Exploring the Intersection of AI Art and Film: A Case Study of *Giant*

Junrong Song
Computational Media and Arts
The Hong Kong University of Science and Technology
Guangzhou, China
jsong799@connect.hkust-gz.edu.cn

David Yip*
Computational Media and Arts
The Hong Kong University of Science and Technology
Guangzhou, China
daveyip@ust.hk

2023 IEEE International Conference on Multimedia and Expo Workshops (ICMEW) | 979-8-3503-1315-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICMEW59549.2023.00066

***Abstract*—Artificial intelligence (AI) has recently been used as a tool for various visual storytelling, but text-to-image models are a stochastic machine learning process that requires human intervention to assist creation better. As we all know, pre-visualization is an important stage in film production, involving subjective choices by the creators. This paper investigates how AI can assist filmmakers during the pre-production stage by generating mood boards from text. We propose a novel pre-production pipeline and guidelines that leverage text-to-image models to create visual previews of film projects. We also conduct a case study to validate and evaluate our approach's effectiveness. Our case study suggests that following the guidelines we have developed can assist filmmakers in generating mood boards that effectively convey the desired atmosphere of their projects and potentially contribute to enhancing the creative process. Our paper aims to contribute to the field of AI art and the film industry.**

Keywords—AI, Pre-visualization, Film, Mood Board, Artificial Creativity, Human-machine collaboration

I. INTRODUCTION

Artificial intelligence (AI) has recently gained significant attention for its ability to generate images from text, with the introduction of Generative Adversarial Networks (GANs). This technology has sparked a wave of interest in using AI to create artworks, from text-generated images to concept art projects [1]. As the capabilities of AI continue to expand, many are now exploring its potential use in the field of film production.

AI-generated images allow for creating a diverse range of images with ease, containing different camera angles, visual elements, and styles. AI technology has also been utilized for various post-production tasks such as editing, dubbing, greenscreen, and visual effects [2]. People also attempted to use AI for AI by using ChatGPT [3] to write prompts for generating AI images. Nevertheless, the randomness of text-to-image generation poses limitations for cinematic storytellers due to its inability to generate sequential images in a consistent, coherent, and purposeful fashion.

This paper aims to investigate how AI can assist filmmakers in the pre-production stage. And will focus on whether AI can act as a co-creator in creating mood boards and enhance the next creative process in storyboarding.

II. BACKGROUND AND MOTIVATION

With the advent of deep learning, researchers have developed several AI-based approaches and tools for generating high-quality images, such as DALL-E [4] and DeepDream [5] which have been successful. Then text-to-image generation using AI gained significant attention on social media platforms in 2022, thanks to the release of open-source models like Disco Diffusion and Stable Diffusion [6].

This study drew inspiration from AI-generated artworks in 2022, such as Giorgio Anselm's series of images created using Midjourney, which depict "Dune" as directed by Sergei Parajanov [7]. And Digiguru's fictional art piece "The Avengers (1980), directed by Wes Anderson" [8]. Anand Gandhi's simple film storyboard on Instagram was also created using AI [9]. Or the short fiction film *Sunspring* [10], which was entirely written by an artificial intelligence bot using neural networks. Despite numerous efforts utilizing AI for creating film-related content, few have tried using AI images to guide actual filming.

However, the randomness of text-to-image generation can limit the storyteller's ability to quickly obtain enough usable sequential images.And for more intricate storyboards, additional adjustments using tools like Photoshop and ControlNet [11] can be necessary, but time-consuming and challenging to handle.

To explore the use of AI-generated images to guide actual filming, our project, *Giant*, will serve as a cross-disciplinary experiment with film and AI technology. Specifically, we will focus on the use of AI-generated mood board. Mood boards are a critical tool for establishing a clear direction for the creative process to ensure that all crew members work towards a shared vision [12]. In the traditional approach of film production, a well-funded team will hire professionals to create their mood board and storyboard that align with the director's vision. However, for low-budget films, mostly, filmmakers may turn to the internet and other media sources for inspiration and use techniques like collages to achieve the desired effects.

Based on previous filming experience, we tried to form a replicated novel pre-production pipeline and guideline as a visualization method for efficiently creating mood boards. This paper will detail our perspectives, methods, and selection process used in the case.

979-8-3503-1315-4/23/$31.00 ©2023 IEEE
DOI 10.1109/ICMEW59549.2023.00066

Authorized licensed use limited to: ExxonMobil. Downloaded on October 27,2025 at 13:19:01 UTC from IEEE Xplore. Restrictions apply.

III. Pipeline Review

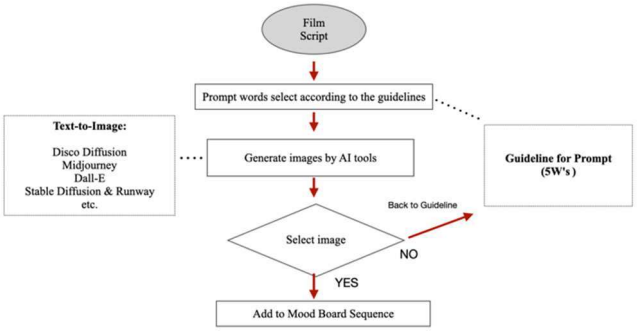


Fig. 1. Pre-production pipeline and guideline for creating text-to-image visual Mood Board as previews of a film project.

As shown in Figure 1, the application was implemented as a pipeline that could be used to quickly visualize different scenes and provide a mood version for filming reference. The pipeline's development was initially based on our prior work experience and subsequently refined through this practical case study. Our objective is to generate usable prompts that can reflect the script and aid the creative and production process. The **Prompt Guidelines** encompass the following critical elements:

1) Shot Composition: This may include describing a long, close-up, medium, high, or low-angle shot, or specifying a particular focal length, such as 35mm or 80mm.

2) Narrative: This involves describing the scene's background, and characters, and providing brief descriptions of the action or events.

3) Emotion: Here, the emotional state of the characters is described.

4) Prompt Assistance: Utilize tools such as ChatGPT to generate complex prompts.

5) Other: Other considerations may include the frame size or artistic style of the image. Like in our cases, a default style has been selected to be used.

An example of a prompt:

Medium shot, a young girl with delicate features and long hair, crying in a wild forest. Tim Burton, Realistic style, Bright Light, 16:9 ...

To begin with, the AI generation platform will be utilized to create 4-8 reference images per scene. The director will then review these images and select the one best suit their specific needs as shooting references. If none of them are satisfactory, the director will refer to the Prompt Assistance part and adjust the narrative accordingly. Possibly with the assistance of an AI chatbot to continue generating suitable prompts for images.

Generating a suitable description may require multiple attempts. In our case study, we utilized ChatGPT to assist us in getting some complex prompts. When the scene is difficult to describe, we will attempt to refine the prompts by providing ChatGPT with the 5W's: "who," "where," "what," "when," and "how/any specific requirements." However, using ChatGPT to iterate prompts for more accurate AI outputs in storytelling is a worthy topic for research, it is not the main focus of our current study. We plan to explore this area in more detail in the future.

IV. Demonstration of Case Study

A script *Giant* written by Junrong is used for our experimental case study. The script is chosen due to its poetic imagery and concise length. It is about a college girl who encounters a giant in the wild on a subway. This encounter motivate her to embark on a journey to find the giant, but despite all her efforts to search in the wilderness, it is in vain. All she finds in the suburbs are the destruction of the forest and the sounds of construction. In the end, on one lonely night, the girl seems to hear the diminishing voice of the giant as it fades into the distance. The script delves into theme of the loss of mother nature due to industrialization, symbolized by the disappearance of the giant, who represents the essence of nature in the film.

According to this script, we simplified and split the text descriptions according to the guideline, input them as prompt into AI generation tools, and get mood boards that can be used as reference images for each shooting scene.

A. Specific Steps

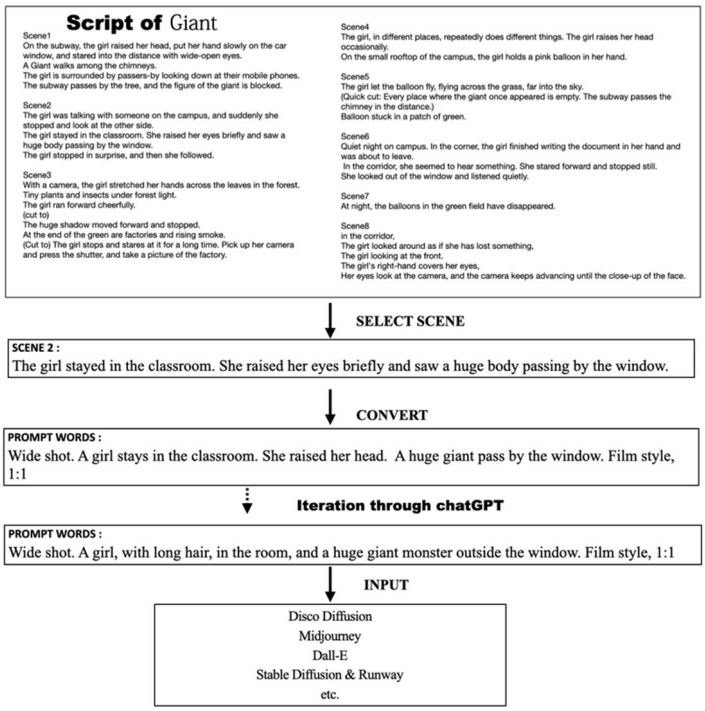


Fig. 2. Illustrates the initial workflow diagram depicting the process from script to prompt, input, and ultimately to the AI-assisted creative platform.

Authorized licensed use limited to: ExxonMobil. Downloaded on October 27,2025 at 13:19:01 UTC from IEEE Xplore. Restrictions apply.

The specific steps are as follows: First, we inputted scene descriptions from script into AI generation tools as prompts. Then, through iterative adjustments to the prompts, we generated reference images as mood boards for desired effects. Finally, the director selected the images to visualize the emotional aspects of their expectations, while also considering the constraints of the actual shooting locations. We initially explored two different styles for the script Giant, using Stable Diffusion and the Runway [13]. However, we found that Runway's default model was more realistic, time-saving, and met our needs. Therefore, during our working process, we mostly utilized Runway to generate our mood boards.

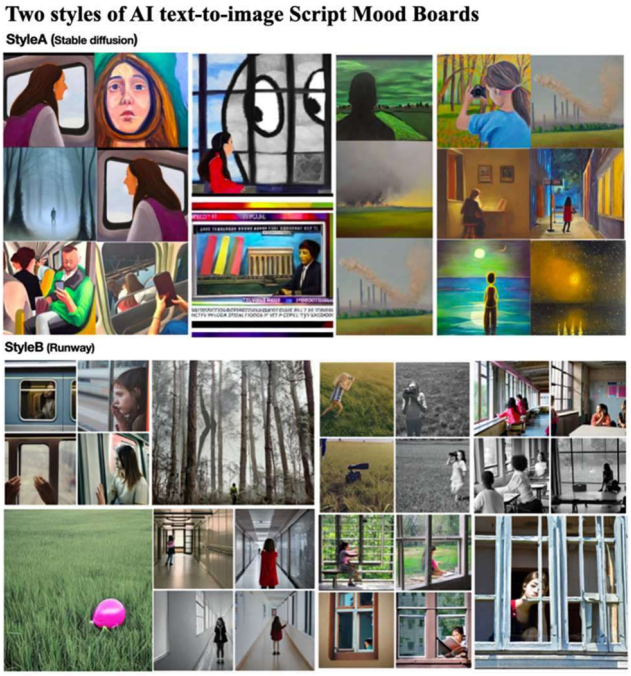


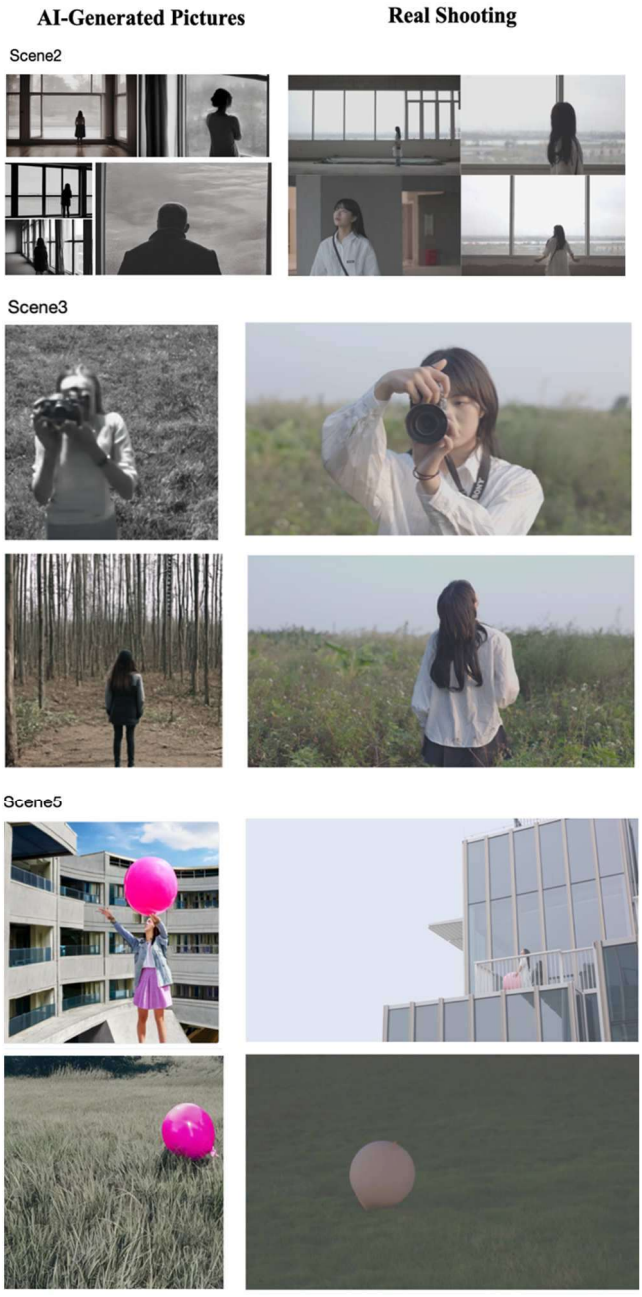
Fig. 3. Display image examples of two different styles generated according to the script *Giant*. Style A was created by using Stable Diffusion, and Style B was created by using Runway.

We created four to eight renderings per prompt to have enough selection to identify the desired outcome. These pictures were selected by the director mainly based on several principles: composition, camera angle, style, and color. If the first four generations did not meet our satisfaction, we would modify the wording to generate the next two or four images.

In addition, we conducted a comparative experiment between the AI-generated method and the traditional method of using Google to search for reference images. The effectiveness of the two methods in creating mood boards was compared, and our pipeline was found to be more effective than inputting individual keyword searches for reference on Google. The AI-generated images were more uniform and referential, which reasonably visualized the emotions that the director expected. We were able to accomplish all the work in three hours, which was half the time taken compared to the traditional one. The final shooting results of our case study confirmed that AI-generated images could serve as an effective guide for establishing the emotional tone of scenes and could truly assist the director in achieving their vision.

B. *Narrative structure and Film language*

Typically, the director would provide specific premises for shot composition, shot size, angle, art style and so on, for AI to generate something that is usable. In this case study, our needs required simple prompts that focus on adjusting the lens and angle, such as long shot, close shot, medium shot, or specific focal lengths like 35mm or 80mm. A narrative description of the scene and a simple characterization of the character's emotions, such as crying or sadness, could also be included.



Authorized licensed use limited to: ExxonMobil. Downloaded on October 27,2025 at 13:19:01 UTC from IEEE Xplore. Restrictions apply.

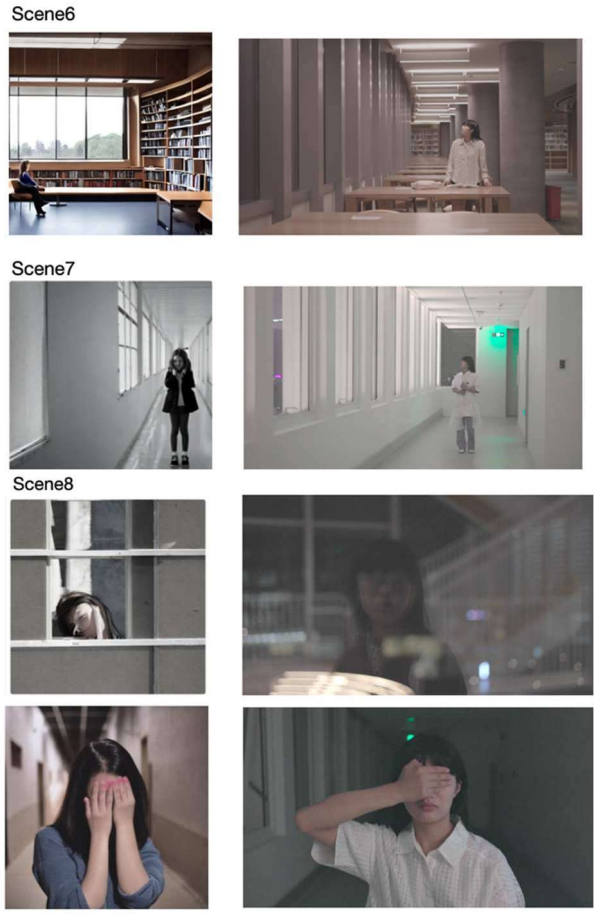


Fig. 4. AI-generated Mood Board for script *Giant* which are arranged in chronological order. Comparison between the pre-production mood board and final shooting results.

Figure 4 presents a partial chronological arrangement of the final filming content, accompanied by the corresponding AI mood board used as a reference. In this particular set, the composition of the structures and colors was found to be inspiring. The visual elements of the selected images can be considered as contributing to the psychological implication of the environment, such as the window pattern of the building and the position of the girl, which conveyed a sense of entrapment in Scenes 2 and 7. Besides, the vibrant colors of nature in Scene 3 represented the girl's vitality, while the lonely pink balloon in Scene 5 symbolized her desire to escape. The contrast between cool and warm tones suggested an unsafe environment in Scene 7. And the depth of the school corridor inspired the selection of a similar location on campus for filming at night. Additionally, the intentional camera angle as a good reference that faced the audience in Scene 8 broke the fourth wall, helping us create a connection between the audience and the girl's emotions.

Here are some reasons mentioned inspired the director, who then selected these images to serve as a source of creative inspiration during the filming process.

C. *Multiple built*

Relatively speaking, guidelines may provide directors with more generated image references for atmosphere and camera angles. However, for some extremely specific details, such as the appearance of the giant in our case, Runway seems could not provide images that meet our requirements. In this case, we have to adopt other methods to solve the problem.

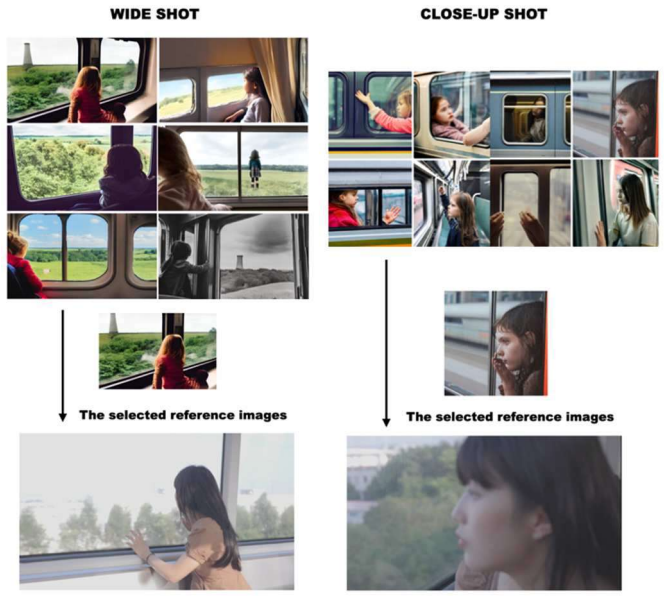


Fig. 5. Illustration of AI-Generated Mood Board Production Pipeline for *Giant*. A comparison of pre-production references to final shooting results for Scene 1.

As an example, we presented a typical scene where the girl in the subway sees the giant in the distance walking among the chimneys. Despite detailed descriptions, generating images of the girl seeing a giant proved challenging. Therefore, we broke down the description and generated different elements separately (Figure 5 and Figure 6). Initially, we used an overview description to establish a style reference as a wide shot of the scene. Then, we inputted a character description to search for a close-up shot of the girl's placement in the frame. During filming, we used the established style as a reference to create an isolated, surreal atmosphere around the girl in the subway, in contrast to the natural world outside (Figure 5). The image seems to meet the director's aesthetic requirements quite accurately and have been considered to express the plot through shot composition, while also meeting the practical filming conditions.

For the exterior window part, we initially attempted to create the mood board solely using AI-generated images. However, as we found that the AI-generated representation did not fully match our expectations, we tried combining traditional methods with AI-generated references to create this anew. First, we used AI-generated images as references to select an appropriate real-world filming location. Additionally, we utilized an image from the web as a reference for the giant's form, which we adjusted

Authorized licensed use limited to: ExxonMobil. Downloaded on October 27,2025 at 13:19:01 UTC from IEEE Xplore. Restrictions apply.

and incorporated into the real location using Photoshop. We then composited the scene based on these (Figure 6).

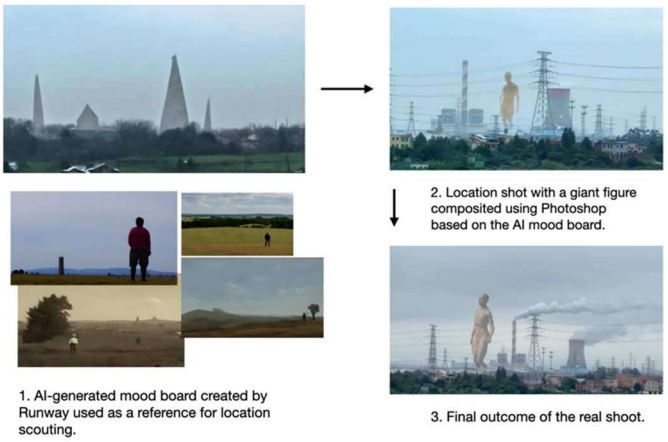


Fig. 6. Comparison of the synthesis process and final shooting outcome for the exterior window section of *S*cene 1.

The materials we obtained helped us to clarify our concept of the giant. In the end, we defined the giant as a brown humanoid with a tall chimney nearby. We also referred to the contrast between the positioning of the giant and the chimneys, which placed the giant and chimney prominently in the center. The contrast between the smoke drifting to the right and the giant walking to the left underscored the theme of opposing forces between nature and industrial development. During filming, we utilized this as our reference to capture the desired atmosphere and subtext, while using the web image of the giant as a reference for our 3D modeling.

D. Prompt assistance by ChatGPT

In our previous tests, we found that the wording of prompts had a significant impact on the generated output, often requiring multiple iterations to arrive at an effective description.

For instance, we struggled to find the right wording for a prompt describing the girl observing a giant passing by the window in an empty classroom. Our initial prompt, "Wide shot. A girl stands in the empty classroom before a big window. A giant passes by outside the window. Realistic, 1:1," resulted in some unsatisfactory mood board. We experimented with various strategies, including adjusting the shot angle, providing more detailed scene descriptions, and substituting synonyms, but were all unable to achieve the desired outcome.

Recently, the use of ChatGPT for creative purposes had become a hot topic [14], so would an AI be better at understanding AI-generated words? With these questions, we tried giving ChatGPT a description of the scene and asking it to generate some key descriptive words. It wrote, "A girl stands in a big empty room. She turns to look out the big window and sees a giant walking past." Surprisingly the new prompt words could generate images that better matched our initial idea. The composition and tone set by the images effectively conveyed the girl's loneliness, and the minimalist composition of the empty environment creates a sense of beauty through the pattern of the windows. We even used them as references for real shooting in storyboarding. This suggests that AI-generated references could at times provide a more accurate description of the script text.

V. IMPLEMENTATION

The method we used proved to be a reliable tool for quickly producing AI-generated mood boards to guide the filming process. Our method differs from previous traditional methods. Through practical experimentation, we also implemented an efficient pipeline and guidelines, incorporating reference images and utilizing prompts to provide useful inspiration for shooting. Although it still shows some limitations, we found that sometimes generating a large number of pictures is not productive and can be a waste of time. It is, however, essential and needs experience to consciously adjust the generation strategy to achieve the best results.

The version generated by AI may not exactly match one's envisioning, but through experimentation, it is possible for filmmakers to make it resemble their desired outcome more closely. As demonstrated, we gradually generate content and utilize tools like ChatGPT to aid in rewriting script prompts. AI-generated mood boards can be a useful reference tool, providing filmmakers with inspiration and new perspectives. They could increase productivity while reducing costs. Also, the use of AI images can assist in visualizing how space and character can be used cinematically in the storyboarding process. Storyboarding is a technique-driven and location-specific process. Different locations have their own unique characteristics and potential for creating moods. During our case study, some AI-generated images showed the potential to transition directly from the mood board to the storyboard.

VI. DISCUSSION AND CONCLUSION

In this study, we explored the potential benefits of utilizing a co-creation approach with AI in the creative process of filmmaking. However, it is important to keep in mind that AI-generated images may not possess the same level of human intuition, emotion, or depth of meaning that is essential to art creation.

Therefore, filmmakers should subjectively develop, choose, and build upon the selected images that best fit their needs. We should view technology as a supplement to human creativity, rather than a replacement for it. Questions regarding the potential homogenization of visual content and ethical considerations also need further exploration to fully understand the potential and limitations of AI in the film creation process. For instance, the use of AI-generated content in filmmaking may lead to a reduction in diversity and originality, which can limit artistic expression and creativity. And Ethical concerns such as intellectual property rights and the perpetuation of harmful stereotypes or biases must be addressed. Hence, it is essential to examine these issues and develop appropriate guidelines and regulations to ensure that the use of AI in film creation is beneficial and ethical for all stakeholders involved.

Besides, AI-generated content may not be as precise as hand-drawn storyboards, the random processing of AI can be seen as a form of divergent thinking or brainstorming in a creative process [15]. This emergent method fills possibilities, generates new and unique ideas for filmmakers, and inspires the creation of distinct frame shots and styles. However, as storytelling requires precision and structure [16], precision in film language is crucial to ensure that the intended message is

Authorized licensed use limited to: ExxonMobil. Downloaded on October 27,2025 at 13:19:01 UTC from IEEE Xplore. Restrictions apply.

effectively conveyed to the audience without the risk of confusion or misinterpretation. Therefore, many researchers are now working on how to improve control over AI-generated images to address these limitations [17], [18], emphasizing the importance of precision in film language to ensure the intended message is effectively conveyed to the audience without the risk of confusion or misinterpretation.

In conclusion, our study has presented the potential of an AI-generated mood board as one example of this approach. Further research and experimentation are needed to fully harness the potential of AI in the film creation process. Additionally, our pipeline and guidelines for using AI in the creative process require more practical studies to be refined and improved. It is important to recognize that while AI can enhance and support the artistic expression of filmmakers, it should still be guided by the subjective vision and creativity of the artist. We hope that in the future, by embracing the potential of AI as a supplement to human artistic vision, we can unlock new avenues of creative expression and push the boundaries of what is possible in the world of cinema.

Link of the film *Giant*:
https://vimeo.com/manage/videos/791246094

REFERENCES

[1] She, James & Cetinic, Eva. (2022). Understanding and Creating Art with AI: Review and Outlook. ACM Transactions on Multimedia Computing, Communications and Applications. 18. 1-22. 10.1145/3475799.

[2] N. Kadner, "Generative AI and the Future of Filmmaking," [Online]. Available: https://virtualproducer.io/generative-ai-and-the-future-of-filmmaking/. [Accessed: Apr. 29, 2023].

[3] "ChatGPT: Optimizing Language Models for Dialogue," OpenAI, 2022. [Online]. Available: https://openai.com/blog/chatgpt/. [Accessed: Apr. 29, 2023].

[4] "DALL·E: Creating Images from Text," OpenAI, 2021. [Online]. Available: https://openai.com/blog/dall-e/. [Accessed: Apr. 29, 2023].

[5] Deep Dream Generator. (n.d.). Retrieved May 30, 2023, from https://deepdreamgenerator.com/

[6] Dehouche, Nassim & Dehouche, Kullathida. (2023). What is in a Text-to-Image Prompt: The Potential of Stable Diffusion in Visual Arts Education .

[7] "Midjourney Imagined Dune Directed By Sergei Parajanov," 80.lv, 2022. [Online]. Available: https://80.lv/articles/midjourney-imagined-dune-directed-by-sergei-parajanov/. [Accessed: Apr. 29, 2023].

[8] M. Oosthuizen, "See The Amazing 80s Avengers By Wes Anderson Created By AI," Fortress of Solitude, 2022. [Online]. Available: https://www.fortressofsolitude.co.za/see-the-amazing-80s-avengers-by-wes-anderson-created-by-ai/. [Accessed: Apr. 29, 2023].

[9] R. Vijayakar, "Anand Gandhi pushes AI to limits to create a stunning storyboard," News India Times. [Online]. Available: https://www.newsindiatimes.com/anand-gandhi-pushes-ai-to-limits-to-create-a-stunning-storyboard/. [Accessed: Apr. 29, 2023].

[10] Sunspring. (2016, June 9). In Wikipedia. Retrieved from https://en.wikipedia.org/wiki/Sunspring

[11] Andrew. (2023, February 28). ControlNet: Control human pose in Stable Diffusion. [Online]. Available: https://stable-diffusion-art.com/controlnet/. [Accessed: Apr. 29, 2023].

[12] Elert Munk, J., Sønderskov, J., & Laursen, L. N. (2020). Visual boards: Mood board, style board or concept board? In L. Buck, E. Bohemia, & H. Grierson (Eds.), *DS 104: Proceedings of the 22nd International Conference on Engineering and Product Design Education (E&PDE 2020), VIA Design, VIA University in Herning, Denmark. 10th -11th September 2020: The Value of Design & Engineering Education in a Knowledge Age* [163727] Design Society. https://doi.org/10.35199/EPDE.2020.47

[13] "Runway | Make the Impossible," Runway. [Online]. Available: https://runwayml.com/. [Accessed: Apr. 29, 2023].

[14] A. Bildea, "Design Your AI Art Generator Prompt Using ChatGPT," [Online]. Available: https://pub.towardsai.net/design-your-ai-art-generator-prompt-using-chatgpt-7a3dfddf6f76. [Accessed: Apr. 29, 2023].

[15] Cheng, Mingyong. (2022). The Creativity of Artificial Intelligence in Art. Proceedings. 81. 110. 10.3390/proceedings2022081110.

[16] D. K. Yip, "The Invisible Art of Storytelling and Media Production," in Advances in Creativity, Innovation, Entrepreneurship and Communication of Design, 2020, pp. 262–266.

[17] Pan, X., Qin, P., Li, Y., Xue, H., & Chen, W. (2022). *Synthesizing Coherent Story with Auto-Regressive Latent Diffusion Models*.

[18] Feng, Z., Zhang, Z., Yu, X., Fang, Y., Li, L., Chen, X., Lu, Y., Liu, J., Yin, W., Feng, S., Sun, Y., Tian, H., Wu, H., & Wang, H. (2022). ERNIE-ViLG 2.0: Improving Text-to-Image Diffusion Model with Knowledge-Enhanced Mixture-of-Denoising-Experts. *ArXiv, abs/2210.15257*.

Authorized licensed use limited to: ExxonMobil. Downloaded on October 27,2025 at 13:19:01 UTC from IEEE Xplore. Restrictions apply.