Museum achieving future sustainable development: A study on the attraction of historical exhibits in museum nursing direction

Zhiyu Zhang[1], Zhuoran Li[2]

1 School of Humanities, Universiti Sains Malaysia, Penang Malaysia

2 School of Nursing, Sun Yat-sen University, Guangzhou, China

Abstract

This research aims to explore the method of combining digital twins (DTs) with Convolutional Neural Network (CNN) algorithms to analyze the attraction of museum historical and cultural exhibits to achieve intelligent and digital development of museum exhibitions. Firstly, the DTs technology is used to digitally model the museum's historical and cultural exhibits, realizing virtual exhibit display and interaction. Then, the Mini_Xception network is proposed to improve the CNN algorithm, which is combined with the ResNet algorithm to construct a human facial emotion recognition model. Finally, the proposed model is used to accurately predict the attraction of museum historical and cultural DTs exhibits by recognizing people's facial expressions when observing them. Comparative experimental results show that the proposed recognition method can greatly improve accuracy and scalability. Compared with traditional recognition methods, the recognition accuracy can be improved by 5.53%, and 2.71s can reduce the model's data transmission delay. The enhanced scalability of the recognition type can also meet the real-time interaction requirements in a shorter time. This research has important reference value for the digital and intelligent development of museum historical and cultural exhibitions.

1. Introduction

There is current research on creating 'nostalgic environments' of exhibits for audiences in museums and institutional care settings to support visitors' memory and recognition. This research identifies care as a 'curatorial' practice: museums act as 'curators', re-establishing and reorganising the 'meaning' of the visit by retaining the audience's personal biographies and sense of social belonging. As curatorial care in nursing homes preserves not only individual but also collective memories of what it takes to be human and to belong to a society, these institutions should be considered important places in society in relation to the production of meaning, value and cultural heritage. Recently, Augmented Reality (AR) has been used in many museum exhibitions to display virtual objects in videos captured from the real world.

The integration of museum intelligence and care is a relatively new direction, and museums can be uniquely effective, particularly in the areas of cultural sensitivity, mental health and emotional healing. Digital twins (DTs) can map entities or systems from the real world to the digital world and perform modeling, simulation, and analysis [1]. Applying this technology in cultural exhibit attraction assessment can help evaluators understand the characteristics, strengths, and weaknesses of cultural exhibits more intuitively.

In order to give full play to the educational function of museums and better adapt to the rapid development of the industry, relevant policy documents need to be effectively implemented. Development, construction, and resource advantages are transformed into market advantages, enriching spiritual connotations, avoiding single-product content, and overcoming the problem of insufficient creativity. The presentation of core culture and the development of cultural products must pay attention to product development and guide cultural opening through cultural creation [2, 3]. This study analyses the current situation of product development in museums and analyses the existing production system and consumer demand. The results of the study can not only guide the intelligent practice of museums and promote design upgrading, but also enrich the theoretical system of design evaluation for nursing exhibits and overcome the randomness of subjective experience evaluation, so as to nourish the practice and promote the development of intelligent museums combined with medical and nursing exhibits.

The purpose of this study was to investigate the appeal of historical and cultural displays to patient audiences with a humanistic orientation. The approach involves several key steps. Firstly, DTs are employed to digitally model historical and cultural museum exhibits, enabling virtual presentation and interaction. Subsequently, enhancements are made to the Convolutional Neural Network (CNN) algorithm, and a visitor facial emotion recognition model is developed by combining Mini_Xception with ResNet. Lastly, a cultural exhibit attractiveness assessment model is established using facial emotion recognition. The innovation of this research lies in the application of DTs technology for the digital modeling of historical and cultural museum exhibits, enabling virtual display and interaction. This represents a significant stride in digital and intelligent advancement. In addition, this study introduces the Mini-_Xception network to refine the CNN algorithm and merges it with the ResNet algorithm to construct a facial emotion recognition model for spectators (care patients), thus improving the prediction accuracy of spectator attraction to exhibits. The overall organisation of this study is described as described below. Section I outlines the historical background of the museum under humanistic care and the

intelligent functions of this museum. Section II arranges the database according to the characteristics of the museum's cultural exhibits by referring to the relevant literature on CNN algorithms. Section 3 combines CNN and facial emotion recognition strategies to build an evaluation model for the attractiveness of cultural exhibits. Section 4 presents the results of the experimental data obtained through data transfer and experiments. Section 5 concludes the experiment by analysing and organising the experimental data. This study has practical value in improving the attractiveness of historical and cultural exhibits within museums and humanities care.

2. Related work

2.1. CNN algorithm and related research on museum exhibits

In the context of relevant CNN research, Sakai et al. (2021) [4] employed graph CNN to anticipate the pharmacological behavior of chemical structures, utilizing the constructed model for virtual screening and the identification of a novel serotonin transporter inhibitor. The efficacy of this new compound is akin to that of in-vitro- marketed drugs and has demonstrated antidepressant effects in behavioral analyses. Rački et al. (2022) [5] utilized CNN to investigate surface defects in solid oral pharmaceutical dosage forms. The proposed structural framework is evaluated for performance, demonstrating cutting-edge capabilities. The model attains remarkable performance with just 3% of parameter computations, yielding an approximate eight fold enhancement in drug property identification efficiency. Yoon et al. (2022) [6] employed CNN and generative adversarial networks to synthesize and explore colonoscopy imagery through the development of a proficiently trained system using imbalanced polyp data. Generative adversarial networks were harnessed to synthesize high-resolution comprehensive endoscopic images. The findings reveal that the system augmented with synthetic image enhancement exhibits a 17.5% greater sensitivity in image recognition. Benradi et al. (2023) [7] introduced a hybrid approach for facial recognition, melding CNN with feature extraction techniques. Empirical outcomes validate the method's efficacy in facial recognition, significantly enhancing precision and recall. Ahmad et al.(2023) [8] proposed a methodology for recognizing human activities grounded in deep temporal learning. The amalgamation of CNN features with bidirectional gated recurrent units, along with the application offeatureselection strategies, enhances accuracy and recall rates. Experimental results underscore the method's substantial practicality and accuracy in human activity recognition.

Furthermore, certain scholars have conducted relevant studies into museum exhibit attributes. Ahmad et al. (2018) [9] investigated the requisites of museums as perceived by the public, deriving insights from museum scholars and experts to outline the trajectory for developing museum exhibitions in Malaysia aimed at facilitating public learning. The results highlight the unique role of museum exhibits in residents' lifelong learning journey. Ryabinin et al. (2021) [10] harnessed a scientific visualization system to explore cyber- physical museum displays rooted in a system-on-a-chip architecture with a customized user interface. This research introduced an intelligent scientific visualization module capable of interactive engagement with and display of museum exhibits. Shahrizoda (2022) [11] delved into architectural and artistic solutions for museum exhibitions, clarifying architectural matters based on existing scientific and historical documents, and assessing prevailing characteristics of architectural and artistic solutions aligned with museum objectives.

These pertinent studies offer valuable points of reference and thought-provoking insights for the current research. They exemplify the diverse applications of CNN across domains such as drug research, image synthesis and recognition, and facial identification. These investigations underscore CNN's robust modeling and predictive capabilities, indicating its potential significance within the realm of cultural exhibits. They present intriguing concepts for image synthesis and enhancement within the museum exhibit domain. Additionally, in the arena of museum product development, the exploration of the integration of deep learning with other technologies holds promise for elevating the precision and effectiveness of product design.

2.2. An overview of the attractiveness of museum historical and cultural exhibits for humanistic care

Developing cultural tolerance in patient audiences based on the appeal of museum historical and cultural exhibits has been shown to be effective. In one study, an art museum educator instructed learners to observe museum artworks using the Visual Thinking Strategies (VTS) method. Of 407 students, 211 responded to a post-session survey (52%). 80% of respondents agreed or strongly agreed that 'observing the art of the event was an effective means of initiating discussion with the interprofessional team'. Qualitative analysis of learner gains identified themes of open-mindedness, listening to other views and perspectives, collaboration/teamwork, patient-centredness and bias awareness [12]. The review [13] also showed that art visualisation training through a local art museum hosted by an arts and culture educator was better able to improve the teaching of: facilitated observation, diagnostic skills, empathy, team building, communication skills, resilience and cultural sensitivity for medical students. The limitations of these studies are that they are single-institution reports, short-term, involve small numbers of students, and often lack control. An analysis [14] of 31 schools in the medical humanities at US medical schools found that 52% of the institutions surveyed had centres/departments focused on the humanities. 65% offered 10 or more medical humanities-assisted courses per year, but only 29% offered a full immersion experience in the medical humanities. The study also found that the number of students in the medical humanities was not sufficiently high to justify a full immersion experience. Thus museum-based training has proven to be effective in the training of medical students, yet the amount of research and implementation is still lacking and is still in the beginning stages of exploration.

The cultural communication and emotional healing functions of museums are highly effective in caregiving. A study [15] of 20 patients with cognitive disabilities (PWCI) and caregivers participated in a focus group at the Andy Warhol Museum of Art in an art engagement programme that included guided tours and art projects. Research has revealed that patient humour and laughter in museum interventions, (Further study may reveal roles of humor and laughter in adaptation to cognitive decline and holistic interventions for improved quality of life.) can improve patients' ability to adapt to cognitive decline and holistic interventions for improved quality of life. Another study [16] examined the feasibility of art museum trips as an intervention for chronic pain patients. The art museum offered 1-hour docent-led tours to 54 individuals with chronic pain. 57% of participants reported pain relief during the tour, with an average pain relief of 47%. Subjects reported less social disconnection and pain unpleasantness before and after the trip. Respondents spoke of the isolating effects of chronic pain and how negative experiences with the healthcare system often exacerbate this sense of isolation. Participants viewed the art

museum trip as a positive and inclusive experience with potential lasting benefits. Art museum trips are feasible for people with chronic pain, and participants reported positive effects on perceived social disconnection and pain.

Research related to the attractiveness of historical and cultural exhibits in museums has shown different developments in many fields, but there are serious gaps in combining intelligent environments to mitigate the impact of nurse audiences. With the development of information technology, the development and improvement of CNN algorithms have enabled the practical implementation of facial expression recognition technology in real life.The utilisation of DTs has facilitated the digital modelling of cultural displays. By applying attractiveness analysis models and artificial intelligence algorithms centred on historical and cultural exhibits, the attractiveness of historical and cultural exhibits in museums targeting humanistic contexts has been analysed, and with this foundation more directions of exploration and research can be expanded. This endeavour aims to increase the attractiveness of cultural exhibits while expanding the museum's cross-disciplinary reach and visibility to other fields.

3. Model research on the attractiveness of historical and cultural exhibits in the National Museum based on CNN

3.1. Functional analysis of cultural products in the National Museum of China

Viewed from a cultural design standpoint, the National Museum of China encompasses a historical journey spanning 5,000 years of traditional Chinese culture. Consequently, cultural products should meticulously select representative elements from the wealth of cultural resources, aligning with their distinctive historical significance. Through these cultural products, one can propagate informational culture within the cultural heritage, uphold esteemed traditional practices, and fortify the educational role of cultural offerings [17-19]. Firstly, the design of museum exhibition booths can consider isolating certain popular exhibits to disperse visitor traffic effectively. Additionally, in booth design, efforts should be made to encourage visitor engagement in interactive activities, fostering experiences deeply rooted in cultural impact. Regularly rotating less popular exhibits can align with the concept of catering to visitors. Given the subjective and challenging nature of exhibit evaluation, different visitors have varying cultural content preferences, inevitably leading to differences. A broader audience can be attracted by contemplating and crafting cultural art works, cultural attributes, and innovative choices, thus cultivating a positive feedback loop. The cultural products and functional attributes of the museum are delineated in Fig. 1:
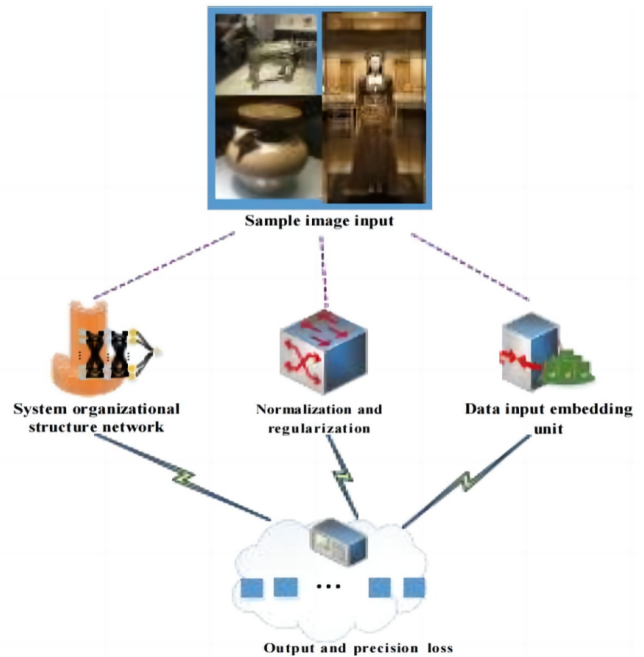

Fig. 1 The structure of the cultural product features of the National Museum

The architectural framework of China National Museum's cultural product characteristics is illustrated in Figure 1. This framework comprises the input layer, data input embedding layer, standardization and normalization layer, output layer, and accuracy loss layer. The input layer receives data from cultural exhibit images as input. The data input embedding layer pre-processes and extracts features from input images, transforming them into a format the model can handle. The standardization and normalization layer processes the data for standardization and normalization, ensuring the model's universality across different datasets. The output layer is the model's final layer, used to predict the features of cultural exhibits. The accuracy loss layer calculates the error between the model's predicted and actual results, updating model parameters through back propagation. This process aims to minimize error loss and enhance model accuracy. Through the collaborative operation of these layers, the essential structural features of cultural products can be effectively extracted and abstracted, achieving automatic recognition, classification, and recommendation capabilities for cultural exhibits.

3.2. Analysis of DTs applied to digital modeling of museum historical and cultural exhibits

DTs can provide more realistic simulations of exhibition venues and exhibits, enabling visitors to understand better the context and meaning of exhibits [20,21]. This research applies DTs to the digital modeling of museum historical and cultural exhibits, including key element entity modeling, DTs virtual modeling, and virtual-real mapping association modeling.

In solid modeling, the elements involved in the activities include exhibit characteristics, functions, performance, etc. Considering the above factors, this research uses a formal modeling language to model the key elements of the exhibit digitization process. The model definition is shown in Eq. (1):

$$PS = PE \triangleleft PP \triangleleft PW \qquad (1)$$

In Eq. (1), PS refers to the physical space of key elements in the

digitization process of exhibits. PE refers to the collection of characteristics of historical and cultural exhibits in physical space. ◁ refers to the collection of functions of historical and cultural exhibits in physical space. PW refers to the collection of historical and cultural exhibits in physical space. refers to the natural connection between PE, PP, and PW, indicating the natural interaction between the three. PE, PP, and PW are all dynamic collections. Collection elements and their status can be continuously updated with the dynamic operation of the digitization process of exhibits.

DTs volume modeling needs high modularity, good scalability, and dynamic adaptability, which can be completed in information space using the parametric modeling method. Virtual models of physical entities are established in Tecnomatrix, Demo3D, Visual Components, and other software [22,23]. In addition to describing the geometric information and topological relationship of the automated production line, the virtual model also contains the complete dynamic engineering information description of each physical object [24,25]. The multiple- dimensional attributes of the model are parametrically defined to realize the real-time mapping of the digital modeling process of exhibits. The specific definition of DTs volume modeling is shown in Eq. (2):

$$CS = DE \triangleleft DP \triangleleft DW \tag{2}$$

In Eq. (2), CS refers to the information space of the key elements of the digitalization process of exhibits. DE refers to collecting characteristics of historical and cultural exhibits in the information space. DP refers to the collection of functions of historical and cultural exhibits in the information space. ◁ refers to collecting historical and cultural exhibits in the information space. refers to the natural connection between DE, DP, and DW, indicating the natural interaction between the three. DE, DP, and DW are all dynamic collections. The collection elements and their states in the information space are updated synchronously with the dynamic operation of the digitalization process of the exhibits in the physical space.

Finally, the virtual-real mapping association is modeled. The virtual-real mapping relationship between the two is further established based on establishing the physical space entity model PS and the information space twin model CS. The formal modeling language is used to model its virtual-real mapping relationship. The model definition is shown in Eq. (3):

$$PS * \overset{1:1}{\Leftrightarrow} CS \tag{3}$$

In Eq. (3), $\overset{1:1}{\Leftrightarrow}$ refers to the real two-way mapping between the physical space entity model and the information space twin model. ◁ refers to the natural connections between different models. Therefore, PE, DE, PP, DP, PW, and DW should maintain asynchronous two-way real

3.3. mapping. Design and research of facial emotion recognition strategy based on CNN

Emotion refers to the subjective emotional experience and internal psychological state that humans have in response to external stimuli. It encompasses various complex experiences and reactions such as happiness, sadness, surprise, anger, fear, and more. Facial expressions are the outward manifestation of emotions, reflecting the emotional states of individuals through muscle movements and neural transmission. Facial expressions can be automatically recognized and classified using computer vision and image processing techniques, helping people gain a more accurate understanding of visitors' emotional states and

responses in front of historical and cultural exhibits in museums. Facial recognition technology can identify facial expressions automatically by analyzing the changes and combinations of facial features, such as the eyes, mouth, eyebrows, and more. For example, when visitors appreciate an exhibit in a museum, they may display different emotions and attitudes, such as liking, disliking, or surprise. By utilizing facial recognition technology, people can monitor visitors'emotional changes and responses in real-time and connect this information with the knowledge and stories related to the displayed historical and cultural artifacts. This enhances visitors' experience and cultural awareness. In conclusion, facial expressions are the outward manifestation of emotions, and facial recognition technology can infer visitors'emotional states and responses by analyzing these expressions. By employing these technologies, people can better understand visitors' emotional experiences and responses in front of historical and cultural exhibits in museums, providing a more scientific and effective reference for exhibition planning and cultural heritage preservation in museums.

Since the input data source of tourist facial expressions is image sequence information collected by surveillance cameras, the model must maintain high recognition accuracy while maintaining high recognition speed [26]. The mini_xception model boasts a lightweight design, enabling it to maintain relatively high accuracy while reducing the network's parameter count and computational load. This lightweight structure makes the mini_xception network well-suited for integration into practical applications, particularly in resource- constrained scenarios [27,28]. Therefore, mini_xception is chosen as the basic network structure. Firstly, layers are added in batches after each convolution to speed up training and improve the model's generalization ability. A regularized dropout rate is used in the model to randomly eliminate some nodes in the network with a given probability. While the network structure is simplified, over-adjustment of the network is avoided. The classification ability of the network is improved. The principle is that in forwarding propagation, some neurons'activation values are randomly stopped in a certain proportion. Therefore, the model is less dependent on some local features. The actual visitors ofthe museum are different, and the lighting intensity of the booth is also different. The main pre-processing methods adopted include grayscale, normalization, and histogram processing. Based on CNN, the structural framework of the designed facial emotion recognition model is shown in Fig. 2:
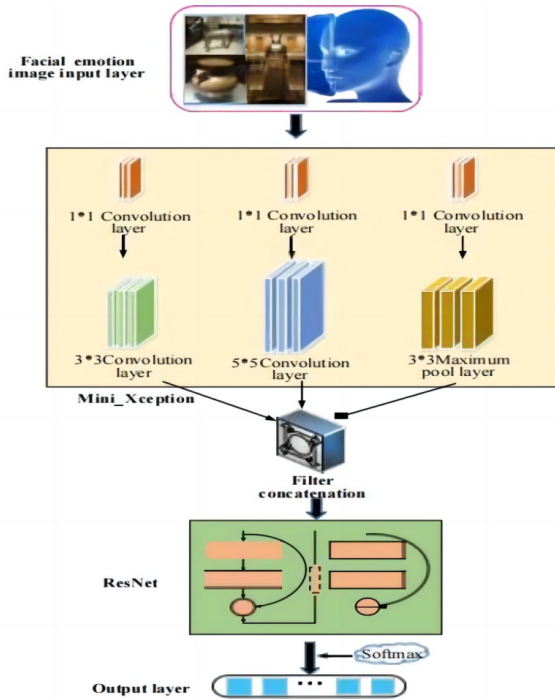
Fig. 2 Framework of facial emotion recognition model structure based on Mini_Xception and ResNet

Assuming there is a sequence offacial images of a group oftourists encompassing diverse emotional expressions as they stand in front of museum exhibits. The objective is to employ a model for recognizing and classifying these emotions while maintaining a balance between speed and accuracy. The mini_xception architecture has been chosen for the network structure. Following each convolutional layer, batch normalization layers have been added to expedite training and enhance model generalization. Additionally, dropout regularization has been employed within the model to deactivate some nodes, mitigating the risk of overfitting randomly. Given a sequence of facial images of different tourists, during the forward propagation phase, a portion of neuron activations is randomly halted to a certain proportion. This ensures the model does not excessively rely on specific localized features. Considering the varying illumination conditions faced by real tourists in the museum, pre-processing is essential for input facial images. Operations such as cropping, scaling, and grayscale conversion are conducted to optimize the image input for network processing. Subsequently, the mini_xception algorithm is applied to extract features from the facial images. This algorithm accelerates recognition speed while reducing the number of parameters and computations. The mini_xception algorithm transforms tourists' facial images into feature-rich representations encompassing spatial and semantic information. Following this, the ResNet algorithm is employed to fuse features extracted by mini_xception. ResNet, being aresidual network, leverages residual connections to reduce training time and enhance accuracy. This algorithm combines the spatial and semantic features derived from mini_xception to elevate emotion classification's accuracy and generalization capabilities. Ultimately, a Softmax classifier is utilized to categorize the fused features, thereby predicting the emotional category of the facial image. This classifier maps the feature vector into a probability space, yielding the predicted probabilities for each category. Suppose the model has been trained. In that case, a facial image of a tourist can be predicted as belonging to the "liking" emotional category. From capturing

facial images to feature extraction and emotional classification, this algorithm aids in the automated recognition of tourists' emotional expressions before historical and cultural exhibits in museums. Consequently, it offers a more scientific and effective basis for museum exhibit planning and cultural heritage preservation.

In this model, W and I represent one network layer's weight tensor and input tensor, respectively. Among them, c represents the channel and $w < w_{in}, h < h_{in}$ represents the width and height, respectively. The network approximates W through a binary tensor B e$\{+1, -1\}^{c \times w \times h}$ cxwxh and a scaling factor C $eR+$, as shown in Eq. (4):

$$W \sim CB \tag{4}$$

In Eq. (4), W refers to the real weight.

In order to obtain the best approximation, it is assumed that vectors w and b represent weights and binarized weights, respectively, as shown in Eq. (5):

$$w, b e R^n, n = cwh \tag{5}$$

The determination of the optimal scale factor and binarization parameters is shown in Eq. (6):

$$w \sim Cb \tag{6}$$

The two sides of the above formula are subtracted, and the optimization function is obtained, as shown in Eqs. (7) and (8):

$$J(\mathbb{C}, b) = \min_{\mathbb{C},b} w \mid \square Cb \tag{7}$$

$$C*, b* = \underset{C,b}{arg min} J(C, b) \tag{8}$$

In Eq. (8), C*, b * represents the scale factor and the weight optimization value after binarization, respectively. Eq. (7) is expanded to obtain Eq. (9):

$$J(\mathbb{C}, b) = \mathbb{C}^2 b^T b - 2\mathbb{C} w^T b + w^T w \tag{9}$$

In Eq. (9), be$\{+1, \square 1\}$ n , b Tb = n isa constant. w refers to a known quantity, so w T w is also a constant. Let q = w Tw, Eq. (9) can be simplified to Eq. (10):

$$J(\mathbb{C}, b) = \mathbb{C}^2 n - 2\mathbb{C} w^T b + q \tag{10}$$

C refers to a positive value. The optimization problem of b is transformed into Eq. (11):

$$b^* = *\underset{b}{arg\ min} \{w^T b\} \quad s.t. \quad b e\{+1, -1\}^n \tag{11}$$

According to the characteristics of the value of b, Eq. (12) is calculated:

$$b* = Sign(w) \tag{12}$$

In order to obtain the optimal value of C* , J takes the partial derivative of C and makes it zero. The optimization result of C is obtained, as shown in Eq. (13):

$$C* = \frac{w^T b^*}{n} \tag{13}$$

The value of b* is plugged into Eq. (13), and Eq. (14) is obtained:

$$C_* = \frac{w^T Sign(w)}{n} = \frac{\sum_i |w_i|}{n} = \frac{1}{n} \| w_{l_1} \| \tag{14}$$

In order to further reduce the quantization error, the network's weight and activation quantization process is introduced into a scaling factor. Quantization weights Kb (w) and activations Kb (C) , including scale factors, are obtained, as shown in Eqs. (15) and (16):

$$K_b(w) = Cb \tag{15}$$

$$K_b(C) = \beta h \tag{16}$$

In Eq. (16), b, h are the results of the weight vector w and the activation vector C after the Sign function. C, β is the corresponding scaling factor. The definition of the convolution operation of the network forward process is shown in Eq. (17):

$$z = \sigma(K_b(w) t K_b(\mathbb{C})) = \sigma(\mathbb{C} B (b - h)) \tag{17}$$

In Eq. (17), z refers to the output and $\sigma(.)$ refers to the activation function. $h \in \{+1, -1\}^n$ refers to the activation value after binarization. β is the scaling factor for the activation. t is the convolution operation. ⊙ points to the inner product of the quantity.

The expression of the approximate point product between C and w is shown in Eq. (18):

$$C^T w \sim \beta h C b \tag{18}$$

The optimization problem to be solved is shown in Eq. (19):

$$C^*, b^*, B^*, h^* = *\underset{C, b, B, h}{\arg \min} \| C \odot w \Box C b B h \|^2 \tag{19}$$

The weight vectors are optimized, and the emotions of faces exhibited by DTs are analyzed precisely.

In order to solve the optimization problem defined by Eq. (19) using swarm intelligence optimization algorithms, the steps are as follows:

Step 1: Randomly generate a set of particles, where each particle represents a possible solution in the solution space. Each particle has its own position and velocity, initialized with appropriate values.

Step 2: Calculate the fitness value of each particle based on the objective function defined in Eq. (19). The fitness value indicates the quality of the solution corresponding to the particle.

Step 3: For each particle, update its individual best position based on its current fitness value and historical best fitness value.

Step 4: Update the global best position of the swarm based on the fitness values of all particles. Record the optimal solution and its fitness value.

Step 5: Update the velocity and position of each particle based on its current velocity, individual best position, and global best position.

Step 6: Check if the termination condition is met, such as reaching the maximum number of iterations or the fitness value reaching a predetermined threshold. If the condition is met, proceed to step 8; otherwise, continue with steps 2 to 6.

Step 7: Repeat steps 3 to 6 until the termination condition is met.

Step 8: The optimal solution recorded during the iteration process is the optimization result of the problem.

3.4. Evaluation model of the attractiveness of cultural exhibits based on face recognition algorithm

The data recognized by the face recognition algorithm is mainly used to associate the subsequent facial expression information with character recognition. The residence time of each visitor in the current state can also be calculated through identity recognition information. Since the representation of features in consecutive images is difficult to change in the actual computational process, continuous recognition inevitably leads to many repeated expression results. The facial expression network information is extracted every 15 frames. Facial expression classification theory and created datasets are used as the basis. The client's facial expressions are divided into six categories. Simple results from facial expression recognition may not accurately reflect visitors' satisfaction with an exhibit. Therefore, information about the expressions of tourists is provided in the index. The attractiveness evaluation model of cultural exhibits is analyzed. The evaluation process of the attractiveness of cultural products is shown in Fig. 3:
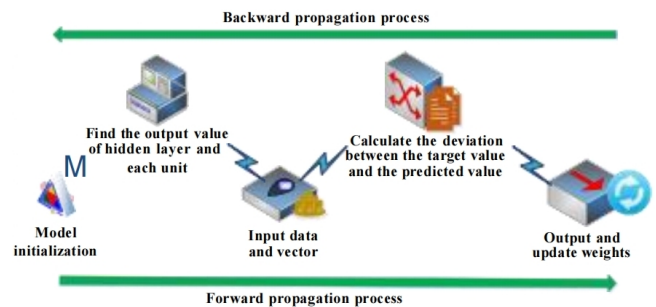

Fig. 3 The evaluation process of the attraction of museum historical and cultural products

In Fig. 3, the evaluation process of the attractiveness of historical and cultural artifacts in museums can be divided into five steps. Step 1: Model Initialization. An appropriate model needs to be selected, and its parameters and network structure are initialized. Commonly used models include CNN and Recurrent Neural Network (RNN). Step 2: Computing Hidden Layer and Unit Outputs. The selected model is used to compute the hidden layer and unit outputs by inputting the relevant data of historical and cultural artifacts in museums. Step 3: Data Input and Vectorization. The relevant data of historical and cultural artifacts in museums is transformed into vector form for computation and processing. Step 4: Calculating Deviation between Target and Prediction. A comparison is made between the predicted results and the actual results to calculate the deviation or error between them. Step 5: Output and Weight Updating. The model parameters are updated through backpropagation based on the magnitude of the error. This aims to minimize the error loss and improve the accuracy of the model. This process can be performed in both forward and backward directions, continuously optimizing the model's performance by updating weights and biases using the backpropagation algorithm. In summary, through the iterative process of these five steps, the accuracy and precision of the

attractiveness evaluation model for historical and cultural artifacts in museums can be gradually improved. This provides a more scientific and effective basis for museum exhibition planning and cultural heritage preservation.

3.5. Data transmission and experimental research

The experiment is conducted in a museum located in City B. The duration of the experiment is one month. Firstly, a random sample of 500 visitors who visited the museum was selected as the subjects. Then, cameras are installed in various exhibition halls to collect facial expression data of the visitors using facial recognition technology in front of different exhibits. The data is acquired by installing cameras in the museum's ten booths. When the file is saved, the identification information ofthe exposure point in the filename is used to distinguish the video data of each exposure point. Ifthe traffic is high, it is limited to the amount of data to be considered, and 10 minutes of exposure input data plus traffic is selected for each program. The CNN algorithm detects faces in each frame of video data. Then, a 160*160-pixel face image is cut out, and the face edge is output at a specific position ofthe frame image. After the visitor obtains the collected face image, each point is input with an identification code. From the data in the MySQL database, quantitative scoring information, human identification information, residence time information, and exposure point information of the program representation labels are imported. Front- end code is used to connect to the back-end. Then, JavaScript is used to access the database data. Finally, an object library that displays attractiveness metrics is used to display the data.

In addition, in order to further analyze the attractiveness of the museum's historical and cultural exhibits, the proposed model algorithm is compared with the performance of Mini_Xception, ResNet, and Benradiet al. (2023). The recognition accuracy, scalability rate, data transmission delay of the system, and the recognition speed of tourists' emotions are analyzed using the human-computer interaction (HCI) system. Additionally, the higher the attraction ofthe exhibits, the longer the visitors stay in the exhibition area. The system analyzes the facial expression scores of tourists in front of different numbered exhibits and the visitor's stay time. It compares the attractiveness of different numbered exhibits according to the facial expression score and the visitor's stay time.

4. Results and Discussion

4.1. Comparison of Recognition Accuracy and Expansion Rate of DTs System

In order to analyze the performance of each algorithm, different algorithms are used in the DTs system to analyze the accuracy of tourist emotion recognition and the system's scalability, as shown in Fig. 4.
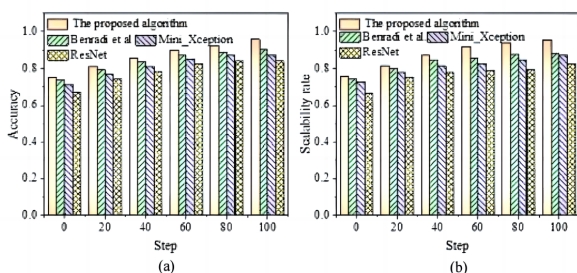

Fig. 4. Recognition accuracy and scalability curves of different algorithms in museum exhibit DTs system (a. recognition accuracy curve; b.scalability rate change curve)

In Fig. 4, as the model step size parameter increases, the accuracy of different algorithms in the DTs system for emotion recognition will slowly increase. When the model step size parameter is 0, the proposed algorithm's recognition accuracy is 75.60%, while the recognition accuracy of other models is less than 73.67%. When the model step size parameter is increased to 100, the recognition accuracy of the proposed model algorithm reaches 95.92%, while other model algorithms do not exceed 90.39%. The recognition accuracy of each model algorithm is sorted from large to small: proposed model algorithm> Benradi et al. (2023)> Mini_Xception> ResNet. In addition, when the model step size parameter is 80, the proposed algorithm improves to 95.75%, which is better than other model algorithms. The results of comparative experiments show that the proposed recognition method can greatly improve the recognition accuracy and extension rate. The recognition accuracy can be increased by 5.53% compared with existing methods. When visitors identify DTs exhibits,the model can achieve more accurate emotion recognition.

4.2. Comparison of results between data transmission delay and tourist emotion recognition speed

The system uses different algorithms to compare the data transmission delay time and the speed of tourist emotion recognition, as shown in Fig. 5.
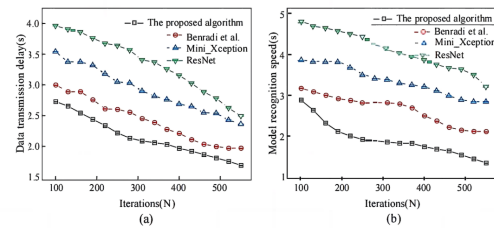

Fig. 5. The data transmission delay time and tourist emotion recognition speed change curve of different algorithms in the system (a. data transmission delay time; b. tourist emotion recognition speed change curve)

In Fig. 5, in the system, as the number ofmodel iterations increases, the data transmission delay time of each model algorithm shows a downward trend. When the model iterates 100 times, the data transmission delay of the traditional algorithm ResNet is 3.96s. The data transmission delay of the improved CNN model is only 2.71s. After 550 model iterations, the data transmission delay of the traditional algorithm ResNet drops to 2.49s. The data transmission delay of the optimized CNN algorithm can be reduced to 1.67s. In addition, the emotion recognition speed of the traditional algorithm ResNet is 3.22s after 550 iterations, and the emotion recognition speed ofthe improved CNN algorithm is only 1.33s. The data analysis results show that the data transmission delay ofthe proposed model algorithm can be reduced to 2.71s. The algorithm is superior to the traditional algorithm regarding emotion recognition speed.

In addition, the resource consumption of the four models is compared, as shown in Fig. 6.
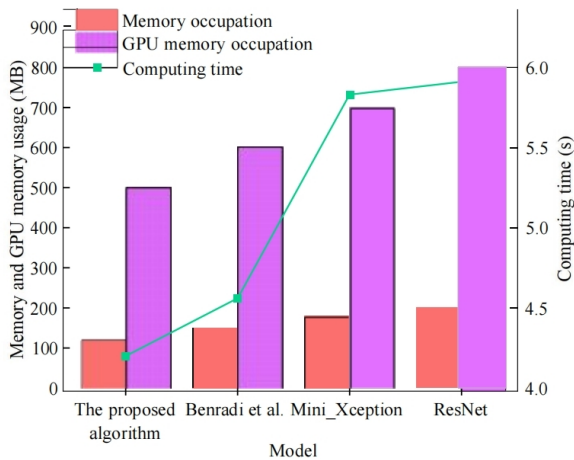
Fig. 6. Comparison of Computational Resource Consumption for Different Algorithms

In Fig. 6, the proposed algorithm's computational time, memory usage, and GPU memory consumption are 4.2s, 120MB, and 500MB, respectively, the lowest values among the four algorithms. This data indicates that the proposed algorithm has advantages over other models in terms of computational time, memory usage, and GPU memory consumption. The result means that the proposed algorithm can efficiently utilize computational resources and has lower resource consumption in the task of recognizing the emotions of museum visitors.

Furthermore, the research also compared the results of different algorithms in terms of data transmission latency and the speed of visitor emotion recognition. The findings revealed that with an increase in the number of model iterations, the data transmission latency for each algorithm exhibited a decreasing trend. For instance, after 100 rounds of model iteration, the data transmission latency for the conventional ResNet algorithm was 3.96 seconds, whereas the improved CNN model demonstrated a data transmission latency of only 2.71 seconds. Upon further augmentation to 550 rounds of model iteration, the data transmission latency of the conventional ResNet algorithm decreased to 2.49 seconds, while the optimized CNN algorithm achieved a reduced data transmission latency of 1.67 seconds. Similarly, the emotion recognition speed of the conventional ResNet algorithm was 3.22 seconds after 550 iterations, whereas the improved CNN algorithm exhibited a recognition speed of merely 1.33 seconds. Based on the results of data analysis, the proposed model algorithm is capable of reducing data transmission latency to 2.71 seconds and outperforms the traditional algorithm in terms of emotion recognition speed.

4.3. Analysis of facial expression scores and visitor stay time results for exhibits with different numbers

The facial expression scores of tourists in front of different numbered exhibits and the data on tourists' staying time in the exhibition area are analyzed, as shown in Fig. 7.
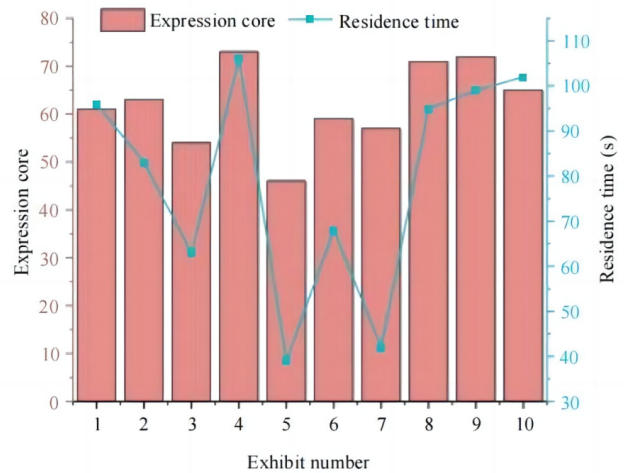

Fig. 7. The average score of facial expressions of museum visitors to different exhibits and the average stay time change

In Fig. 7, the horizontal axis represents the museum's numbered historical and cultural exhibits. The left vertical axis corresponds to the average facial expression scores of 500 visitors in front of various exhibits, while the right vertical axis represents the visitors' average duration of stay and admiration time in front of these exhibits. There is a certain variation in both the average facial expression scores and the average duration of stay for visitors in front of different museum exhibits. Exhibit 4 displays a higher average facial expression score (73) and a relatively longer average stay time (106 seconds). In contrast, exhibit 5 exhibits a lower average facial expression score (46) along with a relatively shorter average stay time (39 seconds). This observation implies a potential correlation between visitors' facial expressions and their duration of stay when encountering different exhibits.

Within this experiment, emotional reactions are the feelings experienced by visitors while observing the exhibits. These emotional reactions may encompass excitement, pleasure, curiosity, and surprise. Such emotional responses could directly impact visitors' preferences for the exhibits. For instance, if a visitor experiences strong excitement and pleasure when observing a particular exhibit, they are likely to favor it. Facial expressions serve as physiological manifestations of emotional experiences conveyed through facial muscle movements like smiling or frowning. Different facial expressions could be associated with distinct emotions and moods, which, in turn, may affect visitors' preferences for the exhibits. For instance, a visitor displaying a noticeable smile while observing an exhibit may indicate their pleasure and satisfaction with it. Positive emotional experiences associated with an exhibit may lead to a higher degree of liking. Comparing visitors' facial expressions in front of different exhibits can help determine which exhibits align better with their interests and preferences. Moreover, visitors might exhibit positive facial expressions such as smiling or widening eyes when they come across exhibits they like. These positive expressions could signify their interest or satisfaction with the exhibit, prompting them to spend more time better appreciating it. Conversely, negative facial expressions like furrowing brows or frowns might indicate dissatisfaction, disinterest, or confusion, leading to reduced interest and shorter stays. Therefore, the experiment compares visitors' average facial expression scores and their average duration of stay. The result shows a significance level (P) of 0.0138, which is less than the conventional threshold of 0.05, indicating a statistically significant positive correlation between visitors' facial expressions and their duration of stay. In this

experiment, the research explored the correlation between facial expression scores and dwell time among visitors by recognizing and analyzing visitors' facial expressions. Facial expression scores were evaluated based on facial muscle movements to assess visitors' emotional experiences, while dwell time referred to the time visitors spent viewing exhibits. The research results demonstrated a statistically significant positive correlation between facial expression scores and dwell time. This implies that when visitors have higher facial expression scores, they tend to spend a longer time in front of exhibits. A higher facial expression score might indicate a positive emotional experience, such as excitement, curiosity, or satisfaction, prompting them to develop a greater fondness for the exhibit and be willing to invest more time in appreciation. Furthermore, the value of the correlation coefficient provides a deeper understanding. Based on the magnitude and sign of the correlation coefficient, the degree and direction ofthe association between facial expression scores and dwell time can be determined. As an illustrative instance ofthis research, the obtained results revealed a correlation coefficient of 0.0138, falling below the established conventional threshold of 0.05. The data indicate a significant positive correlation between the two, implying that as facial expression scores increase, visitors' dwell time also correspondingly increases.

Therefore, the findings of this research reveal a close relationship between visitors' facial expression scores and their dwell time within the exhibition hall. This discovery holds important implications for exhibition design and improvement, aiding in enhancing the attractiveness of exhibits and visitor experiences, thereby elevating visitor satisfaction and the overall success ofthe exhibition.

4.4. Discussion

In summary, the proposed model exhibits low latency in data transmission and improves the accuracy of emotion recognition. Furthermore, a comparison was made with other studies. Razzaq et al.(2022) proposed DeepClass-Rooms, a DTs framework for attendance and course content monitoring in public sector schools in Punjab, Pakistan. It employed high-end computing devices with readers and fog layers for attendance monitoring and content matching. CNN was utilized for on-campus and online courses to enhance the educational level [29]. Sun et al. (2021) introduced a novel technique that formalizes personality as a DTs model by observing users' posting content and liking behavior. A multi-task learning deep neural network (DNN) model was employed to predict users' personalities based on two types of data representations. Experimental results demonstrated that combining these two types of data could improve the accuracy of personality prediction [30]. Lin and Xiong (2022) proposed a framework for performing controllable facial editing in video reconstruction. By retraining the generator of a generative adversarial network, a novel personalized generative adversarial network inversion was proposed for real face embeddings cropped from videos, preserving the identity details of real faces. The results indicate that this method achieves notable identity preservation and semantic disentanglement in controllable facial editing, surpassing recent state- of-the-art methods [31]. In conclusion, an increasing body of research suggests that combining DTs technology with DNN can enhance the accuracy and efficiency of facial expression recognition.

5. Conclusion

With the rapid advancement of science and technology, AI and neural network algorithms have pervaded diverse domains across society. The integration of information technology has profoundly impacted daily life through the meticulous analysis of extensive datasets. Within the context of disseminating historical and cultural narratives, conventional museums often rely on staff members' prolonged visual observations and intuitive assessments to gauge the appeal of exhibits to visitors. This qualitative approach, while competent, poses challenges in terms of quantifiability. This research leverages DTs to digitally model historical and cultural artifacts digitally, enhancing the CNN algorithm to construct a facial emotion recognition model. By coupling this with visitors'dwell time within the exhibit areas, the degree of attraction exerted by each booth upon tourists is assessed and quantified. The functional analysis of historical and cultural artifacts at the museum provides insight. However, certain limitations persist. Foremost among these is the need for enhancement in the accuracy of the facial expression recognition algorithm. Subsequent research endeavors will seek to realize real-time detection of changes in visitors' facial expressions and dwell time through the refinement of the face detection algorithm, thereby addressing the appeal of cultural exhibits to tourists via intelligent HCI approaches.

The combination of museums and medical care is a relatively new and noteworthy direction, especially in terms of cultural sensitivity, mental health and emotional healing, and museums can have a unique effect. The current research is still in the stage of the preliminary attempt to combine history and culture, and the way and effect of the combination of museum and medical treatment are still being explored.

References

[1] Niccolucci F, Felicetti A, Hermon S. Populating the Data Space for Cultural Heritage with Heritage Digital Twins[J]. Data, 2022, 7(8),pp. 105.

[2] Götz F M, Ebert T, Gosling S D, et al. Local housing market dynamics predict rapid shifts in cultural openness: A 9-year study across 199 cities[J]. American Psychologist, 2021, 76(6),pp. 947.

[3] Bebenroth R, Goehlich R A. Necessity to integrate operational business during M&A: the effect of employees' vision and cultural openness[J]. SN Business & Economics, 2021, 1(8), pp. 1-17.

[4] Sakai M, Nagayasu K, Shibui N, et al. Prediction of pharmacological activities from chemical structures with graph convolutional neural networks[J]. Scientific reports, 2021, 11(1), pp. 1-14.

[5] RačkiD, Tomaževič D, Skočaj D. Detection ofsurface defects on pharmaceutical solid oral dosage forms with convolutional neural networks[J]. Neural Computing and Applications, 2022, 34(1), pp. 631-650.

[6] Yoon D, Kong H J, Kim B S, et al. Colonoscopic image synthesis with generative adversarial network for enhanced detection of sessile serrated lesions using convolutional neural network[J]. Scientific reports, 2022, 12(1),pp. 1-12.

[7] Benradi H, Chater A, Lasfar A. A hybrid approach for face recognition using a convolutional neural network combined with feature extraction techniques[J]. IAES International Journal of Artificial Intelligence, 2023, 12(2),pp. 627.

[8] Ahmad T, Wu J, Alwageed H S, et al. Human Activity Recognition Based on Deep-Temporal Learning Using Convolution Neural Networks Features and Bidirectional Gated Recurrent Unit With Features Selection[J]. IEEE Access, 2023, 11, pp. 33148-33159.

[9] Ahmad S, Abbas M Y, Yusof W Z M, et al. Creating Museum

Exhibition: What The Public Want[J]. Asian Journal Of Behavioural Studies (Ajbes), 2018, 3(11),pp. 27-36.

[10] Ryabinin K V, Kolesnik M A. Automated creation of cyber-physical museum exhibits using a scientific visualization system on a chip[J]. Programming and Computer Software, 2021, 47(3), pp. 161-166.

[11] Shahrizoda B. Architectural and Artistic Solution of Museum Exhibition[J]. European Multidisciplinary Journal of Modern Science, 2022, 5, pp. 149-151.

[12] Paul A, Mercado N, Block L, DeVoe B, Richner N, Goldberg GR. Visual thinking strategies for interprofessional education and promoting collaborative competencies[J]. Clin Teach. 2023;20(5):e13644.

[13] Mukunda N, Moghbeli N, Rizzo A, Niepold S, Bassett B, DeLisser HM. Visual art instruction in medical education: a narrative review[J]. Med Educ Online. 2019;24(1):1558657.

[14] Anil J, Cunningham P, Dine CJ, Swain A, DeLisser HM. The medical humanities at United States medical schools: a mixed method analysis of publicly assessable information on 31 schools[J]. BMC Med Educ. 2023;23(1):620.

[15] Liptak A, Tate J, Flatt J, Oakley MA, Lingler J. Humor and laughter in persons with cognitive impairment and their caregivers[J]. J Holist Nurs. 2014;32(1):25-34.

[16] Koebner IJ, Fishman SM, Paterniti D, et al. The Art of Analgesia: A Pilot Study of Art Museum Tours to Decrease Pain and Social Disconnection Among Individuals with Chronic Pain[J]. Pain Med. 2019;20(4):681-691.

[17] Xianying D. Analysis on the Educational Function of Campus Culture in Higher Vocational Colleges[J]. The Theory and Practice of Innovation and Enntrepreneurship, 2021, 4(23), pp. 77.

[18] Chikaeva K S, Gorbunova NV, Vishnevskij VA, et al. Corporate culture of educational organization as a factor of influencing the social health ofthe Russian student youth[J]. Práxis Educacional, 2019, 15(36),pp. 583-598.

[19] DeBacker J M, Routon P W. A culture of despair? Inequality and expectations of educational success[J]. Contemporary Economic Policy, 2021, 39(3),pp. 573-588.

[20] Zhong X, Babaie SarijalooF, Prakash A, et al. A multidisciplinary approach to the development ofdigital twin models ofcritical care delivery in intensive care units[J]. International Journal of Production Research, 2022, 60(13),pp.4197-4213.

[21] Zhang J, Kwok H H L, Luo H, et al. Automatic relative humidity optimization in underground heritage sites through ventilation system based on digital twins[J]. Building and Environment, 2022, pp. 216.

[22] Yang Y, Yang X, Li Y, et al. Application of image recognition technology in digital twinning technology: Taking tangram splicing as an example[J]. Digital Twin, 2022, 2: 6.

[23] Herman H, Sulistyani S, Ngongo M, et al. The structures ofvisual components on a print advertisement: A case on multimodal analysis[J]. Studies in Media and Communication, 2022, 10(2): 145-154.

[24] Kim D, Kim B K, Hong S D. Digital Twin for Immersive Exhibition Space Design[J]. Webology, 2022, 19(1), pp. 47-36.

[25] Tuan, T. H., Seenprachawong, U., & Navrud, S. (2009). Comparing cultural heritage values in South East Asia–Possibilities and difficulties in cross-country transfers of economic values[J]. Journal of cultural heritage, 10(1), 9-21.

[26] Wang Y, Li Y, Song Y, et al. The influence of the activation function in a convolution neural network model of facial expression recognition[J]. Applied Sciences, 2020, 10(5), pp. 1897.

[27] Tian X, Tang S, Zhu H, et al. Real-time sentiment analysis of students based on mini-Xception architecture for wisdom classroom[J]. Concurrency and Computation: Practice and Experience, 2022, 34(21): e7059.

[28] Briceño I E S, Ojeda G A G, Vásquez G M S. Arquitectura mini- Xception para reconocimiento de sexo con rostros mestizos del norte del Perú[J]. MATHEMA, 2020, 3(1): 29-34.

[29] Razzaq S, Shah B, Iqbal F, et al. DeepClassRooms: a deep learning based digital twin framework for on-campus class rooms[J]. Neural Computing and Applications, 2022, pp. 1-10.

[30] Sun J, Tian Z, Fu Y, et al. Digital twins in human understanding: a deep learning-based method to recognize personality traits[J]. International Journal of Computer Integrated Manufacturing, 2021, 34(7-8),pp. 860-873.

[31] LinC, Xiong S. Controllable face editing for video reconstruction in human digital twins[J]. Image and Vision Computing, 2022, 125, pp. 104517.