Application of Intelligent Assisted Technology in Green Environmental Art Design and Analysis of Sustainable Paths

Ruihua Fu

Department of Environmental Art Design, Zhengzhou Academy of Fine Arts, Zhengzhou 451450, Henan, China

Corresponding author: Ruihua Fu (e-mail: 15981937601@163.com)

ABSTRACT The integration of intelligent design technologies with sustainable environmental planning requires effective multimodal information processing and adaptive optimization mechanisms to improve ecological performance while maintaining design quality. This paper proposes an intelligent ecological design framework based on a multimodal generative adversarial network (GAN) for green environmental art design. The proposed model combines convolutional feature extraction, semantic embedding, and attention-guided multimodal fusion to jointly represent visual, structural, and ecological information, while an ecological constraint generator and a dual discriminator incorporating structural similarity and energy consumption prediction enable coordinated optimization of aesthetic expression and environmental performance. An adaptive feedback mechanism further updates model parameters according to ecological indicators to achieve closed-loop optimization. Experimental results demonstrate that the generated designs maintain carbon emissions within 48–53 kg CO, energy consumption between 118 and 123 kWh, and ecological consistency indices of 0.87–0.91 while preserving stable visual coordination. Beyond sustainable environmental design, the proposed multimodal perception and optimization strategy provides methodological support for intelligent spatial planning in electromagnetic sensing environments, wireless monitoring infrastructures, and smart built systems where heterogeneous information fusion and energy-aware decision-making are essential for reliable operation.

INDEX TERMS multimodal generative adversarial network, green environmental art design, intelligent ecological optimization, electromagnetic sensing, multimodal information fusion

I. INTRODUCTION

AS an interdisciplinary field that combines ecology, art, and information technology, green environmental art design has emerged as a key avenue for attaining ecological aesthetic transformation and sustainable urban development, a holistic approach increasingly essential for the industry as it transitions toward sustainable material innovation and ethical, aesthetically driven product design. Design activities are now viewed as ecological regulation and environmental optimization tasks in addition to being aesthetic creation processes due to the growing global ecological crisis and energy consumption issues [1,2]. In this regard, the use of AI technology in environmental art design has emerged as a popular area of study for both industry and academia [3,4]. Intelligent assistance technology has changed the traditional design paradigm through algorithmic modeling and data-driven decision-making, shifting the design process from experience-driven to a scientific decision-making model based on information feedback, thus providing new technical support for the implementation of green design concepts [5,6]. The advancement of this research direction has important theoretical and practical significance for improving the ecological adaptability of design, promoting energy efficiency, and building an intelligent sustainable design system [7,8].

Despite the continuous deepening of relevant research, green environmental art design still faces several key difficulties in integrating intelligence and ecology [9,10]. Existing design models have a low degree of integration of environmental multimodal data, and the coupling relationship between environmental visual characteristics, ecological parameters, and semantic information has not been effectively modeled, resulting in a lack of systematic ecological feedback in design results [11,12]. Secondly, most ecological design optimization processes rely on static assessment models and lack intelligent control mechanisms that dynamically respond to environmental changes and energy consumption, resulting in deviations in the long-term ecological adaptability of design results [13,14]. It is challenging for design solutions to concurrently balance ecological performance and aesthetic expression because current green design assistance systems mostly rely on algorithmic recommendations or parameter constraints without a connection between high-dimensional semantic understanding and generative mechanisms [15,16]. Numerous areas of study reflect these concerns. For instance, the use of generative models in the design of art spaces, energy-efficient building layouts, and urban landscape planning is hindered by limited model generalization and inadequate ecological feedback [17,18]. Scholars have suggested a range of clever modeling and optimization techniques to deal with these problems [19,20]. Hierarchical convolutional structures are used in spatial feature extraction models to encode visual information; however, these models are not flexible enough to accommodate ecological feature constraints [21,22]. Although generative adversarial networks are frequently employed in the domains of spatial reconstruction and art creation, they do not yet have a coupling mechanism that takes ecological performance constraints into account when designing environmental art [23,24]. Although reinforcement learning-based adaptive design optimization techniques can dynamically modify design parameters, their training procedure depends on high-dimensional environmental data, which leads to slow convergence and inadequate ecological response accuracy [25,26]. Existing research still has shortcomings in multimodal data fusion, ecological constraint modeling, and generative feedback optimization, and an intelligent green design system that balances aesthetic expression and ecological performance has yet to be established [27,28]. To address these challenges, this paper proposes an intelligent ecological design modeling method based on a multimodal generative adversarial network. This method, centered on multimodal data fusion, first extracts high-dimensional features of environmental images and ecological parameters through a convolutional coding module, and then uses a semantic embedding layer to achieve a unified mapping of visual, energy, and structural information. Subsequently, an attention weighting mechanism is applied to the generator to strengthen the correlation between the spatial distribution and ecological constraints of different modal features. Simultaneously, a structural similarity metric and energy consumption prediction module are superimposed on the discriminator to achieve joint evaluation of both visual and ecological criteria of the authenticity and ecological performance of the design results. Finally, an adaptive feedback loop based on ecological performance indicators is constructed, and the network parameters are adjusted through gradient inversion, enabling the system to achieve ecologically optimal evolution of the design solution over multiple rounds of iterations. This research takes the intelligent generation model as its technical core, incorporates ecological constraints into the closed-loop feedback system of design generation, and provides a sustainable path for green environmental art design that integrates data-driven, ecological response and aesthetic generation, a methodology highly relevant for the sector to guide the creation of resource-efficient materials and environmentally responsive apparel designs.

II. INTELLIGENT ECO-DESIGN MODELING METHOD

MULTIMODAL DATA FUSION AND REPRESENTATION FOR ECOLOGICAL CONSTRAINTS

Multimodal feature fusion and semantic embedding modeling aim to achieve a unified representation of environmental visual information, ecological parameters, and spatial semantic features to support subsequent generation and discrimination. The model input consists of multi-source data: environmental visual data is derived from a scene image matrix; ecological parameter data consists of continuous features such as carbon emissions, energy density, and light and heat distribution; and spatial semantic data is input in the form of a structured encoding matrix. All inputs undergo normalization and dimension matching before entering the model. Minimum, maximum, and Z-score normalization are used to eliminate dimensional differences between modalities. A description of the spatial structure, ecological performance metrics, and a visual representation of the environment are among the model inputs. The visual image's local and global spatial texture and illumination distribution features are extracted by a convolutional encoder. Each layer of the five-layer convolutional network architecture has a stride of 1 and a kernel size of 3×3.

The semantic embedding stage captures the long-term relationships between ecological semantics and spatial structural features by modeling the sequential dependencies of the fused features using a bidirectional gated recurrent unit structure. A linear mapping layer transforms the encoded results into a single semantic embedding vector. After normalization, the fused features are projected into a single latent space. The general structure of this approach is shown in Figure 1. The vectors output from the semantic embedding layer are aligned with visual and ecological features through a cross-modal projection network. This network consists of a series of fully connected layers that learn to map discrete semantic tags to continuous visual texture feature spaces and ecological parameter spaces. For example, the embedding of material-related terms is associated with specific texture patterns extracted by the convolutional network through a weight matrix obtained during training, and simultaneously mapped to typical carbon emission coefficient ranges in a material database. This mapping relationship is not fixed in advance, but is jointly optimized and learned during end-to-end training by minimizing the overall reconstruction loss of the generated image and its corresponding semantic description, as well as the ecological constraint loss. The projected semantic features, visual features, and ecological parameter features are concatenated along the channel dimension and then input into the subsequent attention fusion module to achieve deep interaction between modalities in a unified representation space.

DESIGN OF ECOLOGICAL CONSTRAINT GENERATOR INTEGRATING ATTENTION MECHANISM

The core goal of the generator is to generate environmental design solutions that combine artistic expression with ecological constraints based on multimodal semantic embeddings. Its architecture employs an improved multimodal generative adversarial network (GAN) generative subnetwork, with feature reconstruction accuracy and ecological adaptability as dual optimization objectives. The generator consists of an input mapping layer, a feature fusion layer, a semantic control layer, and an ecological constraint output layer, forming an end-to-end generation structure from high-dimensional embeddings to design results. The input mapping layer receives the multimodal semantic feature vectors fused in Section "Multimodal Data Fusion and Representation for Ecological Constraints" and maps the embedding dimensions to a 256-dimensional latent feature space via a two-layer fully connected network. The LeakyReLU activation function is used to enhance gradient propagation stability [29,30].

To generate a semantic distribution weight map, the semantic control layer first calculates the autocorrelation matrix between semantic channels and performs element-wise multiplication with the visual feature map to achieve precise semantic constraint control. This layer introduces an ecological parameter mapping submodule, which converts ecological performance indicators such as photothermal response, energy consumption distribution, and carbon emission constraints into ecological constraint vectors via a multi-layer perceptron. These are then fused with visual semantic features during the channel attention phase to achieve feature-level ecological control. After semantic control, the deconvolution upsampling module restores the spatial resolution layer by layer, generating a design result consistent with the input image size [31,32].

The ecological constraint output layer supervises the ecological performance of the generated results by applying an ecological adaptation loss function. This loss function consists of a structural similarity loss, L_{SSIM}, and an ecological performance loss, L_{eco}. L_{SSIM} is used to maintain the rationality of the generated results at the spatial structure and texture level. L_{eco} calculates the ecological deviation based on the predicted carbon emission intensity, energy consumption gradient, and thermal response distribution. Its form is defined as the average squared Euclidean distance between the ecological target and the generated results in the parameter space [33,34]. The total loss function is expressed in a weighted form as:

$$L_{\mathrm{total}} = \lambda_1 L_{\mathrm{SSIM}} + \lambda_2 L_{\mathrm{eco}} \tag{1}$$

In formula (1), the weight coefficients λ_1 and λ_2 are set to 0.6 and 0.4 respectively to balance the optimization strength between visual quality and ecological constraints.

$$L_{\mathrm{eco}} = \frac{1}{N}\sum_{i=1}^{N}\left(\alpha\|C_i - \hat{C}_i\|_2^2 + \beta\|\nabla E_i\|_2^2 + \gamma\|T_i - \hat{T}_i\|_2^2\right) \tag{2}$$

where C_i and $\hat{C}_i$ denote the actual and predicted carbon emission intensity of the i-th region, with data sourced from life cycle assessment databases and material-specific emission factor calculations. The energy consumption gradient ∇E_i is derived from the difference in energy consumption between adjacent cells on a discretized spatial grid; the energy consumption value E_i is computed using a simplified building energy simulation model based on the geometric layout and material properties of the generated design. The difference between the thermal response distribution T_i and the target distribution $\hat{T}_i$ is quantified based on the finite element solution of the steady-state heat conduction equation under identical boundary conditions, where surface thermal properties are assigned according to the material thermal conductivity in the design. The coefficients α, β, γ balance the influence of different physical dimensions and are set to 0.5, 0.3, and 0.2, respectively, in the experiment. During the training phase, the Adam optimizer with a batch size of 32 is used for gradient update. The learning rate is fixed at 2×10^{-4}, and the generator and discriminator are updated alternately in a 1:1 ratio to maintain adversarial stability [35].

A random perturbation noise vector is applied in the generation phase to improve sample diversity and ecological stability. The generator creates an ecological adaptive evolution trend after each training cycle when the ecological parameter feedback module modifies the ecological constraint vector's weight based on the discriminator feedback. The final output is used as the input evaluation module of the ecological design scheme after being verified twice by structural similarity and energy consumption prediction.

The latent spatial randomness introduced by the noise vector aims to enhance the diversity of generated schemes. During training, feature fluctuations caused by noise that deviate from the ecological performance objective are suppressed by corresponding gradient updates. Therefore, the noise does not disrupt the stability of ecological indicators but serves as a means of design exploration within ecological constraints, helping to prevent the model from getting trapped in local optima.

DUAL DISCRIMINATOR MECHANISM COMBINING VISION AND ECOLOGICAL PERFORMANCE

To achieve a synergistic optimization between artistic expression and ecological rationality in generative design,

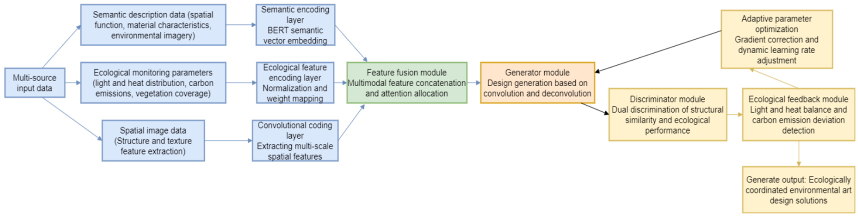

FIGURE 1. Overall framework structure

this paper constructs a dual discriminant mechanism that integrates visual perception and ecological performance assessment. This mechanism operates simultaneously in the image and performance domains, using a joint discriminant strategy to guide the generator to simultaneously improve visual realism and ecological compliance during training.

At the visual discrimination level, the discriminator uses structural similarity as a core metric and combines multi-scale convolutional feature maps to quantify the differences in brightness, contrast, and structural texture between generated images and real samples. By adaptively weighting visually salient regions (material transitions and spatial boundaries) during the SSIM feature extraction stage, the model further improves sensitivity to key regions and increases the discriminator's accuracy in differentiating between texture consistency and detailed structure.

To assess how well the generative solution performs on fundamental ecological metrics like carbon emissions, energy consumption distribution, and spatial efficiency, the discriminator incorporates a differentiable energy consumption prediction subnetwork at the ecological performance level. The generated image and its associated ecological parameters are mapped into an ecological performance score by this subnetwork using a multi-layer, fully connected architecture.

SSIM, as a core indicator for visual discrimination, ensures that the generated solution conforms to the basic physical structure and visual rationality of the real scene in terms of spatial layout and texture continuity, rather than strictly replicating the training samples. This constraint, combined with ecological performance discrimination, guides the generation process. The generator's innovative exploration capability mainly relies on noise injection in the latent space, the reweighting of multimodal features by the attention mechanism, and parameter search driven by ecological constraint loss. The weight coefficient $\lambda 1$ of SSIM loss is set to 0.6. This value has been verified in pre-experiments to strike a balance between maintaining structural rationality and allowing design flexibility, avoiding excessive punishment of structural differences that inhibits innovation in the ecological dimension.

The energy consumption prediction subnetwork is not a statically pre-trained model, but rather embedded within a generative adversarial training framework for end-to-end learning. This subnetwork employs a multilayer fully connected structure, with its weights updated via back-propagation along with the rest of the discriminator. During training, the generator's output design scheme and its corresponding ecological parameters serve as input, and the subnetwork outputs an ecological performance score. This score participates in the discriminator's loss calculation and drives gradient updates. Ecological performance data originates from real-time collected multimodal datasets, including dynamically monitored carbon emissions, energy consumption distribution, and photothermal response indicators. The subnetwork's parameters are iteratively optimized with each training iteration, adjusting its internal representation based on the actual distribution of ecological indicators to adapt to energy consumption patterns under different design scenarios. This design ensures that the ecological assessment module evolves synchronously with the generative adversarial process, avoiding the limitations imposed by a fixed prediction model on design optimization. By continuously comparing the deviation between predicted energy consumption and actual monitoring data, the system dynamically corrects the subnetwork's mapping relationship, achieving coupled evolution between the ecological feedback mechanism and the generation process.

ADAPTIVE FEEDBACK AND ITERATIVE OPTIMIZATION BASED ON ECOLOGICAL INDICATORS

The adaptive feedback mechanism in this study's multimodal generative network adjusts ecological constraints by modifying generator parameters through backpropagation of ecological performance indicators. It uses feedback from the discriminator's ecological branch, which evaluates the light/heat distribution and carbon emissions of generated designs. The deviation vector (ΔE) measures discrepancies in thermal response (ΔH) and carbon load (ΔC), and the performance gradient derived from this vector updates the generator's weights, optimizing the design towards ecological goals over time.

To avoid gradient oscillation and parameter convergence lag, the adaptive feedback module applies a correction strategy based on the momentum term. During each parameter update, the feedback module calculates the cosine similarity between the previous and current gradient directions. When the similarity exceeds a set threshold, the momentum coefficient decays exponentially to prevent over-updates. When the similarity falls below the threshold, the system determines that the current ecological performance adjustment direction is inconsistent with the historical optimization path, automatically increases the adaptive gain term of the learning rate, and redistributes the gradient weights to ensure sufficient flexibility and responsiveness during the optimization process. The dynamic learning rate adjustment rule can be expressed as:

$$\eta_t = \eta_0 \left(1 + \lambda \left(1 - \cos(G_{t-1}, G_t)\right)\right) \tag{3}$$

In formula (3), η_0 is the initial learning rate, λ is the adjustment coefficient, and G_t and G_{t-1} represent the ecological performance gradient vectors of the current and previous rounds, respectively. This mechanism achieves adaptive parameter adjustment by continuously monitoring gradient similarity, allowing the model to maintain sensitivity to ecological constraints at different stages. During the feedback optimization process, the ecological performance loss function is composed of the light and heat distribution loss L_H and the carbon emission constraint loss L_C, and a balance coefficient is applied to achieve dynamic coordination between the two indicators. The loss function is expressed as:

$$L_e = \alpha_H L_H + \alpha_C L_C \tag{4}$$

In formula (4), α_H and α_C are dynamically adjusted based on the convergence rate during the training phase. When the rate of decrease in L_H is significantly faster than that in L_C, the system automatically increases the weight of the carbon emission constraint to strengthen the low-carbon orientation, enabling the model to achieve balanced convergence at the ecological level. After each iteration, the updated generator parameters are re-input into the discriminator for the next round of performance evaluation, forming a closed-loop structure of "generation-discrimination-feedback-optimization."

III. EXPERIMENTAL DESIGN AND RESULT ANALYSIS

The experiment employed a multimodal dataset containing 1250 aligned samples. Each sample consisted of a spatial image, ecological parameter vectors, and semantic annotations. The image data covered 350 indoor ecological spaces and 900 urban public landscape scenes, with an original resolution of at least 2048×2048 pixels, which were then preprocessed to a uniform 256×256 input format. Ecological monitoring parameters included carbon emissions per unit area, regional energy density, illuminance distribution maps, and thermal imaging data. These parameters were obtained through a combination of on-site sensor recordings and EnergyPlus simulations, and a precise spatial correspondence was established between each image and the data using two-dimensional projected coordinates. Semantic description data annotated the materials, functional zones, and ecological design strategies within the scenes using structured labels. All modal data were associated with unique scene IDs, ensuring consistency in feature alignment during training. Table 1 shows the model training and experimental parameter configurations.

TABLE 1. Model training and experimental parameter configuration

Parameter Category	Parameter Name	Value or Setting	Description
Data Split	Training/Validation Ratio	8:2	Ensures feature distribution consistency
Learning Rate Strategy	Initial Learning Rate / Adjustment Strategy	2×10^{-4} / Cosine Annealing	Smooth convergence, avoids overfitting
Batch Size	-	32	Balances memory usage and convergence speed
Training Epochs	Epochs	50	Controls stable convergence of the model
Gradient Clipping	Max Norm	2	Prevents gradient explosion

Table 1 lists some key parameter settings used during model training. 32 samples were used per training session to balance video memory usage with convergence speed. A maximum norm of 2 was set to prevent gradient explosion and maintain training stability.

COMPARATIVE EXPERIMENTAL ANALYSIS

To evaluate the effectiveness of the proposed method, this study compares it with three benchmark models: a model based on a standard generative adversarial network structure but without an ecological constraint module, a non-AI ecological design method based on deterministic parameter optimization, and a model employing reinforcement learning for adaptive design. All comparison models used the same training and testing datasets. The comparison metrics cover core ecological parameters (carbon emission intensity, energy consumption distribution) and visual structure consistency scores.

Table 2 shows the average performance of each method on the same test set in terms of annual carbon emissions per unit area, comprehensive energy consumption, and ecological parameter consistency index. The proposed multimodal generative adversarial network (GAN) method outperforms both standard GANs and non-AI optimization methods in both carbon emission control and energy consumption indicators. Compared with reinforcement learning methods, this method achieves higher values in ecological parameter consistency, while maintaining comparable or better levels in visual structure coordination scores. Standard GANs, lacking explicit ecological constraints, exhibit significant fluctuations in ecological indicators and have significantly higher average values. Although non-AI optimization methods are interpretable in terms of ecological indicators, their generated schemes generally have lower visual structure coordination scores, reflecting their limitations in the aesthetic dimension. Reinforcement learning methods have advantages in dynamic adjustment, but their ecological consistency is still limited by the incomplete exploration of the high-dimensional state space during training.

TABLE 2. Comparative Results of Different Models on Ecological and Visual Performance Indicators (Mean ± Standard Deviation)

Model Method	Carbon Emissions (kg CO_2)	Comprehensive Energy Consumption (kWh)	Ecological Parameter Consistency	Visual Coordination Comprehensive Score
Proposed Method	50.2 ± 1.5	120.5 ± 2.1	0.89 ± 0.02	0.83 ± 0.03
Standard GAN (without ecological constraints)	68.7 ± 5.3	145.2 ± 6.8	0.72 ± 0.05	0.81 ± 0.04
Non-AI Parametric Optimization Method	55.1 ± 1.8	128.3 ± 3.2	0.85 ± 0.03	0.65 ± 0.06
Reinforcement Learning Adaptive Method	52.4 ± 2.2	124.1 ± 3.5	0.82 ± 0.04	0.80 ± 0.04

ECOLOGICAL ADAPTABILITY ANALYSIS

This experiment focused on the ecological adaptability of generative solutions for green environmental art design. By constructing a multimodal generative adversarial network model and jointly training it on spatial imagery, ecological monitoring parameters, and semantic description data, it explored the responsiveness of generative design to carbon emission constraints, energy consumption distribution, and consistency with ecological parameters.

As can be seen in Figure 2, the generative design maintains carbon emissions within the range of 48 kg CO_2 to 53 kg CO_2, energy consumption values range from 118 kWh to 123 kWh, and the ecological parameter consistency index fluctuates between 0.87 and 0.91. These minor fluctuations in carbon emissions and energy consumption are primarily due to the multimodal features' balance between spatial structure and ecological constraints during convolutional encoding and attention fusion. The generator successfully captures the relationship between ecological semantics and environmental constraints during training, as evidenced by the ecological parameter consistency above 0.87. This allows the generative design to sustain a comparatively stable ecological response under various sample conditions. The energy consumption figures above represent the total energy consumption estimates for a single standardized design unit over the entire annual simulation period. The spatial scale of each design unit is fixed at 100 square meters. The simulation uses typical meteorological year data, with a one-year time base. The calculation of carbon emission intensity indicators uses the same spatiotemporal system boundary as the energy consumption range, ensuring consistency in evaluation standards.

VISUAL AND STRUCTURAL COORDINATION ANALYSIS

This part examines the model's structural coordination performance and visual composition under semantic constraints in order to confirm that the generative design strikes a balance between spatial organization and artistic expression. This experiment quantitatively assessed the formal balance, spatial hierarchy, and artistic composition of ten sample groups using a generative model trained with multimodal feature alignment. A quantitative depiction of the design's aesthetic coherence was accomplished by using a structure-aware encoder to record hierarchical variations between design layouts and a visual feature mapping module to extract morphological continuity features. The artistic composition score is calculated by evaluating the symmetry distribution, local contrast variance, and distribution entropy of the visual focal region in the generated image. The spatial hierarchy score is based on the depth map generated by the depth estimation network, quantified by combining the discernibility of object occlusion relationships in the scene and the geometric consistency of perspective structures. The formal balance score is calculated based on the visual weight distribution after image segmentation, evaluated by the weighted distance between the centroid of each region and the image center, as well as the balance of color and texture distribution. All these metrics are automatically extracted and calculated using a pre-trained image analysis model and structural tensor algorithm, and all scores are normalized and mapped to the 0-1 range. This experiment was conducted under uniform lighting conditions and style semantic constraints to eliminate the influence of environmental variables on structural representation. The final visualization results, shown in Figure 3, show the overall score distribution and the local coordination distribution.

As shown in Figure 3, the generative design scores for Artistic Composition remain within a range of 0.83 to 0.89, Spatial Hierarchy between 0.80 and 0.84, and Formal Balance between 0.77 and 0.81. Fluctuations in the Artistic Composition score are primarily due to differences in how the convolutional coder processes edge and texture information when capturing spatial features. The high stability of the Spatial Hierarchy score stems from the attention mechanism's balanced weighting of multi-scale features. The slightly lower Formal Balance score stems from the semantic constraints failing to fully balance local details with overall layout in the complex structures of some samples. As can be seen in the right figure, structural coordination is concentrated between 0.75 and 0.95. The differences in coordination in local areas are due to differences in the processing of local details among different samples.

ECOLOGICAL PERFORMANCE AND DYNAMIC RESPONSE ANALYSIS

The performance of generative design solutions in terms of solar-thermal response and energy consumption distribution was the main focus of this experiment. In order to monitor the shifting patterns of various energy consumption components and the equilibrium of solar-thermal distribution, it iteratively optimized training round samples using a multimodal generative adversarial network model. It measured and documented the energy consumption of the equipment, air conditioning, and lighting during the experiment. In order to maximize solar-thermal balance, it also uses an adaptive feedback mechanism to modify the generator's settings. In order to confirm the model's responsiveness in green design optimization, it also tracked the dynamic changes in energy consumption components during the training phase,

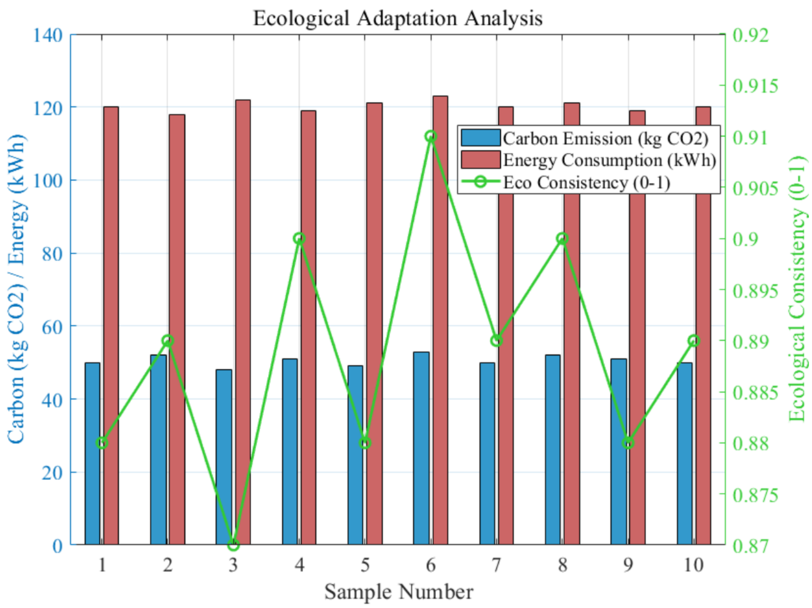

FIGURE 2. Ecological adaptability analysis

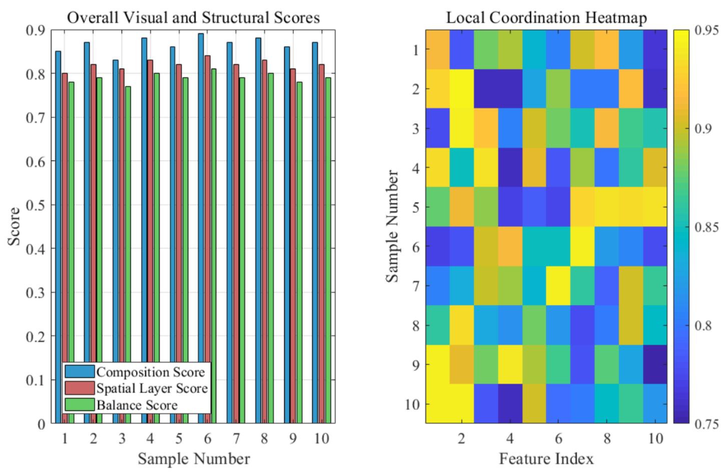

FIGURE 3. Visual and structural coordination analysis

as illustrated in Figure 4.

Figure 4 shows that during the later training phases, the energy consumption of the equipment stayed at about 20 kWh, the lighting at about 30 kWh, and the air conditioning at about 48 kWh. The solar-thermal balance degree steadily stabilized between 0.74 and 0.98. Convolutional coding and attention mechanisms that balance spatial features and maintain a reasonable energy distribution across functional areas are largely responsible for the minimal fluctuations in lighting and equipment energy consumption. Slight fluctuations in air conditioning energy consumption are due to the model optimizing global solar-thermal balance while adjusting local heat loads. The degree of solar-thermal balance gradually improved and stabilized in the high range, reflecting the adaptive feedback mechanism's effective parameter correction during closed-loop iterations, ensuring that the generated design maintains a balance between energy efficiency allocation and solar-thermal control, thus validating the model's responsiveness in green design optimization.

IV. CONCLUSION

This paper addresses the issues of insufficient ecological feedback and weakened sustainability constraints in green environmental art design, and proposes an intelligent ecological design modeling method based on a multimodal generative adversarial network, a framework that offers significant potential for guiding the sector's transition toward aesthetically superior and ecologically constrained product development. The model achieves a unified representation of multimodal features through convolutional feature extraction, attention fusion, and semantic embed-

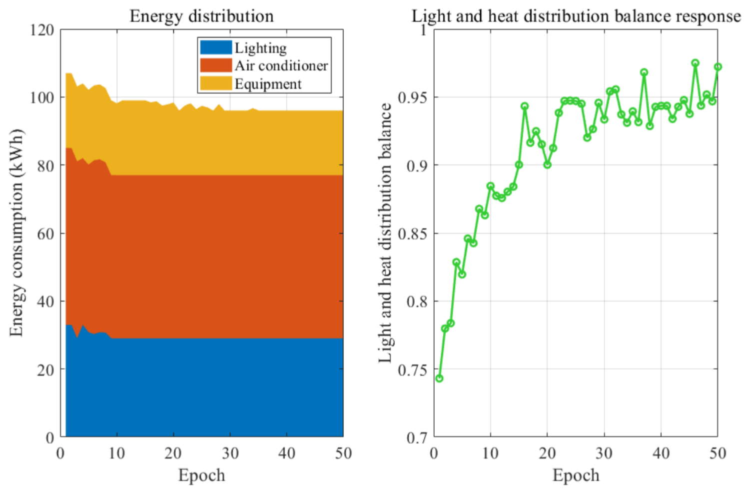

FIGURE 4. Analysis of ecological performance and energy consumption response

ding optimization. It also applies an ecological constraint mechanism and structural similarity supervision into the generator, and combines energy consumption prediction with carbon emission assessment to construct a two-tier discrimination mechanism system of visual and ecological performance. Experimental results show that this method exhibits high stability and ecological adaptability in terms of carbon emission control, energy consumption distribution, and consistency of ecological parameters, verifying the model's ability to synergistically optimize formal aesthetics and ecological performance. This method achieves closed-loop optimization from multimodal perception to ecological generation, providing a new technical path and methodological support for the sustainable application of intelligent auxiliary technology in green environmental art design, and is of great significance for promoting the intelligent and scientific development of ecological aesthetic design, with direct transferability to the industry for guiding automated, sustainable material selection and eco-conscious product styling.

AUTHOR CONTRIBUTIONS

Ruihua Fu designed, collected and analyzed the data, and drafted the manuscript. Ruihua Fu conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Ruihua Fu 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] K. Wang, "Application of green pollution-free materials in environmental art design," Science of Advanced Materials, vol. 15, no. 1, pp. 41-49, 2023, doi: 10.1166/sam.2023.4427.

[2] Y. Sang, "Application in environmental art design practice based on a fuzzy evaluation system," Scientific Reports, vol. 14, no. 1, Art. no. 12441, 2024, doi: 10.1038/s41598-024-62477-2.

[3] H. Liu, G. Song, and L. Yan, "Research on the application mode of green environment design under the background of artificial intelligence," Complexity, vol. 2021, no. 1, Art. no. 8914304, 2021, doi: 10.1155/2021/8914304.

[4] A. E. S. Nosseir, A. Cervone, and A. Pasini, "Review of state-of-the-art green monopropellants: For propulsion systems analysts and designers," Aerospace, vol. 8, no. 1, pp. 20, 2021, doi: 10.3390/aerospace8010020.

[5] W. Fan, B. Yan, Q. Bao, Y. Zhao, and J. Zhou, "Green evaluation for building interior decoration based on BIM-BN technology," Buildings, vol. 13, no. 3, pp. 744, 2023, doi: 10.3390/buildings13030744.

[6] T. Semeraro, A. Scarano, R. Buccolieri, A. Santino, and E. Aarrevaara, "Planning of urban green spaces: An ecological perspective on human benefits," Land, vol. 10, no. 2, pp. 105, 2021, doi: 10.3390/land10020105.

[7] S. S. Mishra and D. Sakharle, "Design of green roof for an energy efficient university campus," Materials Today: Proceedings, vol. 65, pp. 1546-1550, 2022, doi: 10.1016/j.matpr.2022.04.505.

[8] H. Hussein, L. K. Jhung, and H. Zhang, "Green Pause in a City: Design elements of a pocket park in Kuala Lumpur," Environment-Behaviour Proceedings Journal, vol. 7, no. 19, pp. 191-200, 2022, doi: 10.21834/ebpj.v7i19.3206.

[9] A. Balachandran, H. Parayilkalapurackal, S. Rajpoot, and S. Lone, "Bioinspired green fabricating design of multidimensional surfaces for atmospheric water harvesting," ACS Applied Bio Materials, vol. 6, no. 1, pp. 44-63, 2022, doi: 10.1021/acsabm.2c00804.

[10] G. Caprari, G. Castelli, M. Montuori, M. Camardelli, and R. Malvezzi, "Digital twin for urban planning in the green deal era: A state of the art and future perspectives," Sustainability, vol. 14, no. 10, pp. 6263, 2022, doi: 10.3390/su14106263.

[11] J. Yu, Z. Wang, L. Hu, Q. Chen, J. Yang, and Y. Zhou, "Research on the Assistance and Countermeasures of Digital Empowerment in the Environment of Carbon Emission Reduction," Journal of Big Data and Computing (ISSN: 2959-0590), vol. 2, no. 2, pp. 83, 2024, doi: 10.62517/jbdc.202401212.

[12] T. M. Mata, G. M. Oliveira, H. Monteiro, G. V. Silva, N. S. Caetano, and A. A. Martins, "Indoor air quality improvement using nature-based solutions: design proposals to greener cities," International journal of environmental research and public health, vol. 18, no. 16, pp. 8472, 2021, doi: 10.3390/ijerph18168472.

[13] L. Liu, "Scene Classification in the Environmental Art Design by Using the Lightweight Deep Learning Model under the Background of Big Data," Computational Intelligence and Neuroscience, vol. 2022, no. 1, Art. no. 9066648, 2022, doi: 10.1155/2022/9066648.

[14] J. Hutson, J. Lively, and J. Weber, "In the studio with virtual reality: Student perceptions and potential integrations into art and design curriculum," art, design & communication in higher education, vol. 21, no. 2, pp. 235-256, 2022, doi: 10.1386/adch_00056_1.

[15] Y. Lei, Z. Su, X. He, and C. Cheng, "Immersive virtual reality application for intelligent manufacturing: Applications and art design," Matematical Biosciences and Engineering, vol. 20, no. 3, pp. 4353-4387, 2023, doi: 10.3934/mbe.2023202.

[16] X. Chen, "Environmental landscape design and planning system based on computer vision and deep learning," Journal of Intelligent Systems, vol. 32, no. 1, Art. no. 20220092, 2023, doi: 10.1515/jisys-2022-0092.

[17] M. Zahiri, A. Sohrabi S, and S. Dehghan, "Increasing Energy Efficiency During Design Case Study: Building and Green Space of the Museum of Visual Arts," International Journal of Advanced Multidisciplinary Research and Studies, vol. 3, no. 6, pp. 1567-1572, 2023, doi: 10.62225/2583049X.2023.3.6.3917.

[18] J. Ma, T. Inoue, Q. Fang, K. Li, and M. Li, "A Study on Optimal Opening Configuration for Art Museum Exhibition Space Considering Daylight Performance, Indoor Thermal Comfort, and Energy Consumption," Sustainability, vol. 15, no. 23, Art. no. 16431, 2023, doi: 10.3390/su152316431.

[19] D. Auge, J. Hille, E. Mueller, and A. Knoll, "A survey of encoding techniques for signal processing in spiking neural networks," Neural Processing Letters, vol. 53, no. 6, pp. 4693-4710, 2021, doi: 10.1007/s11063-021-10562-2.

[20] C. Wang, H. Yan, W. Huang, W. Sheng, Y. Wang, Y. S. Fan, et al., "Neural encoding with unsupervised spiking convolutional neural network," Communications Biology, vol. 6, no. 1, pp. 880, 2023, doi: 10.1038/s42003-023-05257-4.

[21] G. A. Vargas-Hakim, E. Mezura-Montes, and H. G. Acosta-Mesa, "A review on convolutional neural network encodings for neuroevolution," IEEE Transactions on Evolutionary Computation, vol. 26, no. 1, pp. 12-27, 2021, doi: 10.1109/TEVC.2021.3088631.

[22] G. W. Lindsay, "Convolutional neural networks as a model of the visual system: Past, present, and future," Journal of cognitive neuroscience, vol. 33, no. 10, pp. 2017-2031, 2021, doi: 10.1162/jocn_a_01544.

[23] P. Kumar and V. Gupta, "Restoration of damaged artworks based on a generative adversarial network," Multimedia Tools and Applications, vol. 82, no. 26, pp. 40967-40985, 2023, doi: 10.1007/s11042-023-15222-2.

[24] R. Garozzo, C. Santagati, C. Spampinato, and G. Vecchio, "Knowledge-based generative adversarial networks for scene understanding in Cultural Heritage," Journal of Archaeological Science: Reports, vol. 35, Art. no. 102736, 2021, doi: 10.1016/j.jasrep.2020.102736.

[25] X. Chen and L. Chen, "Exploration of adaptive environment design strategy based on reinforcement learning in cad environment," CAD Computer-Aided Design, vol. 21, pp. 175-190, 2024, doi: 10.14733/cadaps.2024.S23.175-190.

[26] Y. Tian, X. Li, H. Ma, X. Zhang, K. C. Tan, and Y. Jin, "Deep reinforcement learning based adaptive operator selection for evolutionary multi-objective optimization," IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 7, no. 4, pp. 1051-1064, 2022, doi: 10.1109/TETCI.2022.3146882.

[27] N. McCartney and J. Tynan, "Fashioning contemporary art: a new interdisciplinary aesthetics in art-design collaborations," Journal of Visual Art Practice, vol. 20, no. 1-2, pp. 143-162, 2021, doi: 10.1080/14702029.2021.1940454.

[28] S. Sina A and J. Wu, "The effects of retail environmental design elements in virtual reality (VR) fashion stores," The International Review of Retail, Distribution and Consumer Research, vol. 33, no. 1, pp. 1-22, 2023, doi: 10.1080/09593969.2022.2049852.

[29] C. Cucuzzella, "Analysing eco-art installations for their value in affecting change," International Journal of education through art, vol. 19, no. 1, pp. 43-58, 2023, doi: 10.1386/eta_00117_1.

[30] M. Taheri and E. Afzaltousi, "A phenomenological reflection on the aesthetic experience in the contemporary en-vironmental art (Case Study: The Weather Project)," Theoretical Principles of Visual Arts, vol. 8, no. 1, pp. 146-160, 2023.

[31] A. Kadry and E. Hussien, "Applications of Virtual Reality Technologies in the Field of Design and Arts," International Journal of Multidisciplinary Studies in Art and Technology, vol. 6, no. 2, pp. 44-70, 2023, doi: 10.21608/ijmsat.2024.274902.1022.

[32] H. He, M. Wu, and J. Gyergyak, "Intervention and renewal-Interpretation of installation art in urban public space," Pollack Periodica, vol. 16, no. 3, pp. 139-145, 2021, doi: 10.1556/606.2021.00362.

[33] A. S. Arulraj and S. Don Bosco, "Connecting art and development: studying the role of city art on bridges in forwarding urban sustainability," Local Environment, vol. 29, no. 7, pp. 919-931, 2024, doi: 10.1080/13549839.2024.2319639.

[34] T. Gao and L. Kuang, "Cognitive loading and knowledge hiding in art design education: Cognitive engagement as mediator and supervisor support as moderator," Frontiers in Psychology, vol. 13, Art. no. 837374, 2022, doi: 10.3389/fpsyg.2022.837374.

[35] G. Li, Y. Huang, and W. Ye, "Research on the construction of intelligent-aided design platform for green buildings," Journal of Computational Methods in Sciences and Engineering, vol. 25, no. 4, pp. 2907-2917, 2025, doi: 10.1177/14727978251314565.