IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 13 February 2025, accepted 20 March 2025, date of publication 25 March 2025, date of current version 3 April 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3554553

RESEARCH ARTICLE

Interior Innovation Design Using ResNet Neural Network and Intelligent Human–Computer Interaction

KEJING GUO AND JINXIN MA
School of Art Design, Zhengzhou University of Industrial Technology, Xinzheng, Henan 451152, China
Corresponding author: Jinxin Ma (jinxinjf1125@163.com)

ABSTRACT This study aims to optimize the interior design generation process while enhancing personalization and efficiency through the development of a hybrid Residual Network-Human-Computer Interaction (ResNet-HCI) framework. It employs Residual Network (ResNet) neural networks, which leverage residual learning to improve training stability and feature extraction capabilities in deep networks. This allows for efficient feature reuse and optimization of model performance. Additionally, Human-Computer Interaction (HCI) technologies, such as voice commands, gesture control, and Virtual Reality/Augmented Reality (VR/AR), are integrated to enhance user interaction with the design system, thereby improving the intelligence and personalization of the interior design workflow. The Large-Scale Scene Understanding dataset is used for model training to evaluate system performance under varying training steps, hyperparameter configurations, and noise conditions. The experimental results show the following: 1) Significant performance variations are observed across different models under conditions such as increased training iterations, noise interference, and design scoring. In the iteration experiment, model performance generally improves with more training steps. ResNet50 consistently outperforms other models, achieving an F1 score of 0.935 after 20 iterations, demonstrating exceptional feature learning and stability. In the noise robustness analysis, ResNet50 and ResNeXt show minimal performance degradation under Gaussian noise, indicating strong noise robustness. 2) Regarding interior design scoring, hybrid layouts generated using ResNet models combined with HCI technologies excel in multiple dimensions, achieving the highest overall satisfaction scores and emerging as the optimal design solutions. These findings validate the exceptional performance and versatility of ResNet50 and ResNeXt in both deep learning and HCI applications. This study provides both theoretical and practical support for the intelligent transformation of the interior design field, while also offering insights into broader applications in creative industries.

INDEX TERMS ResNet neural networks, human–computer interaction, deep learning, intelligent transformation, innovative interior design.

I. INTRODUCTION

A. RESEARCH BACKGROUND AND MOTIVATIONS

In recent years, the rapid development of Artificial Intelligence (AI) has had a profound impact across various industries, with the architecture and interior design fields being particularly notable [1]. Among the many AI technologies, deep learning (DL) models like convolutional neural networks (CNNs) have excelled in tasks such as image recognition, spatial analysis, and pattern prediction. Notably, Residual Network (ResNet), an advanced CNN architecture, has introduced new possibilities for innovative applications in the design field due to its exceptional capabilities in handling complex data [2]. Meanwhile, advancements in Human-Computer Interaction (HCI) technologies have paved the way for more intuitive and dynamic communication between users and systems. Technologies such as gesture recognition, voice commands, and Augmented Reality (AR)

The associate editor coordinating the review of this manuscript and approving it for publication was Chien-Ming Chen.

have become increasingly widespread in creative fields, bridging the gap between user input and machine-generated results [3].

The interior design field currently faces numerous challenges, and traditional design processes have significant limitations. First, interior design relies heavily on the designer's professional experience and personal style, leading to a subjective process that is difficult to standardize and scale [4]. Second, the design cycle is lengthy, with frequent revisions not only increasing costs but also potentially reducing the overall satisfaction with the final product. Additionally, traditional design methods tend to be slow in responding to client needs, with clients often unable to intuitively grasp the final outcome early in the design process, which affects decision-making efficiency [5]. In terms of data analysis and trend forecasting, traditional methods are often reliant on past experience, lacking scientific, quantitative support, which leads to poor market adaptability.

To address these issues, the integration of AI and HCI technologies offers an innovative solution. AI, through DL models, can analyze vast amounts of design data and automatically generate personalized solutions, improving design efficiency and accuracy. Advanced CNN architectures like ResNet can be used for image recognition and spatial analysis, helping designers quickly optimize layouts and color schemes. The introduction of HCI technologies makes interaction more intuitive, such as allowing clients to adjust and preview design proposals in real-time through gesture control, voice commands, or AR visualization, thus enhancing user experience and satisfaction [6]. Furthermore, AI's data analysis capabilities enable more accurate predictions of market trends and optimize the commercial value of design proposals.

In summary, the motivation of this study is to address the limitations of current interior design methods by constructing an intelligent system that enhances design innovation, efficiency, and user engagement. With the integration of AI and HCI technologies, the future of interior design will be more intelligent, personalized, and interactive, bringing new opportunities for the development of the industry.

B. RESEARCH OBJECTIVES

This study aims to explore the integrated application of ResNet neural network simulations and intelligent HCI in the field of interior design. The specific objectives are as follows: (1) Developing a Hybrid ResNet-HCI Framework: This study aims to design a system that integrates the DL capabilities of ResNet with HCI technologies to analyze spatial data, generate innovative interior design solutions, and enable real-time interaction between users and the system. (2) Enhancing User Engagement in the Design Process: By leveraging an intuitive interactive interface, this study seeks to enable users to directly participate in system operations, creating a more personalized design experience. (3) Validating System Effectiveness through Case Studies: The proposed framework will be applied to real-world interior design projects to assess its impact on design quality, efficiency, and user satisfaction. (4) Providing Insights for Future Intelligent Design Systems: This study highlights the potential for further integration of AI and HCI in creative industries and offers recommendations for broader applications in the design domain. Through these objectives, the aim of this study is to provide both theoretical and practical support for AI-driven design innovation and advance interior design toward more efficient and user-centered practices.

II. LITERATURE REVIEW

In recent years, with the rapid development of technologies such as DL and the Internet of Things (IoT), HCI systems have been widely applied in various fields. Related research has focused on improving interaction experiences and system performance through advanced algorithms. DL models, especially CNNs and their variants (e.g., ResNet), have demonstrated excellent performance in tasks such as image recognition, spatial analysis, and pattern prediction. At the same time, advancements in HCI technology have paved the way for more intuitive and dynamic communication between users and systems. The following will discuss the applications of DL models and HCI technology in interior design.

DL models have shown exceptional capabilities in behavior recognition, emotion recognition, and digital twins (DTs) interaction. Han et al. [7] proposed a visual localization system combining Fast Region-based Convolutional Neural Network (RCNN) and ResNet with 101 Layers (ResNet101), introducing a candidate region extraction algorithm that integrates ResNet and Long Short-Term Memory (LSTM). Experimental results showed that the method demonstrated high efficiency, accuracy, and robustness in behavior recognition and object grasping tasks. The model's accuracy and F1 score were both above 0.98, and the success rate of grasping unmarked objects reached 95%. In the field of Human Activity Recognition (HAR), Ronald et al. [8] investigated a model based on the Inception-ResNet architecture to address the limited computational resources of IoT devices. Experimental results indicated that this model outperformed existing methods in terms of accuracy, F1 score, and other metrics on the University of California, Irvine (UCI) public dataset, while significantly reducing computational resource requirements. Moreover, facial emotion recognition is also a critical area of research in HCI. Zhao et al. [9] proposed a new Cross-modal Image Representation and Dual-stream Feature Enhanced Network (CIR-DFENet) for HAR based on wearable sensors. This method employed three techniques—Markov Transition Field (MTF), Recurrence Plot (RP), and Gramian Angular Field (GAF)—to encode time-series data into colored images, and feature learning was performed by combining CNN with Global Attention Mechanism (GAM) and LSTM self-attention mechanisms. The experimental results demonstrated that the method achieved a high accuracy of 99.40%, effectively improving the task's robustness. On the other hand, Lian et al. [10] addressed the issue of insufficient feature utilization in data-driven fault diagnosis methods by proposing Cross-domain Fusion Image and Lightweight Feature Enhanced Network (CFI-LFENet). This method converted raw time-series data into images and processed them through feature extraction, fusion, and enhancement modules, achieving an average accuracy of 100% on public datasets. Additionally, noise resistance experiments further validated its superiority, showing that the method had high adaptability in complex industrial environments. The system used Bluetooth Low Energy (BLE) units, micro Inertial Measurement Unit (mIMU), and sensor fusion technology to achieve efficient keystroke recognition. This approach provided a new direction for future portable HCI systems.

In addition to specific DL models, current research in Human-Machine Interaction (HMI) is also quite extensive. Wang et al. [11] studied HMI based on the DTs technology and proposed an improved Three-Dimensional Visual Geometry Group (3D-VGG) and 3D-ResNet model. This model can capture human skeletal movements and positions from videos, generating high-quality skeletal data to enhance the interaction between virtual and physical data. The approach indicates that this end-to-end solution promotes a more natural response of machines to human behavior, providing more efficient analytical tools for the development of intelligent interactive systems. Zhang and Wang [12] explored the application of AI in interior design and spatial planning. They proposed a CenterNet algorithm based on attention mechanisms and feature fusion to improve the recognition accuracy of interior design elements. They combined LSTM and CNN for spatial layout feature recognition, significantly improving the automation efficiency of interior design. The experiments showed that this method could complete spatial recognition and vector diagram creation in 5 minutes, compared to the 25 minutes required by traditional manual methods. This resulted in a significant increase in efficiency, with a mean Average Precision (mAP) value of 91.0%. It also performed exceptionally well in recognizing complex structures, such as short walls and door corners. On the other hand, Xiao et al. [13] studied the impact of screen brightness and saturation on HMI emotion perception. They used functional near-infrared spectroscopy to assess hemodynamic changes in the prefrontal cortex. The results showed that high brightness and high saturation screens significantly activated the orbitofrontal cortex and the left dorsolateral prefrontal cortex. These areas are associated with pleasantness and attractiveness, respectively. Low brightness and low saturation increased activation in the left dorsolateral prefrontal cortex. These studies revealed the neurophysiological effects of lighting and color on user emotional perception, providing important theoretical support for optimizing interface design and enhancing user experience.

Therefore, the combination of AI and HCI technology has not only enhanced the accuracy and efficiency of intelligent design but also deepened our understanding of human-computer interface optimization and emotion perception. This offers new research directions for the design of future intelligent interactive systems. The application of DL technology provides strong support for the intelligent development of HCI systems. Various studies have made significant progress in areas such as behavior recognition, emotion detection, and DTs interactions. They have also highlighted future research directions. These include optimizing computational resources, improving model interpretability, and achieving more natural multimodal HMI.

III. RESEARCH MODEL

A. INTRODUCTION TO RESNET NEURAL NETWORKS

1) RESNET

ResNet, a deep CNN architecture, was introduced in 2015 [14]. By incorporating residual modules, it effectively addresses challenges in deep networks, such as vanishing gradients, exploding gradients, and information loss [15]. Figure 1 illustrates a residual unit.

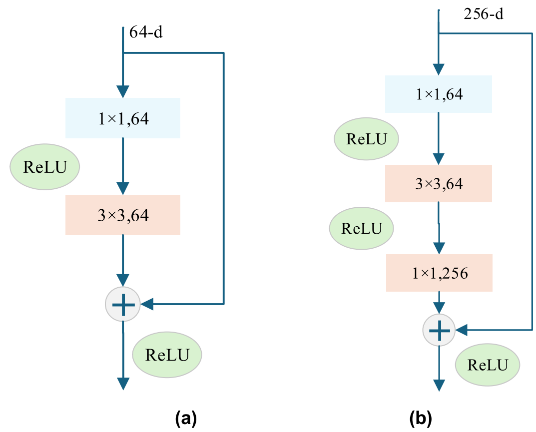

FIGURE 1. The risk zones contain various potential risk sources and corresponding habitat areas (a: standard residual unit; b: "Bottleneck" residual unit).

In ResNet, the residual block is the most fundamental building unit. A typical residual block consists of two convolutional layers and a skip connection [16], [17]. This skip connection structure preserves the integrity of the input features. It also allows gradients to be directly propagated to the shallower layers during backpropagation. As a result, the network's training efficiency and stability are improved [18].

The deep ResNet is constructed by stacking multiple residual blocks, and the final network output can be expressed as shown in Equation (1):

$$H(x) = x + \sum_{i=1}^{n} F_i(x) \tag{1}$$

where n represents the number of residual blocks, and $F_i(x)$ denotes the output of the i-th residual block.

This design offers a novel solution for constructing deeper networks, allowing the number of layers to increase to hundreds or even thousands while maintaining high performance. However, despite ResNet's remarkable success in deep networks, gradient-related issues may still arise in certain situations as the depth of the network increases [19]. Compared to traditional CNNs, ResNet offers significant advantages in the following areas:

(1) Solving the gradient vanishing and degradation problems: Skip connections allow gradients to be more effectively propagated to the shallower layers, improving the training stability of deep networks [20]. (2) Enhancing feature extraction capability: The residual block structure is capable of capturing higher-level features, thus strengthening the representation ability for high-dimensional data. (3) Improved generalization performance: ResNet achieves higher accuracy and robustness in several tasks, such as image classification and object detection. This paves the way for improving future network architectures [21].

2) DENSENET NETWORK

To further address the vanishing gradient problem, Densely Connected Convolutional Networks (DenseNet) introduce a novel connectivity approach. DenseNet establishes direct connections between all layers, allowing each layer to receive feature outputs from all preceding layers. This structure maximizes feature reuse [22]. The core equation for DenseNet is as follows:

$$x_l = H_l([x_0, x_1, \ldots, x_{l-1}]) \tag{2}$$

Where x_l represents the output of the l-th layer.
$[x_0, x_1, \ldots, x_{l-1}]$ refers to the concatenation of all feature maps from the input layer to the $l-1$-th layer.

$H_l(\cdot)$ represents the non-linear transformation at the l-th layer (such as a convolution operation). This design significantly reduces the number of model parameters while effectively avoiding the approach of simply increasing network depth to enhance performance, thus improving training efficiency [23].

3) RESNEXT'S MODULAR DESIGN

Building on ResNet, ResNeXt introduces a more streamlined and efficient network architecture. ResNeXt combines the advantages of ResNet and Inception by incorporating grouped convolutions to further enhance network performance [24].

The output of ResNeXt can be expressed as follows:

$$y = \sum\nolimits_{i=1}^{C} F_i\left(s\right) + s \tag{3}$$

where C represents the number of groups, and $F_i\left(s\right)$ denotes the output of the s-th grouped convolution.

ResNeXt features identical topological structures for each branch, which are combined with ResNet to form the final model architecture. By adjusting the number of groups (cardinality), ResNeXt allows for flexible control over model complexity, reducing the difficulty of hyperparameter tuning while maintaining consistency in the network's topology. This improves transferability and deployment efficiency. Compared to ResNet and DenseNet, ResNeXt achieves a good balance between computational efficiency and model performance through modular design and grouped convolutions [25]. These variant networks have shown outstanding performance in various computer vision tasks and provide significant support for complex applications such as interior design [26]. The success of ResNet and its variants demonstrates that innovation in network architecture is crucial for improving model performance [27]. In the field of interior design, these innovations can enhance the efficiency of spatial data analysis and design generation, driving the intelligent and automated transformation of design processes [28]. Figure 2 illustrates the ResNeXt module.

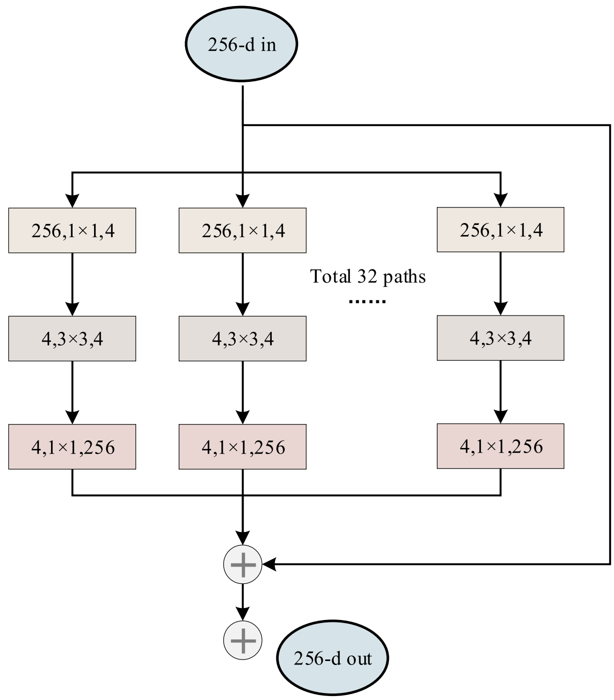

FIGURE 2. ResNeXt module diagram.

B. HCI TECHNOLOGY

1) BASIC PRINCIPLES AND DEVELOPMENT OF HCI

The primary goal of intelligent HCI is to enable computers to better understand and respond to human intentions, facilitating efficient collaboration between humans and machines [29], [30], [31]. Early HCI systems primarily relied on physical input devices, such as keyboards and mice. However, with technological advancements, methods like speech recognition, gesture control, and eye-tracking have gradually become part of the HCI domain, significantly enhancing the naturalness and fluidity of interactions. Figure 3 illustrates the major HCI technologies.

Speech recognition technology allows computers to recognize and understand human language, enabling voice command control. It is widely used in applications such as smart homes and virtual assistants. Using DL models, speech recognition can extract key information from

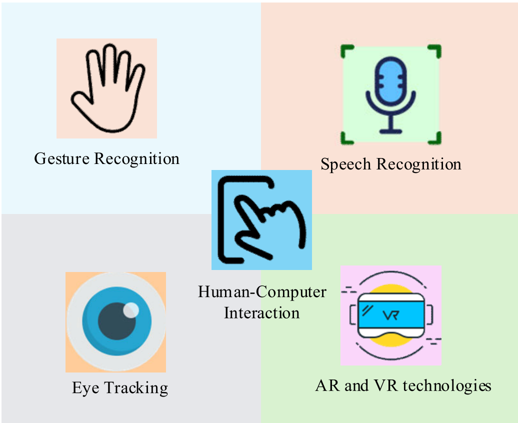

FIGURE 3. **Major HCI technologies.**

complex speech signals and convert it into executable commands for machines [32], [33]. Gesture recognition, which utilizes cameras, sensors, or DL models, enables the identification of user hand movements, facilitating interaction with computers. This technology has significant applications in fields such as Virtual Reality (VR), AR, and interior design. Gesture control is natural and intuitive, providing a more immersive user experience. Eye-tracking technology captures and analyzes eye movements, allowing for the inference of a user's attention and points of interest [34]. It has been widely applied in advertising design, user experience research, and intelligent interaction systems [35], [36]. In interior design, eye-tracking can help designers understand where users focus their attention within a space, enabling them to optimize design solutions. AR and VR technologies provide immersive experiences for users. In interior design, AR overlays virtual elements onto real-world environments, allowing users to preview design effects in real time. VR, on the other hand, creates entirely virtual design spaces, enabling users to interact with and experience design concepts in a simulated environment.

In the field of virtual input devices, customized ring-shaped virtual keyboards and wireless IoT motion recognition rings demonstrate innovative development trends. Zhao et al. [37] proposed a high-precision, layout-optimized customized ring-shaped virtual keyboard aimed at improving input recognition accuracy and user comfort. They showed that improvements, such as arc and segmentation designs, key position adjustments, and letter arrangement optimization, increased the recognition accuracy of the virtual keyboard to 97.57%. These improvements also reduced muscle compression caused by traditional keyboards ($p < 0.05$). The design even halved the fingertip movement distance, significantly enhancing the input experience. On the other hand, Zhao et al. [38] introduced an input solution that combines a wireless IoT motion recognition ring with a paper keyboard. This system uses two to four motion recognition rings on each hand and a 2D keyboard template (e.g., A4 paper) for text input, offering advantages in portability, low cost, and customizability. With two wireless IoT rings worn on each hand, the overall key recognition accuracy reached 94.8%, and when four rings were worn on each hand, the accuracy increased to 98.6%. The system used BLE (Bluetooth Low Energy) units, mIMU (micro Inertial Measurement Unit), and sensor fusion technology to achieve efficient keystroke recognition. This approach provided a new direction for future portable HCI systems.

2) INTEGRATING INTELLIGENT HCI IN INNOVATIVE INTERIOR DESIGN

The intelligent design system in this paper can automatically generate personalized interior design solutions based on real-time user input. For example, the system can generate layouts, color schemes, and furniture choices that meet the user's needs by combining voice or gesture-based inputs with personal preferences and spatial characteristics. Speech recognition technology enables users to interact with the design system through voice commands, performing tasks such as furniture placement, color changes, and style adjustments. For instance, the user can simply say, "Move the sofa to the window" or "Change the wall color to light blue," and the system will automatically recognize and execute the corresponding design adjustments. This process relies on the speech recognition system converting the voice into text commands, which are then analyzed by AI models (such as ResNet) to generate the appropriate solutions. Additionally, the system can confirm the final design through voice feedback, ensuring that it aligns with the user's expectations. To optimize the user experience, the voice recognition interface includes several features. It typically has a speech input button, a real-time speech-to-text display, and a preview window for the design plan. These features greatly enhance the convenience of the operation.

Gesture control technology captures the user's hand movements through cameras or sensors and converts them into input commands for the design system. In practice, users can "grab" virtual furniture and drag it to any location in the room. They can also adjust the lighting and layouts. The system responds in real time, generating the corresponding design. The gesture recognition system interprets these gestures as specific commands, which are then analyzed by AI models to ensure precise interaction. When combined with AR/VR technologies, users can directly manipulate design elements in virtual space, making the design process more intuitive and seamless. The interface design should take into account the sensitivity of gesture recognition and the real-time feedback from the system to ensure smooth interaction.

AR and VR technologies further enhance the immersive design experience. AR technology allows users to overlay virtual design elements onto the real-world environment, enabling real-time previews of design solutions. For example, users can view how different furniture pieces appear in a room through AR glasses or use an AR app to adjust wall colors and materials. In contrast, VR technology provides a completely virtual design space, where users can freely explore and adjust design solutions. Through AR/VR devices, users can interact with the virtual environment using voice commands, gestures, or eye-tracking, with AI models generating personalized design plans based on these inputs. To optimize the immersive experience, the AR/VR interface should be simple and intuitive, incorporating virtual buttons, gesture menus, and eye-tracking interactions to ensure a smooth user experience.

Eye-tracking technology in interior design focuses on analyzing the user's attention distribution to optimize design solutions. By capturing eye movements, this technology identifies the user's focus points and provides valuable data for designers. For example, eye-tracking systems can analyze the user's visual focal points in a space, helping designers adjust layouts, colors, and lighting to better match the user's preferences. When combined with AI models (such as ResNet), the system can automatically optimize design plans based on the user's visual behavior and display the adjusted results in real-time via AR/VR interfaces.

3) METHODS AND SCOPE OF USER DATA COLLECTION

To construct an efficient intelligent interior design system, this paper collects users' voice and gesture data during HCI to optimize the system's ability to understand user inputs and improve the naturalness and responsiveness of interactions. Data Sources: Voice data primarily comes from volunteer experiments and publicly available voice datasets. During the experiments, users of different genders, ages, and linguistic backgrounds were invited to enhance the system's generalizability and robustness. The voice data includes various design-related commands, such as "Move the sofa to the window" and "Change the wall color." Additionally, user feedback voice data is also collected to optimize the system's dialogue interaction capabilities. Automatic speech recognition (ASR) technology is used to transcribe the voice data, and keyword extraction along with semantic analysis is employed to ensure the system can accurately parse user needs.

Gesture data is recorded using depth cameras, such as Leap Motion or Kinect, and sensors built into smart devices, like gyroscopes and accelerometers in smartphones. These capture the user's gesture interactions. A variety of gesture operations in the virtual environment were collected, including actions like "grabbing and moving objects," "rotating furniture," and "scaling objects." These data are used to train the system to accurately identify the user's interaction intentions. The gesture data undergoes preprocessing through computer vision and CNNs. This includes noise reduction, keypoint extraction, and trajectory analysis to enhance recognition accuracy and real-time response capabilities. During data collection, strict adherence to data privacy protection principles is ensured. All user data is collected with informed consent and anonymized to guarantee user privacy and security. Moreover, the data is solely used for optimizing the system in this paper and will not be used for any commercial purposes.

IV. EXPERIMENTAL DESIGN AND PERFORMANCE EVALUATION

A. EXPERIMENTAL DESIGN

In this experiment, the Large-Scale Scene Understanding (LSUN) dataset is used [39], which includes images of interior designs in various styles and layouts, such as bedrooms, living rooms, and study rooms. The dataset consists of 8,000 training images and 2,000 testing images. The hardware environment utilizes an NVIDIA A100 Graphics Processing Unit (GPU) for model training and inference. To ensure the diversity and representativeness of the dataset, images of interior designs covering different styles and functional types are selected from multiple publicly available sources and private databases. The dataset includes various interior design styles, such as modern, minimalist, industrial, and classical, and covers different functional areas, including living rooms, kitchens, offices, and commercial spaces. This ensures that the model can generalize well across various design scenarios.

Data Filtering Process: To improve the quality of the data, a multi-step filtering mechanism is employed. First, low-resolution, blurry, or heavily obstructed images are removed to ensure the clarity of the visual information. Next, images with blurry design elements or excessive noise are excluded after manual review, ensuring the reliability and consistency of the data.

Data Labeling Process: Each image is labeled with two types of tags: (1) Style Labels: Images are categorized according to the primary design style, allowing the model to learn and recognize different design aesthetic features. (2) Function Labels: Images are categorized based on the spatial function (e.g., residential, commercial), helping the model better understand the intended use of the space. The data labeling is performed by domain experts, who manually correct AI-generated preclassification results to ensure high accuracy and consistency. The final dataset provides a solid foundation for the subsequent DL model training, supporting the optimization and improvement of intelligent interior design applications.

To ensure the model performs well across different layers and iteration steps, multiple rounds of hyperparameter tuning are conducted. Table 1 shows the parameter settings.

First, data preprocessing is performed to standardize the interior design samples, ensuring that each input feature is on the same scale. The dataset is divided into a training set (80%) and a test set (20%). Model training is conducted at each training step (5, 10, 15, and 20 steps), using the Adam optimizer and cross-entropy loss function, with the learning rate dynamically adjusted during training. During each experimental phase, model performance is evaluated using metrics such as precision, recall, and F1 score. Next, the impact of noise and image processing on the model's

TABLE 1. Parameter settings.

Parameters	Set Values	Explanation
Training Step	5, 10, 15, 20	Comparison of different training steps to evaluate convergence speed and accuracy.
Batch Size	32	Each training step uses 32 samples for optimization.
Learning Rate	0.001	The initial learning rate is 0.001.
Optimizer	Adam	Adam optimizer is used, combined with learning rate decay.
(Loss Function	Cross-Entropy	Applied to classification tasks.
Regularization	L2 Regularization (0.0001)	It prevents overfitting and enhances the model's generalization ability.

robustness is simulated by assessing the model's performance under conditions such as Gaussian noise, random occlusion, and blurring.

The training process curve is shown in Figure 4.

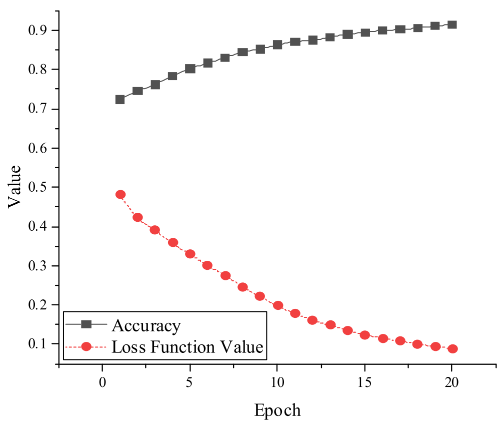

FIGURE 4. Training process curve.

Figure 4 shows that as the number of training steps increases, the training accuracy steadily improves, and the loss function gradually decreases. This indicates that the model is progressively converging. After training for 20 epochs, the performance of the ResNet model stabilizes, with the training accuracy reaching 0.915. Therefore, setting the number of training steps to 20 epochs is sufficient and ensures that the model has fully converged under the given conditions.

B. PERFORMANCE EVALUATION

1) COMPARISON OF MODEL PERFORMANCE AT DIFFERENT ITERATION STEPS

This study evaluates the performance of the models at different training steps using three metrics: precision, recall, and F1 score. The performance comparison results of different models are shown in Table 2.

TABLE 2. Performance comparison of different models.

	Model	Precision	Recall	F1 Score
Training Step = 5	VGG16	0.78	0.75	0.765
	InceptionV3	0.8	0.78	0.79
	DenseNet	0.82	0.8	0.81
	ResNeXt	0.85	0.83	0.84
	MobileNetV2	0.75	0.73	0.74
	ResNet50	0.9	0.88	0.89
Training Step = 10	VGG16	0.82	0.8	0.81
	InceptionV3	0.85	0.83	0.84
	DenseNet	0.87	0.85	0.86
	ResNeXt	0.89	0.87	0.88
	MobileNetV2	0.8	0.78	0.79
	ResNet50	0.92	0.9	0.91
Training Step = 15	VGG16	0.83	0.81	0.82
	InceptionV3	0.86	0.84	0.85
	DenseNet	0.89	0.87	0.88
	ResNeXt	0.91	0.9	0.905
	MobileNetV2	0.82	0.8	0.81
	ResNet50	0.93	0.91	0.92
Training Step = 20	VGG16	0.85	0.84	0.845
	InceptionV3	0.88	0.87	0.875
	DenseNet	0.91	0.9	0.905
	ResNeXt	0.92	0.91	0.915
	MobileNetV2	0.84	0.83	0.835
	ResNet50	0.94	0.93	0.935

At 5 iterations, all models exhibit relatively low performance, with ResNet50 standing out with a precision of 0.90, recall of 0.88, and an F1 score of 0.89, clearly outperforming the other models. ResNeXt follows closely with an F1 score of 0.84, demonstrating strong initial learning capabilities. In comparison, models such as VGG16 and MobileNetV2 perform relatively weaker, with precision and recall values of 0.78 and 0.75, respectively, and an F1 score of 0.765. As the training progresses to 10 iterations, the performance of all models improves significantly, especially for ResNet50, which achieves an F1 score of 0.91, far surpassing the other models. ResNeXt also performs well, with an F1 score of 0.88, demonstrating good stability. DenseNet shows improvement as well, reaching an F1 score of 0.86. In contrast, the performance improvements for VGG16 and MobileNetV2 are relatively slower. At 15 iterations, ResNet50 maintains its leading position, achieving an F1 score of 0.92 and showing stable performance. ResNeXt improves further with an F1 score of 0.905, highlighting its powerful feature learning capability. DenseNet and InceptionV3 reach F1 scores of 0.88 and 0.85, respectively, also demonstrating good stability. By 20 iterations, the model performance peaks. ResNet50, with a precision of 0.94, recall of 0.93, and an F1 score of 0.935, emerges as the strongest model. ResNeXt and DenseNet also maintain excellent performance, with F1 scores of 0.915 and 0.905, respectively, proving their advantages in deeper learning layers.

2) ROBUSTNESS ANALYSIS OF MODELS AT DIFFERENT NOISE LEVELS

To evaluate the robustness of the models under noise interference, various disturbance tests are conducted, including Gaussian noise, random occlusion, and blurring. Figure 5 presents the results of these noise interference tests. It shows that ResNet50 maintains high performance under all levels of noise interference, outperforming the other models. Specifically, under Gaussian noise disturbance, both ResNet50 and ResNeXt show a smaller decrease in F1 score, demonstrating their strong noise robustness. In contrast, VGG16 performs poorly under noise interference, with a significant drop in F1 score, indicating its higher sensitivity to noise.

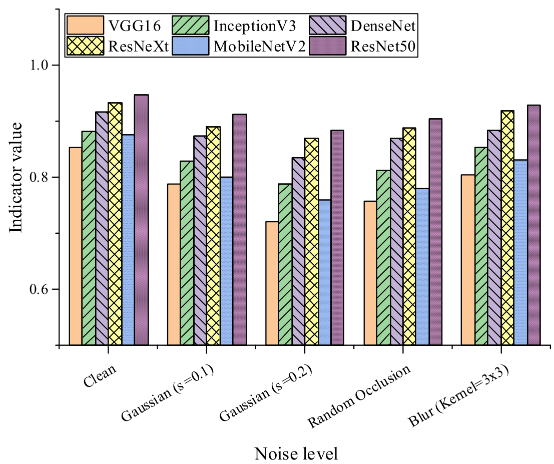

FIGURE 5. The noise interference tests.

3) COMPREHENSIVE EVALUATION FRAMEWORK

In the model performance evaluation, in addition to the F1 score and noise robustness, it is essential to consider multiple dimensions to ensure the model's applicability and stability across different application scenarios. To this end, this study has constructed a multi-dimensional evaluation framework, with the results presented in Figure 6.

The results indicate that ResNet50 and ResNeXt perform best in terms of accuracy and F1 score, reflecting their exceptional feature extraction capabilities. Meanwhile, DenseNet, with its efficient feature reuse mechanism, has an advantage in terms of generalization ability. MobileNetV2, due to its lightweight design, incurs the lowest computational cost but shows a slight decrease in accuracy. Additionally, the noise robustness analysis shows that deep residual structures, such as ResNet50, can effectively mitigate the impact of degraded data quality. Overall, different models exhibit various performance trade-offs across the multi-dimensional metrics, providing valuable insights for model selection in real-world applications.

4) INTERIOR DESIGN EVALUATION

In terms of generating interior design layouts, this study integrates the ResNet model with HCI technology to create multiple design schemes based on user preferences.

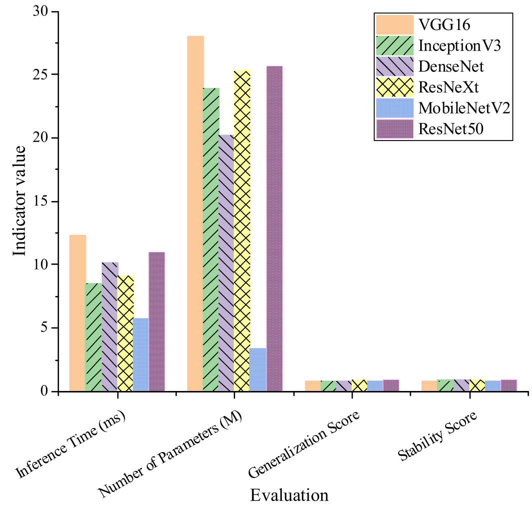

FIGURE 6. Multi-dimensional evaluation results.

To thoroughly evaluate the practical application value of the generated designs, four different layout types are designed and assessed across six dimensions. The experimental data are scored based on design aesthetics, functional rationality, user preference alignment, space utilization efficiency, innovation, and overall satisfaction. Figure 7 displays the performance of the different layout types in practical applications.

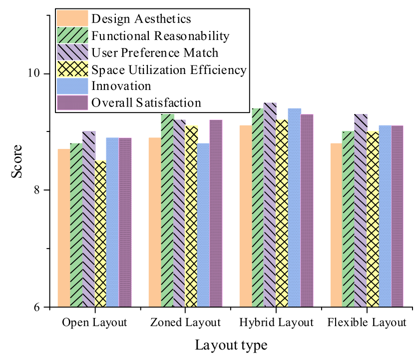

FIGURE 7. Performance of different layout types in practical applications.

The experimental results show that the hybrid layout excels in design aesthetics, functional rationality, user preference alignment, and innovation, with scores of 9.1, 9.4, 9.5, and 9.4, respectively. These results highlight its strong overall advantages, particularly in meeting user needs and fostering innovation. The zoned layout performs exceptionally well in functional rationality (9.3) and space utilization efficiency (9.1), making it ideal for spaces with clear functional distinctions. The flexible layout scores highly in user preference matching (9.3) and innovation (9.1), demonstrating strong adaptability. However, the open layout performs relatively weaker in space utilization efficiency (8.5) and functional rationality (8.8). While it offers good visual appeal, it falls short in terms of space and functionality efficiency. Overall, the hybrid layout emerges as the optimal design scheme due to its balanced performance across multiple dimensions.

C. DISCUSSION

This study combines the ResNet neural network with HCI technology in interior design, exploring its potential to enhance design creativity, efficiency, and user engagement. The experimental results show that the ResNet-based intelligent design system effectively analyzes and generates spatial data, providing efficient and precise design solutions. Additionally, through HCI technology, users can directly interact with the system and participate in the design process in real time, increasing the level of personalization in the design. First, the ResNet model demonstrates superior performance in handling complex data and generating design solutions. Compared to traditional design processes, the intelligent system responds more quickly to design requirements and generates results that align with user preferences based on their input. In particular, ResNet's DL capabilities significantly enhance the accuracy and diversity of designs, especially in spatial layouts and material selection. Moreover, the integration of HCI technology plays a crucial role in improving user experience and participation. Through interactive methods such as voice commands, gestures, and eye tracking, users can express their design needs more naturally, making the design process more intuitive and flexible. This not only strengthens the user's sense of control over the design but also shifts the interaction model between designers and users, giving users more influence in the creative design process.

V. CONCLUSION

A. RESEARCH CONTRIBUTION

The main contribution of this study lies in experimentally validating the effectiveness of the proposed intelligent design framework in real-world design projects. By evaluating key aspects such as design quality, design efficiency, and user satisfaction, the study provides a comprehensive analysis of the framework's application in interior design. The experimental results demonstrate that the framework significantly enhances both the quality and efficiency of designs while improving user satisfaction, thus proving its feasibility and advantages in practical design scenarios. Moreover, this study offers compelling practical evidence for the application of intelligent design systems in the field of interior design, showing that these systems can effectively address challenges present in traditional design processes. Finally, the study provides a fresh perspective for future research in related fields and advances the development of intelligent design technology in real-world applications. It offers both theoretical support and practical insights to further optimize and promote intelligent design systems.

B. FUTURE WORKS AND RESEARCH LIMITATIONS

The continued advancement of HCI technology and multimodal learning (such as combining speech, gesture, and eye-tracking data) holds great potential for providing more personalized and accurate services in interior design. The intelligent design system proposed in this study enhances the interactive experience by integrating VR/AR devices and high-performance GPUs. However, the system's reliance on such advanced hardware could limit its widespread adoption. This is especially true in environments with limited hardware resources. Future research could explore hardware optimization and more universally applicable design solutions to reduce the dependency on high-end hardware.

The impact of cultural differences on design preferences is also a significant factor. Design preferences are often deeply influenced by cultural backgrounds. Future studies could further investigate the needs and preferences of users from different cultural contexts. By analyzing cross-cultural data, design systems could be optimized to better cater to the needs of users from different regions or cultural backgrounds, offering more personalized design solutions.

While the ResNet model performs excellently in standard scenarios, it may face performance degradation when handling extreme or unconventional design requirements. For instance, when confronted with highly complex or creative design requests, the model might struggle to fully understand and generate designs that align with user needs. Future research could improve the system's robustness and generalization capability by incorporating more diverse training data or integrating advanced DL techniques to better handle complex scenarios.

REFERENCES

[1] J. Jiao, W. Liu, Y. Mo, J. Jiao, Z. Deng, and X. Chen, "Dyn-arcFace: Dynamic additive angular margin loss for deep face recognition," *Multimedia Tools Appl.*, vol. 80, no. 17, pp. 25741–25756, Apr. 2021, doi: 10.1007/s11042-021-10865-5.

[2] A. Alnuaim, M. Zakariah, W. A. Hatamleh, H. Tarazi, V. Tripathi, and E. T. Amoatey, "Human–computer interaction with hand gesture recognition using ResNet and MobileNet," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–16, Mar. 2022, doi: 10.1155/2022/8777355.

[3] Q. Zhuang, S. Gan, and L. Zhang, "Human–computer interaction based health diagnostics using ResNet34 for tongue image classification," *Comput. Methods Programs Biomed.*, vol. 226, Nov. 2022, Art. no. 107096, doi: 10.1016/j.cmpb.2022.107096.

[4] A. J. Jalil and N. M. Reda, "Infrared thermal image gender classifier based on the deep ResNet model," *Adv. Hum.-Comput. Interact.*, vol. 2022, pp. 1–11, Jul. 2022, doi: 10.1155/2022/3852054.

[5] T. Tian, L. Wang, M. Luo, Y. Sun, and X. Liu, "ResNet-50 based technique for EEG image characterization due to varying environmental stimuli," *Comput. Methods Programs Biomed.*, vol. 225, Oct. 2022, Art. no. 107092, doi: 10.1016/j.cmpb.2022.107092.

[6] B. Li and D. Lima, "Facial expression recognition via ResNet-50," *Int. J. Cognit. Comput. Eng.*, vol. 2, pp. 57–64, Jun. 2021, doi: 10.1016/j.ijcce.2021.02.002.

[7] X. Han, D. Huang, S. Eun-Lee, and J. Hoon-Yang, "Artificial intelligence-oriented user interface design and human behavior recognition based on human–computer nature interaction," *Int. J. Humanoid Robot.*, vol. 20, no. 6, Dec. 2023, Art. no. 2250020, doi: 10.1142/s0219843622500207.

[8] M. Ronald, A. Poulose, and D. S. Han, "ISPLInception: An inception-ResNet deep learning architecture for human activity recognition," *IEEE Access*, vol. 9, pp. 68985–69001, 2021, doi: 10.1109/ACCESS.2021.3078184.

[9] Y. Zhao, J. Shao, X. Lin, T. Sun, J. Li, C. Lian, X. Lyu, B. Si, and Z. Zhan, "CIR-DFENet: Incorporating cross-modal image representation and dual-stream feature enhanced network for activity recognition," *Expert Syst. Appl.*, vol. 266, Mar. 2025, Art. no. 125912, doi: 10.1016/j.eswa.2024.125912.

[10] C. Lian, Y. Zhao, J. Shao, T. Sun, F. Dong, Z. Ju, Z. Zhan, and P. Shan, ''CFI-LFENet: Infusing cross-domain fusion image and lightweight feature enhanced network for fault diagnosis,'' *Inf. Fusion*, vol. 104, Apr. 2024, Art. no. 102162, doi: 10.1016/j.inffus.2023.102162.

[11] T. Wang, J. Li, Y. Deng, C. Wang, H. Snoussi, and F. Tao, ''Digital twin for human–machine interaction with convolutional neural network,'' *Int. J. Comput. Integr. Manuf.*, vol. 34, nos. 7–8, pp. 888–897, May 2021, doi: 10.1080/0951192x.2021.1925966.

[12] Y. Zhang and J. Wang, ''Artistic sense of interior design and space planning based on human machine intelligent interaction,'' *Intell. Decis. Technol.*, vol. 18, no. 3, pp. 1783–1796, Sep. 2024, doi: 10.3233/idt-240615.

[13] H. Xiao, Z. Zhao, W. Ke, H. Ren, and S. Zhao, ''The influence of screen brightness and saturation on emotional perception in human–machine interactions: A functional near-infrared spectroscopy study,'' *Int. J. Hum.-Comput. Interact.*, vol. 1, pp. 1–12, Nov. 2024, doi: 10.1080/10447318.2024.2425472.

[14] M. F. Qureshi, Z. Mushtaq, M. Z. U. Rehman, and E. N. Kamavuako, ''Spectral image-based multiday surface electromyography classification of hand motions using CNN for human–computer interaction,'' *IEEE Sensors J.*, vol. 22, no. 21, pp. 20676–20683, Nov. 2022, doi: 10.1109/JSEN.2022.3204121.

[15] S. Lu, Q. Hong, B. Wang, and H. Wang, ''Efficient ResNet model to predict protein–protein interactions with GPU computing,'' *IEEE Access*, vol. 8, pp. 127834–127844, 2020, doi: 10.1109/ACCESS.2020.3005444.

[16] Y. Zhang, L. Peng, G. Ma, M. Man, and S. Liu, ''Dynamic gesture recognition model based on millimeter-wave radar with ResNet-18 and LSTM,'' *Frontiers Neurorobot.*, vol. 16, Jun. 2022, Art. no. 903197, doi: 10.3389/fnbot.2022.903197.

[17] S. Auliana, S. Mahrojah, and G. D. P. Aryono, ''Enhanced facial expression recognition through a hybrid deep learning approach combining ResNet50 and ResNet34 models,'' *Kajian Ilm. Inform. dan Komput.*, vol. 4, no. 6, pp. 2676–2685, Jun. 2024. [Online]. Available: https://djournals.com/klik/article/view/1874

[18] S. Mavaddati, ''Voice-based age, gender, and language recognition based on ResNet deep model and transfer learning in spectro-temporal domain,'' *Neurocomputing*, vol. 580, May 2024, Art. no. 127429, doi: 10.1016/j.neucom.2024.127429.

[19] K. S. Yadav, A. M. Kirupakaran, and R. H. Laskar, ''End-to-end bare-hand localization system for human–computer interaction: A comprehensive analysis and viable solution,'' *Vis. Comput.*, vol. 40, no. 2, pp. 1145–1165, Feb. 2024, doi: 10.1007/s00371-023-02837-7.

[20] X. M. Lin, L. Xia, and X. Ye, ''Thermal radiation of tongue surface as a human computer interaction diagnostics technique based on image classification with software interface,'' *J. Radiat. Res. Appl. Sci.*, vol. 17, no. 2, Jun. 2024, Art. no. 100892, doi: 10.1016/j.jrras.2024.100892.

[21] Y. Ren and K. Sun, ''Application effect of human-computer interactive gymnastic sports action recognition system based on PTP-CNN algorithm,'' *Int. J. Adv. Comput. Sci. Appl.*, vol. 15, no. 1, p. 36, Feb. 2024, doi: 10.14569/ijacsa.2024.0150113.

[22] L. Sun and N. Kwak, ''Multimedia human–computer interaction method in video animation based on artificial intelligence technology,'' *Int. J. Inf. Technol. Web Eng.*, vol. 19, no. 1, pp. 1–15, May 2024, doi: 10.4018/ijitwe.344419.

[23] R. Müller, M. Thoß, J. Ullrich, S. Seitz, and C. Knoll, ''Interpretability is in the eye of the beholder: Human versus artificial classification of image segments generated by humans versus XAI,'' *Int. J. Hum.-Comput. Interact.*, vol. 41, no. 4, pp. 2371–2393, Feb. 2025, doi: 10.1080/10447318.2024.2323263.

[24] S. Kaur and N. Kulkarni, ''Recent trends and challenges in human computer interaction using automatic emotion recognition: A review,'' *Int. J. Biometrics*, vol. 16, no. 1, pp. 16–43, Jan. 2024, doi: 10.1504/ijbm.2024.135160.

[25] C. Li, X. Weng, Y. Li, and T. Zhang, ''Multimodal learning engagement assessment system: An innovative approach to optimizing learning engagement,'' *Int. J. Hum.-Comput. Interact.*, vol. 1, no. 2, pp. 1–17, Jun. 2024, doi: 10.1080/10447318.2024.2338616.

[26] S. Bagherzadeh, M. R. Norouzi, S. B. Hampa, A. Ghasri, P. T. Kouroshi, S. Hosseininasab, M. A. G. Zadeh, and A. M. Nasrabadi, ''A subject-independent portable emotion recognition system using synchrosqueezing wavelet transform maps of EEG signals and ResNet-18,'' *Biomed. Signal Process. Control*, vol. 90, Apr. 2024, Art. no. 105875, doi: 10.1016/j.bspc.2023.105875.

[27] J. Park, H. Kang, and H. Y. Kim, ''Human, do you think this painting is the work of a real artist?'' *Int. J. Hum.-Comput. Interact.*, vol. 40, no. 18, pp. 5174–5191, Sep. 2024, doi: 10.1080/10447318.2023.2232978.

[28] H. Zhang, H. Huang, P. Zhao, X. Zhu, and Z. Yu, ''CENN: Capsule-enhanced neural network with innovative metrics for robust speech emotion recognition,'' *Knowl.-Based Syst.*, vol. 304, Nov. 2024, Art. no. 112499, doi: 10.1016/j.knosys.2024.112499.

[29] A. A. Luluh and M. Anandhavalli, ''Identifying voices using convolution neural network models AlexNet and ResNet,'' *Comput. Artif. Intell.*, vol. 2, no. 1, p. 441, Feb. 2024, doi: 10.59400/cai.v2i1.441.

[30] J. Pu and X. Nie, ''Convolutional channel attentional facial expression recognition network and its application in human–computer interaction,'' *IEEE Access*, vol. 11, pp. 129412–129424, 2023, doi: 10.1109/ACCESS.2023.3333381.

[31] H. Zhou, D. Wang, Y. Yu, and Z. Zhang, ''Research progress of human–computer interaction technology based on gesture recognition,'' *Electronics*, vol. 12, no. 13, p. 2805, Jun. 2023, doi: 10.3390/electronics12132805.

[32] A. Dey, S. Biswas, and D.-N. Le, ''Recognition of human interactions in still images using AdaptiveDRNet with multi-level attention,'' *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 10, p. 210, Oct. 2023, doi: 10.14569/ijacsa.2023.01410103.

[33] G. Zellou and N. Holliday, ''Linguistic analysis of human–computer interaction,'' *Frontiers Comput. Sci.*, vol. 6, May 2024, Art. no. 1384252, doi: 10.3389/fcomp.2024.1384252.

[34] S. Chen and J. Wang, ''Virtual reality human–computer interactive English education experience system based on mobile terminal,'' *Int. J. Hum.-Comput. Interact.*, vol. 40, no. 13, pp. 3560–3569, Jul. 2024, doi: 10.1080/10447318.2023.2190674.

[35] W. Li, ''User-centered design for diversity: Human–computer interaction (HCI) approaches to serve vulnerable communities,'' *J. Comput. Technol. Appl. Math.*, vol. 1, no. 3, pp. 85–90, Sep. 2024, doi: 10.5281/zenodo.13506623.

[36] A. Sadeghi Milani, A. Cecil-Xavier, A. Gupta, J. Cecil, and S. Kennison, ''A systematic review of human–computer interaction (HCI) research in medical and other engineering fields,'' *Int. J. Hum.-Comput. Interact.*, vol. 40, no. 3, pp. 515–536, Feb. 2024, doi: 10.1080/10447318.2022.2116530.

[37] Y. Zhao, C. Lian, X. Ren, L. Xin, X. Zhang, X. Sha, and W. J. Li, ''High-precision and customized ring-type virtual keyboard based on layout redesign,'' *IEEE Sensors J.*, vol. 21, no. 22, pp. 25891–25900, Nov. 2021, doi: 10.1109/JSEN.2021.3117948.

[38] Y. Zhao, C. Lian, X. Zhang, X. Sha, G. Shi, and W. J. Li, ''Wireless IoT motion-recognition rings and a paper keyboard,'' *IEEE Access*, vol. 7, pp. 44514–44524, 2019, doi: 10.1109/ACCESS.2019.2908835.

[39] F. Yu, A. Seff, Y. Zhang, S. Song, T. Funkhouser, and J. Xiao, ''LSUN: Construction of a large-scale image dataset using deep learning with humans in the loop,'' 2015, *arXiv:1506.03365*.

KEJING GUO was born in Dengfeng, Zhengzhou, China, in 1983. She received the master's degree from Henan University, China. She is a Teacher with the School of Art Design, Zhengzhou University of Industrial Technology. Her research interests include interior design, landscape design, and furniture design.

JINXIN MA was born in Zhumadian, Suiping, China, in 1989. She received the master's degree from Henan University of Technology, China. She is a Teacher with the School of Art Design, Zhengzhou University of Industrial Technology. Her research interests include interior design, landscape design, and furniture design.