Heliyon 10 (2024) e32706

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Enhancing museum experience through deep learning and multimedia technology

Jingbo Wen [a], Baoxia Ma [b,*]

[a] College of Fine Arts, Capital Normal University, Beijing, 100048, China
[b] Academy of Fine Arts, Beijing Institute of Fashion Technology, Beijing, 100029, China

ARTICLE INFO

Keywords:
Artificial intelligence
Museum
Visitor experience
Image recognition
Multimedia technology

ABSTRACT

Amidst the swift progression of artificial intelligence (AI) technology, the museum sector has witnessed a notable inclination towards its adoption. This manuscript endeavours to amplify the interactive milieu of contemporary museum patrons by amalgamating a deep learning algorithm with multimedia technology. The crux of our investigation is the exploration of an adaptive convolutional neural network (CNN) to enrich the interactive engagement of museum visitors. Initially, we leverage the adaptive CNN for the image recognition chore pertaining to museum artifacts and exhibits, thereby facilitating automatic recognition and categorization. Furthermore, to surmount the constraints of conventional pooling algorithms in image feature extraction, we suggest an adaptive pooling algorithm, grounded in the maximum pooling algorithm paradigm. Subsequently, multimedia algorithms are amalgamated into the interactive apparatus, enabling visitors to immerse in exhibits and avail more profound information and experiences. Through juxtaposition with traditional image processing algorithms, the efficacy of our proposed algorithm within a museum ambiance is assessed. Experimental outcomes evince that our algorithm attains superior accuracy and robustness in artifact identification and classification endeavours. In comparison to alternative algorithms, our methodology furnishes more precise and comprehensive displays and interpretations, accurately discerning and categorizing a myriad of exhibit types. This research unveils innovative notions for the digital metamorphosis and advancement of modern museums. Through the incorporation of avant-garde deep learning algorithms and multimedia technologies, the museum visitor experience is elevated, proffering more enthralling and interactive displays. The elucidations of this manuscript hold substantial merit for the continual evolution and innovation within the museum industry.

1. Introduction

Museums encapsulate and exhibit human culture, history, and artistry, acting as stewards and propagators of cultural heritage. Their core mandate extends beyond preservation to offering visitors immersive and enriching explorations [1–3]. The resonance of museum visitor experience has been ascending, attributed to its bearing on audience engagement, allure, and educational efficacy.

In preceding decades, the incursion of artificial intelligence (AI) technologies within museum spheres has been employed to augment visitor experience. Conventional exhibition displays are progressively being supplemented with AI and interactive facets,

* Corresponding author.
E-mail address: 20190054@bift.edu.cn (B. Ma).

https://doi.org/10.1016/j.heliyon.2024.e32706
Received 7 November 2023; Received in revised form 4 June 2024; Accepted 7 June 2024
Available online 12 June 2024

enabling visitor engagement through mediums like touch screens, virtual reality apparatuses, and guided tour gadgets [4,5]. These technological conduits avail a plethora of information, foundational knowledge, and visuals, fostering a more profound comprehension and valuation of museum assortments. Concurrently, the ubiquitous use of smartphones coupled with mobile internet evolution has propelled museums towards harnessing mobile applications and social media platforms for visitor interaction. Through mobile applications, visitors can access exhibition data, guided tours, interactive engagements, and disseminate their exhibition experiences on social media [6]. This digital and AI-enhanced interaction surpasses temporal and spatial constraints, permitting visitors to maintain engagement with museums and partake in an expansive cultural discourse even beyond the physical museum premises. Nonetheless, despite the ascendant importance of modern technology within museum visitor experience, various challenges linger.

To tackle these challenges and stimulate further progression within museum visitor experience domain, scholars have delved into the efficacious amalgamation of artificial intelligence, deep learning, and multimedia technologies within museum displays and interactions [7,8]. Deep learning algorithms have been extensively employed in undertakings like image recognition, speech recognition, and natural language processing under the artificial intelligence umbrella [9–11]. Concurrently, multimedia technologies have been harnessed to boost audience involvement and interactivity, enabling visitors to submerge into the cultural essence of museums through virtual reality, augmented reality, and interactive displays [12,13]. Significant research milestones have been attained, such as the development of deep learning algorithm-based image recognition systems for cultural relics, which autonomously identify relic categories and furnish visitors with more meticulous and accurate exhibit elucidations [14]. Additionally, through the utilization of virtual reality technology and interactive display devices, visitors can engage with virtual exhibits via gestures, sound, and touch, engendering a deeper comprehension and experience of cultural heritage [15]. Despite these advancements, research on interactive devices for museum visitor experience continues to confront challenges including accurate information presentation from traditional exhibitions to visitors utilizing artificial intelligence, catering to the diverse visitor demographics like children, the elderly, and individuals from varied cultural backgrounds, ensuring sustainable technology development and updating, alongside ameliorating the usability and accessibility of installations. More research and exploration are requisite to propel the development of interactive installations for museum visitor experiences, availing visitors a richer and more profound cultural exploration.

Hence, this manuscript's objective is to propose an interactive device that integrates deep learning algorithms and multimedia technologies to ameliorate the modern museum visitor experience. Through rigorous research and experimentation, we endeavor to explore the employment of advanced algorithms and technologies to augment museum displays and visitor engagement, providing a more personalized and immersive experience. Our study emphasizes designing an adaptive convolutional neural network for artifact and exhibit image recognition, enabling automatic identification and classification. To bolster recognition accuracy, we propose a network model that incorporates a bilinear hybrid attention mechanism based on the ResNet50 network. Moreover, we introduce an adaptive pooling algorithm to surmount limitations inherent in traditional pooling algorithms. To facilitate integration with multimedia algorithms, we design an interactive device that allows visitors to interact with exhibits employing technologies like virtual reality and augmented reality. Through gesture recognition, touch screen interfaces, and voice interaction, viewers can delve into the intricate details, background information, and related narratives of the exhibits, accruing deeper insights into the culture they epitomize.

In summation, this manuscript aspires to contribute to the field of museum visitor experience by proposing an interactive device that integrates advanced deep learning algorithms and multimedia technologies. Through the integration of these technologies, we aim to augment museum displays, visitor engagement, and overall experience. Our study is geared towards developing an adaptive convolutional neural network for accurate artifact and exhibit recognition, supplemented by a bilinear hybrid attention mechanism. By integrating multimedia algorithms and technologies like virtual reality and augmented reality, our interactive device avails visitors the opportunity to interact with exhibits, acquire additional information, and immerse themselves in a more personalized and enriching cultural experience. This research underpins the digital transformation and evolution of modern museums, laying the groundwork for innovative advancements in the museum industry.

2. State of the art

2.1. Applications of deep learning and multimedia technologies related to museum visitor experience

Deep learning and multimedia technologies, regarded as quintessential facets of contemporary technology, are pivotal in augmenting the museum visitor experience. When amalgamated with exhibits, multimedia technology furnishes a richer interactive experience and enhanced information dissemination. The paradigm of museum visitor experience has transitioned from traditional static viewing to a more interactive and diverse format. Museums have embraced multimedia technology, virtual reality (VR), augmented reality (AR), and other avant-garde technologies, facilitating more intuitive and immersive interaction between visitors and exhibits. For instance, employing virtual reality technology, visitors can traverse realistic virtual scenes such as exploring ancient edifices or partaking in historical events, thereby enhancing participatory and immersive sentiments. Furthermore, the utilization of augmented reality technology enables visitors to access exhibit-related information and imagery in real time via smartphones or head-mounted devices, enriching the overall visiting experience.

A plethora of scholars and researchers have delved into profound investigations concerning museum visitor experience, proposing an array of innovative methodologies and technologies. In their seminal work, Zhuang et al. [16] melded virtual reality technology with mobile internet to architect a mobile device-centric museum tour guide system. This system permits visitors to selectively choose exhibit interpretations and exhibition routes aligned with their personal interests and requisites, thus enriching their visiting experience. Vardhan et al. [17] crafted an augmented reality-based system for exhibit display. By harnessing augmented reality technology to amalgamate real exhibits with virtual display content, they facilitated a more intuitive comprehension of the historical backdrop and pertinent information of exhibits, thereby augmenting visitor engagement and learning. Hellou et al. [18] posited an emotion recognition method predicated on emotion computing to gauge museum visitors' emotional states. Utilizing cameras and deep learning algorithms to discern and analyze facial expressions and body language of visitors, they envisaged optimizing exhibition content and services to bolster visitors' emotional experiences. Puspasari et al. [19] scrutinized a museum tour guide system based on intelligent recommendations. They employed machine learning and recommendation algorithms to suggest personalized exhibits and tour routes predicated on visitors' interests and preferences, rendering a more personalized and tailored tour experience. Lee et al. [20] conceived a virtual exhibition system for museums predicated on multimedia technology. By amalgamating multimedia technology with 3D modeling and interactive design, they orchestrated a realistic virtual exhibition environment, facilitating remote visitation and interaction, thereby surmounting temporal and spatial limitations and enhancing the visiting experience. These research endeavors elucidate that the deployment of advanced technological apparatuses, such as virtual reality, augmented reality, and intelligent recommendations, can efficaciously augment the museum visitor experience, enabling visitors to comprehend exhibits better, accrue knowledge, and partake in interactions. In addition to the enhancement of the visitor experience, there is a range of research dedicated to designing games to benefit the museum itself or the environment. For example, a study by Kotsopoulos et al. explored how a serious game could be designed to incentivize museums to achieve energy savings [21]. This research provides opportunities and challenges to introduce game elements into museum environments with the aim of engaging visitors and contributing to energy savings.

Recent strides in the field, particularly in the realm of Quality of Experience (QoE)-based museum touring, have exerted a profound influence on the dynamics of museum interactions. Noteworthy among these advancements is the work by Tsiropoulou et al. which employs a human-in-the-loop approach to customize museum touring experiences based on QoE [22]. This approach leverages social network analysis and mining techniques to enhance the visitor's journey through the museum. Furthermore, Tsiropoulou et al. have introduced a model focusing on personalized adaptation strategies, offering valuable insights into the nuanced facets of visitor satisfaction and overall experience [23]. These recent contributions underscore the growing importance of incorporating Quality of Experience considerations into the conceptualization and assessment of museum visitor interactions. In line with this evolving landscape, our study builds upon these foundational works by seamlessly integrating deep learning algorithms and multimedia technologies, thereby taking another stride toward enhancing the overall museum visitor experience.

Deep learning and multimedia technology, constituting an integral component of modern technology, are instrumental in enhancing the museum visitor experience. The fusion of multimedia technology with exhibits avails a richer interactive experience and information delivery. Furthermore, multimedia technology can be harnessed within the virtual exhibition systems of museums to provide opportunities for remote viewing and interaction. Wu et al. [24] crafted a virtual exhibition system for museums predicated on multimedia technology, employing 3D modeling and interaction design to engender a realistic virtual exhibition environment, providing an immersive visiting experience for audiences.

Nonetheless, the deployment of deep learning algorithms necessitates a substantial corpus of labeled datasets and computational resources, escalating the intricacy and cost of implementation. Additionally, the recognition accuracy of deep learning algorithms is a cardinal factor influencing user experience, warranting further amelioration of recognition accuracy.

2.2. *Bilinear convolutional neural network (B–CNN)*

We explore applications of deep learning and multimedia technologies related to museum visitor experiences. This exploration is chosen because it represents a quintessential facet of contemporary technology pivotal in augmenting the museum visitor experience. The integration of multimedia technology with exhibits fosters a richer interactive experience and enhanced information

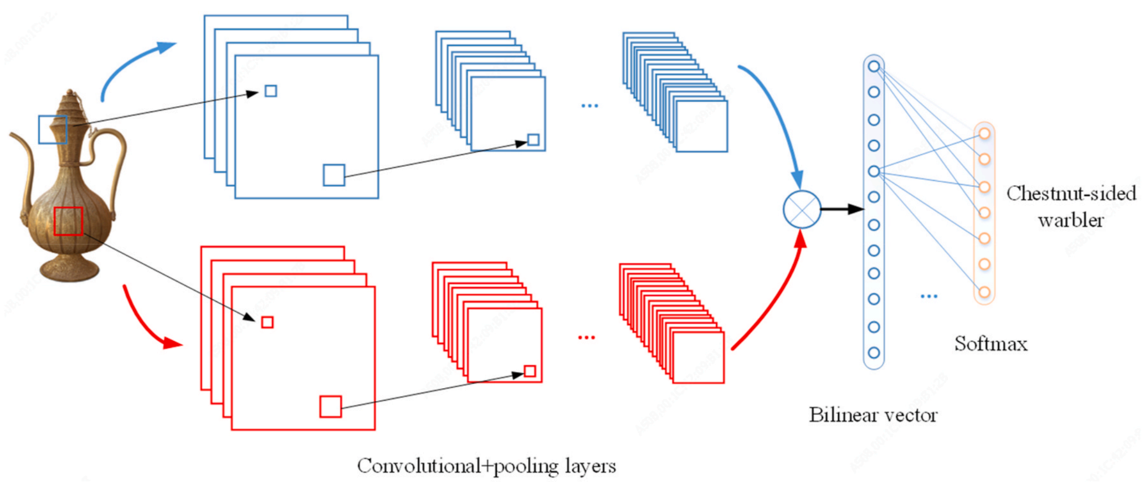

Fig. 1. B–CNN model structure.

dissemination, aligning with the evolving paradigm of museums towards more interactive and diverse formats.

B–CNN is composed of a bifluid CNN which is responsible for completing the feature extraction process [25]. The output results of the two CNN models are then bilinearly manipulated and used as the feature representation of the image. In addition, the model extracts local feature interactions of the image through a translation invariant form. This model has a simple and efficient structure, is a fine-grained image classification model based on weakly supervised information. The B–CNN structure is shown in Fig. 1.

2.3. Attention mechanism module

We introduce the Bilinear Convolutional Neural Network (B–CNN) and the Attention Mechanism module as integral components of our proposed algorithm. The selection of these components is rooted in their proven efficacy in image recognition and feature extraction tasks. The B–CNN structure and Attention Mechanism contribute to the adaptability and accuracy of our algorithm, addressing challenges faced by conventional algorithms in the museum environment.

The attention mechanism is a unique brain signal processing mechanism in humans [26]. Human vision quickly scans the global image to obtain the target area to be focused on, and then devotes more attention resources to this area to obtain more detailed information about the target to be focused on, while suppressing other useless information. The attention mechanism in computer vision is similar to the selective visual attention mechanism in humans, and the core goal is to select the information that is more critical to the current task goal from a large amount of information.

The Hybrid Attention Module (CBAM) consists of a channel attention mechanism and a spatial attention mechanism. This module is designed to enhance the feature representation of an intermediate feature map through a dual-stage process. (1) Channel Attention Mechanism: The first stage involves passing the intermediate feature map through a channel attention map. This channel attention mechanism aims to selectively emphasize or suppress specific channels in the feature map based on their importance to the overall task. The channel attention values are computed element-wise, contributing to the subsequent enhancement of relevant channel information. (2) Spatial Attention Mechanism: Following the channel attention stage, the intermediate feature map undergoes a 2D spatial attention mechanism, resulting in a spatial attention feature map. This stage focuses on capturing relationships between different spatial locations, allowing the model to weigh and combine features across the spatial dimension. Assuming an intermediate feature map $F \in \mathbb{R}^{C\times B\times M}$ is given as input, which first passes through a channel attention map $W_c \in \mathbb{R}^{c\times 1\times 1}$ and then a 2D spatial attention feature map $W_s \in \mathbb{R}^{1\times B\times M}$.

The entire module can be summarized as Formula (1) and (2)

$$F'\& = W_c(F') \otimes F \tag{1}$$

$$F''\& = W_s(F') \otimes F' \tag{2}$$

Where, $\otimes$ denotes element-wise penalty. During the penalty process, the attention values are broadcasted accordingly, i.e., the channel attention values are broadcasted along the spatial dimension. F″ is the feature map with attention.

3. Methodology

3.1. The proposed algorithm in this paper

In this section, we delineate the methodologies employed in our research to enhance museum visitor engagement. The chosen

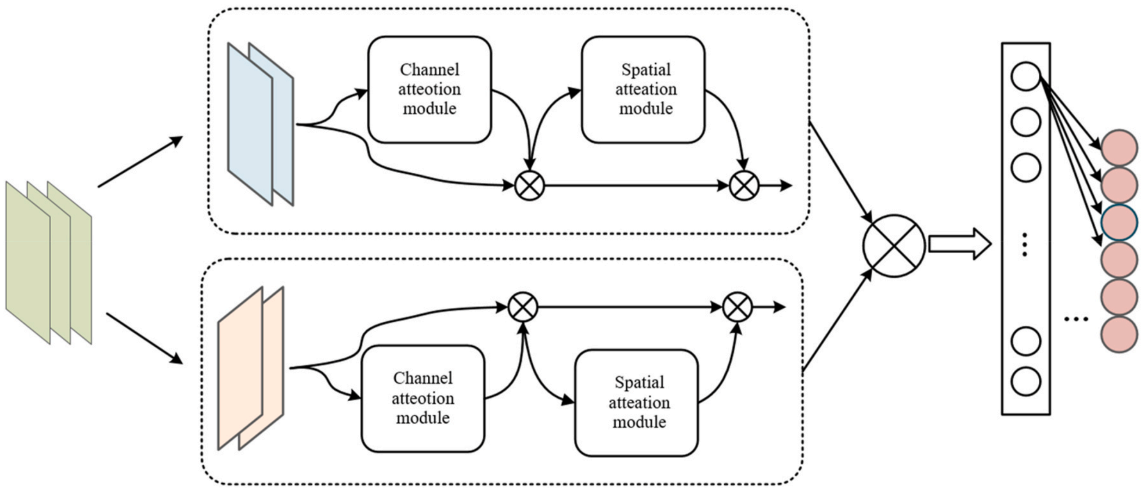

Fig. 2. The proposed network structure.

methodologies, namely the Bilinear Hybrid Attention Module and the Adaptive Pooling Algorithm, have been carefully selected to address specific challenges within the museum context. Figs. 2 and 3 visually represent the architectural components of these methodologies. Fig. 2 illustrates the structure of the Bilinear Hybrid Attention Module, showcasing the convolutional layers, attention mechanisms, and the resulting attention-enhanced features. Fig. 3 provides insights into the Adaptive Pooling Algorithm, depicting the interplay between traditional and adaptive pooling methods for comprehensive feature extraction.

3.1.1. Bilinear hybrid attention mechanism module

Our proposed bilinear hybrid attention mechanism module draws inspiration from established works in image recognition, specifically the mixture-squeeze and excitation module [27]. We adapted this module to the nuances of museum artifact identification by employing two different sizes of convolutional kernels (3 × 3 and 5 × 5) to comprehensively extract features. The resulting features are then blended using the convolutional block attention module, which models relationships between channels and spatial locations. The proposed network structure in this paper is shown in Fig. 2.

In developing the proposed algorithm, we adhered to a systematic and practical process to finalize the formulas described in this subsection. This involved adapting existing published work and tailoring it to suit the specific requirements of our study. The following introductory information sheds light on the practical considerations that influenced the formulation of our bilinear hybrid attention mechanism module. To align with practical challenges, our proposed Bilinear Hybrid Attention Module takes into account real-world dataset variability, encompassing variations in image quality, lighting conditions, and object perspectives commonly encountered in museum environments.

Assume that the inputs and outputs of the network module are Formula (3) and (4)

$$I_{\text{in}} \in \mathbb{R}^{B\times M\times C} \tag{3}$$

$$J_{\text{out}} \in \mathbb{R}^{B\times M\times C} \tag{4}$$

The feature maps produced by the final convolutional layer in the Bottleneck of ResNet50 [28] have a size of B × M × C with C channels. These feature maps are then convolved with 3 × 3 and 5 × 5 convolutional kernels to generate a new set of feature maps denoted as Formula (5) and (6) J_u.

$$J_x = f(x)^{x\times x\times 512\times C'} \otimes I_{in}^{B\times M\times 512} \tag{5}$$

$$J_u = [J_3, J_5] \in \mathbb{R}^{B'\times M'\times C'} \tag{6}$$

The J_u feature maps are then passed through the CBAM (Convolutional Block Attention Module) module, which models the relationships between channels and spatial locations. This completes the integration of the attention mechanism in both the channel and spatial dimensions, resulting in a new set of feature maps denoted as Formula (7):

$$J'_u = [J'_3, J'_5] \in \mathbb{R}^{B'\times M'\times C'} \tag{7}$$

The CBAM module uses a combination of the channel attention module and spatial attention module to weight and combine the feature maps across both dimensions, allowing the network to focus on the most relevant features for each task. The resulting feature maps are then used for subsequent tasks, such as object detection or classification.

The bilinear variation of the eigen maps of the two channel outputs is obtained as Formula (8) $H(l, X, J'_3, J'_5)$.

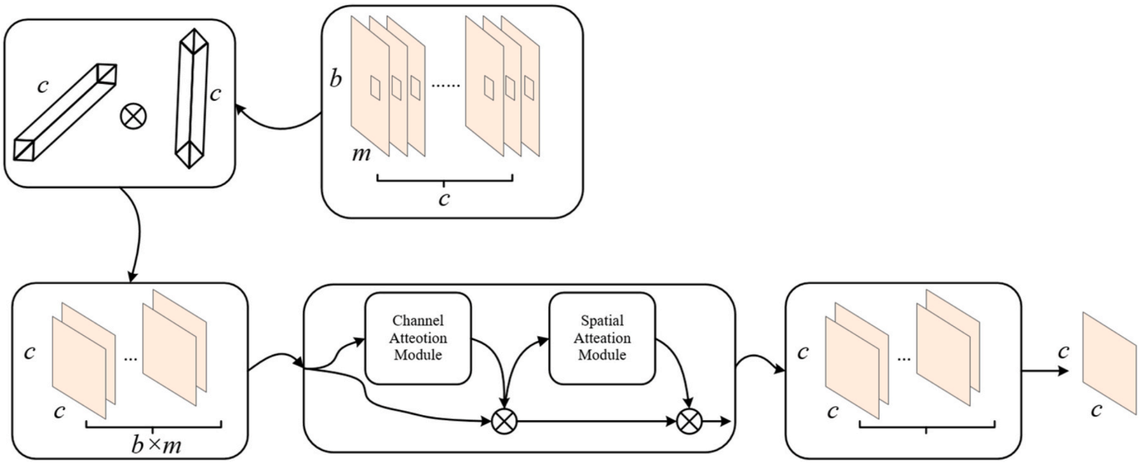

Fig. 3. The overall architecture of proposed method.

$$H\left(l, X, J_3', J_5'\right) = J_3'^{N}(l_3, X) J_5'(l_5, X) \in \mathbb{R}^{C' \times C' \times B'M'} \tag{8}$$

At a given position X, the feature map is denoted as l. This refers to the set of values representing the activations of the neurons at that specific spatial location in the feature map. These values are typically generated by passing the input image through a convolutional neural network (CNN), and represent the features detected by the network at that location. By analyzing the values in the feature map at different spatial positions, the network can identify patterns and objects in the input image.

To obtain $\xi(X)$, the feature maps are first separated by channel, resulting in individual feature maps. Then, the feature values at each spatial position across all feature maps are summed element-wise. This produces a single feature map x, where each element represents the sum of the corresponding elements in the original feature maps. This operation is commonly referred to as channel-wise summation or global summation pooling. The resulting Formula (9) $\xi(X)$ feature map can then be used as input to subsequent layers or modules in the network for further processing or classification.

$$\xi(X) = \frac{1}{B'M'} \sum_{x=1}^{B'M'} H\left(l, X, J_3', J_5'\right) \in \mathbb{R}^{C' \times C'} \tag{9}$$

To obtain the bilinear vector, the $\xi(X)$ feature map is first converted into a vector form. This flattens the spatial structure of the feature map and preserves the channel-wise information. After reshaping, the resulting vector is bilinearly transformed by taking the outer product of itself with a learnable weight matrix. The resulting bilinear vector encodes the interactions between different channels in the feature map. This bilinear vector can be used as a feature representation for subsequent tasks such as image classification or retrieval Formula (10).

$$w = v(\xi(X))w \in \mathbb{R}^{C'^2 \times 1} \tag{10}$$

The L2 regularization operation is performed on m to obtain Formula (11) and (12) J_{out} .

$$t = sign\ (w)\sqrt{|w|}t \in \mathbb{R}^{C'^2 \times 1} \tag{11}$$

$$J_{out} = \frac{t}{\| t \|_2} t \in \mathbb{R}^{C'^2 \times 1} \tag{12}$$

Where, sign is the sign function.

The IPM-SE module differs from the M-SE module in terms of the placement of the attention mechanism. In the IPM-SE module, the attention mechanism is introduced after the bilinear transformation of the pixel values corresponding to the feature map. Specifically, the bilinear features corresponding to each channel and spatial location are weighted by their respective channels and spatial locations using a mixed attention mechanism. This allows the network to highlight the importance of specific regions in the image where important information is located.

The resulting attention-enhanced bilinear features are then passed through a squeeze-and-excitation (SE) module, which further enhances the feature representation by learning channel-wise dependencies and weights. The resulting features are then used for

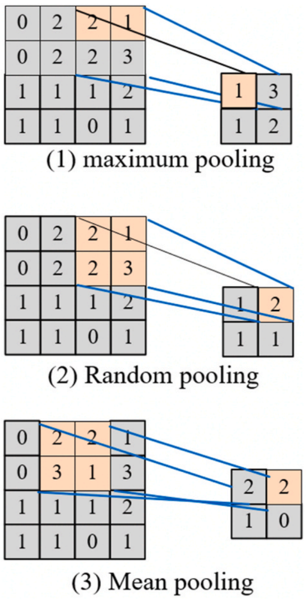

Fig. 4. The pooling process.

subsequent tasks such as image classification or retrieval. The overall architecture of the attention-enhanced bilinear convolutional neural network model incorporating the IPM-SE module is depicted in Fig. 3.

3.1.2. Adaptive pooling algorithm

Pooling is to move several rows or columns to read local features and perform statistical calculations on the local space of the feature map. Therefore, after the convolutional operation, pooling is needed to reduce the dimensionality of the feature map, reduce the risk of over fitting and achieve the translation invariance of the input image. From the current research on convolutional neural networks, there are three mainstream pooling methods: maximum sampling as shown in Fig. 4(1), random sampling as shown in Fig. 4 (2), and mean sampling as shown in Fig. 4(3). The current research on pooling methods has not yet emerged as a complete theoretical basis and is all result-oriented.

This paper leverages the traditional combination of maxima pooling and mean pooling, inspired by well-established pooling methods [29]. The utilization of these traditional methods ensures a comprehensive representation of features, accounting for both edge textures and overall regions. Additionally, the dynamic adaptive pooling algorithm draws theoretical support from the interpolation theorem, adding a layer of adaptability. This strategic integration aims to extract more complete features and enhance the model's capacity to discern feature differences across diverse sub-networks. To enhance user engagement, the dynamic adaptive pooling algorithm is designed to adaptively adjust pooling weights. This adaptation is aimed at providing a more accurate feature representation, considering potential variations in feature maps that may arise from different artifact types. To increase the variability of features available to the two sub-networks, this paper will use the traditional combination of maxima pooling and mean pooling for one sub-network in the model, and a dynamic adaptive pooling algorithm for the other sub-network. In this way, the model can extract more complete features that take into account the edge texture and the overall region, and the diverse pooling methods of the two sub-networks provide more available feature differences.

The dynamic adaptive pooling algorithm is based on the maximum pooling algorithm, and according to the interpolation theorem, a pooling factor μ is added, so that the pooling weights can be adaptively adjusted to extract a more accurate feature representation depending on the feature map. The expression of the algorithm is Formula (13).

$$S_{xy} = \mu \max_{x=1,y=1} \left(F_{xy}\right) + h_2 \tag{13}$$

Where, S_{xy} is the subsampled feature map, $\max_{x=1,y=1} \left(F_{xy}\right)$ is the maximum value extracted from the pooling domain of the input feature map F of size $c \times c$, h_2 is a bias term, and the expression of the pooling factor μ is Formula (14):

$$\mu = \rho \frac{g(q_{max} - g)}{q_{max}^2} + \theta \tag{14}$$

Where, g is the average value in the pooling domain other than the maximum value, q_{max} is the maximum value in the pooling domain, θ is the correction error term, ρ is the characteristic factor, and the expression is Formula (15):

$$\rho = \frac{c}{1 + \left(t_{epo} - 1\right)c^{t_{epo}^2 + 1}} \tag{15}$$

Where, t_{epo} is the iteration factor at training and c is the size of the pooling domain. The pooling factor μ is determined by each parameter in the above two equations and takes a value between 0 and 1, so that both mean pooling and maximum pooling can be taken into account and more accurate features can be extracted.

3.2. Visitor experience interactive system of modern museum based on deep learning technology and multimedia technology

In this section, we outline the system architecture for the visitor experience interactive system of the modern museum. To provide transparency into the design process and justify the choices made in the system's characteristics, we want to elaborate on the steps taken during the development of the architecture.

The development process involved a series of steps aimed at aligning the system with both scholarly considerations and practical requirements. This included pilot visits to modern museums, conducting interviews with museum professionals, and performing a comprehensive requirements analysis. The following details the key aspects of the process: (1) Pilot Visits: To gain insights into the challenges and opportunities faced by modern museums, pilot visits were conducted to several museums. This involved observing visitor interactions, understanding existing technologies in use, and identifying areas that could benefit from technological interventions. (2) Interviews: In-depth interviews were conducted with museum professionals, including curators, exhibition designers, and technology experts. These interviews aimed to gather firsthand information about the specific needs, expectations, and challenges faced by modern museums in enhancing visitor experiences. (3) Requirements Analysis: A thorough requirements analysis was performed to identify the functional and non-functional requirements of the visitor experience interactive system. This process involved considering factors such as the type of exhibits, desired visitor interactions, technological infrastructure, and the integration of deep learning and multimedia technologies.

3.2.1. Requirements analysis

To ensure a systematic and thorough understanding of the museum context and visitor needs, we conducted a detailed requirements analysis employing various methodologies, including pilot visits, interviews, and surveys. In this section, we provide a comprehensive overview of the execution of the requirements analysis phase, shedding light on the methodologies used, tools employed, participant characteristics, survey questions, and data analysis. We conducted pilot visits to a number of modern museums to observe visitors interacting with exhibits, to understand the application of existing technologies, and to identify areas that might benefit from technological interventions.

Prior to the development of our questionnaire and interview guide, an extensive literature review was undertaken to ascertain the gaps and requirements within our research scope. Theoretical foundations and technologies include: 1) User experience and human-computer interaction [30]: it is concerned with how users perceive and interact with technology in museums, and how these interactions affect their experience. 2) Interactive learning and educational technology [31]: It focuses on how educational tools and technologies can facilitate learning, especially in museum environments, and how interactive exhibits can enhance education. 3) Acceptance and use of technology behavior [32]: It is concerned with how visitors accept and use technology in museums, such as mobile apps, augmented reality (AR), and virtual reality (VR). 4) QoE [22]: This concerns QoE-based museum tours, which refers to evaluating and customizing the museum tour experience to meet visitors' expectations and needs. 5) Social network analysis and mining techniques [33]: It is used to enhance the visitor's journey through the museum, related to personalized recommendations and social interaction.

We consider these areas as elements of a theoretical framework, which can be represented in the following diagrammatic form (see Fig. 5).

Informed by these theories and insights, a comprehensive questionnaire was designed to collect feedback and expectations about the museum experience from visitors of different ages and levels of technology exposure. The questionnaires were distributed through multiple online and offline channels, including on-site distribution at the museum, WeChat group distribution, and e-mail distribution. The total number of participants in the questionnaire was 254, with age distribution covering different stages. The questionnaire was crafted to gather feedback and expectations regarding the following aspects.

(1) Demographic Characteristics: Data on gender, age, and educational level were collected through questions in Fig. 6(a, b, c) to inform the design of personalized recommendation algorithms.
(2) Visitation Frequency: Question in Fig. 6(d and e) helped us gauge visitors' familiarity with the museum's content.
(3) Exhibit Preferences: The preferences of visitors for different types of exhibits were directly reflected in the responses to question in Fig. 6(f).
(4) Satisfaction with Current Presentation Methods: Feedback on the satisfaction with the current methods of exhibit presentation was collected via question in Fig. 6(g).
(5) Technology Application Expectations: Questions Fig. 6(h and i) inquired about visitors' expectations and opinions regarding the application of technology in museums, which are vital for the design and integration of our technological solutions.
(6) Perceived Purpose of Technology: The critiques provided by visitors regarding the current state of museum mobile applications, as identified in the question in Fig. 6(j), offered us significant insights.

(2) In-depth interviews: We conducted one-on-one in-depth interviews with museum curators, exhibition designers, and technologists to understand their specific needs, expectations, and challenges in enhancing the visitor experience. Interview

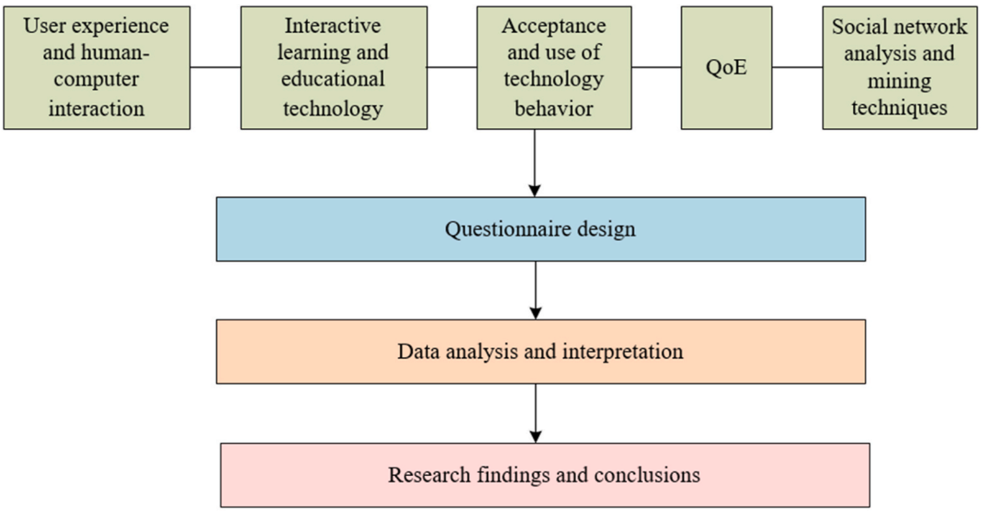

Fig. 5. Structure diagram of questionnaire design theory model.

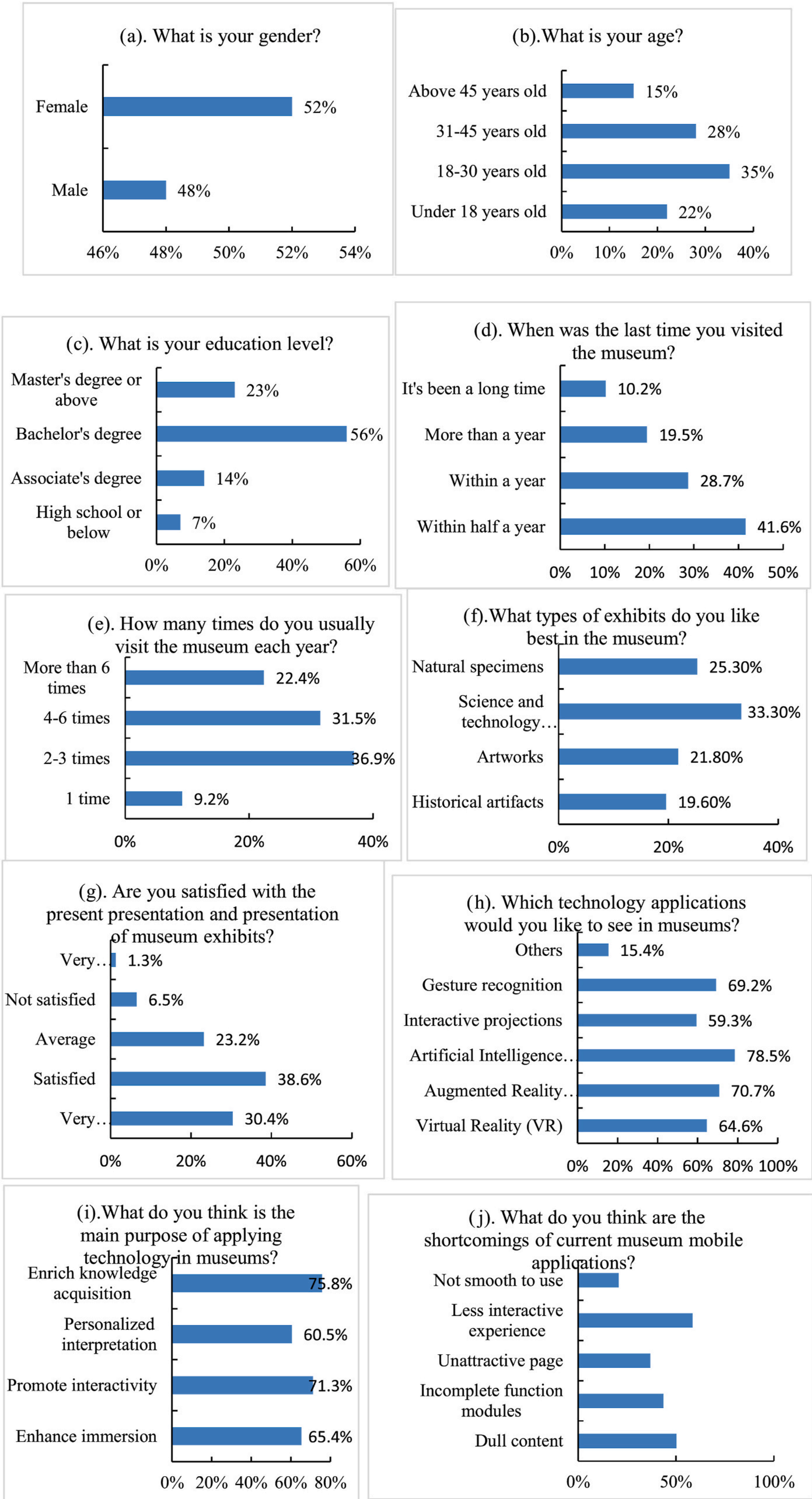

Fig. 6. Museum needs and experience research.

participants included five museum curators, three exhibition designers, and two technologists, with experience ranging from 3 to 15 years. The in-depth interviews were conducted in-house or by videoconference, and the overall implementation time spanned three months. The interviews covered a wide range of topics, including exhibit presentation, interactive experience design, and technology integration needs (see Table 1).

3.2.1.1. Correspondence between survey findings and system design characteristics. Following the presentation of our survey findings in Table 1, we now elucidate the rationale behind our system design characteristics, illustrating how the insights gleaned from our respondents directly influenced the architectural decisions of our museum visitor experience enhancement system. The subsequent discussion maps the survey questions to the corresponding system design elements, showcasing the symbiotic relationship between user feedback and technological innovation.

(1) Demographic Characteristics (Gender, Age, Education Level):

The gender, age, and education level data, collected through questions Fig. 6(a,b,c), have been instrumental in designing a personalized recommendation algorithm capable of catering to a diverse visitor demographic, thereby customizing the museum experience to individual profiles.

(2) Frequency of Museum Visits:

Insights gained from question Fig. 6(d and e) regarding the frequency of visits have been pivotal in understanding visitor familiarity with museum content. This has informed the design of interactive exhibitions and multimedia experiences that engage frequent and first-time visitors alike.

(3) Exhibit Preferences:

Visitor preferences for exhibit types, as expressed in response to question Fig. 6(f), have directly influenced the development of our Adaptive CNN. This technology is now tailored to recognize and suggest exhibits that align with individual visitor interests.

(4) Satisfaction with Current Exhibit Presentation:

Feedback on the satisfaction with current exhibit presentation methods, obtained from question Fig. 6(g), has prompted the integration of a Bilinear Hybrid Attention Mechanism. This feature enhances the system's ability to capture intricate details, providing visitors with a richer and more informative experience.

(5) Desired Technology Applications in Museums:

The expectations visitors hold for technology applications in museums, as revealed in question Fig. 6(h), have been fundamental in our decision to incorporate multimedia technologies and immersive experiences such as Virtual Reality (VR) and Augmented Reality (AR).

(6) Perceived Purpose of Technology in Museums:

Responses to question Fig. 6(i) regarding the perceived purpose of technology in museums have helped us refine the system design to prioritize educational value and visitor engagement, ensuring that technological enhancements complement the museum's

Table 1
Outline of interviews.

Interview topic	Question
Evaluation of existing display modes	What do you think are the shortcomings of modern museums in the presentation and interpretation of exhibits? Is the explanation vivid and comprehensive?
	Is the current visiting experience dull and monotonous? How engaged and interactive are the visitors?
	For different age groups and different backgrounds of visitors, are the existing display methods targeted enough?
Technology needs and expectations	What are the core needs of museums that need to be met with emerging technologies (e.g. VR, AR, AI, etc.)?
	What do you think about the future of using virtual reality, augmented reality and other technologies in museums?
	In what scenarios can artificial intelligence (such as computer vision, natural language processing, etc.) be useful?
Visit experience design suggestions	What are the special needs of visitors of different age groups (such as children, teenagers, middle-aged people, elderly people, etc.) in terms of visiting experience?
	What are your suggestions on accessibility for visitors with different disabilities?
	While integrating new technologies to enhance the visiting experience, how to maintain respect for the core of traditional culture and display forms?
Other suggestions	What other suggestions do you have to help us design a better interactive visitor experience system?

educational mission.

(7) Shortcomings of Current Museum Mobile Applications:

Visitor critiques of current museum mobile applications, as identified in question Fig. 6(j), have been invaluable in designing a user interface and interactive experiences that prioritize ease of use and accessibility.

These survey questions have enabled us to gather comprehensive data on visitor preferences, satisfaction levels, technological expectations, and specific requirements for museum experiences. The data collected is not only affirming of our system's design characteristics but also serves as a valuable input to ensure that our technological solutions are targeted towards addressing the specific needs of our visitors, thereby enhancing the overall museum visitor experience.

3.2.1.2. Analysis and findings. For quantitative data (e.g., Fig. 6), we used statistical analysis methods to calculate the mean, median, frequency distribution and other indicators. For qualitative data (e.g., Table 1), we use code categorization and theme refinement to identify recurring ideas and demand patterns. Through data analysis, we summarized the following key findings.

Finding 1: Visitors generally expect to experience more technological interactions in modern museums, especially the younger generation. This finding prompted us to incorporate a variety of emerging technologies, such as VR, Augmented Reality (AR), and Artificial Intelligence, into our system design to enhance the immersion and interactivity of the visitor experience.

Finding 2: There are significant differences in the needs of visitors of different ages and backgrounds. This led us to design a personalized recommendation system based on artificial intelligence to provide customized content presentation and interaction for different types of visitors.

Finding 3: Museum practitioners place a high value on traditional experiences, but recognize the value of technological innovations. Therefore, our system was designed to complement existing forms of exhibit presentation, maintaining respect for traditional culture while complementing high-tech experiences.

Finding 4: Visitors are positive about the introduction of new technologies in museums, and expect a more engaging and educational experience through technology.

Finding 5: There are still some deficiencies in the existing museum mobile applications, and future development should focus on function expansion, interaction optimization and user experience enhancement.

3.2.2. System architecture overview

By combining insights from pilot visits, interviews, and a comprehensive requirements analysis, the outlined system architecture emerged as a result of a carefully considered and justified design process. This approach ensures that our scholarly work maintains scientific value while also addressing practical considerations in the context of IS design. In this paper, we design a modern museum display system based on deep learning technology and multimedia technology for visitor experience interactivity, as shown in Fig. 7.

As shown in Fig. 7, this paper integrates artificial intelligence, multimedia technology, big data and other information technology on the basis of virtual reality technology, so that the digital museum has the ability of intelligent service, intelligent management and

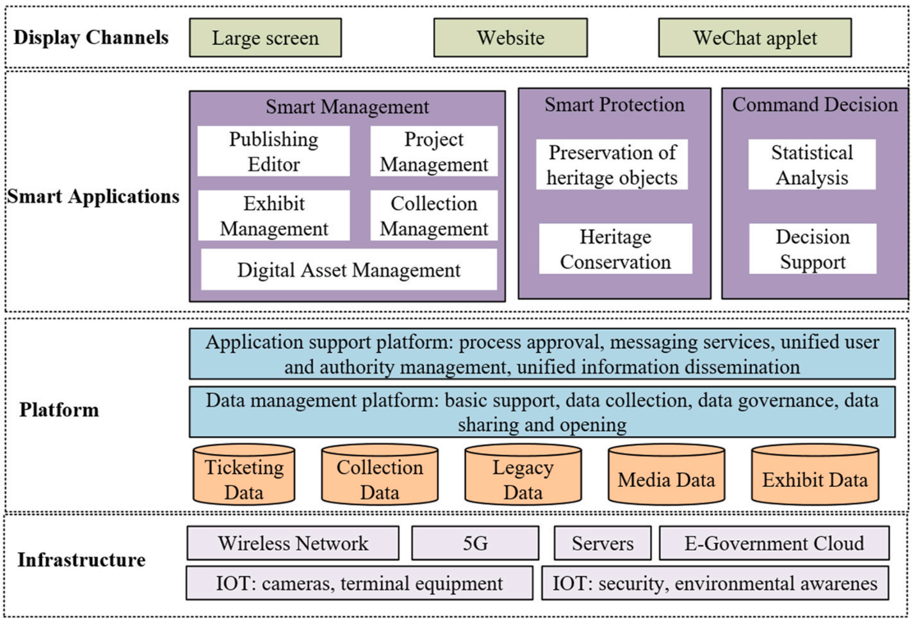

Fig. 7. Interactive system for visitor experience in modern museums based on deep learning technology and multimedia technology.

intelligent decision-making.

(1) Infrastructure includes virtual reality-based network transmission technology, 5G, gigabit broadband and other adaptations to virtual reality, so that digital museums have the ability to build full-scene real-time broadband communication. Virtual reality-based coding and compression, content production, perception and interaction technologies provide the construction of digital museums with ultra-high-definition cameras, sensing terminals and other IoT devices, and virtual reality-based security and trustworthy technologies provide security, safety and environmental perception for digital museums.
(2) The platform includes data management platform and application support platform. The data management platform is based on the fusion application of virtual reality and big data technology, realizing the functions of basic support, data collection, data governance, data sharing and opening. The application support platform is based on the fusion application of virtual reality and network transmission, content production, compression coding and other key technologies to realize the functions of process approval, message service, unified user and authority management, unified information release, etc.
(3) Wisdom applications include wisdom management, wisdom protection and wisdom decision, etc. AR, VR, MR and other interactive technologies play a key role in collection management, exhibition management, heritage conservation, heritage environmental protection, etc., and enhance the interactivity and immersion of digital museums in the visual, auditory and tactile senses. Based on the technical integration of virtual reality + big data + artificial intelligence + cloud computing, it can enhance the wisdom management, wisdom service, wisdom conservation and wisdom decision-making ability of digital museums.
(4) Display channels include large screen, website, and WeChat mini-program. Through these three display channels, users can understand the overall situation of the digital museum and each exhibit. At the same time, these three display channels are also important channels for the digital museum to release news, exhibitions, and event announcements, and users can make reservations for tickets and exhibition activities through the WeChat app. The construction process of display channels is also inseparable from the support of virtual reality technology, through holographic phantom imaging, VR/AR/XR virtual reality, naked eye 3D projection, immersive environmental interaction, panoramic multi-channel, VR animation design, VR interactive experience, XR immersive exhibits, etc. To achieve the unification of the museum theme and interactive interface in text images, interface color, overall layout and other elements, can enhance the digital museum The immersion, interactivity and intelligence level of the digital museum.

These structural modules can be combined with each other in the museum and innovatively designed according to the characteristics and display needs of the exhibits. By providing rich interactive installations, museums can attract a wider audience, making them more actively engaged and immersed, gaining deeper experiences and knowledge.

4. Result analysis and discussion

To evaluate the performance of our proposed algorithm and device in a museum environment, we conducted a series of experiments. We collected a large amount of museum image data and compared it with other traditional image processing algorithms. The experimental results show that our algorithm exhibits higher accuracy and robustness on artifact identification and classification tasks. Compared with other algorithms, our algorithm is able to identify and classify different types of exhibits more accurately, facilitating more accurate and richer displays and interpretations for subsequent processes.

4.1. Dataset acquisition

For our dataset, we gathered samples from various online sources, encompassing a diverse range of traditional artifacts. The dataset

Fig. 8. Some of the images in the dataset.

comprises eight distinct categories, namely gold and silver, bronze, statues, ceramics, jade and stone, enamel, stationery, and carved artifacts. To ensure a comprehensive evaluation, the entire dataset is partitioned into a training set (7248 images) and a validation set (725 images), maintaining a ratio of 9:1. Fig. 8 showcases a random selection of images, while Fig. 9 provides a detailed distribution of categories within the dataset.

4.2. Experimental environment and parameters

The training and testing environments for the experiments in this paper are the same, both for the Windows 11 operating system, and the Pytorch deep learning framework is used to implement the entire model training and testing process, and the specific parameters of the experimental environment are shown in Table 2.

After repeated attempts, the values of the training-related hyperparameters selected in this experiment are shown in Table 3.

4.3. Experiment and analysis

Fig. 10 provides the graphical representation of our analysis strategy in Section 4. Fig. 10(1) included comparative experiments of other algorithms and the proposed algorithm. The mechanisms of Specificity included Accuracy, Precision, Recall, specificity, and Confusion Matrix. Fig. 10(2) includes ablation experiment. Comparison methods include Baseline, +attention mechanism and +adaptive pooling method.

In our pursuit of enhancing recognition accuracy, it is essential to acknowledge the potential limitations and challenges associated with the proposed algorithm. Achieving high accuracy in artifact identification and classification poses certain hurdles that merit consideration.

(1) Dataset Variability: The dataset used in our experiments was collected from the Internet, introducing inherent variations in image quality, lighting conditions, and object perspectives. These factors may impact the algorithm's generalization to diverse museum environments.
(2) Computational Complexity: The proposed algorithm, particularly the bilinear hybrid attention mechanism module, introduces computational complexities due to the integration of attention mechanisms. While this enhances feature extraction, it may also pose challenges in terms of real-time processing, especially in resource-constrained environments.
(3) Sensitivity to Hyperparameters: The performance of deep learning algorithms is often sensitive to hyperparameter choices. In our experiments, we carefully selected hyperparameters; however, further optimization and sensitivity analysis may be required for robust performance across different datasets.
(4) Domain-Specific Challenges: The recognition of artifacts in a museum setting involves domain-specific challenges such as variations in exhibit layouts, lighting conditions, and potential occlusions. These factors may impact the algorithm's performance in real-world museum scenarios.

4.3.1. Experimental comparison analysis of different algorithms

The main evaluation metrics used in this experiment are accuracy, precision, recall and specificity, and confusion matrix.

The confusion matrix is a metric for judging the merit of a model and is often used to judge the merit of classifier models. Table 4 represents the confusion matrix of a binary classification example.

In Table 4, True Positive (TP) denotes instances where the model correctly identifies positive class samples; True Negative (TN) indicates instances where the model correctly excludes negative class samples; False Positive (FP) represents cases where the model incorrectly classifies negative class samples as positive; False Negative (FN) signifies instances where the model erroneously classifies positive class samples as negative. Accuracy (ACC) represents the percentage of samples correctly classified by the model in relation to the total number of samples (including all categories). The calculation formula is as follows Formula (16):

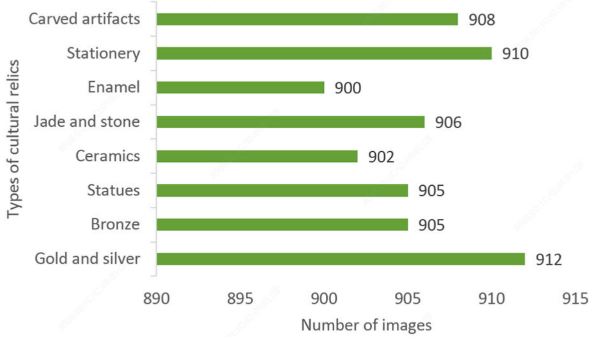

Fig. 9. The specific category distribution of the dataset images.

Table 2
Benchmark model comparison simulation environment.

Simulation Environment	Configuration
CPU	Intel Xeon Silver 4214R@2.40GHzz
GPU	NVIDIA Quadro RTX 3000
Memory	32G
Python version	3.8
Pytorch Versions	1.10.1

Table 3
Network training hyperparameters.

Hyperparameters	Value
Number of categories	8
Optimizer	Adam
Learning rate	1E-04
Training rounds	300
Batch size	16
Pre-training	No

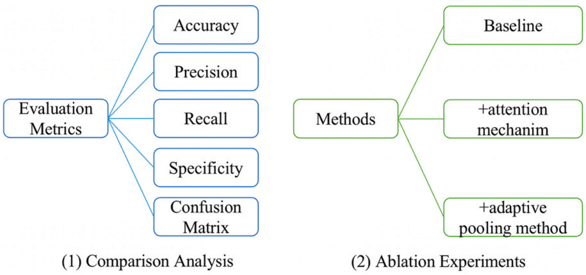

Fig. 10. The graphical representation of our analysis strategy.

Table 4
Dichotomous confusion matrix.

Confusion Matrix		True Value	
		Positive	Negative
Predicted value	Positive	TP	FP
	Negative	FN	TN

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{16}$$

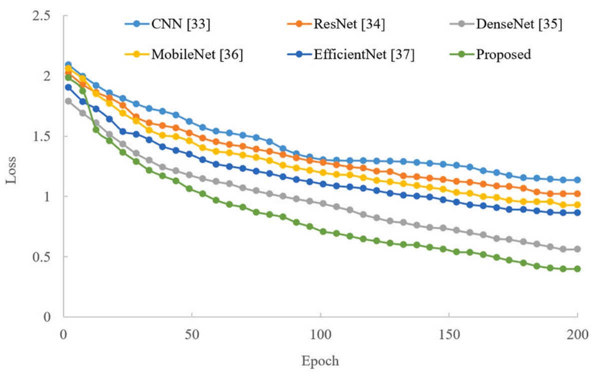

Fig. 11. Comparison of loss convergence of each algorithm.

The process of comparing the loss convergence of the algorithm in this paper with other classical algorithms (CNN [34], ResNet [35], DenseNet [36], MobileNet [37] and EfficientNet [38]) in the training process is shown in Fig. 11.

From Fig. 11, we can see that the loss convergence speed of the method in this paper is faster than the loss convergence speed of other algorithms on the dataset of heritage image classification and recognition.

In order to verify the superiority effectiveness of the algorithm in this paper, the accuracy achieved by each algorithm was compared with other algorithms on the test set of heritage images as shown in Table 5. The accuracy achieved by the improved algorithm using the method in this paper on the test set of heritage images is much higher than that of other benchmark convolutional neural network algorithms, which further proves the effectiveness of the proposed method in this paper.

4.3.2. Ablation experiments

To verify the effectiveness of each module, a series of ablation experiments are conducted in this paper. Experiments are conducted for these three methods: (1) bilinear convolutional neural network based recognition model; (2) attention mechanism; (3) adaptive pooling method. All experiments use the same experimental parameters, and the results are shown in Table 6.

As shown in Table 6, compared with method (1), method (2) uses a hybrid attention mechanism to enhance the model's ability to extract features, and the accuracy is improved by 11.35 %, which proves that the attention mechanism module can improve the accuracy of artifact recognition. Method (3) improves the accuracy of heavy recognition more compared with method (2), and the accuracy is improved by 16.3 %.

4.4. Application scenarios and expected outcomes

In practical museum settings, the application of our proposed methodology and technology could unfold in various scenarios, leading to expected outcomes that significantly impact museum operations and visitor experiences.

(1) Artifact Identification and Classification: Implementing our algorithm in a museum's digital cataloging system could lead to faster and more accurate identification and classification of artifacts, streamlining the cataloging process. The adaptive CNN's capabilities would contribute to efficient artifact management, automating tasks such as inventory management and ensuring proper storage conditions. The expected outcome is increased efficiency in cataloging, resulting in streamlined museum operations, ultimately improving artifact preservation.
(2) Enhanced Visitor Interaction: The integration of our interactive device has the potential to transform visitor engagement. Visitors could use augmented reality to delve into the details of exhibits, fostering a deeper understanding and connection. This interactive experience is anticipated to elevate visitor satisfaction and create a more memorable museum visit. The expected outcome is a more engaged and satisfied visitor base, enhancing the overall visitor experience and encouraging repeat visits.
(3) Dynamic and Educational Exhibits: Multimedia technology integration could be employed to create dynamic and interactive exhibits. For instance, using virtual reality to simulate historical events or utilizing augmented reality for real-time information overlays. This approach is expected to enhance educational programs, offering visitors a more immersive and informative experience. The outcome would be the creation of more captivating and interactive exhibits that contribute to the educational mission of the museum. The expected outcome is an enriched educational experience for visitors, fostering a deeper understanding of cultural heritage.

By envisioning these scenarios and expected outcomes, our research aims to provide practical insights for museum professionals and researchers, demonstrating the positive impact of the proposed methodology on various facets of museum functioning.

5. Discussion

In this section, we aim to strengthen the connection between our research findings and the existing literature, thereby fortifying the contribution of our study to the field. While previous studies have explored various aspects of museum technologies, our work adds valuable insights and advancements in several key areas.

Our proposed algorithm and interactive system for enhancing the visitor experience in museums through the integration of deep learning and multimedia technologies represent a significant step forward in the field of museum technologies. While existing studies have investigated the application of AI and multimedia in museum settings, our work introduces innovative techniques such as the adaptive convolutional neural network (CNN) with a bilinear hybrid attention mechanism and adaptive pooling algorithm. These advancements distinguish our approach from traditional methods, resulting in superior performance in artifact identification and classification tasks, as demonstrated in the experimental results.

Furthermore, our study addresses real-world challenges encountered in museum environments, including varying lighting conditions and exhibit layouts. By prioritizing user feedback obtained during museum trials, we ensure a user-centric approach that enhances usability and overall visitor satisfaction.

The theoretical contributions of our work lie in the development of the bilinear hybrid attention mechanism module and the adaptive pooling algorithm, which improve the accuracy of artifact recognition and feature extraction. These innovations address limitations of traditional methods and provide faster loss convergence, indicating the efficiency of our approach in heritage image classification and recognition datasets.

From a practical perspective, our adaptive CNN, attention mechanism, and pooling algorithm collectively contribute to creating a

Table 5
Comparison of accuracy achieved by algorithms.

Models	Accuracy (ACC)/%
CNN [33]	65.03
ResNet [34]	87.23
DenseNet [35]	90.30
MobileNet [36]	73.64
EfficientNet [37]	81.48
Proposed	93.40

Table 6
Validation of each module.

Methods	Accuracy (ACC)/%
(1) baseline	65. 75
(2) + attention mechanism	77. 10
(3) + adaptive pooling method	93.40

more immersive and interactive museum visitor experience. The accuracy achieved by our algorithm surpasses that of benchmark CNN algorithms, highlighting the practical effectiveness of our approach in real-world museum environments.

In summary, our study fills a gap in the literature by introducing novel methodologies and addressing practical challenges in museum technology. By connecting our findings with past research, we underscore the significance of our contributions to the evolving field of museum technologies. Our work not only enhances visitor experiences but also contributes to the digital transformation and innovation within the museum industry.

6. Conclusion

This study has endeavored to advance the development of an interactive apparatus designed to enhance the visitor experience within contemporary museums. Our approach integrates deep learning algorithms and multimedia technology, focusing on the deployment of adaptive Convolutional Neural Networks (CNN) for automated recognition of artifacts and exhibits. The introduction of a novel bilinear hybrid attention mechanism and an adaptive pooling algorithm contributes significantly to the theoretical framework. Empirical findings demonstrate the heightened accuracy and robustness achieved by the incorporation of adaptive convolutional neural networks in a museum environment. Our proposed algorithms exhibit superior efficacy compared to conventional image processing methods, showcasing precise identification and categorization of diverse exhibits. This study establishes a foundation for the digital transformation and evolution of modern museums, emphasizing the importance of enhancing visitor experience and creating captivating interactive displays. The theoretical contribution lies in the development of innovative algorithms, such as the bilinear hybrid attention mechanism and adaptive pooling algorithm. These advancements address challenges in artifact recognition and feature extraction, setting our approach apart from traditional methods. From a practical standpoint, our work culminates in the creation of an immersive and interactive museum visitor experience.

However, despite the positive outcomes, limitations arise concerning the representativeness of the dataset and potential biases introduced by internet-collected samples. To address these concerns, future research endeavors will prioritize diversifying datasets, collaborating with museums for authentic data, and exploring additional technologies such as augmented reality (AR) and virtual reality (VR) to further enhance the interactive museum experience. In the realm of our requirements analysis, particularly in surveys and interviews, there is acknowledgment of potential methodological biases that may influence participants' responses. To bolster the objectivity of future research methods, we plan to delve into more diversified data collection approaches. This includes considering quantitative surveys, observational studies, and other methodologies to comprehensively and impartially understand visitors' experiences in museums.

Data availability statement

The labeled data set used to support the findings of this study is available from the corresponding author upon request.

CRediT authorship contribution statement

Jingbo Wen: Writing – review & editing, Investigation, Data curation. **Baoxia Ma:** Writing – original draft, Software, Methodology.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing

interests:Baoxia Ma reports administrative support, article publishing charges, statistical analysis, and writing assistance were provided by Beijing Institute of Fashion. Baoxia Ma reports a relationship with Beijing Institute of Fashion that includes: employment, non-financial support, and travel reimbursement. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work is Supported by Ministry of Education, Humanities and Social Sciences project "Research on Museum Experience Design in the Digital Era"(No.22YJA760085).

References

[1] K. Dağ, S. Çavuşoğlu, Y. Durmaz, "The effect of immersive experience, user engagement and perceived authenticity on place satisfaction in the context of augmented reality", Library Hi Tech (2023), https://doi.org/10.1108/LHT-10-2022-04. Vol. ahead-of-print No. ahead-of-print.
[2] S.E. Zaia, K.E. Rose, A.S. Majewski, Egyptian archaeology in multiple realities: integrating XR technologies and museum outreach, Digital Applications in Archaeology and Cultural Heritage 27 (2022) e00249.
[3] Y. Zhou, J. Chen, M. Wang, A meta-analytic review on incorporating virtual and augmented reality in museum learning, Educ. Res. Rev. 100454 (2022).
[4] Z. Ding, Physical interaction and psychological immersion: analysis of digital technologies in museum exhibitions, Academic Journal of Humanities & Social Sciences 6 (6) (2023) 9–14.
[5] K. Yang, H. Wang, The application of interactive humanoid robots in the history education of museums under artificial intelligence, Int. J. Humanoid Rob. (2022) 2250016.
[6] T.L. Chen, Y.C. Lee, C.S. Hung, A study on the effects of digital learning sheet design strategy on the learning motivation and learning outcomes of museum exhibition visitors, Educ. Sci. 12 (2) (2022) 135.
[7] G. Gaia, S. Boiano, A. Borda, Engaging museum visitors with AI: the case of chatbots, Museums and Digital Culture: New Perspectives and Research (2019) 309–329.
[8] E. Ch'ng, S. Cai, F.T. Leow, T.E. Zhang, Adoption and use of emerging cultural technologies in China's museums, J. Cult. Herit. 37 (2019) 170–180.
[9] P. Rani, S. Kotwal, J. Manhas, V. Sharma, S. Sharma, Machine learning and deep learning based computational approaches in automatic microorganism's image recognition: methodologies, challenges, and developments, Arch. Comput. Methods Eng. 29 (3) (2022) 1801–1837.
[10] A. Mukhamadiyev, I. Khujayarov, O. Djuraev, J. Cho, Automatic speech recognition method based on deep learning approaches for Uzbek language, Sensors 22 (10) (2022) 3683.
[11] I. Lauriola, A. Lavelli, F. Aiolli, An introduction to deep learning in natural language processing: models, techniques, and tools, Neurocomputing 470 (2022) 443–456.
[12] V. Shunkov, O. Shevtsova, V. Koval, T. Grygorenko, L. Yefymenko, Y. Smolianko, O. Kuchai, Prospective Directions of Using Multimedia Technologies in the Training of Future Specialists, 2022.
[13] J. Buchner, M. Kerres, Media comparison studies dominate comparative research on augmented reality in education, Comput. Educ. (2022) 104711.
[14] T. Jiang, X.E. Gan, Z. Liang, G. Luo, AIDM: artificial intelligent for digital museum autonomous system with mixed reality and software-driven data collection and analysis, Autom. Software Eng. 29 (1) (2022) 22.
[15] Z. Mao, C. Wong, Y. Zhang, Overview of the application of virtual reality in the field of visual arts and video games. Highlights in Science, Eng. Technol. 44 (2023) 333–340.
[16] X. Zhuang, H. Jiao, K. Lu, Augmented reality interactive guide system and method for tourist attractions based on geographic location, Journal of Electrical and Computer Engineering 2022 (2022).
[17] U. Klentien, Development of virtual museum model for youth, International Journal of Information and Education Technology 12 (4) (2022) 313–317.
[18] M. Hellou, J. Lim, N. Gasteiger, M. Jang, H.S. Ahn, Technical methods for social robots in museum settings: an overview of the literature, International Journal of Social Robotics (2022) 1–20.
[19] S. Puspasari, Machine learning for exhibition recommendation in a museum's virtual tour application, Int. J. Adv. Comput. Sci. Appl. 13 (4) (2022).
[20] H.K. Lee, S. Park, Y. Lee, A proposal of virtual museum metaverse content for the MZ generation, Digit. Creativ. 33 (2) (2022) 79–95.
[21] D. Kotsopoulos, C. Bardaki, T.G. Papaioannou, S. Lounis, G.D. Stamoulis, K. Pramatari, Designing a serious game to motivate energy savings in a museum: opportunities & challenges, in: Games and Learning Alliance: 8th International Conference, GALA 2019, Athens, Greece, November 27–29, 2019, Proceedings 8, Springer International Publishing, 2019, pp. 572–584.
[22] E.E. Tsiropoulou, A. Thanou, S. Papavassiliou, Quality of Experience-based museum touring: a human in the loop approach, Social Network Analysis and Mining 7 (2017) 1–13.
[23] E.E. Tsiropoulou, A. Thanou, S. Papavassiliou, Modelling museum visitors' quality of experience, in: 2016 11th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP), IEEE, 2016, October, pp. 77–82.
[24] Y. Wu, Q. Jiang, H.E. Liang, S. Ni, What drives users to adopt a digital museum? A case of virtual exhibition hall of National Costume Museum, Sage Open 12 (1) (2022) 21582440221082105.
[25] A.J. Prakash, P. Prakasam, An intelligent fruits classification in precision agriculture using bilinear pooling convolutional neural networks, Vis. Comput. 39 (5) (2023) 1765–1781.
[26] D. Soydaner, Attention mechanism in neural networks: where it comes and where it goes, Neural Comput. Appl. 34 (16) (2022) 13371–13385.
[27] B.R. Devassy, J.K. Antony, Histopathological image classification using CNN with squeeze and excitation networks based on hybrid squeezing, Signal, Image and Video Processing (2023) 1–9.
[28] P.S. Satya Sreedhar, N. Nandhagopal, Classification similarity network model for image fusion using Resnet50 and GoogLeNet, Intelligent Automation & Soft Computing 31 (3) (2022).
[29] H. Zang, R. Xu, L. Cheng, T. Ding, L. Liu, Z. Wei, G. Sun, Residential load forecasting based on LSTM fusing self-attention mechanism with pooling, Energy 229 (2021) 120682.
[30] P. Dondi, M. Porta, Gaze-based human–computer interaction for museums and exhibitions: technologies, applications and future perspectives, Electronics 12 (14) (2023) 3064.
[31] C.J. Lin, K.Y. Tang, Y.F. Tu, Advancements and research topics of museum-based mobile learning: a review of journal publications from 2008 to 2019, Interact. Learn. Environ. 31 (8) (2023) 4729–4752.
[32] W. Xu, T.T. Dai, Z.Y. Shen, Y.J. Yao, Effects of technology application on museum learning: a meta-analysis of 42 studies published between 2011 and 2021, Interact. Learn. Environ. 31 (7) (2023) 4589–4604.
[33] D.K.S. Singh, N. Nithya, L. Rahunathan, P. Sanghavi, R.S. Vaghela, P. Manoharan, G.B. Tunze, Social network analysis for precise friend suggestion for twitter by associating multiple networks using ml, Int. J. Inf. Technol. Web Eng. 17 (1) (2022) 1–11.
[34] E. Liu, Research on image recognition of intangible cultural heritage based on CNN and wireless network, EURASIP J. Wirel. Commun. Netw. 2020 (2020) 1–12.
[35] Q. Wang, L. Li, Museum relic image detection and recognition based on deep learning, Comput. Intell. Neurosci. 2022 (2022) 1–11.

[36] L. Cai, H. Li, W. Dong, H. Fang, Micro-expression recognition using 3D DenseNet fused squeeze-and-excitation networks, Appl. Soft Comput. 119 (2022) 108594.
[37] Y. Nan, J. Ju, Q. Hua, H. Zhang, B. Wang, A-MobileNet: an approach of facial expression recognition, Alex. Eng. J. 61 (6) (2022) 4435–4444.
[38] B. Li, B. Liu, S. Li, H. Liu, An improved EfficientNet for rice germ integrity classification and recognition, Agriculture 12 (6) (2022) 863.