



Investigating Semantically-enhanced Exploration of GAN Latent Space via a Digital Mood Board

Qian Wan

City University of Hong Kong
Hong Kong, China
qianwan3-c@my.cityu.edu.hk

Zhicong Lu

City University of Hong Kong
Hong Kong, China
zhiconlu@cityu.edu.hk

ABSTRACT

During past decades, Artificial Intelligence (AI) has been consistently used in Creativity Support Tools (CSTs). Recently, with the development of generative AI models, particularly Generative Adversarial Nets (GAN) in Computer Vision, it became possible that AI directly generates visual ideas. However, there were rarely any work in creativity research that harnessed the design ideas generated by such models directly for design space exploration. To that end we designed a digital mood board as a technology probe to support semantically-enhanced exploration of the StyleGAN latent space. We compared the mood board with traditional gallery display. Preliminary results showed that such design offered a more enjoyable and explicit way of exploring AI generated visual ideas, but gallery display was found to be more straightforward and less demanding.

CCS CONCEPTS

- Human-centered computing → Interactive systems and tools.

KEYWORDS

creativity support, ideation, human-AI interaction, digital mood board

ACM Reference Format:

Qian Wan and Zhicong Lu. 2023. Investigating Semantically-enhanced Exploration of GAN Latent Space via a Digital Mood Board. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3544549.3585740>

1 INTRODUCTION

A Creativity Support Tool (CST) is defined as a tool that “*runs on one or more digital systems, encompasses one or more creativity-focused features, and is employed to positively influence users of varying expertise in one or more distinct phases of the creative process*” by Frich et al. [5]. In the HCI community, the design of creativity support tools plays a fundamental role in the study of creativity. It usually adopts various computational support and interaction

techniques to facilitate creative activities, and recently Artificial Intelligence (AI) and Machine Learning (ML) have come into notice.

For years, Artificial Intelligence has been involved in the design of CSTs in a variety of ways. Previously, AI was used to detect design attributes [10, 15, 16], summarise or retrieve existing designs or materials [2, 10, 15–17], generate semantic stimuli [10, 16], mixing different design styles [14], etc. With the development of Deep Learning, particularly generative models such as Generative Adversarial Network (GAN) [7], the potential of AI for design ideation has been completely opened up by its capability of generating realistic data of high quality and diversity. In character design, for example, AI models like StyleGAN [12, 13] are now able to directly generate diverse and realistic images or visual ideas. Previous works have already used GANs to generate character faces [11] or support the colourisation of a character design sketch [4]. However, it still remains an unanswered question how such model can be used during the *ideation* phase in creativity support, especially for novice users who might lack proper knowledge of the design space. We are particularly motivated to know how generations of AI models like GAN can be integrated into CSTs, and how we can support novice users in exploring and making sense of potential AI generations.

In this paper, we designed a web-based digital mood board as a technology probe that integrates AI generations into the ideation phase. It provides a customisable and interactive system of semantic labels, powered by a pre-trained anime image classifier, to help designers organise and make sense of AI generations. We also tentatively conducted an evaluation of the mood board involving 10 novice users, in comparison with traditional image galleries. Our study manifest that, compared to image galleries, such mood board was thought to be more enjoyable and explicit, while galleries was generally less demanding.

2 RELATED WORK

Previous creativity support tools primarily facilitated ideation by retrieving, analysing, suggesting, and blending existing related materials. For example, Koch et al. [16] proposed a digital mood board that attaches semantic labels to images via AI labeling algorithms, in order to help designers articulate, communicate, and reflect on vague, visual ideas. The tool then translates ideas from designers into search terms, and support serendipitous discovery via online semantic search. In their later work [15], AI text suggestions were also incorporated into the workflow, though novel images were still curated by an image-based search engine. Jeon et al. [10], in their work intended for expert fashion designers, introduced a data-driven creativity support tool. Based on fashion image data labeled with text attributes, they developed AI models capable of fashion attribute detection, style clustering, style forecasting, and style

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHI EA '23, April 23–28, 2023, Hamburg, Germany

© 2023 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-9422-2/23/04.

<https://doi.org/10.1145/3544549.3585740>

merging, to help designers analyse previous fashion trends and set future design direction.

Another line of research in CST also focused on group collaboration in order to boost creativity during ideation. These works often provide cultural and conceptual diversity, usually in a group brainstorming setting. For example, Wang et al. [21] proposed a model that intelligently suggests and selects pictorial stimuli based on group conversation during a brainstorming session. In a follow-up study [22], they also experimented the system with participants of different cultural backgrounds, and proved the significance of cultural diversity within a group in boosting creativity. To facilitate group brainstorming, Shih et al. [20], in an earlier work, proposed a collaborative mind-mapping tool, GroupMind, which was proved in their evaluation to outperform traditional whiteboards.

In order to address the problem of ‘design fixation’ [9] during the ideation phase, computer supported generations were also used in CST research to provide novel designs or ideas. For example, in industrial design, previous CSTs tend to integrate Genetic Algorithms (GE) into their workflow. Kim et al. [14] in their system intended for garden design, allowed users to manipulate the genetic operators between their design and other collaborators’, and substituted the fitness function in GE with user selection. Recently, the development of generative AI models have aroused research interest in the field of HCI, and also in the study of CSTs. Mozafari et al. [17] tentatively proposed an image-based query approach for User Interface Design. Their system could first find a nearest latent code (the high-dimensional input vector) of an input image in the StyleGAN [12, 13] latent space via gradient descent, and then supports synthesis from a source image to a desired style by mixing style vectors fed into the StyleGAN generator.

Our work tentatively integrates StyleGAN generations into a creativity support tool for design space exploration, adopting a human-in-the-loop manner. Unlike image queries, we use a digital mood board with semantic labels as it helps translate the often vague visual ideas of AI for novice users. The mood board also support both serendipitous and convergent explorations of the StyleGAN space based on user feedback, either visual or semantic. Our study compared the design probe with image galleries to inform future design of creativity support tools.

3 THE MOOD BOARD

In this section we elaborate on our design features and briefly explain how they were implemented.

3.1 Sticky Notes

We implemented a semantic tagging system via virtual sticky notes. A sticky note contains a semantic tag that corresponds to multiple images, while one image can also have multiple corresponding sticky notes. A sticky note is used to help designers make sense of AI generations, brainstorm text ideas, and explore more potential generations. Whenever an image is obtained from the backend, and spawned on the front-end interface, its corresponding tags predicted by the classifier will be made into sticky notes and placed on the mood board canvas.

For each sticky note on the mood board canvas, designers can click on it to highlight the tag, and the system will highlight all

images currently assigned to it. Designers can then select or deselect images they believe to be relevant or irrelevant. Designers can also alter the tag of each sticky note by double clicking on the text. Figure 1b provides an illustration of how designers could interact with a highlighted sticky note. If a user likes to define a customised tag, he or she can drag a sticky note from the tile of unused new tags on the tool panel, and input their own text. The user then needs to highlight the new sticky note, and select images he or she believes to have such property.

3.2 Exploring New Generations

The mood board supports exploration of StyleGAN generations via three methods: ‘random sampling’, ‘style mix’, and ‘more similar generations’.

3.2.1 Random Sampling. In the mood board, designers can request more random images by clicking on the ‘request’ icon. Upon clicking, the front-end will randomly sample input vectors from a Gaussian distribution, and send them to the back end. Upon receiving these vectors, the backend server will feed them through the StyleGAN generator, and obtain generated images from its output. These generations will finally be sent to the front-end interface and displayed in the ‘new image’ section.

3.2.2 Style Mix. Inspired by Genetic Algorithm (GA) like methods of design space exploration [14], the mood board also allows users to mix styles between any two groups, either images or semantic tags. To do that a user can draw a selection triangle to select the first group, click on the pop-up ‘mix’ icon, draw another selection triangle to select the second, and then click the ‘mix’ icon again to mix the styles between the two groups (see Figure 1c as an illustration). Each group could include both images and sticky notes, and resulting images will be displayed in the ‘style mix’ section. When mixing styles of two given groups, the system first retrieves all corresponding images and input vectors of two groups. We refer to the two groups of vectors as G_1 and G_2 . It then randomly samples a subset of each group, G_1' and G_2' , to serve as ‘parents’ to breed a new generation $C = G_1' \otimes G_2'$.

3.2.3 More Similar Generations. The mood board also supports requesting generations similar to existing images or semantic labels. Whenever a user finds an image or a sticky note of interest, the mood board allows him or her to explore more similar generations by adding some variations. For an image, if the user clicks the ‘more similar’ icon, the front-end will add some random Gaussian noises to the input vector, and display resulting images from the backend around the image of interest at a smaller scale. For a sticky note, the mood board retrieves all images corresponding to the tag, randomly samples two subsets, and performs a genetic operator same as ‘style mix’ between those subsets, and displays resulting images close to the sticky note.

3.2.4 Implementation. The front-end user interface of the mood board was programmed using React¹ as the web framework. The back-end server of the mood board was programmed in Python using Django framework², and deployed on a computer equipped

¹<https://reactjs.org/>

²<https://www.djangoproject.com/>

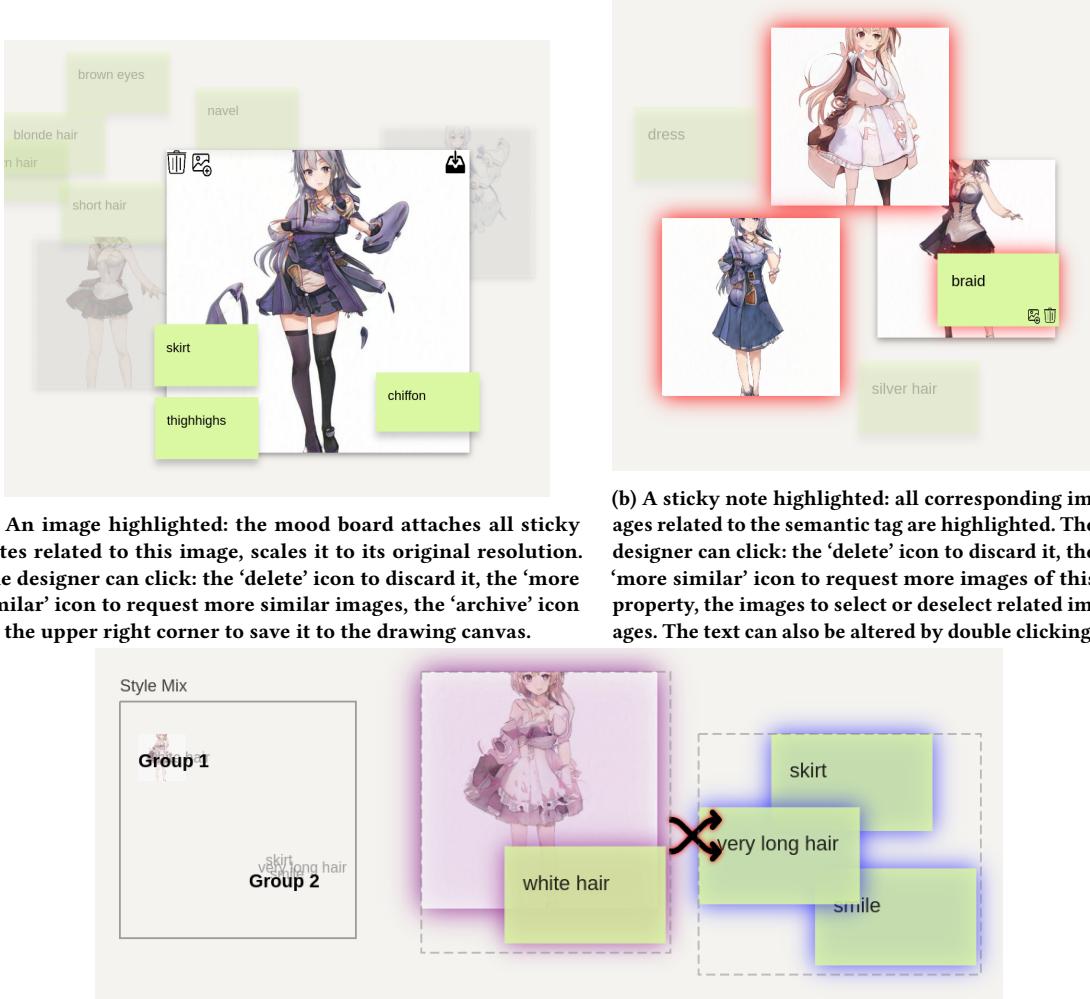


Figure 1: An overview of the mood board features

with an Nvidia GeForce RTX 3080. The backend server of the mood board holds a StyleGAN [12, 13] trained at a resolution of 512x512, and a classifier pre-trained on anime images [19] that extracts semantic labels. It returns a StyleGAN generated image upon an incoming HTTP request attached with an input vector, along with semantic tags predicted by the classifier over the generated image.

4 EVALUATION

We ran a study to evaluate the GAN-driven mood board in terms of creativity support, practical usability, and final results delivered. We intuitively compared our design with a gallery displaying StyleGAN generations, as inspired by previous work [17], and proliferating design gallery platforms, such as Pinterest [18] and Behance [1].

4.1 Participants

In order to conduct our user study with novice users, we recruited 10 participants with no expertise of character design (5 male, 5 female, aged between 18 and 25). For convenience we refer to them as P1 to 10. All participants provided informed consent and agreed to the recording of the session and anonymised publication of the results. For each participant, we offered a coupon equivalent to 50 HKD after the study.

4.2 Study Procedure

At the beginning of our study, all participants were required to complete a short survey regarding their demographic data, and a self-identification of their knowledge of AI, especially generative AI models such as GAN. We then asked our participants to imagine a scenario where they would like to design an anime character

for themselves, either as a social media profile picture, or a virtual avatar in the meta-verse, and that they first would like to browse through some AI generations to draw inspirations.

We conducted a within-subjects study. Each participants were required to use two systems: the GAN-driven mood board, and a simple image gallery, and the order was randomly assigned. The system of image gallery also includes a dummy canvas, and users can click on images they like to bring them to the dummy canvas. When using both systems, participants were required to complete same tasks: 1) archive intriguing and inspirational images to their canvas 2) articulate design ideas they obtain while using the system. Each participant was also required to think aloud and explain their thought process while using these systems. The whole process was screen-recorded and audio-taped for further analysis, and for each system used by each participant, we kept a screenshot of the final results on the web page. After using each of the system, each participant was required to complete a survey adapted from the Creativity Support Index (CSI) [3] and NASA-TLX [8] in 7-point Likert scales (1: Strongly Disagree; 7: Strongly Agree).

4.3 Results

4.3.1 Usability. The NASA-TLX score of two systems are summarised as in Figure 2. After some time of training, participants (e.g., P2, and P4) usually said the GAN-driven mood board was overall ‘simple’ and ‘explicit’. However, participants reported relatively higher level of mental demand (*Median* = 3.5, *IQR* = 2.75) and efforts (*Median* = 3, *IQR* = 1), as compared to frustration (*Median* = 2, *IQR* = 2), and physical demand (*Median* = 2.5, *IQR* = 3.25). The reason was mainly because the mood board required users to read through semantic labels and sort out the mood board when it became cluttered (e.g., “*I felt it was mentally demanding because I had to focus on each sticky notes to make sense of it*” - P10, “*I was too focused on organising things on the mood board* - P7”). These results manifested that participants felt the GAN-driven mood board was overall easy to use, though as expected, an image gallery was perceived to be easier since it was much less complicated.

4.3.2 Creativity Support. We first calculate the mean of each dimension of CSI, and then compute the median and IQR for each dimension and the overall score. As summarised in Table 1, the mood board outperformed the traditional image gallery in *Enjoyment*, *Expressiveness*, and *Immersion*. While using the mood board, most users (e.g., P4-5, P10) said it was ‘more interactive’ and therefore ‘more enjoyable’, and P4 particularly mentioned that “*I prefer using the mood board for exploration since I could kind of tell the AI what to generate*.” The reason that the score of *Exploration* was not significantly higher was partly because image galleries were generally perceived as easier to use, and some participants such as P7 especially preferred a less-demanding way of exploring new generations.

4.3.3 Reflection on Semantic Support. The semantic support is one major difference between our mood board and traditional image galleries. During our study, we observed that all participants spent at least some time browsing through AI suggested tags, highlighting them to see corresponding images, and highlighting images to read through all tags attached. In particular, P2 was found to spend most

	Image Gallery		The Mood Board	
	Median	IQR	MEDIAN	IQR
Enjoyment	5.5	1	6	1
Exploration	6	0.5	6	0.875
Expressiveness	5.5	0.5	6.25	1
Immersion	5.5	1.375	6	0.75
Overall	5.75	1.75	6	0.75

Table 1: Creativity Support Index of the mood board and the image gallery

time exploring semantic features. He used to highlight specific sticky notes to browse and organise images on the mood board, mix styles between sticky notes, and click on ‘more similar’ icon of individual sticky notes. He explained that he preferred using tags to sort out the mood board and explore new generations, because it was more “*explicit*” compared to the image gallery. P7 added that, “*If there would be a stage where I were to draw something (on the canvas), I would pretty much use these (archived sticky notes) as a reference*”. P4 even complimented that, “*The most intriguing thing (of the mood board) is the ability of the AI to illustrate properties from (generated) images. I don’t know what these properties are but the AI already knew (and explained to me), and that impressed me the most*.”

However, P7 noted that it would be more mentally demanding than a simple gallery without semantic support to organise and make sense of the sticky notes on a mood board, especially when it sometimes became cluttered. During our study, there was also a case where P8 came across a semantic tag ‘monochrome’ that intrigued and inspired her, but when she tried telling the AI to generate images of such property, results from the system constantly frustrated her. Furthermore, P1 was the only user that showed obvious inclination towards a system without any semantic tags. She explained that, “*I was more interested in images that are more abstract and sort of inexplicable by semantic tags, images that might not look like a character at all. The sticky notes of the mood board for me seemed distracting*”

4.3.4 Gallery display vs. Interactive mood board. Towards the end of each study, we asked our participants to compare the two systems in terms of exploring new generations. Many of our users (P2, P5-6, P8, P10) showed clear preference for a mood board, describing it as a more enjoyable and customisable form. P5, P6 and P8 all stressed that they would like to interact with the AI via a mood board instead of browsing through static generated images in a gallery, because it would be more enjoyable. P6 particularly noted that it would be quite tedious if she were to explore an image gallery without any interaction for a long time. Besides, P5 and P8 also emphasised that an interactive mood board allowed for more room of customisation. P8 mentioned she could actually ‘fine-tune’ an image using the mood board, while on a gallery you could only click to save it. P5 also said that, “*In an image gallery, if you find an image you like or you dislike, there’s nothing really you can do, you can only scroll down for more. While in another prototype (the mood board), you can filter out very similar images. You can also use various generation (exploration) methods, and that was very effective and helpful*.” Some

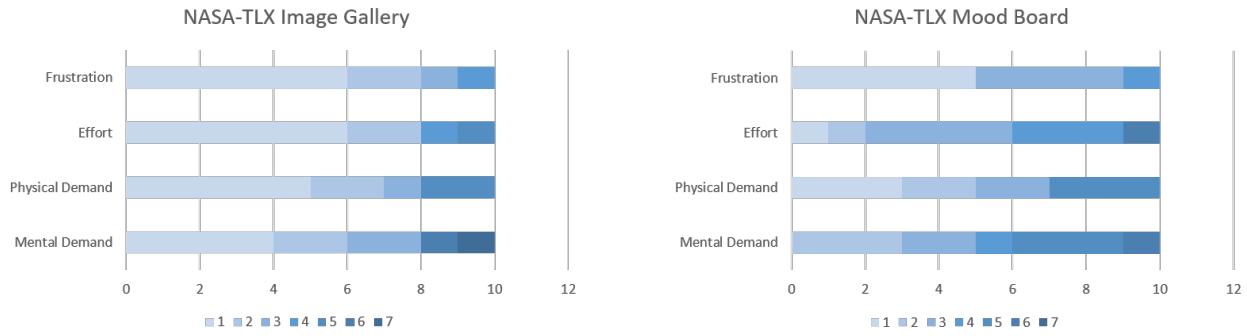


Figure 2: NASA-TLX of the mood board and the image gallery

other users (P4 and P7) also noted the advantage of a image gallery as being more straightforward, though not necessarily negative towards a mood board. P7 said it to be easier to use, and ‘clear at a glance’. P4 also felt that such a gallery brought about more diversity during his study.

5 CONCLUSION

In this paper, we compared a GAN-driven semantic mood board with a traditional image gallery. Our evaluation shows that the mood board is more enjoyable and explicit, but compromises some level of usability. This finding is consistent with previous studies of mood boards, which have claimed that they are fun to create [6] and help translate vague, visual ideas [16]. However, our results reveal that a semantic digital mood board might be more mentally demanding and lack the ‘clarity at one glance’ when exploring StyleGAN generated visual ideas. Furthermore, many of our participants noted the Genetic Algorithm approach (similar to [14]) as more enjoyable and customizable for design space exploration. These implications for CST design and human-AI interaction should inspire future creativity research involving generative AI models.

REFERENCES

- [1] Behance. 2022. *Search the creative world at work*. Retrieved September 16, 2022 from <https://www.behance.net/>
- [2] Carole Bouchard, Jean-francois Omhove, Celine Mougenot, Ameziane Aoussat, and Stephen J Westerman. 2008. TRENDS: a content-based information retrieval system for designers. In *Design Computing and Cognition'08*. Springer, 593–611.
- [3] 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.
- [4] Yuanzheng Ci, Xinzhu Ma, Zhihui Wang, Haojie Li, and Zhongxuan Luo. 2018. User-guided deep anime line art colorization with conditional adversarial networks. In *Proceedings of the 26th ACM international conference on Multimedia*. 1536–1544.
- [5] Jonas Frich, Lindsay MacDonald Vermeulen, Christian Remy, Michael Mose Biskjaer, and Peter Dalsgaard. 2019. Mapping the landscape of creativity support tools in HCL. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [6] Steve Garner and Deana McDonagh-Philp. 2001. Problem interpretation and resolution via visual stimuli: the use of ‘mood boards’ in design education. *Journal of Art & Design Education* 20, 1 (2001), 57–64.
- [7] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2020. Generative adversarial networks. *Commun. ACM* 63, 11 (2020), 139–144.
- [8] Sandra G Hart. 1986. NASA task load index (TLX). (1986).
- [9] David G Jansson and Steven M Smith. 1991. Design fixation. *Design studies* 12, 1 (1991), 3–11.
- [10] Youngseung Jeon, Seungwan Jin, Patrick C Shih, and Kyungsik Han. 2021. FashionQ: an ai-driven creativity support tool for facilitating ideation in fashion design. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–18.
- [11] Yanghua Jin, Jiakai Zhang, Minjun Li, Yingtao Tian, Huachun Zhu, and Zhihao Fang. 2017. Towards the automatic anime characters creation with generative adversarial networks. *arXiv preprint arXiv:1708.05509* (2017).
- [12] Tero Karras, Samuli Laine, and Timo Aila. 2019. A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 4401–4410.
- [13] Tero Karras, Samuli Laine, Miika Aittala, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. 2020. Analyzing and improving the image quality of stylegan. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 8110–8119.
- [14] Kevin Gonyop Kim, Richard Lee Davis, Alessia Eletta Coppi, Alberto Cattaneo, and Pierre Dillenbourg. 2022. Mixplorer: Scaffolding Design Space Exploration through Genetic Recombination of Multiple Peoples’ Designs to Support Novices’ Creativity. In *CHI Conference on Human Factors in Computing Systems*. 1–13.
- [15] Janin Koch, Nicolas Taffin, Michel Beaudouin-Lafon, Markku Laine, Andrés Lucero, and Wendy E Mackay. 2020. Imagesense: An intelligent collaborative ideation tool to support diverse human-computer partnerships. *Proceedings of the ACM on human-computer interaction* 4, CSCW1 (2020), 1–27.
- [16] Janin Koch, Nicolas Taffin, Andrés Lucero, and Wendy E Mackay. 2020. Semantic-Collage: enriching digital mood board design with semantic labels. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 407–418.
- [17] Mohammad Amin Mozaffari, Xinyuan Zhang, Jinghui Cheng, and Jin LC Guo. 2022. GANSpiration: Balancing Targeted and Serendipitous Inspiration in User Interface Design with Style-Based Generative Adversarial Network. In *CHI Conference on Human Factors in Computing Systems*. 1–15.
- [18] Pinterest. 2022. *Discover recipes, home ideas, style inspiration and other ideas to try*. Retrieved September 16, 2022 from <https://www.pinterest.com/>
- [19] Masaki Saito and Yusuke Matsui. 2015. Illustration2vec: a semantic vector representation of illustrations. In *SIGGRAPH Asia 2015 Technical Briefs*. 1–4.
- [20] Patrick C Shih, David H Nguyen, Sen H Hirano, David F Redmiles, and Gillian R Hayes. 2009. GroupMind: supporting idea generation through a collaborative mind-mapping tool. In *Proceedings of the ACM 2009 international conference on Supporting group work*. 139–148.
- [21] Hao-Chuan Wang, Dan Cosley, and Susan R Fussell. 2010. Idea expander: supporting group brainstorming with conversationally triggered visual thinking stimuli. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. 103–106.
- [22] Hao-Chuan Wang, Susan R Fussell, and Dan Cosley. 2011. From diversity to creativity: Stimulating group brainstorming with cultural differences and conversationally-retrieved pictures. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. 265–274.