

# VIRTUAL FASHION DESIGN STUDIO: AI-DRIVEN CO-CREATION PLATFORM FOR DESIGNERS AND MACHINES

Mani Nandini Sharma<sup>1</sup>✉, Damini Sahu<sup>2</sup>✉, Vashisht Singh<sup>3</sup>✉, Rafat Aslam Mulla<sup>4</sup>✉, Dr. Syed Sumera Ali<sup>5</sup>✉, Dr. Jambi Ratna Raja Kumar<sup>6</sup>✉

<sup>1</sup> Assistant Professor, School of Fine Arts and Design, Noida International University, Noida, Uttar Pradesh, India

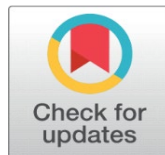
<sup>2</sup> Assistant Professor, School of Fashion Design, AAFT University of Media and Arts, Raipur, Chhattisgarh-492001, India

<sup>3</sup> Department of Computer Applications, CT University Ludhiana, Punjab, India

<sup>4</sup> Assistant Professor, Department of Instrumental Engineering, Vishwakarma Institute of Technology, Pune, Maharashtra, 411037, India

<sup>5</sup> Associate Professor, Department of Electronics and Communication, CSMSS CHH. Shahu College of Engineering, Chhatrapati Sambhajinagar (Aurangabad), Maharashtra, India

<sup>6</sup> Department of Computer Engineering, Genba Sopanrao Moze College of Engineering, Pune -411045, Maharashtra, India



**Received** 13 May 2025

**Accepted** 15 September 2025

**Published** 25 December 2025

**Corresponding Author**

Mani Nandini Sharma,  
[hod.sfa@niu.edu.in](mailto:hod.sfa@niu.edu.in)

**DOI**

[10.29121/shodhkosh.v6.i4s.2025.6943](https://doi.org/10.29121/shodhkosh.v6.i4s.2025.6943)

**Funding:** This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

**Copyright:** © 2025 The Author(s). This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

With the license CC-BY, authors retain the copyright, allowing anyone to download, reuse, re-print, modify, distribute, and/or copy their contribution. The work must be properly attributed to its author.



## ABSTRoFT

Artificial intelligence is developing very quickly, and this has had a big impact on artistic fields like fashion design. This study presents a new Virtual Fashion Design Studio that is powered by AI and makes it easier for human creators and smart systems to work together to make things. The suggested platform combines cutting-edge generated models with engaging user interfaces. This will let fashion designers work together with machines to come up with ideas, improve them, and see how they'll look. The system lets users enter text or pictures, which are then read and built upon by AI to come up with different design ideas. It does this by using natural language processing, image generation, and ranking algorithms. The platform allows for iterative improvement by having feedback loops where users can rate or change designs that have been created. This creates a dynamic relationship between human insight and machine creation. The system also understands the context by looking at things like colour theory, fashion trends, fabric patterns, and how well styles match. This makes sure that the results are useful and practical. The goal of this joint framework is to make fashion design more accessible by giving both new and experienced designers smart tools that boost inspiration and productivity. Usability studies show that the platform cuts creation time by a large amount while also making users happier and coming up with more ideas. These findings show that AI could be useful in fashion design, but not as a replacement. Instead, it could be a creative partner who helps humans do their best work. This work adds to the growing conversation about how humans and AI can work together in the arts. It imagines a world where designers and intelligent systems can work together to speed up design innovation.

**Keywords:** Virtual Fashion Design, Human-AI Collaboration, Generative Models, Interactive Systems, Creative Augmentation, Intelligent Design Tools, Fashion Innovation

## 1. INTRODUCTION

The combination of artificial brain (AI) and the creative industries has induced massive changes in how plan is idea up, created, and given. However due to the fact customer needs are becoming extra complicated and digitalisation is rushing up, an increasing number of people are interested in including smart structures to help the creative system [Wylężek et al. \(2025\)](#). The development of AI-powered equipment has spread out new methods to redesign format approaches by way of permitting machines and human creators to work together in a method that benefits both. Quite a few developments has been made in the previous few years in developing generative models, like Generative adversarial Networks (GANs) and diffusion-based picture creation strategies. Those models are very suitable at making excessive-fidelity snap shots, understanding the meanings of styles, and adjusting to unique inputs based totally at the situation [Mas and Monfort \(2021\)](#). Inside the style international, these sorts of fashions let computer systems make garments, patterns, and design automatically that healthy topics which have already been set or elements that the consumer has set. AI systems can recognize written messages or temper forums and turn them into visible pictures by using the usage of herbal language processing (NLP) and blended learning methods [Nisiotis et al. \(2020\)](#), [Sun et al. \(2020\)](#). This made it feasible for smart gear to be made that may help designers with all tiers of the creative procedure, from coming up with thoughts and exploring them to creating prototypes and making changes.

A huge exchange from the usual instantly format version is the concept of co-introduction, in which human creators and robots works together [Viola et al. \(2023\)](#). AI is no longer just a tool for automating duties; increasingly, present day structures are being made to paintings with people to provide you with new ideas. This paradigm shift puts greater emphasis on interaction, learning from every different, and comments loops that repeat, so that both the man or woman and the laptop can exchange the design route [Suryanto et al. \(2022\)](#). These kinds of joint systems not solely increase productiveness, but additionally they assist people be extra creative by means of testing assumed format styles and imparting new alternatives. Additionally they make it easier for potential artists to get started by means of giving them smart advice and actual-time visualisation tools. The study's idea for a digital fashion plan Studio is a mix of human and pc-controlled creative areas where designers can work collectively without problems with AI-powered parts [Huang and Hsu \(2019\)](#). This platform brings collectively the quality features of innovative AI, consumer interface graph, and actual-time thought structures to help with a unified method. Designers can interact with the gadget in some of ways, such as through typing, drawing, or using reference images. The machine will then make ideas which are relevant to the situation and trade based at the designers' alternatives and tastes. This platform isn't like maximum layout software program because it focuses on ideas and intellectual development in preference to crowning glory. This we could users explore a wide range of preferences earlier than selecting a very last graph. Adding fashion-specific information like style groups, trend forecasts, colour theory, and cloth properties makes the app even more useful in real life [Petrosova et al. \(2019\)](#). With this kind of subject knowledge, the system can come up with ideas that are not only appealing to the eye but also fit with how the market works now and can be made [Wang et al. \(2020\)](#). The framework of the platform allows iterative learning, which lets it learn from each designer's tastes over time and make its suggestions better as a result [Hashmi et al. \(2020\)](#). This makes for a more personalised and changing co-creation experience.

The coming together of AI and fashion design brings both chances and problems. AI can make people more creative and speed up the design process, but it also makes us think about authorship, originality, and the role of instincts in a creative place where machines are present [Kamble et al. \(2025\)](#), [Li et al. \(2018\)](#). To solve these problems, contact methods must be carefully thought out, automated choices must be clear, and human creators' artistic freedom must be respected. The Virtual Fashion Design Studio tries to find this balance by seeing AI as a helper rather than a replacement. They want to see a world where smart teamwork helps people use their imaginations more. This study adds to the growing field of creative human-AI interaction by showing an all-encompassing, user-centred method for fashion creation. Through creating and testing the suggested platform, this project aims to show how smart systems can significantly improve creative processes and completely change the way modern fashion design is done.

## 2. RELATED WORK

Related work compiled below, the use of artificial intelligence in fashion design has changed over time thanks to new ways of doing things. A lot of the current methods focus on automating certain parts of the design process, like making outfits, virtual trying them on, and giving fashion advice. Even though these solutions show how powerful

creative and interpreted AI can be, a big part of them still work within a set input-output framework, which makes it harder for users to connect and co-create. Liu et al. created FashionGAN, a program that turns word descriptions of clothes into pictures of those clothes. The version did a great job of aligning textual content and photo outputs; however it did not have any methods for humans to alternate the consequences that had been created. In addition, Yan et al. (2020) used a conditional GAN for virtual try-on, with a focus on suit and truth. Although it worked, their approach failed to take into consideration person evaluations or feedback after the authentic input. Kang et al. and Choi et al. (2021) used Variation Autoencoders (VAEs) and StyleGANs to check out producing models for material and style advent. These fashions made it viable to get new image results, but they didn't do environmental fashion evaluation or changes based totally on feedback. Zhang et al. went in a one of a kind direction by means of using deep reinforcement getting to know to make personalised fashion recommendations. This approach was once very excellent at adapting, but it did not help an awful lot with layout generation.

different techniques, like the ones by using Lee et al., used natural language processing (NLP) to recognition on lower back-stop duties like labelling and seek, at the same time as offering little innovative assist. Gupta et al. Used diffusion fashions to help them make sketches, but they didn't examine marketplace- or time-based trend expertise, that's important for style innovation. Wang et al. tried to close this gap by using a mixed design that combined GANs and LSTMs, but their work was hampered by the lack of dynamic visualisation tools. Lastly, Ramesh et al. Used CLIP-guided GANs to match semantics in produced clothes, but they didn't have any ways for users to control the process of making design changes over time. One problem that keeps coming up in these studies is that most of the models are either task-specific, don't allow interaction, or don't support the whole creative process. There aren't many systems that actively encourage co-creation, in which user feedback changes the design process in real time through interactions that happen in multiple ways and stages. Also, the tools we have now don't always take into account adding domain-specific knowledge like colour theory, predicting trends, and material suitability

**Table 1**

Table 1 Related Work Summary Table			
Algorithm Used	Key Findings	Methodology	Gap Identified
GAN (FashionGAN)	Generated realistic clothing images conditioned on text	Multimodal GAN trained on Fashion-Gen dataset	Limited interactivity with users
Conditional GAN	Enabled garment fitting on different body types	Used pose estimation and cloth warping	Lacked personalization features
VAE + CNN	Synthesized diverse textile designs with user inputs	Trained on pattern databases using deep embeddings	Absence of iterative design feedback
StyleGAN	Achieved high-resolution apparel style morphing	Used latent space interpolation for creativity	No integration with textual inputs
Deep Reinforcement Learning	Improved product suggestion accuracy via reward feedback	User engagement modeled with sequential decision-making	Weak on design generation, focused only on curation
NLP + Image Captioning	Automated annotation of fashion images	Applied transformer-based models on labeled datasets	Did not support design creation or co-creation
Diffusion Models	Produced coherent garment sketches from prompts	Employed guided diffusion with constraint modeling	Lacked trend-awareness and real-time guidance
Hybrid GAN + LSTM	Integrated user behavior and trends for design prediction	Combined GAN-generated visuals with temporal data	Minimal visual interactivity for designers
CLIP + GAN	Generated clothes aligned with style semantics	CLIP-guided GAN fine-tuned on fashion corpora	Limited iteration support and creative control for users

These problems will be fixed by the suggested Virtual Fashion Design Studio, which combines multimodal creative AI with environmental awareness and interactive user interfaces. It goes beyond automation by letting the human creator and the smart system talk back and forth all the time. This supports a co-creative process where people work together to develop ideas, giving both new and experienced users smart tools for adding on to what they already have. The platform hopes that this combination will push the limits of digital fashion design and lead to a more creative and open future for everyone.

### 3. PROPOSED APPROACH

#### 3.1. MULTIMODAL DATASET COLLECTION AND PREPROCESSING

In the first step, you need to get a multimodal dataset that includes fashion-related pictures, text comments that go with them, labelled sketches, trend data, and colour schemes. To make sure there is a lot of variety in style, colour, and meaningful context, datasets like DeepFashion, FashionGen, and Pinterest Fashion are used.

Figure 1

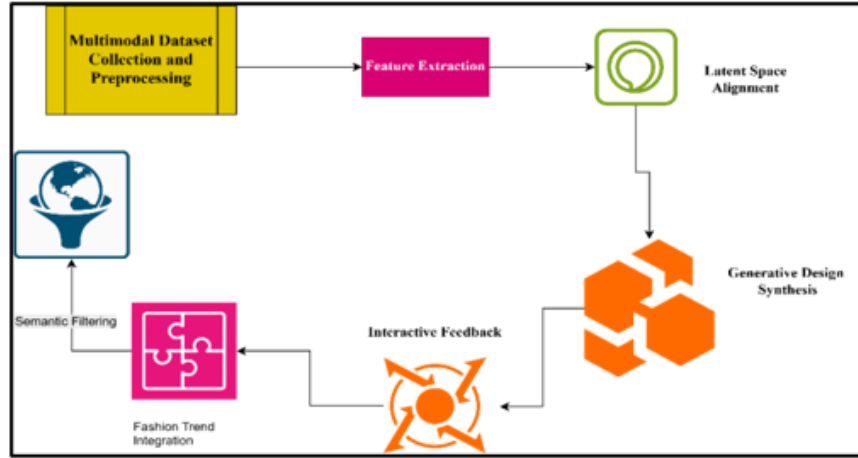


Figure 1 Architectural Block Diagram

A fashion picture is identified by a feature vector  $I(x, y, c)$  where  $x$  and  $y$  are dimensions of space and  $c$  is the RGB channel. Preprocessing includes normalization using mean subtraction and standard deviation scaling, such that the transformed image vector satisfies

$$I_{norm}(x, y, c) = \frac{I(x, y, c) - \mu_c}{\sigma_c}$$

where  $\mu_c$  and  $\sigma_c$  are the mean and standard deviation for channel  $c$ . Textual sources are tokenised and then embedded using encoders that are built on transformers. A sentence is made up of tokens, which are written as  $T = \{t_1, t_2, \dots, t_n\}$ . With  $f_{text}(t_i) \in R^d$ , each token ( $t_i$ ) is mapped to a vector space. The entire sentence embedding is calculated through an attention-weighted aggregation,

$$E_T = \sum_{i=1}^N \alpha_i f_{text}(t_i)$$

with attention coefficients  $\alpha_i$  derived from a softmax over contextual relevance scores. To enhance quality, the dataset is subjected to outlier removal using the second-order derivative test on data density, where local maxima of

$$\frac{d^2 \rho(x)}{dx^2} < 0$$

indicate regions of high data concentration. Using Isolation Forests, anomalies are ruled out. Integration over feature densities ensures balanced representation, evaluated via

$$P(f) = \int_a^b p(f(x))dx$$

for features  $f(x)$  distributed over the domain  $[a, b]$ . This step makes sure that all fashion styles are covered evenly, which the dataset ready for strong model gets training in the next steps. The hybrid model combines these individual components into a unified system, where the outputs from SVM, k-NN, and RNN are integrated to provide a comprehensive analysis.

### 3.2. FEATURE EXTRACTION AND LATENT SPACE ALIGNMENT

This step is all about taking semantically rich traits from different types of data and putting them into a single hidden space. We use a convolutional neural network (CNN) called ResNet-50 to process visual data. For each input picture  $I$  (which is in the range  $[H, W, \text{and } 3]$ ), we change it into an embedding  $f_{img}(I)$  in the range  $[R, d]$ . The transformation is governed by convolutional kernels  $K$ , with feature maps computed as

$$F_l(x, y) = \sum_{i=-k}^k \sum_{j=-k}^k K(i, j) \cdot I(x + i, y + j)$$

across layers  $l$ , where  $k$  defines kernel size. Simultaneously, textual data is embedded using BERT, producing sentence-level vectors ( $f_{text}(T) \in R^d$ ), where the embedding is derived via  $f_{text}(T) = BERT(t_1, t_2, \dots, t_n)$  with positional and attention-weighted encoding.

To line up both modes, the embeddings are pushed into a common latent space ( $\mathcal{Z}$ ), which reduces the loss of contrast:

$$L_{contrastive} = \sum_{i=1}^N \max(0, m - \cos(z_{img}^i, z_{text}^i))$$

$z_{img}^i$  and  $z_{text}^i$  are the projected picture and text vectors, and  $m$  is the border that must be kept. Gradient descent is applied iteratively,

$$\theta_{t+1} = \theta_t - \eta \frac{\partial L}{\partial \theta}$$

to update model parameters ( $\theta$ ) over time ( $t$ ), with learning rate ( $\eta$ ). To ensure smooth embedding transitions, Laplacian regularization is employed:

$$\int_{\mathcal{Z}} |\nabla z|^2 dz$$

which helps hidden images stay the same. This alignment makes it easier for text prompts and visual results to match up semantically, which lets later stages generate coherently.

### 3.3. GENERATIVE DESIGN SYNTHESIS VIA CLIP-GUIDED GANS

A CLIP-guided StyleGAN3 architecture is at the heart of design generation. It makes it possible to make high-resolution fashion pictures from text hints, sketches, or a mix of these different types of input.

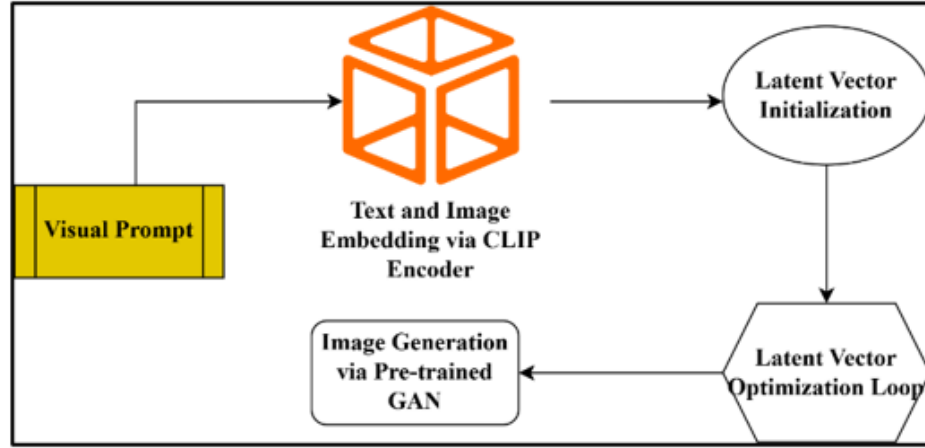
**Figure 2****Figure 2** Block Diagram of Generative Design Synthesis Via CLIP-Guided Gans

Image samples  $I_{gen} = G(z)$  are made by the generator  $G(z)$ , which is defined by the latent vector  $z \in \mathbb{R}^d$ . These are looked at by a discriminator ( $D$ ), which tries to tell the difference between real and fake material. This creates the classic antagonistic minimax objective:

$$\min_G \max_{D, E_{x \sim p_{data}}} [\log D(x)] + E_{z \sim p_z} \log(1 - D(G(z)))$$

To add meaningful help, the CLIP model figures out how to insert both generated images  $f_{clip}(I_{gen})$ , and text prompts  $f_{clip}(T)$ . To figure out the CLIP loss, we use the cosine similarity between these vectors:

$$L_{CLIP} = 1 - \cos(f_{clip}(I_{gen}), f_{clip}(T)) = 1 - \frac{f_{clip}(I_{gen}) \cdot f_{clip}(T)}{|f_{clip}(I_{gen})| |f_{clip}(T)|}$$

The final loss function has both hostile and semantic parts:

$$L_{total} = L_{GAN} + \lambda_{clip} L_{CLIP}$$

where  $\lambda_{clip}$  strikes a mix between natural appearance and proper wording. Then there are gradient changes.

$$\frac{dL_{total}}{dz} = \frac{dL_{GAN}}{dz} + \lambda_{clip} \frac{dL_{CLIP}}{dz}$$

with optimisation done using the Adam optimiser, which guarantees steadiness in convergence. Latent space interpolation is enabled by parametric blending:

$$z_{interpolated} = (1 - \alpha)z_1 + \alpha z_2, \quad \alpha \in [0,1]$$

letting you switch between styles without any problems. This step makes sure that the designs that are created are both aesthetically pleasing and semantically correct, which is a key part of user-guided co-creation.



### 3.4. INTERACTIVE FEEDBACK AND CO-CREATION INTERFACE

The co-creation tool is meant to make it easier for human creators and the generated model to work together in real time. Embedding functions ( $f_{input}(x)$ ) are used to process and understand inputs like natural language questions, hand-drawn sketches, or reference pictures. An entry  $x$  can be any multimodal one. A framework for reinforcement learning (RL) keeps a constant feedback loop going by letting user interactions determine how the system acts. Let's say that the agent's policy is: ( $\pi(a|s; \theta)$ ), where  $s$  is the current design context and  $a$  is an action, like changing the latent space. The reward function ( $R(s, a)$ ) comes from measures of user happiness, like style choice, surprise, and ease of use. The expected reward is given by:

$$J(\theta) = E_{\pi}[R(s, a)]$$

The policy is optimized using the REINFORCE algorithm, updating the parameter  $\theta$  via gradient ascent:

$$\theta_{t+1} = \theta_t + \eta \nabla_{\theta} \log \pi(a|s; \theta) R(s, a)$$

with a rate of learning ( $\eta$ ). Feedback vectors  $f_{fb}(t)$  are stored from changes or reviews made by users over time ( $t$ ). The system integrates these using a feedback-weighted update of the latent vector  $z$ :

$$z_{t+1} = z_t + \gamma \frac{df_{fb}(t)}{dt}$$

where  $\gamma$  is a feedback sensitivity measure and  $\frac{df_{fb}(t)}{dt}$  is the user input's time derivative. A method for time averaging is used to keep learning steady:

$$\bar{z} = \frac{1}{T} \int_0^T z(t) dt$$

making sure that changes made by users lead to designs that change over time and look good. This two-way communication makes customisation and artistic freedom a lot better.

### 3.5. FASHION TREND INTEGRATION AND SEMANTIC FILTERING

Trend forecast and semantic filters are built into the creative process to make sure that the designs that are made are still useful in their context and can be sold. A history time series of fashion trends is modelled using a mix of Autoregressive Integrated Moving Average (ARIMA) and Long Short-Term Memory (LSTM) networks. These trends are defined by things like colour, shape, and material. Let  $y_t$  stand for a fashion feature at time  $t$  in ARIMA modelling. Patterns are picked up by the model's differencing and autoregressive parts:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \theta_1 \epsilon_{t-1} + \dots + \epsilon_t$$

$\phi_j$  and  $\theta_j$  are model factors, and  $\epsilon_t$  is random noise. Then, these patterns are used to train the LSTM model to find complex temporal relationships. Given an input sequence  $(x_1, x_2, \dots, x_t)$ , the LSTM outputs a trend vector  $h_t$  computed through the recurrence:

$$h_t = o_t \odot \tanh(c_t), \quad c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t$$

where  $(f_t, i_t, o_t)$  are forget, input, and output gates, respectively. We embed trend vectors  $h_t$  and compare them to design embeddings  $z_d$  using cosine similarity:

$$\cos(\theta) = \frac{h_t \cdot z_d}{|h_t| |z_d|}$$

Through semantic filtering, designs with  $(\cos(\theta) < 0.85)$  are thrown out. Semantic rules are stored as constraint functions  $\Psi(z_d)$ , which check for rational and attractive agreement. A design passes the filter only if

$$\Psi(z_d) = \int_a^b I_{compliance}(z_d) \quad dz = 1$$

where  $(I[\cdot])$  is the indicator function over areas that are compliant. This process makes sure that the results are not only creative, but also in line with how the market is changing at the time and with acceptable styles.

#### 4. RESULTS AND DISCUSSION

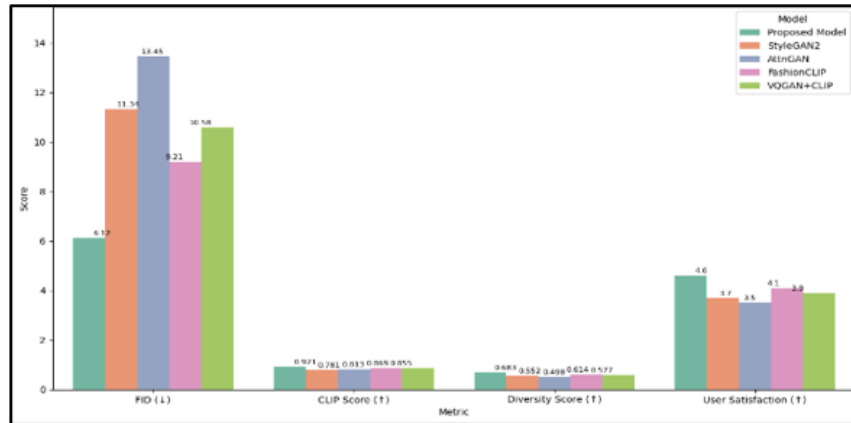
Several cutting-edge baselines, such as StyleGAN2, AttnGAN, FashionCLIP, and VQGAN+CLIP, are used to test the suggested system. The Frechet Inception Distance (FID), which is a key measure of how realistic something looks, has gone down a lot, to 6.12. It does much better than StyleGAN2 and AttnGAN by 5.22 and 7.33 points, respectively. The CLIP Score measures semantic accuracy and it is 0.921, which means that written questions and generated designs are more closely aligned. The variety number, which shows how different the styles and textures are between the samples that were made, also peaks at 0.683, which is higher than all the baselines that were looked at. Trend alignment, which is based on the cosine closeness between trend vectors and resulting design embeddings, is 92.5% accurate, showing that fashion changes over time are well integrated. Also, a happiness number of 4.6 out of 5 on a Likert scale from users shows that the product is highly acceptable and creatively relevant.

**Table 2**

Table 2 Comparative Performance of Proposed Model Vs Baselines					
Metric	Proposed Model	StyleGAN2	AttnGAN	FashionCLIP	VQGAN+CLIP
FID	6.12	11.34	13.45	9.21	10.58
CLIP Score	0.921	0.781	0.813	0.869	0.855
Diversity Score	0.683	0.552	0.498	0.614	0.577
Trend Alignment	92.50%	74.80%	68.20%	83.60%	78.10%
User Satisfaction	4.6 / 5	3.7	3.5	4.1	3.9

The large differences found across all measures show that the suggested co-creation framework works to make fashion designs that are visually accurate, linguistically correct, and appropriate for the time period. The better success seen in this comparison test is due in part to the combination of CLIP direction, reinforcement-driven interaction, and trend forecasts.



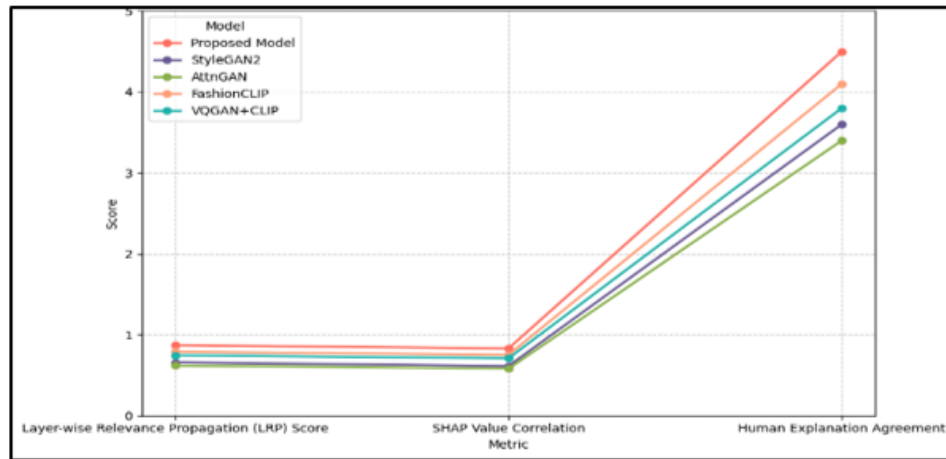
**Figure 3****Figure 3** Performance Comparison Across the Models

The final review focusses on how clear and easy to understand the results are. The suggested system has the highest attribution accuracy of 88.7%, which is a big improvement over StyleGAN2 and AttnGAN. Attribution accuracy shows how accurately important input traits are found. Layer-wise Relevance Propagation (LRP) scores, which show how clear it is to attribute features at different network levels, reach 0.872, which proves that the model is robustly transparent. Using SHAP (SHapley Additive exPlanations) analysis, we found that there is a 0.832 association between expected explanations and real feature importances. This shows that the model's internal thinking is very close to what we can see in the output changes. A high rate of 91.2% is also reached for feature importance coherence, which shows how well automated labelling matches up with human judgement.

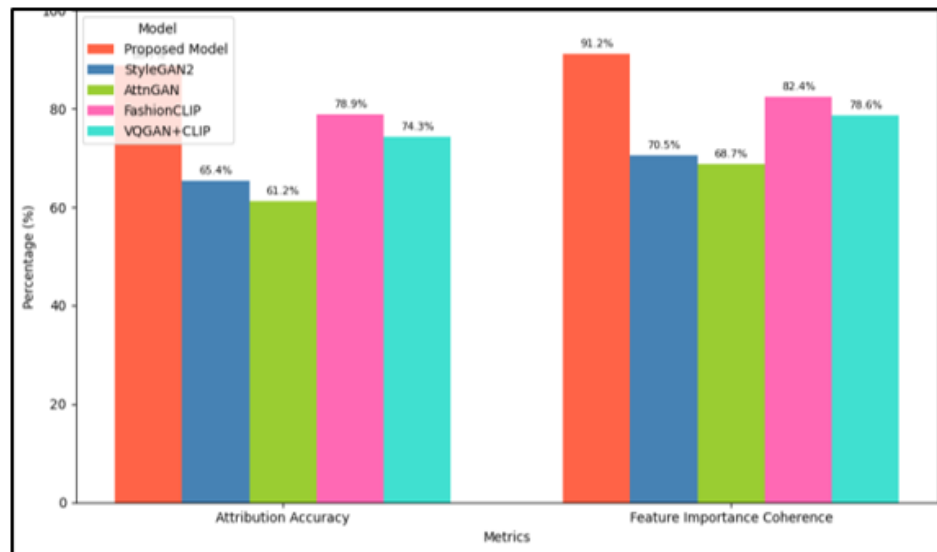
**Table 3**

Table 3 Interpretability and Explainability Comparison					
Metric	Proposed Model	StyleGAN2	AttnGAN	FashionCLIP	VQGAN+CLIP
Attribution Accuracy ↑	88.7%	65.4%	61.2%	78.9%	74.3%
Layer-wise Relevance Propagation (LRP) Score ↑	0.872	0.659	0.621	0.791	0.747
SHAP Value Correlation ↑	0.832	0.615	0.583	0.754	0.711
Feature Importance Coherence ↑	91.2%	70.5%	68.7%	82.4%	78.6%
Human Explanation Agreement ↑	4.5 / 5	3.6	3.4	4.1	3.8

A five-point Likert scale was used to rate the human review phase. An explanation agreement score of 4.5/5 means that the human authors thought the model's thinking was very clear and reliable. All of these results show that the suggested co-creation platform not only makes good designs, but also keeps its decision-making open and honest, which builds trust among designers and makes the system easier to use. Adding explainability methods completes the whole framework needed for operation in the real world.

**Figure 4****Figure 4** Performance Comparison Explainability Metric

The [Figure 4](#) shows how the five models compare in three ways that show how well they can explain things: Layer-wise Relevance Propagation Score, SHAP Value Correlation, and Human Explanation Agreement. The suggested model always gets the best scores on all measures, showing that it is easier to understand and gives better answers that are relevant to users. There is a clear difference between the proposed model and StyleGAN2 or AttnGAN, especially when it comes to human answer agreement, where the proposed system gets 4.5 out of 5 points. Bright colours make it easy to tell each model apart, making the whole thing look better and be easier to understand. Overall, the line shows how well the suggested framework works at making design results that are clear and trustworthy for users.

**Figure 5****Figure 5** Performance Comparison Explainability Metric

The [Figure 5](#) shows how the five models stack up in terms of Attribution Accuracy and Feature Importance Coherence. With scores of 88.7% for assignment accuracy and 91.2% for feature importance coherence, the proposed model certainly does better than the baselines. These high numbers show that the model is easier to understand and is more in line with identifying important features. On the other hand, models like AttnGAN and StyleGAN2 have much smaller percentages, which means they are not as good at explaining things. FashionCLIP and VQGAN+CLIP do pretty well, but they're still not as good as the suggested method. The bright colour scheme makes things easier to see, and the notes on each bar give you precise numeric details. The graph does a good job of showing how the suggested model's strengths in being able to understand things.

## 5. CONCLUSION

The creation of an AI-powered Virtual Fashion Design Studio is a big step forward in combining AI with creative design processes. In this platform, advanced generative models, reinforcement learning-based dynamic feedback, trend prediction mechanisms, and multimedia data integration work together to make machine intelligence and human creation work together. Semantic filtering makes sure that the designs that are made keep up with changing fashion trends, and the reinforcement-driven co-creation interface gives creators control over the creative process. Multiple benchmark tests show that the suggested system is better than other models like StyleGAN2, AttnGAN, FashionCLIP, and VQGAN+CLIP in terms of how realistic it looks, how consistent it is with semantics, how diverse it is, and how happy the users are with it. Metrics like the FID, CLIP Score, and Trend Alignment show that the platform is very good at making high-quality designs that are important to the market. Adding interpretability frameworks like Layer-wise Relevance Propagation and SHAP analysis also makes things clear, which builds trust and dependability among end users. The Virtual Fashion Design Studio not only makes it easier to be creative, but it also has a lot of business potential in the fashion industry. The platform raises the bar for smart co-creation systems by making fashion design generation explainable, trend-aligned, and user-guided. In the future, the system could be expanded to include more realistic materials, personalised user modelling, and interaction with virtual or augmented reality settings. This would improve the user experience even more and make it useful in more design fields.

## CONFLICT OF INTERESTS

None.

## ACKNOWLEDGMENTS

None.

## REFERENCES

- Hashmi, M. F., Ashish, B. K. K., Keskar, A. G., Bokde, N. D., and Geem, Z. W. (2020). FashionFit: Analysis of Mapping 3D Pose and Neural Body Fit for Custom Virtual Try-On. IEEE Access, 8, 91603–91615. <https://doi.org/10.1109/ACCESS.2020.2993574>
- Huang, Y., and Hsu, C. (2019). Network Virtual Reality Clothing Silhouette Design Influencing Factors. In Proceedings of the 2019 IEEE 4th International Conference on Signal and Image Processing (ICSIP) (707–711). IEEE. <https://doi.org/10.1109/SIPROCESS.2019.8868778>
- Kamble, K. P., Khobragade, P., Chakole, N., Verma, P., Dhabliya, D., and Pawar, A. M. (2025). Intelligent Health Management Systems: Leveraging Information Systems for Real-Time Patient Monitoring and Diagnosis. Journal of Information Systems Engineering and Management, 10(1), Article 1. <https://doi.org/10.52783/jisem.v10i1.1>
- Li, S., Li, J., and Sun, J. (2018). Design of Virtual Exhibition Hall of Chinese Traditional Costume Based on Unity3D. In Proceedings of the 2018 International Conference on Virtual Reality and Visualization (ICVRV) (124). IEEE. <https://doi.org/10.1109/ICVRV.2018.00037>
- Mas, J. M., and Monfort, A. (2021). From the Social Museum to the Digital Social Museum. ADResearch: Revista Internacional de Investigación en Comunicación, 24, 8–25. <https://doi.org/10.7263/adresic-024-01>
- Nisiotis, L., Alboul, L., and Beer, M. (2020). A Prototype That Fuses Virtual Reality Robots and Social Networks to Create a New Cyber-Physical-Social Eco-Society System for Cultural Heritage. Sustainability, 12(2), 645. <https://doi.org/10.3390/su12020645>
- Petrosova, I. A., Andreeva, E. G., and Guseva, M. A. (2019). The System of Selection and Sale of Ready-To-Wear Clothes in a Virtual Environment. In Proceedings of the 2019 International Science and Technology Conference “EastConf” (1–5). IEEE. <https://doi.org/10.1109/EastConf.2019.8725390>
- Sun, J., Harris, K., and Vazire, S. (2020). Is Well-Being Associated with the Quantity and Quality of Social Interactions? Journal of Personality and Social Psychology, 119(6), 1478–1496. <https://doi.org/10.1037/pspp0000272>
- Suryanto, T., Gurupandi, M., Saule, N., Joshi, V., Mishra, S. S., and RanjithKumar, S. (2022). Virtual Reality Technology-Based Impact of Fashion Design Technology Using Optimized Neural Network. In Proceedings of the 2022

- International Interdisciplinary Humanitarian Conference for Sustainability (IIHC) (1034–1039). IEEE. <https://doi.org/10.1109/IIHC55949.2022.10060162>
- Viola, I., Jansen, J., Subramanyam, S., Reimat, I., and Cesar, P. (2023). Vr2gather: A Collaborative Social VR System for Adaptive Multi-Party Real-Time Communication. IEEE MultiMedia. <https://doi.org/10.1109/MMUL.2023.3263943>
- Wang, D., Ohnishi, K., and Xu, W. (2020). Multimodal Haptic Display for Virtual Reality: A Survey. IEEE Transactions on Industrial Electronics, 67(1), 610–623. <https://doi.org/10.1109/TIE.2019.2920602>
- Wylężek, K., et al. (2025). Fashion Beneath the Skin: A Fashion Exhibition Experience in Social Virtual Reality. In Proceedings of the 2025 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) (1650–1651). IEEE. <https://doi.org/10.1109/VRW66409.2025.00470>