Smart Media Journal / Vol.14, No.12 / ISSN:2287-1322
https://dx.doi.org/10.30693/SMJ.2025.14.12.40

An Empirical Study on the Applicability of Image Generation AI in Fashion Design: A Comparative Analysis of GPT, LOOK AI, and Diffusion

Sujin Lim

Abstract

This study examines the applicability of generative AI in fashion design through a comparative analysis of GPT-based generators, LOOK AI, and Stable Diffusion. A dress sketch was chosen as a complex test item, and images were evaluated by 15 professional designers using a structured survey across three domains: design reproduction, visual fidelity, and practical usability. ANOVA results showed that LOOK AI achieved the highest accuracy and usability, GPT-based tools performed moderately with strengths in structural interaction, and Stable Diffusion, while creative, showed lower fidelity. By combining expert evaluation with a complex garment type under controlled conditions, the proposed method provides a practical framework for validating AI image tools, offering implications for fashion practice and education as well as potential extension to other design fields.

Keywords : Generative AI | Image Generation Design Fidelity | Practical Usability

I. INTRODUCTION

The fashion industry is rapidly advancing digital transformation, with generative AI emerging as a pivotal visualization tool in the early stages of design. According to global market research, the AI-based fashion market is expected to reach approximately $3.14 billion by 2025, with an annual growth rate of nearly 40%[2]. This growth underscores the practical value of AI tools in visualizing designers' ideas, producing content, and improving communication efficiency. Among them, GPT-based multimodal generators, LOOK AI, and the Stable Diffusion model are the most actively applied in the fashion industry. GPT-based tools enable precise control of style and details through text input [4]; LOOK AI offers functions tailored to apparel design, achieving high adoption among practitioners; and Stable Diffusion is widely used for its capacity to generate creative, high-resolution imagery[1]. For example, Zalando, a leading European online fashion platform, employs AI to automatically generate about 70% of its editorial images, reducing production time from 68 weeks to just 34 days and lowering costs by roughly 90% [3].

Despite such industrial adoption, existing research has concentrated mainly on technical performance and

Manuscript : 2025.09.02
Revised : 2025.09.21
Confirmation of Publication: 2025.10.06
Corresponding Author :Sujin Lim, e-mail : sujin_lim@snu.ac.kr

algorithmic accuracy, with limited exploration of the tools' utility from a designer's practical perspective. Dresses, in particular, pose significant challenges for AI visualization due to their integrated top–bottom structure and complex details such as puffs, pleats, and punching. Prior studies have identified dresses as suitable items for verifying silhouette continuity, volumetric presence, and detail reproduction. Building on this, the present study focuses on comparing GPT, LOOK AI, and Diffusion for their applicability to dress design. Nevertheless, comprehensive comparative analyses of complex garment structures remain scarce. Most prior studies assessed only visual similarity or technical capacity of AI-generated outputs. For instance, [6] compared Midjourney and Runway in educational contexts, but focused narrowly on technical aspects[5] acknowledged strengths in realism but emphasized limitations in creativity and customization. [7] explored Midjourney's commercial potential through a case study of a single tool. In sum, prior work has remained fragmented, without cross-tool evaluations or systematic validation of generative AI's practical applicability in fashion design. In practice, however, designers require more than visual similarity. Key quality factors include realistic texture expression, drape performance, interaction with the human body, and usability in planning documents or sample instructions. Yet systematic discussion of these factors remains limited. For example, in producing a Tech Pack (Technical Specification Package), design teams typically provide a line illustration, list sewing details, and attach fabric swatches. Such illustrations, while functional, cannot fully convey material thickness, sheen, or drape. By contrast, generative image tools can visually reproduce these properties, showing the puff volume of sleeves, the fall of pleats, or the three-dimensionality of punched details.

Consequently, their use in factories may enhance material comprehension, reduce communication errors, and minimize trial-and-error in sampling. Against this backdrop, this study aims to move beyond simple visual matching and instead evaluate the practical applicability of GPT-based generators, LOOK AI, and Diffusion through direct comparison of outputs based on the same dress sketch. The significance of this research is threefold. First, it extends beyond technical performance to assess strengths and limitations of AI tools from a designer's practical perspective. Second, it verifies the feasibility of integrating generative images into industrial documents such as technical drawings. Third, by examining a complex garment type, it highlights the distinctive capabilities of different AI models, offering a foundation for future academic research and educational applications.

Research Question 1. What are the differences in design fidelity (silhouette, details, proportions, etc.) among GPT-based generators, LOOK AI, and Stable Diffusion in images generated from the same dress sketch? Research

Question 2. What differences exist in the visual completeness and realism (texture, drape, light source representation, etc.) of the image generated among the three tools?

Research Question 3. Do the generated images from the three tools show significant differences in terms of practical applicability and user satisfaction (usability in proposals, suitability for communication, future usage intent, etc.)?

II. THEORETICAL BACKGROUND

1. Generative AI Tools and the Digital Transformation of Fashion Design

Generative AI is an artificial intelligence technology that creates new data based on various user inputs such as text, images, and voice. Alongside recent advancements in deep learning technology, it has established itself as a core tool for content generation [4].

Particularly, the advancement of Text-to-Image technology is gaining attention as a novel approach that simultaneously fulfills creativity and efficiency in the design planning stage within the visually-centric fashion design industry. Representative generative image tools include OpenAI's GPT, Stable Diffusion, and LOOK AI, each possessing the following technical characteristics. GPT-based tools generate intuitive and contextually appropriate image outputs based on natural language interpretation capabilities [1], while diffusion-based models excel in high-resolution and detailed texture representation[4]. LOOK AI is trained with a structure specialized for fashion design, characterized by high accuracy in silhouette components and strong practical applicability [4].

The entire fashion industry is rapidly accelerating its digital transformation with the adoption of AI-based tools. Processes that previously relied on manual labor (such as drawing) have evolved with the introduction of CAD systems, 3D simulation, and virtual fitting technology, enabling swift and flexible design work centered around visualization. Furthermore, generative AI, which has recently gained significant attention, is being utilized throughout the entire design process—from initial idea generation and sketch refinement to proposal creation and sample production instructions transcending the mere dimension of digitalization. In particular, generative AI significantly enhances the accuracy and speed of communication through visualization, effectively compensating for the expressive limitations and time constraints inherent in traditional illustration-based systems.

2.Visualization Capabilities and Practical Application Potential of Generative AI

AI image generation tools support fashion design practice in three key ways. First, they transform verbal concepts into images, enabling rapid externalization and expansion of design ideas[4]. Second, they supplement sketches by conveying silhouettes, proportions, details, and material textures, which facilitates more precise, visually based communication. Third, they improve planning and collaboration by serving as practical

visuals in documents, client proposals, and sample production instructions, thereby enhancing efficiency with both internal teams and external partners[7]. In practice, tools such as GPT, Stable Diffusion, and LOOK AI are already being applied in tasks ranging from client presentations and SNS content to technical planning documents, and some brands even employ AI-generated images as official materials before sample production[1]. Despite this growing adoption, notable differences remain among tools in image precision, detail expression, and material reproduction, highlighting the need for comparative analysis and clearer criteria for tool selection.

3. Structure and Visualization Requirements for Dress Design.

A dress is a one-piece garment that combines top and bottom, integrating diverse silhouettes and decorative elements. This structural and visual complexity makes dresses both central to fashion practice and difficult to visualize accurately. In design work flows, sketches usually convey silhouette and details such as sleeves, length, or seam lines, which are then developed into planning documents specifying color, fabric, and styling.

These documents guide the creation of samples, followed by iterative feedback and revisions. Visual materials in this process must go beyond representation to ensure accurate proportions, consistent details, realistic fabric texture and drape, and clear communication with clients and factories.

Because dresses incorporate multiple structural and decorative features simultaneously, precise reproduction is essential when evaluating image generation tools. Traditional methods, such as annotated samples or 2D schematics in Illustrator, often led to miscommunication. Generative AI, however, enables high-resolution visualizations of a sketch from multiple perspectives and styles, enhancing clarity and accuracy in design communication. Based on this rationale, the study compared three representative tools—GPT-based generators, LOOK AI, and Stable Diffusion—selected for their versatility, specialization, and accessibility. The same dress sketch was generated with each tool across multiple views and details, and these outputs were used as stimuli for the designer survey.

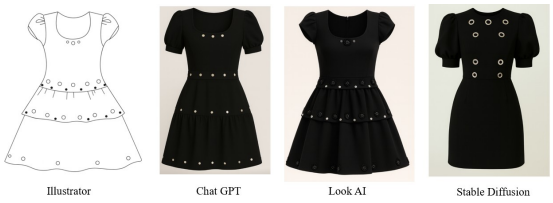

Fig. 1. AI-generated thumbnails of a dress design using GPT, LOOK AI, and Stable Diffusion

III. METHOD

This study designed a survey in which professional designers evaluated images generated by three generative AI tools GPT, LOOK AI, and Stable Diffusion based on the same dress sketch. The evaluation questionnaire comprised 16 items across three domains: (1) Design Reproduction (silhouette, detail, proportion; 4 items), (2) Visual Fidelity (texture, lighting, drape, color; 5 items), and (3) Practical Usability (planning,

sample instructions, communication, satisfaction; 7 items). All items were measured on a 5-point Likert scale. Participants reviewed images from each tool blindly and provided repeated responses using the same criteria. Fifteen designers from domestic and international fashion brands participated in the survey, with an average of 7.4 years of professional experience. This sample size satisfies the recommended range of 10–20 participants for exploratory studies using expert sampling and is considered sufficient for stimulus evaluation. All participants had prior experience in proposal writing, sample production instruction, and the use of digital tools, as well as basic familiarity with generative AI. They were recruited through purposive sampling, and the study' s objectives and procedures were fully explained in advance. Collected data were analyzed using SPSS 27.0. One-way ANOVA was conducted for each item to test for mean differences between tools, with a significance threshold of $p < .05$. For items showing significant differences, Scheffé post-hoc tests were applied to identify statistically significant tool combinations. This process enabled the identification of strengths and weaknesses of each tool, their relative advantages in practical applicability, and designer preferences. To ensure clarity and validity, the F-values, p-values, and comparison results were presented in tables. The proposed method offers clear advantages. By engaging professional designers as evaluators, it moves beyond algorithm-centered testing and ensures that practical applicability is directly assessed. Selecting dresses as the stimulus provides a rigorous benchmark, since their integrated structure and intricate details (e.g., pleats, puffs, perforations) pose greater visualization challenges than other garments. Finally, comparing GPT-based generators, LOOK AI, and Stable Diffusion under identical conditions with the same sketch establishes a controlled framework that systematically highlights tool-specific strengths and limitations.

Ⅳ. RESULT

1.Evaluation of Design Reproducibility

A significant difference was found among the three tools in the 'Overall Silhouette Reproducibility' item (F=30.556, p<.001). Post-hoc tests revealed evaluations in the order LOOK > GPT > Diffusion. LOOK showed the highest match in implementing the original sketch's silhouette, GPT demonstrated an intermediate level, and Diffusion exhibited relatively low reproducibility. This trend was similarly confirmed in the 'Detailed Implementation Capability' category (F=25.844, p<.001). Specifically, LOOK precisely rendered complex garment details like punched eyelets, pleats, and puffs, while GPT succeeded in basic reproduction but showed limitations in delicate details. Conversely, in the 'Unnecessary Element Insertion' category (F=13.288, p<.001), Diffusion significantly included more unnecessary visual elements than other tools. This suggests Diffusion excels at creative and artistic variations but may be

disadvantaged in terms of consistency with garment structure. Additionally, in the 'Naturalness of Proportions and Portions' category (F=15.074, p<.001), Diffusion received the highest proportion of 'Awkward' Responses, highlighting relatively prominent issues with visual distortion and structural imbalance. In summary, LOOK received the highest evaluation for faithful reproduction of silhouettes and details, while GPT achieved moderate stability but showed limitations in fine expression.

Conversely, Diffusion demonstrated high potential for creative expression but exhibited relatively weak characteristics in structural consistency and reproducibility.

2.Visual Fidelity and Realism Evaluation

No statistically significant differences were found among the three tools for items related to material texture, lighting effects, color, and perspective representation. Specifically, for "Material Texture" (F=0.764, p=.473), "Lighting and Shading Effects" (F=0.322, p=.727), "Color Tone Accuracy" (F=1.762, p=.185), and "3D Volume and Perspective Representation" (F=0.918, p=.408). However, in the 'Physical Interaction Between Human Body and Clothing' category (F=5.978, p=.005), LOOK and GPT scored significantly higher than Diffusion. This indicates Diffusion's relative weakness in achieving structural naturalness, particularly when dynamic details like wrinkles and drape interact with the human body. In other words, while Diffusion excels at generating high-resolution photorealistic images, it revealed limitations in reflecting the physical properties of clothing deforming with body movement. Conversely, LOOK and GPT were interpreted as receiving higher practical evaluations for expressing these interactions more realistically. This demonstrates tool differentiation in the crucial practical aspect of achieving 'drape and structural fit.

3.Evaluation of Practical Applicability

LOOK AI also received the highest evaluation in the 'Design Intent Communication Effectiveness' category (F=16.757, p<.001), with GPT at an intermediate level and Diffusion at the lowest level. This clearly demonstrates how well silhouette, detail, and component visual clarity effectively convey the designer's planning intent, signifying their crucial role in practical planning and communication processes. Furthermore, significant differences were observed in the categories of feasibility for use in planning documents/proposals (F=21.348, p<.001), feasibility for sample production instructions (F=11.683, p<.001), and feasibility for communication with pattern teams/factories (F=10.763, p<.001). LOOK and GPT received higher ratings than Diffusion across all three items. LOOK, in particular, demonstrated the highest reliability in terms of accuracy and communicative power in practical communication. This is interpreted as LOOK providing visualization features that reflect practical needs as a fashion-specialized platform.

Consequently, it shows that ensuring visual clarity during the planning stage directly impacts the efficiency of practical collaboration. User Satisfaction and Tool Evaluation LOOK and GPT also scored significantly higher than Diffusion on the 'Overall Quality Satisfaction' (F=11.544, p<.001) and 'Intention to Use in the Future' (F=8.854, p=.001) items. This indicates that designers prioritize structural and practical elements such as silhouette implementation capability, detail expression capability, and practical applicability over simple image generation when selecting tools. Additionally, LOOK received significantly higher ratings than Diffusion in the 'Ease of Use' category (F=7.119, p=.002), while no significant difference was observed compared to GPT.

This result is interpreted as a positive outcome of LOOK providing an intuitive interface and user-centric features as a platform specialized for fashion design. In contrast, while Diffusion excels in creative variation, it appears relatively deficient in terms of ease of operation and intuitiveness required in practical contexts. In summary, designers exhibit higher satisfaction and willingness to use tools that provide structural fidelity and practical utility, suggesting this could become a key criterion for future AI tool adoption strategies in professional settings. Table 1 below summarizes the statistical significance and relative advantages across key evaluation items for the three tools: GPT, LOOK AI, and Stable Diffusion.

Table 1.Summary of Evaluation Results across AI Tools (GPT, LOOK AI, Stable Diffusion)

Category	Significant Difference	Evaluation Ranking
Overall Silhouette Reproduction	Yes (p < .001)	LOOK > GPT > Diffusion
Detail Implementation Accuracy	Yes (p < .001)	LOOK > GPT > Diffusion
Inclusion of Unnecessary Elements	Yes (p < .001)	Diffusion > GPT ≈ LOOK
Naturalness of Proportion and Balance	Yes (p < .001)	Diffusion < GPT ≈ LOOK
Material Texture, Lighting, Color, Depth	No	Similar across tools
Human–Garment Physical Interaction	Yes (p = .005)	GPT ≈ LOOK > Diffusion
Communication of Design Intent	Yes (p < .001)	LOOK > GPT > Diffusion
Practical Utility (Planning/Sampling)	Yes (p < .001)	LOOK > GPT > Diffusion
Overall Satisfaction & Future Use Intention	Yes (p < .001)	LOOK > GPT > Diffusion
Interface Usability	Yes (p = .002)	LOOK > Diffusion

V. CONCLUSION

This study compared and analyzed images generated by three tools—GPT, LOOK AI, and Stable Diffusion—based on the same dress sketch, and verified their practical applicability through evaluations by professional designers. The analysis revealed that LOOK AI received the highest ratings for silhouette and detail reproduction, as well as practical applicability. While GPT-based tools could generate creative images, they remained at an intermediate level due to limitations in detail. Stable Diffusion showed a certain level of visual realism but received the lowest evaluation due to

silhouette mismatches and low practical suitability.

Furthermore, no significant differences emerged among the tools in visual realism metrics like texture, lighting, and color sense. This suggests future competitiveness will likely be determined by accurate reproduction of design specifications and practical applicability. These results provide three key implications. First, fashion practitioners should select tools based on design reproduction accuracy and communication potential rather than simple image quality. Second, fashion education settings should cultivate students' ability to critically evaluate not only image quality but also practical applicability. Third, AI tool developers need to improve algorithms reflecting silhouette and detail implementation capabilities, as well as practical applicability.

However, this study has limitations, being confined to a single item (dresses) and a sample of 15 experts. Future research should deeply explore strategies for integrating AI image generation into design practice, encompassing diverse items, user groups, and tool-specific usage conditions.

REFERENCES

[1] Generative AI in fashion: Use cases and trends, 2024. [Online]. Available: https://research.aimultiple.com/generative-ai-fashion (accessed Aug., 1, 2025)

[2] AI in fashion: Statistics you should know, 2024. [Online]. Available: https://bestcolorfulsocks.com/blogs/news/ai-in-fashion-statistics (accessed Aug., 3, 2025)

[3] Zalando embraces AI for faster fashion content production. *Retail Systems*, 2025. [Online]. Available: https://www.retail-systems.com/rs/Zalando_Embraces_AI_For_Faster_Fashion_Content_Production.php (accessed Jul., 7, 2025)

[4] Jeon, Y. K., & Kim, O. K., *Digital Marketers Chat GPT Utilization Strategy*, Wikibooks, pp. 3–25, 2023.

[5] Jin, Y., Yoon, J., Self, J., & Lee, K., "Understanding fashion designers' behavior using generative AI for early-stage concept ideation and revision," *Archives of Design Research*, vol. 37, no. 3, pp. 1–21, 2024.

[6] Rizzi, M., "Exploring the generative AI potential in the fashion design process: An experimental experience on the collaboration between fashion design practitioners and generative AI tools," *European Journal of Cultural Management and Policy*, vol. 15, no. 1, pp. 87–102, 2025.

[7] Zhang, W., & Liu, H., "Generative AI in fashion design: Opportunities and challenges of Midjourney for creative industries," *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 19, no. 1, pp. 654–670, 2024.

Authors

Lim-Sujin

She received her B.A. degree in Studio Art from the State University of New York at Stony Brook, an M.S. in Fashion Industry Studies from Yonsei University, and a Ph.D. in Clothing and Textiles from Seoul National University.

She is currently a lecturer at Seoul National University, Hongik University, Gachon University, and the Catholic University.

Her main research interests include AI-based fashion content planning, AI-driven design portfolio development, and AI-integrated educational design.