Research on Automatic Recognition and Intelligent Recommendation System for Environmental Art Design Styles Based on Deep Learning

Wen Du
[1]Department of Fine Arts and Craft Design, Yuncheng University, Yuncheng 044000, Shanxi, China
Corresponding author: Wen Du (e-mail: 18203594213@163.com)

ABSTRACT Automatic recognition and recommendation of environmental art design styles require robust visual feature extraction and personalized matching under blurred style boundaries and heterogeneous user preferences. To improve recognition accuracy and recommendation effectiveness, this study constructs a deep-learning-based style recognition and hybrid recommendation system. A dataset of 150,000 environmental art design images is built across 10 style categories, including modern, neoclassical, industrial, Nordic, Zen, neo-Baroque, pastoral, Mediterranean, minimalist, and postmodern styles. Images are cleaned, normalized, augmented, and annotated through expert voting. An improved ResNet-50 network with channel–spatial hybrid attention is designed to enhance style-related discriminative regions and capture both global stylistic semantics and local structural details. Stage-wise transfer learning is used for style classification, and deep 256-dimensional style embeddings are extracted for recommendation. User behavior logs, including browsing, liking, collecting, and forwarding, are weighted to construct long-term and short-term preference vectors. A hybrid recommendation strategy integrates style-similarity recall, item-based collaborative filtering secondary ranking, and diversity penalty adjustment. Experiments show that the improved model achieves 92.7% Top-1 accuracy and 98.4% Top-5 accuracy, outperforming VGG-16, Inception-v3, and original ResNet-50. The dynamic preference-fusion strategy improves click-through rate to 28.1% and diversity entropy to 1.82. The system provides an engineering framework for computer vision, image classification, and intelligent recommendation in visual design applications.

INDEX TERMS computer vision, image classification, attention mechanism, hybrid recommendation, style recognition

I. INTRODUCTION

In the field of environmental art design, with the explosive growth of digital design resources, how to quickly and accurately identify design styles from massive images and achieve personalized recommendations has become a practical problem that urgently needs to be solved [1-2]. Traditional methods mainly rely on manual annotation and keyword-based retrieval. This model has obvious defects: on the one hand, the definition of design style is often highly subjective, and different experts may have different style classifications of the same work, resulting in a lack of unified standards for annotation results; on the other hand, the boundaries between style categories are relatively vague. For example, modern style and minimalism, Nordic style and Zen style have overlapping visual features, and traditional algorithms have difficulty capturing deep discriminative features [3]. The above problems directly lead to the difficulty in efficiently retrieving design works, and users face great difficulties in obtaining content that matches their own aesthetic preferences. A more objective and intelligent solution is urgently needed.

Scholars have carried out a series of studies on the problem of environmental art design style identification and recommendation. Torkashvand et al. further pointed out that although the collaborative filtering method based on deep learning has advantages in feature expression ability, it still has shortcomings in dynamic user interest modeling [4]. Zhang et al. systematically reviewed the development of recommendation systems, pointing out that collaborative filtering, content recommendation, and hybrid recommendation have become the mainstream technical paths, but still face the problems of data sparsity and cold start [5]. In terms of improving methods, Jain et al. introduced temporal context information to optimize the collaborative filtering model, which effectively improved the temporal adaptability of recommendations [6]. Latrech et al. proposed a hybrid recommendation framework that integrates demographic information and deep learning, which alle viated the problem of traditional methods relying on single behavioral data to a certain extent [7]. However, existing research generally has two shortcomings: first, it relies too much on user behavior data and ignores the visual style characteristics of the design work itself; second, it lacks modeling of the deep relationship between content semantics and user preferences in the recommendation process, which makes it difficult to achieve a balance between accuracy and diversity in the recommendation results.

In recent years, deep learning technology has shown strong feature extraction capabilities in the field of image classification, providing new ideas for solving style recognition problems. For example, Jiang et al. combined the attention mechanism with convolutional neural networks and further enhanced the model's ability to express key feature regions by introducing a fusion attention structure [8]. Bichri et al. verified the good generalization ability of convolutional neural networks on custom image datasets through transfer learning [9]; Wan et al. classified complex environmental images based on the improved ResNet-50 model, significantly improving the recognition accuracy [10]. However, most of these works are aimed at paintings or photographs. When directly transferred to environmental art design images, the discriminative power of general networks still has room for improvement because design images contain more structured spatial layouts and material elements. In addition, existing recommendation systems often simply combine the recognition results with the recommendation algorithm, lacking modeling of the deep interaction relationship between style features and user behavior. To address the above shortcomings, this paper proposes a method that combines an improved ResNet-50 network with an attention mechanism with a hybrid recommendation strategy, aiming to improve both the accuracy of style recognition and the personalization level of recommendation.

The research target of this paper is to establish an automatic recognition and intelligent recommendation technical route of environmental art design style based on deep learning, so as to solve the problems of high subjectivity and low recommendation accuracy of traditional methods. The research path is as follows: Firstly, an independent construction and preprocessing or augmentation of dataset of 150k environmental art design images with ten styles. Secondly, design an improved ResNet-50 network which combines channel and spatial hybrid attention mechanism, and realize end-to-end style classification of design images through fine-tuning of transfer learning. Thirdly, take the above detected style feature vectors and user historical behavior data to construct hybrid recommendation strategy based on style matching and collaborative filtering, and explore the dynamic weighted fusion scheme of long-term and short-term preference. Finally, conduct systematic experimental verification on constructed dataset and simulated user behavior logs.

II. IMAGE DATA PREPROCESSING AND STYLE LABELING LIBRARY CONSTRUCTION FOR ENVIRONMENTAL ART DESIGN

The image data used in this study mainly came from mainstream domestic and foreign design websites (such as ArchDaily, Gooood, Pinterest) and publicly available architecture and interior design datasets. A total of more than 180,000 original images were collected. After cleaning operations such as deduplication, water mark removal, and removal of images with low resolution and obviously irrelevant content (such as plain text pages and construction drawings), 150,000 valid images were retained. All images were uniformly adjusted to 224×224 pixels, and the original aspect ratio was maintained while center cropping was performed. In order to enhance the generalization ability of the model, random horizontal flipping (probability 0.5), $\pm 15°$ rotation, random adjustment of brightness and contrast (variation range 0.8–1.2 times), and Gaussian noise addition (standard deviation 0.01) were performed on the training set images in sequence [11].

The 10 style categories include Modern, Neoclassical, Industrial, Nordic, Zen, Neo-Baroque, Pastoral, Mediterranean, Minimalist, and Postmodern. Note that the 'Neo-Baroque' category specifically refers to contemporary interior designs that incorporate classical Baroque elements (e.g., ornate carvings, gold leaf) rather than historical 17th-century photographs. This ensures that all images in the dataset share consistent photographic conditions, lighting, and digital quality, eliminating confounding factors related to temporal discontinuity. There are 15,000 images per style category. The annotation task was performed by three researchers independently, all of whom were specialized in environmental art design. To ensure labeling reliability, the inter-annotator agreement was measured using Fleiss' kappa, yielding a coefficient of 0.84, which indicates a high level of substantial agreement among the experts. Images were included in the data set when at least two researchers agreed on their inclusion; otherwise, images were judged by a senior designer "by majority voting". The data was randomly split into training (120,000), validation (15,000), and test set (15,000) according to 8:1:1 ratio, and the sample ratio of each style category was balanced in each data set as well. The annotation result was saved in the JSON file, and the record includes the image path, one-hot encoded style tag, and the information of original inclusion.

III. DESIGN OF A DEEP STYLE FEATURE EXTRACTION NETWORK BASED ON IMPROVED RESNET-50 AND ATTENTION MECHANISM

The base network is ResNet-50 pretrained on ImageNet. ResNet's residual structure helps to solve gradient vanishing problem in deep networks, and its amount of parameters is acceptable, which is appropriate for the cases that images exhibit rich detail but inter-category relationships are subtle enough, such as environmental art design images. To make the network attend more on style-related discriminative regions, we embed a hybrid attention module after each residual module of ResNet-50, and this hybrid attention includes two branches: channel attention and spatial attention.

Channel attention branch applies global average pooling and global max pooling on the input feature map and obtain two one-dimensional vectors. Then, these two vectors are fed into a shared two-layer fully connected network (the compression ratio of the first layer is 16). Finally, sum and activate by sigmoid to obtain channel weights of two outputs, and use these two channel weights to weight the input feature map on channel dimension. Spatial attention branch applies average pooling and max pooling in channel dimension for each feature map, stack these two two-dimensional feature maps and apply a 7×7 convolutional layer to get one single-channel spatial weight map. Then, activate this weight map by sigmoid and spatially weighted. The final output of hybrid attention is element-wise sum of channel-weighted feature map and spatially weighted feature map. This parallel fusion structure allows the network to capture both global stylistic semantics and local structural details simultaneously without the information bottleneck often introduced by sequential processing. In our preliminary ablation experiments, this sum-based fusion outperformed sequential spatial-channel attention (e.g., CBAM) by 1.2% in Top-1 accuracy for complex design images. The global average pooling layer of the improved ResNet-50 classification head is retained, but the original fully connected layer is removed and replaced with a fully connected layer containing 256 neurons (followed by ReLU activation and Dropout, with a ratio of 0.5) and a fully connected layer with an output dimension of 10 (corresponding to ten styles) [12]. Table 1 lists the main configuration parameters of the improved ResNet-50 network.

TABLE 1. Improved ResNet-50 Network Configuration

Module/ Component	Parameter Setting	Description
Backbone network	ResNet-50 (pretrained on ImageNet)	Alleviate gradient vanishing, accelerate convergence
Attention module	Channel attention (reduction ratio 16) + Spatial attention (7 × 7 conv)	Enhance response of discriminative regions
Classification head	Global pooling → 256-dim FC (ReLU, Dropout 0.5) → 10-dim FC	Replace original classification layer
Training strategy	First 30 epochs freeze lower layers; then 20 epochs unfreeze all	Stage-wise transfer learning
Learning rate	0.0001 for first 30 epochs, 0.00001 for next 20	Coarse-to-fine adjustment
Loss function	Cross-entropy loss	Standard loss for multi-class classification
Optimizer	Adam (weight decay 1e-4)	Adaptive learning rate
Batch size	64	Balance memory and stability
Early stopping	Stop if validation loss does not decrease for 10 epochs, restore best model	Prevent overfitting

A. STYLE CLASSIFIER TRAINING AND AUTOMATIC RECOGNITION PROCESS SETUP

After finishing the network design, the next step was to train the style classifier and the automatic recognition procedure. The training procedure was batch iterative, that is, a mini-batch of 64 images was randomly sampled from the training set and subsequently fed into the network. As a consequence of the forward propagation procedure the predicted probability distribution was obtained and then, after the crossentropy loss computation, weights were backpropagated. In order to avoid overfitting during training an early stopping procedure was used. Basically, every epoch was iterated on the validation set and the loss value was computed. If the validation loss failed to decrease for a duration of 10 consecutive epochs, the training process was terminated, and the model weights were reverted to the parameters that achieved the minimum validation loss. In the actual training process, the model reached convergence early. Following the early stopping criterion (patience = 10), the training was automatically terminated at epoch 42. Specifically,

the optimal performance was reached at epoch 32, which included 30 epochs of frozen-layer training and 2 epochs of full-model fine-tuning. The best validation accuracy recorded was 91.8% at this point.

After the model training was completed, it was deployed as an online recognition module. The automatic recognition process includes the following steps: Input an environmental art design image to be recognized, and first perform the same preprocessing as the training stage, including scaling to 224×224 pixels, center cropping and normalization; then input the preprocessed image tensor into the improved ResNet-50 network for forward computation, and the output layer is a 10-dimensional vector, with each dimension corresponding to the original score of a style class; apply the Softmax function to the vector to convert the score into a probability distribution, take the class corresponding to the maximum probability as the style label of the image, and output the top three candidate styles and their confidence (i.e., Softmax output value). The recognition result is returned in JSON format, including the image name, main style label, confidence, and candidate list. The average processing time of the entire process on a single GPU is 42 milliseconds per image, which meets the real-time requirements. For batch processing scenarios, the system supports inputting up to 256 images at the same time, and uses GPU parallel computing to shorten the overall time consumption [13].

IV. USER BEHAVIOR MODELING AND STYLE PREFERENCE VECTOR GENERATION

The range of user behavior data collection includes four types of system log operation: browse, like, collect and forward. Each behavior record includes user id, design work id, behavior type and time. Because different behaviors reflect varying levels of psychological commitment and intent, we assign weights based on the 'Action-Cost' principle: browsing is a low-cost passive behavior (1 point), while liking (3 points) and collecting (5 points) indicate explicit interest. Forwarding is assigned the highest weight (8 points) as it represents a strong active endorsement and social propagation. To ensure consistency across users with different activity levels, each user's behavioral score is min-max normalized before preference vector aggregation to prevent high-frequency users from skewing the feature space. For each user, all design works which the user had interacted in the past 3 months were selected from behavior log to construct user's historical interaction list. Next, using the style recognition model trained in Part 3, the deep style feature vector of each design work is extracted in advance, with a dimension of 256. For user u, the style feature vectors of all works in their historical interaction list are weighted and summed according to the behavior weights to obtain the initial preference vector. The calculation formula is as follows (1):

$$\mathbf{p}_u = \sum_{i \in I_u} w_{ui} \cdot \mathbf{f}_i \tag{1}$$

where I_u is the set of works that user u has interacted with, w_{ui} is the weight of the corresponding behavior, and $\mathbf{f}_i$ is the 256-dimensional style feature vector of work i. To eliminate the impact of differences in the number of interactions, $\mathbf{p}_u$ is L2 normalized to obtain the user's style preference vector.

To simultaneously capture both the user's long-term stable preferences and short-term dynamic changes, the user's historical interactions are divided into two time windows: a long-term window of the past three months and a short-term window of the most recent week. The long-term preference vector $\mathbf{p}_u^{long}$ and the short-term preference vector $\mathbf{p}_u^{short}$ are calculated using the above method. In subsequent recommendation experiments, the two can be combined using a weighted fusion method, with the fusion formula (2):

$$\mathbf{p}_u^{final} = \alpha \cdot \mathbf{p}_u^{long} + (1-\alpha) \cdot \mathbf{p}_u^{short} \tag{2}$$

where α is the weight of the long-term preference, ranging from 0 to 1. When $\alpha = 1$, only the long-term preference is used; when $\alpha = 0$, only the short-term preference is used. In this paper, $\alpha = 0.6$ is selected as the default parameter for dynamic fusion. The final generated user preference vector is stored in a vector database for subsequent recommendation matching calculations.

V. A HYBRID INTELLIGENT RECOMMENDATION STRATEGY BASED ON COLLABORATIVE FILTERING AND STYLE EMBEDDING

After obtaining the user's final preference vector and the style feature vectors of all design works, a recommendation strategy needs to be designed to generate the final personalized recommendation list. Relying solely on style similarity for recommendation can easily lead to a homogenized recommendation result, while relying solely on collaborative filtering may ignore the style attributes of the design works themselves. Therefore, this paper proposes a hybrid recommendation strategy that integrates these two approaches, specifically including three steps: candidate set recall, secondary sorting via collaborative filtering, and diversity adjustment.

A. STEP 1: CANDIDATE SET RECALL

From the design work library (test set, 15,000 images), $N = 500$ works most similar to the user's final preference vector are selected for each user. From these, the top 200 are passed to secondary ranking. To address the data sparsity problem (5,000 users vs. 15,000 items), co-occurrence similarity is calculated using a style-tag smoothed Jaccard index, which incorporates both user interaction overlaps and the shared style attributes identified by our ResNet-50 model to ensure robust similarity estimates even for items with few interactions. Similarity is calculated using cosine similarity, with the formula (3):

$$\mathrm{sim}(\mathbf{p}_u, \mathbf{f}_i) = \frac{\mathbf{p}_u \cdot \mathbf{f}_i}{\|\mathbf{p}_u\| \|\mathbf{f}_i\|} \tag{3}$$

Where $\mathbf{p}_u$ is the user u's final preference vector, and $\mathbf{f}_i$ is the style feature vector of work i. The top 200 works are sorted from highest to lowest similarity to form the candidate set. This step ensures that the recommended content matches the user's preferences in style.

The second step is secondary ranking using collaborative filtering. Within the candidate set, an item-based collaborative filtering algorithm is used to re-rank the 200 works. First, a user-work interaction matrix is constructed (based on browsing, liking, etc.), and the co-occurrence similarity between works is calculated. For each work i in the candidate set, the collaborative score of user u is predicted, as shown in formula (4):

$$\hat{r}_{ui} = \frac{\sum_{j \in N_u} \mathrm{sim}_{item}(i,j) \cdot r_{uj}}{\sum_{j \in N_u} \mathrm{sim}_{item}(i,j)} \tag{4}$$

where N_u is the set of works that user u has interacted with, $\mathrm{sim}_{item}(i,j)$ is the cosine similarity between work i and j (based on user behavior vectors), and r_{uj} is the normalized behavior intensity of user on work j. After calculating the collaborative score of each candidate work, they are re-ranked from high to low. This step uses the behavior patterns of other users to adjust the ranking, so that candidate works that are more similar to the user's historical interaction works are improved.

The third step is diversity adjustment. To prevent recommendations from clustering in a few styles, a diversity control factor is added before the final output. Specifically, the sorted candidate list is iterated through using a Maximal Marginal Relevance (MMR) inspired diversity penalty. For a candidate work i belonging to style category C_k, its score is adjusted by: $S_i' = S_i - \lambda \cdot \sum_{j \in R} I(\text{style}(j) = C_k)$, where S_i is the initial ranking score, R is the set of already selected works, and λ is the penalty coefficient (set to 0.15 based on empirical testing). This effectively lowers the ranking of redundant styles. In ablation studies, this penalty mechanism improved the diversity entropy from 1.58 to 1.82 compared to a baseline without the penalty. The similarity matrix and collaborative filtering matrix are pre-calculated offline. Although the pairwise matrix for 15,000 items is large, we utilize the Faiss library for Approximate Nearest Neighbors (ANN) search and store the matrix in a sparse format, retaining only the top 500 neighbors per item. This reduces storage requirements to approximately 120MB and ensures that online sorting remains within 15 milliseconds. The offline computation takes approximately 18 minutes on the specified NVIDIA Tesla V100 GPU.

VI. SYSTEM RECOGNITION AND RECOMMENDATION PERFORMANCE EVALUATION

A. EXPERIMENTAL PREPARATION

1) Dataset Composition and Division

This experiment employs the self-constructed “Environmental Art Design Style Image Dataset”, including 150,000 high-resolution images of design works, with 10 styles in total, including modern, neoclassical, industrial, Nordic, Zen, Baroque, pastoral, Mediterranean, minimalist and postmodern, each containing 15,000 images. All images are normalized to be 224×224 pixels and augmented with random horizontal flipping, $\pm 15^\circ$ rotation and dithering of colors. The image dataset is split into training set (120,000 images), validation set (15,000 images) and test set (15,000 images) in 8:1:1 split ratio. User behavior data is extracted from system simulation logs, including browsing, liking, collecting and forwarding behavior sequences of 5,000 virtual users, totally about 600,000 interaction records. The simulation was driven by a Multinomial Logit Model (MNL), where each virtual user’s preference for a specific work was calculated based on style-similarity utility and a Gaussian-distributed noise term to reflect realistic decision uncertainty. These simulated preferences were calibrated against historical category popularity metrics from ArchDaily to ensure the synthetic data closely approximates real-world user distribution patterns.

2) Hardware and Software Environment

Experimental Hardware: Intel Xeon Gold 6248 CPU, NVIDIA Tesla V100 (32GB VRAM) GPU, 256GB RAM.

Software environment: Ubuntu 20.04, Python 3.8, PyTorch 1.10, CUDA 11.3, and Flask (recommended service deployment).

3) Performance Comparison Experiment of Style Recognition Models

To further validate the effectiveness of the improved ResNet-50 model with attention mechanism in the environmental art design style recognition task, a comparative experiment is designed as follows. We choose original ResNet-50, VGG-16, Inception-v3 as the baseline models, and carry out ten style recognition tests on the same test set. The Top-1 accuracy and Top-5 accuracy of three models are recorded respectively. The maximum pool strategy and the same hyperparameter are used.

TABLE 2. Performance Comparison Evaluation of Style Recognition Models

Model	Top-1 Accuracy (%)	Top-5 Accuracy (%)
VGG-16	81.5	93.9
Inception-v3	85.2	95.2
Original ResNet-50	87.4	96.3
Our Model (Improved ResNet-50 + Attention)	92.7	98.4

As can be seen from the data in Table 2, the proposed model achieved the best performance in the ten environmental art design style recognition tasks. The Top-1 accuracy reached 92.7%, an improvement of 5.3 percentage points compared to the original ResNet-50, and improvements of 11.2 and 7.5 percentage points compared to VGG-16 and Inception-v3, respectively. In terms of Top-5 accuracy, our model achieved 98.4%, also significantly higher than the other three comparison models. These results demonstrate that embedding a hybrid attention mechanism on top of ResNet-50 effectively enhances the network’s ability to extract style-related deep features, reduces misclassification, and significantly improves the ability to distinguish between similar styles with blurred boundaries (such as modern and minimalist, Nordic and Zen).

4) Recommendation Strategy Ablation and Effect Comparison Experiment

Construct this experiment to validate the influence of long-term and short-term user preference weighted fusion method on the recommendation performance. We setup 3 fusion schemes as followed: Scheme 1: long-term historical interaction; Scheme 2: Latest 1 week interaction; Scheme 3: Dynamic weighted long-term and short-term user preference fusion scheme, weighted by 0.6:0.4. On the same user behavior dataset, we could calculate the click through rate (CTR) and the entropy value of the recommendation list diversity for above 3 schemes.

Figures 2 (a-b) show the comparison results of CTR and average reciprocal ranking for the three recommendation strategies under different recommendation quantities K. The CTR of group C (hybrid strategy) reached its highest value of 27.1% at K=10, significantly higher than group B (collaborative filtering) and group A (style similarity) at all K values; the peak CTR of group B occurred at K=10 at 21.8%, while the peak CTR of group A was only 18.2%. In terms of Mean Reverse Rank (MRR), group C achieved 0.74 at K=10, which is also better than group B’s 0.61 and group A’s 0.47. These results indicate that the proposed hybrid strategy combines the advantages of style feature

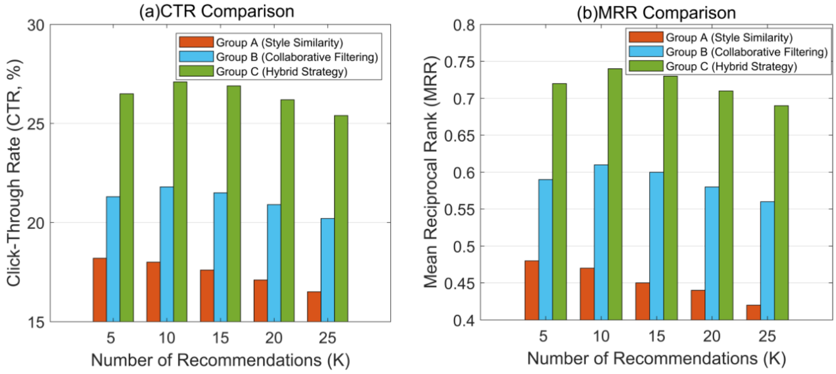

FIGURE 2. Recommendation Strategy Ablation and Effect Evaluation

matching and collaborative filtering, ensuring consistency between recommended content and user visual preferences while leveraging collaborative information to improve ranking accuracy. The overall recommendation performance is significantly better than a single strategy.

5) Experiment on the Impact of Different User Behavior Fusion Methods on Recommendation Performance

We ran an experiment to analyze the impact of different fusion methods of long-term user and short-term user preferences in a recommender system. We designed three schemes as followed: long-term historical behavior only, recent week's behavior only, dynamic fusion of long-term and short-term user preferences and the weighted ratio was 0.6:0.4. We ran the experiment on the same user data set, and then we tested the click through rate (CTR) and diversity entropy of recommendation list.

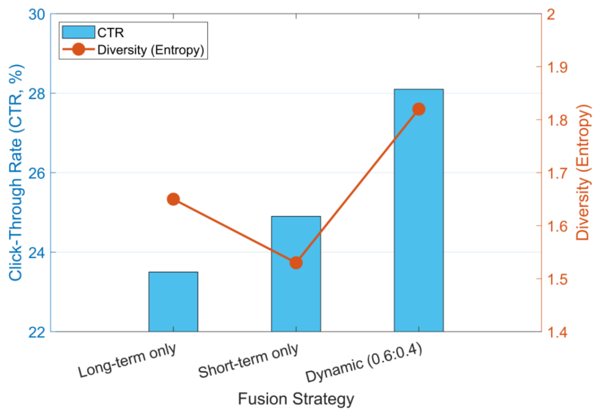

FIGURE 3. Evaluation of the Impact of Different User Behavior Fusion Methods on Recommendation Performance

As shown in Figure 3, different user behavior fusion methods have a significant impact on recommendation performance. The scheme based solely on long-term preferences has a CTR of 23.5% and a diversity entropy of 1.65; the scheme based solely on short-term preferences has a slightly higher CTR of 24.9%, but the diversity entropy decreases to 1.53; while the scheme using a dynamic weighted fusion of long-term and short-term preferences (0.6:0.4) achieved the best results, with a CTR of 28.1% and a diversity entropy of 1.82. This indicates that relying solely on long-term or short-term preferences has limitations. Appropriately fusing both can preserve users' stable style preferences while responding promptly to recent changes in interest, thus achieving a better balance between recommendation accuracy and content diversity.

VII. CONCLUSION

In order to solve the problems of strong subjectivity and low recommendation accuracy existing in the current research on the recognition of environmental art design style, this paper studies the automatic recognition and intelligent recommendation of environmental art design style based on deep learning. The research finds that the improved network applied with attention mechanism can effectively extract the deep style features from design images. The hybrid recommendation strategy improves the accuracy and ranking rationality obviously compared with the single method. The dynamic fusion of users' long-term and short-term preference can achieve the balance between recommendation effectiveness and diversity. Still, there are still some limitations in this paper: the style categories of dataset are limited, user behavior data are from simulated logs and the validation in real scenario is lacking. In the future, we will extend the style categories, introduce online feedback from real users and further improve the performance of the system by fusing the multimodal features (text description, structural sketch, etc.).

AUTHOR CONTRIBUTIONS

Wen Du designed, collected and analyzed the data, and drafted the manuscript. Wen Du conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Wen Du participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] T. Imanishi, “A style classification based on evaluation experiments of residential interiors by experts and identification of explanatory elements and index: A preliminary study towards a classification system of interior styles,” Journal of the Human-Environment System, vol. 27, no. 1, pp. 9-22, 2025, doi: 10.1618/jhes.27.9.

[2] H. Park B and H. Hyun K, “Analysis of pairings of colors and materials of furnishings in interior design with a data-driven framework,” Journal of Computational Design and Engineering, vol. 9, no. 6, pp. 2419-2438, 2022, doi: 10.1093/jcde/qwac114.

[3] B. Wang, S. Zhang, J. Zhang, et al., “Architectural style classification based on CNN and channel-spatial attention,” Signal, Image and Video Processing, vol. 17, no. 1, pp. 99-107, 2023, doi: 10.1007/s11760-022-02208-0.

[4] A. Torkashvand, M. S. Jameii, and A. Reza, “Deep learning-based collaborative filtering recommender systems: a comprehensive and systematic review,” Neural Computing and Applications, vol. 35, no. 33, pp. 24783-24827, 2023, doi: 10.1007/s00521-023-08958-3.

[5] Z. Zhang, B. G. Patra, A. Yaseen, et al., “Scholarly recommendation systems: a literature survey,” Knowledge and Information Systems, vol. 65, no. 10, pp. 4433-4478, 2023, doi: 10.1007/s10115-023-01901-x.

[6] G. Jain, T. Mahara, and C. Sharma S, “Effective time context based collaborative filtering recommender system inspired by Gower’s coefficient,” International Journal of System Assurance Engineering and Management, vol. 14, pp. 429-447, 2023, doi: 10.1007/s13198-022-01813-z.

[7] J. Latrech, Z. Kodia, and N. Ben Azzouna, “CoDFi-DL: a hybrid recommender system combining enhanced collaborative and demographic filtering based on deep learning,” The Journal of Supercomputing, vol. 80, no. 1, pp. 1160-1182, 2024, doi: 10.1007/s11227-023-05519-2.

[8] J. Jiang, H. Xu, X. Xu, et al., “Transformer-Based Fused Attention Combined with CNNs for Image Classification,” Neural Processing Letters, vol. 55, no. 9, pp. 11905-11919, 2023, doi: 10.1007/s11063-023-11402-1.

[9] H. Bichri, A. Chergui, and M. Hain, “Image Classification with Transfer Learning Using a Custom Dataset: Comparative Study,” Procedia Computer Science, vol. 220, pp. 48-54, 2023, doi: 10.1016/j.procs.2023.03.009.

[10] J. Wan, B. Li, K. Wang, et al., “An Improved ResNet50 for Environment Image Classification,” Procedia Computer Science, vol. 242, pp. 1000-1007, 2024, doi: 10.1016/j.procs.2024.08.246.

[11] J. Zhang, X. Yu, X. Lei, et al., “A novel ResNet50-based attention mechanism for image classification,” Journal of Applied Science and Engineering, vol. 27, no. 8, pp. 2961-2969, 2024, doi: 10.6180/jase.202408_27(8).0004.

[12] P. Su, X. Shen, H. Chen, et al., “M-AResNet: a novel multi-scale attention residual network for melting curve image classification,” Multimedia Tools and Applications, vol. 82, no. 27, pp. 42961-42976, 2023, doi: 10.1007/s11042-023-14694-6.

[13] K. Singh, S. Dhawan, and N. Bali, “An Ensemble Learning Hybrid Recommendation System Using Content-Based, Collaborative Filtering, Supervised Learning and Boosting Algorithms,” Automatic Control and Computer Sciences, vol. 58, no. 5, pp. 491-505, 2024, doi: 10.3103/S0146411624700615.