Influencer: Empowering Everyday Users in Creating Promotional Posts via AI-infused Exploration and Customization

Xuye Liu
University of Waterloo
Waterloo, Ontario, Canada
xuye.liu@uwaterloo.ca

Annie Sun
University of Waterloo
Waterloo, Ontario, Canada
a34sun@uwaterloo.ca

Pengcheng An
Southern University of Science and
Technology
Shenzhen, Guangdong, Canada
anpc@sustech.edu.cn

Tengfei Ma
Stony Brook University
Stony Brook, New York, United States
tengfei.ma@stonybrook.edu

Jian Zhao
University of Waterloo
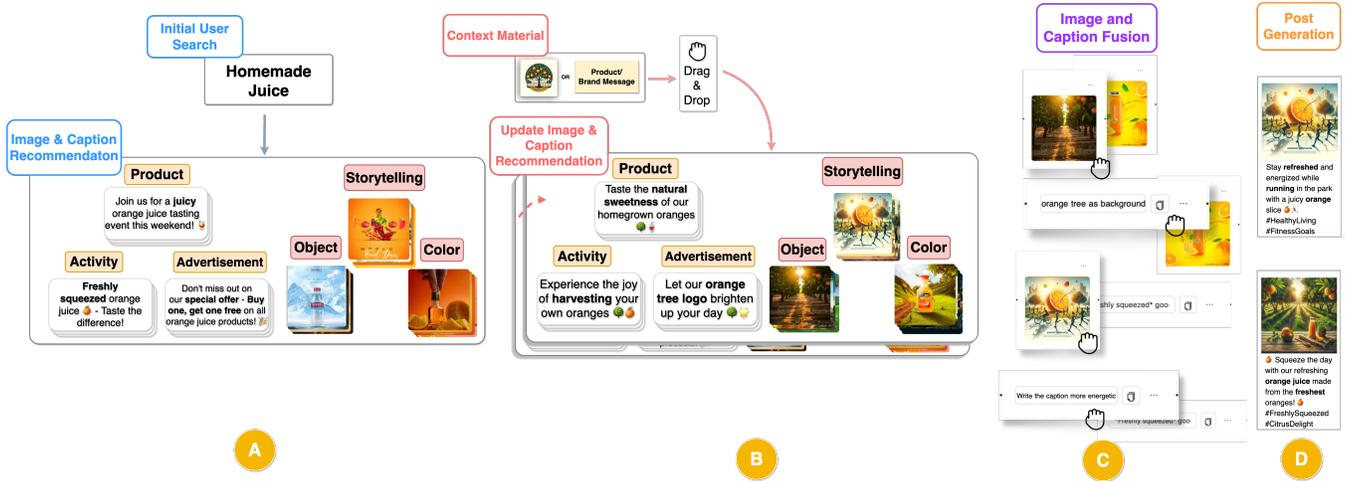Waterloo, Ontario, Canada
jianzhao@uwaterloo.ca

Figure 1: A design novice uses Influencer to ideate and make promotional posts to promote their homemade juice. Influencer has the following core features: (A) The user can input a topic via a text block and explores the related images and captions in three dimensions. (B) Context-aware exploration is supported which updates the image and caption recommendation by dragging a brand/product image or message to the initial image and caption recommendation. (C) Various materials (i.e., image and text) can be flexibly fused to make a new image or caption. (D) Influencer allows the user to not only easily create harmonious promotional posts but also quickly obtain multiple post alternatives. Steps in (A), (B), and (C) can be flexibly combined or skipped; as soon as the user finds satisfied image and/or caption, they can go to (D) for post generation.

ABSTRACT

Creating promotional posts on social platforms enables everyday users to disseminate their creative outcomes, engage in community exchanges, or generate additional income from micro-businesses. However, crafting eye-catching posts with appealing images and effective captions can be challenging and time-consuming for everyday users since they are mostly design novices. We propose Influencer, an interactive tool that helps novice creators quickly generate ideas and create high-quality promotional post designs through AI. Influencer offers a multi-dimensional recommendation system for ideation through example-based image and caption suggestions. Further, Influencer implements a holistic promotional post-design system supporting context-aware exploration considering brand messages and user-specified design constraints, flexible fusion of content, and a mind-map-like layout for idea tracking. Our user study, comparing the system with industry-standard tools, along with two real-life case studies, indicates that Influencer is effective in assisting design novices to generate ideas as well as creative and diverse promotional posts with user-friendly interaction.

CCS CONCEPTS

• **Human-centered computing** → **Systems and tools for interaction design**; • **Computing methodologies** → *Artificial intelligence*; • **Applied computing** → *Arts and humanities*.

KEYWORDS

1 INTRODUCTION

Creating promotional posts on social platforms enables everyday users to share their creativity and engage in micro-entrepreneurship, an emerging economic component representing flexible, small-scale business ventures that offer substantial growth potential [54]. By creating promotional posts on social media, "micro-entrepreneurs" could showcase their products or services (such as selling second-hand items, and homemade cakes [43]), engage with target customers, and build their online presence [1, 41]. This helps individuals, such as low-income groups or college students, to generate additional income for improving their life qualities, thus contributing to the resilience of society [38].

Promotional posts need to be crafted with an eye-catching design and packed with informative content to boost brand visibility, engage customers effectively, and ultimately pave the way for higher conversions and sales [27]. Hence, a successful promotional post requires high-level designer skills. Namely, designers need to iteratively create engaging visuals and effective messages, and thoughtfully blend the two into a harmonious whole.

Therefore, for design novices and everyday users, creating promotional posts is challenging in several aspects. First, in the ideation stage, since users often draw inspiration from existing images [24], it is challenging for them to create novel alternatives from familiar visual narrations [66]. Previous research indicates that exploring examples helps users better understand their design direction and facilitates their ideation step from previous creations than starting from scratch [6, 21, 59]. Thus, many users rely on existing image search engines (e.g., Google Image Search, Pinterest) to collect examples during the ideation stage; however, this process can be very time-consuming [60]. Researchers have developed tools (e.g., MetaMap [32], PopBlend [69]) to support example exploration in the design ideation process, but the search results (design examples) cannot be flexibly edited or modified. Also, users still need to rely on additional retouching tools (e.g., Adobe Photoshop[1], Powerpoint) to create and refine their designs.

Moreover, the coordination between images and captions is essential for creating coherent and persuasive promotional posts that effectively convey the intended brand message and campaign information. However, matching the image to the message, or vice versa, can be a challenge in practice, due to the particular themes of the designed post. Without readily usable image resources, users often need to manually create an image to satisfy design constraints and align them with the theme [26]. Similarly, users need to iterate the caption appropriately to match the image for accurate conveyance

of the intended message. These design hurdles escalate the threshold of promotional post design for non-professionals. While there exists some image editing tools (e.g., Adobe Photoshop) and image captioning tools (e.g., jina[2]), they are either targeted at professional users or lack the flexibility and integrativeness for a streamlined unified workflow: e.g., images and captions need to be created separately and combined manually at a later stage. Furthermore, such a complex, tedious, and time-consuming process makes it difficult for novice users to create a sufficient number of design alternatives to compare and evolve their post, which is essential for having a good design outcome.

To address these challenges, we propose Influencer (Figure 1), an AI-empowered interactive tool for users who do not have professional design skills but need to create promotional posts in their daily lives (e.g., individual creators, college students, small business operators, product managers, freelancers). Drawing design goals from a formative study with five participants with different levels of expertise in design, Influencer is designed to streamline the promotional post design process, featuring four design modules: ideation, context-aware exploration, iterative customization, and delivery of multiple alternatives. Specifically, it provides three modules in the AI-infused pipeline: 1) an example-based multi-dimensional recommendation framework to facilitate the exploration of related seed captions and images to inspire design ideas, 2) a context-aware exploration module that enables users to add complex design constraints, including the requirement of matching color schema of brand image or product message, and 3) flexible fusion for various design materials via LLMs and Generative Image AI models. Besides, it provides an interactive mind-map interface to organize ideas, track thought processes, and present multiple design alternatives. To evaluate Influencer, we conducted a controlled experiment with 12 design novices comparing our system with a baseline resembling the current workflow (Google + ChatGPT + Powerpoint). To further learn how the Influencer can be integrated into a novice's workflow, we conducted two promotional post design cases by inviting two participants to use the system based on their real-life needs in two one-hour hands-on sessions. The results indicate that Influencer's features were appreciated and the system is effective in helping users generate more creative and higher-quality promotional posts. In summary, we make the following contributions:

- A comprehensive AI-infused pipeline for effective ideation and generation of promotional posts, which integrates multidimensional recommendation, context-aware exploration, and conceptual fusion mechanisms for both images and captions.
- A novel interactive canvas-based tool enabled by the pipeline, Influencer, which facilitate users with designing promotional posts and organizing them in a mind-map-like layout to track their thought processes.
- An in-depth comparative evaluation of Influencer and case studies that generate insights into whether and how the tool is effective in supporting promotional post design.

---

[1]https://www.adobe.com/ca/products/photoshop.html

[2]https://jina.ai/

2 RELATED WORK

2.1 Example-based Design and Ideation

Past research indicates that exploring and organizing a wide range of design examples enable designers to gain potential approaches for implementing their ideas [21]. Studies have shown that capturing and managing the context and rationale behind design examples can further enhance the design process and foster innovation [20, 57]. An effective way to generate novel ideas is to utilize example-based exploration rather than starting from scratch [59]. When individuals heavily rely on existing knowledge and examples [30], they may encounter design fixation. In addition, simply providing irrelevant inspirations can not effectively address this issue and can even influence design efficiency [10]. Hence, it is crucial to maintain a balance between interrelatedness and diversity in example-based recommendations to foster designers' creativity [47]. Psychologists have investigated the relationship between human creativity and association [25, 42]. They find that generating creative ideas demands a strong capability for finding an association, rather than only imitating past work [7].

To facilitate this process, researchers have published free association datasets to simplify the ideation process among the general public [18]. The small world project[3] is currently the largest dataset for word associations in English with more than 12,000 cue words, supporting users' conceptual level ideation. It demonstrated the importance of facilitating the ideation process at the concept level in promotional post design [6, 30]. Tools have also been proposed to leverage the example-based approach to facilitate ideation in design activities such as generating compound icons [74] and visual metaphors [32]. Our work is grounded by the findings in the above literature and proposes a novel multi-dimensional ideation method by analyzing the characteristics of images and captions.

2.2 Image and Caption Recommendations

Promotional post designs require the seamless integration of a promotional image and its caption [63]. Jeong's pioneering work [31] mentions that posts with visual metaphors are more persuasive compared to posts with literal (non-metaphorical) images. Therefore, we decide to use visual metaphors as our main image resource in our system.

To ease the design process, various tools have been developed to assist designers in crafting high-quality images. It is generally required to identify the analogy between two images to have a better chance of finding related images [32]. The most common correlation between images is semantic relevance. In addition, both color [33] and object [48] play important roles in establishing metaphorical relationships. Prior research has explored the semantic meanings of colors [29, 33, 71], aiming to establish links between colors and semantics. Besides, objects and storytelling are also important when recommending related images, and designers tend to seek similar elements with similar objects for potential image exploration [23]. Thus, in our work, we seek to use a three-dimensional recommendation framework (storytelling, color, and object) for image exploration in the ideation step.

Moreover, prior research has shown that creating interesting and related captions can improve interaction and the sharing of content on promotional post design [28, 73]. Traditional caption generation methods [53, 61] for posts mainly use natural language techniques to create captions just from the picture, not considering the post's actual use and focusing only on one aspect. However, these methods do not fully meet the varied requirements of different types of posts. Promotional posts can be categorized into three main topics: activity, product, and advertisement. Activity-focused ones entail content that centers on engaging actions or events. They aim to grab the audience's attention by showcasing active happenings, events, or efforts linked to a brand or group [55]. Product-related posts seek to convey the attributes, advantages, and value of particular products or services. This strategy aims to attract potential consumers and encourage purchase decisions [5]. Advertisement-focused posts encompass broader marketing and branding initiatives [34]. They aim to promote the brand as a whole, establish its identity, and convey its unique selling propositions. In our work, we incorporate caption recommendation in three dimensions (activity, product, advertisement) in Influencer to facilitate the caption exploration in promotional post design.

2.3 Promotional Post Design Systems

There are a variety of off-the-shelf tools that support post design with different focuses, such as professional image editing (e.g., Adobe Photoshop), collaborative UI design (e.g., Powerpoint), and user-friendly graphic design (e.g., Canva). While all these tools provide powerful post image editing features from template-level to object-level, they have a steep learning curve for novice designers and require much manual user effort.

The research community has been exploring AI-driven technologies for design assistance. Most has primarily focused on image customization, including masking and editing specific areas, generating new images based on a similar prompt context, and regenerating parts of images for variations [51]. AnyDoor [12] uses a diffusion-based deep learning method to support object-level image customization. Some works also try to incorporate text or image concepts to customize target images. Chen et al. [11] enabled various prompt inputs by incorporating a dual-branch conditioning mechanism to customize images in different concepts. Kumari et al. [35] supported composing multiple concepts to generate images based on the given texts. Beyond image customization, the generation and customization of captions for promotional posts also facilitate the promotional post design process. Large Language Models (LLM) including ChatGPT [44], Llama [65], and Claude [3] have been proposed to customize captions by rewriting the prompt based on the user's need. More recent studies have used Context Sequence Memory Network to customize descriptive captions and predict hashtags based on the query images [16, 46]. In this work, we integrate multiple customization techniques such as regenerating images with similar prompt context and mask editing images into our tool to improve the promotional image design process, empowered by simplified and practical interaction on a mind map.

Additionally, researchers have proposed various intelligent tools to help novices automate parts of the design workflow. Yin et al. [72] developed a system for automatically generating magazine-like

---

[3]https://smallworldofwords.org/en/project/home

visual summaries from traditional social media posts for efficient mobile browsing. Qiang et al. [64] introduced a graphical model that learns to generate scientific posters from research papers. There has also been a variety of work in domain-specific applications with compound design tasks [58, 67, 74]. Some researchers recently developed tools that leverage generative AI to support graphic design ideation and message-driven infographic creation for visual storytelling and ideation [14, 17, 75]. Recent work in aiding creative design [4] also adopts mind-maps to mimic the process of making associations among design materials. However, all of the above tools mainly focus on the ideation step of design, requiring designers to switch to other tools for the actual design work. Also, none of the above tools accommodates user-specified prompt or context images (e.g. brand image) based on design constraints, which are essential to maintain consistency and effectively communicate with the target audience.

3 FORMATIVE STUDY

Our primary target audience is design novices such as product managers, marketers, small business operators, freelance creators, or anyone who needs to quickly make attractive promotional posts without professional training. To better understand their needs, we conducted a formative study, and here we describe its setup and yielded design goals that have guided the development of Influencer.

3.1 Setup

We recruited five participants with different levels of expertise in design (two women and three men, aged 20-30), including two design researchers in the IT company, a marketer experienced in designing promotional posts, and two software engineers who collaborate with product managers, marketers, and designers on product promotion. The study included a 30-minute semi-structured interview with each participant. The interview questions covered how they would normally design promotional posts, including how they got their inspiration, and how they would craft and iterate the posts. They were also asked to identify difficulties they encountered throughout the design process and raise their needs for support. The interviews were audio recorded and then transcribed for further analysis. Five authors used the open-coding method to independently conduct a thematic analysis of the interview transcripts and notes. Then, all five authors collaboratively reviewed and synthesized the results to ensure a comprehensive understanding of the findings.

3.2 Formulating Design Goals

Based on our formative study, we distill the following design goals to inform our development of Influencer. We refer to the participants in our interviews as E#.

3.2.1 R1: Help explore images and captions through multi-dimensional, context-aware recommendations. Traditional methods to collect inspiring examples of images and captions for the ideation stage include searching online resources from search engines (e.g., Google Search), which is usually cumbersome and time-consuming. Exposure to rich information often sparks new ideas [6, 37]. When talking about searching exemplar images and captions for inspiration, two design novices (E1, E3) reported that

they did not have effective techniques for using search engines for this purpose. Four (E1-2, E4-5) suggested a recommendation system that recommends similar images based on an example image provided by the user would be desirable for this task. Three participants (E2-4) looked for diverse stimuli from others' work, including storytelling, color schemes, and object relevance. By implementing example-based multi-dimensional recommendations (i.e., storytelling, color, and object), designers can explore diverse images with any additional knowledge behind the recommendation.

The caption in a promotional post is equally important as images since it reflects the design intention and usage scenarios [28]. E5 expressed frustrations towards the traditional approach to caption recommendation: *"Simple captions aren't enough to engage the audience and tell a great ad story."* Designers also face challenges in accommodating design material (images and captions) under user-specific constraints, such as color schemes and related product or branding requirements [22]. E2 confirmed that *"It's necessary to update material exploration based on the given images from multiple dimensions like narrative or color."* Additionally, E4 emphasized the value of prompt-based ideation: *"Using prompts efficiently helps narrow down the search, so you don't have to start over."* Therefore, a system that integrates both image and caption recommendations, while supporting context and prompt-aware exploration, is essential for promoting effective, brand-aligned design.

3.2.2 R2: Flexible fusion of various images and captions to customize designs. Traditional customization of images or captions for promotional posts involves using ChatGPT to modify prompts or manually customizing in PowerPoint, which can be cumbersome for designers. In ChatGPT, users need to modify prompts to customize contents. Four participants suggested that it would be efficient to conceptually fuse different elements via dragging and dropping to customize elements. E3 confirmed that *"fusing of elements might make it simpler to customize elements quickly"* Since users have varying preferences and expectations depending on different scenarios of the promotional post design processes, the tool must provide sufficient customization in a the interface. The tool should simplify repetitive tasks on customization and offer creative fusion options, enabling rapid design iterations and design alternatives delivery. E5 suggested giving equal importance to both image and text blocks on the canvas. This would allow users to easily create prompts or select reference images to customize captions or images, offering more flexibility. As E2 suggested: *"If text and images were treated equally, I could quickly switch between editing a caption or an image without needing to switch tools or modes."*. Thus, the tool should support flexible fusion of images and captions using either prompts or reference images.

3.2.3 R3: Design a mind-map layout to organize ideas and track the thought process. According to our interview, designers can easily get lost in the iterative exploration of images and captions during ideation without tracking their exploration path. Canvas lets designers freely create by allowing them to make, move, track, and link elements anywhere [62]. Previous work also indicates that tracking the thinking path with mind-maps during exploration is important for ideation [19]. E3 suggested that *"mind-map is one way to connect and recombine materials to generate new ideas."* Maintaining a thorough history of their creative process is a crucial design
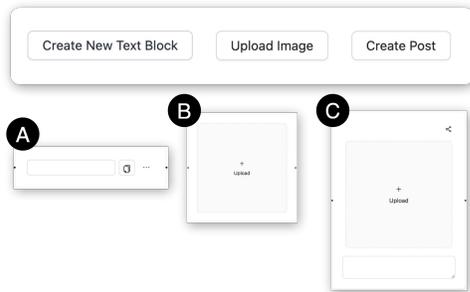
Figure 2: Influencer provides three basic types of interactive blocks on its mind-map canvas. Clicking *Create New Text Block* adds (A) an empty text block. Clicking *Upload Image* generates (B) an empty image block. Clicking *Create Post* produces (C) a blank post block that can fuse selected images and captions.

principle that aids in fostering creativity. E1, E2, and E4 suggested that *"It would be useful to use links to track materials updates. If we don't like a change, we just delete the link to go back."* Thus, the tool should support thought process tracking according to users' design actions.

4 INFLUENCER DESIGN OVERVIEW

Based on the aforementioned design goals, we developed Influencer, an interactive, AI-empowered, and canvas-based tool that helps design novices in four main steps of promotional post design (i.e., ideation, context-aware exploration, customization, and post generation), all through a set of simple interactive blocks (Figure 2). Figure 3 shows an overview of the system backend and workflow, which consists of five parts: (1) *Caption Recommendation*, (2) *Image Recommendation*, (3) *Context-Aware Exploration Mechanism*, (4) *Flexible Fusion Mechanism*, and (5) *Image Regenerate and Mast Edit Module.*

In particular, as shown in Figure 4, users' material (image and caption) searching process is supported through related concepts based on their search input. After identifying an image or caption of interest, users can further explore the recommendation images based on three features: storytelling, color, and object, and explore the recommendation captions in three contexts: product, activity, and advertisement (**R1**). As shown in Figure 5, users can conduct context-aware exploration for captions and images by dragging brand messages (i.e., brand image, product message, and brand value) to the caption and image recommendation block (**R2**). Users can further contextually customize design materials by fusing any images and captions during the promotional post design, as shown in Figure 6 (**R3**). If they already find the satisfied post, they can finalize it directly without proceeding to the additional customization, editing, or fusing steps. All the activities are conducted on a mind-map which records the thinking path generated by users for a quick recollection (**R4**).

4.1 User Interface and Blocks

The user interface of Influencer is based on an interactive canvas allowing users to create three types of blocks: text block (Figure 2.A),

image block (Figure 2.B), and post block (Figure 2.C), as well as offering two types of recommendation blocks: image and caption.

The image block provides four direct functionalities for users to further explore, customize, or utilize images in their promotional post design: 1) "Regenerate Image" (Figure 6.B) to produce a similar image with the same context, 2) "More Images"(Figure 6.B) to browse recommended images based on the selected image in three dimensions (i.e., storytelling, color, object) in a image recommendation block (Figure 4.E), which can further obtain images along the branches (Figure 4.C), 3) "Mask Edit" (Figure 6.B) to edit an image by drawing a mask and providing a prompt for generating a new image, and 4) "Generate Post"(Figure 7.A) to craft a post in a post block directly based on the current image. To minimize user interaction for efficiency, Influencer apply every user-selected or uploaded image with those four options to streamline any design activities (i.e., ideation, context-aware exploration, customization, and post generation) from the block.

The text block offers three direct functionalities (Figure 4.A) to support users in creating captions, exploring images, and generating posts during the design process: 1) "Generate Images": Recommend an assembly of seed image results based on a given topic. 2) "Generate Captions": Recommends captions across three contextual dimensions (i.e., product, activity, advertisement) within a caption recommendation block (Figure 4.F). 3) "Generate Post": Produces a complete post directly based on the user's input text within a post block (Figure 7.B).

Beyond these direct functions from text blocks and image blocks, they also support interactive exploration through drag and fuse interaction. In Influencer, images or texts can served as context to help users do further exploration and customization. Users can create a text block writing a brand message or the requirement of the promotional post design and drag it to the image (Figure 4.E) or caption recommendation block (Figure 4.F). In this way, users can get new images or caption recommendations aligning with the prompt written in the text block. Also, users can create an image block (Figure 2) and upload the context image, like a brand image, and drag it to the image (Figure 4.E) or caption recommendation block (Figure 4.F). Then the user will get a new image and caption recommendation to block aligning with his uploaded image with a similar storytelling meaning and color schema. At the same time, users can also drag a text block or an image block and connect it to any target image block or text block for conceptual customization. Such customization and context-aware exploration are shown in orange arrows to differentiate from the users' initial exploration paths that are shown in blue arrows. Furthermore, users have the option to not only generate posts based on a single image or caption but also select their preferred image and caption directly on the canvas and fuse them to a post block for production.

4.2 Usage Scenario

Before diving into our approach and details, we explain how the design novices use Influencer via a simple usage scenario. Suppose Crystal is a micro-entrepreneur who owns a small grocery store, and she needs to promote a new product, homemade orange juice.

Crystal launches Influencer. She first creates a Text Block (Figure 2.A) on the canvas and wants to explore images related to her
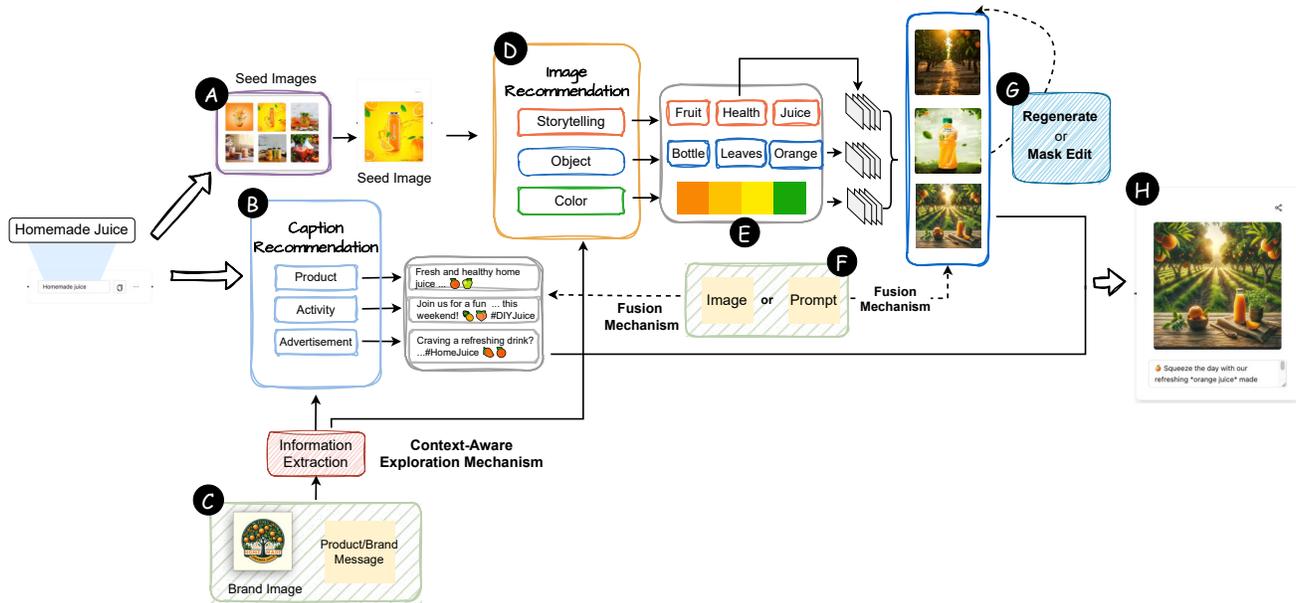
Figure 3: System architecture and workflow of Influencer. From an input topic, the system first analyzes the input to return seed images (A) and the user can select a seed image from the search results. At the same time, the Caption Recommendation (B) digests the input topic and recommends captions from three dimensions (i.e., product, activity, advertisement). Next, Influencer can conduct image recommendation (D) based on the seed image in three dimensions (i.e., storytelling, object, color) and suggest images for each dimension (E). It also supports context-aware exploration (C) based on a brand image or product/brand message to find materials aligning with the brand. Besides, Influencer allows users to customize the images via Regenerate or Mast Edit (G) and conceptually fuse design materials (i.e., images and captions) with an image or prompt reference (F). Together, it generates an aesthetically pleasing promotional post design with harmonious content. A user can further iteratively explore more design recommendations (B, C, D) and generate multiple design alternatives to share on social media (H).

promotional topic: "homemade juice." She clicks the "..." button that shows a list of exploration and post generation options (Figure 4.A). From there, she clicks the "Generate Images" button (Figure 4.A) to create a seed images block (B), consisting of a set of images related to the given topic. After identifying an image of interest, she further explores the recommendation images based on 3-dimensional features (storytelling, color, and object) by clicking on the "More Image" button (Figure 4.C). Influencer then requests her to select a related storytelling keyword from the storytelling keyword panel (Figure 4.D) (**R1**). Crystal wants the focus of the promotion to be on health. Thus she clicks Health (Figure 4.D) and creates an image recommendation block (Figure 4.E) for more images in health storytelling meaning.

Crystal looks at the images and finds a satisfied image (Figure 6) but thinks *"The image's meaning is closely related to my topic, but I need more alternatives with similar meaning and want to change the type of juice shown."* Next, she click the "Regenerate" (Figure 6.B) to generate an image with the same storytelling meaning (Figure 6.C). Using the "Mask Edit" feature, she clicks "Mask" button (Figure 6.D) to frame out the mask she wants to replace and writes the prompt (Figure 6.D). Then, she generates a new image (Figure 6.E).

At the same time, Crystal wants to further explore other alternative paths of design, since she does not have sufficient design

training and wants to see if there are better options. She tries another path to explore useful captions for her promotional post. She thus clicks the "Generate Captions" button (Figure 4.A) to obtain a caption recommendation block (Figure 4.F). She checks this caption recommendation block and finds one caption suitable for her post. She clicks this caption which generates a text block including this caption and edits the text based on her preference (Figure 7.C). Then, she directly generates a post by clicking the "Generate Post" button (Figure 7.C). Through all the above actions, she creates multiple post alternatives for her design and she is quite satisfied (Figure 7).

However, she thinks the promotional image needs to align with the brand image (Figure 5.A) of her grocery and some new descriptions of the promoted product (Figure 5.B). Using new stuff directly to find materials may ignore previous searches and context, potentially leading to inappropriate recommendations. Therefore, Crystal employs a context-aware exploration to find more materials for post design based on the previous search. She links the brand images and the text block with new product descriptions to the caption and image recommendation block. Influencer returns newly recommended images and captions in the image (Figure 6.C) and caption recommendation block (Figure 5.D) from multiple dimensions which align with the color scheme, object, and storytelling meaning of the brand image and the new product descriptions. *"Nice! I got new materials for the post design!"* Crystal and feels happy about the efficiency.
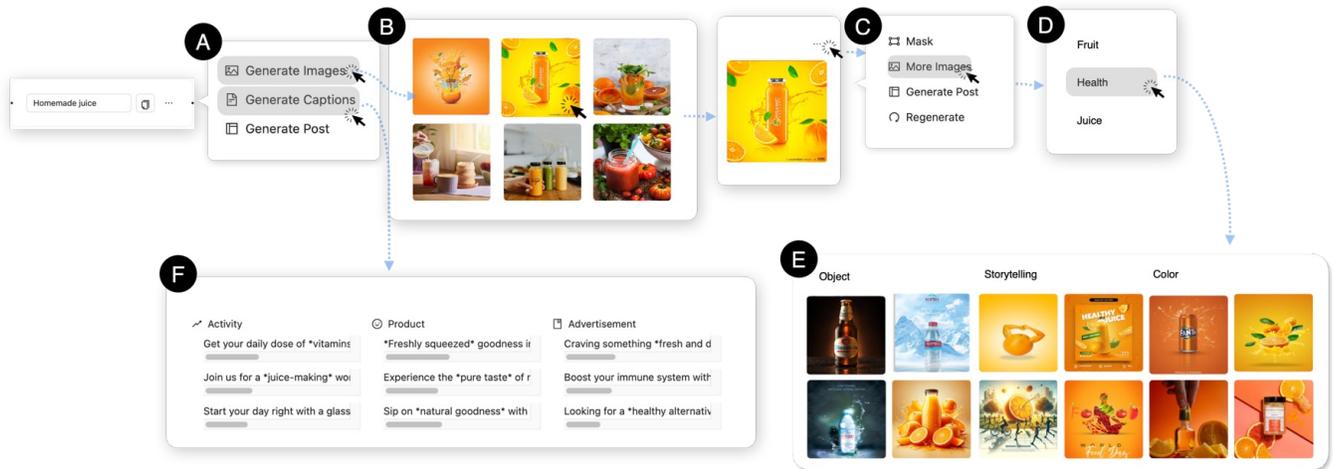
Figure 4: Ideation with Influencer: (A) From a text block, a panel providing three options (*Generate Images, Generate Captions, Generate Post*) for users to ideate captions or images and generate posts; (B) An assembly of seed image results based on users' input topic; (C) From an image block, a panel providing four options for users to customize (*Regenerate, Mask*), conduct image recommendation (*More Images*), or produce post directly (*Generate Post*); (D) A panel providing related storytelling keywords to help users explore diverse materials; (E) An image recommendation block displaying results in three dimensions (i.e., storytelling, object, color); (F) A caption recommendation block displaying results in three dimensions (i.e., product, activity, and advertisement).
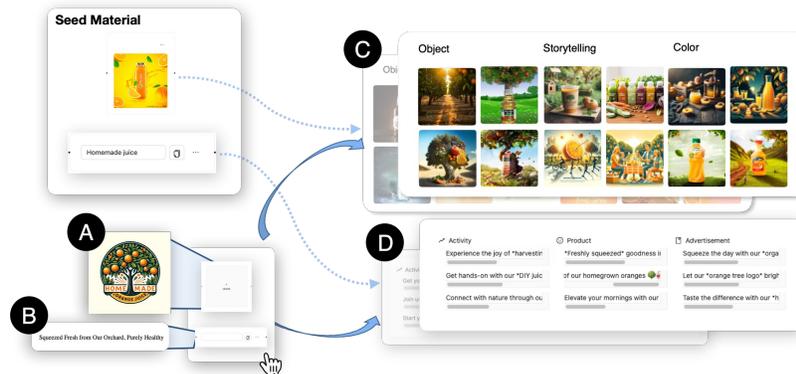

Figure 5: Context-Aware exploration with Influencer: (A) Uploading a brand image and dragging to update the target image recommendation block and caption recommendation block; (B) Creating a text block with product message and dragging to update the target image recommendation block and caption recommendation block; (C) Updated image recommendation block; (D) Updated caption recommendation block.

Using similar interactions above, she continues to customize and generate some alternative designs (Figure 7.ABC) (**R2**).

Crystal is satisfied with the content of the current image design but feels styles are less appealing when compared with another image despite not liking its content (e.g., the left two images in Figure 6.F). She wants to combine these two images; thus she connects this image as a background to that seed image for conceptual fusion (Figure 6.F). A new image is then generated with a style similar to the background image and the same storytelling meaning as the seed image (**R3**). *"Looks better!"* she thinks. If there is no satisfactory image on the canvas, she can also create a text block with the style description and drag it to the seed image for the above conceptual fusion. For example, she drags a text block with the text of "orange tree as background" to the seed image, Influencer generates a new image featuring an orange tree background while retaining the orange juice bottle from the seed image (Figure 6.G). Besides, Influencer supports conceptual fusion on captions by dragging existing images or text blocks with prompts to customize target captions (Figure 6.HI). Thus, Crystal drags a text block with a prompt suggesting "write the caption more energetic" in the target caption (Figure 6.I). She takes a look at the customized caption: *"Awesome! That was quick. This caption is exactly what I needed!"*

During the process, Crystal finds the thinking path very useful. When she get lost in the promotional post design process, the tracked thinking path can remind her of the previous exploration and customization (**R4**).
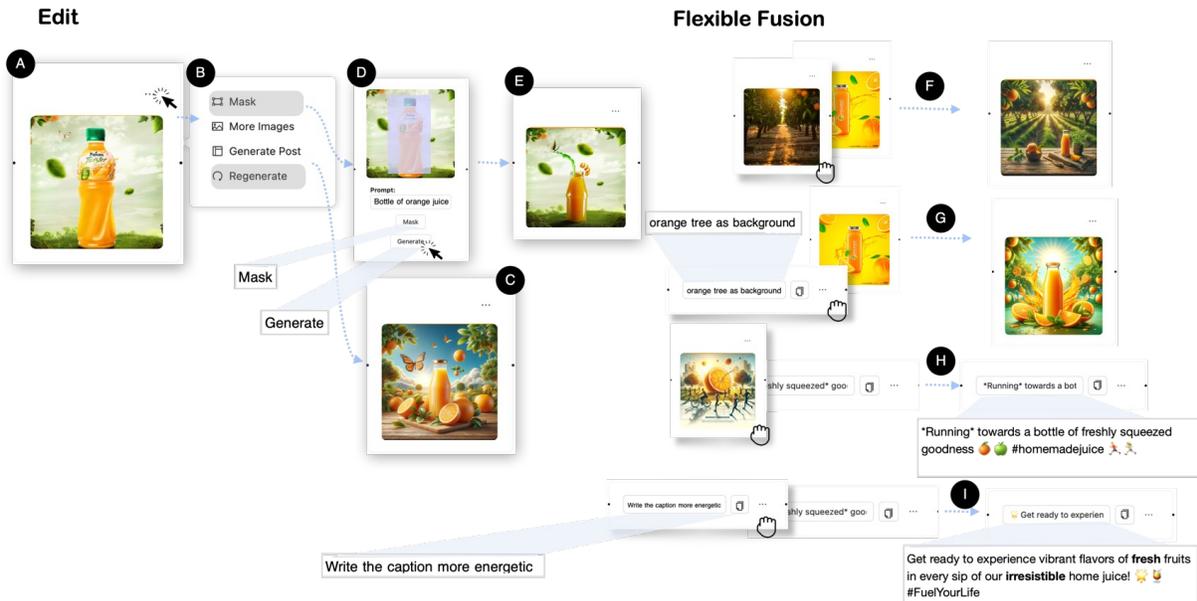
Figure 6: Customization with Influencer: In the *Edit* mode, an image (A) can be modified with "Mask" and "Regenerate" for customization (B). "Regenerate" produces a semantically similar new image (C). "Mask" allows for framing the part to be customized with a prompt (D) and thus produces a new image (E). In the *Fusion* mode, images and/or text can be freely combined by dragging an image or text block with a prompt to the target image or caption (F, G, H, I).
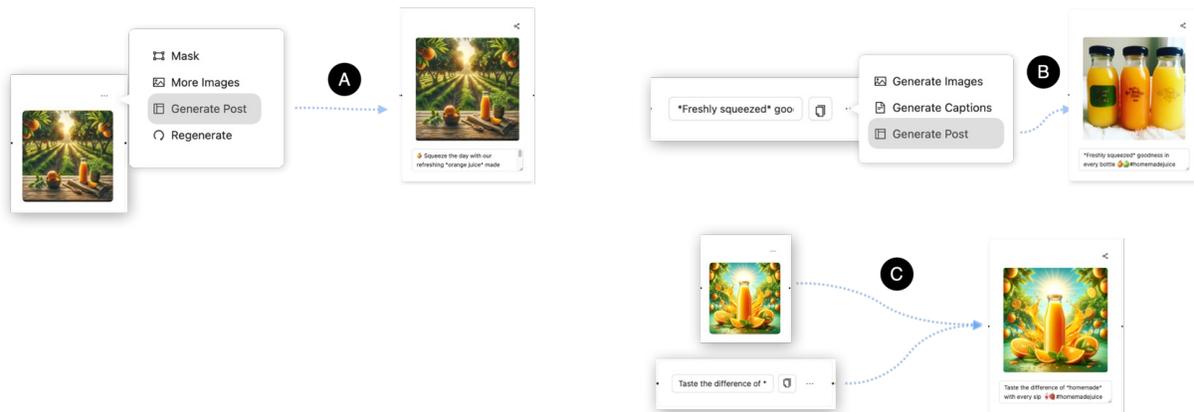

Figure 7: Post generation: A post can be generated based on a selected image (A) or caption (B) by clicking *Generate Post*. Alternatively, a blank post block can be created to combine the selected image and caption (C).

5 INFLUENCER SYSTEM

In this section, we describe the backend of the Influencer system in detail, particularly on the implementation of the multi-dimensional recommendation, context-aware exploration, and multiple fusion mechanisms (Figure 3).

5.1 Dataset Building

Previous works emphasize the importance of using associations from prior examples to create new ideas [6, 70]. Furthermore, our formative study indicates that the creation of promotional images involves analogy between all kinds of information [8], including storytelling [49, 50], color[33], and object [23]. To build the post

image dataset with reasonable associations between each example, we process the Small World Dataset [18] to get a keyword association dataset and collect another Post Image dataset following Kang et al. [32]'s approach with further refinement.

Key Association dataset. The Small World Dataset [18] is the largest free word association resource in English, featuring over 12,000 cue words. It reflects how people remember and recall concepts through word associations [40], collected from 100 participants who provided three associated words for each cue. Participants are fluent speakers from includes participants from over 9 regions (e.g. Australia, Canada, New Zealand, United Kingdom, United States, etc), with American English speakers comprising 50%, followed by British (13%), Canadian (11%), and Australian (5%)

speakers. There are also several words that appeared as cues in multiple forms corresponding to regional spelling variations (e.g., odor and odour), and in such cases, they included only the American English variant [18]. Then association strengths were calculated based on the frequency of each response relative to the total number of responses [9]. We then constructed the keyword association dataset as a directed graph, with forward associations (from responses to cues) and association strengths as weights. It includes words within two association distances from the topic words, creating a dataset with 7,407 words.

Post Image Collection. Our Post Image dataset is built upon [32] where they used Pinterest as main image source for novice designers since it is the most popular website for them to find image examples. Kang et al. [32] built a keyword-image dataset utilizing the keyword association dataset collected before with 4,861 descriptive words and 76,686 images where images mainly focus on the advertising creative images. We incorporated their images into our dataset as a foundational source and further updated the dataset based on feedback from our formative study to include storytelling and object dimensions, which are more valued by users. We found that certain words from the keyword association dataset, such as confine and tribal, are not suitable for promotional posts from our formative study. To address this, we used Sentence-BERT [52] to generate embeddings for all keywords in the dataset and applied Hierarchical-DBSCAN [56] to group similar ones. We then refined these clusters iteratively to ensure the storytelling-related keywords were properly categorized for promotional use. Specifically, two authors iteratively examined each cluster with more than five keywords to manually label the most representative storytelling keywords (e.g., "Journey," "Growth," "Health"). Then we exclude clusters with less than five keywords in it. Finally, we get a new dataset with 4,231 storytelling keywords and 64,396 images. Besides, to identify related objects within our new dataset, we fine-tuned the MMDetection model using SAHI [36] and applied it to each image in the dataset sourced from [32]. As we curated our dataset from a public repository [32] further annotated by two authors, our dataset follows the ethical use of copyright material.

5.2 Image and Caption Recommendation

5.2.1 *Image Recommendation*. Influencer recommends images from the Post Image dataset in three dimensions (i.e., storytelling, color, and object) to provide related and diverse examples, as detailed below. Therefore, we extract the corresponding features from a seed image and retrieve feature-relevant images. Due to space limitations of the block, we display the top four images in each recommended dimension and allow designers to scroll left and right to view more images.

(a) Storytelling. We propose distinct methods for storytelling image recommendation based on two types of images: those available in the dataset and those uploaded by users. For images within the dataset, we have already identified highly related storytelling keywords through the computation of concreteness and imageability scores (ranging from 0 to 1) sourced from [39]. Our approach prioritizes words associated strongly with storytelling concepts. For images uploaded by users, designers mentioned the importance of incorporating user-uploaded images to enrich the exploration of

design materials in promotional post design based on our formative study. To address this, we leverage the state-of-the-art image captioning model in framework OFA [68] to extract key information from the uploaded image. Additionally, we fine-tuned mDeBERTa-v3 using image captions extracted via OFA framework paired with their corresponding storytelling keywords from our dataset. Then, the fine-tuned model classifies OFA-generated captions from uploaded images to identify relevant concepts from our 4,231 storytelling keywords. After getting the highly related concepts, the top four images are randomly selected from the qualified candidates.

(b) Color. We use ColorThief[4] to extract the dominant color values from the image and recommend images with similar color based on color histogram correlation. To ensure a speedy search, we search for images in similar color from related storytelling concepts.

(c) Object. Influencer recommends images with similar objects based on object detection. To facilitate the search speed and make sure that the image has some storytelling relationship with the original image, we only search images in the related concepts. MMDetection models leveraging SAHI [2] is used to extract the main object and retrieve the top probable items from the image. The top four images with similar objects are then recommended. Specifically, We recommend images containing the same objects first. If fewer than four are available, we suggest images with similar objects. Similarity is calculated using cosine similarity of the Euclidean distances between object image embeddings extracted by the MMDetection model. The reason of using the SAHI framework is that there are many small but important objects in the promotional image based on our interview.

5.2.2 *Caption Recommendation*. Addtionally, Influencer recommends captions from three dimensions (product, activity, advertisement) to provide related and diverse examples. Specifically, activity-focused posts emphasize engaging actions or events, product focused posts describe the features and benefits of goods or services, and advertisement-focused posts address broader marketing and branding efforts. Based on the analysis of promotional posts on Instagram [15], there are five themes in total including product, activity, advertisement, text, and others. Based on their statistic of categories and the results of our formative study, we recommend our captions in three contexts (i.e., product, activity, and advertisement). We use GPT-4 to generate caption recommendations based on the user input prompt T_p with the following prompt:

```
Prompt: Generate three promotional captions for each
dimension (product, activity, advertisement) based on the
 given text: <T_p>. please also highlight the keywords
with asterisks and keep rendering icons in each caption.
```

5.3 Context-Aware Exploration

Users already use the seed images I_s and seed text T_s to get the initial image and caption recommendation, Influencer further supports context-aware exploration which is built upon the image or text block dragged by the front-end. As we stated in Sec. 3.2.1, existing off-the-shell recommendation tools cannot fulfill our customization requirements. Based on the findings of our formative study, users find it useful to record their exploration history and update recommendations based on materials like brand images or

---

[4]https://github.com/lokesh/color-thief

product messages, ensuring these align with their chosen themes and design constraints. Two main tasks should be considered for implementation:

(a) Context-Aware Exploration of Images. To enable context-aware exploration of images, users can drag the text prompts (T_c) or image context (I_c) to update the image recommendations created from the seed image I_s. Users can refine the recommendation by searching for similar images with varying colors, objects, or themes. First, we apply the method from Section 5.2.1 to extract keywords for each image in the Post Image dataset across three dimensions (storytelling, object, color). We then collect keywords within each dimension, creating a keyword list that covers all three dimensions.

For text context, we use the OFA image captioning model [68] to extract a description (D_i) from the seed image (I_s) and drag to combine it with the text context (T_c) to create a contextual prompt: $P_t^I = D_i +$ "under the context of" $+ T_c$. This prompt is then classified into the three dimensions using mDeBERTa-v3 [36], returning target keywords (W_s, W_c, W_o) for storytelling, color, and object dimension. These keywords are used to update image recommendations retrieved from the dataset accordingly.

For image context, users can upload images (I_c) like brand logos or personal images and drag to update the image exploration results retrieved from the dataset. We first extract object and color keywords using models like MMDetection with SAHI [2], Color-Thief[5]. To get storytelling keywords, we use image captioning model ofa_imagecaption_coco_large_en [68] and text classification model mDeBERTa-v3 [36] to get related storytelling keywords W_s from the storytelling keyword list. The recommendations are prioritized based on the similarity between the extracted keywords from both the seed image and the context image, ensuring consistency in the recommendations across these dimensions.

(b) Context-Aware Exploration of Captions. Captions can be updated by dragging both text and image context to caption recommendations to align with promotional goals.

For text context (T_c), users can drag it to caption recommendations by changing the target product, activities, or advertisement focus. We leverage GPT-4 [45] to generate new captions based on the seed text (T_s) which is used to generate initial caption recommendation and text context (T_c) across three dimensions: product, activity, and advertisement.

For image context, users can upload personal or brand images (I_c) and drag to the caption recommendation, and the system uses the OFA image-captioning model [68] to extract a description (D_i). This description is combined with the seed text (T_s) to update captions in the same three dimensions utilizing the GPT-4. The prompt is shown below:

```
Prompt: Generate three promotional captions for each
dimension (product, activity, advertisement) based on the
given text: <T_s> under the context of [<T_c> or D_i].
Please highlight the keywords with asterisks and include
rendering icons in each caption.
```

---

[5]https://pypi.org/project/colorthief/

5.4 Image and Caption Fusion

In promotional design, texts and images can be used to customize and regenerate images by fusing them into the target images according to their promotional requirements.

(a) Image Fusion. For text-based imagefusion, users can provide and drag text prompts (T_p) to fuse and modify images. We first run the image captioning model ofa_imagecaption_coco_large_en [68] to extract contextual information D_i from the target image I_t. Then we combine this with the text prompt (T_p) to generate a new image with DALL·E [51].

For image-based image fusion, users can upload a personal interest image (I_i), and we extract descriptions (D_p) from the personal images using the similar method with OFA [68]. Then we combine it with the D_i extracted from the target image to generate a fused image in the same method with DALL·E [51]. The prompt for the two above condition is shown below:

```
Prompt: <D_i> + under the context of: + [<T_p> or <D_p>]
```

(b) Caption Fusion. Text and image contexts can also be used to drag and fuse with the caption to create more relevant and engaging promotional captions under the post requirements. In text-based caption fusion, users can provide a text prompt (T_p) alongside the target caption (C_t) to regenerate a more refined caption with GPT-4 API [45].

For image-based caption fusion, users upload a context image (I_i). Similarly, we use the model ofa_imagecaption_coco_large_en [68] for image captioning to extract the relevant description (D_i) from the image and combine it with the target caption (C_t) to generate a new caption through GPT-4 that matches the image's context. The prompt is shown below:

```
Prompt: Regenerate the following caption <C_t> based on
the context: [<T_p>/<D_i>]. Highlight important keywords
with asterisks and include icons.
```

6 USER EVALUATION

We conducted a within-subjects controlled experiment comparing Influencer with a Baseline. According to our formative study, Google Search stands as the most commonly utilized platform for seeking inspiration, and ChatGPT and Powerpoint are commonly used by users, making it a suitable baseline system for promotional post design. We excluded specialized tools like Canva and Adobe Express from the baseline comparison, as they are generally preferred by experienced designers. Novice users tend to favor more familiar tools like ChatGPT and PowerPoint due to the steep learning curve of specialized platforms. The baseline (Figure 8) featured a Google Search-like interface with essential functions such as image searching (utilizing the same database) and image saving capabilities. Users can employ ChatGPT to customize the image and write the caption to create a promotional post, and use PowerPoint to finalize the promotional post design.

6.1 Participants

We recruited 12 participants (7 women and 5 men, age: $M = 27.3$, $SD = 3.6$) via social media and mailing lists. The majority (10 out of 12) need to engage with graphic design in their daily tasks,
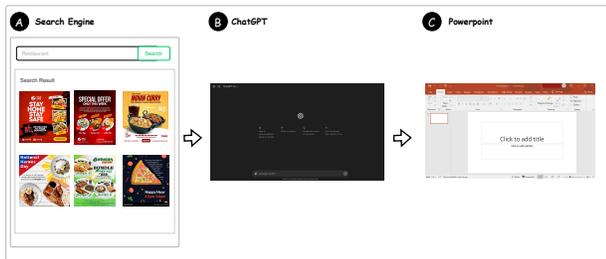
Figure 8: The Baseline system includes two parts: (A) search engine, (B) ChatGPT, and (C) Powerpoint, which resembles the current workflow with industry standard tools.

although they are not proficient designers. Their occupations include: 4 marketers, 4 product managers, 2 graduate students, and 2 small business owners. With a pre-study questionnaire, they were considered novice designers based on their self-reported design experiences on a 5-point Likert Scale, including expertise in promotional post design ($MD = 1$, $SD = 0.42$) as well as familiarity with color theories ($MD = 1.5$, $SD = 0.78$), ChatGPT ($MD = 1.5$, $SD = 0.63$), and Powerpoint ($MD = 1$, $SD = 0.56$), where 1 indicates less expertise/ familiarity. Thus, they are representatives of the typical target users for Influencer (e.g., marketers, product managers, UI designers, and small business owners) who want to easily and quickly produce a few promotional post designs for broadcasting events or products.

6.2 Tasks and Design

We employed a within-subjects design for the study. We designed three tasks to compare Influencer with the Baseline on multiple aspects of creativity and assistance in promotional post design tasks. All tasks require users to design a promotional post from the provided topic which is the most common case for novices to not start from scratch. We set a 10-minute time limit for each task with each system (Baseline and Influencer) to simulate a timing design scenario. This also allowed the study length to be reasonable in a within-subject design.

Task 1 focuses on a product promotion scenario. Given a product topic "Fruit Juice", users need to design relevant images to comply with the topic. They also need to generate the related captions to be compatible with the images and make the final post design harmonic overall.

Task 2 focuses on the scenario of activity promotion. Provided with an activity theme: "Basketball Game", participants need to design the promotional image to achieve a design matching the theme. Participants should also generate reasonable captions to express a particular design need and suitably describe the image. There should be at least two versions of the final product with different colors to accommodate different sentiments or occasions.

Task 3 aims to test the brand-based promotion design scenario. Given a brand image and topic: "Ramen," participants need to design an image compatible with this brand image on the canvas, which includes the color of the image, and the object of the image. In addition, the final design should be aesthetically pleasing.

6.3 Procedure

During the study, participants complete the above three tasks using both of the study systems (i.e., Influencer and Baseline), one after another. The order and combination of study data and tools were counterbalanced. For each condition, participants were first introduced to the study system (i.e., Baseline or Influencer). During the Influencer condition, participants were given a brief tutorial demonstrating its basic functionality, followed by a period of free exploration. Once users were comfortable using the tool, tasks were presented sequentially with the final deliverables for each task. Once participants finished their session, we asked them to provide a binary (yes/no) answer for each task's completion status. After completing those tasks , each participant were asked to complete a questionnaire related to their experience, comparing the Baseline and the Influencer. The questionnaire included the Creative Support Index (CSI) assessment [13] regarding users' experience in exploration, expressiveness, enjoyment, workload, etc. Besides, we conducted a short semi-structured interview with each participant to gain their qualitative feedback. The entire study lasted approximately 90 minutes including three tasks per condition, questionnaire, and semi-structured interview, and participants were remunerated with $20.

7 RESULTS

7.1 Quantitative Results

In the following, we report our results on task completion rate and participants' ratings for Influencer and Baseline.

Completion rate. With Influencer, all participants claimed they had completed all the required promotional posts within the given time for the three tasks. With the Baseline, 91.67% (11/12) of participants completed the post design in Task 1, but only 75.00% (9/12) of the post design in Task 2 and 83.3% (10/12) in Task 3 were fully completed. Participants struggled to complete tasks with the baseline due to the time-consuming process of finding suitable images through Google Search and the iterative nature of creating posts with tools like ChatGPT and PowerPoint. Counting all the completed tasks, the average task completion times across all participants for Influencer were 5.2 minutes ($SD = 0.29$) for Task 1, 7.3 minutes ($SD = 0.39$) for Task 2, and 8.6 ($SD = 0.37$) minutes for Task 3. For Baseline, the average completion times were notably higher: 6.2 minutes ($SD = 0.52$) for Task 1, 8.1 minutes ($SD = 0.45$) for Task 2, and 9.6 minutes ($SD = 0.41$) for Task 3, all remaining under the 10-minute limit. This highlights the efficiency of our system compared to Baseline. Task 3 required the most time due to its emphasis on aligning design elements with a specific brand image, which involved additional customization. This indicates that Influencer has the potential to accelerate design processes when designers need to edit images or quickly generate alternative designs. Here, we present six representative drafts among 72 drafts generated with Influencer and the baseline system in Figure 9.

Design Experience. We utilized the Creativity Support Index (CSI) [13] to measure the degree of creativity support for Influencer and the Baseline in the study. Participants rated five creativity support factors with scores on a Likert scale from 0 (strongly disagree) to 5 (strongly agree). Figure 10 shows the individual CSI score for each factor, i.e., enjoyment, exploration, expressiveness,

P1 with Influencer | P1 with Baseline | P4 with Influencer | P4 with Baseline | P10 with Influencer | P10 with Baseline
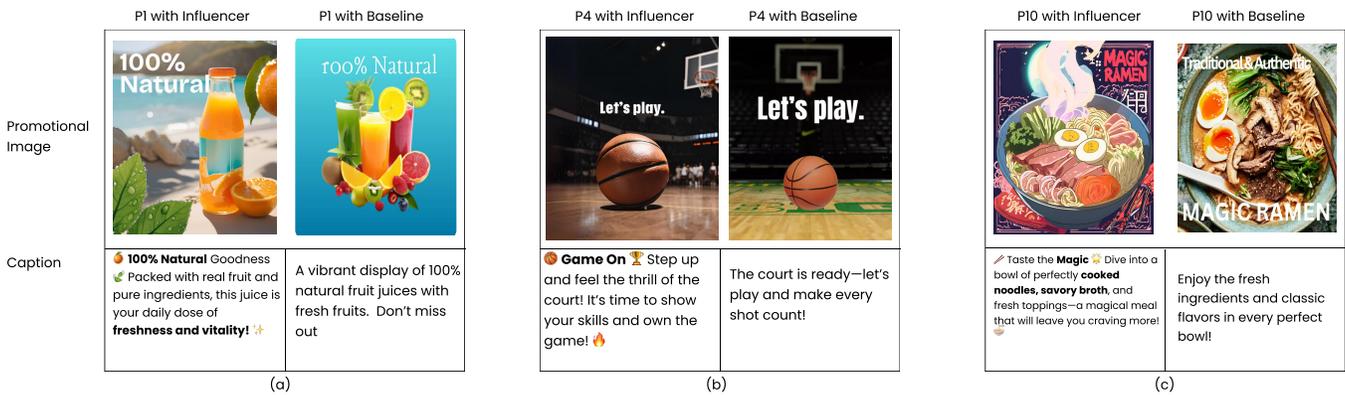
Promotional Image

Caption

(a)　　　　(b)　　　　(c)

Figure 9: Examples of six drafted posts generated by three participants (P1, P4, P10) with Influencer and Baseline system in three tasks: (a) a fruit juice promotional post for a small grocery store, (b) a post for promoting basketball game activity, and (c) a ramen advertisement for a Japanese restaurant.
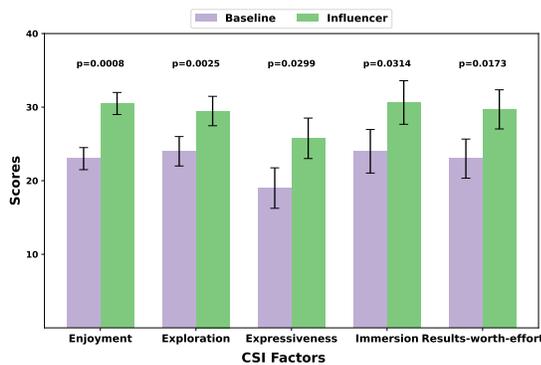

Figure 10: Results of Creative Support Index (CSI) for Influencer and Baseline (the higher the better) on the factors of enjoyment, exploration, expressiveness, immersion, and results worth effort.

immersion, and results-worth-effort. Overall, Influencer achieved a CSI score of 76.8 (SD = 13.52), which is higher than the Baseline with a score of 66.5 (SD = 18.29). The paired t-tests show that the overall CSI score of Influencer is significantly higher than that of Baseline (t = 3.87, p = .0009). It was found that Influencer generated statistically significant improvements in enjoyment (t = 3.59, p = .0008), exploration (t = 3.94, p = .0025), expressiveness (t = 4.14, p = .0299), immersion (t = 3.97, p = .0314), and results worth effort (t = 3.18, p = .0173). The results indicate that participants enjoyed their overall experience with Influencer. Influencer supported graphic design explorations effectively, enhanced users' expressiveness during the creative process and increased their satisfaction with their design outcomes.

Usefulness of Functions. We also evaluated different functionalities within Influencer to see which ones are comparatively more useful with a 5-point Likert scale: mind-map based Visual Representation (M = 4.1, SD = .71), Materials Fusion (M = 3.6, SD = 1.24), Post Generation (M = 3.9, SD = .84), Caption Recommendation (M = 3.4, SD = .65), Storytelling Recommendation for Image (M = 3.8, SD = 1.24), Color Recommendation for Image

(M = 2.4, SD = 1.57), Object Recommendation for Image (M = 3.1, SD = 1.16), Regenerate Image (M = 2.3, SD = .89), Mask Edit Image (M = 2.8, SD = 1.26). As we can see, mind-map based Visual Representation, Post generation, and Dragging Interaction were recognized as the most useful functions in Influencer. However, we also find that some functions have deviations, including materials fusion, storytelling/color/object exploration, and mask edit images. P2 commented: *"Sometimes mask edit cannot generate the exact object as I expected"*. P5 also added: *"I can't get the exact color I want through image recommendations or fusion feature, especially when I want colors that are slightly different, like dark blue and light blue"*. According to our observations, this kind of disagreement in usefulness evaluation reflects different user preferences with different post design habits and the limitation of color extraction and object detection models (see Section 9 for more discussion)

7.2 Qualitative Results

In general, participants appreciated various novel functions of Influencer. We summarize their feedback from the semi-structured interviews based on the following themes.

General preference. Overall, 10 out of 12 participants stated that they would like to use Influencer as their promotional post design tool over the baseline system. Two other participants preferred to combine the two systems to do image searches and get image and caption recommendations but did image customization on another interface.

Image/ caption recommendation Influencer employs multi-dimensional recommendation framework for image/caption exploration while the baseline uses Google. All participants had a clear preference for the multi-dimensional recommendation framework for design materials searching during the interview. Eight noted that the images retrieved were creative and useful compared to their prior experiences using a search engine (P1-2, P4, P10-12). They also mentioned the usability of three-dimensional recommendation for image and caption and stated *"I would appreciate using this function when lacking initial ideas on post design. The recommendations fit my design goal well."* (P6). In addition, participants mentioned that even though there are some recommended images are not suitable

for the topic, it is still useful since Influencer already saves them more time in designing a final version of the post design (P8). Further, *"recommendations make post design process easier and reduce my mental effort required for brainstorming."* (P7). Meanwhile, three participants (P4, P10, P11) recognized the three recommendation directions on caption based on different contexts: *"I can quickly change my post captions to fit different scenes, making it much easier than starting from scratch."* (P9). The feedback indicates that the multi-dimensional image and caption recommendation is useful to facilitate the promotional post design with desired content.

Context-Aware Exploration quality All participants found that the context-aware recommendation is helpful. The utility of context-aware exploration was highlighted when exploring the images and captions based on the given materials: *e.g.,* *"it swiftly guides users to find post images and captions that align with their given brand logo or product message""* (P1-3), *"I can get many captions following my preferences or product relevance through context awareness on caption exploration""* (P5). Moreover, the logical interaction of context-aware exploration was also recognized by 3 participants(P6, P7, P10): *"It is user-friendly to allow me to update the recommendation stuff by dragging the created text prompt to the target recommendation block."* Another participant also appreciated the linkage during the exploration: *"The linkage in the mind-map tells me why this recommendation block is updated, so I can track and know what to focus on."*

Flexible fusion of various images and captions. Influencer employs flexible fusion for customization, while baseline and only use ChatGPT to iteratively regenerate the materials. Powerpoint also provides limited customization features. Eight participants echoed a clear preference for flexible fusion feature. They found it useful when they wished to directly customize an image or caption by writing a prompt or uploading the image they possess (P1-2, P8, P10-12). By flexibly fusing the image or caption, they could *"freely customize the material based on my preference and iteratively brainstorm"*(P12). P1 also mentioned *"It is helpful when I want to use my image to directly rewrite a caption since it helped me integrate the information from my images smoothly."* P4 echoed, *"I feel it convenient since flexible fusion helps me get multiple alternatives quickly for easy comparison later."*

Mind-map layout to track thinking path and record design alternatives. Four participants found it very useful to trace back their thinking path (P6-8, P10). They expressed that it is very inconvenient to switch from different tools to edit images or captions. By enabling tracking design history, they can *"check their design process, freely fuse different materials, and review multiple design alternatives."* P7 also added that *"I feel secure since it records my thoughts and helps me track my thinking process."* Another participant also mentioned different colors for dragging features helped them clearly differentiate the exploration process and customize the process (P11). He recognized the Influencer *"showing how the initial thoughts have been refined and making it easier for him to follow the thought process and rationale behind decisions"*. P12 added, *"I can personally upload an image or write a prompt to customize the image or caption exploration path, which is very flexible and helps me find the related image efficiently."* However, P4 mentioned a potential problem: *"It could become cluttered when handling too many*
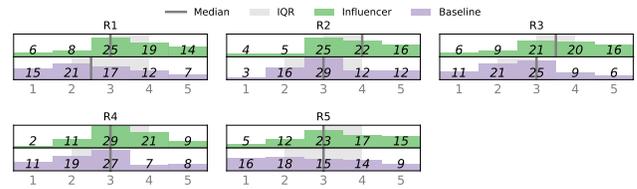

Figure 11: Experts' ratings on the quality of participant-generated promotional posts on a 5-point Likert scale (the higher the better) for five aspects: overall satisfaction (R1), clarity of content (R2), audience engagement (R3), content consistency (R4), and aesthetics (R5).

materials at once, making it harder to navigate and fuse elements effectively."

7.3 Expert Assessment

We recruited three expert designers to help us rate and analyze participant promotional post designs. We also invited them to share their views regarding Influencer afterward. The experts' backgrounds are as follows: 1) an assistant professor specializing in graphic design, with 15 years of experience; 2) a visual designer in the industry, focusing on advertising design, with eight years of experience; 3) an art teacher, teaching art design in high school and usually create design contents for the school's social media, with 10 years of experience. We denote them as EA1-3, respectively.

Post Quality Assessment. While the CSI assessment in our questionnaire indicates participants' self-experience of the created promotional post, these do not reflect how the posts are received by the audience. To assess the quality of the promotional posts created with Influencer, we invited the three experts to rate all the post drafts from 12 participants in two hours. For a fair comparison, we removed the trials from the two conditions if the participants did not complete that task with either of the tools. Each expert rated the promotional posts by participants on a 5-point Likert Item and in random order. The results of their ratings on five different aspects are shown in Figure 11. The Wilcoxon signed-rank tests found that the outcomes of Influencer had significantly better ratings than those of the Baseline for all aspects: overall satisfaction ($t = 632.0, p = .0001$), clarity of content ($t = 984.5, p = .0621$), audience engagement ($t = 754.0, p = .0015$), content consistency ($t = 674.0, p = .0002$), and aesthetics ($t = 803.5, p = .0037$). The results indicate that Influencer could effectively facilitate users to generate better quality promotional posts, more consistent with the given topic, and more engaging, aesthetic, and clear for audiences.

7.4 External User Assessment

To evaluate the quality of the promotional posts created by the participants from a "customer" perspecive, we conducted an online survey with another 20 users (12 women and 8 men; age: *26.4*, *SD* = 2.8) recruited via social media. The goal of this survey was to capture third-person perceptions on the promotional posts. Each user was asked to rate 40 posts (20 from Influencer; 20 from Baseline) randomly selected from the pool of design drafts obtained from our user study, to ensure the survey in a reasonable length. They rated
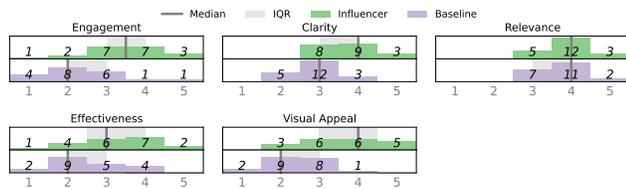
Figure 12: Customer feeling rating on the quality of participant generated promotional posts on a 5-point Likert scale (the higher the better) five aspects.

these posts along five dimensions on a 5-point Likert scale and in randomized order.

The results of their ratings are shown in Figure 12. The Wilcoxon signed-rank tests found that the outcomes of Influencer had significantly better ratings than those of the Baseline for all aspects except relevance: engagement ($t = 1.500, p = .0005$), clarity ($t = 3.000, p = .0005$), relevance ($t = 76.50, p = .2504$), effectiveness ($t = 10.00, p = .0002$), and visual appeal ($t = 0.500, p = .0003$). The results indicate that both Influencer and baseline can create relevant posts, likely due to the sufficient customization options available. However, Influencer leveraging AI-infused feature can produce more engaging, clear, and visually appealing posts, which are more effective in encouraging users to buy products.

8 CASE STUDIES

We conducted an in-depth review of two promotional post design cases by inviting two participants to use the system based on their real-life needs in two one-hour hands-on sessions. These two cases span diverse settings, including public welfare activities and small business marketing. In so doing, we intend to: (1) concretely illustrate how the Influencer can assist everyday users in crafting effective promotional content, and (2) refer back to these cases in the following sections to contextualize the Influencer framework through practical examples.

The case study started with a tutorial of Influencer. The participant was then presented with a partially completed promotional post for a local community event and asked to continue the design process. The think-aloud protocol was employed during the user's exploration and interaction with the tool. We assisted with any technical questions regarding the system's functionality. At the end of the session, additional feedback was collected from the user regarding their experience with Influencer. Figure 13 shows some example outputs from the participants.

8.1 Small Businesses Promotion

To learn how Influencer can be used within a micro-entrepreneur workflow, a one-hour hands-on session was conducted with a milk tea shop owner (TO) who had just built her milk tea shop for two months. As part of her daily job, she needed to promote her new shop on her own, as she was just starting and did not have the budget for advertising.

Exploring Promotional Content. After launching Influencer, the TO browsed the Image Recommendation, where she was presented with various image options related to her milk tea shop.

She noted that the images covered different themes, such as vibrant backgrounds, minimalistic designs, and playful elements (Figure 4.E). She then used the Caption Recommendation (Figure 4.F) to get an overview of the suggested captions generated by the system, which were based on the initial keywords she provided. At a glance, she noticed the product dimension captions effectively highlighted the milk tea's flavors and ingredients, as well as the enjoyment of the drink, i.e., *"The product captions describe the taste and freshness of our teas, while the captions in advertisement and the event focus on the special deals and fun gatherings at the shop."* Besides, she also mentioned the effectiveness of generated captions: *"the captions captured our signature milk flavor and the chewy tapioca pearls perfectly which saves much time thinking about what to write for the post."*

Next, she selected a few images and captions by clicking on them to explore further. Together, with the help of the context-aware exploration (Figure 5), she further wrote a text prompt and uploaded his shop logo to refine the exploration result. She noted that the text prompt seemed more helpful in finding the expected image examples, especially when there is no well-recognized object from the uploaded logo. She mentioned that *"The text prompts helped me out since my logo isn't that famous. It was great for finding the right kind of images that fit my shop's style."*

Customizing Promotional Content. The TO then wished to continue customizing her selected captions and images. She noticed the system's flexible fusion feature (Figure 6) for captions automatically adjusted the text to wrap aesthetically around the image, enhancing the storytelling aspect of her promotional posts. For example, she dragged an image of a vibrant, freshly poured milk tea next to a caption that read, "Burst of Flavor in Every Sip," which dynamically integrated the text into the lively visual scene. She appreciated this flexible fusion feature: *"It is interesting to blend text and image. I can share the stories visually and textually at once."* She added: *"When I dragged 'Fresh milk tea with golden ratio of pearls' next to the cup image, it pops out against the dark background—much better than just plain product descriptions!"*. She then uploaded her shop's logo and dragged it onto the selected image to adjust the color palette to match her shop's logo. She noted that the image featuring a close-up of a milk tea cup with a colorful background aligns with the fun and energetic tone she wanted to convey, commenting that *"This image expresses our shop identity."*

She also noticed that the system allows her to save and compare different versions of the promotional post, making it easier to choose the most satisfying design for her promotion. She commented: *"The captions gave me good options—one highlights our milk tea and chewy pearls, and another creates a cozy tea time mood."* The final design she selected was then ready to be posted on her shop's social media account. She suggested adding automatic suggestions of similar designs would be helpful for the decision-making process.

Posting and Observing the Impact. The TO selected the second promotional post showing the lively poured milk tea. She then shared the promotional post on her shop's social media account. She wanted to see how her followers would respond to her crafted promotional post. She immediately noticed positive engagement, with around 40 likes and multiple comments from loyal customers and new followers in two hours. She observed that the post's bright colors and energetic tone aligned well with her target audience's
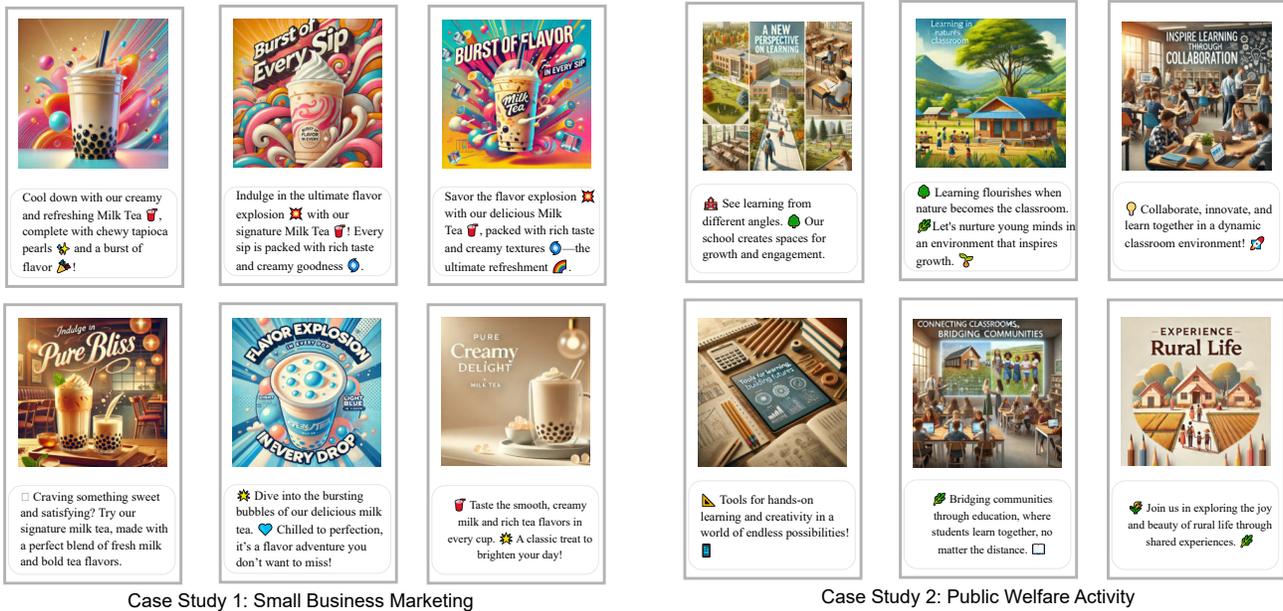
Figure 13: Participant-generated promotional post examples during the two case studies.

expectations, and the caption "Burst of Flavor in Every Sip!" resonated strongly. Some customers commented: *"Looks delicious! I can't wait to try this!"*

The TO was particularly excited by the spike in engagement, which she had not experienced before. *"This post is catching people's attention,"* she remarked. The sudden influx of interactions gave her the confidence to continue using Influencer for future promotional campaigns. She realized that by leveraging the system's intuitive design tools and targeted content suggestions, she could consistently create posts aligned with her brand's identity and customer preferences.

8.2 Public Welfare Activities Promotion

To learn how Influencer can be used within a public welfare activity organizer's workflow, a one-hour hands-on session was conducted with a teacher (TH) with eight years of experience organizing public welfare initiatives to support impoverished students in the countryside. As part of her work, she had to create a promotional post for this event, but she had limited experience in design and social media promotion.

Understanding the Community's Needs. After launching the Influencer, the TH started by exploring the Image (Figure 4.E) and Caption Recommendation features (Figure 4.F). She was presented with various thematic images and text options that resonated with the spirit of community service and education. She noted a collection of images ranging from candid classroom scenes to detailed shots of educational materials, all conveying support for impoverished students. At a glance, she said that these images conveyed reasonable messages, i.e., *"These images capture the reality of the students' daily educational experiences."*

Next, the TH selected specific images to explore further. She noted that Influencer's dynamic user interface allowed her to click

on and enlarge images for a better view. She could further scrutinize each image's context and composition to ensure they matched the message she wanted in her promotional posts. She noticed that the caption *"highlighted the classroom's collaborative spirit and the students' active engagement."* She also commented that the recommended images under the storytelling dimension seem more suitable for welfare activity promotion. *"Storytelling dimension captures the essence of our welfare activities, focusing on the educational environment and student interaction. But images under product and activity seem to focus more on commercial aspects."*

Designer Alternatives. The TH then wished to create multiple promotional post designs on the Influencer platform since she was uncertain about what styles would most appeal to her audience. She generated multiple design alternatives on the canvas by using the flexible fusion feature to blend images and captions and presented these to a group of students, aiming to estimate their reactions and preferences: *"The fusion feature allows me to quickly generate multiple images, like blending students studying in a rural school setting with a vibrant background of local landscapes."* She also appreciated the regenerate feature, commenting that *"It helps me find similar images with related meanings, like variations of the school setting or different angles of the same landscape."* She noticed that the generated caption from the system is useful because it *"captures the essence of rural students immersed in their studies with a lively, beautiful background."*

She appreciated the canvas to showcase these design alternatives clearly and dynamically. With canvas, she got some post designs that were more popular with the students. *"The regenerate feature then helps me find similar images with related meanings, such as variations of the school setting or different angles of the same landscape."*

Continue Mobile sharing and impact evaluation. The TH then posted the final promotional designs on various social media platforms using Influencer's direct publishing features. She observed huge responses from both her colleagues and students. Particularly, they responded positively to images that showed interactive classroom activities and the direct impact of donations, clearly highlighting the students' needs and how could improve their learning environments. The event received about 320 sign-ups out of the school's total of 347 students. Additionally, over 250 students inquired about how to donate and contribute supplies via her promotional post sharing channel.

9 DISCUSSION

This paper focuses on promotional post design, enabling everyday users to express their creativity such as in micro-entrepreneurship. As indicated by the results, using Influencer could streamline the promotional post design process, enabling users to craft engaging and effective promotional content with relative ease. Beyond presenting a novel design case and contextually confirming the benefits of our tool, we discuss the extracted underlying patterns and implications for future research and design below.

9.1 Examples and Context as Key Components in Design

Studies have been conducted to explore the idea of a recommendation framework on how users could get inspiration from single visual elements [29, 32, 33, 48]. However, little has been done to explore and use context-aware exploration for specific design needs. Our work shows that it can improve the process beyond simple recommendations, especially in complex design situations that need customized solutions. Our findings from Section 7 reveal distinct views of three dimensions in the context-aware exploration: *"Storytelling dimension is easy to use and understand."* (P2 and P8). *"Color is useful in most cases especially when I upload a brand image and want to find related images in a similar color schema."* (P10), *"Object dimension can help me find images having similar objects which are helpful when I need to design a product-type promotional post."* (P11). But P6 and P7 expressed concern about color and objects in one of the tasks they did. They mention the image recommendation on color sometimes is slightly inaccurate. The reason may be the inaccurate color extraction caused by the limitation of the object detection model, leading to colors that do not match the users intend to use in their uploaded brand image. To better support this approach in designing, one design expert (E2) suggested that adding a color picker to let users choose preferred colors, and then recommend related color palettes may be helpful to improve image recommendation accuracy. Moreover, while Influencer demonstrates the feasibility of context-aware image exploration, its approach could be extended to different image datasets. Future work could further investigate how dataset selection influences recommendations and consider potential ethical concerns, such as dataset ownership and copyright constraints, when applying this system more broadly.

9.2 Flexible Fusion and Generation with Personalization

Our formative study data suggests that Influencer offers a valuable feature for arranging and combining design elements through an interactive drag-and-drop function. This simplifies the design process and provides a new experience not fully available in current creative tools. They usually offer template-based designs, where users select from pre-defined layouts and styles, limiting the scope for personalized interaction and creativity [64, 67, 72]. The future design could build upon this experiential interaction and make the design process more immersive. For example, in 3D design, when a user moves one 3D model near another, the system can automatically align and merge the shapes into a unified, complex entity. Further, when lighting elements are added to the scene, their effects can intelligently blend to create a cohesive atmosphere that enhances the visual impact of the 3D space. Besides, based on our user studies, we learned that users may have different preferences on the design styles of post images and captions. We can use data analytics and machine learning to analyze user behavior, preferences, and past design choices, providing customized recommendations for images and captions. Moreover, This is also an interesting avenue to consider in the future development of more personalized recommendation models in design.

9.3 Trade-off between Automation and Autonomy

To facilitate the creation of promotional posts by design novices, Influencer employs a nuanced recommendation system that carefully balances guidance with creative freedom. The examples in the findings show that a multi-dimensional recommendation system for images and captions enhances the design process by providing a wide range of visual and text options, allowing users to explore different design ideas. Further, users were able to tailor the design process to the user's specific brand narrative and aesthetic preferences with context-aware exploration. The results indicate that Influencer outperforms the baseline on various design tasks. However, this customization is limited by the performance of the Generative Image AI models. For example, *"The image fusion feature sometimes doesn't work as I expected. Newly generated images didn't change the style of my dragged image."* (P12). Fine-tuning the Generative Image AI model in a post image dataset could improve the visual quality of generated post images. Allowing users to edit and polish the images later would also be a helpful addition to the system, such as adding text to the image, adjusting image saturation, and other basic features like those in Photoshop. In summary, this suggests that an ideal design system should balance automation and autonomy to provide inspiration with recommendation, generation, and customization abilities.

9.4 Limitations and Future Work

This work has several limitations which we plan to address in the future. First, even though our system supports auto-layout for each block, it struggles to maintain flexibility when users create many alternatives, often requiring manual adjustments to resolve overlapping blocks. While users can modify design elements like

images, captions, and color schemes, the flexibility to fully customize layouts and structures remains limited. Future work could focus on developing a more flexible mind-map editor that supports self-defined layouts, allows users to save useful blocks based on their needs, and enables seamless collaboration through real-time feedback and enriched design workflows.

Second, Influencer recommends images based on three dimensions: storytelling, object, and color. However, the current visual algorithm has limitations in accurately identifying highly related images within complex images. Additionally, when users upload their own images and attempt to update the image recommendation block, existing vision models cannot correctly extract the storytelling concept of the uploaded image. To improve the accuracy of recommendation algorithms, we could try to expand the dataset and use a more advanced vision model in the future.

Third, since there are many kinds of information including texts, photos, typography design, and layout that can inspire users, our three-dimensional exploration would benefit by some extension. To adapt our system to more creative scenarios (e.g., banner design, photography, etc), we can include more images in the dataset or recommend images in more dimensions (e.g., typography, mood, style). Also to help professional designers, we can incorporate Influencer with Photoshop to perform more complex tasks or leverage advanced image editing features from Photoshop to fulfill more advanced user needs.

Fourth, the validation is constrained by limited cases, which may not capture the diversity of real-world creative scenarios. Broader case studies across various domains are needed to validate whether the results can be broadly applied. Long-term deployment and testing are also needed to assess how the system adapts to evolving user needs and supports creative workflows over time.

10 CONCLUSION

In this paper, we have introduced Influencer, a tool designed to support novice designers such as freelance creators, marketers, and product managers with promotional post design. By seamlessly integrating captivating images and well-crafted captions, Influencer addresses the challenges faced by design novices in generating attention-grabbing content. Our system offers a mindmap-like layout for ideation and incorporates multidimensional AI-powered exploration and customization of related images and captions to revolutionize the design process. Additionally, Influencer allows flexible fusion of various design elements. Through a comprehensive evaluation, including controlled experiments, expert assessment, external user assessment, and case study, we have demonstrated that Influencer significantly enhances the ideation and design process. It enhances the effectiveness of promotional post design in different task scenarios by fostering engaging interactions, diverse explorations, and trackable thought processes.

ACKNOWLEDGMENTS

REFERENCES

[1] Bamidele Adekunle and Christine Kajumba. 2020. The Nexus between Instagram and digital entrepreneurship. *Journal of African Development* 21, 1 (2020), 14–40.

[2] Fatih Cagatay Akyon, Sinan Onur Altinuc, and Alptekin Temizel. 2022. Slicing Aided Hyper Inference and Fine-Tuning for Small Object Detection. In *2022 IEEE International Conference on Image Processing (ICIP)*. IEEE, Bordeaux, France, 966–970. https://doi.org/10.1109/icip46576.2022.9897990

[3] Anthropic. 2023. Model Card and Evaluations for Claude Models.

[4] Suyun Sandra Bae, Oh-Hyun Kwon, Senthil Chandrasegaran, and Kwan-Liu Ma. 2020. Spinneret: Aiding Creative Ideation through Non-Obvious Concept Associations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376746

[5] Subhojit Banerjee. 2009. Effect of product category on promotional choice: comparative study of discounts and freebies. *Management Research News* 32, 2 (2009), 120–131.

[6] Nathalie Bonnardel and Evelyne Marmèche. 2005. Towards supporting evocation processes in creative design: A cognitive approach. *International journal of human-computer studies* 63, 4-5 (2005), 422–435.

[7] David C Brown. 2008. Guiding computational design creativity research. *Studying Design Creativity, Springer* N/A, N/A (2008).

[8] Noel Carroll. 1994. Visual metaphor. In *Aspects of metaphor*. Springer, N/A, 189–218.

[9] Andrew Cattle and Xiaojuan Ma. 2017. Predicting word association strengths. In *Proceedings of the 2017 conference on empirical methods in natural language processing*. Association for Computational Linguistics, Copenhagen, Denmark, 1283–1288.

[10] Joel Chan, Pao Siangliulue, Denisa Qori McDonald, Ruixue Liu, Reza Moradinezhad, Safa Aman, Erin T Solovey, Krzysztof Z Gajos, and Steven P Dow. 2017. Semantically far inspirations considered harmful? accounting for cognitive states in collaborative ideation. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition*. Association for Computing Machinery, New York, NY, USA, 93–105.

[11] Li Chen, Mengyi Zhao, Yiheng Liu, Mingxu Ding, Yangyang Song, Shizun Wang, Xu Wang, Hao Yang, Jing Liu, Kang Du, and Min Zheng. 2023. PhotoVerse: Tuning-Free Image Customization with Text-to-Image Diffusion Models. arXiv:2309.05793 [cs.CV]

[12] Xi Chen, Lianghua Huang, Yu Liu, Yujun Shen, Deli Zhao, and Hengshuang Zhao. 2023. AnyDoor: Zero-shot Object-level Image Customization. arXiv:2307.09481 [cs.CV]

[13] Erin Cherry and Celine Latulipe. 2014. Quantifying the creativity support of digital tools through the creativity support index. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 4 (2014), 1–25.

[14] DaEun Choi, Sumin Hong, Jeongeon Park, John Joon Young Chung, and Juho Kim. 2024. CreativeConnect: Supporting Reference Recombination for Graphic Design Ideation with Generative AI. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–25.

[15] Kar-Hai Chu, Jon-Patrick Allem, Tess Boley Cruz, and Jennifer B Unger. 2017. Vaping on Instagram: cloud chasing, hand checks and product placement. *Tobacco control* 26, 5 (2017), 575–578.

[16] Cesc Chunseong Park, Byeongchang Kim, and Gunhee Kim. 2017. Attend to You: Personalized Image Captioning With Context Sequence Memory Networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Honolulu, HI, USA, 895–903.

[17] Richard Lee Davis, Thiemo Wambsganss, Wei Jiang, Kevin Gonyop Kim, Tanja Käser, and Pierre Dillenbourg. 2024. Fashioning Creative Expertise with Generative AI: Graphical Interfaces for Design Space Exploration Better Support Ideation Than Text Prompts. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–26.

[18] Simon De Deyne, Danielle J Navarro, Amy Perfors, Marc Brysbaert, and Gert Storms. 2019. The "Small World of Words" English word association norms for over 12,000 cue words. *Behavior research methods* 51 (2019), 987–1006.

[19] Yenan Dong, Shangshang Zhu, and Wenjie Li. 2021. Promoting sustainable creativity: An empirical study on the application of mind mapping tools in graphic design education. *Sustainability* 13, 10 (2021), 5373.

[20] Volodymyr Dziubak and Andrea Bunt. 2018. Prism: Enhancing Graphic Designers' Visual Research with Interactive Search Trails. In *Proceedings of the 2018 designing interactive systems conference*. Association for Computing Machinery, Hong Kong,

China, 1283–1294.

[21] Claudia Eckert and Martin Stacey. 2000. Sources of inspiration: a language of design. *Design studies* 21, 5 (2000), 523–538.

[22] Eric Evans. 2004. *Domain-driven design: tackling complexity in the heart of software.* Addison-Wesley Professional, Boston, MA, USA.

[23] Lampros Gkiouzepas and Margaret K Hogg. 2011. Articulating a new framework for visual metaphors in advertising. *Journal of Advertising* 40, 1 (2011), 103–120.

[24] Milene Gonçalves, Carlos Cardoso, and Petra Badke-Schaub. 2014. What inspires designers? Preferences on inspirational approaches during idea generation. *Design studies* 35, 1 (2014), 29–53.

[25] Harrison G Gough. 1976. Studying creativity by means of word association tests. *Journal of Applied Psychology* 61, 3 (1976), 348.

[26] Nichakan Horphet and Aranya Srijongjai. 2020. An Analysis of Multimodal Text Design in Social Media Marketing: A Case Study of Starbucks Posts on Instagram. *Vacana* 8, 1 (2020), 119–154.

[27] Laura Hummell. 2006. Synectics for creative thinking in technology education. *Technology and Engineering Teacher* 66, 3 (2006), 22.

[28] Roope Jaakonmäki, Oliver Müller, and Jan Vom Brocke. 2017. The impact of content, context, and creator on user engagement in social media marketing. In *Proceedings of the Annual Hawaii International Conference on System Sciences*, Vol. 50. IEEE Computer Society Press, IEEE Computer Society Press, Los Alamitos, CA, 1152–1160.

[29] Ali Jahanian, Shaiyan Keshvari, SVN Vishwanathan, and Jan P Allebach. 2017. Colors–Messengers of Concepts: Visual Design Mining for Learning Color Semantics. *ACM Transactions on Computer-Human Interaction (TOCHI)* 24, 1 (2017), 1–39.

[30] David G Jansson and Steven M Smith. 1991. Design fixation. *Design studies* 12, 1 (1991), 3–11.

[31] Se-Hoon Jeong. 2008. Visual metaphor in advertising: Is the persuasive effect attributable to visual argumentation or metaphorical rhetoric? *Journal of marketing communications* 14, 1 (2008), 59–73.

[32] Youwen Kang, Zhida Sun, Sitong Wang, Zeyu Huang, Ziming Wu, and Xiaojuan Ma. 2021. MetaMap: Supporting visual metaphor ideation through multi-dimensional example-based exploration. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Yokohama, Japan, 1–15.

[33] EunJin Kim and Hyeon-Jeong Suk. 2017. Thoughts and tools for crafting colors: Implications from designers' behavior. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. Association for Computing Machinery, New York; NY, 321–331.

[34] Mikyoung Kim, Doori Song, and Ahnlee Jang. 2021. Consumer response toward native advertising on social media: the roles of source type and content type. *Internet Research* 31, 5 (2021), 1656–1676.

[35] Nupur Kumari, Bingliang Zhang, Richard Zhang, Eli Shechtman, and Jun-Yan Zhu. 2023. Multi-Concept Customization of Text-to-Image Diffusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Vancouver, Canada, 1931–1941.

[36] Moritz Laurer, Wouter van Atteveldt, Andreu Casas, and Kasper Welbers. 2022. Less Annotating, More Classifying: Addressing the Data Scarcity Issue of Supervised Machine Learning with Deep Transfer Learning and BERT-NLI. *Political Analysis* N/A, N/A (2022), 1–33.

[37] Dorothy Leonard and Jeffrey F Rayport. 1997. Spark innovation through empathic design. *Harvard business review* 75 (1997), 102–115.

[38] Patrick Wendell Barbosa Lessa, Beatriz Gondim-Matos, and Antonio Messias Valdevino. 2023. Micro-entrepreneurs in the creative industry: how resilience overcomes the impacts of the pandemic. *Revista Brasileira de Gestão de Negócios* 25 (2023), 353–372.

[39] Ting Liu, Kit Cho, George Aaron Broadwell, Samira Shaikh, Tomek Strzalkowski, John Lien, Sarah M Taylor, Laurie Feldman, Boris Yamrom, Nick Webb, et al. 2014. Automatic Expansion of the MRC Psycholinguistic Database Imageability Ratings.. In *LREC*. European Language Resources Association (ELRA), Reykjavik, Iceland, 2800–2805.

[40] Xiaojuan Ma. 2013. Evocation: analyzing and propagating a semantic link based on free word association. *Language resources and evaluation* 47, 3 (2013), 819–837.

[41] RG Vishnu Menon, Valdimar Sigurdsson, Nils Magne Larsen, Asle Fagerstrøm, Herborg Sørensen, Helena Gunnars Marteinsdottir, and Gordon R Foxall. 2019. How to grow brand post engagement on Facebook and Twitter for airlines? An empirical investigation of design and content factors. *Journal of Air Transport Management* 79 (2019), 101678.

[42] Thomas Merten and Ines Fischer. 1999. Creativity, personality and word association responses: Associative behaviour in forty supposedly creative persons. *Personality and Individual Differences* 27, 5 (1999), 933–942.

[43] Wendell O Metcalf. 1973. *Starting and managing a small business of your own.* Small Business Administration, Washington, D.C., USA.

[44] OpenAI. 2023. GPT-4 Technical Report. https://doi.org/10.48550/arXiv.2303.08774 arXiv:2303.08774 [cs].

[45] OpenAI. 2023. GPT-4 Technical Report. arXiv:2303.08774 [cs.CL]

[46] Cesc Chunseong Park, Byeongchang Kim, and Gunhee Kim. 2018. Towards personalized image captioning via multimodal memory networks. *IEEE transactions on pattern analysis and machine intelligence* 41, 4 (2018), 999–1012.

[47] Matti Perttula and Pekka Sipilä. 2007. The idea exposure paradigm in design idea generation. *Journal of Engineering Design* 18, 1 (2007), 93–102.

[48] Matthew O Peterson. 2019. Aspects of visual metaphor: an operational typology of visual rhetoric for research in advertising. *International Journal of Advertising* 38, 1 (2019), 67–96.

[49] Savvas Petridis and Lydia B Chilton. 2019. Human errors in interpreting visual metaphor. In *Proceedings of the 2019 Conference on Creativity and Cognition*. Association for Computing Machinery, San Diego, CA, USA, 187–197.

[50] Barbara J Phillips and Edward F McQuarrie. 2004. Beyond visual metaphor: A new typology of visual rhetoric in advertising. *Marketing theory* 4, 1-2 (2004), 113–136.

[51] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. 2021. Zero-Shot Text-to-Image Generation. arXiv:2102.12092 [cs.CV]

[52] Nils Reimers and Iryna Gurevych. 2019. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. arXiv:1908.10084 [cs.CL] https://arxiv.org/abs/1908.10084

[53] Jinjin Ren, Liming Lin, and Wei Zheng. 2024. Product promotion copywriting from multimodal data: New benchmark and model. *Neurocomputing* 575 (2024), 127253.

[54] Amanda Dale Richards. 2006. *Sustainable micro-entrepreneurship to ensure positive economic growth in the Western Cape.* Ph. D. Dissertation. Cape Peninsula University of Technology.

[55] Carsten D Schultz. 2017. Proposing to your fans: Which brand post characteristics drive consumer engagement activities on social media brand pages? *Electronic Commerce Research and Applications* 26 (2017), 23–34.

[56] Khaled Sellami, Mohamed Aymen Saied, and Ali Ouni. 2022. A hierarchical DBSCAN method for extracting microservices from monolithic applications. In *Proceedings of the 26th international conference on evaluation and assessment in software engineering*. Association for Computing Machinery, Gothenburg Sweden, 201–210.

[57] Moushumi Sharmin, Brian P Bailey, Cole Coats, and Kevin Hamilton. 2009. Understanding knowledge management practices for early design activity and its implications for reuse. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Glasgow Scotland, UK, 2367–2376.

[58] Xinyu Shi, Ziqi Zhou, Jing Wen Zhang, Ali Neshati, Anjul Kumar Tyagi, Ryan Rossi, Shunan Guo, Fan Du, and Jian Zhao. 2023. De-Stijl: Facilitating Graphics Design with Interactive 2D Color Palette Recommendation. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (<conf-loc>, <city>Hamburg</city>, <country>Germany</country>, </conf-loc>) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 122, 19 pages. https://doi.org/10.1145/3544548.3581070

[59] Pao Siangliulue, Kenneth C Arnold, Krzysztof Z Gajos, and Steven P Dow. 2015. Toward collaborative ideation at scale: Leveraging ideas from others to generate more creative and diverse ideas. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. Association for Computing Machinery, Vancouver, BC, Canada, 937–945.

[60] Pao Siangliulue, Joel Chan, Krzysztof Z Gajos, and Steven P Dow. 2015. Providing timely examples improves the quantity and quality of generated ideas. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*. Association for Computing Machinery, Seoul, Korea, 83–92.

[61] Nikita Srivatsan, Sofia Samaniego, Omar Florez, and Taylor Berg-Kirkpatrick. 2023. Text Conditional Alt-Text Generation for Twitter Images. *arXiv preprint arXiv:2305.14779* N/A, N/A (2023), N/A.

[62] Fabio Staiano. 2022. *Designing and Prototyping Interfaces with Figma: Learn essential UX/UI design principles by creating interactive prototypes for mobile, tablet, and desktop.* Packt Publishing Ltd, Birmingham, England.

[63] Marika Tiggemann. 2022. Digital modification and body image on social media: Disclaimer labels, captions, hashtags, and comments. *Body Image* 41 (2022), 172–180.

[64] Yu ting Qiang, Yanwei Fu, Xiao Yu, Yanwen Guo, Zhi-Hua Zhou, and Leonid Sigal. 2017. Learning to Generate Posters of Scientific Papers by Probabilistic Graphical Models. arXiv:1702.06228 [cs.CV]

[65] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971* N/A, N/A (2023).

[66] Khai N Truong, Gillian R Hayes, and Gregory D Abowd. 2006. Storyboarding: an empirical determination of best practices and effective guidelines. In *Proceedings of the 6th conference on Designing Interactive systems*. Association for Computing Machinery, University Park PA, USA, 12–21.

[67] Anjul Tyagi, Jian Zhao, Pushkar Patel, Swasti Khurana, and Klaus Mueller. 2022. Infographics Wizard: Flexible Infographics Authoring and Design Exploration. *Computer Graphics Forum* 41, 3 (June 2022), 121–132. https://doi.org/10.1111/

cgf.14527

[68] Peng Wang, An Yang, Rui Men, Junyang Lin, Shuai Bai, Zhikang Li, Jianxin Ma, Chang Zhou, Jingren Zhou, and Hongxia Yang. 2022. OFA: Unifying Architectures, Tasks, and Modalities Through a Simple Sequence-to-Sequence Learning Framework. *CoRR* abs/2202.03052 (2022).

[69] Sitong Wang, Savvas Petridis, Taeahn Kwon, Xiaojuan Ma, and Lydia B Chilton. 2023. PopBlends: Strategies for conceptual blending with large language models. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Hamburg, Germany, 1–19.

[70] Merryl J Wilkenfeld and Thomas B Ward. 2001. Similarity and emergence in conceptual combination. *Journal of Memory and Language* 45, 1 (2001), 21–38.

[71] Ziming Wu, Zhida Sun, Taewook Kim, Manuele Reani, Caroline Jay, and Xiaojuan Ma. 2018. Mediating color filter exploration with color theme semantics derived from social curation data. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–24.

[72] Wenyuan Yin, Tao Mei, and Chang Wen Chen. 2013. Automatic generation of social media snippets for mobile browsing. In *Proceedings of the 21st ACM International Conference on Multimedia* (Barcelona, Spain) *(MM '13)*. Association for Computing Machinery, New York, NY, USA, 927–936. https://doi.org/10.1145/2502081.2502116

[73] Joanne Yu, Wilson Cheong Hin Hong, and Roman Egger. 2024. The Art of Post Captions: Readability and User Engagement on Social Media. *Journal of Travel Research* 0, 0 (2024), 00472875241228822.

[74] Nanxuan Zhao, Nam Wook Kim, Laura Mariah Herman, Hanspeter Pfister, Rynson WH Lau, Jose Echevarria, and Zoya Bylinskii. 2020. Iconate: Automatic compound icon generation and ideation. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Honolulu HI USA, 1–13.

[75] Tongyu Zhou, Jeff Huang, and Gromit Yeuk-Yin Chan. 2024. Epigraphics: Message-Driven Infographics Authoring. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Honolulu, HI, USA, 1–18.