# AESMOOD: AN INTELLIGENT SYSTEM FOR GENERATING MOOD BOARDS WITH AESTHETIC COMPUTING.

**Rong Rong[a], Shuyi Liu[a], Hanling Zhang*[a]**

a Hunan University, School of Design, China
* jt_hlzhang@hnu.edu.cn

**DOI: 10.63442/SMFX9427**

**KEYWORDS | MOOD BOARD, IMAGE GENERATION, AESTHETIC COMPUTING, DEEP LEARNING**

**ABSTRACT |** In the design process, manually constructing mood boards is a difficult, multi-step task requiring significant time and effort to gather and select images. To improve the efficiency of mood board creation, we introduce AesMood, an intelligent mood board generation system based on image generation and aesthetic computing. AesMood utilises the Stable Diffusion 2 image generation algorithm, which enhances the diffusion model by introducing a cross-attention layer into the model architecture and substantially improves visual fidelity. To mitigate the issue of inconsistent aesthetic quality among generated images, AesMood integrates the Style-specific Art Assessment Network (SAAN). SAAN efficiently extracts both style-specific and generic aesthetic features to evaluate and ensure the high aesthetic quality of the images for mood boards. In the preliminary phase of designing the system, interviews were conducted, which yielded three design requirements: meet user needs, editability, and simplicity of use. Three important functionalities are encoded in the AesMood system, including intelligent mood board generation, aesthetic scoring, and mood board editing. We invited 20 designers to employ our system, and its performance was assessed across four dimensions—usefulness, ease of use, learnability, and satisfaction—using the Likert scale in conjunction with semi-structured interviews. The results demonstrate high satisfaction levels, with participants noting the system's ability to broaden inspiration, stimulate creativity and imagination among designers, and enhance the efficiency of the ideation capture phase.

# 1. Introduction

Mood boards are an important tool for designers in the design process and are a source of inspiration. The construction of mood boards is a complex, multi-step task requiring designers to spend a lot of time and effort gathering and selecting images. Additionally, this process is confronted with the challenge of accessibility to high-quality image resources. Consequently, designing an automated system capable of generating high-quality mood boards is important for optimising the design workflow. Such a system could reduce designers' cognitive load and time demands by adjusting the resource acquisition phase and ensuring a consistent supply of premium visual materials.

In an era where data is not merely a collection of numbers but a rich canvas of human experience, aesthetic computing is redefining the boundaries of data interpretation and representation. Aesthetic computing enables designers and artists to experiment with different identities and forms of expression, exploring how data can shape and reshape design expression and self-perception. Aesthetic computing refers to the application of computational methods by computers to simulate human aesthetic decision-making (Xu, et al, 2022), and it represents a quantitative study of the aesthetic attributes of artistic forms such as images. Aesthetics assessment, as one of the primary research directions in aesthetic computing, principally involves extracting aesthetic perceptual features to assess the aesthetic quality of visual inputs. Prior research has applied aesthetic computing to the domain of photography recommendation (Wu, 2022). However, studies focusing on utilising aesthetic computing for mood board construction assistance are currently absent. Integrating aesthetic computing into an intelligent mood board generation system is anticipated to alleviate the issue of inconsistent aesthetic quality among generated images. The most employed datasets for aesthetics assessment include AVA (Murray, et al, 2012), AADB (Kong, et al, 2016), and CUHK-PQ (Tang, et al, 2013). Given that mood boards are a crucial referential component in the design process for designers and are often characterised by their aesthetic and abstract nature, we have incorporated the Style-specific Art Assessment Network (SAAN) (Yi, et al, 2023), trained on Boldbrush Art Image Dataset (BAID), into our intelligent mood board generation system as an auxiliary tool for designers. The SAAN algorithm effectively extracts specific styles and general aesthetic features to evaluate and ensure the high aesthetic quality of mood board images.

This paper proposes ***AesMood***, an intelligent mood board generation system based on image generation and aesthetic computing. The primary objective of AesMood is to enhance the efficiency of mood board creation by automatically generating images corresponding to specific keywords input by designers and conducting an aesthetics assessment of these images. The image generation algorithm used by AesMood is stable diffusion 2 (Rombach, et al, 2022), which improves visual fidelity significantly by incorporating cross-attention layers into the model structure. Through a series of user studies, we have evaluated AesMood's performance in assisting designers, overall user satisfaction with the system, and the effectiveness of its various functions. The contributions of this work are as follows:

- The mood board intelligent generation system, named AesMood, has been developed, which generates images corresponding to specific keywords input by designers to construct mood boards. The intelligently generated mood boards feature a fixed layout and are endowed with editable capabilities.
- We have creatively integrated an aesthetics assessment algorithm into the AesMood system, referencing public aesthetics during the mood board construction process and assisting designers in their decision-making.
- Experimental results indicate that AesMood performs well in assisting designers with mood board generation, enhancing design efficiency, providing a broader source of inspiration, and stimulating designers' creativity and imagination.

# 2. Related Work

## 2.1 Mood Board

The mood board is a collection of visual images arranged to express the emotional response to a design brief (Garner, et al, 2001). Lucero (2012), through empirical research, identified five main functions that mood boards can serve in the early stages of the design process: framing, aligning, paradoxing, abstracting, and directing. Current research on mood boards primarily focuses on innovation in the form of mood boards. Koch et al. (2020) investigated the innovative use of digital mood boards to enrich designers' creative process by attaching semantic labels to images. Their study introduced SemanticCollage, a digital tool that employs advanced semantic labelling algorithms to assist designers in transforming vague visual concepts into searchable terms, thereby enhancing the understanding and communication of design ideas without disrupting their creative thinking. Zabotto et al. (2019) explored how to use Kansei engineering to connect users and designers, employing automated mood boards to convey emotions and aid clients in analysing ideas for product development. Their paper examined the potential of a Kansei engineering system based on rough set probability statistics, which could link customers' affective words with sustainably collected images online. They proposed a new Kansei engineering process comprising five cycles that capture user opinions at all stages of the design process. Ivanov et al. (2022) introduced the MoodCubes system, a system for rapidly creating and manipulating multimedia content to address the challenges of gathering and combining inspirational materials in the early stages of the creative process. MoodCubes supports designers by deconstructing objects (such as extracting colour palettes), suggesting new materials (such as 3D models, images, and lighting effects), and providing filters to alter the aesthetic of a scene. These studies have expanded creative research on constructing mood boards by incorporating new elements such as text and 3D models, yet they still utilise existing collected materials. Wan et al. (2023) proposed a digital mood board, GANCollage, driven by StyleGAN and supported by a pretrained anime image classifier, which aids designers in organising and understanding these generated ideas in the form of “sticky notes” on the mood board canvas, with a primary application in character design. Our research, however, attempts to explore the automatic generation of mood boards using the stable diffusion 2 image generation algorithm by generating a diverse range of mood board images to support designers in various design fields (e.g., product design and vehicle design).

## 2.2 Image Generation

Image generation algorithms are an artificial intelligence technology that typically relies on sophisticated mathematical models to produce images that appear authentic by emulating the characteristics of real-world imagery. These algorithms have utility across various applications, including artistic creation, game development, film special effects, and aiding designers in conceptualising new product designs. Currently, the mainstream algorithms for image generation include GANs (Generative Adversarial Networks) (Goodfellow, et al, 2020), VAEs (Variational Auto-encoders) (Kingma, et al, 2019), and diffusion models (Ho, et al, 2020). Due to their superior image quality, diffusion models have gradually replaced GANs and have become the most popular image generation models. Among the research based on diffusion models, Avrahami et al. (2022) proposed an innovative method for the local editing of natural images using natural language descriptions and region of interest (ROI) masks. This method combines a pretrained language-image model (CLIP) with a Denoising Diffusion Probabilistic Model (DDPM) to guide the editing process toward the textual prompts provided by the user, ensuring seamless integration of the edited area with the rest of the image. Kim et al. (2022) introduced DiffusionCLIP, a powerful method for text-guided image manipulation using diffusion models, addressing the limitations of GAN inversion methods in diverse real-world image applications. Their work highlighted the full inversion capability of diffusion models, enabling zero-shot image manipulation across unseen domains and diverse content. They introduced a novel noise composition method to facilitate direct multi-attribute manipulation. Additionally, Gu et al. (2022) proposed the vector quantised diffusion (VQ-Diffusion) model for text-to-image generation, which combines vector quantised variational autoencoder (VQ-VAE) with the recently developed conditional variant of DDPM. Their approach aims to overcome unidirectional bias and error accumulation in existing methods, demonstrating significant improvements in text-to-image generation tasks. The VQ-Diffusion model handles more complex scenes and significantly enhances the quality of the synthesised images. Our research employs the stable diffusion 2 algorithm, which enhances the diffusion model by incorporating cross-attention layers into the model structure, thereby significantly improving visual fidelity.

### 2.3 Aesthetics Assessment

Research on aesthetics assessment currently focuses on two main areas: Image Aesthetics Assessment (IAA) and Personalized Image Aesthetics Assessment (PIAA). In Image Aesthetics Assessment, Ke et al. (2023) proposed a novel method to learn image aesthetics from user comments, employing vision-language pre-training to develop multimodal aesthetic representations. He et al. (2022) conducted a comprehensive study on image aesthetics assessment, introducing the TAD66K dataset containing 66K images across 47 themes. They also developed a Theme and Aesthetics Network (TANet), a model that adapts to the image assessment rules of different themes, achieving state-of-the-art results on several datasets. Furthermore, He et al. (2023) proposed An Enhancer for Aesthetics-Oriented Transformers (EAT) to improve the performance of transformers in IAA tasks. EAT utilises a deformable, sparse, and data-dependent attention mechanism, refining attention through offsets to balance foreground and background, thus outperforming previous methods on various datasets.
In the domain of Personalized Image Aesthetics Assessment, Li et al. (2022) addressed the challenges of PIAA by proposing a meta-learning-based algorithm—Transductive Aesthetic Preference Propagation (TAPP-PIAA), which avoids the need for fine-tuning personal data, thereby reducing training costs and preventing underfitting/overfitting. Yang et al. (2022) tackled the challenges of PIAA by introducing a new personalised image aesthetic database—Personalized image Aesthetics database with Rich Attributes (PARA). They presented a conditional PIAA model that uses thematic information as a prior condition, surpassing existing methods and offering insights into the complex interplay of image aesthetics and personal characteristics that generate personalised aesthetic preferences. Considering mood boards' aesthetic and abstract attributes, our research has incorporated the Aesthetics Assessment Algorithm SAAN, trained on the Boldbrush Art Image Dataset (BAID) with 60,337 art images, into our intelligent mood board generation system. This provides designers with a quantified reference to public aesthetics.

## 3. Design of a Mood board Generation System

### 3.1 Pre-study Interview

To comprehensively understand the significance of mood boards to designers, their challenges in creating mood boards, and their needs for an automated mood board generation system, we interviewed seven designers (aged 26-30, with professional experience ranging from six months to six years). The positions of the participants were as follows: interior designer (3), game designer (1), spatial designer (1), fashion designer (1), and CMF designer (1). As the participants were in various cities, the interviews were conducted online and audio-recorded for subsequent analysis. We designed a survey questionnaire comprising eleven questions that covered basic information, the importance of mood boards, experiences and challenges with using mood boards, and the needs related to an automated mood board generation system. The synthesised feedback from the interviews is summarised as follows:

***Importance of Mood Boards.*** All designers concurred on the paramount importance of mood boards in the design process, describing them as perpetually essential. One designer stated, *“You cannot leap from an idea in your head directly to a sketch or a model. There is a chasm in between.”* Designers regard mood boards as a source of inspiration, an intuitive guide, and a visual summary. Regarding communicating with clients, a spatial designer mentioned, *“Mood boards can better help users grasp the general feel of a proposal before any concrete design is in place.”*

***Difficulty in Finding Images.*** All designers identified the most significant challenge in constructing mood boards as difficulty in sourcing images. They often spend considerable time searching online for images that meet their requirements but fail to find suitable ones. "*No matter how you search, you cannot find what you want,*" one designer illustrated, "*For instance, if you have a creative idea that might not exist in the market currently, you actually cannot find corresponding images. I think finding images is a significant challenge, especially when making creative-type mood boards.*" Another fashion designer said, "*The images found may not match the needed tone, requiring modifications to the images, such as removing certain elements, which is time-consuming.*"

***Automated Generation of Mood Boards.*** All designers strongly desired a system that could automatically generate mood boards, saving time seeking out and editing images. One designer stated, *"If it could automatically generate mood boards based on the textual information we have researched, that would significantly reduce our workload."* Regarding the functionality of the automated mood board generation system, one designer believed, *"The simpler the system, the better, focusing on the core function of generating mood boards."* Most designers hoped that the layout of the final mood board could be flexible enough, with added editing capabilities. One designer had specific requirements for the precision of generated images, hoping for a more accurate match of suitable images to the input keywords.

## 3.2 AesMood System Design

Informed by interview feedback, we developed AesMood, an intelligent mood board generation system designed to create mood boards that align with designers' needs. This system encompasses three primary functions: automatic mood board generation, aesthetic scoring, and mood board editing. The entire system was developed using Python. The operational workflow and system framework of the AesMood system are shown in Figure 1.

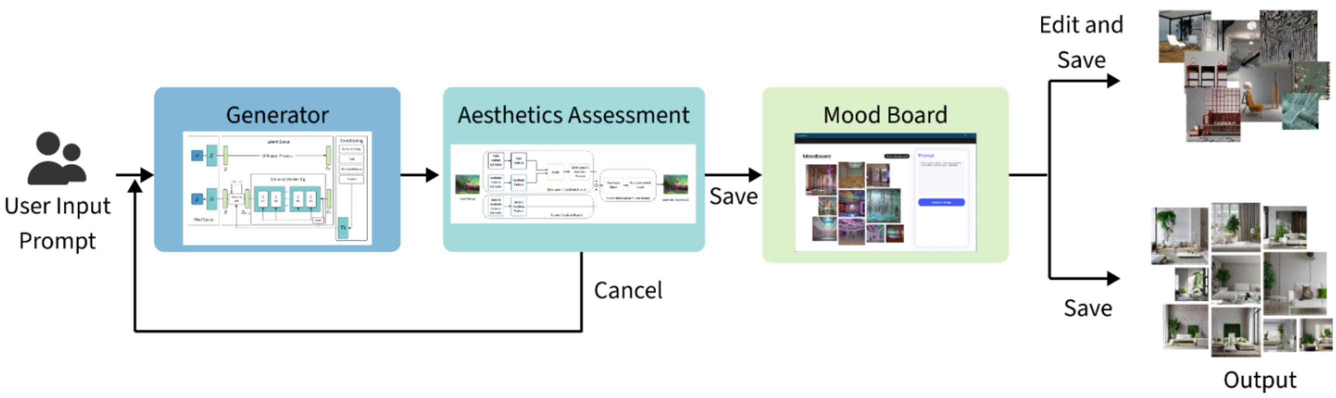


Figure 1. Operational workflow and system framework of AesMood.

Based on the interview feedback, we developed the AesMood to address three core requirements of designers, as shown in Table 1. Beyond following these requirements, to address the issue of inconsistent image quality commonly associated with current image generation algorithms, we incorporated an aesthetic scoring feature into the AesMood system. This feature is intended to assist designers in building mood boards by providing aesthetic evaluations.

Table 1. Design Requirements for Constructing an Intelligent Mood Board Generation System.

| | **Design Requirements** |
|---|---|
| **1)** | Develop a system capable of automatically generating mood boards that more accurately align with the requirements of designers, thus conserving the time they spend searching for images. |
| **2)** | Create an intelligent mood board generation system with image editing capabilities, enabling designers to utilise this system to complete the construction of mood boards directly. |
| **3)** | Design an intelligent mood board generation system with a simple interface and precise interactions, facilitating ease of learning and use for designers. |

Figure 2 presents the main interface of the AesMood system. The left side is a blank space as the Mood Board display area, where designers can modify the size and position of selected images after saving them. On the right is the prompt input area, where designers can enter keywords related to the images they wish to generate into the text box, separating each term with a comma. Upon entering the desired keywords, click the “Generate Image” button to initiate the image generation process. As depicted in Figure 2, we entered a set of keywords “future interior design, renaissance painting installation, digital visual, cyberspace” as an example of usage.

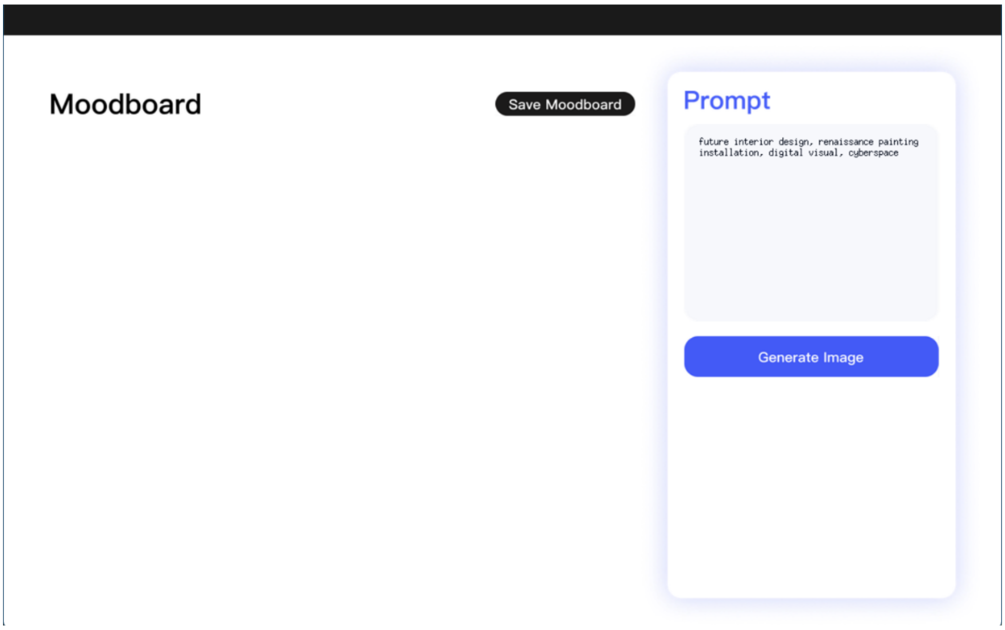


Figure 2. The main interface of the AesMood system.

The interface for image generation using the prompt is depicted in Figure 3. The current image's prompt is displayed at the top of the interface, with the generated image showcased in the centre. Concurrent with the image generation, an aesthetics assessment algorithm calculates an aesthetic score for each image displayed below the image. This aesthetic score offers a quantified evaluation of public aesthetic appeal, serving as a reference for designers when deciding whether to incorporate an image into the mood board. If the designer opts to include the image in the mood board, they select the “Save” button; if not, they choose the “Cancel” button.

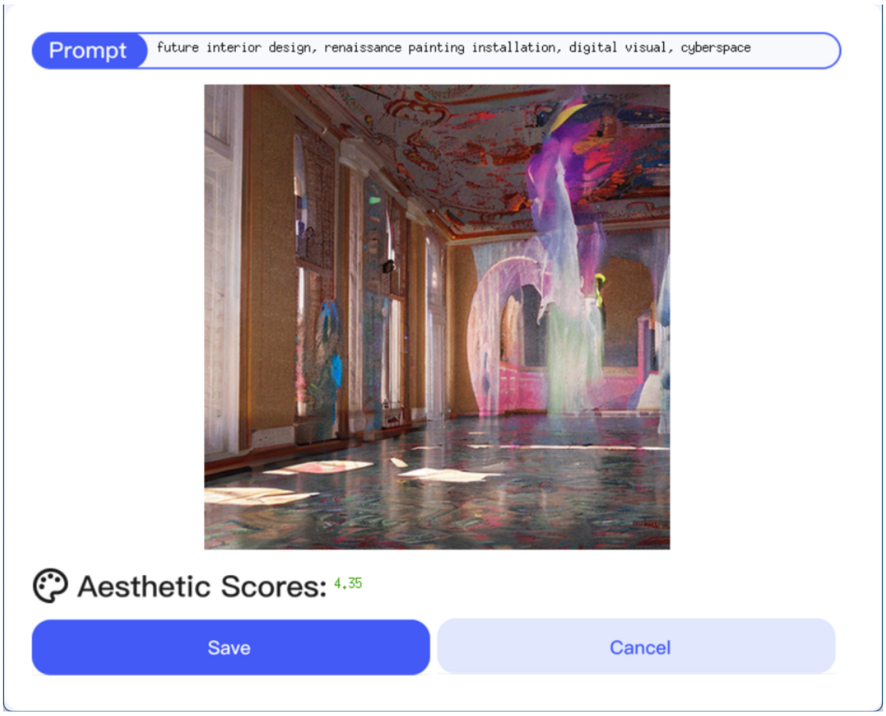


Figure 3. Image generation interface.

As seen in Figure 4, after saving an image, it is directly displayed on the mood board and arranged in a preset sequence with a fixed layout, with a maximum of ten images for the designer's reference in layout planning. If the designers are satisfied with the fixed layout, they can choose to save the mood board directly. If the designers feel adjustments are necessary, the AesMood system also supports editing functionality. The specific operations are as follows: images can be moved using a drag-and-drop action with the mouse's left button, and the size of the images can be adjusted by placing the cursor over the image, holding down the control key, and scrolling the mouse wheel. The designer can flexibly adjust the position and size of the images on the mood board while adding images that meet their requirements until the final composition is determined. The designer can select the “Save Moodboard” button to save the mood board. Additionally, the images generated during the construction of the mood board are saved on the computer, allowing the designer to edit them further using other specialised design software if needed.

## 3.3 Algorithm of AesMood

The image generation algorithm AesMood uses is Stable Diffusion 2, as illustrated in Figure 5. It enhances the diffusion model by incorporating cross-attention layers into the model structure, thereby significantly improving visual fidelity. The differential model learns data distribution by denoising standard distribution variables, effectively reverse-engineering a fixed-length Markov chain. In image synthesis, these models are trained through a series of denoising autoencoders to predict the denoised version of the input. The algorithm's perceptual compression model allows access to a lower-dimensional latent space that is more suitable for likelihood-based generative models. This algorithm uses image-specific inductive biases and 2D convolutional layers to construct a UNet, focusing on the perceptually most relevant parts. Additionally, by incorporating a cross-attention mechanism, the algorithm transforms the differential model into a more flexible conditional image generator capable of handling inputs across various modalities, preprocessing these inputs through domain-specific encoders.

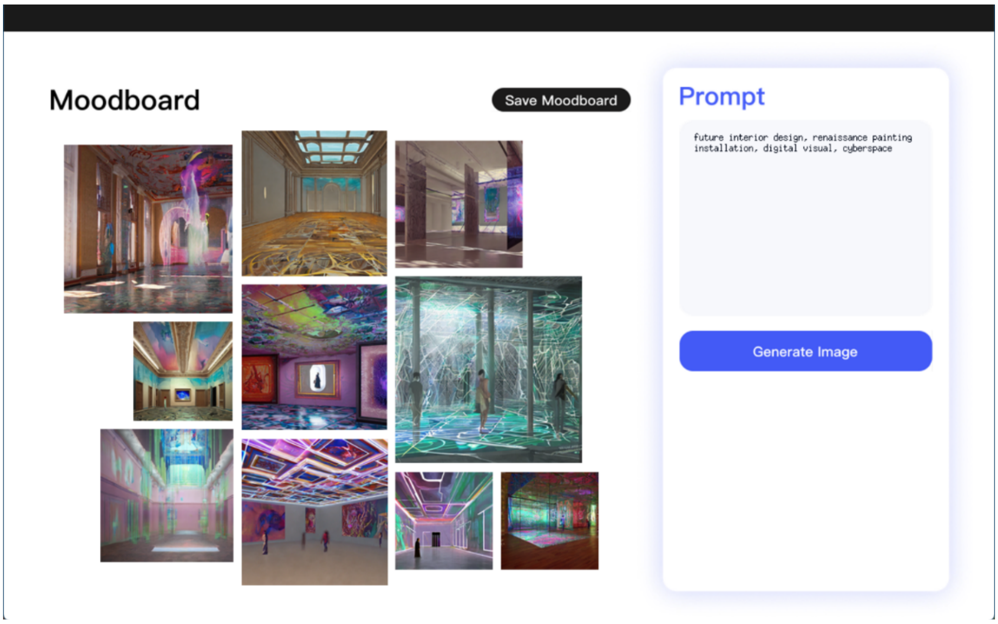


Figure 4. Fixed layout of the mood board generated by AesMood.

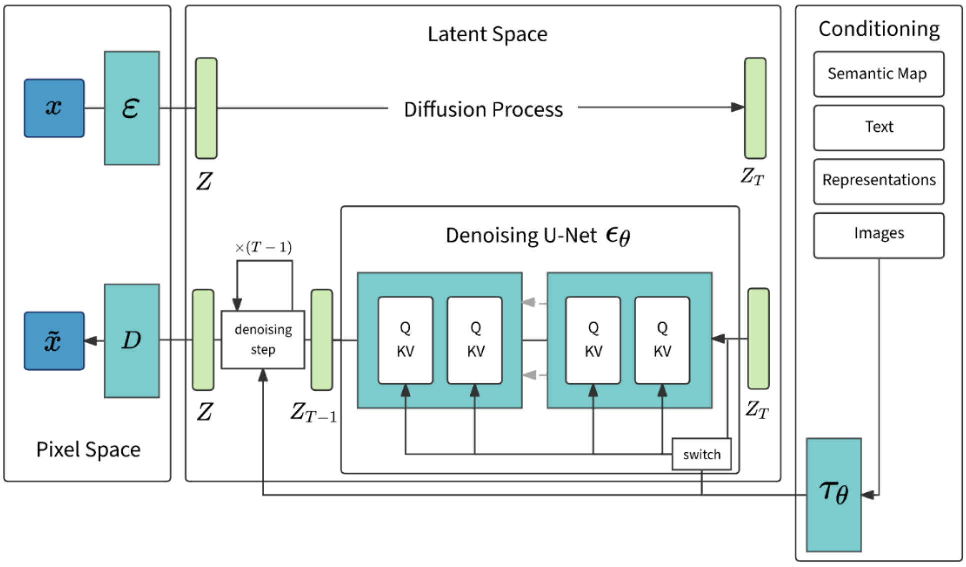


Figure 5. Network structure of the image generation algorithm.

AesMood also introduces an aesthetics assessment algorithm, the Style-specific Art Assessment Network (SAAN), trained on the Boldbrush Artistic Image Dataset (BAID), which comprises 60,337 artistic images. As shown in Figure 6, the SAAN algorithm contains three modules: (1) Style-specific Aesthetic Branch (SAB): This extracts aesthetic features related to artistic style, using a pre-trained VGG-19 to extract style features and a pre-trained ResNet-50 to extract aesthetically relevant features. The AdaIN layer integrates style features into aesthetic features while preserving the spatial structure of style features. (2) Generic Aesthetic Branch (GAB): This branch extracts general aesthetic features based on self-supervised learning, such as the integrity of prominent parts and frame layout. It uses ResNet-50 as the backbone network and applies self-supervised learning methods for pretraining. (3) Spatial-information Fusion Module: This module uses non-local blocks to fuse spatial information and incorporates the composition of artworks into the assessment. It merges the features extracted by SAB and GAB and then uses non-local blocks to fuse spatial information.

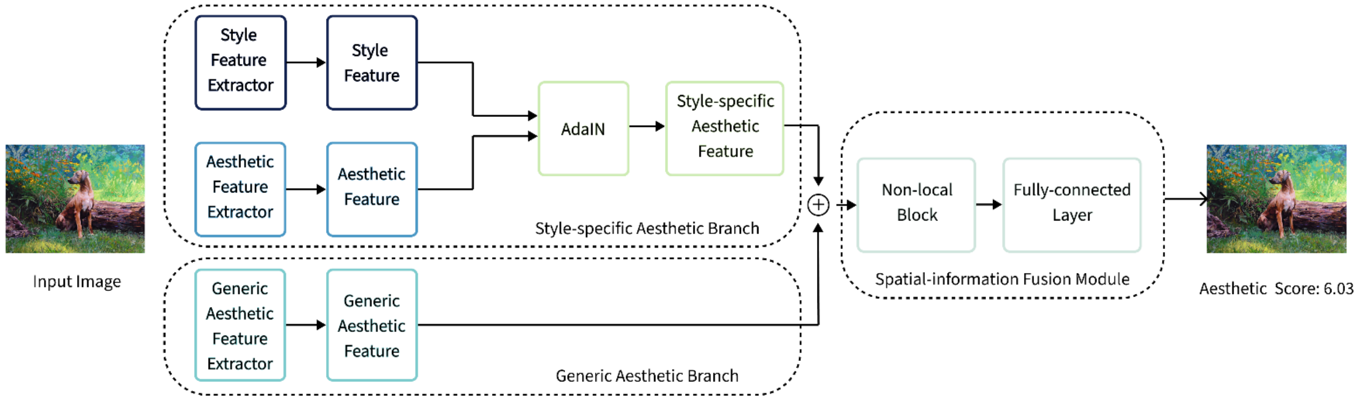


Figure 6. Network structure of the aesthetics assessment algorithm.

# 4. Evaluation

## 4.1 Experimental Setup

To evaluate the performance of the AesMood system in assisting designers and to assess overall user satisfaction and the effectiveness of each system function, we conducted a series of user studies. We recruited 20 participants (age 22-30, M=25, SD=5.94) as the human subjects in this study. All of them were designers (average of 6 years of design experience), specialising in product design, landscape design, interior design, etc. The experiment involved asking designers to conceive designs based on the fixed theme of "smart home" and to create mood boards using the AesMood system. The experiment included three main parts: (1) pre-experiment, (2) creation of mood boards, and (3) post-test.

1. Pre-experiment: Designers initially familiarised themselves with the AesMood system by trying out its main functions and operating procedures. After several adjustments, they preliminary determined the prompts to be used in creating mood boards to minimise the impact of prompt adjustments during the subsequent experimental process.

2. Creation of mood boards: In the formal experimental phase, we set up four control experiments to verify the effectiveness of the functions in the AesMood system and their impact on designers during the mood board creation process. We asked designers to create mood boards using four different versions of the system, each with varying functionalities: AesMood, which includes the full functionalities of image generation, aesthetic scoring, and image editing capabilities (dragging and resizing); AesMood-edit, which provides for only image generation and aesthetic scoring; AesMood-score, which includes only image generation and image editing; and AesMood-edit-score, which provides for only image generation. Designers used these four systems in a random order to create complete mood boards to prevent the order of the experiments from affecting the results.

3. Post-test: We designed a user questionnaire based on a 7-point Likert scale across four dimensions—usefulness, ease of use, learnability, and satisfaction (Lund, 2001)—and asked designers to rate their agreement after each experimental group (1: strongly disagree, 7: strongly agree). Designers were also asked to score the final mood boards they created (on a 10-point scale). Finally, we conducted a semi-structured interview to gain more specific feedback from designers about the AesMood system.

## 4.2 Results

The normality of the Likert scale data was assessed using the Shapiro-Wilk test, which yielded significant results ($p < 0.01$), indicating that the data did not follow a normal distribution. Consequently, we employed the Kruskal-Wallis non-parametric test to determine if there were significant differences between the various systems across the four dimensions of usefulness, ease of use, learnability, and satisfaction. Dunn's post hoc analysis was conducted to further discern specific differences between pairs of systems. The mood boards created by designers using the AesMood system are shown in Figure 7.

Figure 7. The mood boards created by designers using AesMood.

1. Usefulness

As shown in Figure 8, significant variances were detected through the Kruskal-Wallis test in the areas of usefulness ($H=14.89$, $p<0.01$), efficiency ($H=9.00$, $p<0.05$), and requirement satisfaction ($H=23.70$, $p<0.001$). Pairwise comparisons revealed that AesMood significantly outperformed the other three systems in terms of usefulness, with p-values less than 0.01 when compared to AesMood-edit, less than 0.05 when compared to AesMood-score, and less than 0.001 when compared to AesMood-edit-score. This indicates that the functionalities for aesthetic scoring and mood board editing substantially impact the enhancement of system usefulness.

Regarding the fulfilment of needs, AesMood was significantly superior to both AesMood-edit and AesMood-edit-score ($p<0.001$), and AesMood-score also significantly outstripped AesMood-edit and AesMood-edit-score ($p<0.01$), indicating that the capability to edit mood boards is of paramount importance to designers from the perspective of meeting their needs. Concerning the enhancement of mood board creation efficiency, the evaluation (M=5.60, SD=0.50) indicates that designers generally perceive the AesMood system as having improved the efficiency of mood board creation. Furthermore, from the assessment results concerning the acquisition of design inspiration (M=5.85, SD=0.59), it can be inferred that the image generation capabilities of the AesMood system have aided designers in obtaining more design inspiration.

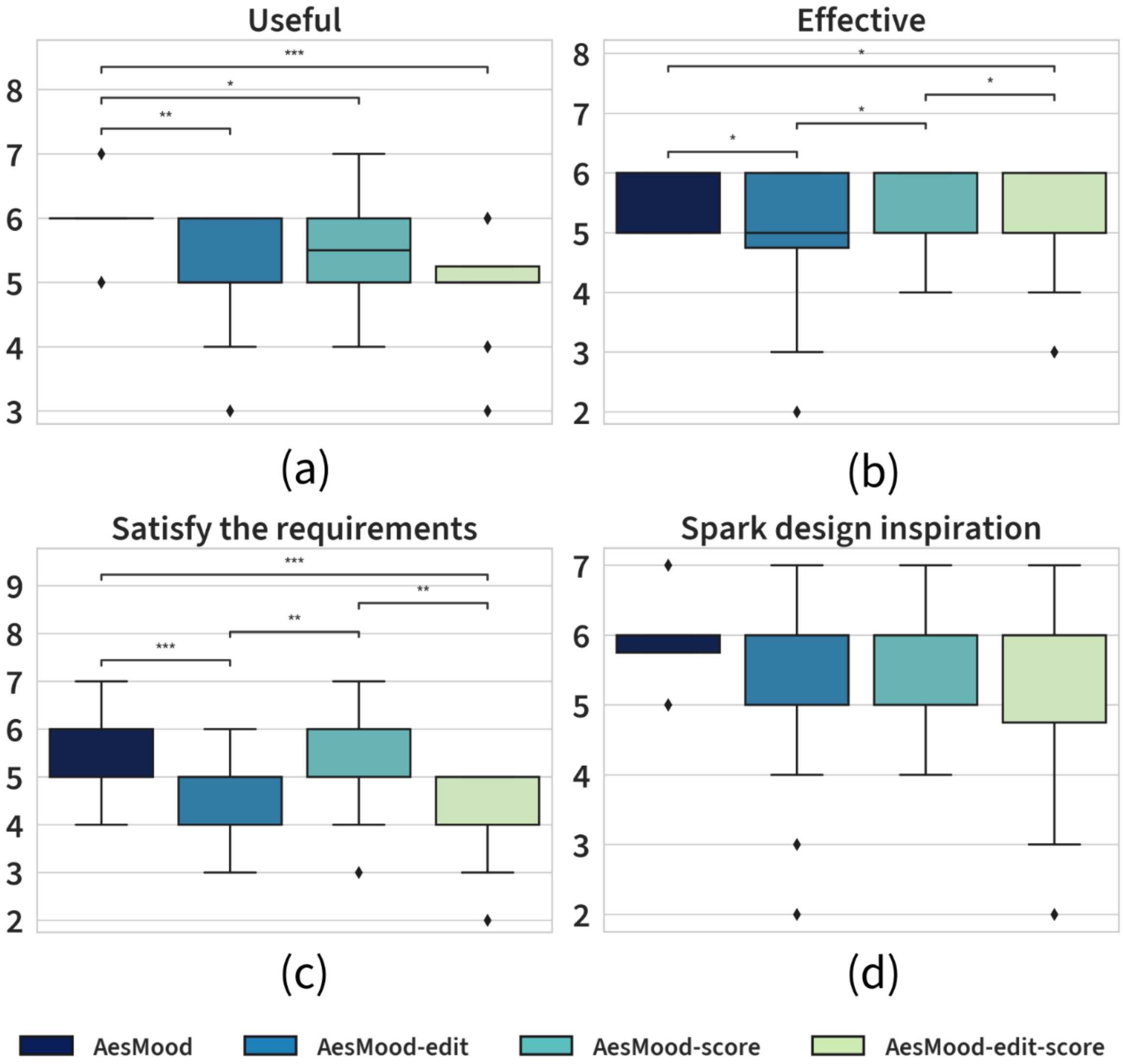


Figure 8. The evaluation results for system usefulness (*: 0.01 < p < 0.05, **: 0.001 < p < 0.01, and ***: p < 0.001). (a) the system's usefulness in facilitating the creation of mood boards; (b) the improvement in efficiency for creating mood boards compared to usual practices; (c) the system's fulfilment of needs concerning the creation of mood boards; and (d) the system's role in augmenting design inspiration.

2. Ease of use

As depicted in Figure 9, there is a statistically significant variance in user-friendliness among different systems (H=8.09, p<0.05), with pairwise comparisons revealing that systems without an editing feature are rated lower in user-friendliness. The AesMood system scores favourably in ease of use (M=6.20, SD=0.62), indicating that it is relatively easy to use, and it also receives a positive rating for user-friendliness (M=6.05, SD=0.83).

3. Learnability

Figure 10 indicates no significant difference in learnability among the various systems. The AesMood system demonstrates commendable learnability in the overall assessment (M=6.40, SD=0.60), and designers can achieve proficiency in utilising the AesMood system with relative swiftness (M=6.55, SD=0.61).

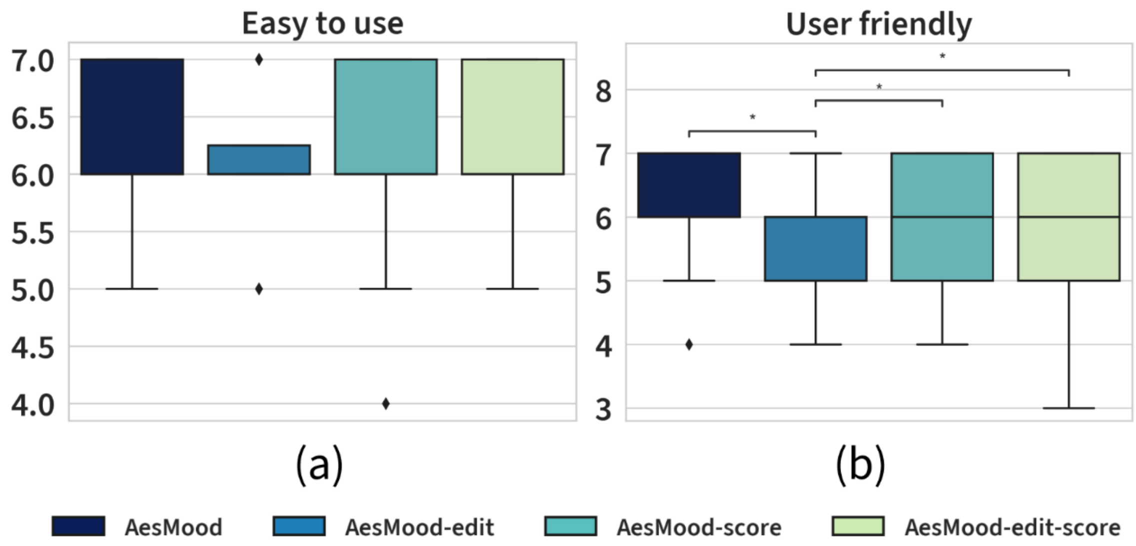


Figure 9: The evaluation results for system ease of use (*: 0.01 < p < 0.05, **: 0.001 < p < 0.01, and ***: p < 0.001). (a) This system is very simple to use. (b) This system is user-friendly.

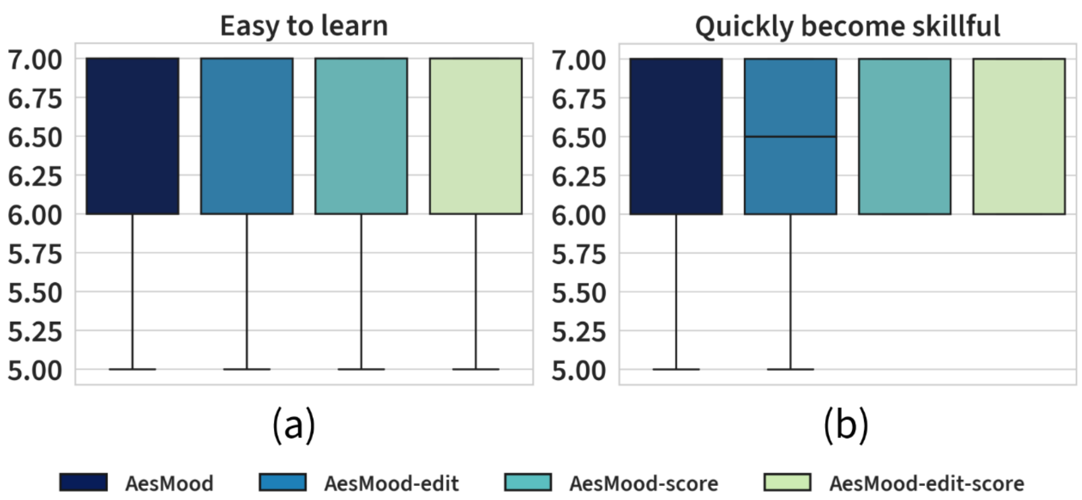


Figure 10. The evaluation results for system learnability. (a) It is easy to learn how to use this system. (b) I was able to become proficient in operating this system quickly.

4. Satisfaction

Figure 11 reveals that there are significant differences in the system satisfaction assessment, specifically in the domains of user enjoyment (H=13.20, P<0.01), overall system satisfaction (H=8.92, P<0.05), willingness to recommend the system to other designer friends (H=10.16, P<0.05), and satisfaction with the mood board (H=12.04, P<0.01). Through pairwise comparisons, it was found that AesMood significantly outperformed both AesMood-edit and AesMood-edit-score in terms of user enjoyment (P<0.05), and AesMood-score was significantly superior to AesMood-edit and AesMood-edit-score (P<0.01), indicating that the addition of mood board editing features can enhance designers' pleasure when using the AesMood system. Regarding the willingness to recommend the system to other designer friends, AesMood was significantly preferred over AesMood-edit ($P<0.05$) and AesMood-edit-score ($P<0.01$), suggesting that designers are more inclined to recommend a system with a more diverse and comprehensive set of features to their peers.

The overall satisfaction with the AesMood system was evaluated from two perspectives: satisfaction with the system and the generated Mood board. The results showed that systems with editing capabilities, AesMood and AesMood-score, were evaluated significantly higher than those without editing features in both system and Mood board satisfaction assessments. Analysing the AesMood evaluation scores, it is noted that AesMood received high ratings in terms of system satisfaction (M=5.65, SD=0.67) and Mood board satisfaction (M=5.30, SD=0.66), indicating that, overall, designers have a favourable satisfaction rating for the AesMood system.

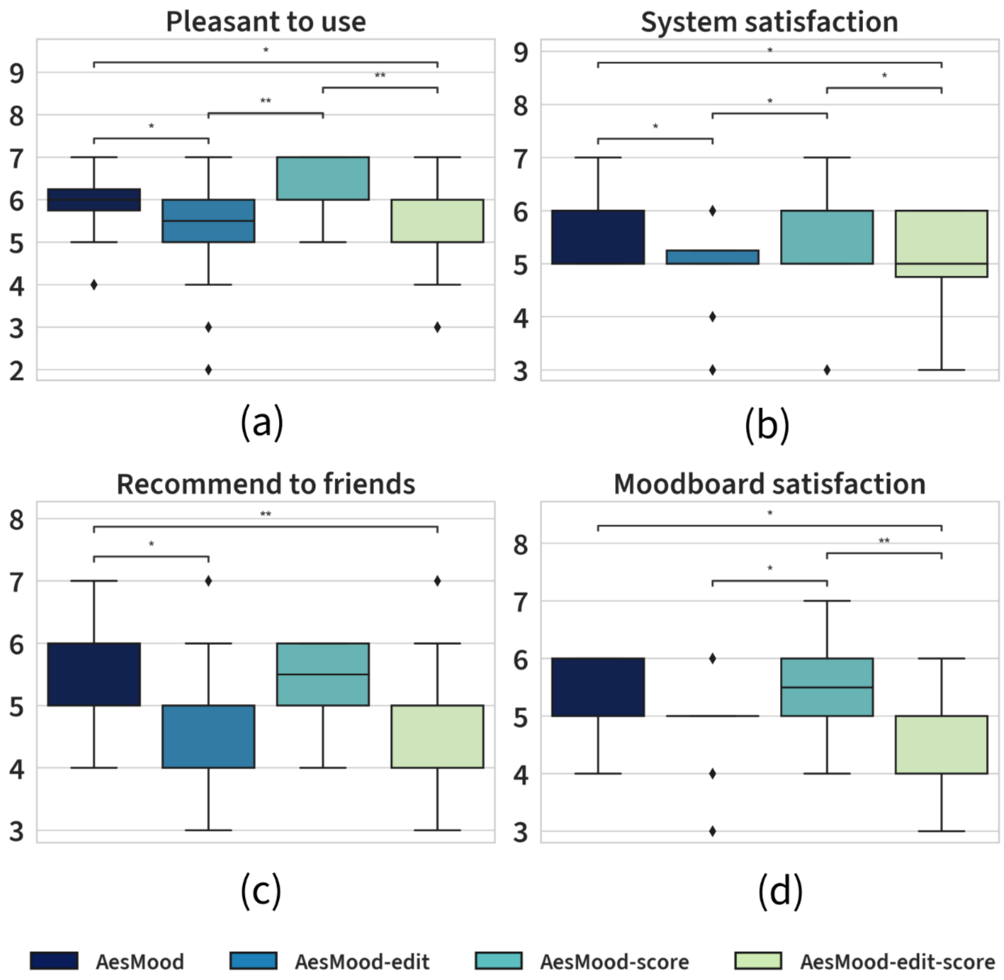


Figure 11. The evaluation results for system satisfaction (*: $0.01 < p < 0.05$, **: $0.001 < p < 0.01$, and ***: $p < 0.001$). (a) Utilising this system is a pleasant experience for me. (b) I am satisfied with this system. (c) I am willing to recommend this system to other designer friends. (d) I am content with the Mood board generated by using this system.

We also collected designers' scores for the final mood boards, as shown in Figure 12, which indicated significant differences between systems ($H=20.48$, $P<0.001$). Pairwise comparisons revealed that the mood board scores for AesMood were significantly higher than those for AesMood-edit and AesMood-edit-score ($P<0.001$), and AesMood-score was also considerably higher than AesMood-edit and AesMood-edit-score ($P<0.01$). This suggests that the mood board editing feature is critical to the designers' evaluation of the final generated mood boards. The AesMood mood board generation system, which has editing capabilities, better responds to the needs of designers and helps them create more satisfactory mood boards. Additionally, the size of the P-values and a comparative analysis of the scores for AesMood (M=7.70, SD=1.03) and AesMood-score (M=7.40, SD=1.00) suggest that the aesthetic scoring can enhance the mood board scores to a certain extent.

Furthermore, we used the Mann-Whitney test to analyse whether the use of other AI image generation tools affected satisfaction and found no significant differences in satisfaction with the system (U=710.50, P>0.05), satisfaction with the mood board (U=742.50, P>0.05), and mood board scores (U=801.00, P>0.05). This indicates that designers who have used other AI image generation tools are also delighted with our designed AesMood system.

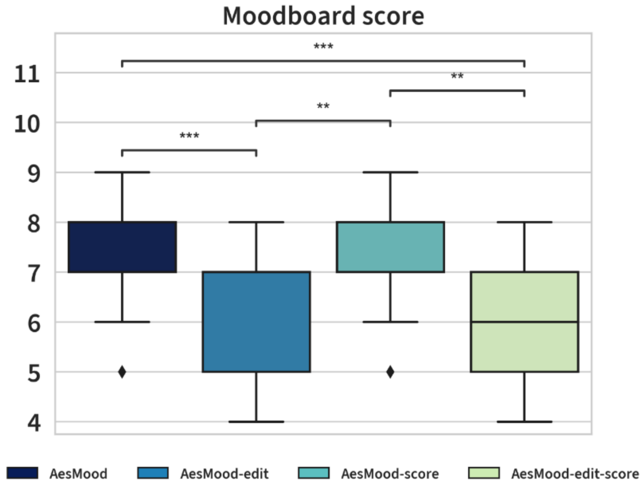


Figure 12. Mood board score (Ten-point scale) (*: $0.01 < p < 0.05$, **: $0.001 < p < 0.01$, and ***: $p < 0.001$).

In the semi-structured interviews, we posed some open-ended questions to understand designers' more specific opinions about the AesMood system. Regarding which system they liked the most, 80% of the participants chose AesMood. At the same time, 75% of the participants least liked the AesMood-edit-score system, indicating that most designers prefer a mood board generation system with more diverse functionalities. We have also compiled some representative responses in Table 2. Most designers considered the aesthetic scoring feature to be of certain assistance in creating mood boards. They believe that “*Aesthetic scoring plays a role in assisting decision-making.*” “*With a high aesthetic score, I feel more confident about the mood board I create.*” Some designers noted no significant impression when first using a system with aesthetic scoring. However, the difference became quite pronounced compared to systems lacking this feature, leading them to favour mood board creation systems with aesthetic scoring. The designers expressed satisfaction with the functionality and interface of AesMood, noting that “*The innovation in generating mood boards is stronger and more aligned with my needs.*” and “*The user interface is simple, and I am quite satisfied overall.*” One designer highlighted that “*The generated images exhibit proficient style integration.*” demonstrating one of the prominent advantages of applying an image generation algorithm to mood board creation. Additionally, the randomness of the image generation algorithm provides designers with a broader source of inspiration, as one designer mentioned, “*The randomness in the generated images is quite interesting and helps to inspire creativity.*”

Designers have offered several suggestions to address the shortcomings of AesMood. Regarding aesthetic scoring, they propose that “*Aesthetic scoring based on the client's preferences is more referential.*” This means that in the design process, by allowing the aesthetics assessment algorithm to learn the client's aesthetic preferences in advance, the system's aesthetic scoring suggestions would align more closely with the client's tastes, thereby better fulfilling the client's needs. Regarding system functionality improvements, designers have indicated a desire for “*some help available when entering prompts.*” During our experiments, we observed that designers with more experience could easily conceive the keywords for their prompts, quickly obtaining the desired mood board images. In contrast, other designers required more time to consider or search for relevant keywords using search engines. Therefore, incorporating prompt suggestions into AesMood would significantly enhance user-friendliness.

Table 2. Examples of User Responses from Semi-Structured Interviews.

| Cluster | Notes |
|---|---|
| **Aesthetic Scoring** | "Aesthetic scoring plays a role in assisting decision-making." "With a high aesthetic score, I feel more confident about the mood board I create." |
| **System Functionality** | "The innovation in generating mood boards is stronger and more aligned with my needs." "Once I have the images for the mood board, I can layout directly, which is very convenient." "The generated images exhibit proficient style integration." |
| **System Interface** | "The user interface is simple, and I am quite satisfied overall." |
| **Inspiration Stimulation** | "The randomness in the generated images is quite interesting and helps to inspire creativity." |
| **Improvements in Aesthetic Scoring** | "Aesthetic scoring based on the client's preferences is more referential." |
| **Improvements in System Functionality** | "I hope there could be some help available when entering prompts." |

## 5. Discussion

We designed the AesMood intelligent mood board generation system based on user needs derived from preliminary interviews. This system enhances the efficiency of mood board creation by automatically generating images corresponding to specific keywords input by designers. Integrating an aesthetics assessment algorithm into the AesMood system allows designers to reference public aesthetics during the mood board construction process. Feedback from designers indicates that AesMood increases the efficiency of mood board creation, a development that saves time and introduces new dynamic elements into the creative process. Overall, our proposed AesMood system can allow designers to reference public aesthetics, increase creative efficiency, and provide a broader range of inspiration. Additionally, the AesMood system excels at integrating styles, which can stimulate designers' imagination and creativity.

From both quantitative analysis and qualitative feedback, we observe that designers perceive aesthetic scoring as a supportive tool for decision-making, which can enhance their confidence while creating mood boards. Furthermore, including the aesthetic scoring feature can, to some extent, elevate the designers' final appraisal of their mood boards. However, aesthetic scoring also presents an issue of insufficient interpretability. As Wu (2020) has indicated, the issue of insufficient interpretability accompanies aesthetic scoring, which more closely resembles a black box that directly maps inputs to aesthetic scores. In practical application, designers may be more concerned with understanding how aesthetic scores are calculated and how to obtain higher aesthetic scores. Future work could consider incorporating suggestions on enhancing aesthetic scores and guiding designers to input more appropriate prompts to improve the aesthetic quality of the generated images.

Our experiments show that AI-based image generation algorithms possess an element of randomness. This randomness can aid designers by expanding their creative boundaries and offering a broader source of inspiration. However, it can also lead to issues of excessive divergence. Such divergence may prevent designers from constructing mood boards within a more defined scope. The algorithms perform notably well in aesthetic-related directions but exhibit certain limitations in displaying design products and specific product functions. Future work could contemplate fine-tuning the image generation algorithms on datasets related to design products and specific product functions to compensate for the current deficiencies.

Designers' unfamiliarity with prompt writing can affect their use of the AesMood system, necessitating additional tools to help designers better translate their requirements into suitable and accurate prompts.

In this context, the SemanticCollage digital tool (Koch, et al, 2020) serves as a pertinent example, employing advanced semantic tagging algorithms to aid designers in converting vague visual concepts into searchable terms. Future work could integrate prompt suggestions into the AesMood system to facilitate more precise prompt descriptions by designers, ensuring that the generated images align more closely with their needs.

In exploring the development of automated design assistance tools, it is essential to reflect deeply on their ethical implications for design practice. The design process is not merely an output of functionality but also an integral part of a designer's personal expression and experience, characterised by exploratory, intuitive, and emotional investment. Although automated design tools can significantly enhance efficiency, they may also deprive designers of intuitive responses and emotional experiences during the creative process, potentially leading to a reduced depth of understanding of design materials and elements. Zhang et al. have determined in their research that using artificial intelligence in the design process can induce an illusion of success in human designers, contributing to their complacency. The study states, "*Once human designers follow AI suggestions, they give up the opportunity to explore the design space by themselves.*" (Zhang, et al, 2021, p.20) Therefore, in the development of automated design assistance tools, consideration should be given to how to preserve the designer's creative involvement while increasing efficiency. Future work could enhance interactivity by learning the personal aesthetics of designers and catering to their individualised customisation needs.

Our work also has limitations: The aesthetics assessment algorithm we selected is somewhat limited, as it was trained on a limited artistic image dataset and may provide lower scores for complex design products and other types of images. Furthermore, due to constraints in model size and computational resources, we chose the Stable Diffusion 2 algorithm for image generation.  However, if computational resources permit, selecting algorithms like Stable Diffusion XL could generate images of higher quality and clarity.

In future work, the AesMood system is anticipated to incorporate many innovative features, such as multimodal creation capabilities. Beyond images, it aims to automatically generate text, 3D models, sound, video, and other components constituting novel mood board forms (Ivanov, et al, 2022). Additionally, we plan to introduce real-time collaboration tools for use by multiple designers simultaneously, which is expected to enhance their workflow and creative expression. It is also worth considering training the aesthetics assessment algorithm on different aesthetic evaluation datasets to see if this leads to higher satisfaction with the aesthetic scoring feature among designers. Future research could further explore how to refine the algorithms to accommodate a broader range of design contexts and personal preferences. Additionally, considering AI algorithms cannot understand overly abstract design descriptions, future work could focus on developing a design corpus and optimising image generation algorithms to meet specific design needs more accurately.

## 6. Conclusion

In this paper, we design, construct, and evaluate AesMood, an intelligent mood board generation system based on image generation and aesthetic computing. Utilising the Stable Diffusion 2 image generation algorithm and the SAAN aesthetics assessment algorithm, AesMood automatically generates images corresponding to specific prompts by designers and references public aesthetics to assist designers in constructing mood boards and facilitating their judgment. User studies indicate that AesMood performs commendably in aiding designers with creating mood boards, enhancing design efficiency, broadening the scope of inspiration, and stimulating the designers' creativity and imagination. We aspire for our work to contribute to AI-assisted design, and future work will aim to refine the system further to accommodate the increasingly diverse needs of designers.

## References

Avrahami, O., Lischinski, D., & Fried, O. (2022). Blended diffusion for text-driven editing of natural images. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 18187–18197. https://doi.org/10.1109/CVPR52688.2022.01767

Garner, S., & McDonagh-Philp, D. (2001). Problem interpretation and resolution via visual stimuli: The use of 'mood boards' in design education. *Journal of Art & Design Education, 20*(1), 57–64. https://doi.org/10.1111/1468-5949.00250

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM, 63*(11), 139–144. https://doi.org/10.1145/3422622

Gu, S., Chen, D., Bao, J., Wen, F., Zhang, B., Chen, D., Yuan, L., & Guo, B. (2022). Vector quantized diffusion model for text-to-image synthesis. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 10686–10696. https://doi.org/10.1109/CVPR52688.2022.01043

He, S., Ming, A., Zheng, S., Zhong, H., & Ma, H. (2023). EAT: An enhancer for aesthetics-oriented transformers. *Proceedings of the 31st ACM International Conference on Multimedia*, 1023–1032. https://doi.org/10.1145/3581783.3611881

He, S., Zhang, Y., Xie, R., Jiang, D., & Ming, A. (2022). Rethinking image aesthetics assessment: Models, datasets and benchmarks. *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*, 942–948. https://doi.org/10.24963/ijcai.2022/132

Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems, 33*, 6840–6851.

Ivanov, A., Ledo, D., Grossman, T., Fitzmaurice, G., & Anderson, F. (2022). MoodCubes: Immersive spaces for collecting, discovering and envisioning inspiration materials. *Designing Interactive Systems Conference*, 189–203. https://doi.org/10.1145/3532106.3533565

Ke, J., Ye, K., Yu, J., Wu, Y., Milanfar, P., & Yang, F. (2023). VILA: Learning image aesthetics from user comments with vision-language pretraining. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 10041–10051. https://doi.org/10.1109/CVPR52729.2023.00968

Kim, G., Kwon, T., & Ye, J. C. (2022). DiffusionCLIP: Text-guided diffusion models for robust image manipulation. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2416–2425. https://doi.org/10.1109/CVPR52688.2022.00246

Kingma, D. P., & Welling, M. (2019). An introduction to variational autoencoders. *Foundations and Trends® in Machine Learning, 12*(4), 307–392. https://doi.org/10.1561/2200000056

Koch, J., Taffin, N., Lucero, A., & Mackay, W. E. (2020). SemanticCollage: Enriching digital mood board design with semantic labels. *Proceedings of the 2020 ACM Designing Interactive Systems Conference*, 407–418. https://doi.org/10.1145/3357236.3395494

Kong, S., Shen, X., Lin, Z., Mech, R., & Fowlkes, C. (2016). Photo aesthetics ranking network with attributes and content adaptation. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), *Computer vision – ECCV 2016* (Vol. 9905, pp. 662–679). Springer International Publishing. https://doi.org/10.1007/978-3-319-46448-0_40

Li, Y., Yang, Y., Li, H., Chen, H., Xu, L., Li, L., Li, Y., & Guo, Y. (2022). Transductive aesthetic preference propagation for personalized image aesthetics assessment. *Proceedings of the 30th ACM International Conference on Multimedia*, 896–904. https://doi.org/10.1145/3503161.3548244

Lucero, A. (2012). Framing, aligning, paradoxing, abstracting, and directing: How design mood boards work. In *Proceedings of the designing interactive systems conference* (pp. 438–447).

Lund, A. M. (2001). Measuring usability with the USE questionnaire. *Usability Interface, 8*(2), 3–6.

Murray, N., Marchesotti, L., & Perronnin, F. (2012). AVA: A large-scale database for aesthetic visual analysis. *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2408–2415. https://doi.org/10.1109/CVPR.2012.6247954

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 10674–10685. https://doi.org/10.1109/CVPR52688.2022.01042

Tang, X., Luo, W., & Wang, X. (2013). Content-based photo quality assessment. *IEEE Transactions on Multimedia, 15*(8), 1930–1943. https://doi.org/10.1109/TMM.2013.2269899

Wan, Q., & Lu, Z. (2023). GANCollage: A GAN-driven digital mood board to facilitate ideation in creativity support. *Proceedings of the 2023 ACM Designing Interactive Systems Conference*, 136–146. https://doi.org/10.1145/3563657.3596072

Wu, X. (2022). Interpretable aesthetic analysis model for intelligent photography guidance systems. *27th International Conference on Intelligent User Interfaces*, 661–671. https://doi.org/10.1145/3490099.3511155

Yang, Y., Xu, L., Li, L., Qie, N., Li, Y., Zhang, P., & Guo, Y. (2022). Personalized image aesthetics assessment with rich attributes. *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 19829–19837. https://doi.org/10.1109/CVPR52688.2022.01924

Yi, R., Tian, H., Gu, Z., Lai, Y.-K., & Rosin, P. L. (2023). Towards artistic image aesthetics assessment: A large-scale dataset and a new method. *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 22388–22397. https://doi.org/10.1109/CVPR52729.2023.02144

Xu, Q., Wu, Q., & Gong, W. (2022). 基于美学计算的智能设计方法分析与启示 [Analysis and enlightenment of smart design methods based on aesthetic computing]. *装饰, 09*, 17–22. https://doi.org/10.16272/j.cnki.cn11-1392/j.2022.09.031

Zabotto, C. N., Sergio Luis da, S., Amaral, D. C., Janaina Mascarenhas Hornos, C., & Benze, B. G. (2019). Automatic digital mood boards to connect users and designers with kansei engineering. *International Journal of Industrial Ergonomics, 74*, 102829. https://doi.org/10.1016/j.ergon.2019.102829

Zhang, G., Raina, A., Cagan, J., & McComb, C. (2021). A cautionary tale about the impact of AI on human design teams. *Design Studies, 72*, 100990. https://doi.org/10.1016/j.destud.2021.100990

**About the Authors:**

**Rong Rong** is a PhD student in Design at Hunan University. Her research interests are focused on human-computer interaction and artificial intelligence, with a current emphasis on aesthetic computing.

**Shuyi Liu** is a master's student in Design at Hunan University. She holds a bachelor's degree in Industrial Design from the same institution. Her research interests are in human-computer interaction and emotional regulation.

**Hanling Zhang** is a professor at Hunan University. He has led over ten national and provincial projects and published over 60 papers in international conferences and journals, including TIP, CVPR, and ECCV. His research interests include human-computer interaction, user experience intelligence evaluation, and the analysis of human factors data.

**Acknowledgements:** We would like to thank all the study participants and the reviewers for their valuable comments. This work was supported in part by the National Key R&D Program of China(2021YFF0900602), in part by the Key R&D Program of Hunan(2022SK2104), in part by the Leading plan for scientific and technological innovation of high-tech industries of Hunan(2022GK4010).

# P/REFERENCES OF DESIGN

**This contribution was presented at Cumulus Budapest 2024: P/References of Design conference, hosted by the Moholy-Nagy University of Art and Design Budapest, Hungary between May 15-17, 2024.**

**Conference Website**
cumulusbudapest2024.mome.hu

**Conference Tracks**
Centres and Peripheries
Converging Bodies of Knowledge
Redefining Data Boundaries
Bridging Design and Economics
Speculative Perspectives
The Power of Immersion
The Future of Well-being
Taming Entropy: Systems Design for Climate and Change
Ways of Living Together
Cumulus PhD Network

**Full Conference Proceedings**
https://cumulusbudapest2024.mome.hu/proceedings

ISBN Volume 1: 978-952-7549-02-5 (PDF)
ISBN Volume 2: 978-952-7549-03-2 (PDF)

DOI Volume 1: https://doi.org/10.63442/IZUP8898
DOI Volume 2: https://doi.org/10.63442/TADX4016

**Conference Organisers**
Moholy-Nagy University of Art and Design Budapest (MOME)
mome.hu
Cumulus Association
cumulusassociation.org