RESEARCH ARTICLE
Published: 12 April 2026
doi: 10.15627/jd.2026.11

Volume 13 | Issue 1 | 196-213

OPEN ACCESS

Journal of Daylighting

Journal homepage: www.solarlits.com/jd

Adopting Generative Artificial Intelligence in Daylighting Design Process

Mario Wagdy Berty, Ayman Assem, Fatma Fathy*

Department of Architecture, Faculty of Engineering, Ain Shams University, Cairo, 11517, Egypt

Received 10 February 2026; Revised 11 March 2026; Accepted 16 March 2026; Published online 12 April 2026

Citation: *Mario Wagdy Berty, Ayman Assem, Fatma Fathy, Adopting Generative Artificial Intelligence in Daylighting Design Process, Journal of Daylighting, 13:1 (2026) 196-213. doi: 10.15627/jd.2026.11*

ABSTRACT

Generative Artificial Intelligence (GenAI) is increasingly adopted in architectural design and widely investigated in the research field. However, several limitations hinder its broader integration into design processes. Among these, the lack of validation beyond aesthetic assessment remains a research gap, especially regarding the environmental performance of Artificial Intelligence (AI)-generated designs. While current research primarily applies GenAI to exterior and urban-scale environmental assessments, this study focuses on GenAI-driven active control of interior daylighting. An experimental study explored the fine-tuning of a Low-Rank Adaptation (LoRA) model to generate Daylight Autonomy (DA) heatmaps with controlled window-to-wall ratio (WWR). Experiment scope was bounded by the control of WWR as the sole variable, while space type, geographic location, dimensions, orientation, finishing materials, and other design parameters were kept constant. The proposed workflow enabled the generation of interior designs that achieve target Spatial Daylight Autonomy (sDA) ranges, classified as "High," "Moderate," or "Low." As a result, Structural Similarity Index Measure (SSIM) confirmed strong similarity between AI-generated and simulation-based DA maps, with average values of 0.851, 0.740, and 0.694 for the "High," "Moderate," and "Low" categories, respectively. In addition, a one-sample t-test was conducted on 120 generated samples—40 per category—verifying that generated designs fit target sDA ranges beyond random chance (33.33%). 83% of generated designs using "High" prompt keyword fit target range ($t(39) = 8.086$), while fitting rates reached 68% for "Moderate" ($t(39) = 4.560$) and 75% for "Low" category ($t(39) = 6.014$). Overall, the findings confirm that the proposed GenAI-driven workflow can reliably control daylighting performance, supporting the transition of GenAI from a purely aesthetic visualization tool to a validated, performance-oriented method for early-stage architectural design.

Keywords: Generative artificial intelligence, Daylighting, Interior design, Low-rank adaptation

1. INTRODUCTION

Architecture, engineering, and construction (AEC) industry, in parallel with many other fields, is greatly evolving nowadays, in accordance with the fourth industrial revolution (Industry 4.0). It is now hosting new cutting-edge technologies, such as 3D printing, Internet of Things (IoT), Building Information Modeling (BIM), Blockchain, cloud computing, big data and data analytics, augmented reality (AR), digital twins, and predictive modeling. The massive evolution of the AEC industry has beneficial applications, allowing the generation of innovative architectural and structural designs, increased construction and operational safety, reduction of embodied and operational energy requirements, saving in construction and operational costs, increased construction speeds and enhancement in sustainability [1].

One of these new technologies that has increasingly emerged in several fields through the last decade is AI. It is defined as the technology of increasing intelligence of the machine in a way it can interpret, learn, communicate, operate, solve problems, and provide future plans [2].

AI was formally introduced in 1956 [3], and has been applied in the AEC industry since the 1970s, with significant growth from the late 1990s onward [4]. As AI evolved, subfields such as Machine Learning (ML) emerged, enabling algorithms to analyze data, recognize patterns, and make predictions.

*Corresponding author.
mario.wagdy@eng.asu.edu.eg (M. W. BERTY)
ayman.assem@eng.asu.edu.eg (A. ASSEM)
fatma.fathy@eng.asu.edu.eg (F. FATHY)

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

NOMENCLATURE

Abbreviation	Definition
AEC	*Architecture, Engineering and Construction*
AI	*Artificial Intelligence*
AR	*Augmented Reality*
BIM	*Building Information Modeling*
CFD	*Computational Fluid Dynamics*
cGANs	*Conditional Generative Adversarial Networks*
CI	*Confidence Interval*
CNNs	*Convolutional Neural Networks*
DA	*Daylight Autonomy*
DL	*Deep Learning*
GANs	*Generative Adversarial Networks*
GenAI	*Generative Artificial Intelligence*
GUI	*Graphical User Interface*
IES	*Illuminating Engineering Society*
IoT	*Internet of Things*
LEED	*Leadership in Energy and Environmental Design*
LoRA	*Low-Rank Adaptation*
LRV	*Light Reflectance Value*
ML	*Machine Learning*
MSE	*Mean Squared Error*
NSGA-II	*Non-Dominated Sorting Genetic Algorithm II*
Parti	*Pathways Autoregressive Text-to-Image*
PLW	*Pedestrian-Level Wind*
PSNR	*Peak Signal-to-Noise Ratio*
PSQA	*Patient-Specific Quality Assurance*
sDA	*Spatial Daylight Autonomy*
SD	*Stable Diffusion*
SPSS	*Statistical Package for the Social Sciences*
SSIM	*Structural Similarity Index Measure*
UDI	*Useful Daylight Illuminance*
UI	*User Interface*
UTCI	*Universal Thermal Climate Index*
WWR	*Window-to-Wall Ratio*

While Deep Learning (DL), a subfield of ML, provides models with the ability to learn through the process where the system rewards all appropriate behaviors and punishes inappropriate ones [5].

More recently, DL has led to the development of GenAI, which can produce diverse content including text, images, and 3D models. In architecture, GenAI tools are increasingly used to support concept generation, massing studies, floor plans, façades, and other design tasks, either as a collaborative assistant or even a primary author [6,7].

Generative Adversarial Networks (GANs), Diffusion-Based models, Convolutional Neural Networks (CNNs), Pathways Autoregressive Text-to-Image (Parti) models, and other DL models form the basis for one of the most recent GenAI features: text-to-image generation. These models generate images from textual prompts and are widely implemented in tools such as DALL-E, Midjourney, and Stable Diffusion [5].

Most widely used text-to-image generators rely on diffusion technique, where an image of a reference object and a text description are given to the AI model. Then, model progressively adds noise to images during training and then reverses the process to produce new visuals. It generates a new image that aligns with the given prompt by learning patterns from large training datasets [8].

To improve GenAI application in specialized architectural tasks, LoRA fine-tuning is employed to adapt large pre-trained diffusion models to specific tasks. Rather than retraining all model parameters, LoRA inserts low-rank matrices into existing weight layers, enabling targeted task customization, with reduced computational requirements [9]. Additionally, ControlNet extends diffusion models by enabling greater control over outputs through conditional inputs such as depth maps, edges, or segmentation data [10].

Based on these technological foundations, text-to-image generation and other related techniques, such as image-to-image and text-to-3D, have increasingly been introduced into research field. Ongoing research aims to explore their broader application across different stages of architectural design, and to enhance the accuracy of generated outputs. Consequently, diffusion models, LoRA fine-tuning, training workflows, prompt engineering, and ControlNet configurations remain under continuous development.

As highlighted in recent research, GenAI tools applied in architecture mainly focus on producing fast and visually attractive design concepts. They prioritize aesthetic quality over functional, environmental, and structural performance [11].

Although text-to-image models are very effective during the early conceptual stage and generate persuasive images, their outputs usually lack built-in validation and do not ensure structural realism or performance-based intelligence [12]. This strong emphasis on visual appeal creates a plausibility gap, where designs that look realistic are not technically resolved or constructively feasible. Such generated designs may not be fully understood or properly assessed based on appearance alone. As a result, architects must rely on subsequent engineering analysis to confirm whether the proposed concepts are functional and buildable [13]. Moreover, recent studies show that AI-generated spatial layouts often fail to respond properly to environmental factors such as solar orientation and daylight performance, indicating a gap between generative representation and environmental parameters [12].

Among the validation parameters that extend beyond purely aesthetic assessment, environmental performance of AI-generated designs represents a significant research gap. Current studies have begun applying GenAI for environmental assessment across multiple scales, including architectural massing, façade design, and urban-scale analysis. However, limited research investigated not only the assessment but also the control of environmental

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

parameters within AI outputs. In addition, the validation of interior environments generated using AI remains underexplored. Accordingly, this research focuses on the active environmental control of AI-generated interior designs, with a strong emphasis on one of the most critical environmental factors in interior spaces: daylighting performance.

Daylighting was selected as the focal point for this study not only for its role in reducing energy consumption, but also for its significant human-centered benefits, enhancing occupants' psychological and physiological well-being. Exposure to natural light has been linked to improved mood, reduced fatigue and eyestrain, enhanced morale, and a stronger connection to the outdoor environment. Therefore, daylighting is essential in interior spaces, as it directly influences comfort, health, and the overall quality of the built environment [14].

Several metrics are used to assess daylighting performance, such as DA, sDA, and Useful Daylight Illuminance (UDI). DA at each measurement sensor represents the percentage of annual occupied hours when the target illuminance at that point is achieved using daylight alone [15,16]. While sDA assesses the effectiveness of daylight for the entire interior space, throughout the year. It measures the percentage of the occupied floor area that achieves a minimum illuminance level for a defined portion of the annual occupied hours. The standard threshold is 300 lux for at least 50% of the occupied hours (sDA300,50%), as defined by the Illuminating Engineering Society (IES) [17]. UDI extends this approach by evaluating the percentage of occupied hours in which illuminance levels fall within predefined ranges, distinguishing between insufficient (<100 lux), useful (100–2000 lux), and excessive (>2000 lux) daylight conditions [18,19]. Recent refinements define the useful range from 100 to 3,000 lux, subdivided into supplementary (100–300 lux), autonomous (300–3,000 lux) levels, and values above 3,000 lux which are considered excessive and cause potential visual discomfort [20].

Among these daylight performance metrics, this research focuses on sDA of interior spaces through the measurement of DA at each analysis sensor. An experimental study was applied to a representative office room, where a GenAI-aided design workflow is employed to generate multiple design iterations, each intended to meet targeted sDA values.

Experiment narrowed its scope to a single interior design parameter that affects sDA: WWR. It is the target variable to be modified using GenAI to achieve desired values. Space type, location, finishing materials, façade orientation, and other effective elements were kept fixed to be able to clearly measure the effect of WWR modification on sDA. This was implemented through the fine-tuning of a LoRA model to generate DA heatmaps with controlled WWR, aligned with target sDA ranges.

To conclude, this study aims to develop and assess a GenAI-driven workflow for generating interior design alternatives with sDA values aligned with target ranges. The workflow utilized LoRA fine-tuning as a control mechanism for modifying WWR as a key variable affecting sDA and then validated the generated outputs against simulation-based daylighting results. Its overall effectiveness was evaluated through visual and statistical analyses, with the goal of demonstrating the workflow's applicability as a performance-based support tool in early-stage interior design.

To achieve these objectives, the paper first reviews the relevant literature on GenAI applications in architectural design, highlighting the critical gap in performance-based daylighting validation. It then presents the research methodology, detailing the experimental workflow: base case modeling, dataset preparation, LoRA fine-tuning process, and validation procedures. This is followed by the presentation and discussion of the results, including visual inspection, image-based analysis, annual daylight simulation, and statistical evaluation. The paper then demonstrates the application of the validated workflow within an interior design process to generate iterations targeting different sDA ranges. Finally, it concludes by summarizing the main findings, highlighting the study's contributions and limitations, and outlining directions for future research.

2. LITERATURE REVIEW

2.1. Plausibility gap in AI-generated architectural designs

As previously stated, the validation of AI-generated architectural designs beyond aesthetic qualities remains underexplored, as ongoing research focuses primarily on visually pleasing outputs. Cervantes and Morales [13] argued that GenAI produces images that may look highly convincing and realistic, but they may not be structurally, tectonically, or environmentally feasible. In other words, AI-generated images act more like visual predictions than true representations of a buildable design. Therefore, they require further validation to turn visual plausibility into an accountable architectural system.

The gap between aesthetic appearance and environmental functionality is clearly evident in recent empirical applications. For instance, Çelik [12] investigated the daylighting performance of sustainable housing plans generated by text-to-image diffusion models across five distinct climate zones. By reconstructing AI-generated plans in AutoCAD and conducting climate-based daylight simulations in Velux Daylight Visualizer, the study revealed that these generative models heavily prioritize spatial and aesthetic representation over environmental logic. The AI models mostly failed to integrate critical parameters like solar orientation or seasonal daylight modulation. The study concluded that, although AI tools show better performance in schematic visualization, their outputs remain environmentally inadequate without validation.

Hu et al. [9] also noted that despite the notable advancements in the automation of floor plans generation, the assessment of the design performance remains a major challenge, especially regarding natural lighting, ventilation and energy efficiency.

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

To address this gap, recent research starts investigating the application of GenAI in performance-oriented architectural processes. Several studies focus on AI-aided environmental assessment and monitoring, while others apply GenAI to performance-related data prediction, alongside a limited number of attempts to actively control the environmental performance of GenAI outputs.

2.2. Performance prediction and assessment models

Mokhtar et al. [21] trained a cGAN model on datasets generated from computational fluid dynamics (CFD) simulations of created urban geometries. This model helped approximate pedestrian wind flow patterns around buildings with similar accuracy to CFD simulations, but at a much faster and more efficient rate.

Another study conducted by Kastner and Dogan [22] introduced a GAN-based framework as a fast surrogate for CFD simulations in urban design. Authors trained a Pix2Pix GAN model on 564 CFD-generated urban geometries, enabling instant wind flow predictions directly within Rhino and Grasshopper. The model achieved SSIM of 0.75–0.97, producing results within four seconds compared to several hours for CFD. SSIM is a metric used for image comparison that provides a comprehensive assessment of pixel-level accuracy, perceptual quality, and structural consistency between generated and reference images [9]. While the model was less accurate in complex regions, it provided sufficient precision for early design feedback, facilitating real-time performance-driven urban planning.

Furthermore, Huang et al. [23] presented an automated environmental performance-driven framework, on urban scale, that employed a Pix2Pix GAN model, to accelerate the traditional time-consuming simulation process. The GAN was trained on simulated datasets to predict Pedestrian-Level Wind (PLW), annual cumulative solar radiation, and the Universal Thermal Climate Index (UTCI) in real time. Integrated with a Non-Dominated Sorting Genetic Algorithm II (NSGA-II), this framework enabled the optimization of urban block morphology to fit the previously mentioned environmental parameters. Results demonstrated high prediction accuracies with coefficients of determination (R^2) of 0.70 for PLW, 0.86 for radiation, and 0.80 for UTCI, accelerating simulations by 120–240 times compared to conventional numerical methods.

Previously discussed studies introduced promising GenAI-driven tools for environmental prediction, focusing mainly on wind flow patterns as well as other performance-related parameters. These approaches demonstrated reliable results in shorter time compared to conventional simulation methods. However, they lacked active control mechanisms that can adapt GenAI outputs to meet specific environmental conditions. Regarding the reduced accuracy in complex regions reported by Kastner and Dogan [22], this limitation may be addressed in relevant studies by limiting the research scope and bounding the experimental workflow to a single or a small set of analyzed parameters, under specific conditions. This bounded setup reduces geometric and environmental variability, allowing a clearer evaluation of how effectively the GenAI workflow can control performance outcomes before extending to more complex, multi-variable cases.

Other studies focused on the application of GenAI for daylighting performance prediction and assessment. For instance, He et al. [24] introduced proxy models, developed using a CNN and a GAN, specifically Pix2Pix. The aim was to predict daylighting performance in general building floorplans, while reducing the time and complexity of traditional simulations. Their CNN model was able to accurately predict daylight distribution uniformity in spaces, with a high accuracy score (R^2 = 0.959), while the GAN generated lighting visuals very close to simulation results, with SSIM reaching 0.90.

Hu et al. [9] introduced one of the most recent applications of Gen AI, proposing a two-stage GenAI-driven workflow. This workflow integrated a fine-tuned diffusion model using LoRA, with Conditional Generative Adversarial Networks (cGANs) to automate the generation of floor plans, followed by the evaluation of their daylighting performance. cGAN is a neural network that generates images depending on input conditions, allowing more targeted outputs compared to standard GANs. In the first stage of the study, a LoRA was trained on a dataset of parametrically generated residential floor plans created in Grasshopper, labeled by room types and topological relationships. This enabled the automated generation of diverse and realistic floor plans. Then, the study employed Pix2PixHD, a cGAN model designed for high-resolution image synthesis, to assess daylighting performance. The model was trained on paired datasets of floor plans and their corresponding daylight simulations, generated using the Ladybug and Honeybee plugins. It successfully learned to predict sDA directly from floor plan images, achieving a deviation of less than 5% from simulation results and an average SSIM of 0.98, indicating a significant similarity between the predicted and simulated DA heatmaps.

However, these studies primarily frame GenAI as a prediction and assessment tool, rather than a control-oriented design system that can actively adapt outputs to predefined daylighting targets. This gap is particularly evident in the approach proposed by Hu et al. [9], as it relied on a decoupled, sequential process in which generation and evaluation happened in separate stages. As a result, cGAN functioned as a post-generation assessment tool. While it accurately evaluated an already generated plan, it lacked the capacity to actively optimize initial generation phase towards targeted daylighting parameters.

2.3. AI-aided active environmental control

One of the few attempts to control generated designs is presented by Li and Li [25]. It consisted of involving a diffusion model in a daylighting-driven design workflow. Several massing models were parametrically generated, conducting daylight simulations to produce light distribution maps for each mass plan view, then fine-

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

tuning a LoRA using these maps. Trained LoRA was then employed to predict window placements of any new mass. This happened by inserting the plan view as a base image, then through image-to-image, LoRA allowed generating a light distribution map on the plan view, hence predicting the exact windows allocation on the façade.

While this approach showed promising AI-driven workflow that guides façade development during early design stages, it still lacked clearly defined and measurable criteria for describing maps used for LoRA fine-tuning and for evaluating the daylighting performance of the generated outputs.

2.4. Research gap

Previous studies demonstrate GenAI's strong potential for environmental prediction and assessment, yet a clear gap remains in the environmental performance control of AI-generated outputs. Furthermore, while existing research predominantly covers the urban scale, building massing, façades design, and general architectural floor plans generation, environmental performance of AI-generated interior designs remains insufficiently explored.

2.5. Research aim and objectives

This research investigates actively controlling the environmental performance of AI-generated interior designs. It emphasizes daylighting performance, with a particular focus on one of the effective metrics: sDA. Daylighting was chosen as the target environmental performance parameter for its contribution to reducing energy consumption and enhancing occupants' health, comfort, and the overall quality of the built environment [14].

Accordingly, this study aims to investigate the effectiveness of a GenAI-driven workflow for generating interior designs with controlled daylighting performance. An experimental study is applied to a representative office room, where multiple design iterations are generated through the proposed workflow. The scope is bounded to WWR as a key sDA-related variable, while space type, location, finishing materials, façade orientation, and other effective factors are kept fixed to clearly measure the effect of WWR variation on sDA values. A LoRA fine-tuning is then employed to generate DA heatmaps with controlled WWR, aligned with desired sDA ranges.

To achieve this aim, the study pursues the following objectives: (1) to propose a GenAI-driven workflow for generating interior design alternatives aligned with target sDA ranges; (2) to investigate LoRA fine-tuning as a control mechanism for modifying WWR; (3) to validate the AI-generated outputs against simulation-based daylighting results; (4) to visually and statistically evaluate the effectiveness of the proposed workflow in achieving predefined sDA ranges; and (5) to demonstrate the potential of the proposed workflow as a performance-based decision-support tool in early interior design stages.

3. METHODOLOGY

As outlined in the objectives, this study followed a structured experimental methodology to employ and test a proposed GenAI-driven workflow, enabling the achievement of target sDA ranges through AI outputs. The process involved applying a series of detailed procedures using the required tools, followed by validating the experimental results.

3.1. Used tools

To support the experimental procedures, various specialized software tools and AI models were employed. Table 1 presents a summary of these tools, their versions and brief descriptions. More detailed descriptions, along with the justification for their selection, are provided within the corresponding procedures in which they were used.

3.2. Procedures

Experimental procedures included four main sequential steps, starting with the base case modeling, followed by the training dataset preparation, fine-tuning of the LoRA model, and finally the results validation process. Each step is explained in detail in the following sections.

3.2.1. Experiment base case modeling

Experiment was conducted on a specific type of rooms: an office room, with fixed dimensions, orientation, floor level, location, and contextual conditions. All interior design elements such as finishing materials were also fixed, except for the WWR, which was the varying parameter that was subject to further studies through LoRA training in subsequent steps.

Figure 1 illustrates the schematic plan and section of the suggested office room. As shown in (Fig. 1(a) and (b)), it is a room with defined dimensions: 5 m length, 5 m width, and 3.5 m clear height. The only parameters that were left undefined are the window dimensions represented by the variables W (width) and H (height), as well as the number of window openings, as they were subject to variable parametric iterations during the following experiment procedures.

A 3D model of the base case was created using Rhinoceros 3D v.8.8, the latest accessible version at the start of the experiment. The software also provides smooth integration with its built-in parametric design tool, Grasshopper v1, in which modeled base case was defined for the subsequent steps.

It was assumed to be east-oriented and located on the fifth floor. This placement was a specific design choice intended to isolate the room from ground-level contextual obstructions and external reflections, ensuring that WWR remained the sole variable affecting daylight exposure. Modeled room with all its elements including walls, floor, ceiling and windows was then defined as a Honeybee model, with specified properties. Honeybee is a

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

Table 1. Summary of used tools and AI models, including their versions and brief descriptions.

Used Tool / Model	Version	Brief Description
Stable Diffusion (SD)	SD–1.5 / Automatic 1111 Web User Interface (UI)	Text-to-image and image-to-image AI model
Rhinoceros 3D	v8.8	3D modeling tool
Grasshopper	v1 (Built-in Rhinoceros v8.8)	Parametric modeling tool
Honeybee Plugin	Ladybug Tools for Grasshopper v1.8.0	Building performance simulation plugin
Kohya_SS Graphical User Interface (GUI)	v22.3.0	Fine-tuning interface for diffusion models
Light-based ControlNet	v1.0 (Base Model: SD-1.5)	ControlNet model guiding image generation via a reference light distribution
IBM SPSS Software	v27	Statistical analysis software

Table 2. Proposed finishing materials for the office room and the corresponding LRV assigned in the Honeybee model.

	Used Material	Light Reflectance Value (LRV) (%)
Flooring	Wooden floor	40
Walls	Grey paint on plaster	50
Ceiling	White-painted Gypsum Boards	80

Fig. 1. (a) Schematic plan of the proposed dimensions for the base case room, W indicates the varying width of the parametrically generated windows. (b) Internal elevation showing proposed clear height, H indicates the varying windows height.

Grasshopper plugin that creates, runs, and visualizes daylight simulations, through Radiance simulation engine.

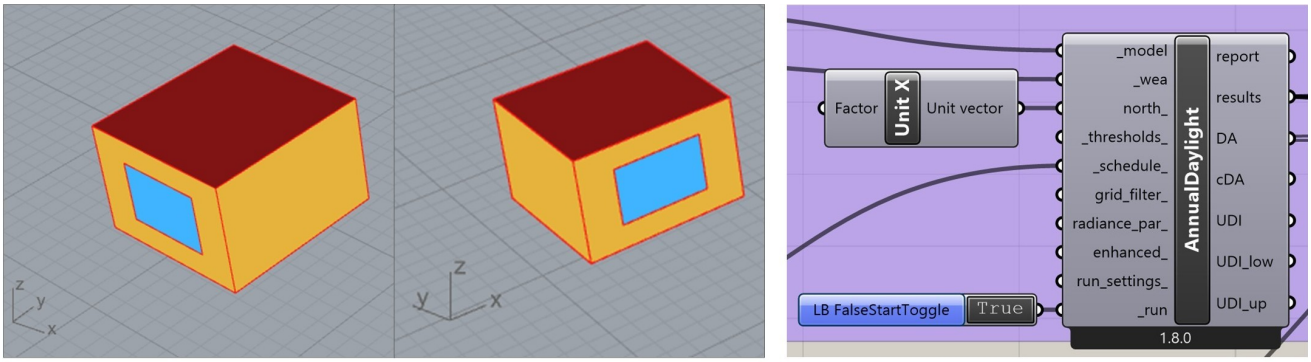
First, room type was defined as an office, located in Cairo, Egypt, with the corresponding CAI weather data file applied. Finishing materials were then assigned in the Honeybee model, with their respective light reflectance values (LRVs) provided in Table 2.

The glazing consisted of a single-pane system with a visible transmittance of 85%. The daily occupancy schedule was set from 8:00 AM to 5:00 PM, with weekends on Friday and Saturday. These specific properties were chosen to be representative of typical local offices, ensuring the simulations reflect realistic daylighting conditions for the specified region. For the daylighting analysis, a grid size of 0.05 by 0.05 m was employed. Within the simulation setup, the north direction was reassigned from the y-axis (Rhinoceros default) to the x-axis. This adjustment, shown in (Fig. 2), ensured that the model windows, aligned with the negative y-axis, were correctly defined as east-facing, as intended.

3.2.2. Training dataset preparation

After modeling the base case office room and defining all its elements as a Honeybee model, Grasshopper was then used to parametrically generate 40 WWR iterations. The WWR values ranged from 1% to 99% and were randomly generated to provide a wide spectrum of opening sizes. Each case was then subject to an annual daylighting simulation. Instead of being constructed at the working plane level, the sensor grid was constructed on all space planes: walls, floor, ceiling, and windows. The annual daylight simulation enabled the evaluation of DA at each analysis sensor, which was represented through a color range, generating a 3D DA heatmap for each WWR iteration.

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.



(a) (b)

Fig. 2. (a) Screenshots from the Rhinoceros interface of the base case Honeybee model, showing the studied façade orientation with respect to the standard XYZ axes. (b) The annual daylight simulation component in the Grasshopper interface, illustrating the north orientation defined for the simulation setup.

Table 3. Categorization of sDA groups applied in experiment procedures based on achieved percentage values.

sDA Group	Value range (%)
Low sDA	0-54
Moderate sDA	55-74
High sDA	75-100

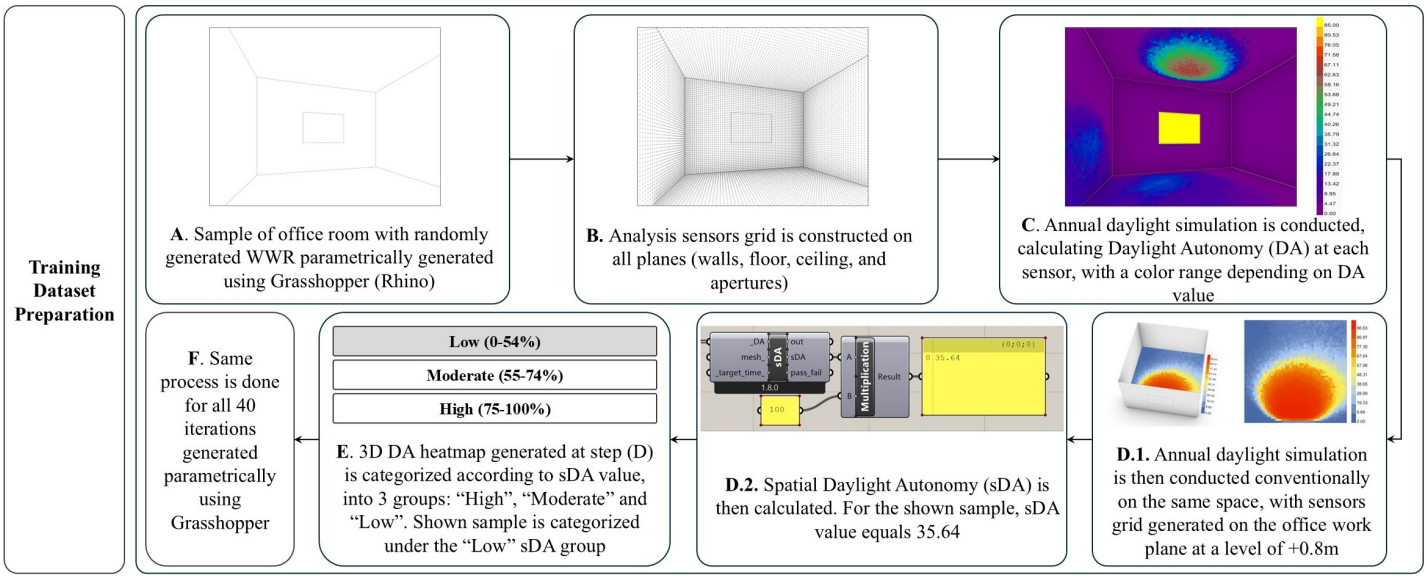


Fig. 3. Diagram showing the step-by-step process of training dataset preparation, illustrated through the presented sample.

The sDA percentage was then calculated for each case through a conventional simulation conducted on a 2D sensor grid located on the office working plane level at +0.8 m.

As noted earlier, sDA300,50% represents the percentage of regularly occupied floor area receiving at least 300 lux for a minimum of 50% of the annual occupied hours [17]. According to the U.S. Green Building Council's Leadership in Energy and Environmental Design (LEED) v4.1 certification [26], specific sDA thresholds are associated with daylight credit scoring. The rating system awards 1 point when 40% or more of the occupied area meets the sDA daylight criterion, 2 points when 55% or more is achieved, and 3 points for values starting from 75%.

Accordingly, an analytical grouping method inspired by LEED daylight thresholds was adopted in the experimental study. Specifically, cases with sDA values below 55% were grouped as "Low," corresponding to the lower LEED credit band (from 40% up to 55%). Cases with sDA values greater than 55% and up to 75% were grouped as "Moderate," reflecting the intermediate LEED threshold range, while cases with sDA values above 75% were grouped as High, consistent with the highest LEED daylight performance threshold. This classification, presented in Table 3, was first applied to categorize the DA heatmaps used for training. It was then applied within the GenAI-driven workflow as a reference grouping for controlling sDA values of generated interior designs.

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

Based on the adopted sDA categorization, the 40-image dataset was divided into 17, 14, and 9 DA heatmaps, belonging to the "High," "Low," and "Moderate" sDA groups, respectively. Figure 3 shows the detailed training dataset preparation workflow applied to one sample of the random WWR iterations that were parametrically generated. After applying the analysis sensors grid on the room planes and conducting the annual daylight simulation, the resulting DA values were represented on a three-dimensional heatmap. As the sDA value of this sample was 35.64%, this sample was classified within the "Low" sDA group according to previously defined categorization groups. Same process was then repeated for all 40 randomly generated WWR iterations in order to construct the training dataset, to be ready for the following training processes.

3.2.3. LoRA fine-tuning

Each DA heatmap was set to be 512 by 512 pixels, then paired with its textual description and labeled according to its sDA group. For example: "High SDA heatmap, of a 5 by 5 meters office room, clear height 3.5 meters, east-facing, at the 5th floor, in Cairo Egypt, grey painted walls, white gypsum ceiling, oak wooden flooring." The 40 images and textual description were then used as a training dataset for LoRA fine-tuning, based on SD-1.5 model.

Many researchers have discussed the ideal dataset size for training a LoRA model, but there is no single agreed number. In general, LoRA requires only a small number of images, usually from one image to a few dozen, as it is very data efficient. For basic object-focused personalization, LoRA models can learn high-quality concepts from only 3 to 5 images. In practice, using around 30 carefully selected object images is often considered a good balance. This amount provides enough variation to train a stable LoRA model [27]. Other informal technical sources suggest that color-gradient–based LoRA models typically require a training dataset consisting of an optimal range of 30 to 50 highly diverse samples. This enables the model to effectively decouple color patterns from underlying geometric structures.

Accordingly, the use of 40 heatmaps for LoRA training in this study can be considered relatively small but sufficient. This dataset size achieves a practical balance between dataset diversity and the limitations of computational resources and experiment time. The fine-tuning process was conducted using Kohya_SS GUI v22.3.0, an accessible user interface that facilitates diffusion models training, fine-tuning parameters configuration, and LoRA implementation. Table 4 shows the exact LoRA training parameters, used in Kohya_SS GUI.

To reduce the risk of overfitting during LoRA fine-tuning, several controlled training strategies were applied. Given the relatively small dataset size, the model capacity was limited by reducing the LoRA rank, setting the network dimension to 64 with a corresponding network alpha of 64. This reduced the number of trainable parameters and helped prevent the model from memorizing specific heatmap patterns. The learning rate was lowered to 8×10^{-5} to ensure gradual parameter updates and avoid rapid convergence to narrow local minima. The total number of training steps was limited to 1000, consisting of 25 epochs with a batch size of 1, which provided sufficient exposure to the dataset to learn spatial patterns and color variations, while avoiding excessive iterations that could lead to memorization and overfitting. No regularization images were used, as the task focuses on structured spatial color distributions rather than identity or concept learning, making additional regularization unnecessary. The resolution was fixed at 512 × 512 pixels, which is standard for the SD-1.5 base model, and mixed precision FP16 was used to maintain stable optimization without increasing model complexity. Overall, these settings balance model capacity and training duration, promoting generalizable learning of spatial illumination gradients instead of simple replication of training samples.

Figure 4 illustrates the process of LoRA fine-tuning using samples from the dataset of 40 DA heatmaps employed for training. The diagram shows three representative heatmaps from each sDA category: "High," "Moderate," and "Low," paired with a text description indicating their categorization group and used for LoRA fine-tuning.

As a result, trained LoRA was able to generate DA heatmaps through text-to-image generation with varying WWR, depending on the inserted prompt keywords. It could generate new heatmaps that are aligned with target sDA groups, as specified by prompt keyword, whether "High," "Moderate," or "Low." The base AI model used in this study for DA heatmap generation through the trained LoRA in Stable Diffusion was SD-1.5.

Stable Diffusion is a text-to-image and image-to-image AI model that generates images based on textual prompt. It operates using a latent diffusion architecture, a deep learning approach that enables efficient and high-quality image synthesis. Among the most common user interfaces to interact with Stable Diffusion are Web UI and Comfy UI [28].

For experimental procedures, Automatic 1111 Web UI was used for images generation for its accessibility, ease of use, and direct control over generation settings, which supported a consistent and repeatable workflow. Figure 5 shows the image generation process of a new heatmap using trained LoRA.

To summarize the experiment procedures, (Fig. 6 shows the detailed steps starting with the base case modeling, then the preparation of the DA heatmaps training dataset, which was finally used for LoRA fine-tuning. Then results validation took place to validate the effect of the trained LoRA on adjusting sDA values of generated designs.

3.3. LoRA results validation process

LoRA-generated outputs were tested and analyzed through several procedures. First, trained LoRA was used to generate 120 new DA heatmaps, 40 maps using each of the following prompt keyword: "High," "Moderate," and "Low" SDA maps.

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

Table 4. LoRA fine-tuning parameters, dataset settings, and computational specifications adopted in the Kohya_SS GUI.

Parameters	Experiment Values
Dataset Size	40
Epochs	25
Batch Size	1
Training Steps	1000
Base Model	SD-1.5
Max Resolution	512 by 512 pixels
Learning Rate	0.00008
Network Dimension	64
Network Alpha	64
Mixed Precision	Fp16
Save Precision	Fp16
xformers	True
Regularization	No regularization images
GPU (VRAM)	RTX4060 (8GB)
Training Time	About 3 Hours

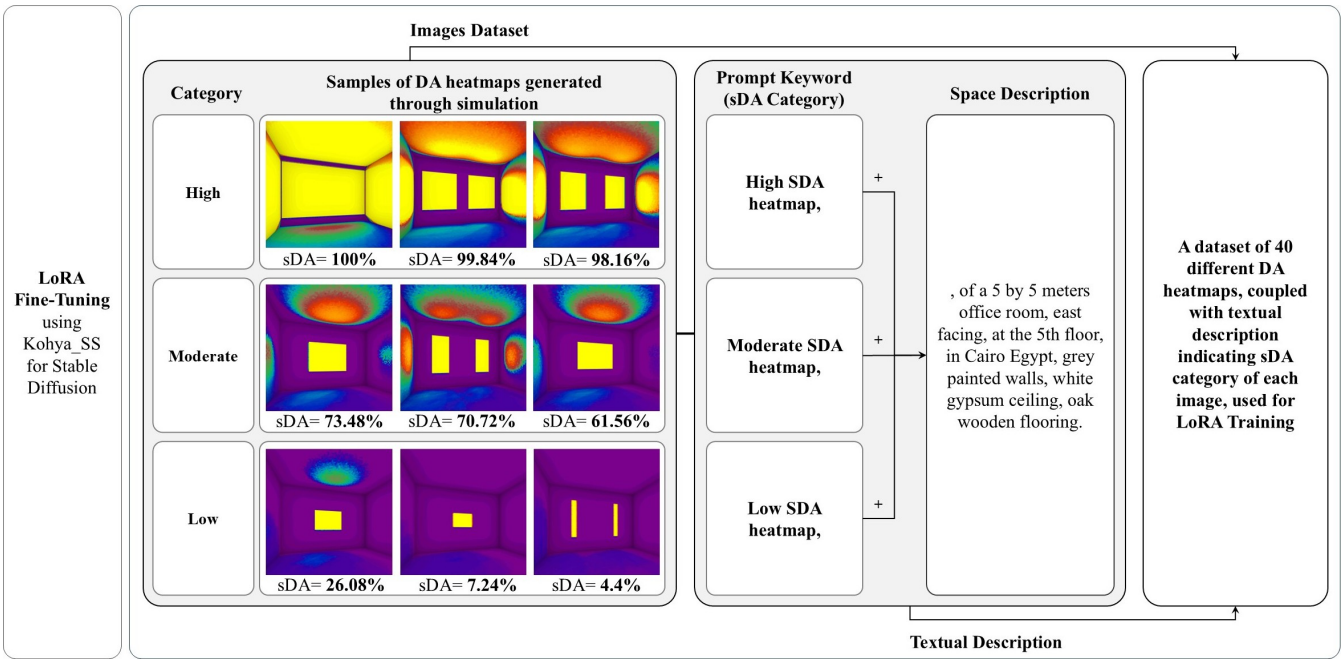


Fig. 4. Diagram illustrating detailed steps of the LoRA fine-tuning through the presented samples.

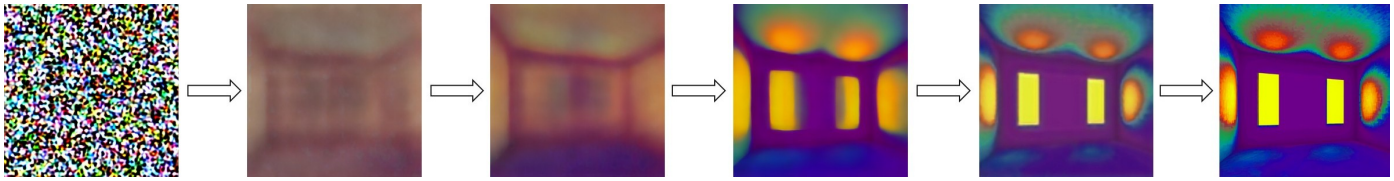


Fig. 5. Denoising process for DA heatmap generation, using trained LoRA.

The validation sample consisted of 40 heatmaps per category to effectively assess whether the model generalized the spatial daylighting patterns rather than overfitting. This validation setup reduced the likelihood of memorization of the original 40-image training dataset. The generated maps were then subject to four analysis steps, including visual inspection, image analysis, annual daylight simulations and statistical analysis tests.

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

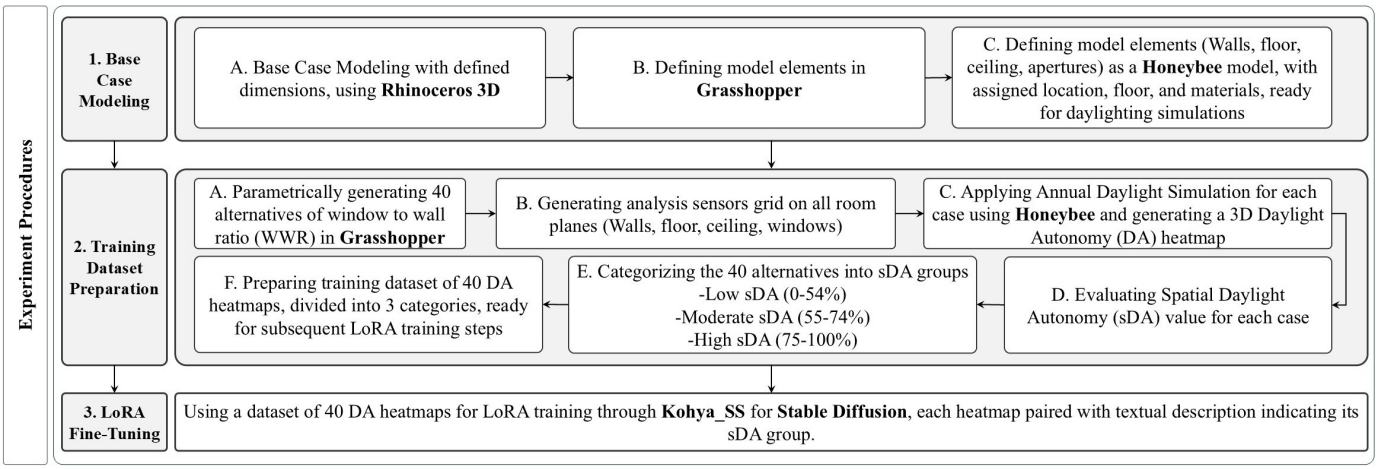


Fig. 6. Workflow diagram of the experimental study, including base case modeling, training dataset preparation and LoRA fine-tuning.

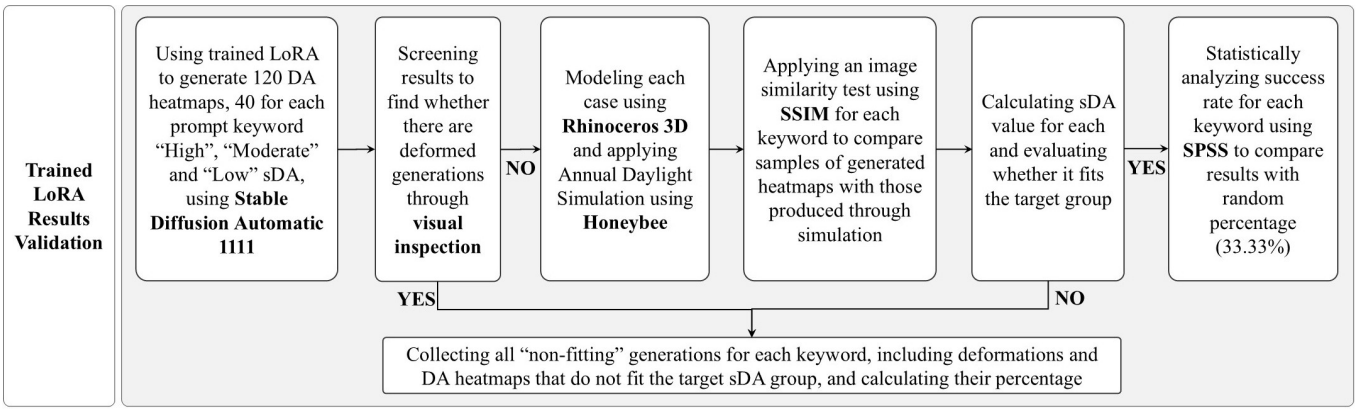


Fig. 7. Diagram illustrating detailed steps of the trained LoRA results validation.

3.3.1. Visual inspection

As noted by Ploennigs and Berger [29], text-to-image tools sometimes generate deformed output, failing to accurately match inserted prompts. Therefore, DA maps generated by trained LoRA were first screened via visual inspection to exclude irrelevant results. The criteria for identifying deformed outputs included the presence of irregular room or windows geometric boundaries, or unrealistic color gradients that were inconsistent with physics-based daylight distribution patterns.

3.3.2. Image analysis

To evaluate each generated heatmap after the exclusion of deformed results, the modeled base case WWR was first adjusted to correspond with the generated image. An annual daylight simulation was then performed to conventionally produce a DA heatmap for the same WWR. Used metric to compare the DA heatmaps generated by the trained LoRA model with those obtained from the simulation was the SSIM. SSIM was selected as the image analysis metric because the objective is to evaluate similarity in spatial daylighting patterns, not only pixel-wise numerical differences. Unlike error-based metrics such as Mean Squared Error (MSE) or Peak Signal-to-Noise Ratio (PSNR), SSIM assesses image similarity through luminance, contrast, and structural correlation. Rather than comparing images solely on a pixel-by-pixel error basis, these three aspects are combined into a single score that describes how the test image preserves the visual structure of the reference image. The SSIM value ranges from 0 to 1, where 1 indicates that the two images are identical. As the SSIM value becomes closer to 1, the images are considered more similar, meaning that any distortions are minor and difficult to notice [30]. In addition, SSIM is one of the most widely used reference-based image similarity metrics in relevant studies, including those comparing AI-generated heatmaps with simulation-based results, as identified in the reviewed literature.

3.3.3. Annual daylight simulation

Similarly, after the exclusion of the deformed images, DA maps generated by trained LoRA were used for adjusting the WWR of the base case 3D model, followed by annual daylight simulation using Honeybee plugin. Then sDA percentage was calculated for each generation to verify whether it meets the target sDA group. The 40 maps generated using each prompt keyword were categorized into two groups: "Fitting" and "Not Fitting." Maps were considered "Not Fitting" if they were deformed or reached sDA values outside

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

the target group range. The percentage of each category was then recorded to evaluate LoRA accuracy in reaching target sDA.

3.3.4. Statistical analysis

A one-sample t-test was conducted to evaluate the correspondence between the inserted prompt keyword and the sDA classification of the generated designs using SPSS. It is a statistical analysis software that plays an essential role in modern research by helping researchers organize, process, and analyze data to reach clear conclusions. It is widely used in fields such as social science, health, business, and education due to its easy interface and strong analytical tools [31]. Given that there are three possible sDA categories: "High," "Moderate," and "Low," the random chance percentage is 33.33%. Between the 40 generated heatmaps for each prompt keyword, statistical test compared the percentage of maps that fall within the intended sDA category against the random-chance baseline. As a result, this validation stage allowed assessing the effectiveness of the LoRA model in generating outputs that match target performance.

Figure 7 shows the detailed validation procedures as discussed, starting with the visual inspection, followed by the image-based analysis using SSIM. Annual daylight simulation for each generated design was then conducted to calculate the percentage of "Fitting" and "Non-Fitting" heatmaps. Finally, the one-sample t-test allowed comparing each prompt keyword generations with the random-chance percentage.

4. RESULTS

Following the validation process, results were systematically recorded and analyzed by various means, including visual and statistical methods. This analysis enabled the effectiveness assessment of the fine-tuned LoRA in achieving sDA ranges of the generated designs through the control of WWR.

4.1. Visual inspection

Visual analysis of the AI-generated DA heatmaps enabled the identification of deformed results, including irregular outputs or unrealistic light distribution patterns. Once detected, these heatmaps were excluded from the proposed GenAI-driven design workflow.

Table 5 shows the deformed heatmaps generated using the trained LoRA model for each prompt keyword. For the "High" keyword, 4 out of 40 generated heatmaps were identified as deformed, representing 10% of the outputs. Meanwhile, the "Moderate" and "Low" keywords each produced 3 deformed heatmaps, with a percentage of 7.5% of the 40 generated samples at each case.

The slight variation in deformation may indicate that generating bright, high-intensity light distributions may introduce an increased susceptibility to visual deformation during the diffusion process. However, the overall deformation rate remained relatively low across all categories.

4.2. Image analysis

Since there is no universally accepted SSIM threshold for determining image similarity, several studies have attempted to identify an appropriate range of values that can reliably indicate structural homogeneity between compared images. As an example, Lee et al. [32] proposed a practical value based on performance in patient-specific quality assurance (PSQA) application. In their study, SSIM was used to compare planned and measured radiation dose distributions. The researchers evaluated the SSIM Passing Rate (SPR), which represents the percentage of pixels that reach a certain similarity level, across SSIM values from 0.00 to 1.00. Their results showed that a threshold of 0.65 was the most suitable for distinguishing between minor differences and clinically important errors. Based on their findings, they suggested that an SSIM value of 0.65 or higher can be considered sufficient for clinical evaluation. It should be noted that the 0.65 threshold originated from a medical application and may not directly apply to daylighting heatmap evaluation. This reference was adopted only as an indicative benchmark, as no studies in the reviewed literature were found to define an accepted SSIM threshold for architectural image analysis or simulation heatmaps comparison.

After applying the SSIM test on the 40 generated DA heatmaps for each prompt keyword, the average value was calculated for each case. The results indicated that the heatmaps generated using the "High" keyword achieved the highest similarity, with an average SSIM value of 0.851, followed by the "Moderate" keyword, which recorded an average SSIM of 0.740. The lowest similarity was recorded for the "Low" keyword generations, with an average value of 0.694.

These values indicate a clear decrease in SSIM as the target sDA category shifts from "High" to "Low." This decline reflects lower accuracy in generating subtle illumination gradients under low-light conditions. A contributing factor may be the training dataset size, as the number of training samples in the "High" category was greater than in the "Moderate" and "Low" categories. Increasing dataset size and diversity, particularly for the "Moderate" and "Low" categories, may help mitigate this issue and improve generation consistency.

Table 6 presents examples of the performed SSIM tests, showing the 3D heatmaps generated using the trained LoRA, and their corresponding heatmaps generated through simulation. The resulting SSIM values for the shown samples are also provided.

4.3. Annual daylight simulation

After conducting an annual daylight simulation for each generated heatmap, 82.5% of the 40 iterations generated using "High" prompt keyword were correctly falling within the High sDA group.

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

Table 5. Deformed samples generated using different prompt keywords through trained LoRA model.

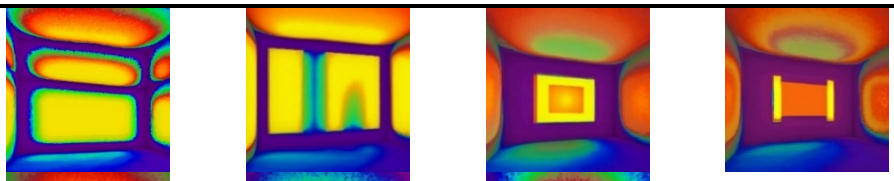
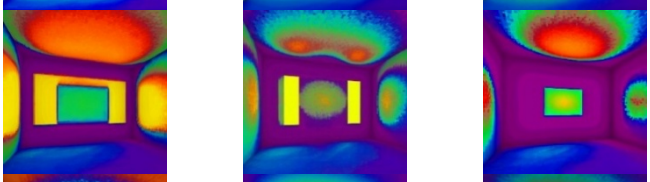
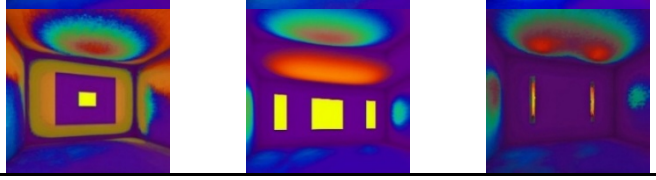
Case	Prompt Keywords	Deformed Outputs			
		1	2	3	4
1	"High sDA maps"				
2	"Moderate sDA maps"				-
3	"Low sDA maps"				-

Table 6. SSIM measurement test applied to samples of heatmaps generated using trained LoRA and their corresponding heatmaps generated through simulation.

DA heatmaps category	**"High"**			
Sample of LoRA-generated DA heatmaps				
Same WWR heatmaps generated through simulation				
SSIM	0.858	0.923	0.936	0.958
DA heatmaps category	**"Moderate"**			
Sample of LoRA-generated DA heatmaps				
Same WWR heatmaps generated through simulation				
SSIM	0.811	0.681	0.739	0.835
DA heatmaps category	**"Low"**			
Sample of LoRA-generated DA heatmaps				
Same WWR heatmaps generated through simulation				
SSIM	0.746	0.702	0.698	0.739

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

The rest were whether deformed results or within other sDA groups. "Moderate" prompt keyword achieved a fitting percentage of 67.5%, while the "Low" keyword achieved a percentage of 75%.

Table 7 provides a detailed categorization of the 40 iterations generated for each prompt keyword, classifying them as "Fitting" and "Not Fitting." "Not Fitting" category is divided into "Deformed" or "Outside Target Range" outputs. Exact percentage of each category is also provided.

The relatively limited training dataset size may have contributed to the lower fitting rate of the "Moderate" sDA category. In addition, as previously stated, the number of "Moderate" training images was lower than that of the "High" and "Low" categories, which may have reduced category separability and increased the possibility of boundary overlap between adjacent sDA ranges. This limitation could potentially be addressed by expanding the training dataset and increasing the number of "Moderate" samples.

For a deeper analysis of the non-fitting samples for each sDA category, a confusion matrix was also conducted on the trained LoRA outputs, presented in Table 8. It is defined as a contingency table used to evaluate the performance of a classification model by directly comparing its predicted outcomes against the actual results [33]. Instead of just providing an overall accuracy score, the matrix provides a transparent breakdown of where the model is succeeding and where it is getting "confused" by different categories. This is why it is essential for deep analysis of category misclassifications [34].

The results indicate that for the "High" and "Low" categories, non-fitting generations are approximately evenly distributed between the other two categories. However, the "Low" category shows a relatively higher percentage of misclassified outputs, with a high concentration near the "Low" category. This may suggest a slight boundary overlap between these performance levels, indicating that a larger training dataset could be beneficial to improve category separation and model discrimination accuracy.

4.4. Statistical analysis

For each prompt keyword, results were compared with random percentage, to ensure trained LoRA has a positive impact on daylight performance, not just a random effect. For the "High" category, the success rate of fitting target group was about 83%, significantly higher than the random baseline chance of 33.3% ($t(39) = 8.086$, $p < .001$, Cohen's $d = 1.279$), with a mean difference of 0.492 and a 95% Confidence Interval (CI) of 0.369 to 0.615 (approximately 0.702 to 0.948 for the mean fitting rate). This indicates a very strong effect of trained LoRA on matching "High sDA maps" prompt keywords. The Moderate category achieved a 68% fitting rate, lower than "High" category, but still greater than random percentage ($t(39) = 4.560$, $p < .001$, Cohen's $d = 0.721$), with a mean difference of 0.342 and a 95% CI of 0.19 to 0.494 (approximately 0.524 to 0.827 for the mean fitting rate), indicating a medium LoRA effect. Lastly, the "Low" category achieved a 75% fitting rate, with a strong statistical significance ($t(39) = 6.014$, $p < .001$, Cohen's $d = 0.951$), with a mean difference of 0.417 and a 95% CI of 0.277 to 0.557 (approximately 0.61 to 0.891 for the mean fitting rate), indicating a large effect of trained LoRA on generating maps with low sDA values. Statistical results of the three sDA categories are presented in Table 9.

5. LORA IMPLEMENTATION IN THE PROPOSED DESIGN WORKFLOW

5.1. Workflow application through a case study

Generally, previous results confirm that the trained LoRA produces indicative and homogenous DA maps, similar to those generated by daylight simulations, as shown in the generated images analysis. It is also proved statistically that maps conform to the inserted prompt keyword. After the validation of the trained LoRA being able to generate DA maps that fit target sDA groups, it can be implemented in the GenAI-driven interior design process to integrate daylighting performance as a controlled parameter through design generations.

For instance, proposed workflow was applied to the previously defined office space to generate four interior design alternatives per sDA category. First, trained LoRA was applied in SD Automatic 1111 Web UI to create 3 DA heatmaps, one for each prompt keyword. Each map was then used as an input to the light-based ControlNet, functioning as a guiding condition. Then a detailed description for the intended room design was inserted as text prompt. For the 3 cases, the same prompt was used: "Interior Design of a 5 by 5 meters office room, east-facing, at the 5th floor, in Cairo, Egypt, grey painted walls, white gypsum ceiling, oak wooden flooring." Finally, four design iterations were generated at each case, all fitting target sDA values range.

The chart presented at (Fig. 8) shows the application of the proposed GenAI-driven workflow, generating interior designs for each case "High," "Moderate," and "Low" sDA values. It must be noted that the evaluation of their architectural quality and realism remains highly subjective. The assessment of aesthetics was considered outside this study scope.

5.2. Comparison with conventional optimization

To compare the computational time and output of the proposed GenAI-driven workflow with a conventional optimization approach, the previous case study was tested using Galapagos in Grasshopper. It is an optimization tool used to find optimal design solutions by varying input parameters according to a defined fitness objective. An optimization process was conducted on the previously described office room, with the objective of achieving high sDA values. Window parameters, including window height, width, sill height, and horizontal separation distance, were defined as genomes to be varied throughout the optimization process, while the resulting sDA value was assigned as the fitness to be maximized.

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

Table 7. Fitting rates of the 40 generated samples per inserted prompt keyword (High, Moderate and Low), showing Fitting and Not Fitting maps percentage to the target sDA range.

AI-Generated DA maps Category	"Fitting" Percentage	"Not Fitting" Total Percentage	Deformed	Outside Target Range
"High sDA maps"	82.5%	17.5%	10%	7.5%
"Moderate sDA maps"	67.5%	32.5%	7.5%	25%
"Low sDA maps"	75%	25%	7.5%	17.5%

Table 8. Confusion matrix of the non-fitting LoRA-generated outputs for the three prompt keywords: "High," "Moderate," and "Low.".

	High	Moderate	Low	Deformed
High	33	2	1	4
Moderate	3	27	7	3
Low	4	3	30	3

Table 9. Comparison of the statistical results for each prompt keyword with the random baseline percentage (33.3%).

Prompt keyword	Number of Samples	Mean category-match value	t statistic t(39)	p (2-tailed)	Cohen's d effect size	Mean Difference from Random Baseline (0.333)	95% CI for Mean Difference
High	40	0.83	8.086	< 0.001	1.279	0.492	0.369 to 0.615
Moderate	40	0.68	4.560	< 0.001	0.721	0.342	0.19 to 0.494
Low	40	0.75	6.014	< 0.001	0.951	0.417	0.277 to 0.557

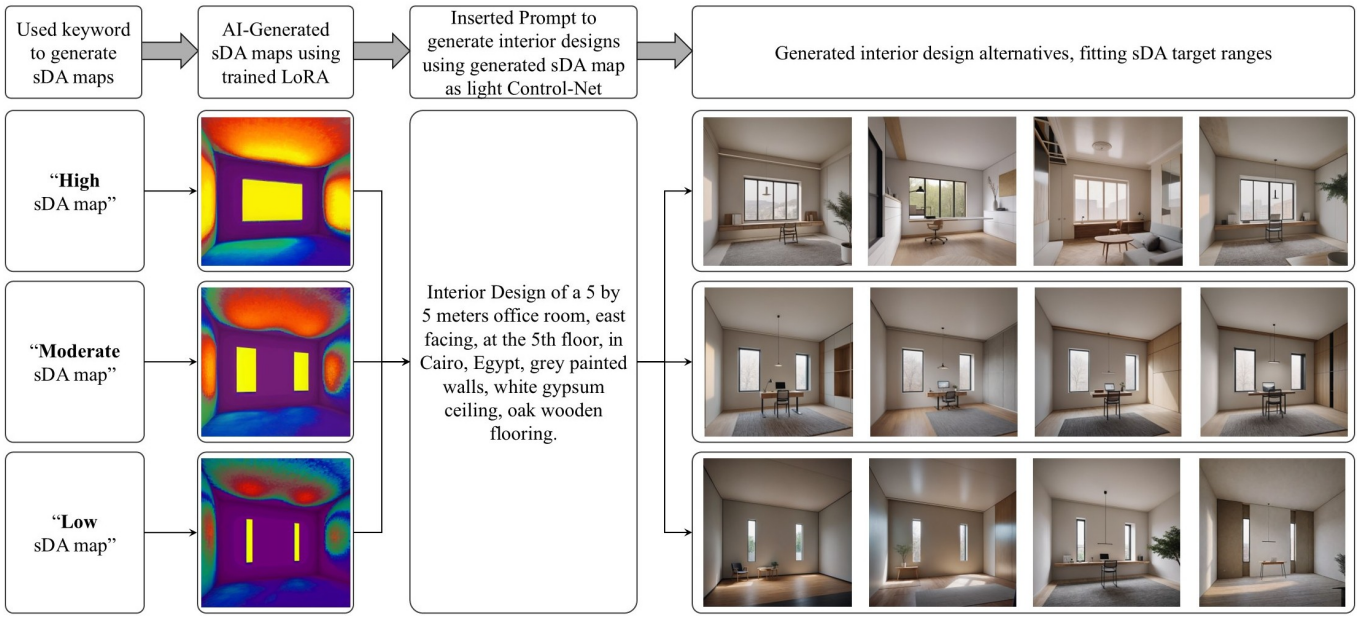


Fig. 8. Diagram illustrating proposed GenAI-driven interior design workflow, adapted to daylighting performance.

Figure 9 illustrates the selected genome parameters, fitness definition, Galapagos settings, and the base case Honeybee model.

After running the optimization, Galapagos processed a large number of iterations. Since each iteration required an annual daylight simulation to calculate the resulting sDA value and identify higher-performing solutions, the process involved high computational and time costs. In the conducted trial, the runtime exceeded 12 hours before the solver was manually stopped at approximately 700 iterations. Figure 10 presents resulting iterations and a sample of the sDA values displayed in the Galapagos interface.

Three high-sDA iterations were selected as representative examples. The first achieved an sDA value of 100% with a window height of 2.1 m, a window width of 4.6 m, and a sill height of 0.7 m. The second achieved an sDA value of 93.28% with a window height of 1.7 m, a window width of 3.3 m, and a sill height of 0.3 m. The third reached an sDA value of 89.96% with a window height of 1.8 m, a window width of 3.0 m, and a sill height of 0.2 m. In all three cases, no horizontal separation was applied.

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

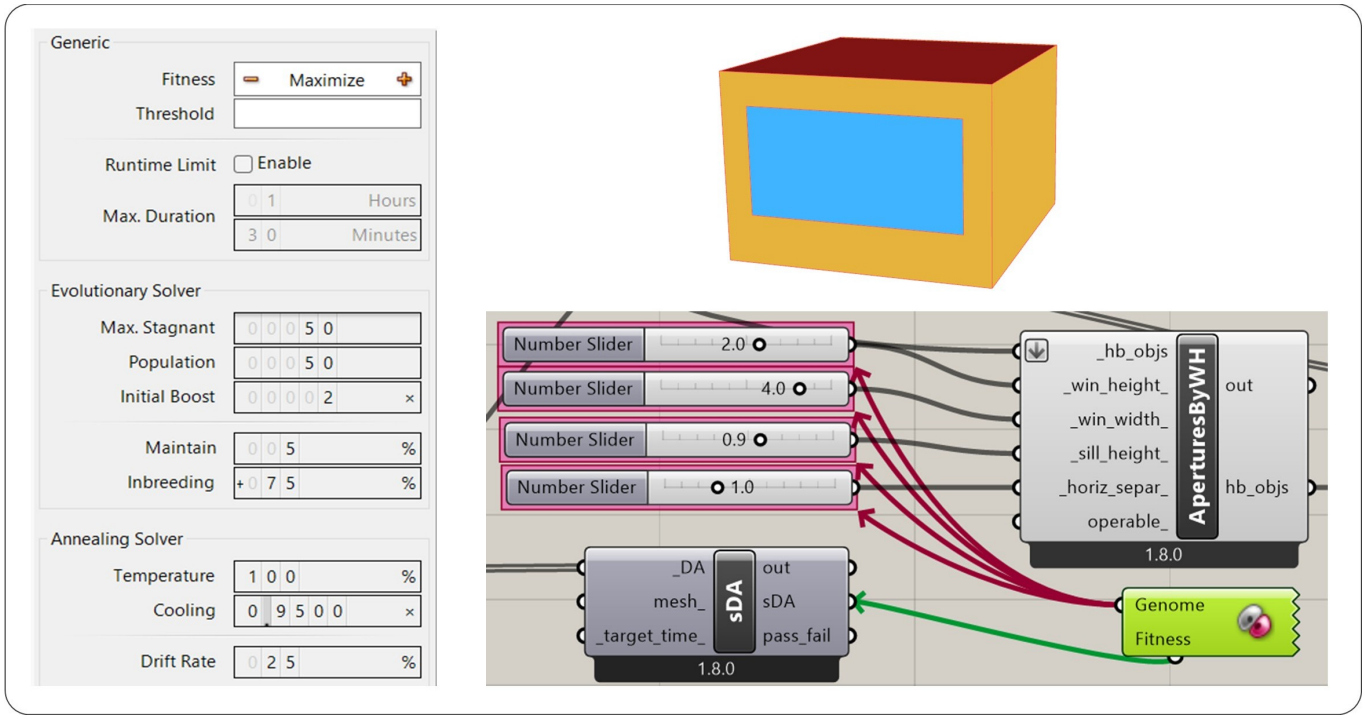


Fig. 9. Galapagos optimization setup in Grasshopper, showing the selected genome parameters, fitness definition, optimization settings, and the base case Honeybee model.

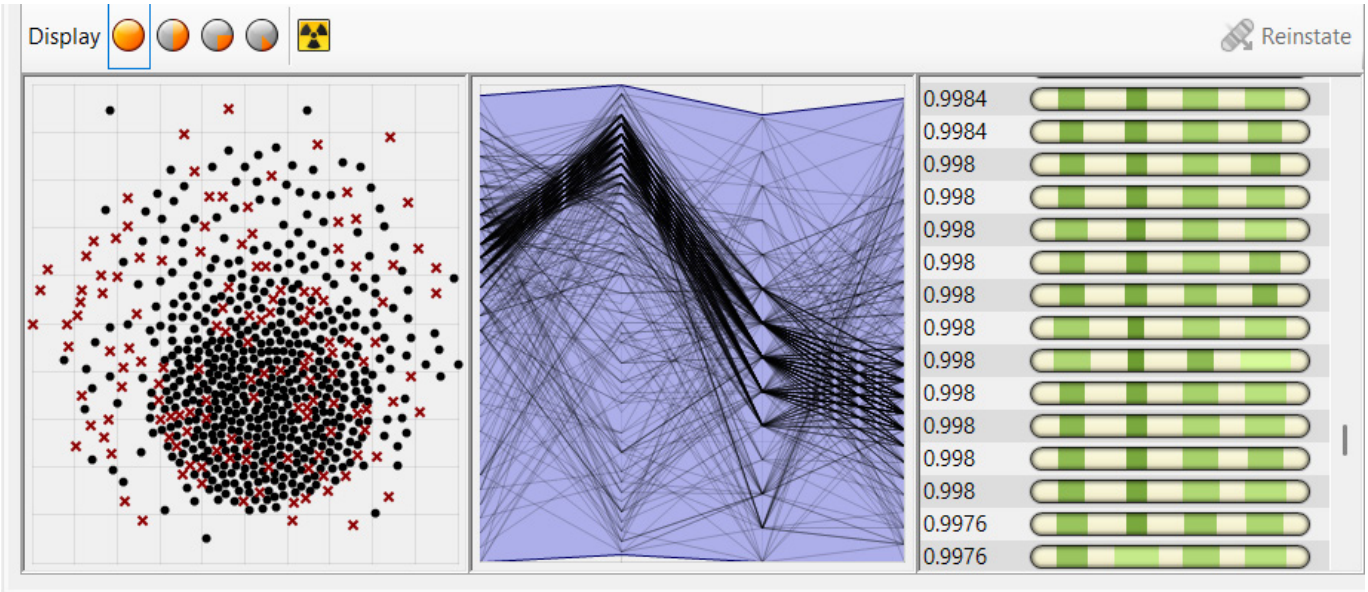


Fig. 10. Optimization results, showing the generated iterations and a sample of the corresponding sDA values in the Galapagos interface.

Three-dimensional perspectives, DA heatmaps, and sDA values of these examples are shown in Fig. 11.

Despite the long computational time, the output was limited to a set of solutions achieving high sDA values, but only in the form of generic geometric models rather than complete interior design proposals. To proceed beyond this stage, the selected options required further interior design development, followed by 3D modeling and realistic rendering. These additional processes required approximately 1–2 more hours, depending on the intended render quality. Moreover, if other WWR configurations were selected to better suit the interior design composition, the design, modeling, and rendering stages would need to be repeated. Similarly, if a different target sDA range was intended, the workflow would need to restart from the optimization stage.

By contrast, the GenAI-driven workflow required approximately 6.5 hours, excluding the base case modeling, which was identical in both workflows.

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

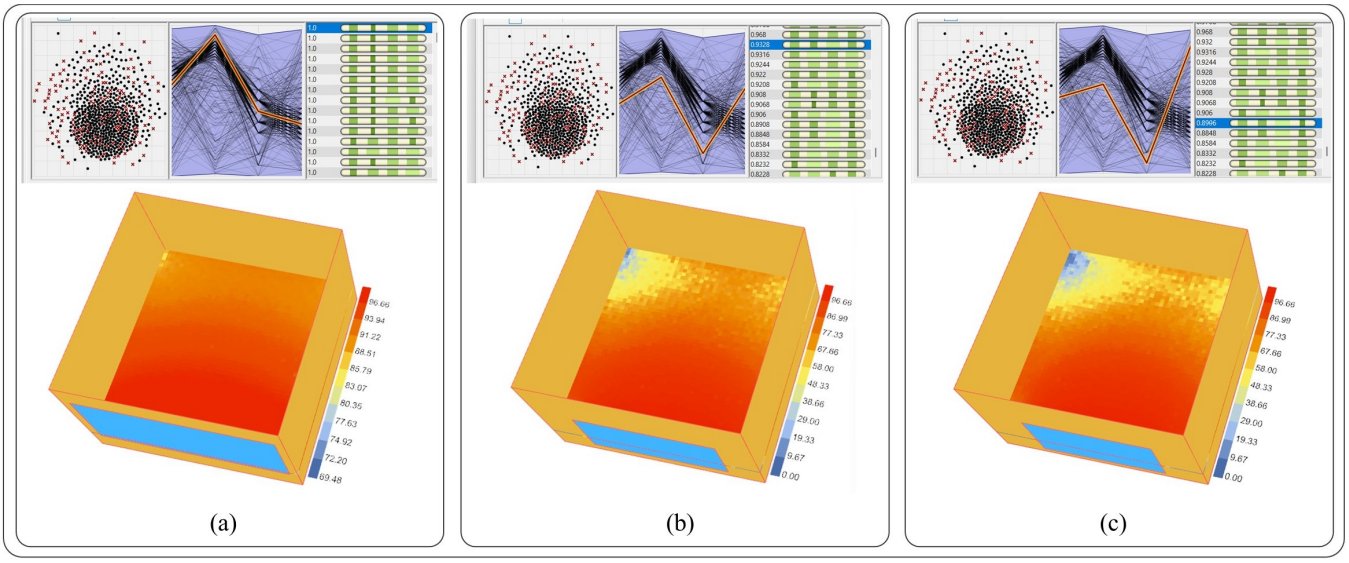


Fig. 11. (a) High-sDA optimized solution with sDA = 100%. (b) sDA = 93.28%. (c) sDA = 89.96%. Corresponding 3D perspectives and DA heatmaps are shown for each case.

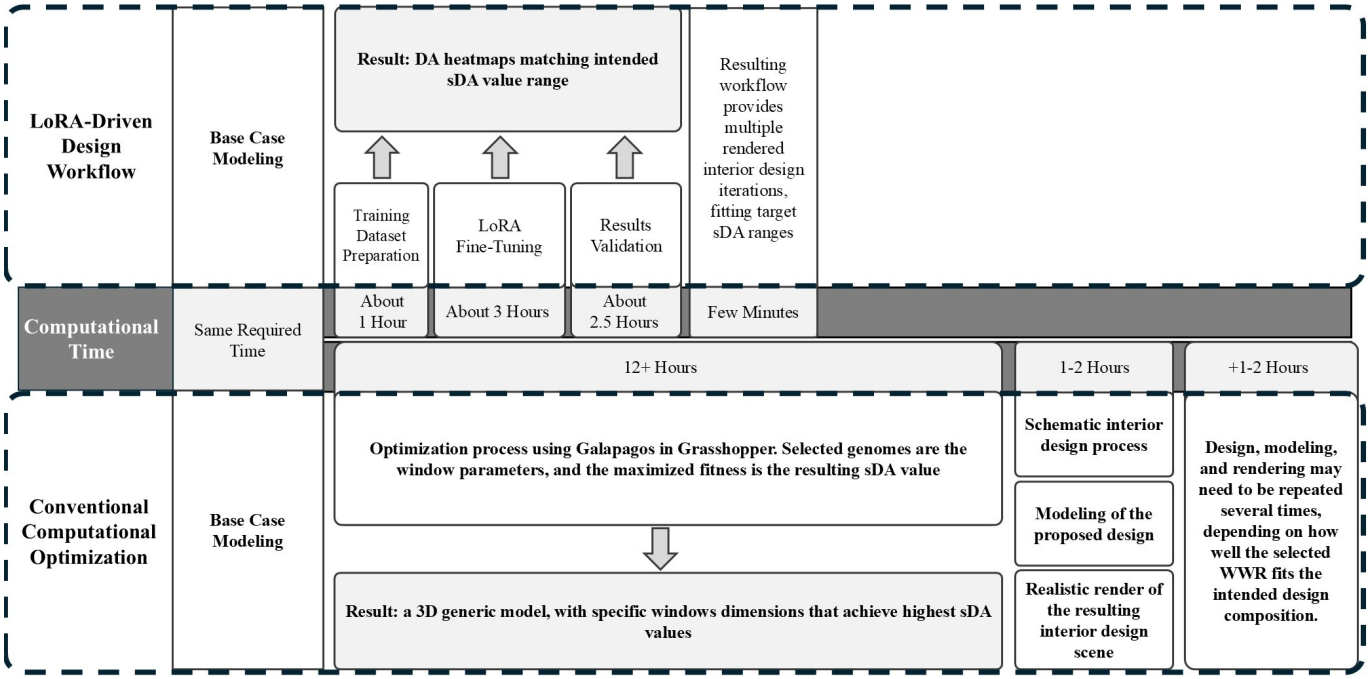


Fig. 12. Comparison between conventional optimization and proposed GenAI-driven workflows in terms of computational time and output type.

After dataset preparation, fine-tuning, and validation, the trained LoRA model became capable of generating multiple iterations corresponding to various target sDA ranges within minutes. In addition, the generated output was not a preliminary geometric model, but a complete interior design iteration delivered as a realistic rendered scene. The comparison between the two workflows is shown in Fig. 12. It highlights the advantage of the GenAI-driven workflow in reducing computation time and enabling a faster transition from daylighting targets to final visualized design iterations.

5.3. Guidance for practitioners

The proposed Gen-AI driven workflow can be similarly adapted to other interior spaces under different conditions, as long as the workflow setup and dataset are adjusted accordingly. This allows architects and interior designers to turn their AI-driven inspirational outputs into more refined proposals, fitting the target daylighting performance, through an easy user-friendly process.

To practically implement this workflow within a target interior space, practitioners may follow the following steps: (1)Base Case Modeling: creating a 3D model of the intended space; (2)Dataset Preparation: generating WWR alternatives and running annual

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

daylight simulations, resulting in a 3D DA heatmap for each case; (3) LoRA Fine-Tuning: using generated maps as a dataset for LoRA fine-tuning after categorizing them into sDA groups; (4) Results Validation: validating outputs through visual inspection, simulations, image-based and statistical analyses; (5) Target Definition: determining the desired sDA target for the interior space, whether High, Moderate, or Low; (6) Heatmap Generation: using the fine-tuned LoRA model via SD to create a DA map matching the selected target; (7) Conditional Design Generation: applying the generated DA map to the light-based ControlNet as a guiding condition; (8) Prompt Insertion: entering a detailed descriptive prompt for the interior space; and (9) Iterative Generation: running the generation process to produce multiple interior design alternatives until a preferred composition is obtained.

6. LIMITATIONS AND WORKFLOW GENERALIZABILITY

While the experimental results demonstrate proof-of-concept success, several limitations related to available computational resources and time should be acknowledged and may be addressed in future studies.

6.1. Dataset size and validation process

As stated in the methodology, LoRA model was trained on a relatively small dataset of 40 DA heatmaps derived from parametric WWR variations. Although this size is generally considered sufficient for LoRA fine-tuning, it remains limited for establishing strong generalizability.

In addition, results validation strategy was applied consistently across the dataset; however, future studies should incorporate more explicit cross-validation or hold-out validation procedures, to better evaluate predictive stability and boundary sensitivity. While the experimental study was limited by some factors, the same methodology can be systematically extended to other conditions, including different orientations, floor levels, space functions, internal finishing materials, and external obstructions or reflections. This expansion, however, would require a significantly larger dataset for LoRA training, requiring more time and computational resources. Nevertheless, with broader training data coverage, LoRA can be fine-tuned to respond accurately to a wider range of keywords related to interior design, hence achieving higher qualities of daylighting performance.

6.2. Experimental scope

The experimental setup kept all interior design parameters fixed, including space type, location, orientation, surface reflectance, glazing properties, external obstructions, and contextual conditions, while WWR was the only selected variable. Although this controlled setup allowed a clear assessment of the WWR effect on sDA, the trained model currently learns daylight behavior under a highly constrained set of conditions rather than a generalized strategy. Accordingly, broader applicability would require retraining the LoRA model on multi-variable datasets incorporating a wider range of environmental and geometric conditions. Such an expansion would significantly increase training complexity and computational requirements, but it would be necessary to improve the workflow generalizability.

6.3. Broader implications

The proposed workflow offers a structured approach for integrating environmental performance targets into GenAI-driven design systems. It could also be extended to other environmental aspects, such as thermal performance or indoor environmental quality, if sufficiently diverse and well-structured training datasets are developed.

7. CONCLUSION

Despite the mentioned limitations, experiment results demonstrate that the trained LoRA positively influenced the sDA category of the generated DA heatmaps, thereby improving one of the most critical environmental parameters in the designed interior space: daylighting performance. SSIM analysis confirmed strong similarity between AI-generated and simulation-based DA maps, with average values of 0.851, 0.740, and 0.694 for the "High," "Moderate," and "Low" categories, respectively. In addition, statistical validation conducted on 120 generated samples (40 per category) confirmed that the generated outputs matched the target sDA ranges beyond random chance (33.33%). Fitting rates reached 83% for the "High" category, 68% for "Moderate," and 75% for "Low."

These results indicate that the proposed workflow successfully integrates GenAI tools into a more comprehensive interior design process. Rather than limiting AI applications to the production of architectural mood boards, the methodology enables the generation of multiple design iterations that respond to targeted daylighting performance.

This research therefore highlights the potential of fine-tuned diffusion models to bridge visual quality and environmental validation. The findings suggest that GenAI tools can move beyond aesthetics and stylistic consistency toward performance-aware design generation. With the development of broader and more diverse datasets, the proposed workflow may support the efficient achievement of aesthetic, functional, structural, and environmental goals in the AI-generated designs.

FUNDING

This research received no external funding.

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to everyone who contributed through valuable discussions and support during the research process and experimental procedures.

AUTHOR CONTRIBUTIONS

Conceptualization: M.W.B., A.A. and F.F.; Methodology, M.W.B., A.A. and F.F.; Software , M.W.B.; Validation, A.A. and F.F.; Formal Analysis, F.F.; Investigation, M.W.B. and F.F.; Resources, M.W.B., A.A. and F.F.; Data Curation, M.W.B. and F.F.; Writing – original draft, M.W.B.; Writing – review and editing, A.A. and F.F.; Visualization, M.W.B.; Supervision, A.A. and F.F.; Project Administration, A.A. and F.F. All authors have read and agreed to the published version of the manuscript.

DECLARATION OF COMPETING INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] S. K. Baduge, S. Thilakarathna, J. S. Perera, M. Arashpour, P. Sharafi, B. Teodosio, A. Shringi, and P. Mendis, Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning methods and applications, Automation in Construction, 141 (2022) 104440.
[2] C. Debrah, A. P. C. Chan, and A. Darko, Artificial intelligence in green building, Automation in Construction, 137 (2022) 104192.
[3] J. McCarthy, M. L. Minsky, N. Rochester, and C. E. Shannon, A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence, AI Magazine, 27:4 (2006) 12-14.
[4] A. Darko, A. P. C. Chan, M. A. Adabre, D. J. Edwards, M. R. Hosseini, and E. E. Ameyaw, Artificial intelligence in the AEC industry: Scientometric analysis and visualization of research activities, Automation in Construction, 112 (2020) 103081.
[5] E. Yildirim, Text-to-Image Generation A.I. in Architecture, in: Art and Architecture: Theory, Practice and Experience, H. H. Kozlu, Ed., Livre de Lyon: Lyon, France, 2022, pp. 97-119.
[6] C. Li, T. Zhang, X. Du, Y. Zhang, and H. Xie, Generative AI models for different steps in architectural design: A literature review, Frontiers of Architectural Research, 14:3 (2025) 759-783.
[7] D. Newton, Generative Deep Learning in Architectural Design, Technology|Architecture + Design, 3:2 (2019) 176-189.
[8] P. Fernandez, Technology behind text to image generators, Library Hi Tech News, 39:10 (2022) 1-4.
[9] X. Hu, H. Zheng, and D. Lai, Prediction and optimization of daylight performance of AI-generated residential floor plans, Building and Environment, 279 (2025) 113054.
[10] L. Zhang, A. Rao, and M. Agrawala, Adding Conditional Control to Text-to-Image Diffusion Models, in: 2023 IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France, 1 October 2023, pp. 3813-3824.
[11] Y. Zhou and Y. Pan, A Generative AI Framework for Adaptive Residential Layout Design Responding to Family Lifecycle Changes, Buildings, 15:22 (2025) 4155.
[12] T. Çelik, Evaluating daylight performance of AI-generated housing plans via diffusion models and climate-based simulation, Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 39 (2025) e19.
[13] J. C. L. Cervantes and C. E. S. Morales, Design in the Age of Predictive Architecture: From Digital Models to Parametric Code to Latent Space, Architecture, 6:1 (2026) 25.
[14] L. Edwards and P. Torcellini, A Literature Review of the Effects of Natural Light on Building Occupants, National Renewable Energy Laboratory: Golden, Colorado, Technical Report NREL/TP-550-30769, Jul. 2002.
[15] Association Suisse des Electriciens, Éclairage intérieur par la lumière du jour, SN 418911, Zürich, 1989.
[16] C. F. Reinhart and O. Walkenhorst, Validation of dynamic RADIANCE-based daylight simulations for a test office with external blinds, Energy and Buildings, 33:7 (2001) 683-697.
[17] Illuminating Engineering Society of North America, Approved Method: IES Spatial Daylight Autonomy (sDA) and Annual Sunlight Exposure (ASE), IES LM-83-12, New York, NY, 2012.
[18] A. Nabil and J. Mardaljevic, Useful daylight illuminance: a new paradigm for assessing daylight in buildings, Lighting Research & Technology, 37:1 (2005) 41-57.
[19] A. Nabil and J. Mardaljevic, Useful daylight illuminances: A replacement for daylight factors, Energy and Buildings, 38:7 (2006) 905-913.
[20] J. Mardaljevic, M. Andersen, N. Roy, and J. Christoffersen, Daylighting Metrics for Residential Buildings, in: Proceedings of the 27th Session of the CIE, Sun City, South Africa, 9-16 July 2011, pp. 93-110.
[21] S. Mokhtar, A. Sojka, and C. C. Davila, Conditional Generative Adversarial Networks for Pedestrian Wind Flow Approximation, in: Proceedings of the 11th Annual Symposium on Simulation for Architecture and Urban Design, Online, 25-27 May 2020, pp. 1-8.
[22] P. Kastner and T. Dogan, A GAN-Based Surrogate Model for Instantaneous Urban Wind Flow Prediction, Building and Environment, 242 (2023) 110384.
[23] C. Huang, G. Zhang, J. Yao, X. Wang, J. K. Calautit, C. Zhao, N. An, and X. Peng, Accelerated environmental performance-driven urban design with generative adversarial network, Building and Environment, 224 (2022) 109575.
[24] Q. He, Z. Li, W. Gao, H. Chen, X. Wu, X. Cheng, and B. Lin, Predictive models for daylight performance of general floorplans based on CNN and GAN: A proof-of-concept study, Building and Environment, 206 (2021) 108346.
[25] P. Li and B. Li, Generating Daylight-driven Architectural Design via Diffusion Models, 20 April 2024, arXiv: 2404.13353. doi: 10.48550/arXiv.2404.13353.
[26] U.S. Green Building Council, LEED v4.1 Building Design and Construction, U.S. Green Building Council: Washington, DC, Jan. 2020.
[27] A. Patricio, A. Dehban, and R. Ventura, FLORA: Efficient Synthetic Data Generation for Object Detection in Low-Data Regimes via finetuning Flux LoRA, 29 August 2025, arXiv: 2508.21712.
[28] Y. Cao, A. Abdul Aziz, and W. N. R. Mohd Arshard, Stable diffusion in architectural design: Closing doors or opening new horizons?, International Journal of Architectural Computing, 23:2 (2025) 339-357.
[29] J. Ploennigs and M. Berger, AI Art in Architecture, AI in Civil Engineering, 2:1 (2023) 8.
[30] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE Transactions on Image Processing, 13:4 (2004) 600-612.
[31] P. Jain and S. Sengar, Unraveling The Role Of IBM SPSS: A Comprehensive Examination Of Usage Patterns, Perceived Benefits, And Challenges In Research Practice, Educational Administration: Theory and Practice, 30:5 (2024) 9523-9530.
[32] J. C. Lee, H. W. Park, and Y. N. Kang, Feasibility study of structural similarity index for patient-specific quality assurance, Journal of Applied Clinical Medical Physics, 26:3 (2025) e14591.
[33] S. Yang and G. Berdine, Confusion matrix, The Southwest Respiratory and Critical Care Chronicles, 12:53 (2024) 75-79.
[34] G. Varoquaux and O. Colliot, Machine Learning for Brain Disorders, Springer US: New York, NY, 2023.

2383-8701/© 2026 The Author(s). Published by solarlits.com. This is an open access article distributed under the terms and conditions of the Creative Commons Attribution 4.0 License.