coimbatore, India

DOI 10.4108/eai.17-11-2023.2342820

three-dimensional, intuitive and realistic perception of surrounding things can be established.

Although China started research on VR platforms based on artificial intelligence relatively late, in recent years, the rapid development in information and communication technology, data processing, and pattern recognition has led to significant improvements in related research. Some scholars have proposed a virtualization interface design method for user interaction. This method used an embedded visual model and multi-dimensional encoding method to fuse the content contained in the video captured from the image, and generated a stream frame representation as a three-dimensional format file for easy storage of spatial relationships. It also utilized DL process training algorithms to classify the original image dataset and display it within a visible range, achieving interaction [1-2]. Some scholars have proposed the concepts of knowledge based learning and heuristic interaction design. The core idea was to utilize a multi user collaborative work environment, collaborative platform, and well visible tasks for training, and to make judgments on problems by obtaining relevant information. Multiple different functional modules were used in virtual scenes to achieve interactive behavior [3-4]. This article adopted intelligent perception and DL technology to design intelligent interaction on the VR platform.

This article first analyzed the characteristics of VR scenes in traditional media environments. Then, an interaction design method based on the combination of intelligent perception and DL was proposed, and relevant research was conducted on this basis. After in-depth reading and sorting out existing literature, it was determined to use a hierarchical structure model to construct a user experience evaluation index system. Next, the modeling, parameter setting, and data training of the virtual roaming system were completed using the programming language toolbox. Finally, by using the MySQL database management system, a VR platform interaction design method based on the combination of intelligent perception and DL was implemented.

2. Discussion on Intelligent Interactive Design of VR Platform

2.1 VR Technology

VR technology is an advanced video compression, digital animation, and image processing and interaction system based on projectors. It utilizes computer vision and image synthesis to achieve, consisting of computers, sensors, and communication devices. It can obtain multi angle image data through a control machine, and process and analyze it to obtain a set of image or video information. Using a digital signal collector, analog signals can be converted into digital quantization values as evaluation parameters for user viewing experience. Based on these data, VR scenarios can be established and attribute variables of the model can be measured [5-6]. This technology is widely used in fields such as 3D reconstruction and stereo photogrammetry. In 3D reconstruction, virtual scenes are created by geometrically modeling objects in space and simulating camera motion. Subsequently, the corresponding coordinate data was constructed based on the model to generate a viewpoint trajectory map. By using a projector, relevant actions and parameter calculations can be completed to form a complete image work. This enables VR technology to be photographed and applied in real environments. VR technology utilizes advanced computer control methods such as digital signal processing and image recognition to achieve the description and analysis of three-dimensional information of objects in space. In the VR environment, users can perform various operations through virtual machines, such as watching videos, browsing web pages, and obtaining real-time dynamic images. When users cannot directly access it, they can divide the content of interest into objects within different visual ranges. During the interaction process using VR technology, sensors collect on-site image and sound data and convert them into electrical signals. Then, it is transmitted to the computer through a serial port for processing and analysis, and parameter values such as viewpoint relative to camera time and screen distance at corresponding positions are obtained. At the same time, cameras are used to capture videos generated from different angles and color regions in real scenes, and computational encoders are used to simulate the true temperature of objects during camera motion and heating [7-8]. Unlike traditional media reporting, VR big data analysis platforms use multimedia formats such as images and animations to present a large amount of news information, video images, and other content for users to watch. This approach presents a more intuitive overview of news events and provides in-depth reporting, building a virtual world based on computer platforms. The digital media data is converted into analog signals and transmitted to the user terminal. The information is presented to the user through software processing, enabling it to be recognized and understood by the human eye. The system provides control instructions and determines interactive operation modes and corresponding execution strategies to achieve the application of VR technology in practical engineering.

2.2 Intelligent Interaction System

By accessing the application software system in the VR platform, data related to interface content, operation methods, etc., can be obtained. At the same time, corresponding solutions can be provided according to different needs. For example, when a user needs to modify their personal information, they can add a new password and current time on the client input page. If a historical record of a certain date is found in the backend database or an abnormal situation occurs in a certain location, the system would pop up a feedback dialog box, prompting the administrator to update the latest date in a timely manner and handle emergencies, in order to better understand the user's feelings and needs for the intelligent interaction system [9-10]. Visual interaction refers to the use of computer, image processing, and other technologies to combine users' senses of hearing, taste, smell, and touch, forming complete sensory information. The intelligent interface mainly consists of the following three parts: user interface, operation channel, and data storage and call module, used to obtain control command signals, execute response commands, and output related functional content. Fig. 1 shows the visual interaction workflow.

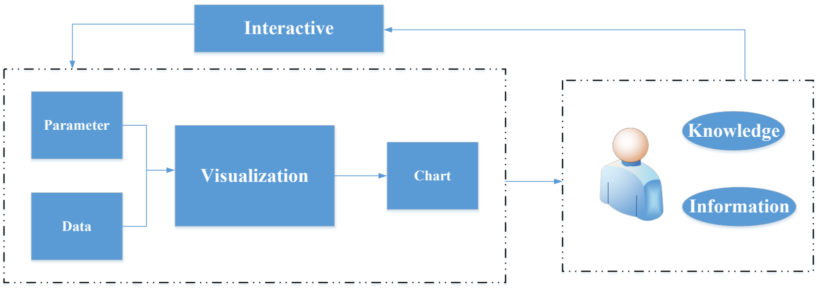

Fig. 1. Visual interaction workflow

The interaction system is the most core and important part of the entire VR platform, connecting users and servers. When specific tasks need to be completed, the processed information is sent to the central control room through the computer, and then corresponding instructions are obtained from the computer room and sent to the client. At the same time, it can also be connected to the backend through network technology, wireless communication, and other methods, and interact in real-time to meet the operational requirements of interactive systems in different application scenarios. From a user perspective, the reason for choosing this type of product is that it can directly express the desired content through images or text descriptions, and can communicate and share knowledge points or keywords with others through video conferencing or audio editing. Therefore, it has become an intelligent application tool with visual interaction function and auditory experience characteristics. It mainly includes server platforms and application software on the client and server sides. Among them, service terminals and application software are implemented through embedded operating systems, while desktop applications are developed using object-oriented structures, ultimately completing relevant functional operations [11-12]. Interactive interfaces are the medium for information exchange between users and application systems, which can transmit user input data to the user interface for processing. The front-end device stores the collected data for later analysis and maintenance. The interface control module is responsible for implementing the functions of various parts within the entire VR platform and controlling the workflow between different modules to complete various tasks.

3. Experimental Process of Intelligent Perception and DL

3.1 Intelligent Perception

Intelligent perception technology uses various sensors, such as GPS (Global Position System), navigation systems, and LiDAR, to influence the behavior of robots. Visual sensors can detect obstacles and targets in the environment and make judgments. The auditory receiving device can respond accordingly when there are obstacles around it. The tactile system controls the robot's actions by measuring its position and state in real-time, in order to improve reaction speed and accuracy. The application of these intelligent perception technologies has improved the perception ability of robots, providing users with a better experience and comfort [13-14]. On the VR platform, intelligent interaction design utilizes existing technology to analyze and process the information obtained during user perception, and transmits it to the terminal through algorithms. Users model the environment, weather, and other information based on their own situation, and then generate corresponding parameters through the model, which is then fed back to the user. The entire process involves multiple sensory involvement, such as auditory neural networks and visual centers. Through the application of VR technology, the quality of information services has been significantly improved. Users can control and interactively operate the 3D scene roaming system in real-time through mobile devices, and can also choose suitable locations for mobile communication roaming according to their needs. Compared to the simple and single functional client interface of traditional media, the application of intelligent perception technology on VR platforms can bring various sensory experiences, multi angle information feedback, and rich and interesting advantages. Applying VR technology in virtual reality space can perceive natural scenes through digitization and simulation, and achieve interaction based on computers. This interactive method can use projection devices to capture image sequence information from different orientations through two cameras, thereby real-time collecting and monitoring parameters such as object movement and environmental temperature [15-16].

3.2 DL Algorithms

This algorithm is based on a heuristic online recommendation system in DL. When task driven users, the project data was first compared and analyzed with the current feature parameters. Then, the weights of the corresponding attributes were calculated based on different feature values. Finally, project reasoning was used to determine whether the project would be the ultimate goal, and the next candidate object was identified as an instance of the test results. This technology uses a specific type of artificial neural network structure to simulate the cognitive process and regularity of the human brain on things in the natural world. By using multi-layer neurons to connect with each other, a DL algorithm is formed, and its advantages are utilized to complete the task of collaborative processing between multidimensional data [17-18]. From a mathematical perspective, the logarithm of probability comes from a linear combination model of predictive variables, and the process of logistic regression is divided into three major steps. The first step is to construct the prediction function h. Logistic regression is essentially a classification algorithm mainly used for binary classification and partial multi classification problems. Therefore, it uses the sigmoid regression function, which is expressed as:

$$G(z) = \frac{1}{1 + \mathrm{e}^{-z}} \quad (1)$$

From this, a prediction function h (x) is constructed, which represents the probability of the result taking 1.

$$h_\theta(x) = g(\theta^T x) = \frac{1}{1 + e^{-\theta^T x}} \quad (2)$$

The actual output vector of the output layer neuron is:

$$a_i = \sigma(z^o) = V\sigma\sigma(Tx + b1) \quad (3)$$

It can handle nonlinear problems by analyzing the relationships between input datasets, hidden parameters, and output results, with good stability and robustness. At the same time, it also has the characteristics of reducing computational complexity to a certain extent and improving computational speed. In the model, the classification results are obtained by training the matrix one by one from left to right. Then A-B tests are conducted on samples of different categories at each location. If the sample has poor linearity, it can be directly subjected to a pull-down slope detection experiment. On the contrary, it is necessary to observe whether there is a correlation and a directional change relationship between them, in order to determine whether they have the same or similar characteristics, and provide evaluation conclusions [19-20]. It is based on a large amount of distributed memory and describes typical knowledge points in the network by classifying and merging various types of deep objects. Under this algorithm, user input features can be sorted. The model mainly consists of three parts. The first part is the small world structure, which is used to target massive sample information. The second part is the large spatial information library, which provides dimensional information from the dataset obtained from each training set to all dimensions to be analyzed, thereby achieving DL. DL algorithm is a parallel multidimensional search method based on data mining. When designing a model, the core idea is to first construct a rule library that contains multiple information vectors such as sample size, input dimension, and output dimension to describe the problem attributes. Secondly, based on the selection of different parameters in the training set, appropriate initial probability values are selected as the iteration starting point.

3.3 Intelligent Perception and Interaction Evaluation Indicators

This system focuses on the scene design of VR platforms, converting user perceived content into formats such as codes, text boxes, or barcodes, and sending it wirelessly to the computer to receive data information. In this process, software and hardware resource management and testing tools play an important role. Users use the most basic, minimum difficulty, best convenience, fastest response speed, and strongest security requirements in a virtual environment. In addition, it can make corresponding scene switching adjustments according to different needs, while conducting effect verification and analysis of good and bad factors. After completing platform modeling, it is necessary to conduct extensive experiments on the model to ensure its normal operation and verify user experience. By collecting and inputting data into the user interface, it is possible to read and retrieve information such as scene variables, perceptual content, and time used in the complete interactive system, and convert this information into machine language for processing. By interacting with users and obtaining data, the collected information can be processed and users can receive feedback in this project. Based on the preliminary test results, the comprehensive performance of the VR platform is evaluated to meet the target requirements, with main indicators including perceived quality and availability. Based on relevant factors, an evaluation system is established and a comparison between the final effect and the evaluation value was determined.

4. Performance Experiment of Intelligent Interactive VR platform

Table 1. Basic data set of the VR platform

Data set	Point cloud scene	Target class	Record
A	5896	5423	6853
B	6358	4893	7543
C	8758	6543	6749
D	7467	6784	7849

In order to design the required core information, Table 1 provides the necessary data. Next is the mobile terminal database, where mobile devices interact with cloud platforms and computers through wireless networks for access control, real-time query, and retrieval functions. The computer sends SMS request instructions to the user through the General Packet Radio Service (GPRS) protocol; After receiving the corresponding software program from the user end, it is returned to the system database for processing, and then data is obtained from the web server. In order to enhance the visibility and real-time performance of the model, as well as facilitate and efficiently collect data, relevant workflow and preparation work have been carried out. In addition, considering the differences in interface functional requirements among different types of users, it is necessary to fully reflect the people-oriented concept in interface design, and combine it with practical application scenarios to achieve a good operating experience.

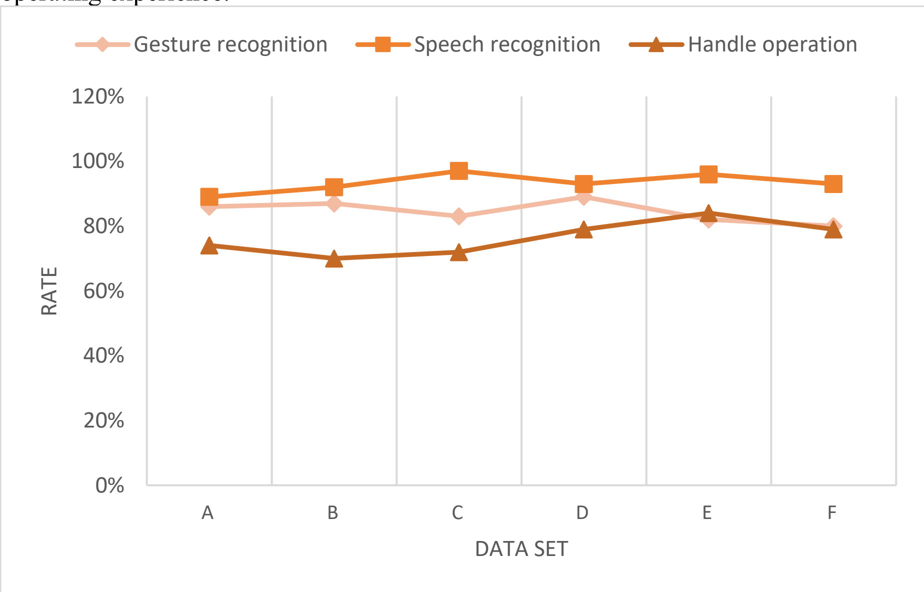

Fig. 2. Flexibility

In order to achieve real-time interactive experience for users, a flexible design principle for interaction systems has been proposed. The intelligence of VR platform is a 3D scene model structure and visualization tool interface constructed based on a virtual reality technology. It needs to meet the requirements of storing, processing, and displaying control of information such as data and images to ensure its ease of use. At the same time, maintaining consistency in user experience throughout the entire process is also crucial. Therefore, the interactive system should be as simple and easy to operate as possible, as well as easy to maintain and upgrade. According to the data in Fig. 2, the flexibility of gesture recognition was over 80%; speech recognition was above 89%; handle operation was above 70%.

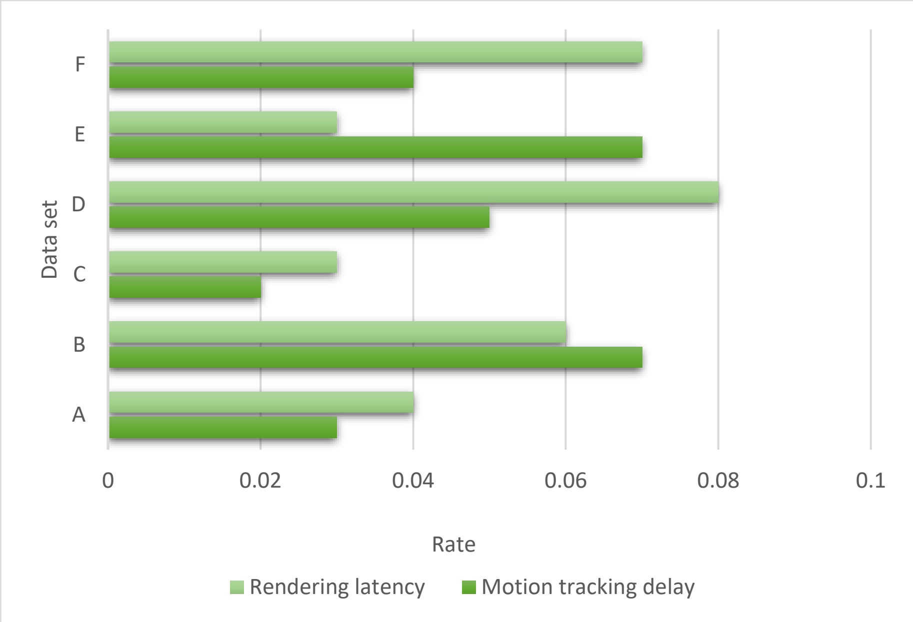

Fig. 3. Delay rate

In this study, a method based on intelligent interaction model algorithm was used to calculate the latency rate of users on the VR platform. By processing and analyzing system data, it is possible to accurately calculate the perceived time that users encounter problems during use and need to solve them in a timely manner. To achieve this goal, the device to be tested was placed in the scene environment to simulate triggering location information. By sending the set response parameters to the mobile terminal, real-time feedback results and delay data were processed and analyzed to identify the problems that users encounter during use. From the delay rate test data in Fig. 3, it can be seen that the motion tracking delay rate was within 0.07; the rendering delay rate was within 0.08. This intelligent interaction model based algorithm provided accurate data, helped better understand users' experience on the VR platform and provided guidance for problem solving.

In virtual reality systems, moving objects are subject to certain limitations in terms of perception time and spatial dimensions due to their actual existence. Therefore, users need to accurately obtain real-time scene information and dynamic changes in the surrounding environment. This article mainly studied, analyzed, and evaluated the process of interactive interface design, data acquisition, and algorithm optimization based on intelligent visual mapping technology, for extracting trajectory information of moving objects with multiple attributes. By calculating multiple observation samples, the optimal parameter values were determined and their accuracy results were evaluated at the corresponding positions. According to the data in Fig. 4, the tracking accuracy of the head was over 90%; the tracking accuracy of the hand was above 89%; the tracking accuracy of body movements was over 90%.

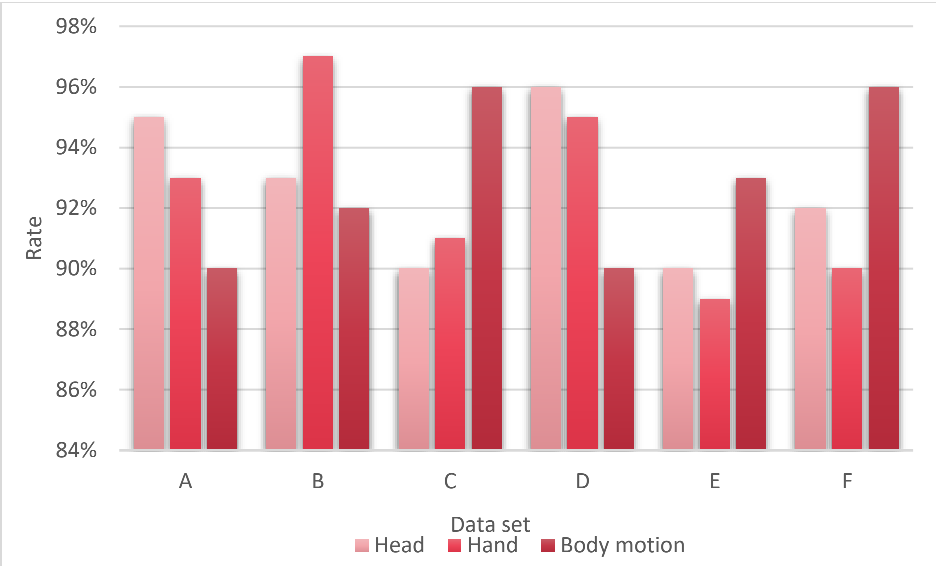

Fig. 4. Motion tracking accuracy

5. Conclusion

With the rapid development of digital media technology, VR platforms based on DL and intelligent interaction design are also steadily updating. This article aimed to study the construction of a 3D spatial perception, feature analysis, and evaluation system based on web pages in a virtual preview environment, and to discuss it from two aspects: multi user participation and collaborative recommendation. Through a review and summary of relevant literature at home and abroad, a method for building models based on artificial intelligence algorithms was proposed, and experimental testing was conducted. Finally, a complete interactive work was generated by combining the preview scheme.

References

[1] Xuan Li, Hengxin Chen, Shengdong He, Xinrun Chen, Shuang Dong, Ping Yan, Bin Fang:Action recognition based on multimode fusion for VR online platform. Virtual

Real. 27(3): 1797-1812 (2023).
[2] Lucas Diniz da Costa, Paulo Victor de Magalhães Rozatto, André Luiz Cunha de Oliveira, Rodrigo Luis de Souza da Silva:VR Tools - A Free Open-source Platform for Virtual and Augmented Reality Applications on the Web. J. Mobile Multimedia 18(1): 27-42 (2022).
[3] Chengyuan Zheng, Jinyu Yin, Fangzhen Wei, Yu Guan, Zongming Guo, Xinggong Zhang:STC: FoV Tracking Enabled High-Quality 16K VR Video Streaming on Mobile Platforms. IEEE Trans. Circuits Syst. Video Technol. 32(4): 2396-2410 (2022).
[4] Tetsunari Inamura, Yoshiaki Mizuchi, Hiroki Yamada:VR platform enabling crowdsourcing of embodied HRI experiments - case study of online robot competition. Adv. Robotics 35(11): 697-703 (2021).
[5] Zeljko Lukac, Ivan Kastelan, Mario Vranjes, Branislav M. Todorovic:AMV ALPHA Learning Platform for Automotive Embedded Software Engineering. IEEE Trans. Learn. Technol. 14(3): 292-298 (2021).
[6] Na Liu, Quanlin Pu, Yan Shi, Shengtai Zhang, Luyi Qiu:Older Adults' Interaction With Intelligent Virtual Assistants: The Role of Information Modality and Feedback. Int. J. Hum. Comput. Interact. 39(5): 1162-1183 (2023).
[7] Shengbo Sun, Hao Wang, Chong Li, Yi Wang, Bing Li:An adaptive fusion method of multi-mode human-computer interaction information in intelligent warehouse. Int. J. Reason. based Intell. Syst. 15(2): 148-155 (2023).
[8] Ping Sun, Rui Shan, Shuoyu Wang:An Intelligent Rehabilitation Robot With Passive and Active Direct Switching Training: Improving Intelligence and Security of Human-Robot Interaction Systems. IEEE Robotics Autom. Mag. 30(1): 72-83 (2023).
[9] Mesfer Alrizq, Shauban Ali Solangi, Abdullah Alghamdi, Muhammad Ali Nizamani, Muhammad Ali Memon, Mohammed Hamdi:An Architecture Supporting Intelligent Mobile Healthcare Using Human-Computer Interaction HCI Principles. Comput. Syst. Sci. Eng. 40(2): 557-569 (2022).
[10] Sara Moussawi, Marios Koufaris, Raquel Benbunan-Fich:The role of user perceptions of intelligence, anthropomorphism, and self-extension on continuance of use of personal intelligent agents. Eur. J. Inf. Syst. 32(3): 601-622 (2023).
[11] Zhichuan Tang, Yidan Hu, Weining Weng, Lekai Zhang, Lingtao Zhang, Jichen Ying:An Intelligent Shadow Play System With Multi-dimensional Interactive Perception. Int. J. Hum. Comput. Interact. 39(6): 1314-1326 (2023).
[12] Yongtao Ma, Dianfei Su, Yuting Li, Chenglong Tian, Weijia Meng:Spatial Perception of Tagged Cargo Using Fused RFID and CV Data in Intelligent Storage. IEEE Internet Things J. 10(2): 1574-1587 (2023).
[13] Siddique Latif, Heriberto Cuayáhuitl, Farrukh Pervez, Fahad Shamshad, Hafiz Shehbaz Ali, Erik Cambria:A survey on deep reinforcement learning for audio-based applications. Artif. Intell. Rev. 56(3): 2193-2240 (2023).
[14] Pingli Ma, Chen Li, Md Mamunur Rahaman, Yudong Yao, Jiawei Zhang, Shuojia Zou, Xin Zhao, Marcin Grzegorzek:A state-of-the-art survey of object detection techniques in microorganism image analysis: from classical methods to deep learning approaches. Artif. Intell. Rev. 56(2): 1627-1698 (2023).
[15] Riza Rae Pineda, Takatomi Kubo, Masaki Shimada, Kazushi Ikeda:Deep MAnTra: deep learning-based multi-animal tracking for Japanese macaques. Artif. Life Robotics 28(1): 127-138 (2023).
[16] Taira Watanabe, Kensuke Tanioka, Satoru Hiwa, Tomoyuki Hiroyasu:Performance comparison of deep learning architectures for surgical instrument image removal in gastrointestinal endoscopic imaging. Artif. Life Robotics 28(2): 307-313 (2023).
[17] Chengxi Yan, Xuemei Tang, Hao Yang, Jun Wang:A deep active learning-based and

crowdsourcing-assisted solution for named entity recognition in Chinese historical corpora. Aslib J. Inf. Manag. 75(3): 455-480 (2023).

[18] Lars Wagner, Sven Kolb, Patrick Leuchtenberger, Lukas Bernhard, Alissa Jell, Dirk Wilhelm:Deep learning assisted intraoperative instrument cleaning station for robotic scrub nurse systems. Autom. 71(7): 572-580 (2023).

[19] Mostafa Saneii, Ali Kazemeini, Sania Esmaeilzadeh Seilabi, Mohammad Miralinaghi, Samuel Labi:A methodology for scheduling within-day roadway work zones using deep neural networks and active learning. Comput. Aided Civ. Infrastructure Eng. 38(9): 1101-1126 (2023).

[20] Jueon Park, Hyojin Sung:XLA-NDP: Efficient Scheduling and Code Generation for Deep Learning Model Training on Near-Data Processing Memory. IEEE Comput. Archit. Lett. 22(1): 61-64 (2023).