 https://doi.org/10.24191/jeesr.v29i1.020

Lightweight GAN and Stable Diffusion in Batik Design Generation: A Review

Darni Darmin, Ihsan Mohd Yassin*, Azlee Zabidi, Fairul Nazmie Osman, and Nooritawati Md Tahir

***Abstract*—Generative artificial intelligence has emerged as a transformative tool for traditional textile innovation, including batik design synthesis. Among dominant paradigm, Lightweight Generative Adversarial Networks (LGANs) and diffusion-based models such as Stable Diffusion represent two contrasting architectural approaches in batik image generation. However, their relative suitability for batik motif generation remains insufficiently systematized. This paper presents a comparative review of LGAN and Stable Diffusion in the context of batik design. The analysis examines architectural mechanisms, training stability, computational efficiency, optimization strategies, and evaluation protocols, incorporating quantitative metrics including FID, IS, PSNR, SSIM, LPIPS, KID and CLIP-score. The review indicates that LGANs reduce computational requirements once trained, enabling real-time experimentation with batik motif variations on less powerful hardware, whereas Stable Diffusion demonstrates greater semantic controllability and high-resolution motif synthesis through latent denoising processes. The study further argues that quantitative metrics alone are insufficient to assess culturally grounded textile generation, emphasizing the need for integrated human-centred and authenticity-aware evaluation. This review outlines the basis for a domain-specific evaluative framework to guide architecture selection and future research in sustainable AI-assisted batik innovation.**

***Index Terms*—Batik design generation, diffusion models, generative adversarial networks, lightweight GAN, stable diffusion**

I. INTRODUCTION

Batik constitutes both a cultural heritage artifact and a strategic economic asset within Southeast Asia, particularly in Malaysia, where it contributes to tourism, small-medium enterprises (SMEs), and community-based industries [1], [2], [3]. Despite its cultural prominence, the Malaysian batik sector faces persistent challenges, including limited digital integration, repetitive motif development, scalability constraints, and increasing competition from industrial textile production [2], [3]. In parallel with national digital transformation initiatives, the integration of artificial intelligence (AI) into creative industries has emerged as a viable pathway to enhance productivity while preserving cultural authenticity. Recent AI-driven textile design studies demonstrate that generative models can support motif innovation without undermining heritage characteristics [4], [5], thereby positioning generative AI as a strategic enabler for sustainable batik modernisation.

Generative image modelling has advanced significantly with the emergence of Generative Adversarial Networks (GANs) and diffusion models as dominant paradigms for high-fidelity image synthesis [6], [7]. GANs generate realistic images through adversarial optimisation but are frequently affected by instability, mode collapse, and convergence sensitivity [8], [9], [10]. To mitigate computational and deployment limitations, Lightweight GANs (LGANs) incorporate parameter reduction strategies such as pruning, architectural simplification, and quantisation, enabling efficient operation on resource-constrained hardware [11], [12]. In contrast, diffusion models formulate synthesis as a progressive denoising process, enabling the generation of high-quality samples [13]. Stable Diffusion, which operates in latent space with U-Net-based denoising and cross-attention conditioning, demonstrates strong semantic controllability and high-resolution output capability across creative domains [14], [15]. These distinct architectural philosophies reflect a fundamental trade-off between computational efficiency and generative stability.

In batik design generation, both approaches have shown promising yet fragmented developments. Lightweight adversarial architectures such as Light Wasserstein GAN with Gradient Penalty (LWGAN-GP) have demonstrated efficiency in synthesising batik motifs under constrained computational settings [16], [17], while diffusion-based approaches enhanced with Low-Rank Adaptation (LoRA) and conditioning mechanisms have achieved improved stylistic fidelity and semantic alignment in traditional textile synthesis [18], [19], [20]. Within the Malaysian context specifically, a LGAN approach has been proposed to automate batik motif generation as a means of revitalising the industry and creating economic opportunities for the B40 (bottom 40% household income group) community [21]. However, existing studies evaluate these models independently, often relying on heterogeneous datasets and inconsistent metrics including Fréchet Inception

This manuscript is submitted on February 18, 2026, revised on September 18, 2026, accepted on September 30, 2026, and published on October 31, 2026.

Darni Darmin, Ihsan Mohd Yassin and Fairul Nazmie Osman are from the Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia.

Azlee Zabidi is from the Faculty of Computing, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan, Pahang, Malaysia.

Nooritawati Md Tahir is from the Faculty of Electrical Engineering and the Institute for Big Data Analytics and Artificial Intelligence (IBDAAI), Universiti Teknologi MARA, Shah Alam, Selangor, Malaysia. She also with the Department of Cybersecurity International Information Technology University, Almaty, Kazakhstan.

*Corresponding author
Email address: ihsan_yassin@uitm.edu.my

1985-5389/© 2026 The Authors. Published by UiTM Press. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Distance (FID), Inception Score (IS), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), Kernel Inception Distance (KID), Learned Perceptual Image Patch Similarity (LPIPS) and Contrastive Language-Image Pretraining (CLIP) score measures [22], [23], [24], [25]. The absence of a unified analytical framework limits objective benchmarking and obscures architecture-specific suitability for culturally grounded, resource-aware batik innovation, particularly within the Malaysian context.

Despite growing research on AI-driven textile synthesis, current works emphasise either adversarial optimisation and computational efficiency improvements [10], [16] or diffusion-based stability and semantic controllability enhancements [15], [26], with no study systematically comparing both architectures for batik design generation. This fragmentation hinders evidence-based decisions regarding the most appropriate generative architecture for sustainable, scalable batik development in resource-constrained environments.

To address this gap, this paper presents a comparative review of LGANs and Stable Diffusion models for batik design generation, identifying key research gaps and future directions to guide architecture selection in this domain.

The main contributions are: (1) a comparative synthesis of the architectural, computational, and generative characteristics of LGANs and Stable Diffusion in the context of batik design; (2) an integrated evaluation perspective encompassing computational feasibility, generative fidelity, semantic alignment, and cultural authenticity alongside common quantitative metrics; and (3) a domain-specific analytical perspective that links model characteristics to the practical requirements of batik design generation, including deployment feasibility, stylistic control, motif fidelity, and cultural authenticity.

The remainder of this paper is organised as follows. Section II presents the theoretical foundations of generative image modelling. Section III reviews LGAN architectures and parameter reduction strategies, the Stable Diffusion framework and its optimisation mechanisms, and compares both architectures in batik design generation from architectural, computational, and evaluation perspectives. Section IV discusses research gaps and future research directions. Finally, Section V concludes the paper with key findings and recommendations for advancing generative AI in culturally grounded textile design.

II. Theoretical Foundation of Generative Image Modelling

Generative models for image synthesis are commonly categorised according to their underlying learning paradigms and architectures, with GANs, Variational Autoencoders (VAEs), and diffusion-based models forming the core classes [27], [28], [29]. GAN-based approaches employ an adversarial learning framework in which a generator and a discriminator are jointly optimised to produce visually realistic images [8], [9], [30], [31], [32], [33], [34]. VAEs represent another category of generative models and are often utilised for data augmentation and representation learning [35], [36], [37]. More recently, diffusion models have emerged as a distinct and increasingly dominant category, formulating image synthesis as an iterative denoising process that enables stable training and high-fidelity generation [13], [27], [28], [38]. Fig. 1 illustrates the categorisation of generative models.

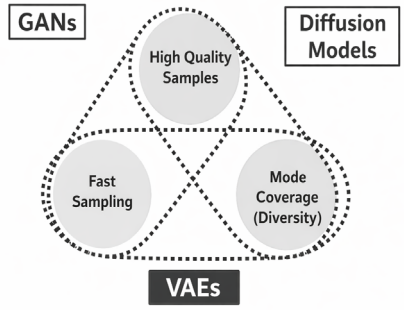


Fig. 1. Categorisation of Generative Models.

GAN-based and diffusion-based image generation methods are primarily distinguished by their learning mechanisms, training stability, and computational characteristics. GANs employ an adversarial framework, enabling efficient synthesis of high-resolution and visually realistic images for applications such as image translation, data augmentation, and medical imaging [8], [9], [30], [31], [32], [33], [34]. However, adversarial training is often associated with instability issues, including mode collapse and non-convergence, which may limit output diversity, although architectural improvements [9], [28], [39].

In contrast, diffusion models generate images through an iterative denoising process, resulting in more stable training dynamics and consistently higher image fidelity and diversity, particularly in high-quality and text-conditioned image synthesis tasks [28], [32], [40], [41], [42], [43], [44]. These benefits come at the expense of increased computational complexity and longer inference times due to multi-step sampling procedures [28], [32], [41], [42].

Consequently, existing studies characterise GAN-based approaches as efficiency-oriented and suitable for fast image generation, whereas diffusion-based methods prioritise stability and perceptual quality, reflecting a fundamental trade-off that underpins their complementary roles in contemporary image synthesis research [28], [32], [45].

III. Comparison of Lightweight Generative Adversarial Networks (LGANs) and Stable Diffusion Model

A. Lightweight Generative Adversarial Networks (LGANs)

LGANs are a subset of GANs designed to address the high computational demands of traditional GANs, making them suitable for deployment on mobile and edge devices, while maintaining the quality of the generated data [12], [46], [47]. This is typically achieved through structured parameter reduction and architectural simplification strategies, including

model pruning to remove redundant network parameters [48], the adoption of depthwise separable convolutions to decrease computational operations [46], [47], and ternary weight representations, a form of weight quantisation that restricts parameters to {-1, 0, +1} in order to simplify calculations and reduce memory usage [49]. Table 1 summarises the parameter-reduction strategies for LGAN models.

Compared to traditional GANs that are often deployed in high-performance computing environments for high-resolution image synthesis, lightweight variants are optimised for computational efficiency and low memory consumption, enabling real-time applications and resource-constrained deployment scenarios [12], [46], [47], [48]. Representative implementations include B-DCGAN and Ternary GAN, which employ restricted weight representations to enhance efficiency [49], LSRGAN, which integrates depthwise separable convolutions and optimised feature modules [47], and MobileARF-GAN, which combines efficient discriminator design with high-detail image generation for mobile environments [50].

TABLE I. PARAMETER REDUCTION STRATEGIES IN LGAN

Strategy	Function	Models
Model Pruning	Removes redundant weights or filters to reduce parameter count and computational load	PPCD-GAN
Depthwise Separable Convolution	Replaces standard convolution with factorized operations to reduce multiply-accumulate (MAC) operations	LSRGAN
Ternary Weight Representation	Restricts weights to {-1, 0, +1} to reduce memory usage and simplify arithmetic operations	B-DCGAN; Ternary GAN architectures

These approaches demonstrate that LGAN architectures can significantly reduce computational requirements without severely compromising generative performance, thereby extending GAN applicability to embedded and mobile platforms [12], [46], [47], [48], [50].

B. Stable Diffusion Model

Stable Diffusion is grounded in the established diffusion model framework, which is known for stable optimisation dynamics and the ability to produce high-quality outputs [51]. It has been widely used for text-to-image generation due to its flexibility, controllability, and adaptability across diverse application domains. The architecture comprises several core components that jointly enable high-quality text-to-image generation. First, the text encoder processes input prompts and converts them into latent textual embeddings, ensuring that semantic information is effectively represented for conditioning the generative process [52]. At the core of the architecture lies the latent diffusion model, which operates in a compressed latent space rather than pixel space. This component models the data distribution through a forward process that incrementally adds noise, a reverse denoising process that reconstructs structured representations, and an iterative sampling procedure used to generate new samples from learned distributions [26]. The resulting latent representations are then passed to an autoencoder decoder, which reconstructs them into full-resolution images aligned with the textual input [52]. Fig. 2 illustrates the core components in Stable Diffusion.

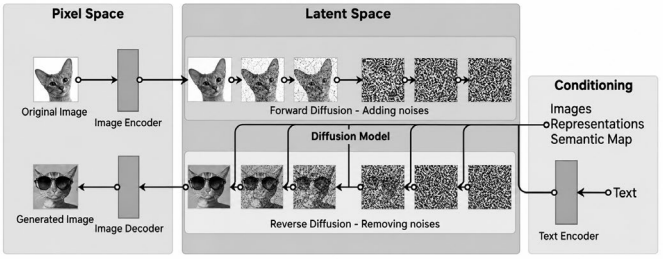


Fig. 2. Core components in Stable Diffusion.

Architecturally, the denoising backbone typically adopts a U-Net structure, which facilitates effective noise removal and preserves fine-grained details through skip connections, thereby enhancing image fidelity [14]. To improve computational efficiency and scalability, various optimisation strategies have been proposed, including model distillation, feature inheritance mechanisms, and architectural search techniques that reduce redundancy while maintaining performance [15], [53]. In addition, Stable Diffusion can be extended by integrating complementary modules such as LoRA, which enables efficient style-specific fine-tuning, and ControlNet, which enhances structural guidance and spatial consistency during generation [22], [54]. These architectural components and enhancement strategies together define the functional framework that enables Stable Diffusion to generate semantically coherent and visually detailed images from textual prompts [22], [52].

Beyond standard prompt-based generation, various fine-tuning and conditioning strategies have been proposed to enhance semantic alignment and output specificity, including techniques such as DreamBooth and textual inversion [23] and have also been shown to implicitly support image-to-prompt conversion through latent alignment with vision-language models, facilitating image editing, variation, and prompt refinement with minimal additional supervision [55]. In addition, discriminative adaptations of Stable Diffusion have been explored for few-shot image-text matching, leveraging cross-attention mechanisms to capture fine-grained interactions between visual and textual representations [56]. Application-driven studies further highlight its effectiveness in specialised domains, including the synthesis of clinically relevant medical images to address data scarcity [57], the generation of context-aware illustrations for literary works [58], and personalised face image generation that balances identity preservation with text-guided editability [18].

C. LGANs and Stable Diffusion Comparisons in Batik Design Generation

GANs are recognized for their ability to replicate and enhance traditional batik designs, addressing challenges in

manual preservation. LGANs make it feasible to deploy these generative tools in local workshops, supporting both cultural preservation and creative innovation [59], [60].

Stable Diffusion models enable the generation of high-quality, diverse, and innovative images from text prompts, which can be directly applied to batik pattern and textile design. This supports rapid prototyping and creative exploration, allowing designers to experiment with new motifs and styles efficiently [61], [62].

In exploring the comparison between LGANs and Stable Diffusion for batik design generation, relevant GAN-based studies in the batik domain are also considered to provide contextual support, particularly given the limited number of studies specifically focusing on LGANs. The comparison examines several key aspects, including architectural differences, training methodologies, computational efficiency, implications for artists and designers, and evaluation metrics used or considered for assessing the quality of the generated designs.

1) Architectural Differences

LGANs models are designed to be more efficient than traditional GANs, which consist of a generator and a discriminator that compete against each other to produce realistic outputs. The aim is to reduce computational demands while maintaining quality, often through architectural modifications and optimisations [12], [63]. They face challenges like mode collapse and training instability, which can affect the diversity of generated designs [64], [65].

Stable Diffusion operates differently by using a diffusion process to generate images. It has shown greater performance in generating high-quality images without the adversarial training that characterises GANs. The architecture of Stable Diffusion allows for more stable training and better handling of complex patterns, making it particularly effective for intricate designs like batik [4], [5].

2) Training Methodologies

LGANs training process involves adversarial training, where the generator and discriminator improve iteratively. However, this can lead to instability and require careful tuning of hyperparameters [64]. The training methodologies often focus on enhancing stability and reducing convergence issues, which can impact the quality and diversity of outputs [12], [63].

Stable Diffusion model employs a different training approach that does not rely on adversarial methods. Instead, it uses a denoising process to generate images, which has been shown to yield more consistent and high-fidelity results [4], [5]. The training methodology allows for greater flexibility in generating diverse batik motifs, as it can leverage large datasets effectively.

3) Computational Efficiency and Resource Requirements

Although LGANs are designed to reduce the computational burden of conventional GANs, their application to batik design generation can still require substantial computational resources. A batik-specific LGAN study trained the model for 200,000 epochs using an NVIDIA RTX 3090 Ti and reported approximately 168 hours of computing time [21]. The study generated 15,000 synthetic images, of which 13,011 (86.74%) were classified as batik or batik-like patterns, while 1,989 images were rejected due to unclear details or artefacts [21]. These findings indicate that lightweight architectures can reduce model complexity while still requiring considerable computational time for batik image generation. In comparison, a Stable Diffusion-based batik study demonstrated the generation of 512 × 512 images using 30 sampling steps and a batch size of 8 on an NVIDIA RTX 3060 Laptop GPU with 16 GB RAM [5].

However, the reported studies employed different hardware, datasets, model configurations, and experimental settings; therefore, their results do not establish a direct computational-efficiency advantage of either approach. A controlled comparison using consistent hardware, image resolution, and evaluation settings is therefore necessary to determine the relative computational efficiency of LGANs and Stable Diffusion for batik design generation.

4) Implications for Artists and Designers

The use of LGANs can enhance creativity by enabling artists to explore new design possibilities, but the challenges of training stability and output diversity may limit their effectiveness [64], [65].

On the other hand, Stable Diffusion offers a more reliable platform for generating intricate batik designs, enabling artists to push the boundaries of traditional motifs while ensuring cultural authenticity [5]. The integration of Stable Diffusion into the batik design process can streamline workflows, particularly during the ideation stage, by efficiently generating a variety of design options [5]. This can lead to increased productivity and innovation within the batik industry, allowing designers to focus more on creative aspects rather than technical challenges.

5) Evaluations Metrics

In batik design generation, evaluation primarily relies on quantitative metrics that assess image quality, diversity, and distributional similarity. The FID is among the most widely adopted measures, quantifying the distance between feature distributions of real and generated images using a pre-trained Inception network, where lower values indicate better generative fidelity and diversity [24], [66], [67]. It has been used for quantitative evaluation in batik motif generation, where a reduction in FID was associated with improved visual fidelity of the generated motifs [68]. It is also widely used for assessing generative models, including Stable Diffusion [69]. The IS is also commonly employed to evaluate how confidently generated images can be classified into distinct categories, with higher scores reflecting improved realism and diversity, although it remains sensitive to sample size and dataset characteristics [24], [66], [70]. PSNR and SSIM assess image similarity at the pixel and structural levels, respectively, with higher PSNR and SSIM values reflecting better preservation of texture, motif boundaries, and structural consistency in

generated batik patterns [19], [68], [71]. Both metrics have also been applied for image quality assessment in batik crack pattern generation [60].

KID is similar to FID but uses a different statistical approach to measure the similarity between real and generated image distributions, with lower KID scores indicating greater distributional similarity [4]. In addition, precision measures the proportion of generated images that are realistic, and recall measures the proportion of real images that are captured by the generator, with higher precision and recall values indicating better performance. While non-redundancy measures evaluate output diversity, ensuring that the model does not suffer from limited motif variation or mode collapse [4]. LPIPS assesses perceptual similarity between generated and reference images, with lower values indicating closer visual correspondence [24].

For Stable Diffusion-based batik generation, evaluation integrates perceptual, semantic, and alignment-oriented measures. The CLIP-score measures image-text alignment by quantifying the correspondence between generated motifs and their descriptive prompts, with higher values indicating better cross-modal consistency [25], [54]. In Bakaran batik pattern analysis, it was used alongside the IS to compare the generated results of different Stable Diffusion versions [72]. The Semantic Visual Consistency Score (SVCS) and User-Centric Image Coherence Metric (UCICM) were proposed to evaluate the semantic and visual coherence of Stable Diffusion-generated images with respect to textual prompts and user requirements [73]. They may also be applicable to batik image generation for assessing prompt-image semantic consistency and visual alignment with reference batik designs. Table 2 summarizes the evaluation metrics used or considered for

LGAN and Stable Diffusion-based batik image generation.

Human evaluation in both architectures was also incorporated to capture aesthetic judgment, cultural authenticity, and perceived design quality factors that automated metrics alone may not fully represent [19], [74]. These quantitative and qualitative metrics provide a comprehensive evaluation framework for assessing fidelity, diversity, semantic alignment, and cultural relevance in batik design generation [4], [19], [25], [54], [68], [71], [73], [74], [75].

Several studies have explored LGAN-based approaches for batik generation. For example, the LWGAN-GP model demonstrated the potential to generate diverse and stable batik motifs with reduced training times [16]. Additionally, advancements in GAN architectures, such as DefGAN-Im and PT-CLGAN, further illustrate ongoing efforts to enhance image generation capabilities [17], [76]. LGANs are also capable of producing distinctive and complex batik designs with greater efficiency and accuracy but also demonstrate the potential of this technology to attract young artists and provide sustainable income opportunities for the B40 community [21].

Research has utilized different versions of the Stable Diffusion algorithm (Versions 1.5, 2.0, and SDXL) to generate Bakaran batik motifs. Version 1.5 and SDXL were found to be the most effective, with SDXL pushing the boundaries of traditional design and Version 1.5 preserving cultural authenticity [5]. This demonstrates the versatility of Stable Diffusion in both maintaining traditional patterns and innovating new designs. Stable Diffusion v1.5, combined with LoRA and Latent Consistency Model (LCM), has been used to generate Miao batik patterns. This approach effectively transfers the unique structural and textural features of Miao batik, improving style fidelity and generation efficiency [19]. This method overcomes the limitations of existing generative models in capturing complex semantic and stylistic attributes.

Overall, the comparison between LGANs and Stable Diffusion models in batik design generation demonstrates that both architectures offer distinct strengths shaped by their underlying generative mechanisms. LGANs prioritise computational efficiency and deployment feasibility through parameter reduction and adversarial optimization, making them suitable for resource-constrained environments, while Stable Diffusion provides stable training dynamics, strong semantic controllability, and greater capability in generating intricate, high-resolution batik motifs.

TABLE II. THE EVALUATION METRICS USED FOR LGANS AND STABLE DIFFUSION

Metrics	LGANs	Stable Diffusion
FID	Yes	Yes
IS	Yes	Yes
PSNR	Yes	Yes
SSIM	Yes	Yes
KID	Yes	No
Precision and Recall	Yes	No
Non-redundancy	Yes	No
LPIPS	No	Yes
SVCS	No	Yes
UCICM	No	Yes
CLIP-score	No	Yes

IV. RESEARCH GAPS AND FUTURE DIRECTIONS

Several research gaps remain in the application of LGANs and Stable Diffusion models for batik design generation, particularly in terms of training stability, cultural fidelity, and dataset adequacy. Although LWGAN-GP demonstrates improved diversity and stability compared to conventional GAN architectures, challenges related to training efficiency, convergence instability, and mode collapse persist, thereby constraining the consistent generation of diverse and high-quality batik motifs [4], [16], [77]. In parallel, while Stable Diffusion has shown strong capability in producing batik patterns, limitations remain in accurately inheriting stylistic attributes particularly in colour composition and fine-grained detailing, and performance variability across different model versions has been observed, with certain configurations underperforming in batik-specific tasks [5], [78].

Future research should therefore prioritise enhanced training methodologies and culturally informed model adaptation. For LGANs, this includes improving training stability and computational efficiency through refined optimization strategies and potentially hybrid architectural exploration [4], [16]. For Stable Diffusion, more effective fine-tuning

mechanisms, such as LoRA and domain-specific conditioning should be further investigated to improve stylistic precision and cultural authenticity in batik generation[5], [20]. In addition, integrating curated cultural datasets, incorporating user-centric design interfaces, and enabling real-time feedback mechanisms can strengthen practical usability and creative control for designers [20], [79]. Cross-disciplinary collaboration among AI researchers, cultural scholars, and batik artisans is also essential to ensure that technological advancements align with artistic integrity and heritage preservation objectives [4] Addressing these gaps and pursuing these directions will advance both the technical robustness and cultural relevance of generative models in batik design synthesis.

V. CONCLUSION

This review has comparatively examined LGANs and Stable Diffusion models for batik design generation from architectural, computational, and evaluative perspectives. The analysis confirms that LGANs are optimised for parameter efficiency and deployment feasibility through pruning, quantisation, and architectural simplification, making them suitable for resource-constrained and mobile environments. However, their adversarial nature introduces instability risks that may affect motif diversity and the consistency of convergence. In contrast, Stable Diffusion models provide more stable training dynamics, stronger semantic controllability, and greater capability in synthesising intricate, high-resolution batik motifs through latent denoising and cross-attention conditioning. The comparison underscores a clear efficiency-stability trade-off that shapes their respective applicability in culturally grounded textile synthesis.

Importantly, this study highlights that performance assessment in batik generation must extend beyond common quantitative indicators such as FID, IS, KID, PSNR, SSIM, LPIPS, and CLIP-score. While these metrics measure distributional similarity and perceptual quality, they do not fully capture symbolic coherence, regional stylistic identity, or cultural authenticity embedded in traditional batik patterns. Therefore, architecture selection should be guided by a multidimensional framework integrating computational feasibility, generative fidelity, semantic alignment, and human-centred cultural evaluation. Such an integrated perspective is essential for ensuring that AI-driven batik innovation remains technically robust and culturally responsible.

Future work should prioritise three strategic directions: (1) development of hybrid or optimised frameworks that balance lightweight efficiency with diffusion-based stability for scalable deployment; (2) establishment of standardised, culturally annotated batik datasets to enable reproducible benchmarking and stylistic precision; and (3) incorporation of culturally informed conditioning mechanisms and human-in-the-loop refinement systems to strengthen authenticity and creative control. Advancing these directions will position generative AI not merely as an automation tool, but as an enabling technology that supports sustainable heritage preservation and digital transformation within the batik industry.

REFERENCES

[1] R. G. Tiwari, A. K. Agarwal, V. Jain, and A. Kumar, "Batik Classification in Indonesia: Exploring its Significance on Tourism and Economy," in *Proc. 2023 Int. Conf. Sustaining Heritage: Innov. Digital Approaches, ICSH 2023*, 2023, pp. 119–124, doi: 10.1109/ICSH57060.2023.10482828.

[2] N. H. M. Akhir, N. W. Ismail, S. R. Kaliappan, and R. Said, "Intersectoral linkages of Malaysian Batik Industry: An application of input-output analysis," *Jurnal Ekonomi Malaysia*, vol. 51, no. 1, pp. 57–68, 2017, doi: 10.17576/JEM-2017-5001-5.

[3] H. M. Razali, M. Ibrahim, M. Omar, and S. F. M. Hashim, "Current challenges of the batik industry in Malaysia and proposed solutions," in *AIP Conf. Proc.*, 2021, doi: 10.1063/5.0055651.

[4] O. Octadion, N. Yudistira, and D. Kurnianingtyas, "Synthesis of batik motifs using a diffusion-generative adversarial network," *Multimed. Tools Appl.*, vol. 84, no. 7, pp. 3407–3438, 2025, doi: 10.1007/s11042-025-20620-9.

[5] P. Ardhianto, Y. P. Santosa, and G. Anandhita, "Enhancing the Traditional Batik Design Practices: An Approach to Batik Motif Design Using Artificial Intelligence," *IAENG Int. J. Comput. Sci.*, vol. 52, no. 2, pp. 317–324, 2025, Accessed: Sep. 25, 2026. [Online]. Available: https://www.iaeng.org/IJCS/issues_v52/issue_2/IJCS_52_2_05.pdf

[6] A. Kumar *et al.*, "Navigating the Realm of Generative Models: GANs, Diffusion, Limitations, and Future Prospects—A Review," in *Lect. Notes Netw. Syst.*, 2024, pp. 301–319, doi: 10.1007/978-981-97-2550-2_23.

[7] Z. Shen, M. Mao, and P. Fan, "A Primary Comparison of Diffusion Models and Generative Adversarial Networks for Image Synthesis," in *ACM Int. Conf. Proc. Ser.*, 2024, pp. 225–234, doi: 10.1145/3696271.3696307.

[8] W. Chi, Y. H. Choo, and O. S. Goh, "Review of Generative Adversarial Networks in Image Generation," *J. Adv. Comput. Intell. Intell. Informat.*, vol. 26, no. 1, pp. 3–7, 2022, doi: 10.20965/jaciii.2022.p0003.

[9] J. Biju, R. Jeevitha, T. Richard, R. Chowdhury, P. Kumar, and K. Radhika, "Evaluating Generative Adversarial Networks Performance in Image Synthesis with Graphical Analysis of Loss Functions and Quality Metrics," in *Proc. 2nd IEEE Int. Conf. Integrated Intell. Commun. Syst., ICIICS 2024*, 2024, doi: 10.1109/ICIICS63763.2024.10859959.

[10] Y. Dogan and H. Yalim Keles, "Stability and diversity in generative adversarial networks," in *Proc. 27th Signal Process. Commun. Appl. Conf., SIU 2019*, 2019, doi: 10.1109/SIU.2019.8806356.

[11] A. Yadav, B. S. Avinash, and B. Kumar, "Lightweight Generative Adversarial Network (LGAN) Architectures for Resource-Constrained Hardware Applications," in *Proc. INDISCON 2024 5th IEEE India Council Int. Subsections Conf.: Sci., Technol. Soc.*, 2024, doi: 10.1109/INDISCON62179.2024.10744337.

[12] M. S. Alexiou and J. S. Mertoguno, "A Survey on Recent Advancements in Lightweight Generative Adversarial Networks, their Applications and Datasets," in *Proc. Int. Conf. Tools with Artif. Intell., ICTAI*, 2023, pp. 269–278, doi: 10.1109/ICTAI59109.2023.00047.

[13] Y. Huang *et al.*, "Diffusion Model-Based Image Editing: A Survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 47, no. 6, pp. 4409–4437, 2025, doi: 10.1109/TPAMI.2025.3541625.

[14] C. Si, Z. Huang, Y. Jiang, and Z. Liu, "Boosting Diffusion U-Net with Free Lunch for Text-to-Image and Text-to-Video Generation," in *Adv. Comput. Vis. Pattern Recognit.*, vol. Part F886, 2026, pp. 315–339, doi: 10.1007/978-3-031-94969-2_14.

[15] J. Zhu, Y. Wang, S. Pan, P. Wan, D. Zhang, and G. Huang, "A-SDM: Accelerating Stable Diffusion Through Model Assembly and Feature Inheritance Strategies," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 36, no. 10, pp. 18478–18491, 2025, doi: 10.1109/TNNLS.2025.3587724.

[16] A. Mujhid, A. Wibowo, and R. R. Siswanto, "Light Wasserstein Generative Adversarial Network with Gradient Penalty (LWGAN-GP) for Batik Nitik Synthetic," in *Proc. 7th Int. Semin. Res. Inf. Technol. Intell. Syst. (ISRITI)*, 2024, pp. 132–136, doi: 10.1109/ISRITI64779.2024.10963461.

[17] X. Nan, Y. Zhu, J. Fan, W. Li, Y. Wang, and X. Jiang, "DefGAN-Im: Adversarial Industrial Defect Synthesis with Conditional Feature Disentanglement on Imbalanced Datasets," in *Lect. Notes Comput. Sci.*, 2025, pp. 137–148, doi: 10.1007/978-981-96-9891-2_12.

[18] J. Ming and S. Qiu, "CoFaDiff: Coordinating Identity Fidelity and Text Consistency in Diffusion-Based Face Generation," *Appl. Sci. (Switzerland)*, vol. 16, no. 1, 2026, doi: 10.3390/app16010414.

[19] Q. Hu, Y. Peng, J. Xu, Z. Shao, Z. Tian, and J. Chen, "Adaptive Stylized Image Generation for Traditional Miao Batik Using Style-Conditioned LCM-LoRA Enhanced Diffusion Models," *Mathematics*, vol. 13, no. 12, 2025, doi: 10.3390/math13121947.

[20] T. Ma, J. Zhang, and Y. Jiang, "Innovative Design of Miao Ethnic Batik Patterns Based on Stable Diffusion," in *Proc. 17th Textile Bioeng. Informat. Symp. (TBIS), Hung Hom, Hong Kong, 2024*, pp. 451–458, doi: 10.52202/076989-0055.

[21] M. A. Latiff *et al.*, "Harnessing the Power of Generative Adversarial Networks (GANs) for Novel Batik Designs: An Exploration of Lightweight GANs (LGANs) for Automatic Batik Design," Qeios, May 2024, doi: 10.32388/HHB1CA.

[22] Y. Zhao, J. Zhang, D. Ding, and J. Liu, "Research on the Construction of an AI Platform for Haipai New Year Paintings Based on the Stable Diffusion Model," *Int. J. Gaming Comput. Mediat. Simul.*, vol. 17, no. 1, 2025, doi: 10.4018/IJGCMS.383051.

[23] T. Massaglia, B. Vacchetti, and T. Cerquitelli, "DreamShot: Teaching Cinema Shots to Latent Diffusion Models," in *CEUR Workshop Proc.*, 2024. Accessed: Sep. 25, 2026. [Online]. Available: https://iris.polito.it/handle/11583/2988712

[24] Ł. Tymoszuk and R. Kasprzyk, "Evaluating GAN Metrics for Image Generation: Challenges, Limitations, and Future Perspectives," in *Lect. Notes Netw. Syst.*, 2026, pp. 285–299, doi: 10.1007/978-3-032-14163-7_20.

[25] K. N. Sai, U. Wable, A. Singh, N. V. S. S. Koundinya, H. Gonuguntla, and C. Jain, "Harnessing Multimodal AI for Creative Design: Performance Evaluation of Stable Diffusion and DALL-E 3 in Fashion Apparel and Typography," in *Proc. Int. Conf. Elect., Comput., Energy Technol., ICECET 2024*, 2024, doi: 10.1109/ICECET61485.2024.10698398.

[26] Z. Chang, G. A. Koulieris, H. J. Chang, and H. P. H. Shum, "On the design fundamentals of diffusion models: A survey," *Pattern Recognit.*, vol. 169, 2026, doi: 10.1016/j.patcog.2025.111934.

[27] D. Dhillon, S. P. Yadav, and A. Varshney, "Generative Models for Image Synthesis," in *AI ML Techniques Image Process. Object Detection*, 2025, pp. 159–180, doi: 10.1007/978-981-96-7445-9_9.

[28] A. Kumar *et al.*, "Navigating the Realm of Generative Models: GANs, Diffusion, Limitations, and Future Prospects—A Review," in *Lect. Notes Netw. Syst.*, 2024, pp. 301–319, doi: 10.1007/978-981-97-2550-2_23.

[29] P. Gupta and S. Shukla, "Image Synthesis Using Machine Learning Techniques," in *Lect. Notes Data Eng. Commun. Technol.*, vol. 38, 2020, pp. 311–318, doi: 10.1007/978-3-030-34080-3_35.

[30] Q. Zhou, "Research on deep learning framework for multi scale information graph generation and visualization enhancement based on self attention generative Adversarial Network," *Discover Appl. Sci.*, vol. 7, no. 7, 2025, doi: 10.1007/s42452-025-07306-5.

[31] N. Kulkarni, S. Jagtap, H. Patil, and S. Deshpande, "Facial Image Generation using fully trained Generative Adversarial Network," in *Proc. 2024 8th Int. Conf. Comput., Commun., Control Autom., ICCUBEA 2024*, 2024, doi: 10.1109/ICCUBEA61740.2024.10775059.

[32] Z. Shen, M. Mao, and P. Fan, "A Primary Comparison of Diffusion Models and Generative Adversarial Networks for Image Synthesis," in *ACM Int. Conf. Proc. Ser.*, 2024, pp. 225–234, doi: 10.1145/3696271.3696307.

[33] P. Kumar, L. Raja, and A. Kumar, "Computational Analysis for Candidate X-ray Images Using Generative Adversarial Network," in *Lect. Notes Netw. Syst.*, 2023, pp. 853–867, doi: 10.1007/978-981-99-0981-0_64.

[34] C. Zhang, L. Lin, and Y.-W. Chen, "Deterministic Guided Progressive Medical Image Cross-Modal Generation Based on Deep Learning," in *Lect. Notes Comput. Sci.*, 2025, pp. 47–57, doi: 10.1007/978-981-96-2644-1_4.

[35] F. Aymen, A. Pester, and F. Andres, "Synthetic Water Crystal Image Generation Using VAE-GANs and Diffusion Models," in *Lect. Notes Netw. Syst.*, 2024, pp. 95–104, doi: 10.1007/978-3-031-54327-2_10.

[36] A. Jain, D. K. Roy, F. T. Ayasrah, P. William, and G. P. Lakshmi, "Learning Based on Generative AI with Image Synthesis and Data Augmentation Techniques," in *Proc. Int. Conf. Soc. Netw. Analysis, Manag. Secur., SNAMS*, 2024, pp. 41–49, doi: 10.1109/SNAMS64316.2024.10883800.

[37] P. Celard, E. L. Iglesias, J. M. Sorribes-Fdez, R. Romero, A. S. Vieira, and L. Borrajo, "A survey on deep learning applied to medical images: from simple artificial neural networks to generative models," *Neural Comput. Appl.*, vol. 35, no. 3, pp. 2291–2323, 2023, doi: 10.1007/s00521-022-07953-4.

[38] T. Khattar, S. Bare, S. Kuyate, and V. Wangikar, "A Comparative Study of Image Synthesis Models: Stack GANs and Diffusion Based Text to Image Generation," in *Lect. Notes Netw. Syst.*, 2026, pp. 30–44, doi: 10.1007/978-3-032-06253-6_3.

[39] H. Zhang, "Image Generation Method Based on Improved Generative Adversarial Network," *Recent Adv. Comput. Sci. Commun.*, vol. 16, no. 7, pp. 43–50, 2023, doi: 10.2174/2666255816666230330153428.

[40] C. Li, Y. Qi, Q. Zeng, and L. Lu, "Comparison of Image Generation methods based on Diffusion Models," in *Proc. 2023 4th Int. Conf. Comput. Vis., Image Deep Learn., CVIDL 2023*, 2023, pp. 1–4, doi: 10.1109/CVIDL58838.2023.10167336.

[41] Z. Yang, M. Astaraki, Ö. Smedby, and R. Moreno, "Efficient Generation of Synthetic Breast CT Slices By Combining Generative and Super-Resolution Models," in *Lect. Notes Comput. Sci. (including subseries Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, 2025, pp. 65–74, doi: 10.1007/978-3-031-77789-9_7.

[42] X. Tang, H. Zhu, and J. Tong, "Image generation method based on improved diffusion models," in *Proc. SPIE Int. Soc. Optical Eng.*, 2025, doi: 10.1117/12.3076349.

[43] Md. Y. Hossain, Md. M. H. Rakib, and I. R. Nijhum, "Usability of Pre-trained Diffusion Models in Generating Novel Datasets and Its Performance Evaluation," in *ACM Int. Conf. Proc. Ser.*, 2023, pp. 86–91, doi: 10.1145/3594409.3594430.

[44] D. Mehta, A. Mehta, and P. Narang, "LDFaceNet: Latent Diffusion-Based Network for High-Fidelity Deepfake Generation," in *Lect. Notes Comput. Sci. (including subseries Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, 2025, pp. 386–400, doi: 10.1007/978-3-031-78389-0_26.

[45] Z. Yu, J. Dai, Y. Zhang, J. Yang, and L. Luo, "SSAIM: Not All Self-Attentions Contain Effective Spatial Structure in Diffusion Models for Text-to-Image Editing," in *MM 2025 Proc. 33rd ACM Int. Conf. Multimedia, Co-Located with MM 2025*, 2025, pp. 9472–9480, doi: 10.1145/3746027.3754904.

[46] Z. Liu *et al.*, "Automatic intelligent recognition of pavement distresses with limited dataset using generative adversarial networks," *Autom. Constr.*, vol. 146, 2023, doi: 10.1016/j.autcon.2022.104674.

[47] N. Jiang, W. Zhao, H. Wang, H. Luo, Z. Chen, and J. Zhu, "Lightweight Super-Resolution Generative Adversarial Network for SAR Images," *Remote Sens. (Basel).*, vol. 16, no. 10, 2024, doi: 10.3390/rs16101788.

[48] S.-R. Su, S.-C. Hsia, and S.-K. Yang, "Old Video Quality Enhancement Technology Based on Generative Adversarial Networks," in *Proc. 11th IEEE Int. Conf. Consumer Electron. Taiwan, ICCE-Taiwan 2024*, 2024, pp. 425–426, doi: 10.1109/ICCE-Taiwan62264.2024.10674122.

[49] K. Nakamura and H. Nakahara, "Optimizations of Ternary Generative Adversarial Networks," in *Proc. Int. Symp. Multiple-Valued Logic*, 2022, pp. 158–163, doi: 10.1109/ISMVL52857.2022.00031.

[50] E. N. Banguis, J. J. C. Dela Cruz, R. M. Mollida, R. B. Nakazato, and T. V. Nakazato, "GAN-Based Hybrid Image Generation: ARF-GAN and MobileGAN Synthesis," in *Proc. 9th Int. Artif. Intell. Data Process. Symp., IDAP 2025*, 2025, doi: 10.1109/IDAP68205.2025.11222340.

[51] V. Amar and H. Shyan, "Text-to-Image Generator using GANs," in *Proc. Int. Conf. Integrated Intell. Commun. Syst., ICIICS 2023*, 2023, doi: 10.1109/ICIICS59993.2023.10421307.

[52] M. Liu and Y. Hu, "Application Potential of Stable Diffusion in Different Stages of Industrial Design," in *Lect. Notes Comput. Sci. (including subseries Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, 2023, pp. 590–609, doi: 10.1007/978-3-031-35891-3_37.

[53] H. Huang, X. Chang, and L. Yao, "Flexiffusion: Training-Free Segment-Wise Neural Architecture Search for Efficient Diffusion Models," in *CIKM 2025 Proc. 34th ACM Int. Conf. Inf. Knowledge Manag.*, 2025, pp. 949–959, doi: 10.1145/3746252.3761406.

[54] J. Park, S. Hong, and S. Choo, "A Generation Method and Evaluation of Architectural Facade Design Using Stable Diffusion with LoRA and ControlNet," *J. Architectural Inst. Korea*, vol. 41, no. 8, pp. 85–96, 2025, doi: 10.5659/JAIK.2025.41.8.85.

[55] Y. Ding, C. Tian, H. Ding, and L. Liu, "The CLIP Model is Secretly an Image-to-Prompt Converter," in *Adv. Neural Inf. Process. Syst.*, 2023, doi: 10.52202/075280-2456.

[56] X. He *et al.*, "Discffusion: Discriminative Diffusion Models as Few-shot Vision and Language Learners," *Trans. Mach. Learn. Res.*, vol. 2024, 2024, Accessed: Sep. 25, 2026. [Online]. Available: https://openreview.net/forum?id=GtnipgAomT

[57] A. Abaid, M. Ali Farooq, N. Hynes, P. Corcoran, and I. Ullah, "Synthesizing CTA Image Data for Type-B Aortic Dissection using Stable Diffusion Models," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., EMBS*, 2024, doi: 10.1109/EMBC53108.2024.10782969.

[58] F. R. P. Mahlow, A. F. Zanella, W. A. C. Castaneda, and R. A. Sarzi-Ribeiro, "Illustrating Classic Brazilian Books using a Text-To-Image Diffusion Model," *IEEE Latin America Trans.*, vol. 22, no. 12, pp. 1000–1008, 2024, doi: 10.1109/TLA.2024.10789626.

[59] A. E. Minarno, S. Riyadi, G. W. Wicaksono, and Y. Munarko, "Batik Pattern Generation with Enhanced GAN Architectures and Multi-Scale Discriminators," in *2025 IEEE 15th Int. Conf. Syst. Eng. Technol., ICSET 2025, Conf. Proc.*, 2025, pp. 132–137, doi: 10.1109/ICSET65917.2025.11284213.

[60] Y. Huang, J. Su, J. Wang, and S. Ji, "BatIK-DG: Improved deblurgan for batik crack pattern generation," in *Proc. IOP Conf. Ser.: Mater. Sci. Eng.*, 2020, doi: 10.1088/1757-899X/790/1/012034.

[61] A. Adityan, K. Vaiapury, T. Palanisamy, and S. K. Tangavel, "Fashion Design Transformation using Text-Guided Diffusion Models," in *4th Int. Conf. Sentiment Analysis Deep Learn., ICSADL 2025 Proc.*, 2025, pp. 1558–1563, doi: 10.1109/ICSADL65848.2025.10933465.

[62] G. R. Naidu, T. Rajamani, Y. V. Chowdhary, M. Abdulrahim, N. A. Munaf, and S. P. Padma Priya, "Prompt-Driven Acceleration of Text-to-Image Synthesis: A CUDA-Enhanced Stable Diffusion Framework," in *Proc. 2025 IEEE 1st Int. Conf. Recent Trends Comput. Smart Mobility, RCSM 2025*, 2025, doi: 10.1109/RCSM67767.2025.11507877.

[63] S. Senthilkumar, P. Kumaresan, K. Brindha, and Y.-C. Hu, "Advanced Progress in Optimized Generative Adversarial Network Applications Across Domains: A Comprehensive Survey," *Archives Comput. Methods Eng.*, 2025, doi: 10.1007/s11831-025-10361-1.

[64] S. Ardebili, M. Kargar, and N. S. Hoseini, "An analytical review of GANs: technical evolution, architectures, applications, datasets, and challenges," *Iran J. Comput. Sci.*, vol. 9, no. 1, 2026, doi: 10.1007/s42044-025-00369-y.

[65] R. Bhat and R. Nanjundegowda, "A Review on Comparative Analysis of Generative Adversarial Networks' Architectures and Applications," *J. Robotics Control (JRC)*, vol. 6, no. 1, pp. 53–64, 2025, doi: 10.18196/jrc.v6i1.24160.

[66] K. D. Kumar, S. Srang, and D. Valy, "Evaluating Text-to-Image GANs Performance: A Comparative Analysis of Evaluation Metrics," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 11, pp. 618–627, 2023, doi: 10.17762/ijritcc.v11i8s.7248.

[67] P. Kokate, A. D. Joshi, and P. S. Tamizharasan, "An Empirical Comparison of Generative Adversarial Network (GAN) Measures," in *Lect. Notes Elect. Eng.*, 2021, pp. 1383–1396, doi: 10.1007/978-981-15-5341-7_105.

[68] A. E. Minarno, I. Soesanti, and H. A. Nugroho, "Batik GAN for Generating Motif Synthesis using Multi-Discriminator and Self-Attention," in *Proc. Int. Conf. Innov. Artif. Intell. Internet Things, AIIT 2025, 2025*, doi: 10.1109/AIIT63112.2025.11082971.

[69] S. Jamal, H. Wimmer, and C. M. Rebman, "Perception and evaluation of text-to-image generative AI models: a comparative study of DALL-E, Google Imagen, GROK, and Stable diffusion," *Issues Inf. Syst.*, vol. 25, no. 2, pp. 277–292, 2024, doi: 10.48009/2_iis_2024_123.

[70] A. Shah, N. Rakesh, M. Gulhane, P. Agrawal, M. Kaur, and S. Dixit, "Evaluating Generative Models: A Review of Challenges and Opportunities in Assessment of Generative Adversarial Networks," in *Proc. 2025 2nd Int. Conf. Comput. Intell., Commun. Technol. Netw., CICTN 2025*, 2025, pp. 12–16, doi: 10.1109/CICTN64563.2025.10932502.

[71] A. E. Minarno, D. R. Akbi, and Y. Munarko, "Enhanced BatikGAN SL Model for High-Quality Batik Pattern Generation," *Int. J. Informat. Visualization*, vol. 9, no. 3, pp. 1142–1148, 2025, doi: 10.62527/joiv.9.3.3096.

[72] K. A. Septemedi and Y. Santosa, "Usability Analysis of Stable Diffusion-based Generative Model for Enriching Batik Bakaran Pattern Synthesis," *Jurnal Informatika Proxies*, vol. 7, no. 2, pp. 128–146, 2024, doi: 10.24167/proxies.v7i2.12472.

[73] R. Gabriel Joshua, K. Khadar Nawas, and S. R. Varsini, "Enhancing Text-to-Image Model Evaluation: SVCS and UCICM," in *Proc. 2023 6th Int. Conf. Recent Trends Advance Comput., ICRTAC 2023*, 2023, pp. 537–547, doi: 10.1109/ICRTAC59277.2023.10480770.

[74] W. Yang, C. Wang, L. Liu, S. Dong, and Y. Zhao, "Advancing Interior Design with AI: Controllable Stable Diffusion for Panoramic Image Generation," *Buildings*, vol. 15, no. 8, 2025, doi: 10.3390/buildings15081391.

[75] A. E. Minarno, T. D. Antoko, and Y. Azhar, "Generation of Batik Patterns Using Generative Adversarial Network with Content Loss Weighting," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 13, no. 1, pp. 348–356, 2023, doi: 10.18517/ijaseit.13.1.16201.

[76] H. Zhang, X. Zhao, J. Dong, X. Shi, S. Ge, and T. Wu, "PT-CLGAN: An intelligent garment generation method based on progressive training and contrastive learning strategy," *Textile Res. J.*, 2025, doi: 10.1177/00405175251400282.

[77] P. Ghasemi, C. Yuan, T. Marion, and M. Moghaddam, "DCG-GAN: Design concept generation with generative adversarial networks," *Design Sci.*, vol. 10, 2024, doi: 10.1017/dsj.2024.13.

[78] J. Zhang and Y. Jiang, "Style Transfer Technology of Batik Pattern Based on Deep Learning," *J. Fiber Bioengineering Informat.*, vol. 16, no. 1, pp. 57–67, 2023, doi: 10.3993/jfbim02171.

[79] X. Li, L. Zhao, and L. Wang, "Design of the Three Rabbit Caisson Pattern Based on the Stable Diffusion Model," in *Proc. 2024 5th Int. Conf. Intell. Design, ICID 2024*, 2024, pp. 276–280, doi: 10.1109/ICID64166.2024.11024561.