

GenColor: Generative Color-Concept Association in Visual Design

Yihan Hou

yhoud073@connect.hkust-gz.edu.cn
The Hong Kong University of Science
and Technology (Guangzhou)
Guangzhou, China

Manling Yang

myang838@connect.hkust-gz.edu.cn
The Hong Kong University of Science
and Technology (Guangzhou)
Guangzhou, China

Xingchen Zeng

xzeng159@connect.hkust-gz.edu.cn
The Hong Kong University of Science
and Technology (Guangzhou)
Guangzhou, China

Xiaojiao Chen

chenxiaojiao@zju.edu.cn
Zhejiang University
Hangzhou, China

Yusong Wang

ywang677@connect.hkust-gz.edu.cn
The Hong Kong University of Science
and Technology (Guangzhou)
Guangzhou, China

Wei Zeng*

The Hong Kong University of Science
and Technology (Guangzhou)
Guangzhou, China
The Hong Kong University of Science
and Technology
Hong Kong SAR, China

Abstract

Existing approaches for color-concept association typically rely on query-based image referencing, and color extraction from image references. However, these approaches are effective only for common concepts, and are vulnerable to unstable image referencing and varying image conditions. Our formative study with designers underscores the need for primary-accent color compositions and context-dependent colors (e.g., ‘clear’ vs. ‘polluted’ sky) in design. In response, we introduce a generative approach for mining semantically resonant colors leveraging images generated by text-to-image models. Our insight is that contemporary text-to-image models can resemble visual patterns from large-scale real-world data. The framework comprises three stages: *concept instancing* produces generative samples using diffusion models, *text-guided image segmentation* identifies concept-relevant regions within the image, and *color association* extracts primary accompanied by accent colors. Quantitative comparisons with expert designs validate our approach’s effectiveness, and we demonstrate the applicability through cases in various design scenarios and a gallery.

CCS Concepts

- Human-centered computing; • Computing methodologies
→ Artificial intelligence;

Keywords

Color-concept association, Visual design, Generative AI

*Wei Zeng is the corresponding author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI '25, Yokohama, Japan

© 2025 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-1394-1/25/04
<https://doi.org/10.1145/3706598.3713418>

ACM Reference Format:

Yihan Hou, Xingchen Zeng, Yusong Wang, Manling Yang, Xiaojiao Chen, and Wei Zeng. 2025. GenColor: Generative Color-Concept Association in Visual Design. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26-May 1, 2025, Yokohama, Japan*. ACM, New York, NY, USA, 19 pages. <https://doi.org/10.1145/3706598.3713418>

1 Introduction

Color is a fundamental element in visual design, serving as a powerful approach to convey meaning and represent specific concepts [9, 23]. Many colors have well-established associations with particular semantics, enabling designers to enhance visual communication [77]. These associations are valuable in domains such as data visualization [20, 35, 67] and graphic design [18, 23, 69], where color associations are essential to represent categories or themes. For instance, as illustrated in Figure 1 (A), a set of tourist pins for different cities can use colors linked to landmarks, such as red for the Golden Gate Bridge or sandy yellow for the Pyramids. Similarly, in data visualization, these concept-associated colors help categorize information effectively, enhancing clarity and engagement. However, selecting the satisfied color can be laborious, especially when designers are dealing with unfamiliar concepts or striving to align colors with specific emotions or contextual conditions. For instance, when selecting colors for a group of historical landmarks, designers might generally infer the desired tone for the “*Statue of Liberty*” based on common sense (e.g., green or blue for the widely recognized appearance). Careful selection is crucial to ensure that the chosen colors not only reflect the intended concept (such as the “*Statue of Liberty*”) but also evoke the desired context of emotions (e.g., “*bringing light and hope*”). The process remains challenging, particularly for novice designers.

This motivated the development of various methods to facilitate color-concept associations by providing external evidence and references to support designers’ color design decisions [9, 18]. Crowdsourcing efforts [44] leverage collective intelligence to determine color associations; however, these methods can be time-consuming and resource-intensive. In response, computational approaches quantifying color-concept associations have emerged. These methods typically involve either mining color concept co-occurrences from large text corpora or identifying colors from image references

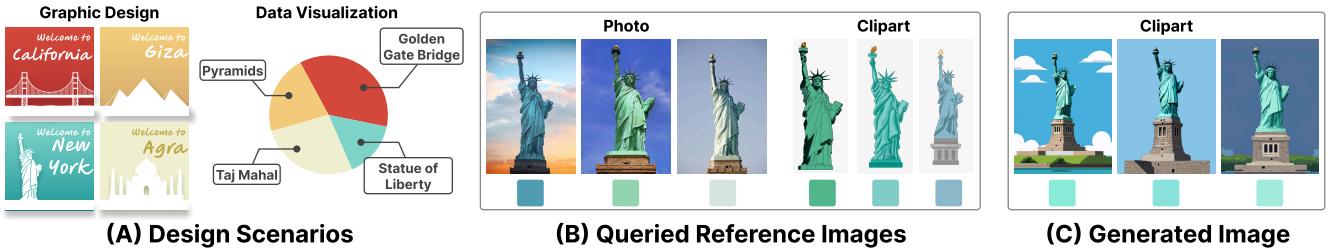


Figure 1: Design scenarios where appropriate color-concept associations are effective to enhance visual communication: (A) graphic design and data visualization. These scenarios require color-concept association, which is preferable in generated images. When comparing the queried and generated images for the “Statue of Liberty”, (B) colors extracted from the queried images (both photos and clipart) show significant variations, whereas (C) the generated images provide a more consistent color representation.

retrieved from image databases [20, 35, 57, 67]. Notably, this approach mirrors conventional practices used by designers who often search for color-referencing images from online resources.

However, these methods encounter significant limitations from various perspectives. First, the linguistic methods in the initial stage are limited to common concepts and may find it challenging to capture context-dependent associations, as N-gram may be sparse for less frequent concepts [57]. For instance, while the concept of the “Statue of Liberty” is popular, the specific phrase “Statue of Liberty with the feeling of light and hope” could be rare. Secondly, obtaining representative images with consistent color distributions from image databases for a given concept can be challenging. Queried images can display significant visual discrepancies due to various external factors such as lighting conditions, image quality, and stylistic variations. For instance, photos of the statue retrieved from Google Image Search exhibit notable differences, as illustrated in Figure 1. Some studies [35, 67] attempt to refine the query results by incorporating terms like “clipart”, yet the outcomes still show considerable variations. These complex color disparities in the queried results make it challenging to determine the most representative color for the “Statue of Liberty”. A flexible and robust computational approach is needed for mining color-concept associations.

We present a novel framework, *GenColor*, which utilizes a generative approach to extract semantically meaningful colors from images produced by text-to-image (T2I) models. By providing designers with a flexible and supportive framework, *GenColor* enables them to explore color concepts and select contextually appropriate color palettes, serving as a reference for their design decisions. Our work is motivated by the insight that contemporary T2I models are trained on large-scale real-world data, and the generated high-quality images can resemble visual patterns in the real world [8, 33, 50]. As such, T2I models can serve as a data mining tool to uncover real-world visual patterns [72]. The design of *GenColor* is informed by design goals from a preliminary study on designers’ workflows and their expectations for a supportive tool for color-concept association (Sect. 3). The *GenColor* framework is implemented in three stages: 1) *Conceptual Instancing* (Sect. 4.1) flexibly generates representative image samples of the corresponding context-dependent concept through the stable diffusion model,

2) *Text-guided Image Segmentation* (Sect. 4.2) identifies semantic-related regions within the image, and 3) *Color Association* (Sect. 4.3) extracts primary-accent color compositions.

To evaluate the effectiveness of *GenColor*, we collect a designer baseline dataset for concept coloring of 36 common concepts (Sect. 5.1), and conduct a quantitative comparison with existing query-based approaches (Sect. 5.2). The results show that *GenColor* produces color compositions that closely align with the designers’ choices and outperform query-based approaches in terms of representativeness and designer preference. This indicates that the generative framework offers two primary advantages over retrieval-based approaches: 1) *feasibility and flexibility* in generating representative images for context-dependent concepts; and 2) *robustness and effectiveness* in managing variations in the generated images, such as controlling lighting conditions and style. We also demonstrate the application scenario of the *GenColor* in tasks such as identifying colors for given concepts and utilizing the identified colors to color clipart, highlighting its potential to support designers in real-world scenarios (Sect. 5.3). A gallery of color-concept associations has been developed and will be made publicly available to facilitate research in this direction (Sect. 5.4).

In summary, our work makes the following contributions:

- We identify and distill design considerations of associating color with concepts, emphasizing the need for context-dependent concepts and primary-accent color composition.
- We propose *GenColor*, a novel framework that leverages generative models for obtaining representative image samples and enabling robust post-process for handling image variation.
- We collect a coloring dataset from professional designers, and the effectiveness of the *GenColor* is evaluated through quantitative comparison with existing query-based approaches, showing the potential of the *GenColor* in supporting designers in real-world design tasks.

2 Related Work

2.1 Color in Visual Design

Color is a fundamental element in visual design, with applications in various domains, including graphic design [27, 56, 69], visualization design [67, 70], and interior design [19, 34]. Effectively using

colors is essential for conveying information, evoking emotions, and enhancing appeal. Various previous literatures have focused on color perception, including contrast [63], harmony [34, 68], and psychological impacts [30, 52]. Adobe has integrated these perception guidances into their Spectrum color system to enhance usability and accessibility [1]. Besides, color semantics, which refers to the way people associate colors with specific concepts, plays a crucial role in enhancing communication beyond effectiveness and aesthetics [35, 67, 70], such as “red” with “apples” or “blue” with the “sky”. Thoughtful application of color semantics can enhance the conveyance of meaning and emotion, enhancing communication beyond mere aesthetics. For example, using semantic-resonant colors can improve cognitive performance in interpreting information visualization, allowing users to understand data categories without constantly referencing a color legend [35]. Moreover, these colors can convey abstract concepts like style or theme, aiding in the communication of the overall ambiance of a design. As such, color semantic mapping is a prevalent strategy in various design disciplines like industrial design [29, 32] and dashboard design [70].

Conventional approaches to associating color semantics rely on designers’ expertise and personal experience, which may be influenced by cultural background, personal preferences, and past experiences [31]. However, clients may have different preferences for color-concept associations compared to designers [47]. There is a growing need to provide designers with references on color-concept associations [18]. To address this gap, computational approaches for mapping color-concept associations have emerged, including crowdsourcing efforts like the XKCD color survey [44] and data mining from large image or language databases [67]. In this study, we introduce a novel approach that leverages generative models, instead of searching for references, to investigate color-concept associations. Drawing inspiration from generative models trained on extensive real-world image datasets, our approach can capture various color concepts, align well with designer perceptions, and offer designers reference colors for diverse concepts.

2.2 Computational Approaches for Color-Concept Association

Associating colors with concepts traditionally relies on human ratings, where participants are tasked with selecting colors that best match given concepts [52, 79], providing names for given colors [44], or assessing the strength of these associations [24, 65, 74]. While effective, these methods are time-consuming, labor-intensive, and limited in scope, covering only a small number of concepts and colors. To address these limitations, researchers have developed automated methods that extract color-concept associations through statistical analysis of language and image databases. Among them, linguistic-based approaches use the frequency of basic color terms (e.g., “red” and “blue”) in language as indicators of color-concept associations [67]. These linguistic methods are often combined with image-based approaches to provide more comprehensive color information [35, 36]. For example, Setlur and Maureen [67] developed a two-stage framework that leverages N-gram [40] and Google Image Search to map a word or phrase to a color. The method pairs a given term with each of the eleven Berlin & Kay basic color terms [4] and calculates their co-occurrences based on querying them in the

Google N-gram corpus¹. Once the highly related basic color terms are identified (e.g., associating “red” and “green” to “apple”), a triple-tuple such as (“apple”, “red”, “clipart”) is submitted to Google Image Search API to retrieve images and extract the dominant colors from these reference images.

However, these language or image database-based approaches face significant challenges. Language databases, such as Google N-gram, fall short in covering the combinations of context-dependent concepts and color terms, thus compromising the accurate justification of their association. For example, “quiet forest” and “polluted sky” are common descriptions, while their combinations with basic color items are inaccessible in the N-gram corpus. As the complexity of concept descriptions increases, the limitations of N-gram-based approaches become more pronounced [10, 57]. On the other hand, the search-based reference images are unstable and involve varying image conditions, making further color extraction challenging [20]. For example, the searched images may contain co-occurring concepts irrelevant to the target concept, leading to misleading color-concept association. To mitigate the issues, existing work limits the search corpus to simple clipart [35, 67] or only derives color histogram rather than dominant colors [20]. However, these constraints inherently limit the application scope. In this work, we introduce a generative approach to color-concept association. We leverage the powerful generative capabilities and controllability of diffusion models to flexibly synthesize contexts and concepts, and control image conditions (e.g., fidelity and illumination), thereby overcoming the limitations of traditional language and image database methods.

2.3 Generative Models as Tools for Data Mining

Generative diffusion models have demonstrated powerful capabilities in producing high-fidelity images from prompts [59, 61]. Research has further enhanced their applicability in scenarios like controllable generation [81], image editing [5, 25] and inpainting [39]. Building on these advancements, the HCI community has explored use cases that utilize generative models for human-AI co-creation [14, 75], such as typographical design [80] and product personalization [71]. In contrast to previous works focusing on creative design, our work employs generative models as a **data mining tool to uncover real-world visual patterns, beyond mere image synthesis**. Our motivation stems from the ability of generative models to capture complex patterns and variations in the training data, thus generating new, realistic samples that reflect the characteristics of the original data. Recent studies have employed pretrained diffusion models to generate synthetic data for training models, thereby improving their performance on classical computer vision tasks such as image classification [2, 64] and image segmentation [48]. For example, Sariyildiz *et al.* [64] leverage Stable Diffusion [59] to generate synthetic ImageNet [13] clones, *i.e.*, datasets with synthetic images for the ImageNet classes, using class names as prompts. Close to our work, researchers have trained generative models (e.g., generative adversarial networks [16]) on faces [7] and cars [11] image datasets to analyze the evolution of their characteristics over history.

¹<https://books.google.com/ngrams/>

Leveraging diffusion models to generate data for training and analysis in vision-centric tasks represents a novel paradigm. The approach can reduce manual labor and mitigate potential data collection issues, such as poor image quality and mislabeled data [43]. To efficiently explore visual patterns of specific interest from the large visual representation space of pre-trained models, fine-tuning [60] and prompt engineering [17] are often required to achieve precise control over the generated content. For example, Siglidis *et al.* [72] introduced an approach that fine-tunes diffusion models on specific datasets to facilitate the generation of image samples that align closely with their labels. However, fine-tuning is more appropriate for closed-set concepts (e.g., classification tasks with finite concept labels), while our target is open-set color-concept associations. To achieve a precise analysis of the intended visual concepts, we introduce a robust post-processing pipeline along with prompt tuning that achieves flexible control. Specifically, we integrate open-set detection [37] and promptable segmentation [28] to precisely segment concept-relevant regions within the image and then conduct color association extraction for primary and accent colors.

3 Formative Study

3.1 Study Design

To gain insights into how professional designers apply color-concept associations, we conducted semi-structured interviews with four experienced designers, each lasting around 45 minutes. The participants brought diverse experiences: E1 had over 10 years of experience in visual design, E2 had 6–10 years, and E3 and E4 had 3–5 years of experience. Their expertise spanned graphic design, UI design, game art design, and product design, offering valuable insights into the role of color in different design contexts. Two participants had formal academic backgrounds in design-related fields, while the others combined academic knowledge with industry practice, offering a well-rounded perspective on how theoretical concepts are applied in professional settings.

During the interviews, we explored the following topics: (1) the workflows designers use when selecting colors for specific concepts, (2) the guiding principles and considerations for semantic color associations and (3) designers' expectations for a supportive tool that integrates color-concept associations into their processes. All interviews were recorded and transcribed to facilitate a detailed analysis process.

3.2 Designer Feedback

In line with previous research, associating colors with specific concepts to convey meaning is a common practice in visual design [21, 51, 53]. Here, we use the example of color-concept association in clipart design, where the task is to design a clipart of “mountain”, as illustrated in Figure 2. Given different contexts (“pristine” vs. “polluted”), the anticipated designs and corresponding colors are different. In the following, we use this example and summarize the designers’ feedback into key considerations in design practice and workflow.

3.2.1 Considerations in Design Practice. In color-concept association for visual design, the concept is *context-dependent*, and the

choice of color typically involves a composition of both *primary* and *accent colors*.

- **Context-dependent concept.** Designers often associate colors with various meanings [9], such as emotions, styles, cultures, or specific items. We refine this broad notion of “*meaning*” into two distinct components: *concept* and *context*. The *concept* refers to a concrete design element, such as a mountain, tree, or car, while the *context* reflects the abstract emotional tone or condition to be conveyed [18, 23]. For example, in a clipart designed to raise awareness about environmental pollution, the “*mountain*” serves as the concept, while its “*clean*” or “*polluted*” state represents the context, conveying a sense of environmental purity or degradation, as shown in Figure 2. Bright greens can evoke an uplifting atmosphere of a pristine mountain, while grayish greens can depict the degradation of a polluted mountain. This context-dependent color choice adjusts the design’s mood without altering the core concept. Proper colors align messages with both concept and context. A mismatch in color-concept association, such as using bright green for a polluted mountain, could confuse viewers and weaken the intended message.
- **Primary-accent color composition.** Color composition rules are considered a fundamental principle in visual design. When associating with concepts, the primary color is closely associated with the concept and the context being depicted. It typically occupies the largest area in the design, ensuring that the concept remains central and easily recognizable. Accent colors complement the primary color to create a balanced and meaningful visual representation. They provide decorative or contextual elements to support the overall design. In the example, the primary color is bright green for “*pristine mountain*” and grayish green for “*polluted mountain*”. The accent colors, such as darker and lighter shades of green, are used to suggest distance, depth, and shading.

Concept and Context-Dependent Color Selection. All participants emphasized that the concept is the primary consideration when selecting colors to convey specific meanings. There are some common color-concept associations used in design, such as blue for the sky, green for mountains, and brown for soil. Besides, they also pointed out that color selection is influenced by factors beyond the concept itself. These factors can be considered as contexts, such as feelings, styles, and target audiences. While all participants mentioned feelings, they acknowledged that feelings can be vague. For instance, E1 described designing a poster about lime, explaining that instead of using the saturated green of the lime peel, they would opt for a fresher green to convey a “*refreshing taste*.” E1 elaborated, “*If I look at the context, I would definitely consider its sour feeling, and I think it must be related to the green and yellow color*.” E3 also highlighted that sometimes the customer’s requirements are vague, such as “*changing the color to be more active*.” Regarding style, E2 pointed out that color usage differs between flat and realistic designs, with flat designs typically using more saturated colors. Additionally, participants may consider the audience of their designs. For instance, E3 noted that designs aimed at children tend to use brighter, more vivid colors, while business designs employ more subdued tones to convey professionalism. There is a need for a flexible approach that

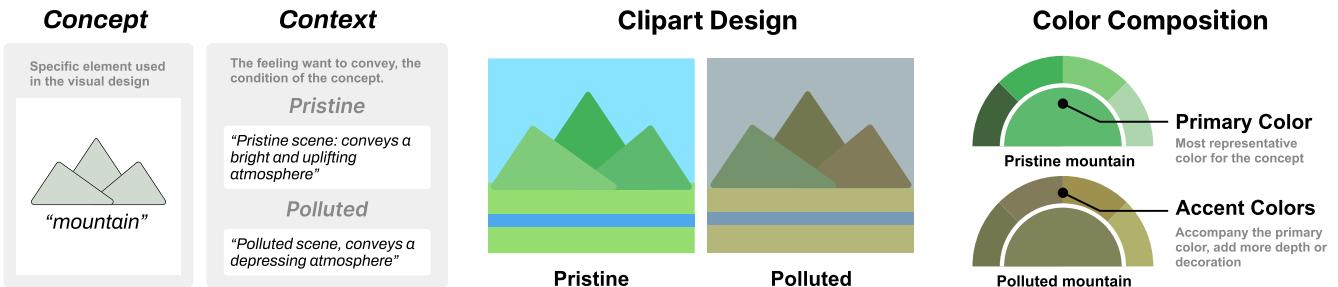


Figure 2: Usage scenario for designing a clipart to raise awareness about mountain pollution. (a) Illustration of the design elements, including the concept (“mountain”) and the context (“pristine” vs. “pollution”). (b) Examples of clipart design: the clipart for “pristine mountain” uses bright green, whilst the one for “polluted mountain” uses grayish green. (c) Primary-accent color composition is used by designers, with the primary color representing the concept and the context, while accent colors provide depth and decoration.

can adapt to these contextual factors, providing designers with a more nuanced and tailored color recommendation.

Intuitive Representation for Color Composition. Designers consistently expressed the need for a well-defined color composition, with a highlight on the primary color and some accent colors. E2 gave an example of food packaging for different flavors, explaining, “*The color should be representative of the flavor, such as tomato-flavored potato chips are always red. So when buyers see it, they can quickly know what flavor it is.*” E4 also preferred starting with a primary color tied to the concept, complemented by other colors working together in the composition. Participants were critical of existing color representations of possibility distributions, with E1 describing them as “*not intuitive*” and “*too many colors unrelated to the concept*.” A palette containing around five colors, as E1 mentioned, was deemed sufficient, similar to those provided by design tools. While the current color palette is insufficient, as usually treated equally with all colors. They mentioned it would be helpful to make the primary color salient, in the center, with accent colors surrounding it, to provide a clear visual hierarchy. Besides, they also suggested that the color ratio should be provided, even if not strictly followed, to give designers a sense of the relative importance of each color, which aligns with previous research [69].

3.2.2 Design Workflow. The practice of color-concept association can generally be categorized into conventional and computational approaches.

- **Conventional approach.** Conventionally, designers begin by gathering references to identify color associations with specific concepts. They search for relevant images, photography, or game scenes that effectively convey the intended concept, using image searching engines like Google Images² or Pinterest³. Many designers habitually collect references, often maintaining large and high-quality personal libraries for inspiration. From these image references, designers draw inspiration for primary colors and iteratively identify accent

colors. Various tools, such as eyedroppers for color extraction and color wheels for modifying color compositions, are utilized to aid this process.

- **Computational approach.** Recent advancements in linguistic analysis and image processing have boosted the development of computational approaches for color-concept association. This approach typically involves several steps: identifying relevant concepts in text corpora (e.g., Google Ngrams⁴), finding reference images from an image dataset, and extracting colors from those reference images. These processes can all be automated to a large extent, leveraging machine learning techniques. The CIELAB color space [12] is often used in the color extraction process due to its computational simplicity and relatively perceptually uniform properties.

Universal and Objective Color References. In the conventional approach, we found that the process of associating colors with concepts varies among designers. Each step relies heavily on intuition and is influenced by personal preferences. While general principles such as color composition rules (E3) and color psychology (E1 and E4) provide guidelines, the final color choice remains subjective. Personal experience and painting styles also play a role. For instance, E4 prefers Morandi colors with lower saturation, while E1 favors brighter colors. This variability extends to how designers approach different concepts, depending on their familiarity and personal feelings. E3 mentioned that what design school actually teaches is aesthetics, which is “*when you see something good, recognize the feeling, and know how to recreate it*”. While courses may teach step-by-step techniques, effective design work also benefits from objective references that help designers understand the public’s perception. As E3 explained, “*Sometimes they (designers) focus only on what they want to express, but what others perceive may be different from their intent, and the result is not what their customer wants.*” Participants believe there is a need for an objective reference to provide a useful benchmark, even if it isn’t directly used in the final design. Such a reference, aligned with public perception, would help guide designers toward more universally understood choices, mentioned by E3 as “*move closer to objective data.*” This could also benefit novice designers, as E1 noted, who may not yet

²<https://images.google.com/>

³<https://www.pinterest.com/>

⁴<https://books.google.com/ngrams/>

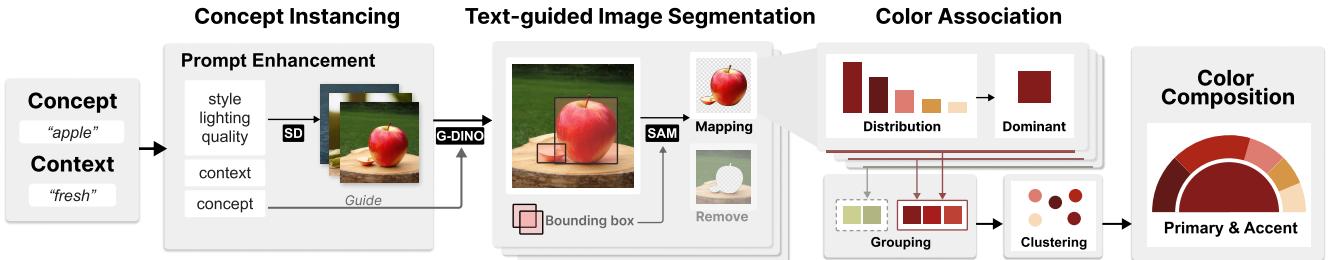


Figure 3: Overview of the *GenColor* framework. The framework includes three stages: Concept Instancing for generating representative image samples, Text-guided Image Segmentation for identifying relevant regions, and Color Association for extracting the primary-accent color composition.

be familiar with or sensitive to certain design concepts and need a starting point for color selection.

Flexible and Robust Image-Based Computational Methods. The computational approach can be limited by the number and quality of images available in the dataset. While designers perceive computational methods as more objective than conventional approaches, the results are often unsatisfactory due to the limitations of image referencing. E2 mentioned that the retrieved images are not always relevant or fail to capture the desired atmosphere. This limitation is attributed to the search engine's algorithm's inability to consider context. For example, while there may be numerous image references for a generic concept like “*mountain*”, there are often fewer relevant references for a more specific concept like “*polluted mountain*”. Moreover, the search results may not use color to reflect the context of “*polluted*”; they might simply add visual elements like garbage without modifying the color to convey degradation. Such retrieved image results are not helpful for finding color references that align with the intended concept and context. Additionally, the retrieved images are often of low quality, with cluttered backgrounds or inconsistent lighting, further complicating the task of identifying suitable color references.

3.3 Design Goals

Facing the gap between designer requirements and existing approaches, we aim to address these issues by leveraging generative models for color-concept association. Our work is inspired by contemporary T2I models, which can resemble visual patterns from large-scale real-world data. Based on insights from the formative study, we identified the following design goals:

- (1) **Provide Justified and Objective Color References.** The framework should be both justifiable and transparent, offering a clear rationale for each suggested color, ensuring designers feel confident about the result. To align with designers' practice of extracting colors from image references, the framework shall compute colors from image references, rather than directly give a coloring result. Additionally, these color references should be derived from large-scale datasets, helping to minimize personal biases and providing objective guidance in the color selection process.
- (2) **Maximize Feasibility and Flexibility in Concepts and Context.** The framework should accommodate a wide range

of concepts and contexts encountered in design. It must dynamically adapt to various conceptual entities and contextual factors, such as emotional tone, artistic style, and target audience. The generative approach should flexibly adjust to these variations, ensuring the generated colors align with the core concept while capturing subtle contextual nuances, offering a tailored response to each unique design scenario.

- (3) **Facilitate Robustness and Effectiveness in Associating Primary-accent Color.** The framework should establish robust color-concept associations by providing a primary-accent color composition that reflects the concept and context. It should identify semantically related regions in reference images and minimize noise during the process. The color composition should consist of a primary color that closely reflects the concept, with accent colors providing additional depth and variety while still maintaining relevance to the overall design.

4 Generative Approach for Color-Concept Association

This section presents *GenColor*, an innovative framework that automatically associates a given concept with primary-accent color composition mined from images by generative models. As depicted in Figure 3, the framework comprises three stages: (1) *Conceptual Instancing* to generate representative images samples (Sect. 4.1), (2) *Text-guided Image Segmentation* to identify conceptual-relevant image regions for color analysis (Sect. 4.2), and (3) *Color Association* to extract primary-accent color compositions (Sect. 4.3). Such a framework mirrors the professional design process of sourcing references and selecting visually appealing concept colors, as discussed in *Design Practice* (Sect. 3.2.2).

4.1 Conceptual Instancing

The initial stage of our framework is dedicated to leveraging diffusion models, specifically Stable Diffusion 3, to generate images that accurately represent a given concept. Given the critical role of prompts and hyperparameters in the image generation process and the potential risk of producing duplicate images, this stage involves two primary steps: a universal prompt refining strategy and a generation strategy varying the hyperparameters, ensuring

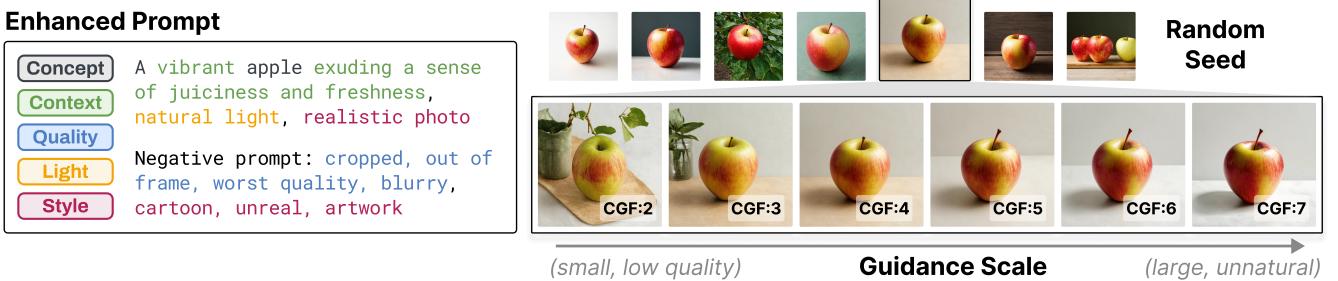


Figure 4: Prompt design and parameter setting in the Conceptual Instancing stage. The prompt is refined based on the core principles of concept, context, style, lighting, and quality control. The guidance scale and seed are adjusted to control image quality and diversity.

the diffusion model to generate images that reflect real-world color-concept associations.

4.1.1 Enhancing the Concept Prompt. Prompt design is critical for the quality and relevance of generated images, as highlighted in previous works [38, 54]. An effective prompt should strike a balance between specificity and flexibility, ensuring accurate visual interpretation while avoiding unnecessary artifacts or distortions. To ensure realistic and consistent color reproduction, we must avoid introducing unnecessary color variance that diverges from the intended concepts and contexts. At the same time, developing a flexible prompt template enables more effective control over image conditions, allowing for the seamless synthesis of concepts across various contexts, as shown in Figure 4 (left).

Specifically, we focus on the following core principles when refining our prompts:

- **Concept:** Clearly define the core concept to be visually represented. Not introducing extra elements or irrelevant objects ensures that the generated images are focused and accurate.
- **Context:** Context refers to the conditional or emotional state of a concept. For example, “ripe banana” versus “unripe banana”, or “sunny sky” versus “stormy sky”. By providing context, we guide the model to generate images that accurately capture the intended mood or setting. It is important to avoid descriptions that directly reference specific color tones, as this could introduce bias into the model.
- **Style:** Style should be defined based on the intended use of the image. We differentiate between two common categories: “realistic photos” and “clipart.” However, terms like “clipart” can be misleading to the model, as Stable Diffusion may misinterpret it. To ensure clarity, we recommend using phrases like “colored flat design” for simpler illustrative styles and “realistic photo” for high-detail, lifelike outputs. By setting clear stylistic expectations, we can mitigate unwanted style mismatches, such as cartoonish or 3D elements when realism or flatness is intended.
- **Lighting:** Lighting should be carefully controlled to maintain realism and color consistency across images. Natural lighting is generally recommended avoiding distortions from artificial light sources, which can introduce color shifts or unnatural shading [78].

- **Quality Control:** Ensuring the generation of high-quality images involves a combination of carefully chosen parameters and the use of negative prompts. Negative prompts allow us to exclude unwanted characteristics—such as overly cropped images or incorrect styles [3]. For example, when generating realistic photos, we would specify in the negative prompt to avoid cartoon or 3D styles, while for flat design, we would exclude realistic 3D rendering styles. This targeted exclusion helps ensure that the generated objects are fully within the frame and match the intended quality.

4.1.2 Image Generation with Stable Diffusion. After refining the concept prompts, we proceed to the image generation phase using the Stable Diffusion 3 model [59]. This model is selected for its ability to produce detailed and visually appealing images that align well with specified prompts, effectively reflecting real-world color associations. To introduce variability and creativity in the generated samples, we adjust certain parameters within the model:

- **Guidance Scale:** This parameter manages the balance between strict adherence to the prompt and creative exploration. Randomizing the guidance scale within an appropriate range helps generate diverse images. Extremely low values can lead to distortions or incomplete images, while high values can result in unnatural contrasts (see Figure 4 (right)). Based on Stable Diffusion recommendations and our empirical adjustments, we set the value to vary randomly between 3 and 6.
- **Seed:** Randomizing the seed parameter ensures that each image generation is unique, even when the prompt remains consistent. This randomness is crucial for producing diverse images and avoiding repetitive outputs.

We generate a set of 50 images for each concept, each with a resolution of 1024×1024 pixels. The quantity of images is chosen following previous literature on obtaining color-concept association [20, 57] to provide a diverse and comprehensive representation of the concept. This resolution is chosen to provide a high level of detail, which is essential for comprehensive color analysis in subsequent stages. Additionally, generating images at a higher resolution increases the likelihood that the model will place the object in the center of the image, resulting in more consistent and varied compositions. By carefully adjusting these parameters, we

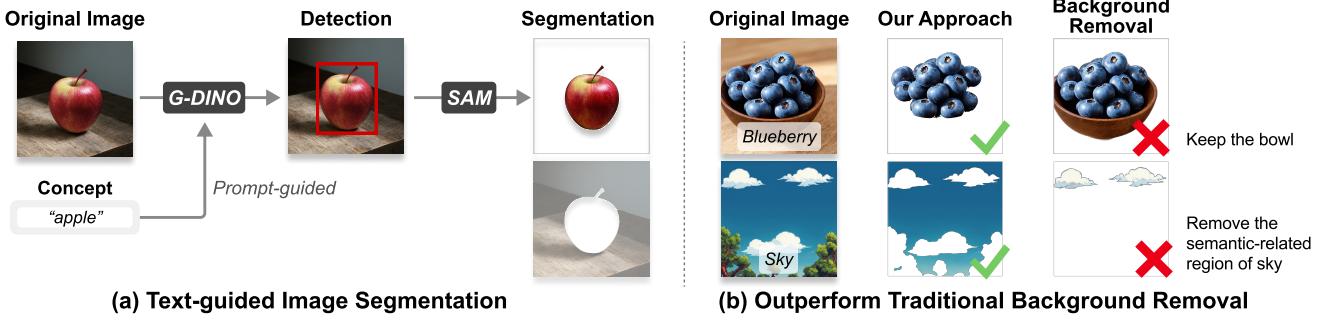


Figure 5: The text-guided image segmentation process. (a) The pipeline utilizes G-DINO for prompt-guided detection and SAM for concept-based segmentation. (b) Comparison of traditional background removal and text-guided segmentation.

ensure that the generated images adhere to the specified concepts and exhibit a wide range of variations, enhancing the overall quality and diversity of the image set.

4.2 Text-guided Image Segmentation

After obtaining representative images of a specified concept, accurately identifying the concept-related regions is crucial for further precise color information extraction. Previous query-based approaches consider unstable search results and varying image conditions, thus restricting the image type to clipart [35, 67] or using background removal techniques to obtain image segments [20]. However, traditional background removal has several drawbacks. One major limitation is the insensitivity to specified concepts. Background removal techniques may retain irrelevant co-occurring elements, leading to misleading color-concept associations. For example, in the “blueberry” case, background removal will still retain unrelated elements (the bowl contains the blueberry), making subsequent color extraction difficult, as shown in Figure 5 (b). Another issue is that common concepts, such as “sea” and “sky,” often appear as part of the image background. Background removal techniques might eliminate these elements, which are common in visual design. For example, Figure 5 (b) presents a case of “sky,” where the background removal technique retains the foreground clouds while removing the “sky” we are interested in. To address the issues, we employ a text-guided image segmentation approach that seamlessly integrates with previous text-guided image generation, as depicted in Figure 5. This integration enables the effective and efficient segmentation of concept-relevant regions in the generated images, thereby enhancing the accuracy of our analysis.

We aim to use specified concepts as text prompts to get the corresponding image segment directly. However, in computer vision, determining the mask in an image that matches the region described by text remains challenging. To effectively address this open-set segmentation challenge [58], we employ an ensemble approach, dividing it into two main components: open-set detection [37] and promptable segmentation [28]. Figure 5 (a) illustrates the process, which begins with GroundingDINO [37], an open-set object detector that uses textual input to generate bounding boxes around regions of interest within the image. To refine these bounding boxes, we apply Non-Maximum Suppression (NMS) to eliminate overlapping regions, retaining only the most confident and relevant

detections. Subsequently, the annotated boxes obtained through GroundingDINO serve as the box prompts for the Segment Anything Model (SAM) [28], which generates precise mask annotations for the identified regions. After segmentation, all segmented regions are merged into a single mask, representing the concept-relevant areas within the image. By leveraging the capabilities of these two robust expert models, we can more effectively accomplish open-set detection and segmentation tasks, thereby improving the accuracy of color-concept association analysis.

4.3 Color Association

The final stage of our framework focuses on extracting color information from the segmented regions of generated images and deriving a primary-accent representative color composition for given concepts. The objective is to identify the primary and accent colors that best represent the concept, ensuring that the color composition is both accurate and meaningful for design applications. Figure 6 illustrates the color association pipeline. Our pipeline begins with color discretization, where the color values within the segmented regions are mapped into a discretized $16 \times 16 \times 16$ RGB color space. This step reduces the complexity of the color data by categorizing close color values into defined bins, simplifying identifying the most frequent colors within each image. The color corresponding to the bin with the highest count is then selected as the dominant color at the image level, such as in Figure 6.

Considering that a concept can be associated with multiple reasonable colors across different images, it is essential to group colors appropriately. Our approach aims to balance conciseness and diversity in color representation, thereby providing an effective color abstraction. We merge similar colors, while still allowing diversity at the color palette level. Specifically, displaying multiple reasonable colors together should be avoided, as they may not typically appear simultaneously in a design. For instance, an “apple” is represented by different dominant colors like red or green, but it would be unusual to see both colors in one apple. To preserve diversity within the same color category, we distinguish between different shades of similar colors. Take “mountains”, for instance; they can appear in deep green or deep blue-green tones. Although these shades are somewhat similar, separating them into distinct groups ensures a varied and comprehensive color palette and provides diverse references for designers. Thus, we conduct image grouping

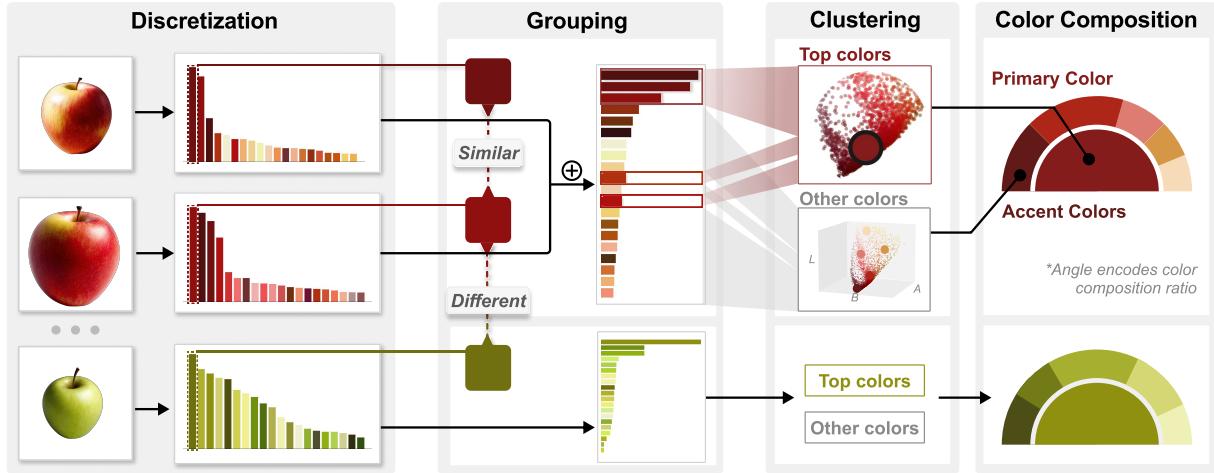


Figure 6: The process of extracting color information from segmented regions of generated images to derive a primary-accent color composition for a given concept. The framework includes color discretization, grouping similar colors into same groups, and clustering to identify the primary and accent colors for the final color composition.

based on their dominant colors. To address this, we group images by calculating CIEDE2000 color distances in the CIELAB space and applying a threshold of 12. Such a threshold allows for noticeable differences between colors, such as variations within the same color category (e.g., different shades of red), while still considering them part of the same color group. This ensures that similar colors are clustered together while distinctly different colors remain separate. After grouping, we sort the color groups by the number of images they contain and select the top five most frequent color groups. To ensure the robustness of the pipeline, we exclude any color groups that contain fewer than three images. After grouping, we aggregate the color bin values of each image in the same group.

Similar to image-level color extraction, we apply color discretization to the aggregated color bins of each image group. The top colors are considered the dominant color for the group. To deliver the color composition of each image group, we separate their aggregated color bins into two classes: *top colors* and *other colors*. Specifically, using the CIEDE2000 color distance threshold ≤ 7 , commonly recognized as the range distinguishable by the human eye [73], we filter the colors surrounding the group-level dominant color as the *top colors*. Colors not falling within this threshold are categorized as *other colors*. Then, we set the centroid of *top colors* as the primary color. For accent colors, we apply k -means clustering to compute the typical colors for the concept across all images, where we set k to 5 following previous color palette research [6]. This grouping-then-clustering approach ensures the full exploration of color-concept association and the delivery of the primary and accent colors to represent the concept, thus providing a coherent and comprehensive color composition for design purposes.

For the final color composition, we design an intuitive visual representation that combines the primary and accent colors. Inspired by the traditional painting palette, we create a radial glyph, placing the primary color in the center and the accent colors around it, as shown in Figure 6. The central position of the primary color emphasizes its importance, allowing designers to quickly identify the

main color associated with the concept. Accent colors with greater variance are positioned on the outer ring, helping designers explore a broader range of complementary colors. The angular range of each accent color represents its proportion within the color palette, reflecting its frequency as determined by the clustering algorithm. This visual representation provides a clear and intuitive way to present the color composition of a concept, aiding designers in selecting appropriate colors for their design tasks.

5 Evaluation and Application

This section presents a comprehensive evaluation of *GenColor*, demonstrating that it generates more representative and user-preferred colors compared to previous computational methods. To quantitatively assess the different methods in color space using color distances, we first gather a baseline dataset from designers (Sect. 5.1). Then, we compare different approaches using color difference comparison (Sect. 5.2.2) and user ratings (Sect. 5.2.3). Furthermore, we present various downstream application scenarios of *GenColor* (Sect. 5.3). Finally, we construct a large-scale color gallery where designers can explore and retrieve relevant color compositions by performing text-based search (Sect. 5.4).

5.1 Designers' Coloring Dataset

The lack of baseline data hinders the quantitative assessment of different approaches in the color space. Here we aim to collect a dataset that shows how professional designers associate colors with various concepts and contexts, and then use this dataset as a baseline for evaluating the effectiveness of our generative approach.

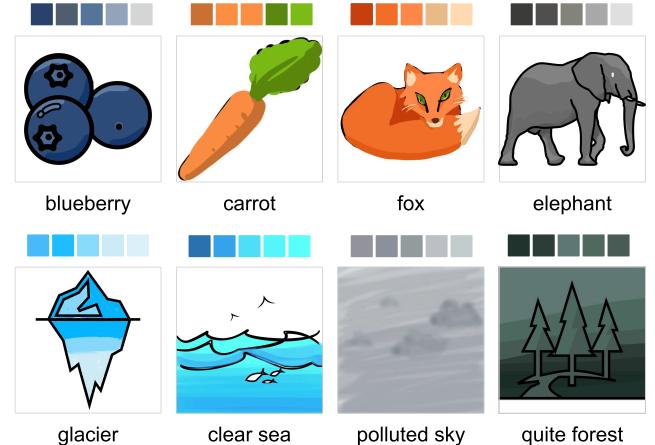
5.1.1 Scope of the Concepts and Contexts. We selected a diverse range of concepts and contexts commonly used in design practice, drawing from previous studies and designer interests. The selected concepts fall into two main categories: *Concept* and *Context*, as summarized in Table 1.

Table 1: Examples of entities classified by type and subtype.

Type	Subtype	Entity
Basic Concept	Fruit	apple, banana, blueberry, cherry, grape, peach, tangerine
	Vegetable	carrot, celery, corn, eggplant, mushroom, olive, tomato
	Environment	mountains, sky, ocean, desert, grassland, forest, lake, glacier
	Animal	fox, tiger, lion, elephant, crocodile, flamingo, penguin, peacock
Context	Conditional	polluted, clear, pure, at springtime, at sunset
	Emotional	lively, quiet, peaceful, depressed, happy, sad, energetic

- **Concept:** We primarily focused on natural elements, as they tend to exhibit more stable color associations compared to artificial concepts, which can vary more widely (e.g., color of plastic). Based on previous studies [42, 57, 67], we included simple, standalone entities such as common fruits (e.g., apple, banana) and vegetables (e.g., carrot, corn), with seven items in each category. To broaden the scope, we also incorporated more complex elements, including animals and environmental components, adding eight items per category. Examples include animals like fox, tiger, and lion, and environmental elements like mountains, sky, and ocean.
- **Context:** This category pairs concepts with different contexts that can influence their color associations. We considered both the conditional and emotional aspects of each concept, selecting six items in this category. Examples include the sky in different states (e.g., polluted vs. clear) or scenarios evoking different feelings (e.g., quiet vs. lively).

5.1.2 Participants and Procedure. We recruited 24 professional designers from online design communities, representing a diverse demographic in visual design. Participants ranged from 20 to 30 years old ($M = 25.5$, $SD = 2.57$), with 14 males and 10 females. Professionally, 4 worked in industry, 11 held related degrees and were employed in academia, and 8 had experience in both. Their backgrounds spanned graphic design, UI design, interior design, illustration, product design, and animation. In terms of expertise, 5 had 1-2 years of experience, 10 had 3-5 years, 7 had 6-10 years, and 1 had over 10 years. This study received approval from the university's ethics committee. The study began with obtaining informed consent and collecting background information through a questionnaire. All participants were explicitly informed that their work contributed to the development of a Generative AI tool, specifically in the evaluation stage, and that their data would be used solely for research purposes. They then received detailed instructions and completed a tutorial session featuring three example concepts. After confirming their understanding, they proceeded with the main task. Participants were asked to complete a coloring task for 36 different concepts, using the most representative colors they associated with each concept. Designers were asked to color the drawings without any reference, using around 1-5 representative colors of their choice. They were encouraged to reflect the importance of each color by adjusting the area it occupied in their design. To simplify the task, we provided participants with a line drawing of each concept. The line drawings were found online and verified by an expert designer.

**Figure 7: Examples of the dataset collected from professional designers.**

to ensure that the concepts were easily recognizable and unambiguous. Designers were allowed to use any tools they were familiar with, and no time limits were imposed on the task. Upon completion, designers confirmed that their chosen colors matched their understanding of each concept. The study was conducted online, taking approximately 1-2 hours, with participants compensated around 15 USD.

5.1.3 Designers' Coloring Dataset. We collected a total of 36 concepts, each of which was colored by 24 designers, resulting in 864 color samples. The colorings are of high quality, with consistent color choices across designers. Figure 7 illustrates examples of the coloring results, and the corresponding extracted color compositions are shown above. The whole dataset is available in the supplementary material, and will be released in the future.

5.2 Quantitative Evaluation

To evaluate the effectiveness of our approach, we conducted a quantitative evaluation comparing the color compositions generated by our generative method with those query-based results.

5.2.1 Experiment settings. For the evaluation, we constructed a dataset of color compositions for 36 distinct concepts aligned with the designers' coloring dataset. We included two types of image

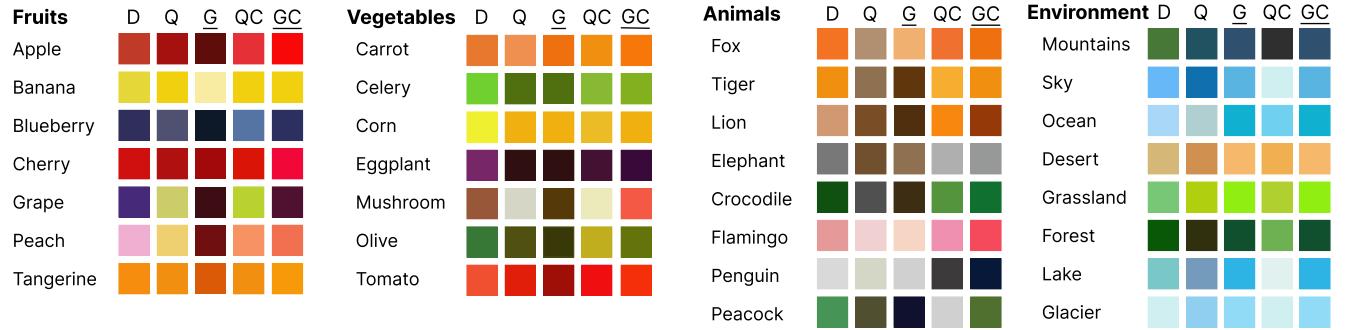


Figure 8: Our result compares with previous works on basic concepts. D stands for the designers's ground truth, G stands for the generated result by our method, Q stands for the query-based result, and QC stands for the query-based clipart result.

sources: photos and clipart. Previous studies suggest that, for certain concepts, human-made illustrations may better capture people's color associations [35]. For instance, the association between "money" and the color green is more pronounced in clipart images than in photos. It is important to determine when to use photos and when to use clipart based on the specific concept. For each concept, we applied the following conditions to obtain the image dataset and process them with the pipeline in Sect. 4:

- **Queried Image (Q):** Image datasets were obtained by querying the concept in Google Image Search. We use the top 50 images returned from the search engine, consistent with previous studies [20]. The search engine is expected to return the most relevant images for the concept. Most of the images are photos, but some clipart images may also be included.
- **Generated Image (G):** Realistic photo datasets generated by our method are designed to reflect the natural color patterns associated with photographic representations of the concept. The prompt is set to generate "*realistic photo*" with "*natural light*". Contexts have been enhanced in prompt to generate more context-dependent color, such as "*exuding a sense of depression and heaviness*" for "*polluted*". Note that the prompt is not directly related to the color, but the context can influence the color association.
- **Query-based Clipart (QC):** Image dataset obtained by querying the concept with the term "*clipart*" appended in Google Image Search. Color compositions are generated by querying the concept with the term "*clipart*" appended in Google Image Search. This method filters the results to focus on more simplified and symbolic representations, emphasizing key colors typically used in clipart images.
- **Generated Clipart (GC):** Clipart datasets generated by our method, based on simplified and stylized representations that often use more distinct and exaggerated colors. This dataset is generated with the prompt "*simple flat design*". The context is also enhanced in the prompt.

We observed that the T2I model responds more effectively to visual contexts than to descriptions targeting other senses. For instance, context like "*clear*" versus "*polluted*" are captured more easily than "*lively*" versus "*quiet*." This likely stems from the model's training on visual data, making it less sensitive to certain auditory

terms. To improve this, we enhanced prompts by using more detailed descriptions related to atmosphere and emotions, such as replacing "*quiet*" with "*evoking feelings of silence and lonely*." Further details are available in the Appendix.

5.2.2 Color Difference Comparison. To demonstrate the effectiveness of our approach, we compared the color generated by our method with those obtained from the query-based methods. First, we visually examined the primary colors of each basic concept, as shown in Figure 8. Then, we used CIEDE2000 color difference to quantify the similarity between the primary colors extracted from the generated images and the designers' ground truth.

The results in Figure 8 show that the primary colors extracted from our generated approach are generally consistent with those produced by designers. While there are minor visual differences in brightness and saturation, the color differences remain within an acceptable range, and the obtained colors are reasonable. In contrast, the query-based approach exhibits more variation, particularly in environmental concepts.

These observations are further supported by the quantitative color distance comparison in Figure 9. We grouped the comparisons based on the image source type—whether generated from photos or clipart—since it's not reasonable to directly compare colors between photos and clipart. For basic concepts, our generated approach outperformed the query-based approach and aligned more closely with the designers' results. Specifically, the generated clipart ($M = 15.18$, $SD = 14.90$) had a smaller color distance than the queried clipart ($M = 19.97$, $SD = 18.92$). For photos, the generated image ($M = 22.26$, $SD = 13.03$) was also closer to the ground truth compared to the general queried results ($M = 24.19$, $SD = 17.61$).

When examining context-dependent concepts, the advantages of the generative approach are even more pronounced, with the small color distance and more stable results, as shown in Figure 9. The generated clipart ($M = 13.37$, $SD = 3.26$) had a much smaller and more stable color distance compared to the query-based clipart ($M = 27.08$, $SD = 18.56$). Similarly, the generated image ($M = 17.91$, $SD = 6.19$) was closer to the ground truth compared to the general queried results ($M = 22.77$, $SD = 15.12$). This advantage is likely due to the query-based approach's failure to properly link colors with context. For instance, as shown in Figure 9 (right), when

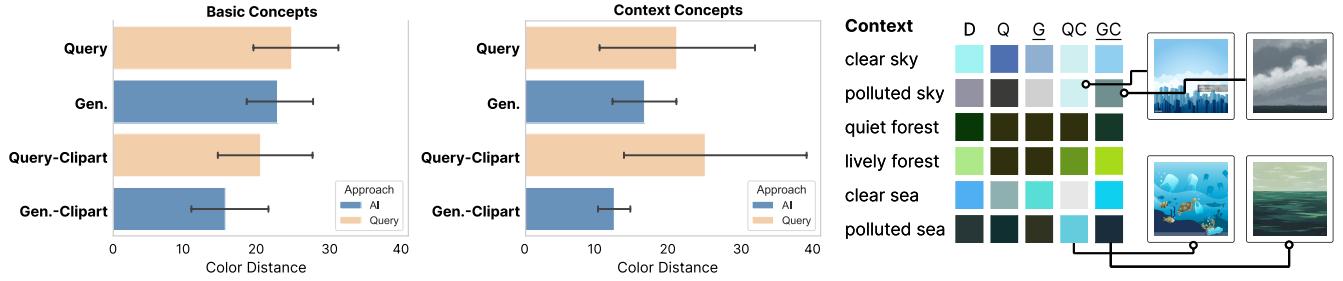


Figure 9: Color distance between different approaches and the designer’s ground truth. Left shows the color distance of our generative approach and two query-based methods, across basic concepts and context-dependent concepts. Right shows the dominant color of each context, with example generated and queried images.

querying for clipart of a polluted sky, the results often add visual elements such as chimneys and smoke to convey pollution but make little effort to modify the actual color of the sky. Similarly, in the clipart of a polluted sea, garbage is added to the image, but the sea remains blue. In contrast, the colors generated by our approach are closely aligned with the context, effectively conveying the intended atmospheric or environmental feel, such as the somber tone of a polluted sky or the murky green of a polluted sea.

5.2.3 User Rating. We conducted a user study to evaluate the representativeness and preference of the color-context association across four different methods and four groups of concepts. The results are shown in Figure 10.

Study Design. We recruited 30 participants (21 females, 9 males) through posters and social media, with an average age of 27.07 years ($SD = 6.90$). Among the participants, 12 participants have design backgrounds, including graphic design, UI design, visual design, etc. The remaining 18 participants came from various fields, such as engineering, computer science, and psychology, with no professional design experience. This participant pool provided a wide range of perspectives for the evaluation. The experiment assessed two primary factors: *representativeness* and *preference*, both rated on a 7-point Likert scale. Representativeness was defined as the extent to which the color composition was perceived to align with the concept and context, while preference referred to participants’ subjective liking of the color composition, both are primarily influenced by color harmony and visual appeal. Participants evaluated color compositions generated by four different methods: Query (Q), Generated Image (G), Query Clipart (QC), and Generated Clipart (GC). A total of 36 concepts were tested across these four methods and two metrics, resulting in 8640 trials (36 concepts \times 4 methods \times 2 metrics \times 30 participants). The order of the trials was fully randomized, and participants were blinded to the method used to generate each color composition. The test took approximately 20 minutes to complete, and participants were compensated approximately 3 USD for their time.

Results. We conducted a one-way ANOVA to compare the ratings across the four methods and four groups of concepts, followed by a post-hoc T-test to identify significant differences between these methods.

For the representativeness evaluation, the generated clipart received the highest rating ($M = 4.71, SD = 1.46$), while the generated

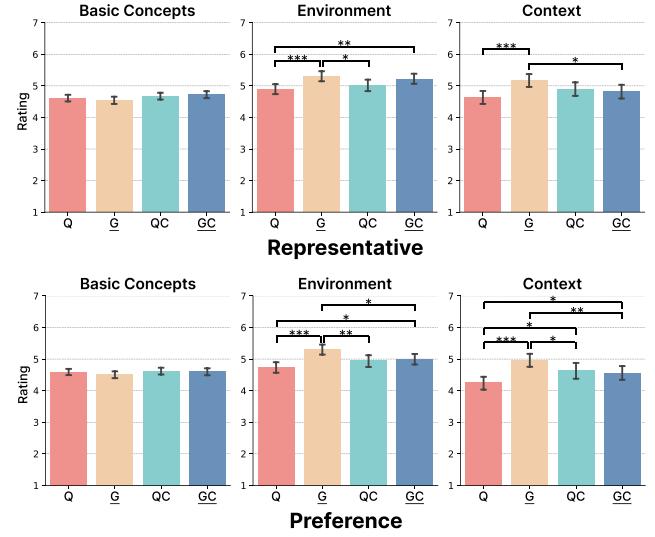


Figure 10: User rating for the representative and preference of the color-context association across four different methods and four groups of concepts.

image ($M = 4.55, SD = 1.47$) had the lowest rating for basic concepts. The differences between these ratings and those of the query clipart ($M = 4.68, SD = 1.41$) and query ($M = 4.61, SD = 1.37$) were minimal. In the environmental concepts group, the generated image ($M = 5.31, SD = 1.26$) received the highest rating, significantly higher than both the query ($M = 4.90, SD = 1.27, p < .001$) and the query clipart ($M = 5.02, SD = 1.41, p < .05$). Additionally, the generated clipart ($M = 5.23, SD = 1.23$) was rated significantly higher than the query ($p < .01$). Similarly, in the context-dependent concepts group, the generated image ($M = 5.17, SD = 1.44$) was rated significantly higher than both the query ($M = 4.63, SD = 1.41, p < .001$) and the generated clipart ($M = 4.82, SD = 1.43, p < .05$). Although the query clipart ($M = 4.89, SD = 1.50$) received a higher score than the generated clipart, the difference was not statistically significant. These findings suggest that participants perceive the generated images as more representative than other methods, particularly for natural environments and context-dependent concepts.



Figure 11: Example of identifying the representative color of holiday and landmark.

In the preference evaluation, no significant differences were found between the four approaches in the basic concepts group. However, the query clipart approach ($M = 4.62, SD = 1.41$) received the highest preference rating, followed closely by the generated clipart ($M = 4.60, SD = 1.47$), query ($M = 4.59, SD = 1.33$), and generated image ($M = 4.50, SD = 1.41$). For the environmental concepts group, participants significantly preferred the generated image ($M = 5.30, SD = 1.27$) over all other methods. This difference was statistically significant when compared to the generated clipart ($M = 5.00, SD = 1.37, p < .05$), query ($M = 4.73, SD = 1.30, p < .001$), and query clipart ($M = 4.94, SD = 1.42, p < .01$). In the context-dependent concepts group, the generated image again received the highest preference rating ($M = 4.97, SD = 1.42$), with significant differences observed compared to the query clipart ($M = 4.63, SD = 1.61, p < .05$), generated clipart ($M = 4.56, SD = 1.51, p < .01$), and query methods ($M = 4.24, SD = 1.43, p < .001$). These findings suggest that participants prefer the generated image for natural environments and context-dependent concepts. The generated clipart received the second-highest rating in the environmental concepts group, significantly outperforming the query method ($p < .05$). In contrast, the query method was rated lowest in the context-dependent concepts group, with statistically significant differences compared to all other methods: generated image ($p < .001$), generated clipart ($p < .05$), and query clipart ($p < .05$). This indicates that participants tend to dislike the query method the most for context-dependent concepts.

Overall, no significant differences were observed in representativeness or preference among the four methods for basic concepts. However, for natural environments and context-dependent concepts, the generated image outperformed the other methods in both representativeness and user preference, while the query method was rated as the least representative and least preferred.

5.3 Application Scenario

We demonstrate two application scenarios where our method can support designers in color-concept association tasks: identifying representative colors for design elements and clipart coloring.

5.3.1 Identifying Representative Colors for Design Elements. In scenarios such as designing posters for holiday promotions or city

branding campaigns, identifying the primary colors associated with each design element is a critical first step. Our method offers an efficient solution for designers to pinpoint representative colors that align with the intended theme. Figure 11 illustrates examples of identifying the representative colors for both holidays and landmarks. For instance, considering key elements of Christmas, such as the Christmas tree, the method recommends a color composition with green as the primary color, which is strongly associated with the holiday. Similarly, when branding iconic landmarks like the Taj Mahal or the Statue of Liberty, the method generates distinct color compositions that reflect each landmark's unique visual identity—beige and gold for the Pyramids or teal for the Statue of Liberty. This ensures consistency and visual cohesion across a series of promotional materials.

5.3.2 Clipart Coloring. In another scenario, designers tasked with coloring illustrations for promotional graphics can benefit from the color compositions generated by our method. These results serve as a valuable reference, aligning closely with public perception and providing guidance for accurate and effective coloring. Figure 12 illustrates how designers apply the recommended colors to various conceptual entities and contexts. For basic design elements, such as apples or penguins, the method offers intuitive and recognizable color compositions that enhance the clarity and visual appeal of the design. In more complex scenarios, the method adapts its color recommendations based on environmental or temporal contexts. For instance, a farm field during harvest is characterized by warm browns and yellows, while the same field in spring might feature fresh greens. Similarly, a desert scene at midday could be depicted in harsh, hot, and bright tones, whereas a desert at sunset might be represented with cooler, darker tones. This flexibility in adapting to diverse contexts not only streamlines the design process but also provides designers with a reliable starting point or reference for their work. By offering context-dependent color recommendations, our method helps designers create more engaging and visually coherent designs that resonate with their intended audience.

5.4 Gallery of Color-Concept Association

To enhance research and practical design applications, we developed a gallery of color-concept associations, offering a diverse range

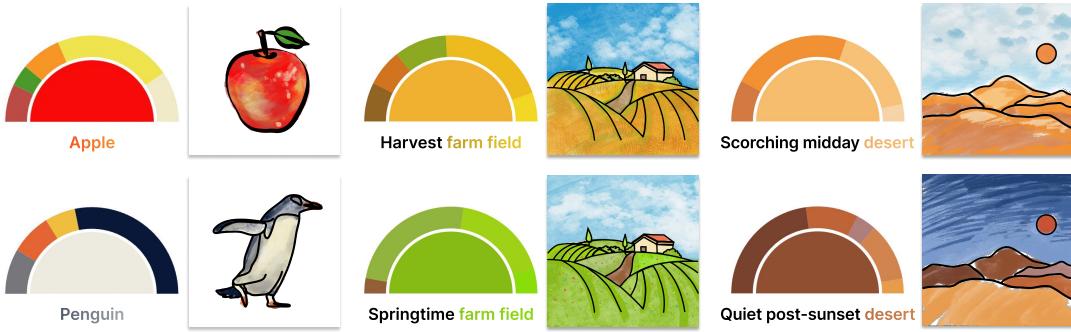


Figure 12: Result of user coloring task, demonstrating how the recommended colors are applied to different conceptual entities and contexts, such as seasonal or time variations.

of concepts and contexts (see Figure 13). This aligns with previous studies [26, 44], providing a visual representation showing the outcome of studies on color-concept associations. Our gallery includes common design elements and extends to nuanced categories such as different styles, emotions, and times of day. Designers can perform text-based searches to easily explore and retrieve relevant color compositions, enabling them to find colors that best suit their specific design needs. This functionality provides designers with an extensive and adaptable resource to support various creative projects.



Figure 13: Gallery of color-concept associations, showcasing a wide range of concepts and contexts.

6 Discussion

6.1 Personalized Color-concept Association

Generative AI's ability to produce universal color-concept associations that align closely with human judgments demonstrates its reliability. It is valuable for designers to effectively convey intended meanings and emotions to the general public, as well as to ensure that their designs are universally understood. However, generative AI can be extended beyond universal associations to cater to more personalized design solutions. Acting as an adaptive agent, generative AI can simulate more nuanced audience perspectives by factoring in personality traits, cultural contexts, and emotional responses, which are essential for achieving personalized design solutions [18]. This adaptability is particularly impactful when designing for diverse audiences where a one-size-fits-all approach

to color selection is insufficient. We initially tested how different audiences, such as elderly individuals, women, young adults, and children, perceive colors in the contexts of farms and cityscapes, as shown in Figure 14. The results reveal that color related to children is bright and saturated, while for elderly audiences, it is darker and muted. These findings highlight the capability of generative AI to tailor color associations in ways that resonate more deeply with specific user groups.

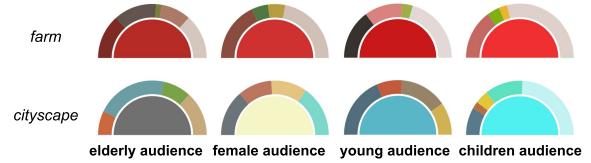


Figure 14: Associated colors for farm and cityscape concepts as perceived by different audience groups.

These findings highlight generative AI's potential to transcend traditional design practices, offering nuanced color associations for personalized, context-sensitive designs that align with diverse audiences' needs.

6.2 Balancing Creativity and Consistency

The integration of generative AI into design workflows reveals a critical tension between creativity and consistency. On the one hand, AI's capacity to generate consistent color-concept associations ensures designs are broadly accessible and effectively communicate intended messages, aligning with general audience expectations. This is particularly beneficial for conveying widely accepted meanings or evoking predictable emotional responses, enhancing efficiency in reaching diverse audiences. While consistency ensures baseline communicability, controlled diversity at the palette level allows creative exploration within defined boundaries. Design is not solely about meeting audience expectations; it also involves innovation and reflecting the designer's unique perspective [46]. Thus, it is essential that generative AI supports diverse color-concept associations without hindering designers' creativity.



Figure 15: Exploring diverse color-concept associations using GenColor. (A) Prompt level: GenColor allows diverse input exploration, allowing designers to specify varying attributes. (B) Palette level: GenColor provides a spectrum of palettes linked to each concept, offering designers a broad range of options.

By introducing diversity at the prompt and color palette levels, GenColor facilitates the exploration of diverse color-concept associations. At the prompt level, we enhance creative flexibility through a variety of context-aware concepts (e.g., vibrant and desolated), styles (e.g., photo, flat design, painting), and lighting (e.g., tones and intensity) options, as shown in Figure 15 (A). Different prompts can influence image generation, capturing diverse design intentions even with a fixed seed. We acknowledge that a concept may be linked to multiple color palettes for color variation at the palette level. As illustrated in Figure 15 (B), we provide a spectrum of palettes associated with each concept, giving designers a wide array of options. Concrete concepts like “*fox*” or “*clear sky*” may show slight shade variations. In contrast, broader concepts like “*mountains*” or “*park in spring*” lead to more diverse palettes featuring a wider range of shades and tones. For instance, mountains might include various greens and blues hues, while spring parks could showcase pinks, greens, and yellows.

Moreover, generated color palettes are intended to support designers’ decisions by serving as inspirational references, not as final solutions. Importantly, designers should always retain the freedom to select, modify, and refine these suggestions according to their vision. Generative AI should be seen as offering a starting point or reference rather than as a substitute for the designer’s voice. In this collaborative process, the role of the designer is crucial in refining and personalizing AI suggestions, ensuring that the final design embodies both consistency and creativity. In this way, AI can facilitate the efficient generation of typical color associations while still allowing designers to express their unique creative identity.

6.3 Limitation and Future Work

Expand the Scope of Concepts. Our current framework primarily addresses concrete concepts within specific contexts. However, abstract concepts, such as “*danger*” or “*freedom*” are challenging to represent visually and remain outside the scope of this study. To improve, incorporating metaphorical or symbolic imagery could better capture these abstract ideas. Additionally, our approach currently handles single concepts in isolation and does not address the

complexities of managing multiple concepts simultaneously. In the formative study, E4 highlighted the difficulty of selecting colors for multiple concepts while maintaining harmony and conveying the intended context effectively. Future work could focus on methods to support the composition of colors representing multiple concepts, ensuring cohesive and harmonious design outcomes.

Enhance Contextual Adaptability. Our current testing has focused on conditional and emotional contexts. While our framework is theoretically generalizable to more complex contexts—such as audience preferences, cultural influences, or artistic styles—these areas have not been sufficiently explored. Future work could investigate how generative AI can adapt to these richer contexts and offer relevant color recommendations. An intriguing area for future research is to explore which contexts are interpretable by generative AI. Understanding this could provide valuable insights into optimizing AI for diverse and nuanced design needs.

Biases and Limitations in T2I Model. Despite generating high-quality images, T2I models exhibit North-centric biases, where AI systems are disproportionately influenced by data and perspectives from developed regions [45, 55, 76]. This bias can also affect color-concept associations. For instance, the color generated for “*money*” is typically green, reflecting its association with the US dollar, and “*taxis*” are generally portrayed as yellow. However, as artificial concepts, these colors can vary across different countries. In our approach, we address these variations by associating each concept with multiple color palettes, allowing the model to capture a broader range of color associations. In future work, thoroughly exploring model biases from the perspective of color is essential. Another limitation of the T2I model is its sensitivity to different prompts. While we have made efforts to enhance the model’s responsiveness by refining prompt structures and incorporating more detailed context, as mentioned in Sect. 5.2, the model still doesn’t perform perfectly across all concepts. Therefore, it is essential to recognize that generative AI is a supportive tool that assists designers by offering suggestions, rather than replacing their expertise.

Expanding and Comparing Generative Models. This work focuses on T2I models for generating images to extract color-concept

associations, which are preferred by designers as they provide direct references for color selection and align with design workflows. Additionally, large language models also show the capability to generate color-concept associations based on text descriptions. Previous research has demonstrated that color-concept associations derived from language models generally correlate with human judgments [41]. Exploring the differences between image-based and language-based references is valuable. Future work should compare these approaches to gain insights into their interpretability and investigate the potential benefits of combining them.

Ethical Considerations of GenAI Tools. While *GenColor* aims to support designers in color exploration, we acknowledge the broader ethical implications of GenAI tools. One concern is GenAI tools may lead to the homogenization of creative output, producing works that conform to mainstream designers and diminishing diversity [15, 49, 62]. Marginalized designers, whose works often reflect unique personal and cultural elements, may be disproportionately affected. Additionally, the automation of some tasks traditionally performed by human designers could lead to job losses and exacerbate inequality in the design industry, especially for those marginalized ones [62]. The HCI community has recognized that GenAI tools should serve as supportive rather than replacement tools for human creativity [49]. They advocate for keeping designers in the loop, enabling them to leverage GenAI tools to enhance efficiency in routine tasks and thus devote more time and energy to creative and expressive aspects of their work [22, 66]. Aligning with these principles, *GenColor* is positioned as an assistive tool that empowers designers rather than replacing their creative role. Future work should focus on fostering diversity, ensuring transparency, and establishing ethical guidelines to promote responsible and inclusive use of GenAI in design practices.

7 Conclusion

In this study, we introduced *GenColor*, a generative framework designed to improve the association of colors with concepts in visual design. By leveraging T2I models, *GenColor* overcomes the limitations of traditional retrieval-based methods, offering flexibility in generating context-dependent images and robustness in managing variations. Our evaluation shows that *GenColor* aligns closely with designer preferences and outperforms existing approaches. We also provide a publicly available gallery of color-concept associations to support research and practical applications in design. Overall, *GenColor* enhances designers' ability to make informed and contextually relevant color choices, facilitating more effective visual communication.

Acknowledgments

The authors wish to thank the anonymous reviewers for their valuable comments, as well as the time and effort of all the study participants. This work is partially supported by the Guangzhou Basic and Applied Basic Research Foundation (No. 2024A04J6462).

References

- [1] Adobe. [n.d.]. Reinventing Adobe Spectrum's colors. <https://adobe.design/stories/design-for-scale/reinventing-adobe-spectrum-s-colors>. Accessed: 2024-12-01.
- [2] Shekoofeh Azizi, Simon Kornblith, Chitwan Saharia, Mohammad Norouzi, and David J Fleet. 2023. Synthetic Data from Diffusion Models Improves ImageNet Classification.
- [3] Yuanhao Ban, Ruochen Wang, Tianyi Zhou, Minhao Cheng, Boqing Gong, and Cho-Jui Hsieh. 2024. Understanding the Impact of Negative Prompts: When and How Do They Take Effect? *arXiv preprint arXiv:2406.02965* (2024).
- [4] Brent Berlin and Paul Kay. 1991. *Basic Color Terms: Their Universality and Evolution*. Univ of California Press.
- [5] Tim Brooks, Aleksander Holynski, and Alexei A Efros. 2023. InstructPix2Pix: Learning to Follow Image Editing Instructions. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 18392–18402. doi:10.1109/CVPR52729.2023.01764
- [6] Huiwen Chang, Ohad Fried, Yiming Liu, Stephen DiVerdi, and Adam Finkelstein. 2015. Palette-based Photo Recoloring. *ACM Trans. Graph.* 34, 4 (2015), 139–1. doi:10.1145/2766978
- [7] Eric Ming Chen, Jin Sun, Apoorv Khandelwal, Dani Lischinski, Noah Snavely, and Hadar Averbuch-Elor. 2023. What's in a Decade? Transforming Faces Through Time. In *Computer Graphics Forum*, Vol. 42. 281–291. doi:10.1111/cgf.14761
- [8] Minshuo Chen, Song Mei, Jianqing Fan, and Mengdi Wang. 2024. Opportunities and challenges of diffusion models for generative AI. *National Science Review* 11, 12 (10 2024), nwae348. doi:10.1093/nsr/nwae348 arXiv:<https://academic.oup.com/nsr/article-pdf/11/12/nwae348/60677321/nwae348.pdf>
- [9] Yun Chen, Luwen Yu, Stephen Westland, and Vien Cheung. 2021. Investigation of Designers' Colour Selection Process. *Color Research & Application* 46, 3 (2021), 557–565. doi:10.1002/col.22631
- [10] Alexander Clark, Gianluca Giorgolo, and Shalom Lappin. 2013. Statistical Representation of Grammaticality Judgements: the Limits of N-Gram Models. In *Proceedings of the Annual Workshop on Cognitive Modeling and Computational Linguistics (CMCL)*. 28–36. <https://aclanthology.org/W13-2604>
- [11] Théophile Dalens, Mathieu Aubry, and Josef Sivic. 2019. Bilinear Image Translation for Temporal Analysis of Photo Collections. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43, 4 (2019), 1197–1212. doi:10.1109/TPAMI.2019.2950317
- [12] Commission Internationale de l'Eclairage. 1978. Recommendations on uniform color spaces, color-difference equations, psychometric color thermics. CIE.
- [13] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. 2009. ImageNet: A Large-scale Hierarchical Image Database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*. 248–255. doi:10.1109/CVPR.2009.5206848
- [14] Xianzhe Fan, Zihan Wu, Chun Yu, Fenggui Rao, Weinan Shi, and Teng Tu. 2024. ContextCam: Bridging Context Awareness with Creative Human-AI Image Co-Creation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–17. doi:10.1145/3613904.3642129
- [15] Sanjana Gautam, Pranav Narayanan Venkit, and Sourojit Ghosh. 2024. From melting pots to misrepresentations: Exploring harms in generative ai. *arXiv preprint arXiv:2403.10776* (2024).
- [16] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative Adversarial Nets. In *Advances in Neural Information Processing Systems*, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K.Q. Weinberger (Eds.), Vol. 27. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2014/file/5ca3e9b122f61f8f06494c97b1acf3-Paper.pdf
- [17] Yaru Hao, Zewen Chi, Li Dong, and Furu Wei. 2023. Optimizing Prompts for Text-to-Image Generation. In *Advances in Neural Information Processing Systems*, Vol. 36. Curran Associates, Inc., 66923–66939. https://proceedings.neurips.cc/paper_files/paper/2023/file/d346d91999074dd8d6073d4c3b1373b-Paper-Conference.pdf
- [18] Lena Hegemann and Antti Oulasvirta. 2024. Palette, Purpose, Prototype: The Three Ps of Color Design and How Designers Navigate Them. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Article 147, 19 pages. doi:10.1145/3613904.3641976
- [19] Yihan Hou, Manling Yang, Hao Cui, Lei Wang, Jie Xu, and Wei Zeng. 2024. C2Ideas: Supporting Creative Interior Color Design Ideation with a Large Language Model. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Article 172, 18 pages. doi:10.1145/3613904.3642224
- [20] Ruizhen Hu, Ziqi Ye, Bin Chen, Oliver van Kaick, and Hui Huang. 2023. Self-Supervised Color-Concept Association via Image Colorization. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2023), 247–256. doi:10.1109/TVCG.2022.3209481
- [21] Nicholas Humphrey. 2019. *Colour for Architecture Today*. Taylor & Francis. 9–12 pages.
- [22] Angel Hsing-Chi Hwang. 2022. Too Late to be Creative? AI-Empowered Tools in Creative Processes. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI EA '22). Association for Computing Machinery, New York, NY, USA, Article 38, 9 pages. doi:10.1145/3491101.3503549

- [23] Ali Jahanian, Shaiyan Keshvari, S. V. N. Vishwanathan, and Jan P. Allebach. 2017. Colors – Messengers of Concepts: Visual Design Mining for Learning Color Semantics. *ACM Trans. Comput.-Hum. Interact.* 24, 1 (2017). doi:10.1145/3009924
- [24] Domicèle Jonauskaite, Jörg Wicker, Christine Mohr, Nele Dael, Jelena Havelka, Marietta Papadatou-Pastou, Meng Zhang, and Daniel Oberfeld. 2019. A Machine Learning Approach to Quantify the Specificity of Colour–emotion Associations and Their Cultural Differences. *Royal Society Open Science* 6, 9 (2019), 190741. doi:10.1098/rsos.190741
- [25] Bahjat Kawar, Shiran Zada, Oran Lang, Omer Tov, Huiwen Chang, Tali Dekel, Inbar Mosseri, and Michal Irani. 2023. Imagic: Text-Based Real Image Editing with Diffusion Models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 6007–6017. doi:10.1109/CVPR52729.2023.00582
- [26] Chris Kim, Uta Hinrichs, Saif M. Mohammad, and Christopher Collins. 2020. Lexichrome: Text Construction and Lexical Discovery with Word–Color Associations Using Interactive Visualization. In *Proceedings of the ACM Designing Interactive Systems Conference* (Eindhoven, Netherlands) (DIS '20). Association for Computing Machinery, New York, NY, USA, 477–488. doi:10.1145/3357236.3395503
- [27] Hye-Rin Kim, Min-Joon Yoo, Henry Kang, and In-Kwon Lee. 2014. Perceptually-based Color Assignment. *Computer Graphics Forum* 33, 7 (2014), 309–318. doi:10.1111/cgf.12499
- [28] Alexander Kirillov, Eric Mintun, Nikhila Ravi, Hanzi Mao, Chloe Rolland, Laura Gustafson, Tete Xiao, Spencer Whitehead, Alexander C. Berg, Wan-Yen Lo, Piotr Dollar, and Ross Girshick. 2023. Segment Anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. 4015–4026. doi:10.1109/ICCV51070.2023.00371
- [29] Shigenobu Kobayashi. 1981. The Aim and Method of the Color Image Scale. *Color Research & Application* 6, 2 (1981), 93–107. doi:10.1002/col.5080060210
- [30] Sevinc Kurt and Kelechi Kingsley Osueke. 2014. The Effects of Color on the Moods of College Students. *Sage Open* 4, 1 (2014). doi:10.1177/2158244014525423
- [31] Inseok Lee and Won-Gue Hwang. 2015. Effects of Personal Experiences on the Interpretation of the Meaning of Colours Used in the Displays and Controls in Electric Control Panels. *Ergonomics* 58, 12 (2015), 1974–1982. doi:10.1080/00140139.2015.1047803
- [32] Young-Jin Lee and Joohyeon Lee. 2006. The Development of An Emotion Model Based on Colour Combinations. *International Journal of Consumer Studies* 30, 2 (2006), 122–136. doi:10.1111/j.1470-6431.2005.00457.x
- [33] Xinghui Li, Jingyi Liu, Kai Han, and Victor Adrian Prisacariu. 2024. SD4Match: Learning to Prompt Stable Diffusion Model for Semantic Matching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 27558–27568.
- [34] Juncong Lin, Pintong Xiao, Yinan Fu, Yubin Shi, Hongran Wang, Shihui Guo, Ying He, and Tong-Yee Lee. 2022. C3 Assignment: Camera Cubemap Color Assignment for Creative Interior Design. *IEEE Transactions on Visualization and Computer Graphics* 28, 8 (2022), 2895–2908. doi:10.1109/TVCG.2020.3041728
- [35] Sharon Lin, Julie Fortuna, Chinmay Kulkarni, Maureen Stone, and Jeffrey Heer. 2013. Selecting Semantically-Resonant Colors for Data Visualization. *Computer Graphics Forum* 32, 3pt4 (2013), 401–410. doi:10.1111/cgf.12127
- [36] Albrecht Lindner, Nicolas Bonnier, and Sabine Süßstrunk. 2012. What is the Color of Chocolate? – Extracting Color Values of Semantic Expressions. In *Conference on Colour in Graphics, Imaging, and Vision*, Vol. 6. Society of Imaging Science and Technology, 355–361. <https://api.semanticscholar.org/CorpusID:2082872>
- [37] Shilong Liu, Zhaoyang Zeng, Tianhe Ren, Feng Li, Hao Zhang, Jie Yang, Chunyuan Li, Jianwei Yang, Hang Su, Jun Zhu, et al. 2023. Grounding DINO: Marrying DINO with Grounded Pre-Training for Open-Set Object Detection. *arXiv preprint arXiv:2303.05499* (2023).
- [38] Vivian Liu and Lydia B Chilton. 2022. Design Guidelines for Prompt Engineering Text-to-Image Generative Models. In *Proceedings of the CHI conference on human factors in computing systems*. 1–23. doi:10.1145/3491102.3501825
- [39] Andreas Lugmayr, Martin Danelljan, Andres Romero, Fisher Yu, Radu Timofte, and Luc Van Gool. 2022. RePaint: Inpainting using Denoising Diffusion Probabilistic Models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 11461–11471. doi:10.1109/CVPR52688.2022.01117
- [40] Jean-Baptiste Michel, Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, The Google Books Team, Joseph P. Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, Steven Pinker, Martin A. Nowak, and Erez Lieberman Aiden. 2011. Quantitative Analysis of Culture Using Millions of Digitized Books. *Science* 331, 6014 (2011), 176–182. <http://www.sciencemag.org/content/331/6014/176.full>
- [41] Kushin Mukherjee, Timothy T Rogers, and Karen B Schloss. 2024. Large Language Models Estimate Fine-grained Human Color-concept Associations. *arXiv preprint arXiv:2406.17781* (2024).
- [42] Kushin Mukherjee, Brian Yin, Brianne E. Sherman, Laurent Lessard, and Karen B. Schloss. 2022. Context Matters: A Theory of Semantic Discriminability for Perceptual Encoding Systems. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2022), 697–706. doi:10.1109/TVCG.2021.3114780
- [43] Nicolas M Müller and Karla Markert. 2019. Identifying Mislabeled Instances in Classification Datasets. In *2019 International Joint Conference on Neural Networks (IJCNN)*. 1–8. doi:10.1109/IJCNN.2019.8851920
- [44] Randall Munroe. 2010. Color Survey Results. <https://blog.xkcd.com/2010/05/03/color-survey-results/>
- [45] Ranjita Naik and Besmira Nushi. 2023. Social Biases through the Text-to-Image Generation Lens. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (Montréal, QC, Canada) (AIES '23). Association for Computing Machinery, New York, NY, USA, 786–808. doi:10.1145/3600211.3604711
- [46] Harold G Nelson and Erik Stoltzman. 2014. *The Design Way: Intentional Change in An Unpredictable World*. MIT press.
- [47] Annie WY Ng and Alan HS Chan. 2018. Color Associations among Designers and Non-designers for Common Warning and Operation Concepts. *Applied Ergonomics* 70 (2018), 18–25. doi:10.1016/j.apergo.2018.02.004
- [48] Quang Nguyen, Truong Vu, Anh Tran, and Khoi Nguyen. 2024. Dataset Diffusion: Diffusion-based Synthetic Data Generation for Pixel-level Semantic Segmentation. In *Proceedings of the International Conference on Neural Information Processing Systems*. Vol. 36. Curran Associates Inc.
- [49] Jonas Oppenlaender. 2022. The Creativity of Text-to-Image Generation. In *Proceedings of the 25th International Academic Mindtrek Conference* (Tampere, Finland) (Academic Mindtrek '22). Association for Computing Machinery, New York, NY, USA, 192–202. doi:10.1145/3569219.3569352
- [50] Hadas Orgad, Bahjat Kawar, and Yonatan Belinkov. 2023. Editing Implicit Assumptions in Text-to-Image Diffusion Models. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. 7030–7038. doi:10.1109/ICCV51070.2023.00649
- [51] Charles E Osgood. 1957. *The Measurement of Meaning*. University of Illinois press.
- [52] Li-Chen Ou, M. Ronnier Luo, Andrée Woodcock, and Angela Wright. 2004. A Study of Colour Emotion and Colour Preference. Part II: Colour Emotions for Two-colour Combinations. *Color Research & Application* 29, 4 (2004), 292–298. doi:10.1002/col.20024
- [53] Stephen E. Palmer and Karen B. Schloss. 2010. An Ecological Valence Theory of Human Color Preference. *Proceedings of the National Academy of Sciences* 107, 19 (2010), 8877–8882. doi:10.1073/pnas.0906172107
- [54] Guy Parsons. 2022. The DALL-E 2 Prompt Book. <https://dall-e.gallery/the-dalle-2-prompt-book/>
- [55] Shah Prerak. 2024. Addressing Bias in Text-to-Image Generation: A Review of Mitigation Methods. In *Proceedings of the International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN)*. 1–6. doi:10.1109/ICSTSN61422.2024.10671230
- [56] Qianru Qiu, Mayu Otani, and Yuki Iwazaki. 2022. An Intelligent Color Recommendation Tool for Landing Page Design. In *Companion Proceedings of the International Conference on Intelligent User Interfaces*. Association for Computing Machinery, 26–29. doi:10.1145/3490100.3516450
- [57] Ragini Rathore, Zachary Leggon, Laurent Lessard, and Karen B. Schloss. 2020. Estimating Color-Concept Associations from Image Statistics. *IEEE Transactions on Visualization and Computer Graphics* 26, 1 (2020), 1226–1235. doi:10.1109/TVCG.2019.2934536
- [58] Tianhe Ren, Shilong Liu, Ailing Zeng, Jing Lin, Kunchang Li, He Cao, Jiayu Chen, Xinyu Huang, Yukang Chen, Feng Yan, Zhaoyang Zeng, Hao Zhang, Feng Li, Jie Yang, Hongyang Li, Qing Jiang, and Lei Zhang. 2024. Grounded SAM: Assembling Open-World Models for Diverse Visual Tasks. *arXiv preprint arXiv:2401.14159* (2024).
- [59] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. 2022. High-Resolution Image Synthesis With Latent Diffusion Models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 10684–10695.
- [60] Nataniel Ruiz, Yuanzhen Li, Varun Jampani, Yael Pritch, Michael Rubinstein, and Kfir Aberman. 2023. DreamBooth: Fine Tuning Text-to-Image Diffusion Models for Subject-Driven Generation. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 22500–22510. doi:10.1109/CVPR52729.2023.02155
- [61] Chitwan Saharia, William Chan, Saurabh Saxena, Lala Lit, Jay Whang, Emily Denton, Seyed Kamyar Seyed Ghasemipour, Burcu Karagol Ayan, S. Saro Mahdavi, Raphael Gontijo-Lopes, Tim Salimans, Jonathan Ho, David J Fleet, and Mohammad Norouzi. 2022. Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding. In *Proceedings of the Advances in Neural Information Processing Systems*. 36479–36494. https://proceedings.neurips.cc/paper_files/paper/2022/file/ec795aeadae0b7d230fa35cbaf04c041-Paper-Conference.pdf
- [62] Pamela Samuelson. 2023. Generative AI meets copyright. *Science* 381, 6654 (2023), 158–161. doi:10.1126/science.adf0656 arXiv:<https://www.science.org/doi/pdf/10.1126/science.adf0656>
- [63] Frode Eika Sandnes. 2021. Inverse Color Contrast Checker: Automatically Suggesting Color Adjustments that meet Contrast Requirements on the Web. In *Proceedings of the International ACM SIGACCESS Conference on Computers and Accessibility (Virtual Event, USA) (ASSETS '21)*. Association for Computing Machinery, New York, NY, USA, Article 72, 4 pages. doi:10.1145/3441852.3476529
- [64] Meri Bülent Sarıylıdz, Karteek Alahari, Diane Larlus, and Yannis Kalantidis. 2023. Fake it Till You Make it: Learning Transferable Representations from Synthetic ImageNet Clones. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 8011–8021. doi:10.1109/CVPR52729.2023.00774

- [65] Karen B Schloss, Laurent Lessard, Charlotte S Walmsley, and Kathleen Foley. 2018. Color Inference in Visual Communication: The Meaning of Colors in Recycling. *Cognitive Research: Principles and Implications* 3 (2018), 1–17. doi:10.1186/s41235-018-0090-y
- [66] Matthew Law Seo-young Lee and Guy Hoffman. 2024. When and How to Use AI in the Design Process? Implications for Human-AI Design Collaboration. *International Journal of Human-Computer Interaction* 0, 0 (2024), 1–16. doi:10.1080/10447318.2024.2353451
- [67] V. Setlur and M. C. Stone. 2016. A Linguistic Approach to Categorical Color Assignment for Data Visualization. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2016), 698–707. doi:10.1109/TVCG.2015.2467471
- [68] Yu-Chuan Shen, Wu-Hsiung Yuan, Wen-Hsing Hsu, and Yung-Sheng Chen. 2000. Color Selection in The Consideration of Color Harmony for Interior Design. *Color Research and Application* 25, 1 (2000), 20–31. doi:10.1002/(SICI)1520-6378(200002)25:1<20::AID-COLA>3.0.CO;2-5
- [69] Xinyu Shi, Ziqi Zhou, Jing Wen Zhang, Ali Neshati, Anjul Kumar Tyagi, Ryan Rossi, Shunai Guo, Fan Du, and Jian Zhao. 2023. De-Stijl: Facilitating Graphics Design with Interactive 2D Color Palette Recommendation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Article 122, 19 pages. doi:10.1145/3544548.3581070
- [70] Yang Shi, Siji Chen, Pei Liu, Jiang Long, and Nan Cao. 2022. ColorCook: Augmenting Color Design for Dashboarding with Domain-Associated Palettes. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW2, Article 433 (nov 2022), 25 pages. doi:10.1145/3555534
- [71] Yang Shi, Yechun Peng, Shengqi Dang, Nanxuan Zhao, and Nan Cao. 2024. Personalizing Products with Stylized Head Portraits for Self-Expression. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–18. doi:10.1145/3613904.3642391
- [72] Ioannis Siglidis, Aleksander Holynski, Alexei A Efros, Mathieu Aubry, and Shiry Ginosar. 2024. Diffusion Models as Data Mining Tools. *arXiv preprint arXiv:2408.02752* (2024).
- [73] Danielle Albers Szafrań. 2018. Modeling Color Difference for Visualization Design. *IEEE Transactions on Visualization and Computer Graphics* 24, 1 (2018), 392–401. doi:10.1109/TVCG.2017.2744359
- [74] Diana Su Yun Tham, Paul T Sowden, Alexandra Grandison, Anna Franklin, Anna Kai Win Lee, Michelle Ng, Juhyun Park, Weiguo Pang, and Jingwen Zhao. 2020. A Systematic Investigation of Conceptual Color Associations. *Journal of Experimental Psychology: General* 149, 7 (2020), 1311. doi:10.1037/xge0000703
- [75] Matthias Peter Verheyden and Matthias Funk. 2023. Collaborative Diffusion: Boosting Designerly Co-Creation with Generative AI. In *Extended abstracts of the 2023 CHI conference on human factors in computing systems*. 1–8. doi:10.1145/3544549.3585680
- [76] Yixin Wan, Arjun Subramonian, Anaelia Ovalle, Zongyu Lin, Ashima Suvarna, Christina Chance, Hritik Bansal, Rebecca Pattichis, and Kai-Wei Chang. 2024. Survey of Bias In Text-to-Image Generation: Definition, Evaluation, and Mitigation. *arXiv preprint arXiv:2404.01030* (2024).
- [77] ANNA WIERZBICKA. 1990. The meaning of color terms: semantics, culture, and cognition. *Cognitive Linguistics* 1, 1 (1990), 99–150. doi:doi:10.1515/cogl.1990.1.1.99
- [78] Sam Witteveen and Martin Andrews. 2022. Investigating Prompt Engineering in Diffusion Models. *arXiv preprint arXiv:2211.15462* (2022).
- [79] Benjamin Wright and Lee Rainwater. 1962. The Meanings of Color. *The Journal of General Psychology* 67, 1 (1962), 89–99.
- [80] Shishi Xiao, Liangwei Wang, Xiaojuan Ma, and Wei Zeng. 2024. TypeDance: Creating semantic typographic logos from image through personalized generation. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–18. doi:10.1145/3613904.3642185
- [81] Lymii Zhang, Anyi Rao, and Maneesh Agrawala. 2023. Adding Conditional Control to Text-to-Image Diffusion Models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 3836–3847.

A Sensitivity of the Diffusion Model

In our study, we observed that the T2I model exhibits varying sensitivity to different types of prompts. Specifically, the model is more responsive to prompts with visual descriptions compared to abstract or non-visual concepts. To address this, we enhanced the prompts by incorporating related visual descriptions, which helped the model better capture the intended context. Instead of only use “quiet”, we added “*evoking feelings of silence and lonely*” to provide a more concrete visual description.

Figure 16 demonstrates this effect. The top row, prompted with “quiet forest,” includes images that, despite the prompt, exhibit open and sunlit scenes (highlighted in red), which contradict the expected mood. The second row, with the enhanced prompt, shows a more consistent adherence to the quiet and solitary atmosphere. Similarly, the bottom two rows compare “lively forest” with its enhanced version; the latter produces images with a more vibrant and hopeful tone, as intended, with the red boxes indicating images that still convey a sense of depth and solitude rather than liveliness.

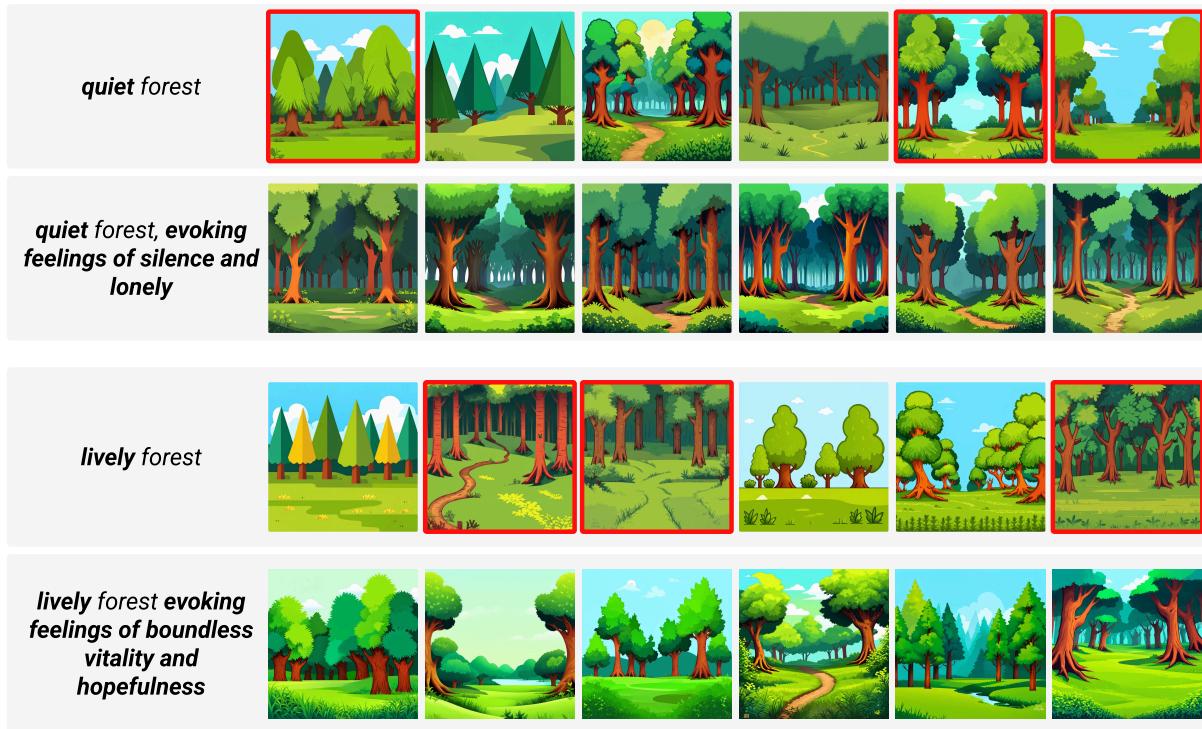


Figure 16: Sensitivity of the T2I model to different types of prompts, comparing the original prompts with enhanced prompts. The red boxes highlight images that do not align with the intended concept.