ISSN:1738-6764, eISSN:2093-7504

A Study on the Use of AI Image Generation Tools for Picture Book Illustration Development

Zheng-Yi Jin [1] and Ji-Hyun Lee [2,*]

[1] Graduate School of Kunsan National University, Korea, PhD Student; xiaojin1996220@gmail.com
[2] Kunsan National University, Korea, Professor; illust@kunsan.ac.kr
* Correspondence

https://doi.org/10.5392/IJoC.2025.21.3.060

Manuscript Received 25 April 2025; Received 24 July 2025; Accepted 18 August 2025

Abstract: *The Fourth Industrial Revolution, driven by digital innovations, has brought significant changes to core technologies, industrial systems, and job structures. Against this backdrop, this study explores the application of AI image generation models, specifically those based on diffusion architectures, in the development of picture book illustrations. Focusing on two widely adopted tools—DALL·E 3 and Midjourney—which represent state-of-the-art implementations of diffusion-based image synthesis, the study adopts a case study approach to evaluate their stylistic characteristics, usability, and associated challenges. The research compares the visual styles generated by both tools, outlines a five-stage workflow for AI-assisted picture book creation, and analyzes limitations in AI creativity, data bias, and ethical concerns. Findings indicate that DALL·E 3 produces whimsical and intuitive illustrations suitable for narrative clarity, while Midjourney excels in artistic expression and visual depth. By highlighting the capabilities and constraints of these AI models, this study offers practical insights into how such technologies can support creative illustration processes. It also emphasizes the need for further research using objective data and statistical methods to assess user satisfaction and broader applicability.*

Keywords: Picture Book Illustration; AI Image Generation Tools; Prompts; DALL·E 3; Midjourney

1. Introduction

1.1 Background and Objectives

The Fourth Industrial Revolution has driven significant transformations in core technologies, industrial systems, and job structures, primarily through innovations in digital technologies. Artificial Intelligence (AI), alongside the Internet of Things (IoT), big data, and mobile technologies, has become a central force in reshaping society. According to Kim et al. (2021), AI and smart factories are increasingly being integrated manufacturing sector [1]. The progress of deep learning technologies has marked the beginning of a new era for AI, allowing systems to adopt self-learning mechanisms modeled after the neural networks of the human brain.

The launch of ChatGPT in November 2022 garnered widespread attention for providing an interactive experience that mimics a conversation with a real person, representing a significant advancement over previous technologies. Since then, generative AI has expanded its reach into various modalities, including text, images, coding, audio, video, and even 3D modeling. AI-based image generation, in particular, has demonstrated the ability to produce high-quality visuals that rival human-created illustrations and photographs, having a profound impact on both social and economic landscapes [2]. One of the most notable developments in image generation technology is the use of Generative Adversarial Networks (GANs) in AI. By inputting simple text prompts, users can generate visually striking and highly original images, highlighting the creative capabilities of this technology [3]. These advancements not only open up new possibilities for various industries but also raise important ethical and social considerations that must be carefully addressed.

Against the backdrop of technological advancements, research into the potential applications of AI image generation tools has been steadily expanding. Park (2023) examined the feasibility of using AI image generation

tools, particularly focusing on Midjourney, for fashion design development [4]. Lim and Yoon (2024) investigated the expressive potential of generative AI tools for visual identity design, with a specific emphasis on Midjourney's capabilities [5]. Park and Kim (2024) conducted a comparative analysis of two leading image-generation AI tools, DALL·E 3 and Midjourney, to create front, side, and rear views of 3D characters, proposing an optimized AI generation process based on their findings [6]. Yang and Kim (2024) explored the use of Midjourney in redesigning packaging for Suzhou-style rice wine, with the aim of enhancing its traditional regional cultural identity [7]. While existing studies have largely focused on fashion and brand design, there is still limited research on the application of AI image generation tools in the field of picture book illustration.

Picture book illustrations play a vital role in complementing the narrative by visually interpreting the story, and in some cases, even conveying the entire story without text. They stimulate young readers' imaginations, convey the mood and emotions of the story, and emphasize key narrative elements [8]. Picture book illustration is not merely about drawing; it is an art form that serves as a central component of the book. By employing various styles and techniques, it offers visual enjoyment for readers and deepens their engagement with the story. The use of AI image generation tools offers numerous advantages in the creation of picture book illustrations. These advantages include increased efficiency, enhanced creativity, accessibility, cost-effectiveness, the ability to implement diverse styles, the facilitation of collaboration, and the creation of customized content. AI tools assist illustrators by automating tasks such as sketch generation, color palette suggestions, style transformations, background creation, adding interactive elements, and streamlining repetitive tasks, thereby allowing illustrators to focus on the more creative aspects of their work [9].

Among the various generative AI tools available today, DALL·E 3 and Midjourney were selected for this study because they represent two of the most prominent and stylistically distinct image generation platforms. DALL·E 3 is notable for its natural language prompt processing, integration with conversational AI, and capacity for narrative clarity and emotional warmth. In contrast, Midjourney is widely adopted for its unique artistic aesthetic, rich textures, and abstract visual expressions. These contrasting characteristics make them ideal for comparative analysis in the context of picture book illustration, where both clarity of storytelling and visual creativity are crucial.

This study aims to analyze how AI image generation tools, specifically DALL·E 3 and Midjourney, can be applied in the development process of picture book illustrations through case studies. Unlike prior studies that focus primarily on output quality or design aesthetics, this research takes a process-oriented approach that examines not only visual characteristics but also the workflow, ethical concerns, and tool usability within the creative pipeline. By doing so, it seeks to fill the gap in existing literature and offer foundational insights for the practical integration of AI tools into picture book production from both a technical and narrative design perspective.

1.2 Scope and Methodology

Previous studies identify DALL·E 3 and Midjourney as two of the most prominent AI image generation tools, widely utilized as creative resources across various fields. This research focuses on these two tools, narrowing its scope to case studies where they are applied in the development process of picture book illustrations, specifically within a sequence of four consecutive spreads, with the aim of conducting an in-depth analysis of their practical applications. Specifically, this study will employ case studies utilizing DALL·E 3 and Midjourney to examine how AI tools contribute to the planning, creation, and revision stages of picture book illustration development. Additionally, it aims to identify challenges associated with the use of these AI tools and propose solutions to address them, providing a comprehensive evaluation of both the potential and limitations of AI image generation tools in this context.

2. Literature Review

2.1 Picture Book Illustration

2.1.1 Overview of Picture Book Illustration

Picture books are primarily composed of illustrations and simple text, designed to visually convey stories while focusing on stimulating the imagination and creativity of children. By combining images and text, they present clear and straightforward narratives that young readers can easily understand. Moreover, picture books often integrate educational elements, subtly imparting moral lessons or introducing important concepts in an engaging manner [10]. From a publishing perspective, picture books are a literary genre defined by the

complementary relationship between text and illustrations, working together to form a unified narrative. The origins of picture books can be traced back to *Orbis Sensualium Pictus (The Visible World in Pictures)*, created by J.A. Comenius, the founder of sensory-based education [11]. This early educational work used illustrations to depict the names of objects, aiming to enhance children's learning outcomes and is widely considered a precursor to modern picture books.

Children, being in the early stages of cognitive development, find it more effective to understand objects or phenomena when visual elements are paired with textual descriptions, rather than relying on text alone. This is one reason picture books remain a widely used tool for enhancing comprehension. Educators such as Martin Luther and Friedrich Froebel utilized and created picture books as part of children's education in early modern Europe. Although children remain the primary readership of picture books, recent trends indicate a notable growth in titles designed for adult readers. Picture books can convey narratives solely through illustrations or by combining images with text. However, even without text, a picture book can still fulfill its purpose, whereas it cannot exist without illustrations. Thus, illustrations are an indispensable and essential component of picture books [12].

2.1.2 Expression Styles in Picture Book Illustration

Like other art forms, picture book illustration encompasses a broad range of expressive styles. These styles are thoughtfully selected to align with the text and intended message, creating a unified atmosphere where all elements work together harmoniously. Illustrations that complement the narrative not only deepen its meaning but also help young readers retain the content for a longer period. The artistic style of picture book illustrations reflects the illustrator's subjective interpretation, profoundly influencing both content and mood of the story. The expressive styles of picture book illustrations can be categorized in various ways. However, they are generally classified into four main artistic styles: realism, expressionism, surrealism and abstraction(Table 1).

Realism focuses on portraying subjects as they appear in real life, without distortion, and emphasizes an accurate and detailed representation of characters, objects, and events. This approach prioritizes faithful reproduction of reality, making the content easily understandable and fostering familiarity and emotional connection. Realism is frequently used in educational picture books, such as textbooks, and is particularly effective in conveying intimate or emotionally resonant themes [13]. Expressionism, which emerged as a response to realism, emphasizes subjective emotions and moods over representational accuracy. This style incorporates deliberate distortions, exaggerations, and various techniques involving lines, spaces, colors, forms, and textures to create impactful visual effects. Through the illustrator's unique personal expression, expressionism effectively highlights and conveys the themes, content, and emotional tone of the narrative [14]. Common techniques in expressionism include elongation, simplification, exaggeration to intensify emotions, and magnification of particular shapes. In picture book illustration, realism and expressionism are the most commonly applied styles.

Abstract illustrations can be classified into several styles, including geometric abstraction, organic abstraction, and informel. Geometric abstraction uses basic shapes such as lines, circles, and triangles, often focusing on mathematical relationships. In contrast, organic abstraction features graceful, freeform curves, spontaneous or accidental shapes, and visual symbols [15]. Informel abstraction is characterized by its non-representational nature, expressed through spontaneous brushstrokes and textured surfaces. Surrealist illustrations are particularly effective in conveying exaggeration and imagination, linking and combining unrelated forms to produce images that defy logical interpretation and broaden perceptual boundaries. This approach enhances emotional impact by stimulating the viewer's imaginative capacity and sensory perception. Both abstract and surrealist styles are widely employed in postmodern picture book illustration.

Table 1. Classification of expressive styles in picture book illustration

Style of expression	Characteristic	Effect
realism	accurate representation of real-world subjects without distortion	facilitates reader comprehension and empathy, particularly effective in educational picture books
expressionism	emphasis on subjective emotions, utilizing deliberate distortions and exaggerations	highlights the themes, content, and tone of the narrative
abstraction	geometric abstraction, organic abstraction, and informel styles	conveys complex ideas or concepts in a simplified manner
surrealism	combinations of images that defy logical understanding	stimulates emotions and introduces a new sensory experience

2.2 Artificial Intelligence and Picture Book Illustration

Historically, picture book illustrations have been created through conventional artistic media, including watercolor, acrylic, and pencil, supplemented by digital techniques such as computer graphics. With the advent of artificial intelligence, however, sophisticated techniques for image generation have emerged, substantially transforming this domain. Among these, diffusion models stand out as a leading approach due to their capacity to produce visually rich and stylistically coherent images from textual descriptions [16]. These models underpin widely used generative AI platforms such as DALL·E 3 and Midjourney, enabling the creation of imaginative, narrative-aligned visuals across a range of artistic styles [17].

Diffusion models, a class of generative algorithms in machine learning, are designed to synthesize new data—such as images or audio—by learning the statistical patterns of existing datasets [18]. They internalize the structure of their training data to generate new outputs that mirror those characteristics.

The process begins with forward diffusion, in which structured data such as images is gradually corrupted through the sequential addition of random Gaussian noise. This occurs via a Markov chain mechanism, a probabilistic process in which each step introduces noise based on the previous state [19]. During training, the model learns to approximate how this noise distorts the original data over time. Once the model has sufficiently captured this transformation, it proceeds to the reverse diffusion phase.

In the reverse process, the model starts from noisy inputs and learns to iteratively remove the noise, reconstructing data that resembles the original. This learned denoising capability forms the foundation for generating novel content. When generating new images, the model initiates with random noise and is guided by a text prompt that serves as a semantic blueprint for the desired output [20]. Through successive iterations of denoising, the model refines the noise into a coherent image, aligning the emerging visual content with the descriptive features outlined in the prompt [21]. This alignment is achieved by minimizing the discrepancy between the learned visual representation and the semantic cues embedded in the prompt. As a result, diffusion models are capable of generating outputs that are contextually appropriate for picture book illustration.

2.3 Prompt Design in AI-based Picture Book Illustration

2.3.1 Text Prompt

In AI-based image generation, text commands and prompts serve as the core input methods through which users guide the creation of visual content. These inputs allow models to interpret natural language and generate corresponding images, making them essential tools in creative fields such as picture book illustration. Text commands typically involve concise, descriptive phrases—e.g., "a red apple on a table"—that instruct the AI to generate an image with specific visual elements. These commands are processed by multimodal models, which are deep learning systems trained to understand and link both text and image data. A prompt, on the other hand, is often more complex and expressive. It may include stylistic, emotional, or contextual information—e.g., "a small boat breaking apart in the waves on a sunset beach, watercolor style". Compared to simple text commands, prompts offer greater flexibility and adaptability, enabling the AI to produce outputs that reflect nuanced creative intent [22, 23].

In the picture book illustration process, prompts allow creators to iteratively refine visual outputs by adjusting text inputs. This interaction reflects not only the artist's intuitive ideas but also their accumulated experiences, emotions, and narrative goals. By selecting key words that capture the essence of a concept, creators can craft prompts that result in customized and emotionally resonant illustrations. Thus, prompt design becomes a central part of the ideation and prototyping stages, bridging human creativity with AI's generative capabilities to support storytelling in a visually compelling way.

2.3.2 GPT-Based Prompt Generation

In AI-based picture book illustration research, narratives are first written and then transformed into image-specific prompts through GPT-based prompt conversion tools, a method that has been employed to concretize picture book stories into visual scenes. The difference between AI-generated image prompts and general command descriptions primarily lies in the details, contextual expression, language style, and clarity of the resulting output. AI image prompts assist in creating more accurate and detailed images by providing rich, descriptive details. For instance, "A tranquil forest with tall green trees and sunlight filtering through the leaves" gives a precise and vivid depiction of the forest. Furthermore, AI image prompts can offer the entire narrative related to a specific scene. For example, "Snowball discovers a path covered with colorful, glittering petals" not only describes the scene but also advances the storyline, helping the AI better understand and reproduce the flow and mood of the scene.

In terms of language style, AI image prompts employ vibrant language and rhetorical expressions. For instance, "The forest is lush and green with sunlight filtering through the trees, creating a magical and inviting atmosphere" conveys not only visual elements but also the mood and emotions of the scene. In contrast, general command descriptions use more direct, functional language, often simplified as in "Generate a rabbit running through the forest." Regarding the clarity of the outcome, AI image prompts offer a clear set of instructions for the image the AI needs to generate through detailed descriptions. For example, "They play happily together in the garden, surrounded by vibrant blooms and a serene, magical atmosphere" provides clear instructions for all the elements of the final scene. As a result, AI image prompts are better suited for creating complex and detailed images, effectively conveying both the story and atmosphere. Table 2 presents example test prompts for picture book illustration.

Table 2. Picture book illustration test prompt examples

Part	Picture Book Story	AI Prompt Content
(a)	in a quiet forest, there lived a little white rabbit named Snowball. Snowball had long ears and soft white fur and loved hopping around and exploring the forest.	a tranquil forest with tall green trees and sunlight filtering through the leaves. A white rabbit named Snowball with long ears and fluffy fur is happily hopping and exploring the forest, surrounded by colorful flowers and soft grass. The scene is peaceful and idyllic.
(b)	one day, while Snowball was playing in the woods, he discovered a path he had never seen before. The path was covered with colorful flower petals that sparkled like tiny gems. Filled with curiosity, Snowball decided to follow the magical trail.	a white rabbit named Snowball with long ears and fluffy fur playing in a tranquil forest. Snowball discovers a path covered with colorful, glittering petals. The rabbit looks curious and decides to follow the path to see where it leads. The forest is lush and green with sunlight filtering through the trees, creating a magical and inviting atmosphere.
(c)	as Snowball hopped along the petal-covered path, he came across a squirrel who was gathering pinecones. The squirrel greeted Snowball warmly and asked, "Would you like to join me on an adventure?" Snowball nodded happily, and the two new friends continued down the enchanting path together.	snowball, a white rabbit with long ears and fluffy fur, follows a path covered with colorful, glittering petals in a tranquil forest. Snowball meets a squirrel picking pinecones. The squirrel is animated, speaking to Snowball, who happily nods. They continue on the petal path together. The forest is lush and green with sunlight filtering through the trees, creating a magical and friendly atmosphere.
(d)	soon, they arrived at a beautiful garden. The garden was filled with flowers of every color, and butterflies fluttered gracefully between the blooms. Snowball and the squirrel played joyfully in the garden, chasing butterflies and enjoying the stunning scenery. As the sun began to set, the two friends decided it was time to head home. Snowball was thrilled about the day's adventure and looked forward to discovering more exciting journeys in the days to come.	snowball, a white rabbit with long ears and fluffy fur, and a friendly squirrel arrive at a beautiful garden filled with colorful flowers and fluttering butterflies. They play happily together in the garden, surrounded by vibrant blooms and a serene, magical atmosphere. The sunlight bathes the garden in a warm glow, making the scene feel joyful and idyllic.

3. Case Analysis Utilizing Prompts

3.1 DALL·E

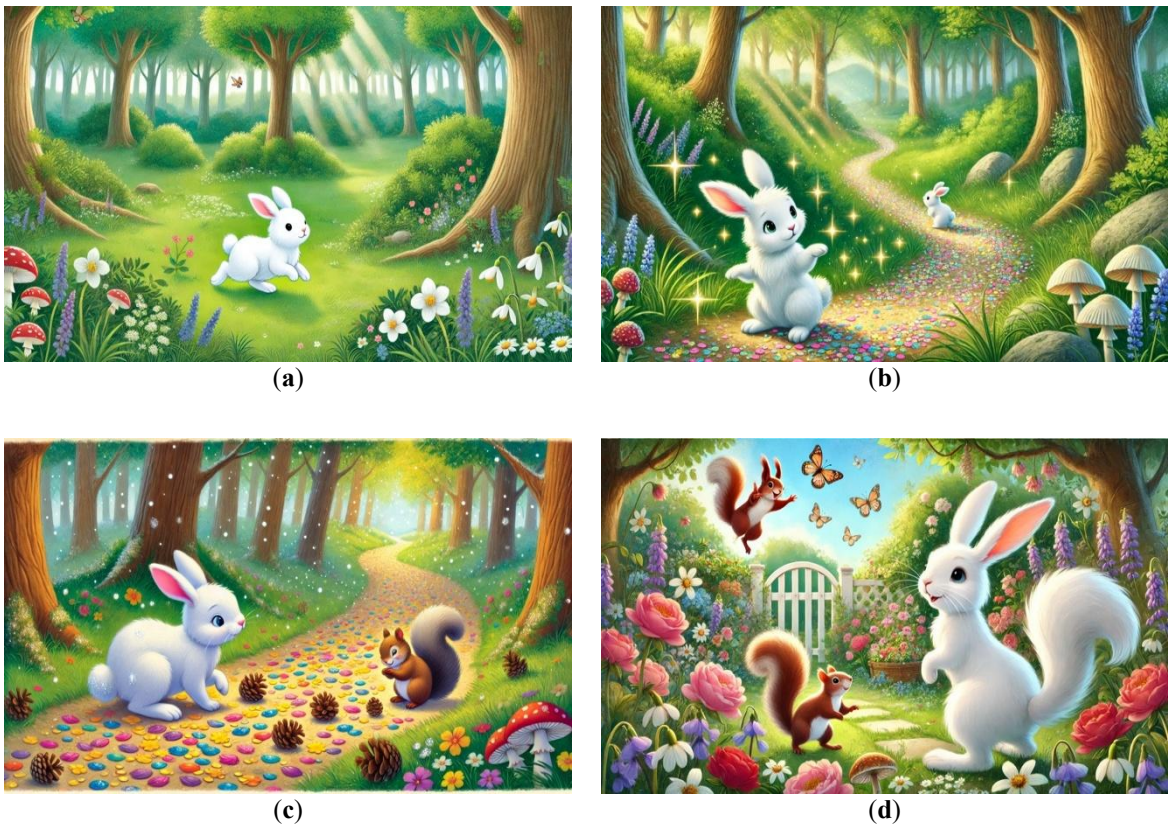
3.1.1 Characteristics of Tools

DALL·E 3, developed by OpenAI, is the latest text-to-image generation AI model, released in 2023. The name DALL·E 3 is a combination of the Pixar animated robot "WALL·E" and the surrealist artist Salvador Dalí [2]. DALL·E 3 offers intuitive and powerful image generation capabilities based on user-inputted text. While the quality of the generated images may decrease if the prompt is not detailed, the model excels at producing creative and original results. It reflects complex and specific user descriptions through enhanced text-image associations, supporting a variety of styles and themes. This allows users to obtain diverse and imaginative visual outputs. DALL·E 3 can be accessed via a subpage on OpenAI's website and is currently available for use

in the free version of GPT-4. These functional characteristics place DALL·E 3 at the forefront of modern text-to-image generation models, playing a significant role in both research and practical applications.

3.1.2 Utilization of Test Prompts

As part of the case analysis, the test prompts shown in Table 2 were entered into the DALL·E 3 interface. The resulting images are summarized in Figure 1. The illustrations created by DALL·E 3 are characterized by a gentle, fairy-tale-like style, using soft lines and vibrant colors. The illustrations incorporate rich details such as flowers, butterflies, and pinecones, with the overall composition full of life and energy. The use of color combines bright hues with a soft, balanced overall palette, providing an appealing visual effect. In terms of content reproduction, the illustrations accurately portray the scene of Snowball, the white rabbit, running and exploring in the forest. It depicts Snowball discovering a path covered with colorful petals, meeting a squirrel picking up pinecones, and finally arriving at a beautiful garden filled with flowers of various colors and fluttering butterflies. The key elements of each scene and the narrative details are accurately represented, demonstrating excellent storytelling ability. In terms of creativity, the illustration showcases unique character designs. The visual appeal of Snowball is enhanced by its large eyes and long ears, adding to the charm of the characters. The colorful petal path and sparkling effects contribute a fantastical element, stimulating the reader's imagination. The interactions between the characters are designed to be natural and warm, reinforcing the emotional depth of the story. The environmental design is intricate, with the interplay of light and shadow in the forest and the color combinations in the garden effectively showcasing creativity and artistic quality.



(**a**) (**b**)

(**c**) (**d**)

Figure 1. Example Results Using Test Prompts for Picture Book Illustrations with DALL·E 3; (a)-(d) correspond respectively to the 'AI Prompt Content' in Table 2, reflecting the respective 'Picture Book Story'

3.2 Midjourney

3.2.1 Characteristics of Tools

Midjourney, first released in March 2022, is an AI image generation tool developed by a team led by David Holz, co-founder of Leap Motion, and is accessible exclusively through the Discord chat and community app [24]. The generated images are provided in high resolution, making them suitable for commercial use in fields such as marketing, advertising, and content creation. However, since it operates via a chat platform like Discord, users

must navigate the inconvenience of joining Discord and interacting through chat commands. When commands are issued through chat, new images continuously push previous ones up the screen, requiring users to scroll through posts to find earlier images. While Discord is a familiar text and voice chat platform for gamers, new users must learn specific processes, such as the chat command format for payments, prompt entry, and server configuration. Despite these challenges, even beginners who are not accustomed to prompts can generate high-quality images, leading to a steady increase in users [25]. Initially, Midjourney offered free image generation experiences, but now free usage is no longer available. Individuals must pay $30 per month to access the service without restrictions.

3.2.2 Utilization of Test Prompts

The prompt presented in Table 2 was entered into the chat window of Midjourney, and the resulting images are summarized in Figure 2. The illustrations generated by Midjourney showcase a rich artistic style combining various techniques such as watercolor and oil painting. Overall, the style evokes a sense of dreams and childhood innocence, making it suitable for children's picture books. The color palette is bright and layered, creating a fantastic and vibrant world through delicate light and shadow effects and a rich combination of colors.

In terms of content reproduction, the illustration accurately depicts the key scenes described in the prompt. The first scene, where Snowball explores the forest, is successfully portrayed through detailed plant depictions and natural light and shadow effects, capturing the image of the white rabbit frolicking freely in the woods. Additionally, the scene where Snowball discovers a path covered in colorful petals is depicted precisely, with the vibrant colors of the petals and sparkling effects enhancing the fantastical atmosphere. Next, the interaction between Snowball and the squirrel collecting pinecones is vividly captured, with the characters' movements and expressions appearing natural and conveying strong narrative expression. Finally, the scene where Snowball and the squirrel arrive at a beautiful garden, filled with flowers of various colors and flying butterflies, is recreated with a sense of liveliness and energy, effectively conveying the vibrancy of the scene.

From a creativity perspective, the illustrations feature unique character designs. The images of Snowball are enhanced by its cute, long ears and large eyes, adding to their visual appeal. The colorful petal path and sparkling effects introduce a fantastical element to the scene, making the story more engaging and sparking the reader's imagination. The interactions between the characters are depicted in a natural and warm manner, deepening the emotional impact of the story.

(**a**) (**b**)

(**c**) (**d**)

Figure 2. Example Results Using Test Prompts for Picture Book Illustrations with Midjourney; (a)-(d) correspond respectively to the 'AI Prompt Content' in Table 2, reflecting the respective 'Picture Book Story'

To support a clearer understanding of the practical characteristics and differences between the two leading AI-based image generators, Table 3 compares DALL·E 3 and Midjourney across key features such as platform access, user interface, stylistic tendencies, and licensing. This comparison highlights how each tool supports different creative workflows and user needs in the context of picture book illustration or other visual content creation.

Table 3. Comparison of key features of DALL·E 3 and Midjourney

Category	DALL·E	Midjourney
developer	OpenAI	Midjourney team (led by David Holz)
access	ChatGPT / OpenAI site	Discord platform only
interface	natural language (chat-style)	command-based (Discord chat)
style	versatile, narrative-friendly	artistic, surreal, painterly
prompt control	high, supports detailed descriptions	moderate, stylized output even with short prompts
ease of use	Intuitive for general users	requires learning DSiscord commands
commercial use	allowed (with subscription)	allowed (with $30/month plan)

3.3 Comparative Analysis of Image Generation by DALL·E 3 and Midjourney

The illustrations created by DALL·E 3 feature a warm, fairy-tale-like style characterized by soft lines and vibrant colors, effectively conveying a dream-like atmosphere. This style is highly suitable for picture books, easily capturing readers' attention. The use of color is overall bright and soft, providing intuitive and appealing visual effects. In terms of content reproduction, DALL·E 3 accurately represents the scenes and story, focusing particularly on intuitive expressions of the narrative. The character designs are cute and charming, with long ears and large eyes as distinguishing features. Fantastical elements like the colorful petal path and sparkling effects stimulate the reader's imagination. The interactions between characters are depicted naturally and warmly, adding emotional depth to the story. The environmental design is intricate, showcasing high artistic value and creativity.

In contrast, the illustrations generated by Midjourney boast a wide range of styles, employing various techniques such as watercolor and oil painting. These illustrations offer unique visual appeal and artistic depth. The use of color is rich and profound, enhanced by delicate and striking light and shadow effects that emphasize artistic expression. In terms of content reproduction, Midjourney provides more detailed depictions, with characters' movements and expressions vividly portrayed, demonstrating strong artistic quality. The character designs are distinctive and engaging, with rich details. Fantastical elements are further heightened through the use of diverse lighting and detailed rendering. The lively interactions between characters strengthen the emotional expression of the story, while the environmental design is delicate and creative.

In sum, DALL·E excels at conveying intuitive, fairy-tale imagery, while Midjourney emphasizes artistic depth and detail. Nevertheless, the illustrations generated by DALL·E 3 demonstrate strengths in artistic style and narrative fidelity. Some character details, however, appear inconsistent across consecutive illustrations, which may reduce their suitability for maintaining a coherent narrative flow. Similarly, the illustrations generated by Midjourney also have limitations: the consecutive illustrations reveal inconsistencies in character depiction and shifts in style, medium, and technique, making them less suitable for sustaining a continuous narrative. Table 4 provides a comparative summary of the characteristics of illustrations generated by the two AI tools.

Table 4. Comparison of characteristics of illustrations generated by DALL·E 3 and Midjourney

Classification	DALL·E 3	Midjourney
painting style	a unified, warm, fairy-tale-like style with soft lines	diverse painterly styles
color palette	bright and vibrant colors with an overall soft tone	a rich and profound palette with strong contrasts, refined by delicate light and shadow
content representation	accurate depiction of scenes and narrative flow, focusing on clarity and fidelity	more refined details with dynamic and vivid depictions of characters' actions and expressions, demonstrating strong artistic expressiveness
fantastical elements	magical and imaginative atmosphere with colorful elements	rich fantasy elements with enhanced magical effects through light and shadow and intricate details

emotional elements	natural and warm character interactions, enhancing emotional depth	lively character interactions with strong emotional expression
appropriateness	suitable for children's picture books requiring intuitive clarity	suitable for single-scene illustrations requiring high artistic expression and detailed depiction
narrative consistency	inconsistent character details across scenes	inconsistent character depiction, accompanied by variations in style, medium, and technique between scenes

4. Analysis Results on the Potential of AI Tools in Illustration Development

4.1 Illustration Development Process Through AI

As examined in the preceding case, the process of creating picture book illustrations using AI generation tools can be structured into five stages. The first step is concept development and initial planning, where the picture book story is established and the initial ideas are sketched. Setting clear goals is essential to define the final purpose of the illustrations, which in turn guides the selection of the appropriate style and atmosphere for the story. For example, the characteristics of the illustrations will differ depending on the purpose, such as whether the book is intended for young children, educational purposes, storytelling, or interactivity.

The second step is AI prompt creation. In this stage, detailed prompts are written in a way that the AI can understand. The prompt must clearly describe the specific details of the illustration and the desired style. For example, a prompt like "A tranquil forest with tall green trees and sunlight filtering through the leaves. A white rabbit named Snowball with long ears and fluffy fur is happily hopping and exploring the forest, surrounded by colorful flowers and soft grass" includes specific descriptions. Additionally, reference images that resemble the desired style are collected and provided to the AI tool. This helps the AI generate more accurate results.

The third step involves utilizing AI image generation tools. In this phase, AI tools are used to generate images, and the results are reviewed. When selecting AI tools like DALL·E 3 and Midjourney, it is important to consider the characteristics and strengths of each tool. For example, DALL·E 3 excels in the clear and detailed depiction of color and form, making it suitable for cases that require a unified style and clear visual representation. On the other hand, Midjourney excels in various styles and creative expression, making it ideal for picture books that require artistic sensibility and intricate details. Designers should use these tools to generate multiple versions of images and review the results to ensure they align with the desired style.

The fourth step is revision and refinement. In this stage, any missing or necessary modifications in the generated images are carefully addressed. Graphic software such as Photoshop can be used for manual adjustments, such as fine-tuning colors and adding additional details. The fifth and final step is design and distribution. Once image quality is verified, all revisions and refinements are completed, and a final review ensures consistency with the initial objectives. The illustrations are then combined with text using editorial design software such as Adobe InDesign, formatted into book layout, and either printed or distributed digitally.

4.2 Challenges and Constraints of AI-Driven Illustration

This study has examined the potential of DALL·E 3 and Midjourney as tools for picture book illustration and identified their respective strengths and weaknesses. However, AI-based tools are fundamentally constrained by their reliance on existing data. These limitations can be examined in more detail as follows.

4.2.1 Limitations of Creativity

The limitations of AI models' creativity stem from their reliance on data, the absence of imagination and intuition, and constraints in understanding context. While AI recognizes patterns within the learned data and generates transformed results, it struggles to create entirely new styles or concepts. This is because AI operates solely based on training data, restricting its capacity to produce truly novel ideas. Diffusion models, which are widely used in current state-of-the-art image generation systems, excel at producing highly detailed and coherent images from text prompts [6]. However, they too are constrained by the statistical distributions learned from their training sets. While they can recombine visual elements in creative ways, their outputs are essentially extrapolations of what already exists, lacking the spontaneous originality and intuitive leaps seen in human creativity [16]. To address these limitations, ongoing research focuses on enriching training data, incorporating

human-in-the-loop feedback, and exploring hybrid approaches that blend human and machine creativity. Therefore, although diffusion-based AI tools offer powerful support for creative tasks, they cannot fully replace human originality and artistic intuition. A balanced, collaborative approach between AI and human creators remains essential.

4.2.2 Limitations of Data Bias

Bias in illustration datasets arises during the processes of data collection, processing, and labeling, and can significantly impact the performance and fairness of AI models. During data collection, selective bias may occur, favoring specific sources or styles, while representational bias can arise if the dataset fails to adequately represent a diverse range of styles, cultures, and themes. In the data preprocessing phase, preprocessing bias may emerge if certain types of data are excessively altered or removed, while bias in the refinement process may involve selectively removing images with certain styles or elements. In the data labeling stage, cognitive bias can occur due to subjective judgments by labelers, and nominal bias can arise if too much emphasis is placed on specific classes or categories [3]. These biases reduce the generalization ability of AI models, leading to inconsistent performance across different contexts, and may result in discriminatory outcomes for specific groups or styles, thereby undermining fairness. To mitigate these biases, efforts are needed to increase dataset diversity, apply fair labeling methodologies, and carefully review the data processing stages to build reliable AI models.

4.2.3 Ethical and Intellectual Property Issues

The use of AI in picture book illustration raises several ethical and intellectual property concerns. Key issues include potential copyright infringement, reduced roles for human creators, biased representations due to unbalanced training data, privacy violations, and lack of transparency or accountability. AI models, particularly diffusion models used in tools like DALL·E 3 and Midjourney, generate images by learning from vast datasets that may contain copyrighted works, raising legal questions about ownership and originality [16]. Since users cannot confirm whether outputs are influenced by protected material, using AI-generated images in commercial contexts may carry legal risks, especially when licensing terms are unclear. In addition, the increasing reliance on AI may diminish opportunities for illustrators and creative professionals. Data bias can lead to stereotypical or culturally insensitive depictions, while training on real-world images may result in privacy violations if identifiable features are reproduced without consent [18]. Finally, without clear records of how images are generated, it becomes difficult to ensure accountability or defend authorship. To address these challenges, it is important to choose AI platforms with transparent licensing policies, document the creative process, and integrate ethical and legal considerations from the outset to ensure responsible use of AI in picture book illustration.

5. Conclusion

This study aimed to analyze how AI image generation tools, specifically DALL·E 3 and Midjourney, are applied in the development of picture book illustrations through case studies, while also exploring the illustration creation process and associated challenges. The findings reveal that DALL·E 3 produces illustrations with soft lines and vibrant, pastel-like colors that evoke a warm, fairy-tale atmosphere suited for children's content, whereas Midjourney demonstrates strength in artistic expression through stylistic diversity, rich color depth, and detailed light and shadow effects. DALL·E 3 emphasizes emotional warmth and accessibility, while Midjourney enhances narrative depth through dynamic visual storytelling. The creative workflow was identified in five key stages: concept development, prompt preparation, AI tool usage, image refinement, and final production. Despite their strengths, both tools exhibit limitations rooted in their reliance on training data, lack of imagination, and difficulty in context comprehension, which hinder their ability to create genuinely novel or intuitive content. Addressing these shortcomings requires human-AI collaboration, use of multimodal systems, and expanded, fairly labeled datasets to reduce bias. Furthermore, ethical issues such as copyright infringement, labor displacement, data bias, and privacy concerns necessitate early-stage legal and ethical integration in the design and use of AI systems. Although this study offers practical insights, it remains constrained by its reliance on subjective analysis without objective user data. Future research should therefore adopt empirical methods and statistical modeling to more rigorously evaluate user satisfaction and the practical effectiveness of AI tools in picture book illustration.

Conflicts of Interest: The authors declare no conflict of interest.

References

[1] B. C. Kim, Y. J. Ahn, and K. K. Yoo, "Incorporation of Artificial Intelligence and Smart Factories in Domestic Manufacturing Industries," Journal of the Korean Society of Mineral and Energy Resources Engineers, vol. 58, no. 6, pp. 620-625, 2021, doi: http://doi.org/10.32390/ksmer.2021.58.6.620.

[2] D. H. Kwon, "Analysis of Prompt Elements and Use Cases in Image-Generating AI: Focusing on Midjourney, Stable Diffusion, Firefly, DALL·E," Journal of Digital Contents Society, vol. 25, no. 2, pp. 341-354, 2024, doi: https://doi.org/10.9728/dcs.2024.25.2.341.

[3] Y. C. Zhang and S. H. Park, "A study on the social value of digital art using an AI image generator -Digital art research using AI image generator in the Post-Human Era," Journal of Communication Design, vol. 82, no. 1, pp. 107-118, 2023, doi: https://doi.org/10.25111/jcd.2023.82.09.

[4] K. S. Park, "Study on the feasibility of using AI image generation tools for fashion design development -Focused on the combined use of Midjourney," The journal of Convergence on Culture Technology, vol. 9, no. 6, pp. 237-244, 2023, doi: https://doi.org/10.17703/JCCT.2023.9.6.237.

[5] J. H. Lim and S. J. Yun, "A Study on the Possibility of Expressing Visual Identity Design using Image Generation AI- Focusing on Midjourney -," A Journal of Brand Design Association of Korea, vol. 22, no. 1, pp. 265-276, 2024, doi: https://doi.org/10.18852/bdak.2024.22.1.265.

[6] H. S. Park and K. S. Kim, "Comparative Study of AI for Generating Basic Images of 3D Characters Using DALL·E 3 and Midjourney," Journal of Digital Art Engineering & Multimedia, vol. 11, no. 2, pp. 149-157, 2024, doi: https://doi.org/10.29056/jdaem.2024.06.01.

[7] N. Yang and C. S. Kim, "A Study on the Renewal Design of Rice Wine Packaging in Suzhou Bridge to Strengthen Traditional Regional Culture -Focusing on the Use of Generative Artificial Intelligence Tool-," Journal of the Korean Society of Design Culture, vol. 30, no. 2, pp. 195-209, 2024, doi: https://doi.org/10.18208/ksdc.2024.30.2.195.

[8] J. C. Horne, "Six decades of picture book illustration: The art of Barbara Cooney," Children's Literature in Education, vol. 32, no. 2, pp. 91-109, 2001, doi: http://doi.org/10.1023/A:1010382411727.

[9] Z. Chen, L. Chen, Z. Zhao, and Y. Wang, "AI illustrator: Art illustration generation based on generative adversarial network," presented at the 2020 IEEE 5th International conference on image, vision and computing (ICIVC), Beijing, pp. 155-159, Jul. 2020, doi: http://dx.doi.org/10.1109/ICIVC50857.2020.9177494.

[10] P. Nodelman, Intl Comp Ency Child Lit E2 V1, Routledge, 2018.

[11] J. K. Ahn, J. H. Lee, E. K. Kim, Picture Books for College Students, Knowledge Community, 2016.

[12] A. Villarreal, S. Minton, and M. Martinez, "Child illustrators: Making meaning through visual art in picture books," The Reading Teacher, vol. 69, no. 3, pp. 265-275, 2015, doi: http://doi.org/10.1002/trtr.1405.

[13] S. C. Kang, "Early Picture Book Illustrations of Walter Crane," Journal of the Korean Society of Design Culture, vol. 20, no. 61, pp. 5-16, 2019, doi: https://doi.org/10.37379/JKSIR.2019.61.1.

[14] Y. G. Han and K. S. Kim, "Early Picture Book Illustrations of Walter Crane," Research on the Spatial Expression Characteristics of Illustration in Picture Books, vol. 21, no. 3, pp. 131-142, 2021, doi: https://doi.org/10.5392/JKCA.2021.21.03.131.

[15] L. Mason, P. Pluchino, M. C. Tornatora, and N. Ariasi, "An eye-tracking study of learning from science text with concrete and abstract illustrations," The Journal of Experimental Education, vol. 81, no. 3, pp. 356-384, 2013, doi: http://doi.org/10.1080/00220973.2012.727885.

[16] Altexsoft, AI Image Generation Explained: Techniques, Applications, and Limitations, 2011. [Online] Available: https://www.altexsoft.com /blog/ai-image-generation

[17] V. Sorin, Y. Barash, E. Konen and E. Klang, "Creating artificial images for radiology applications using generative adversarial networks (GANs)–a systematic review," Academic radiology, vol. 27, no. 8, pp. 1175-1185, 2020, doi: https://doi.org/10.1016/j.acra.2019.12.024.

[18] C. Zhang, C. Zhang, M. Zhang and I. S. Kweon, "Text-to-image diffusion models in generative ai: A survey," Academic radiology, arXiv preprint, vol. 27, no. 8, pp. 1175-1185, 2020, doi: https://doi.org/10.48550/arXiv.2303.07909.

[19] Q. Xu, H. Wang, L. Meng, Z. Mi, J. Yuan, and H. Yan, "Exposing fake images generated by text-to-image diffusion models," Pattern Recognition Letters, vol. 176, pp. 76-82, 2023, doi: https://doi.org/10.1016/j.patrec.2023.10.021.

[20] Q. Cai, M. Ma, C. Wang, and H. Li, "Image neural style transfer: A review," Computers and Electrical Engineering, vol. 108, p. 108723, 2023, doi: https://doi.org/10.1016/j.compeleceng.2023.108723.

[21] J. S. Ubhi and A. K. Aggarwal, "Neural style transfer for image within images and conditional GANs for destylization," Journal of Visual Communication and Image Representation, vol. 85, p. 103483, 2022, doi: https://doi.org/10.1016/j.jvcir.2022.103483.

[22] H. H. Park and H. W. Han, "Analysis of Prompt Writing in Generative AI," The Journal of the Korea Contents Association, vol.24, no.4, pp. 59-67, 2024, doi: https://doi.org/10.5392/JKCA.2024.24.04.059.

[23] L. Kim and N. Y. Kim, "Context-Based Prompt Selection Methodology to Enhance Performance in Prompt-Based Learning," Journal of The Korea Society of Computer and Information, vol. 29, no. 4, pp. 9-21, 2024, doi: https://doi.org/10.9708/jksci.2024.29.04.009.

[24] P. Jie and X. Shan, "Comparative Analysis of AI Painting Using [Midjourney] and [Stable Diffusion]-A Case Study on Character Drawing," International Journal of Advanced Culture Technology, vol. 11, no. 2, pp. 403-408, 2023, doi: https://doi.org/10.17703/IJACT.2023.11.2.403.

[25] D. M. Hanna, "The use of artificial intelligence art generator "midjourney" in artistic and advertising creativity," Journal of design sciences and applied arts, vol.4, no.2, pp. 42-58, 2023.