



Bridging Structure and Emotion: A Generative Framework for Accessible and Expressive Knit Design

Jiayin Fan, Li Li[†]

Division of Integrative Systems and Design, The Hong Kong University of Science and Technology, Hong Kong 999077, China

[†]*E-mail: lillyli@ust.hk*

Received: December 4, 2025 / Revised: January 22, 2026 / Accepted: January 29, 2026 / Published online: February 10, 2026

Abstract: Knit pattern design is challenging, requiring both technical skill and creative vision. We introduce an accessible generative framework using Stable Diffusion fine-tuned with LoRA and an interactive web interface. Our system utilizes a dual-label dataset (over 3000 images) annotated with both technical attributes (stitch type, complexity) and semantic/emotional descriptors (e.g., cozy, elegant) to ensure outputs are structurally coherent and stylistically diverse. The LoRA-tuned model exhibits strong domain specialization, producing clearer stitch definition and fewer non-knitting artifacts than the baseline. We develop an interactive platform featuring an emotion-technical balance slider. This framework positions generative AI as a creative partner that lowers technical barriers and expands the expressive possibilities of textile design.

Keywords: Knit Pattern Generation; Generative AI in Fashion; LoRA Fine-Tuning; Semantic Conditioning; AI-Driven Design Tool

<https://doi.org/10.64509/jdi.11.58>

1 Introduction

Knitting has long represented a unique intersection of craft, engineering, and design. Traditional knit pattern creation requires not only artistic vision but also deep technical expertise in stitch types, yarn properties, and structural constraints. While digital tools and online design communities have made sharing and documenting patterns more accessible, the process of designing original knitting patterns remains labor-intensive and heavily reliant on domain knowledge. As the demand for personalized and expressive textile design grows, there is an increasing need for computational systems that democratize the creative process and lower the technical barriers for designers.

Recent computational approaches have begun to address the unique challenges and multi-scale complexity of textile and knitting structures. Knitted textiles exhibit intricate hierarchical geometry [1], motivating the development of specialized representations for simulation and design. For example, Kapllani et al.'s TopoKnit introduced a process-oriented data structure that captured yarn-level topology for efficient simulation of weft-knitted fabrics [1]. Earlier explorations, such as Underwood's work on 3D knitted preforms, demonstrated how shaping knitted shells can enable lightweight

composite reinforcement applications [2]. Building on such foundations, several deep-learning based studies focused on automating the translation between knit patterns and manufacturing instructions. Kaspar et al. developed a neural inverse knitting system capable of generating machine instructions directly from a single product image [3], while deepKnit by Scheidt et al. used learning-based models to produce low-level knitting instructions from design specifications [4]. Complementary work by Sheng et al. proposed a vision-based robotic pipeline to reverse-engineer fabric structures [5], and Zheng et al. created a two-stage CNN approach for deriving knitting instructions from photographs [6]. Together, these works established the feasibility of computational knitting pipelines, though most emphasized technical correctness, manufacturability, and automation rather than creative pattern ideation.

In contrast, an emerging body of research in fashion and textile design explored how generative AI can enhance creativity rather than production workflows. Early generative models in fashion focused broadly on garments and high-level styling [7], but more recent efforts targeted textile-level pattern synthesis. Wu and Li demonstrated that a tailored GAN model can generate knitted textile patterns with creative and practical utility for design applications [7]. Similarly, Hu

[†] Corresponding author: Li Li

* Academic Editor: Qiaoling Zhang

© 2026 The authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

et al.'s MSEmbGAN significantly improved the realism of multi-stitch embroidery textures [8]. Karagoz et al. showed that fine-tuning a text-to-image diffusion model on textile pattern data yielded superior quality and coherence compared to generic diffusion models [9]. Beyond contemporary fashion prototypes, generative models have also supported cultural preservation through interactive systems that regenerated traditional embroidery motifs, such as Wang et al.'s work on Yi minority embroidery [10]. These developments indicated a growing trend in using generative models as creative partners in textile and fashion design.

Diffusion models combined with efficient fine-tuning techniques have become central in these workflows. Latent diffusion frameworks like Stable Diffusion can be adapted to specialized material domains through parameter-efficient strategies. Zhou et al., for example, fine-tuned Stable Diffusion XL on a custom material dataset and used reinforcement learning to enhance realism [11]. A major contributor to the accessibility of such domain adaptation is LoRA, which inserts lightweight trainable matrices into frozen model weights and has become widely adopted due to its efficiency [12]. Ongoing research continued to refine these adaptation methods, such as tensor-decomposition variants proposed by Tao et al. that achieved improved fidelity with even fewer trainable parameters [13]. Additional work explored attribute-level control: Chen et al.'s AttriCtrl demonstrated how mapping user-adjustable attribute intensities into latent space allowed diffusion models to modulate aesthetic or emotional properties without extensive retraining [14].

Another growing direction in generative design is the incorporation of semantic and emotional conditioning. Standard text prompts often struggled to express abstract concepts such as atmosphere or mood [15], motivating specialized conditioning techniques. Yang et al.'s EmoGen introduced an emotion embedding space and alignment strategy to help diffusion models generate emotionally faithful images [15]. Current design workflows rely heavily on mood boards - static collages that establish an aesthetic direction but remain disconnected from technical realization. This creates a gap where the "mood" is often lost when translating concepts into manufacturable structures. Our framework bridges this gap by treating emotion as a tunable parameter alongside technical constraints, differentiating our approach from generic attribute control methods by specifically enforcing the structural coherence required for knitting. Treating emotion as a first-class conditioning variable thus offered new opportunities for expressive textile and pattern design.

Parallel to these algorithmic advances was a shift toward human-centered interactive tools. Rather than fully automatic black-box generators, human-AI interaction research emphasized co-creation, transparency, and iterative control [16]. Real-time interactive systems such as Nimi et al.'s gesture-driven Ukiyo-e diffusion model illustrated how immediate feedback and intuitive controls enhance creative engagement [16]. Similarly, interactive textile design systems, such as the embroidery regeneration platform by Wang et al. [10], highlight the value of designer-in-the-loop workflows. These examples reinforced that generative AI tools are most effective when they empower designers rather than replace them.

Motivated by these developments, this study aims to expand generative knitting research toward design-level exploration rather than production automation. We developed a user-friendly knitting pattern generation system powered by diffusion models with LoRA fine-tuning, trained on a curated dataset of over 3000 knit images annotated with both technical attributes and semantic/emotional descriptors. Our system incorporated these descriptors into the conditioning process, enabling the generation of visually coherent patterns aligned not only with stitch structures but also with stylistic or emotional intent. To support iterative design exploration, we introduced a browser-based interactive interface that allows designers to adjust technical and emotional controls and preview pattern variations in real time. Through experiments and qualitative analysis, we demonstrated that our system can produce diverse, high-quality knit patterns shaped by both technical prompts and emotional guidance, ultimately positioning generative AI as a creative collaborator in textile design.

2 Methodology

The proposed framework aims to generate knit patterns that are both visually coherent and semantically expressive by adapting a latent diffusion model with low-rank adaptation (LoRA) fine-tuning. The methodology consisted of four main components: (1) dataset curation and preprocessing; (2) model architecture and fine-tuning strategy; (3) semantic and emotional conditioning; and (4) interactive user interface design. An overview is shown in [Figure 1](#).

2.1 Dataset Preparation and Preprocessing

To enable domain-specific adaptation of diffusion models, we curated a dataset of 3000+ knit images, each with a resolution of 256×256 pixels. To ensure technical accuracy and consistent image quality, the dataset was curated from an institutional research archive comprising digitized knit samples from past and current laboratory projects. Unlike unverified web-scraped data, this source ensured that all samples were captured under controlled lighting conditions with verifiable stitch information. The dataset represents a wide range of stitch structures, color palettes, and design complexities, covering both traditional motifs and modern patterns. Representative examples are shown in [Figure 2](#).

2.1.1 Image Preprocessing

All images were normalized to a fixed resolution of 256×256 and converted to RGB format. To enhance diversity and mitigate overfitting, standard augmentation techniques were applied, including random rotations, horizontal flips, brightness adjustments, and mild color jittering. This step ensured that the model learns robust representations of knitting textures across different contexts.

2.1.2 Labeling Scheme

Each image was paired with two categories of labels:

1. Technical attributes - stitch type (e.g., rib, cable, lace), pattern complexity (simple vs. intricate), and motif category (geometric, floral, figurative).

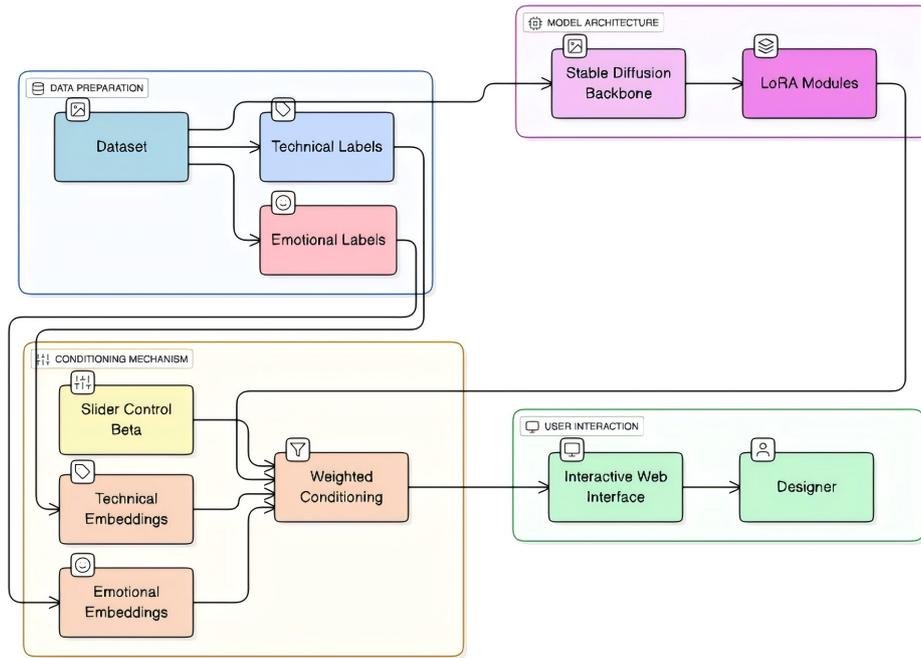


Figure 1: Framework Overview

**Representative Dataset Overview Figure
(36 Unique Patterns Displayed, Database Total: 3046)**

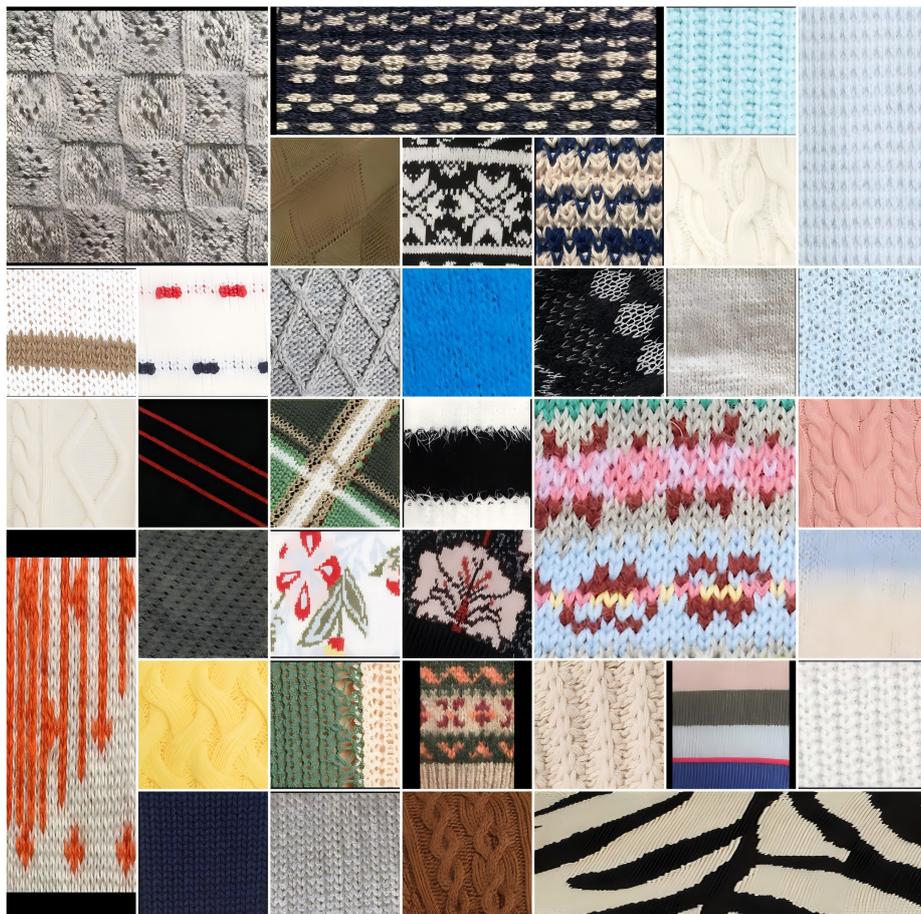


Figure 2: Dataset Examples

2. Semantic/emotional descriptors - mood-oriented tags such as cozy, elegant, playful, calm, energetic. These descriptors were selected based on their frequent use in textile and fashion discourse and validated through a pilot exercise involving three textile design researchers, achieving a consensus rate of 85%.

This dual-label structure enabled conditioning the model on both functional and affective aspects of knitting design, allowing outputs to be technically meaningful while also emotionally resonant.

2.2 Diffusion Model and LoRA Fine-Tuning

We employed Stable Diffusion (SD) as the generative backbone due to its proven effectiveness in producing high-quality, diverse images. SD operates as a latent diffusion model (LDM), where images are encoded into a lower-dimensional latent space using a variational autoencoder (VAE). The denoising process is trained to iteratively reconstruct images from noise by learning the reverse of a forward diffusion process.

Formally, the forward process at time step t is defined as:

$$q(z_t|z_0) = \mathcal{N}(\sqrt{\bar{\alpha}_t}z_0, (1 - \bar{\alpha}_t)I), \quad \bar{\alpha}_t = \prod_{s=1}^t \alpha_s, \quad (1)$$

where z_0 is the latent representation of the input image and α_t is a noise schedule. The denoising network ε_θ is optimized by minimizing:

$$\mathcal{L}_{\text{denoise}}(\theta) = \mathbb{E}_{z_0, t, \varepsilon} \|\varepsilon - \varepsilon_\theta(z_t, t, c)\|^2, \quad (2)$$

with conditioning input c representing text or label embeddings.

2.2.1 LoRA Fine-Tuning

To adapt the model efficiently to the knit dataset, we employed Low-Rank Adaptation (LoRA). LoRA introduces lightweight trainable matrices $A \in \mathbb{R}^{d \times r}$ and $B \in \mathbb{R}^{k \times r}$ such that:

$$W' = W + \Delta W, \quad \Delta W = \alpha BA, \quad r \ll \min(d, k), \quad (3)$$

where W is the frozen pre-trained weight matrix and α is a scaling factor. Only A and B are updated during fine-tuning, significantly reducing computational cost while preserving the representational power of the pre-trained model.

This strategy allows training on a relatively small dataset while maintaining generalization and avoiding catastrophic forgetting. In practice, LoRA was applied to attention and feedforward layers of the U-Net within the diffusion model, with a rank $r = 32$ and a scaling factor $\alpha = 32$. The model was based on Stable Diffusion v1.5. We used a standard noise scheduler with 1000 diffusion steps and sampled using the DPMSolver++ scheduler. An illustration of the LoRA fine-tuning procedure is shown in Figure 3.

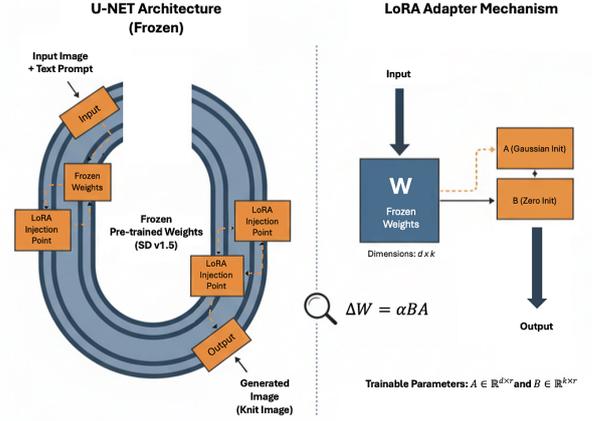


Figure 3: LoRA Fine-Tuning Illustration

2.3 Semantic and Emotional Conditioning

A core innovation of this framework is the integration of semantic/emotional conditioning alongside technical labels. While existing textile-focused generative models often emphasize manufacturability, we aim to broaden the expressive capacity of pattern generation by embedding *design mood* and *aesthetic intent*.

To integrate the condition vectors, we employ learnable linear projection matrices $W_t \in \mathbb{R}^{d \times d}$ and $W_e \in \mathbb{R}^{d \times d}$ to map the encoded text features into the shared latent space. The combined conditioning vector c is computed as:

$$c = W_t \phi_t(t) + \beta W_e \phi_e(e), \quad (4)$$

where t and e denote technical and emotional labels, ϕ_t, ϕ_e are text encoders, W_t, W_e are projection matrices, and β controls the relative influence of emotional descriptors. We utilize a linear combination to facilitate smooth interpolation, where β allows the user to continuously shift the generation bias from structural fidelity to emotional expression.

For sampling, we adopted classifier-free guidance (CFG) to balance unconditional and conditional predictions:

$$\varepsilon = (1 + w) \varepsilon_\theta(z_t, t, c) - w \varepsilon_\theta(z_t, t, \emptyset), \quad (5)$$

where w is the guidance scale. This mechanism sharpened alignment with conditioning inputs, ensuring generated patterns reflect both the technical and emotional aspects specified.

2.4 Interactive User Interface

To maximize accessibility for designers, we implemented a web-based interactive platform rather than an OS-specific application. The platform integrates model inference via a lightweight backend with a responsive frontend built in React.

The interactive interface was designed to support intuitive and iterative design exploration. It includes: (1) a prompt panel that allows designers to choose stitch types, motifs, and emotional descriptors through dropdown menus or natural-language input; (2) adjustable control sliders, such as the CFG scale w and emotion weight β , which offer clear handles for navigating technical–emotional trade-offs; (3) a preview grid

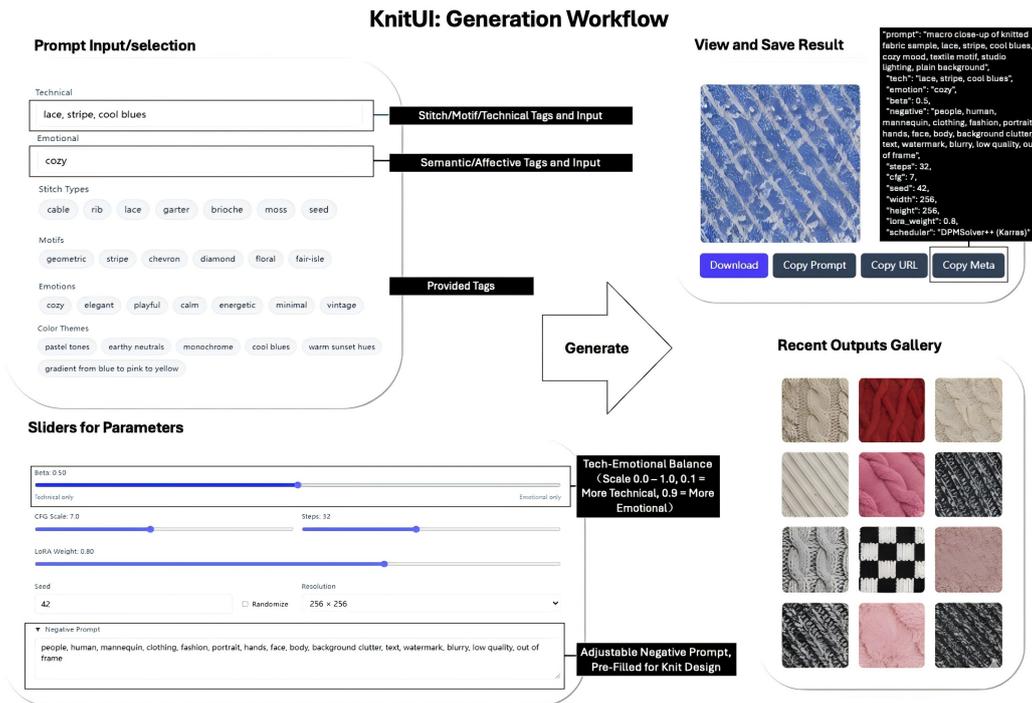


Figure 4: Interactive UI Screenshot

that displays real-time or batched generation results, enabling users to compare variations and refine their selections; and (4) export options that allow generated patterns to be downloaded as high-resolution PNG images, while associated generation parameters (including prompts, seeds, and model weights) are simultaneously exported as JSON metadata to facilitate reproducibility and future iteration.

This interface design emphasized ease of use and creative exploration, reducing the need for specialized technical knowledge and empowering designers to focus on aesthetics and intent. An example interface screenshot is shown in Figure 4.

3 Experiments

The experimental evaluation was designed to assess the effectiveness of the proposed framework in generating diverse, high-quality knit patterns that align with both technical attributes and semantic or emotional descriptors. We presented two categories of experiments: (1) baseline comparisons between the pre-trained Stable Diffusion model and our LoRA-fine-tuned version; and (2) ablation analysis of conditioning strategies.

3.1 Experimental Setup

3.1.1 Hardware and Training Environment

All experiments were conducted on a workstation equipped with an NVIDIA RTX 4090 GPU (24 GB VRAM), 256 GB RAM, and AMD EPYC 9555 64-Core Processor. Training and inference were implemented in PyTorch 2.1 using the diffusers library, with mixed precision (fp16) enabled for efficiency. The LoRA fine-tuning was trained for 15000 steps

with a batch size of 16, AdamW optimizer ($lr = 1e-4$, $\beta_1 = 0.9$, $\beta_2 = 0.999$), and cosine learning rate scheduling.

3.1.2 Dataset Splits

The curated dataset of 3000 knit images was divided into three subsets to support model training and evaluation: (1) an 80% training set consisting of 2400 images, (2) a 10% validation set containing 300 images, (3) a 10% test set with 300 images. Labels for stitch type, complexity, motif, and semantic or emotional descriptors were evenly distributed across all subsets to maintain balanced representation throughout the splits.

3.1.3 Baselines

Two baseline models were used for comparison: (1) the original Stable Diffusion v1.5 model without any fine-tuning, (2) our LoRA-tuned Stable Diffusion model trained on the curated knit dataset. The evaluation centered on differences in image quality, domain relevance, and the controllability of conditioning inputs.

3.2 Baseline Comparison

We first compared generation results between the original Stable Diffusion (v1.5) model and the LoRA-tuned version under identical prompts. The baseline model produced visually diverse outputs but frequently depicted non-knitting textures such as woven fabrics or printed surfaces. In contrast, the LoRA-tuned model generated outputs that exhibited recognizable knitted loops, coherent motifs, and textile-like structures, indicating effective domain adaptation.

Quantitative evaluation was conducted using the Fréchet Inception Distance (FID), Inception Score (IS), and Learned Perceptual Image Patch Similarity (LPIPS) on paired test

samples (Table 1). For clarity, \downarrow indicates that lower is better (FID, LPIPS), while \uparrow indicates that higher is better (IS). The LoRA-tuned model achieved an FID of ≈ 195.0 and an IS of 2.54 (± 0.43) compared with 3.17 (± 0.61) for the baseline. Although the IS value decreased and the FID remained high, reflecting reduced global diversity and color variation, the results indicated that the fine-tuned model specialized toward the knitted-pattern domain rather than general photographic imagery. This specialization yielded more structurally accurate and domain-coherent outputs, even if overall visual diversity is lower.

Qualitatively, this trade-off suggests that LoRA fine-tuning successfully redirected the generative model toward the intended design space, replacing generic variety with higher structural fidelity and realistic knit-texture formation.

Table 1: Quantitative Metrics

Model	FID (\downarrow)	IS (\uparrow)	LPIPS (\downarrow)
SD 1.5 Baseline	220.50	3.10	0.7405
LoRA-Tuned SD (Ours)	195.00	2.54	0.6970

3.3 Conditioning Ablation

To assess the impact of semantic and emotional conditioning, we conducted ablation experiments across three configurations: (1) technical-only conditioning that included stitch type, motif, and complexity; (2) emotion-only conditioning using descriptors such as cozy, elegant, or playful; and (3) hybrid conditioning that combined both types of inputs through a user-adjustable weighting parameter β .

Results showed that technical-only conditioning yielded structurally accurate patterns but limited stylistic variation, whereas emotion-only conditioning produced highly diverse patterns with reduced structural coherence. The hybrid approach balanced both, producing outputs that align with structural prompts while reflecting emotional tone (e.g., a “cozy cable knit” with softer textures and muted palettes, or an “energetic geometric knit” with bold contrasts and sharp motifs).

These findings confirmed that emotional labels enhance creative controllability without undermining technical correctness.

4 Results and Discussion

The results of our experiments demonstrated that the proposed framework successfully generated knit patterns that balance structural fidelity, aesthetic diversity, and emotional expressiveness. This section presents qualitative and quantitative results, followed by a discussion of broader implications for design practice.

4.1 Visual Comparisons

Figure 5 illustrates representative outputs under different experimental conditions. In baseline comparisons, the pre-trained Stable Diffusion model often produced images resembling woven fabrics, printed textures, or abstract shapes with limited stitch definition. By contrast, the LoRA-tuned model consistently generated outputs exhibiting recognizable knit

structures such as cables, lace motifs, and ribbing patterns. When conditioned only on technical attributes, the generated results closely adhered to the specified stitch types and motifs but lacked variation in mood and style. For example, prompts specifying “cable knit” yielded structurally correct outputs but with repetitive textures and limited stylistic diversity. Emotion-only conditioning produced highly varied visual appearances, with noticeable differences in color palette, contrast, and motif arrangement depending on descriptors such as cozy, playful, or elegant, though structural accuracy was less reliable.

The hybrid conditioning approach combined the strengths of both, yielding structurally coherent yet stylistically diverse patterns. For instance, a “cozy cable knit” output displayed muted color tones, soft contrasts, and uniform stitch spacing, while an “energetic geometric knit” displayed sharper motif edges and bold color contrasts, aligning with both technical and emotional intent.

4.2 Quantitative Outcomes

Quantitative evaluations complemented the qualitative observations. The LoRA-tuned model achieved a Fréchet Inception Distance (FID) of ≈ 195.0 and an Inception Score (IS) of 2.54 (± 0.43) compared with 3.17 (± 0.61) for the baseline Stable Diffusion model.

Although the absolute FID value remains high, reflecting the large stylistic gap between knitted-texture imagery and the photographic domain used by the Inception network, the change indicates that the LoRA model learned a distinct, domain-specific representation. The decrease in IS suggests reduced global color and compositional diversity, a typical effect of domain specialization.

Importantly, visual inspection confirms that this specialization yields more structurally consistent stitches and fewer non-knitting artifacts, aligning with the system’s design-level objective of realism and coherence rather than sheer variety.

4.3 Discussion of Design Implications

The experimental findings suggest several key implications for design practice. (1) Lowering technical barriers: by embedding stitch type, motif, and structural complexity directly into the conditioning pipeline, the system enabled non-experts to produce structurally coherent knit patterns without needing specialized textile-engineering vocabulary, thereby reducing entry barriers for students and hobbyists. (2) Expanding creative exploration: the inclusion of emotional descriptors introduced an additional dimension of control, allowing designers to generate variations aligned with mood or thematic intent. Although quantitative diversity decreased, the perceptual and affective range broadened, supporting exploratory workflows in which aesthetic expression is as central as structural accuracy. (3) User-centered interaction: the emotion-technical weighting mechanism illustrated how abstract model parameters can be translated into intuitive interface controls, helping bridge the gap between algorithmic complexity and creative decision-making. (4) Limitations: despite encouraging results, several constraints remain. As illustrated in Figure 6, generated samples occasionally exhibit structural incoherence, such as disconnected cable stitches or ‘melting’

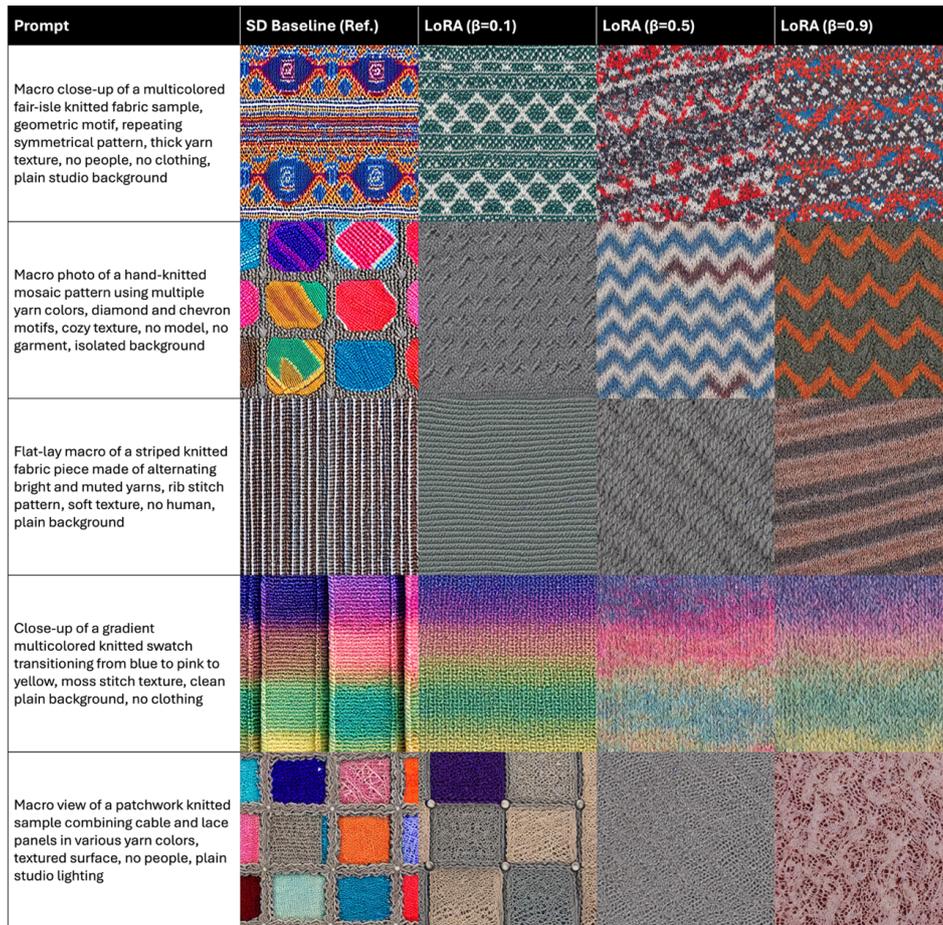


Figure 5: Visual Comparison (Baseline, LoRA, Ablation)

yarn topology (Figure 6(a)), and color bleeding in complex motifs (Figure 6(b)). Additionally, using high guidance scales can introduce unnatural artifacts or rigid, plastic-like textures (Figure 6(c)). Finally, emotional descriptors were inherently subjective and may vary across users, and manufacturability considerations were intentionally excluded. Specifically, the model generates raster-based pixel data, which lacks the vector-based needle instructions (e.g., racking, loop transfers, yarn carrier movements) required by industrial knitting machines (e.g., Stoll or Shima Seiki). Consequently, while the visual fidelity is high, the output requires manual transcription or further processing to be machine-ready.

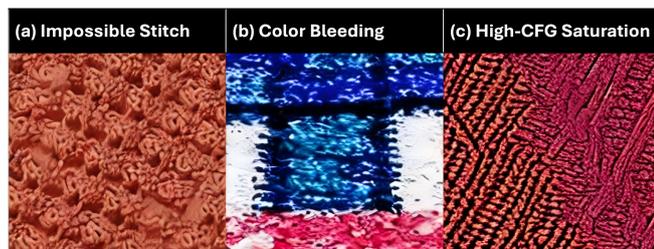


Figure 6: Typical Failure Cases. (a) Structural Inconsistency: Complex cable intersections may merge or disconnect. (b) Color Bleeding: Distinct colorwork motifs can blur into undefined noise. (c) Over-Guidance Artifacts: Using an excessively high classifier-free guidance scale (CFG ± 12.0) often results in unnatural contrast.

5 Conclusions

This study introduced a generative framework for knit-pattern design that integrates Stable Diffusion with lightweight LoRA fine-tuning and a browser-based interactive interface. In contrast to prior research focused primarily on manufacturability or abstract fashion ideation, this work emphasizes design-level exploration, empowering designers to create expressive, semantically rich knit patterns without extensive technical expertise.

The framework contributed on three fronts: (1) dataset curation and dual-labeling that combine technical and emotional attributes; (2) efficient domain adaptation of a large diffusion model via LoRA fine-tuning; and (3) a user-friendly interface enabling intuitive control over stylistic and emotional dimensions. Compared with the baseline Stable Diffusion model, the LoRA-tuned version produced outputs with clearer stitch definition, coherent motifs, and fewer non-knitting artifacts. Quantitatively, the LoRA model achieved an FID of approximately 195 and an IS of 2.54 (± 0.43) versus 3.17 (± 0.61) for the baseline, signifying a focused specialization toward knit-specific imagery rather than general photographic diversity. These trends, supported by qualitative visual analysis, confirmed that targeted fine-tuning can redirect a general diffusion model toward a coherent design domain.

Beyond technical outcomes, the research underscored important design implications. Embedding emotional descriptors within the conditioning process allows designers to treat mood, style, and intent as tangible design variables. This

aligns with contemporary textile workflows, where affective and aesthetic considerations drive early-stage ideation. Simultaneously, maintaining structural coherence ensured that generated outputs remain interpretable within the knitting domain, balancing creative exploration with technical relevance.

Nevertheless, limitations persist. The dataset's modest scale constrains stylistic coverage, and emotional descriptors require refinement to reduce subjectivity. Future work will address these aspects by scaling data collection, enriching the emotional vocabulary, incorporating multimodal inputs such as sketches or reference images, and integrating manufacturability constraints to connect conceptual design with machine-ready pattern generation.

In summary, this research demonstrates the potential of generative AI as a creative partner in knit design. By combining structural knowledge, emotional nuance, and human-centered interaction, diffusion models can act as catalysts for accessible, expressive, and exploratory design workflows, advancing the broader vision of democratizing textile design through creativity, personalization, and innovation.

Funding

The authors wish to thank the Hong Kong General Research Fund for funding support the project (Project No. 15602323).

Author Contributions

Conceptualization, Jiayin Fan; methodology, Jiayin Fan; software, Jiayin Fan; formal analysis, Jiayin Fan; data curation, Jiayin Fan and Li Li; writing - original draft preparation, Jiayin Fan; writing - review and editing, Li Li; supervision, Li Li; project administration, Li Li; funding acquisition, Li Li. All authors have read and agreed to the published version of the manuscript.

Conflict of Interest

All the authors declare that they have no conflict of interest.

Data Available

The data and materials used in this study are available upon request.

References

- [1] Kapllani, L., Amanatides, C., Dion, G., Shapiro, V., Breen, D.E.: TopoKnit: A process-oriented representation for modeling the topology of yarns in weft-knitted textiles. arXiv preprint arXiv:2101.04560 (2021). <https://doi.org/10.48550/arXiv.2101.04560>
- [2] Underwood, J.: The design of 3D shape knitted preforms. PhD thesis, RMIT University (2009). <https://core.ac.uk/download/pdf/15614587.pdf>
- [3] Kaspar, A., Oh, T.H., Makatura, L., Kellnhofer, P., Aslarus, J., Matusik, W.: Neural inverse knitting: From images to manufacturing instructions. arXiv preprint arXiv:1902.02752 (2019). <https://doi.org/10.48550/arXiv.1902.02752>
- [4] Scheidt, F., Ou, J., Ishii, H., Meisen, T.: deepKnit: Learning-based generation of machine knitting code. *Procedia Manufacturing* **51**, 485–492 (2020). <https://doi.org/10.1016/j.promfg.2020.10.068>
- [5] Sheng, H., Cai, S., Zheng, X., Lau, M.: Knitting robots: A deep learning approach for reverse-engineering fabric patterns. *Electronics* **14**(8), 1605 (2025). <https://doi.org/10.3390/electronics14081605>
- [6] Zheng, X., Sheng, H., Cai, S., Lau, M.C., Zhao, K.: Automated knitting instruction generation from fabric images using deep learning. In *2024 3rd International Conference on Automation, Robotics and Computer Engineering (ICARCE)*, pp. 197–201 (2024). <https://doi.org/10.1109/ICARCE63054.2024.00044>
- [7] Wu, X., Li, L.: An application of generative AI for knitted textile design in fashion. *The Design Journal* **27**(2), 270–290 (2024). <https://doi.org/10.1080/14606925.2024.2303236>
- [8] Hu, X., Yang, C., Fang, F., Huang, J., Li, P., Sheng, B.: MSEmGAN: Multi-stitch embroidery synthesis via region-aware texture generation. *IEEE Transactions on Visualization and Computer Graphics* **31**(9), 5334–5347 (2025). <https://doi.org/10.1109/TVCG.2024.3447351>
- [9] Karagoz, H.F., Baykal, G., Eksi, I.A., Unal, G.: Textile pattern generation using diffusion models. arXiv preprint arXiv:2304.00520 (2023). <https://doi.org/10.48550/arXiv.2304.00520>
- [10] Wang, Z., Xia, Y., Nie, K., Guo, M.: Alive Yi: Interactive preservation of Yi minority embroidery patterns through digital innovation. In *SIGGRAPH Asia 2024 Posters*, pp. 1–2 (2024). <https://doi.org/10.1145/3681756.3697919>
- [11] Zhou, X., Figueiredo, P., Hašan, M., Deschaintre, V., Guerrero, P., Hu, Y., Kalantari, N.K.: RealMat: Realistic materials with diffusion and reinforcement learning. arXiv preprint arXiv:2509.01134 (2025). <https://doi.org/10.48550/arXiv.2509.01134>
- [12] Hu, E.J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., Chen, W.: LoRA: Low-rank adaptation of large language models. arXiv preprint arXiv:2106.09685 (2021). <https://doi.org/10.48550/arXiv.2106.09685>
- [13] Tao, Z., Takida, Y., Murata, N., Zhao, Q., Mitsufuji, Y.: Transformed low-rank adaptation via tensor decomposition and its applications to text-to-image models. arXiv preprint arXiv:2501.08727 (2025). <https://doi.org/10.48550/arXiv.2501.08727>

- [14] Chen, D., Duan, Z., Li, Z., Chen, C., Chen, D., Li, Y., Chen, Y.: AttriCtrl: Fine-grained control of aesthetic attribute intensity in diffusion models. arXiv preprint arXiv:2508.02151 (2025). <https://doi.org/10.48550/arXiv.2508.02151>
- [15] Yang, J., Feng, J., Huang, H.: EmoGen: Emotional image content generation with text-to-image diffusion models. arXiv preprint arXiv:2401.04608 (2024). <https://doi.org/10.48550/arXiv.2401.04608>
- [16] Nimi, H., Lu, M., Chacon, J.C.: Embodied co-creation with real-time generative AI: An Ukiyo-e interactive art installation. *Digital* **5**(4), 61 (2025). <https://doi.org/10.3390/digital5040061>