AIGC-Driven Digital Fashion Performance Content Creation and Multi-Modal Communication Path Exploration

Lizhen SHEN
Zhejiang Fashion Institute of Technology, Ningbo 315000, Zhejiang, China
Corresponding author: Lizhen SHEN (e-mail: shenlizhen1688@163.com)

ABSTRACT To address the inconsistency between the physical properties of eco-friendly textile materials and their visual representation in digital fashion generation, as well as the limited adaptability of multimodal communication under dynamic performance scenarios, this study proposes an Artificial Intelligence Generated Content (AIGC) framework integrating a Physics-Informed Neural Network (PINN) with a diffusion-based generation model and multimodal recommendation strategy. Material microstructure information and physical constraint equations are embedded into the generation process to jointly model permeability, glossiness, and elastic behavior, while motion-aware texture evolution and style-guided optimization improve the realism and physical consistency of virtual garments. A crossmedia adaptation mechanism further enables efficient deployment across AR fitting, virtual performances, social media , and metaverse applications through dynamic parameter mapping and low-latency transmission. Experimental results demonstrate superior recommendation accuracy, visual consistency, and computational efficiency, achieving an 89.7% Top-5 accuracy, an 83.2% Recall@10, and an 85 ms generation latency while maintaining stable physical perception of material properties. By bridging physically constrained content generation with multimodal communication, the proposed framework provides a reliable solution for intelligent digital fashion and sustainable textile visualization. Furthermore, its integration of real-time transmission and multimodal perception offers valuable insights for electromagnetic-enabled wireless media delivery, antenna-assisted immersive display systems, and future smart communication platforms requiring high-fidelity visual information propagation.

INDEX TERMS aigc, physics-informed neural network, multimodal communication, electromagnetic-enabled transmission, digital fashion

I. INTRODUCTION

THE rapid development of AIGC technology has promoted the application of digital virtual clothing in ecommerce, fashion design, and other fields [1– 3]. In particular, in dynamic scenes such as virtual catwalks and digital stage performances, clothing generation technology needs to meet complex requirements such as motion-driven animation, camera angle changes, and stage lighting synchronization. Traditional static texture generation can no longer cover the dynamic changes of real performances, making the expansion of AIGC in the field of digital clothing performance a research hotspot [4–6]. However, current digital clothing generated by AIGC exhibits deviations in simulating the physical properties of environmentally friendly materials, and key properties such as breathability and texture are difficult to accurately convey through visual content [7–9]. This disconnect between users' expectations of the environmentally friendly clothing experience and actual performance has become a bottleneck in the efficiency of transforming sustainable consumption models and a technical obstacle to the low-carbon transformation of the textile industry [10]. Existing studies have discussed the challenges of digitally representing the physical properties of environmentally friendly textile materials [11–13]. Building on these insights, this paper further focuses on the creation of digital clothing performance content [14] and the exploration of multimodal communication paths [15]. With the development of diversified media channels, such as metaverse spaces, AR fitting, and social short video platforms, the dissemination mode of digital clothing has evolved from single-image display to immersive and interactive scenes. Regarding the challenges of digital modeling of environmentally friendly materials, existing studies have investigated their physical properties and mechanical behaviors, such as bending stiffness, tensile responses, and weaving parameters [16–18], and have also explored visualization through methods including generative adversarial networks (GANs) [19–21]. Dan R, Liu Z, and Shi Z attempted to simulate the elastic coefficient of recycled fibers through finite element analysis, but the computational complexity resulted in low generation efficiency [22–24]. Malabadi et al. focused on the sustainable potential of plant-based leather made from diverse natural resources [25], and Mandal et al. reviewed the application of plant-derived natural dyes in leather dyeing, with attention to color fastness and environmental impact [26]. Although these studies advance the understanding of eco-friendly leather materials, they do not address the unified modeling of physical properties and visual effects. In this study, these properties are not merely generic attributes but sustainability indicators, as they affect the energy use, life cycle durability, and recyclability of eco-friendly textiles.

To resolve the above contradictions, scholars have proposed various improvement plans. Xu et al. proposed Fabric-GAN, an enhanced generative adversarial network designed for data augmentation in fabric defect detection, which improved the generalization ability and realism of generated defect images, thereby significantly boosting the accuracy of textile defect recognition [27]. However, its optimization goal is still limited to visual features and ignores physical property constraints. There are studies addressing the multiscale challenges of fiber-reinforced thermoplastic composites, where homogenization techniques are applied to describe consolidation processes and predict fiber disorientation during press molding [28]. While these models provide valuable insights for composite manufacturing, they do not focus on the visualization of ecofriendly textile materials or their tactile properties. Some studies have applied neural networks in textile material analysis. Li et al. employed near-infrared spectroscopy combined with a back-propagation neural network to achieve accurate qualitative identification of waste textiles [29]. Zulfiqar et al. provided a review of artificial-neural-network-based models for predicting seam strength in denim jeans, emphasizing their potential for improving garment durability assessment [30]. Xu et al. proposed a convolution-augmented Transformer model for fiber identification from fabric images, which significantly enhanced textile classification accuracy [31]. Wu et al. developed the ClothGAN framework based on generative adversarial networks and style transfer, and, by constructing a Dunhuang clothing dataset, generated clothing designs that incorporated Dunhuang elements [32]. Although these methods have advanced either the recognition or generative aspects of textiles, they do not achieve the coordinated optimization of environmentally friendly material physical properties with visual generation and multimodal communication.

This study builds a collaborative generation system around the dual goals of performance content creation and multimodal communication path optimization for digital clothing: by embedding motion capture data (gait cycle, rotation angle) into the Stable Diffusion model, real-time synchronization of texture dynamic response and wrinkle distribution is achieved, enhancing physical consistency and visual realism in stage performance. For communication scenarios such as AR fitting and virtual live broadcast, a feature vector fusion model is constructed to achieve cross-media recommendation adaptation of user preferences, body shape, and material parameters. This research establishes a complete link from dynamic content generation to media-intelligent distribution, promoting digital clothing from watchable to performable and usable. While this study focuses on eco-friendly materials due to their increasing importance in sustainable fashion, the proposed PINN-diffusion framework is not limited to such materials. The method simulates physical properties shared by eco-friendly and conventional textiles, such as permeability, gloss, and elastic modulus. Therefore, the method has general applicability, and this paper selects eco-friendly materials as representative case studies to demonstrate its practical value.

II. ALGORITHM DESIGN

ENVIRONMENTALLY FRIENDLY MATERIAL MODELING BASED ON PHYSICAL INFORMATION NEURAL NETWORK (PINN)

This study employs a PINN framework to develop a cross-scale modeling method that incorporates physical constraints for the microstructural characteristics of two typical environmentally friendly materials: recycled fibers and plant-based leather. These two materials are selected as representative examples due to their commercial relevance and distinct microstructures, but the proposed framework is also adaptable to other sustainable textiles, such as organic cotton, hemp, and bio-based composites, provided their physical parameters are available. By integrating experimental measurement data with theoretical physics equations, a joint optimization model that includes material geometric parameters, mechanical response, and mass transfer performance is established. The experiment includes three-dimensional reconstruction data of recycled polyester fiber samples and plant-based leather samples, which are meshed to form a voxelized input matrix. For the air permeability simulation, the unsteady Darcy–Forchheimer equation is established:

$$\nabla \cdot \left(\frac{\rho K}{\mu} \nabla p \right) + \nabla \cdot \left(\beta \rho^2 K^2 |\nabla p| \nabla p \right) = \frac{\partial (\rho \epsilon)}{\partial t} \quad (1)$$

K denotes the permeability tensor, β is the non-Darcy coefficient, and ϵ represents the porosity. The mechanical response is modeled using a modified Mooney–Rivlin hyperelastic formulation:

$$\sigma_{ij} = C_1 \left(\lambda_i^2 - \lambda_k^{-2} \right) \delta_{ij} + C_2 \left(\lambda_i^4 - \lambda_k^{-4} \right) \delta_{ij} + p \delta_{ij} \quad (2)$$

C_1 and C_2 are material parameters; λ is the stretch ratio; p is the hydrostatic pressure. The material parameters, including the permeability tensor K, non-Darcy coefficient, and

Mooney–Rivlin constants (C_1, C_2), were obtained through preliminary sampling and physical response calibration on small representative specimens of recycled fibers and plant-based leather. These calibration results were used to fit the constitutive equations, and the finalized values were stored in the material microstructure database as inputs for the PINN. Surface glossiness was measured using a BYK micro-TRI-gloss meter at a 60° incident angle. To ensure comparability across different materials, the raw gloss values were normalized and reported as arbitrary units (a.u.).

The neural network architecture adopts an improved structure of the deep implicit field network, which includes four fully connected hidden layers. The input layer dimensions comprise material feature vectors and physical parameters, and the output layer analyzes the air permeability and elastic modulus (in units of MPa).

The loss function consists of three parts: the data fitting term L_{data}, the physical constraint term $L_{physics}$, and boun dary matching L_{BC}.

A two-stage optimization process is used during training: first, pre-training is performed with frozen physical constraints, followed by fine-tuning using the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) solver. The dynamic visualization module uses NeRF (Neural Radiance Fields) technology to map the physical parameters output by PINN to a three-dimensional material field. By constructing a continuous mapping from spatial coordinates (x, y, z) to physical properties (air permeability, elastic modulus), real-time rendering of the microstructure is achieved. The user interface supports parameter adjustment, and the system can complete physical field reconstruction and visualization updates to meet the real-time requirements of the virtual fitting scene. The mapping relationship is shown in Figure 1.

As shown in Figure 1, the dynamic visualization module integrates physical parameter prediction with realtime interactive rendering. The system further provides an interactive parameter adjustment interface, allowing users to tune key properties in real time. These adjustments are directly linked to the rendering parameters of the diffusion model, enabling personalized visualization output.

DIFFUSION MODEL-DRIVEN VIRTUAL CLOTHING GENERATION

An improved Stable Diffusion architecture is proposed to construct a diffusion-based generative model for eco-friendly clothing materials. By fusing microscopic material physical parameters (derived from PINN) with multimodal text descriptions, it generates high-resolution images (1,024 × 1,024) with the texture of recycled fibers and the optical properties of plant-based leather, addressing the ambiguity inherent in traditional GAN models in fine-grained material representation. New dynamic texture generation and style transfer modules are added to synergistically optimize clothing visual effects and stage scene requirements.

MODEL ARCHITECTURE AND CONDITIONAL INPUT DESIGN

The pre-trained Stable Diffusion v2.1 model (UNet architecture) is used, with the last three decoding layers replaced to enhance local texture generation. Input conditions include: a 768-dimensional vector τ that encodes the text description, a 128-dimensional physical parameter vector ϕ output by PINN (embedded in UNet layers 12, 18, and 24), motion parameters (model rotation angle θ and gait period t, embedded in layers 9 and 15 to drive dynamic wrinkle generation), and an artistic style code s. The choice of these layers follows the functional hierarchy of the UNet architecture: mid-to-late decoding layers (12, 18, 24) are responsible for restoring fine-grained texture and optical details, making them suitable for integrating physical properties, while intermediate layers (9, 15) better capture large-scale geometric deformation, making them appropriate for embedding motion dynamics.

TRAINING STRATEGY AND LOSS FUNCTION

A two-stage fine-tuning strategy is adopted. In the first stage, the first 10 layers of UNet are frozen, and the last 22 layers are fine-tuned with L_{simple} loss on the environmentally fri endly material subset. In the second stage, all parameters ar e unfrozen, the physical condition vector is applied, and th e hybrid loss function is switched:

$$\mathcal{L} = \mathbb{E}_{x_0,\epsilon,t,c}\left[\|\epsilon - \epsilon_\theta(x_t, t, \tau, \phi)\|_2^2\right] + \lambda_{\mathrm{vlb}}\mathcal{L}_{\mathrm{vlb}} \quad (3)$$

In formula (3), $\mathcal{L}_{\mathrm{vlb}}$ is the variational lower bound loss, ϵ denotes the noise, ϵ_θ is the noise prediction network, x_t is the noisy image at time step t, τ is the text embedding, and ϕ is the physical parameter vector. A temporal consistency loss is introduced to constrain the inter-frame continuity of dynamic textures, and a style transfer loss is introduced to ensure that the artistic style aligns with the text description.

PHYSICAL PERCEPTION GENERATION OPTIMIZATION

Physically-aware optimization is achieved through a differentiable rendering module: a normal map N is generated based on BRDF parameters and fused into the output. The specular reflectance ρ = 0.5 × (1 – G) + 0.5 × G (G is the glossiness parameter) is dynamically adjusted to adapt the glossy/matte effects to stage lighting variations. Using the fiber orientation O output by PINN, a directional gradient constraint is applied via a Steerable CNN layer to control texture alignment. The inference phase utilizes a DDIM sampler, with input conditions expanded to include textual prompts, a physical parameter vector ϕ, and motion parameters θ and t. This generates dynamic sequences synchronized with the runway motion, ensuring a negative correlation between the skirt's hem amplitude and the elastic modulus E, thereby enhancing physical consistency.

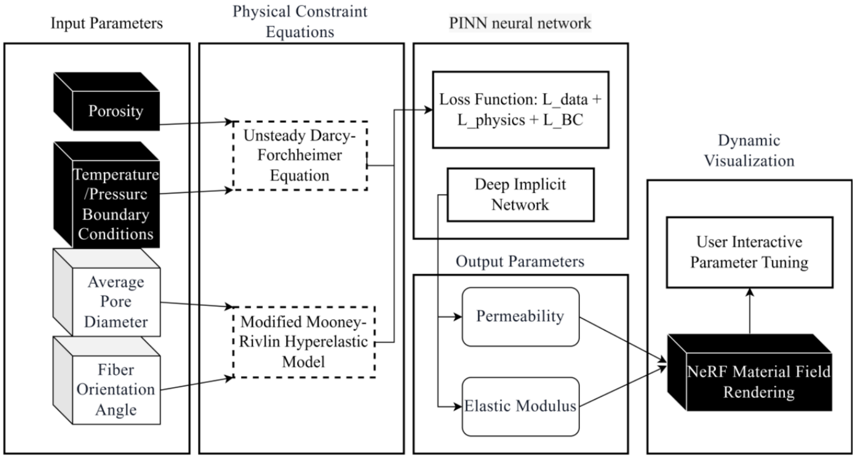

FIGURE 1. Mapping the relationship between material properties and PINN parameters

MULTIMODAL USER NEEDS MODELING

This study develops a multimodal recommendation model based on the Transformer architecture, enabling the dynamic matching of clothing features and user needs by integrating user behavior data, environmental preference labels, and body parameters. The data input layer contains three types of heterogeneous data sources: user browsing records are extracted through a Temporal Convolutional Network (TCN) to form a 128-dimensional dynamic preference vector; environmental labels are embedded and mapped into 64-dimensional semantic vectors through a knowledge graph; PointNet++ encodes the body data to generate 32-dimensional geometric features. All feature vectors are aligned to scale through the adaptive normalization layer (LayerNorm) and concatenated into a 224-dimensional multimodal feature vector $\mathbf{v} \in \mathbb{R}^{224}$. The 3D physical parameters are not concatenated into the 224-dimensional feature vector but are used as key/value inputs in the cross-attention mechanism. The user features and recommendation weight distribution are shown in Table 1.

TABLE 1. User Characteristics and Recommendation Weight Distribution

Feature Type	Dimension	Processing Method	Weight Allocation	Function Description
User Browsing History	128D	Temporal Convolutional Network (TCN)	Dynamic Preference Weight: 0.4	Captures recent preference trends
Eco-Friendly Labels	64D	Knowledge Graph Embedding	Semantic Association Weight: 0.3	Establishes semantic linkage between eco-labels and material properties
Body Shape Data	32D	PointNet++ Encoding	Fit Accuracy Weight: 0.2	Optimizes garment fit based on 3D scan geometry
Physical Parameters (PINN Output)	3D	Cross-Modal Attention Embedding	Material Property Weight: λ = 0.3	Enhances attention weights to align breathability/gloss with visual features
Diversity Enhancement Module	-	Maximum Marginal Relevance	Relevance/Diversity Trade-off: λ = 0.6	Balances novelty and relevance; optimizes coverage

The core of the model utilizes an improved Transformer encoder, comprising eight layers of stacked self-attention modules and feedforward networks. The improvements include: (1) we apply a cross-modal attention mechanism by embedding PINN-predicted physical parameters (air permeability ϕ_p and glossiness ϕ_g) into the key vectors, so that the attention weights α_i are conditioned on material properties:

$$\alpha_i = \text{Softmax}\left(\frac{\mathbf{Q}_i \cdot (\mathbf{K}_i + \lambda\phi_p \oplus \phi_g)}{\sqrt{d_k}}\right) \tag{4}$$

In formula (4), λ is the physical parameter weight, and $\oplus$ represents vector concatenation. (2) The user behavior temporal offset Δt is applied in the position encoding to enhance the influence of recent preferences on the recommendation results. The feedforward network uses a bilinear gating unit to control the feature transfer strength through the gating signal $\sigma(\mathbf{W}_g\mathbf{v})$:

$$\mathbf{h}' = \sigma(\mathbf{W}_g\mathbf{v}) \otimes \tanh(\mathbf{W}_1\mathbf{v}) + \mathbf{b}_1 \tag{5}$$

In the training phase, a contrastive learning strategy is employed to construct positive and negative sample pairs. The positive sample is the clothing feature $\mathbf{v}^+$ that the user clicks and purchases, while the negative sample is the unclicked sample $\mathbf{v}^-$ of the same category. Negative samples are sampled with category balancing to ensure fair representation across all material types. The loss function consists of two parts:

$$\mathcal{L} = \mathcal{L}_{\text{contrastive}} + \beta\mathcal{L}_{\text{reg}} \tag{6}$$

The contrast loss term is defined as:

$$\mathcal{L}_{\text{contrastive}} = -\log\frac{\exp(\mathbf{h}^T\mathbf{h}^+/\tau)}{\exp(\mathbf{h}^T\mathbf{h}^+/\tau) + \sum_{k=1}^{N}\exp(\mathbf{h}^T\mathbf{h}_k^-/\tau)} \tag{7}$$

The regularization term $\mathcal{L}_{\text{reg}}$ constrains the distribution consistency of the recommended results and the PINN physical parameters:

$$\mathcal{L}_{\text{reg}} = \|\mu_\phi - \hat{\mu}_\phi\|_2^2 + \|\sigma_\phi - \hat{\sigma}_\phi\|_2^2 \tag{8}$$

In formula (8), μ_ϕ and σ_ϕ are the mean and standard deviation of the physical parameters of the training set; $\hat{\mu}_\phi$

and $\hat{\sigma}_\phi$ are the model output statistics; the weight coefficient is $\beta = 0.2$.

In the inference phase, a recommendation list is generated through a dynamic threshold screening mechanism: first, the matching score $s_i = \mathbf{v}_u^T \mathbf{v}_i$ between the user feature vector $\mathbf{v}_u$ and each sample $\mathbf{v}_i$ in the candidate clothing library is calculated, and the top 20 candidates are selected after sorting by score; then, an environmental parameter filter is applied to remove samples whose air permeability ϕ_p is lower than the user's historical preference average; finally, through the diversity enhancement module, the Maximum Marginal Relevance (MMR) algorithm is used to balance the novelty and relevance of the recommendation results:

$$\mathrm{MMR} = \arg\max_{i\in\mathcal{R}} \left[\lambda s_i - (1-\lambda)\max_{j\in\mathcal{S}} \mathrm{Sim}(\mathbf{v}_i, \mathbf{v}_j)\right] \quad (9)$$

CROSS-MEDIA COMMUNICATION PATH ADAPTATION

Based on the user demand modeling results of physical property perception, this study designs a cross-media communication path adaptation mechanism that addresses the differentiated demand problems of multimodal content in AR fitting, social media short videos, digital stage live broadcasts, and other scenarios through dynamic parameter mapping and communication efficiency optimization. For AR fitting scenarios, real-time parametric mapping technology based on physically based rendering (PBR) is used to map the air permeability P and glossiness G output by PINN into the surface normal map weight ω_{PB} and the specular reflection coefficient ρ_{SP}. The conversion relationships are:

$$\omega_{PB} = \sigma\left(\frac{\partial P}{\partial t} + \nabla\cdot\mathbf{J}_G\right) \quad (10)$$

$$\rho_{SP} = \frac{G - G_{\min}}{G_{\max} - G_{\min}} \quad (11)$$

In formula (10), σ is the Sigmoid activation function, and $\mathbf{J}_G$ is the gloss gradient flow field.

In the dissemination of short videos on social media, a dynamic cropping and style transfer fusion algorithm is designed. Based on the intermediate layer feature map F_{mid} of the Diffusion model, a key frame sequence is generated, and the attention mask M_{att} is constructed in combination with the user's environmental preference label T_{eco}:

$$M_{\text{att}}(x,y) = \sum_{i=1}^{N} w_i \cdot \exp\left(-\frac{\|F_{\text{mid}}(x,y) - c_i\|^2}{2\sigma^2}\right) \quad (12)$$

In formula (12), c_i is the label embedding vector, and w_i is the semantic weight.

For the live broadcast scenario of the digital stage, a low-latency transmission protocol based on 5G edge computing is developed. By quantifying the joint entropy $H(\theta_{\text{diff}}, \Phi_{\text{phys}})$ of the diffusion model parameter θ_{diff} and the physical field data Φ_{phys}, a dynamic bit rate allocation model is established:

$$R_{\text{bit}}(t) = \alpha\cdot H(\theta_{\text{diff}}, \Phi_{\text{phys}}) + \beta\cdot\frac{\partial Q}{\partial t} \quad (13)$$

In formula (13), α and β are empirical coefficients, and Q is the network jitter parameter.

Cross-media communication optimization adopts a multi-objective planning model to balance click-through rate (CTR), content coverage (Cov), and energy consumption (E):

$$\begin{aligned}&\min_x\{\lambda_1(1-CTR) + \lambda_2(1-Cov) + \lambda_3 E/E_{\max}\}\\ &\quad\text{s.t.}\quad \sum_{i=1}^{3}\lambda_i = 1\end{aligned} \quad (14)$$

To further enhance communication efficiency and visual consistency across different media environments, a media feature library Mc is constructed, encompassing indicators such as frame rate upper limit, interactivity requirements, resolution requirements, and physical consistency priority.

PINN–DIFFUSION COUPLING GENERATION OPTIMIZATION

This study achieves end-to-end coupling between the PINN model output and the Diffusion generation process through physical parameter embedding and a joint training strategy, solving the simulation distortion problem caused by the separation of physical properties and visual effects of environmentally friendly materials. Specific methods include:

(1) Physical parameter embedding mechanism: In the UNet noise prediction module of Stable Diffusion v2.1, the physical parameter vector $\phi = [\phi_p, \phi_e, \phi_g]$ (air permeability ϕ_p, elastic modulus ϕ_e, gloss ϕ_g) generated by PINN is used as a conditional input. In the noise prediction network at time step t, ϕ is mapped to the scale factor γ and offset β through a linear transformation, and the intermediate feature map $\mathbf{F}_t$ is affine transformed:

$$\mathbf{F}_t' = \gamma(\phi)\cdot\mathbf{F}_t + \beta(\phi) \quad (15)$$

(2) Joint training strategy: A two-stage fine-tuning framework is adopted. In the first stage, the parameters of the first 10 layers of UNet are frozen; the 22 layers after are fine-tuned with single physical conditions on the environmentally friendly material subset; mean square error loss is used to optimize noise prediction accuracy. In the second stage, all parameters are unfrozen; multi-physics field joint constraints are applied; a hybrid loss function is used:

$$\begin{aligned}\mathcal{L} = \mathbb{E}_{x_0,\epsilon,t,\phi}&\left[\|\epsilon - \epsilon_\theta(x_t, t, \tau, \phi)\|_2^2\right]\\ &+ \lambda_{\text{phys}}\mathcal{L}_{\text{phys}} + \lambda_{\text{freq}}\mathcal{L}_{\text{freq}}\end{aligned} \quad (16)$$

In formula (16), $\mathcal{L}_{\text{phys}}$ is the physical consistency loss, and the physical property distribution of the image is generated by the partial differential equation residual constraint of the PINN model:

$$\mathcal{L}_{\text{phys}} = \big\|\nabla\cdot(\rho K/\mu\nabla p) + \nabla\cdot\big(\beta\rho^2 K^2|\nabla p|\nabla p\big) - \partial(\rho\epsilon)/\partial t\big\|_2^2 \tag{17}$$

The weight coefficient $\lambda_{\text{phys}} = 0.15$; $\mathcal{L}_{\text{freq}}$ is the frequency domain consistency loss, which calculates the difference in low-frequency components between the generated image and the real sample under the Daubechies-4 wavelet basis.

(3) Dynamic visualization enhancement: During the latent space iteration process, a physical parameter-guided attention mechanism is designed: ϕ is encoded as a key vector $\mathbf{K}_\phi$ and combined with the query vector $\mathbf{Q}_\tau$ of the text embedding τ to calculate the attention weight α and dynamically adjust the text-image alignment direction of the noise prediction:

$$\alpha = \text{Softmax}\left(\frac{\mathbf{Q}_\tau\cdot(\mathbf{K}_\tau + \eta\mathbf{K}_\phi)}{\sqrt{d_k}}\right) \tag{18}$$

In formula (18), η is the physical parameter control coefficient. This mechanism ensures that the pore structure details are enhanced when generating recycled fiber textures, while the sharpness of the specular reflection area is enhanced in the generation of plant-based leather.

(4) Optimization of the inference stage: When using the DDIM sampler (with 200 steps) for inference, a dynamic physical parameter interpolation strategy is employed. In the noise inversion stage, linear interpolation is performed to generate a transition state physical field, driving the continuous evolution of texture details. The process is shown in Figure 2.

PERFORMANCE ACTION–CLOTHING DYNAMIC COUPLING ALGORITHM

The algorithm in this section embeds motion capture data in the Stable Diffusion infrastructure and performs phase control and physical coupling on the noise prediction process by applying key timing parameters such as the performer's step frequency f, rotation angle θ, and acceleration vector $\vec{a}$. First, an action timing encoder $\mathcal{E}_{\text{motion}}(t)$ is constructed to convert the action signal into a time embedding vector $\vec{e}_t$ synchronized with the diffusion step number t, and then applied to the UNet middle layer as a control item, so that the fold direction of the generated texture is synchronized with the performance rhythm. The texture phase is mapped to the channel attention weight of the noise prediction intermediate layer, allowing for dynamic control over the distribution frequency and directionality of texture wrinkles. For the stress–deformation relationship, a response function based on the Mooney–Rivlin hyperelastic model is constructed:

$$W = C_1(\bar{I}_1 - 3) + C_2(\bar{I}_2 - 3) \tag{19}$$

In formula (19), C_1 and C_2 are dynamically adjustable parameters, and their range of variation is controlled by the performance action amplitude $\|\vec{a}\|$ to ensure that the stress–strain relationship is consistent with the actual motion amplitude.

To improve the visual consistency of style and action synchronization, a style–action cross-attention module is applied to fuse the artistic style embedding S with the action encoding $\vec{e}_t$ to generate a style dynamic adjustment vector:

$$\vec{s}_t = \text{Attention}(S, \vec{e}_t) \tag{20}$$

This vector controls the activation of the style channel of the Style Injection layer, ensuring style consistency under intense motion.

Based on the above algorithm framework, an action-driven local physical field refinement mechanism is applied to further enhance the local accuracy of fabric response in highly complex action scenes. First, the angular velocity $\vec{w}_i(t)$ and local acceleration $\vec{a}_i(t)$ of each part of the human body during the performance are collected and mapped to each vertex v_j on the 3D clothing mesh model. The action–vertex mapping relationship is established through the weight function w_{ij}:

$$F_j(t) = \sum_i w_{ij}\cdot\vec{a}_i(t) \tag{21}$$

In formula (21), $F_j(t)$ represents the local inertial driving force acting on vertex v_j, which serves as the driving input of the physical constraint term in PINN to update the physical state of the node.

To achieve the matching of the action cycle and the vibration frequency of the fabric, the phase matching function in the Fourier domain is applied to synchronize the gait main frequency f_0 and the natural vibration frequency f_n of the fabric in the frequency domain:

$$\Delta\phi = \arg\min_\phi \|\mathcal{F}_{\text{motion}}(f_0) - \mathcal{F}_{\text{cloth}}(f_n + \phi)\| \tag{22}$$

In formula (22), $\mathcal{F}_{\text{motion}}$ and $\mathcal{F}_{\text{cloth}}$ are the frequency spectrum representations of the action signal and the cloth response, respectively. The phase compensation term is used to fine-tune the time sampling points of the cloth details in the generation stage, thereby realizing the visual action-leading-cloth response mechanism.

In the model training stage, the action-guided loss function is added:

$$\mathcal{L}_{\text{motion-align}} = \sum_t \left\|\nabla I_t^{\text{gen}} - \nabla I_t^{\text{ref}}(\vec{e}_t)\right\|^2 \tag{23}$$

In formula (23), I_t^{gen} is the generated image frame, and I_t^{ref} is the real reference frame corresponding to the action encoding. The gradient difference is used to constrain the image motion edge to align with the action and suppress asynchronous creases and non-physical deformations.

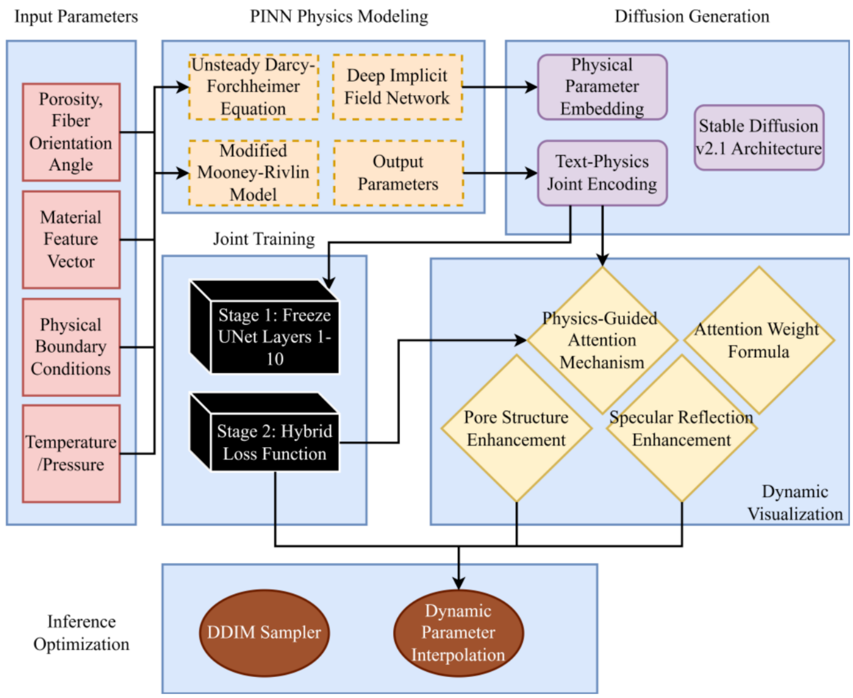

FIGURE 2. PINN–Diffusion Coupling Architecture Flow

COLLABORATIVE OPTIMIZATION OF STAGE SCENES

This study integrates the style transfer module with physical parameters in digital costume performance, dynamically adjusting glossiness, texture, and artistic style in response to stage lighting and music rhythm. Glossiness parameters predicted by the PINN are mapped to reflection coefficients, while rhythm synchronization based on Transformer analysis aligns costume surface features with beat signals. A diffusion model with cross-attention fuses artistic style vectors and lighting information, enabling continuous tonal and gloss variation. Motion response algorithms and physical field coupling further ensure coordination between costume dynamics and the stage environment, as summarized in Table 2.

TABLE 2. Dynamic Mapping of Stage Environment Parameters to Clothing Physical Properties and Visual Effects

Environmental Variable	Adjusted Physical Property	Affected Visual Effect	Control Method
Light Intensity	Glossiness G	Reflection strength, gloss changes	Lighting–material mapping + PINN adjustment
Music Tempo	Vibration Frequency f	Dynamic texture changes	Tempo synchronization + motion interpolation
Motion Amplitude	Elastic Modulus E	Flowing effect, wrinkle depth	Motion–physical field coupling
Style Encoding	Texture Mapping Weight α	Color tone and detail rendering	Cross-attention + style feature library
User Keyword Input	Parameter Adjustment Signal	Style (lightweight)	Semantic–physical embedding module

Table 2 illustrates how stage environment parameters such as light, rhythm, movement, style encoding, and user input influence clothing properties like gloss, vibration, and elasticity. It shows that lighting mapping, movement synchronization, and cross-attention enable real-time regulation of clothing dynamics for coordinated and realistic stage effects.

III. EXPERIMENT AND VERIFICATION

EXPERIMENTAL DESIGN

This study constructs multimodal data, including recycled fiber clothing images and user behaviors, to verify the effectiveness of the proposed AIGC system. The dataset consists of 6,000 clothing images (resolution: 1,024 $\times$ 1,024), covering dresses, coats, and jackets made of recycled polyester, organic cotton, and recycled wool, collected from the 2020–2023 spring/summer collections of brands in Europe (45%), East Asia (35%), and North America (20%). User behavior data was collected through an e-commerce API, consisting of 8,500 users over a 24-month period (2022–2023), with browsing, click, and purchase records included. The Hue–Saturation–Value (HSV) color space correction is used to eliminate lighting differences. Specifically, a global normalization method is used to adjust brightness and color balance across all samples, ensuring consistency in texture representation. Physical property parameters such as air permeability and glossiness are annotated. Glossiness values presented in this study are expressed in normalized arbitrary units (a.u.). User behavior data is obtained through the e-commerce platform Application Programming Interface (API), including browsing time, click sequences, and purchase records, covering user groups from different regions and age groups. The data is divided into a training set, a validation set for virtual catwalk scene performance test verification, and a test set, in a 4:1:1 ratio, to ensure statistical significance. To further validate the independent prediction performance of the PINN model, this study compared its predictions of material permeability and elastic modulus with finite element simulations and experimental measurements. The results are shown in Table 3.

TABLE 3. Comparison of PINN Predictions with Finite Element Simulations and Experimental Measurements

Material / Property	PINN Prediction	Finite Element Simulation	Experimental Measurement	Relative Deviation
Recycled fiber – Permeability ($\times 10^{-12}$ m^2)	2.35	2.28	2.31	$\leq$3.0%
Recycled fiber – Elastic modulus (MPa)	92.8	95.1	94.3	$\leq$2.5%
Plant-based leather – Permeability ($\times 10^{-12}$ m^2)	0.62	0.65	0.64	$\leq$3.2%
Plant-based leather – Elastic modulus (MPa)	118.4	122	120.5	$\leq$3.0%
Plant-based leather – Glossiness (a.u.)	0.72	0.7	0.71	$\leq$2.8%

As shown in Table 3, PINN results are highly consistent with reference values, with relative deviations typically within 2–4%. Specifically, the permeability values predicted by PINN for regenerated fibers and plant-based leather follow the same trends as those obtained from physics-based simulations and laboratory testing. Similarly, elastic modulus and gloss values exhibit consistent accuracy across materials. These results confirm that PINN can provide reliable approximations of key material properties even before integration with a diffusion model, highlighting its independent contribution to the overall framework.

The experimental process is divided into two stages: model training and system verification. In the model training stage, PINN modeling uses some 3D reconstructed images for pre-training and the remaining images for fine-tuning; the Stable Diffusion v2.1 model freezes the parameters of the first 10 layers of UNet and finetunes the last 22 layers with mean square error loss on a subset of environmentally friendly materials; the recommendation algorithm is based on the Transformer architecture and performs end-to-end training on user behavior data. A double-blind test is used in the system verification phase, and fashion designers are invited to score the perceptual quality of generated images using a 5-point Likert scale, while Peak Signal-to- Noise Ratio/Structural Similarity (PSNR/SSIM) metrics are computed programmatically to assess pixel-level differences. The real-time test statistically analyzes the generation delay and recommendation throughput of the model under five hardware configurations: NVIDIA A100, RTX 3090, Tesla V100, NVIDIA RTX 2080 Ti, and GTX 1660 Super.

RECOMMENDATION PERFORMANCE EVALUATION

This study verifies the matching degree of the recommendation system to user needs through a quantitative evaluation of multimodal recommendation algorithms, utilizing Top-5 accuracy and Recall@10 indicators to measure the coverage and accuracy of the recommendation results. Table 4 summarizes the comparison between the model in this study and other methods in terms of Top-5 accuracy, Recall@10, recommendation coverage, and response delay, verifying the core value of physical property perception recommendation.

TABLE 4. Recommended Performance Indicator Evaluation

Model Name	Top-5 Accuracy (%)	Recall@10 (%)	Recommendation Coverage (%)	Response Latency (ms)
AIGC–PINN	89.7	83.2	81.5	85
GMF	68.5	61.2	54.7	130
Conventional Transformer	79.2	74.1	70.3	105

Table 4 compares the proposed AIGC–PINN with GMF and the traditional Transformer in recommendation accuracy, coverage, and response delay. AIGC–PINN achieves 89.7% Top-5 accuracy and 83.2% Recall@10, surpassing GMF and Transformer, with a coverage rate of 81.5% and optimized response delay of 85 ms, meeting real-time fitting requirements. These results reflect the synergistic gain of physical parameter embedding and multimodal feature fusion. The advantage comes from modeling material microstructure with PINN and embedding dynamic physical parameters into the Transformer via cross-modal attention, where physical parameter vectors align with user instructions for adaptive recommendations, balanced by the MMR algorithm. Figure 3 further verifies the mechanism through a heat map of correlations between physical parameters and environmental labels.

Figure 3 shows the heatmap of attention weight distribution between eco-labels and physical parameters. For example, the correlation between organic cotton and elastic modulus reaches 0.777, while plant-based leather is strongly linked to glossiness at 0.710. Carbon neutral certification shows the highest weight with air permeability (0.757). These results demonstrate that the model successfully captures meaningful semantic–physical associations for recommendation.

USER PERCEPTION CONSISTENCY VERIFICATION

This study verifies the perceived consistency of garment physical properties by comparing double-blind evaluations with lab measurements. Thirty designers scored the breathability of recycled fiber and plantbased leather on a 5-point scale across three stages: real samples as benchmarks, anonymous scoring of generated images, and mapping scores to lab data with MSE calculation. A dynamic parameter control mechanism supports real-time interaction. Figure 4 shows the permeability error distribution of AIGC–PINN, GMF, and Transformer, confirming the error suppression effect of physical property perception recommendation.

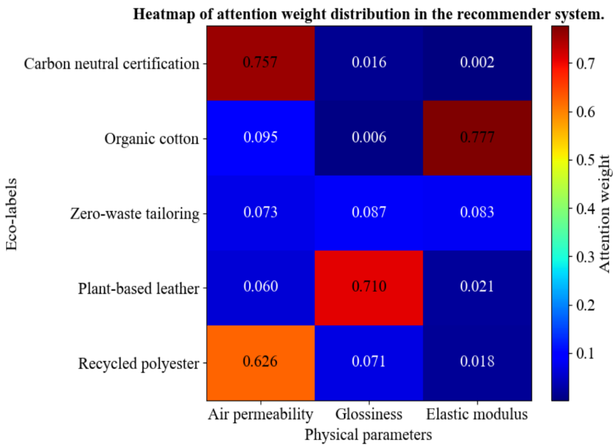

FIGURE 3. Heat Map of Attention Weight Distribution of the Recommendation System

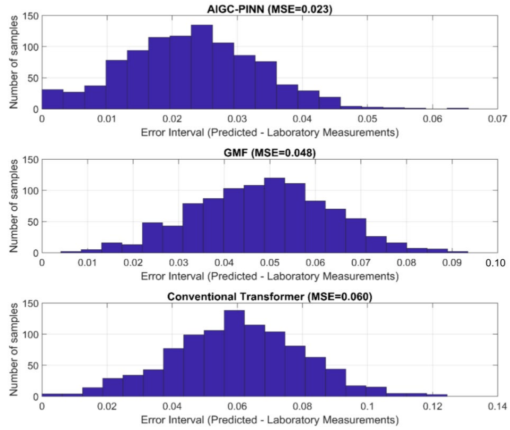

FIGURE 4. Histogram of Air Permeability Error Distribution

Figure 4 presents the permeability error distribution of AIGC–PINN, GMF, and Transformer. AIGC–PINN achieves the lowest error, concentrated at 0.01–0.03 (MSE = 0.023), while GMF extends to 0.03–0.06 (MSE = 0.048) and Transformer to 0.04–0.08 (MSE = 0.060). The superior accuracy of AIGC–PINN results from embedding physical parameters into the diffusion process via a material microstructure database and constraint equations, combined with mixed loss and cross-modal attention to balance novelty and relevance.

SYSTEM RESPONSE EFFICIENCY TEST

This study verifies the real-time performance of the AIGC system through hardware performance comparison and resource optimization strategies. The test indicators include generation delay (in milliseconds) and recommendation throughput (times per second). The experiment covers five typical hardware configurations: NVIDIA A100, RTX 3090, GTX 1660 Super, Tesla V100, and NVIDIA RTX 2080 Ti. The test process is divided into four stages: 1. Generation delay test; 2. Recommendation throughput test; 3. Video memory occupancy test;

4. Response delay test. The performance comparison of different devices is shown in Table 5.

TABLE 5. Performance Comparison of Different Equipment

Hardware Configuration	Generation Delay (ms)	Recommendation Throughput (items/sec)	Memory Usage (GB)	Response Latency (ms)
NVIDIA A100	85	42.3	1.8	47
NVIDIA RTX 3090	112	31.7	2.4	63
Tesla V100	105	28.4	2.1	55
NVIDIA RTX 2080 Ti	118	25.6	3.1	81
GTX 1660 Super	130	21.5	4.2	112

Table 5 presents the performance of AIGC–PINN under five hardware configurations. On NVIDIA A100, the model achieves 85 ms generation delay, 42.3 items/sec throughput, 1.8 GB memory usage, and 47 ms response latency, outperforming other devices. These gains stem from physical property embedding and resource optimization: mixed-precision computing and multi-level caching compress memory and control delay, while stage training jointly optimizes physical constraint and frequency-domain losses. The reported 85 ms reflects full-link latency, including parameter embedding, diffusion inference, and recommendation adaptation.

VIRTUAL CATWALK SCENE PERFORMANCE TEST

Based on the constructed virtual catwalk environment, this section uses the PINN–Diffusion model and the traditional GAN model for comparative testing to verify the physical accuracy and computational performance of the model in dynamic clothing performance. The test selects two typical actions—model turning and fast walking—and utilizes motion capture data to drive the model, generating dynamic texture and deformation effects of clothing. The performance test uses the NVIDIA A100 environment to evaluate the single-frame generation time and fabric drape angle deviation. Inference was conducted with a DDIM sampler, 200 steps, batch size = 1, using FP16 precision and cuDNN 8.5.0. Timing was measured with PyTorch Profiler. Figure 5 shows the performance comparison of different generative models in a 50-frame action in a virtual catwalk scene. Figure 5(a) reflects the delay of traditional GAN and PINN–Diffusion in generating each frame. It can be seen that the average generation time of the GAN model is about 25 ms, and some frames fluctuate, exceeding 28 ms; while PINN–Diffusion is more stable overall, averaging only about 13 ms, with a smaller fluctuation range, all controlled within 15 ms. The 13 ms reported here represents the isolated single-frame rendering latency of the PINN–Diffusion model under virtual catwalk conditions. Figure

5(b) indicates the error deviation angle of each frame of fabric simulation. The GAN model achieves a maximum of approximately 18°, with values above 11°, while the PINN–Diffusion error is concentrated between 8° and 12°, which is better than the traditional method overall. PINN–Diffusion applies physical information such as material elastic modulus and porosity during the generation process, decouples constraints through neural networks, and is driven by motion capture parameters, so that the fabric maintains high physical consistency during rapid deformation.

COMPARISON OF CROSS-MEDIA COMMUNICATION EFFECTIVENESS

To comprehensively evaluate the performance of digital clothing content generated by AIGC in different dissemination paths, this study selects three typical media scenarios for comparative analysis: TikTok (2D), Metaverse Show (3D), and AR Fitting (interactive), and compares indicators such as interaction rate, rendering efficiency, content adaptation rate, user engagement improvement rate, propagation delay fluctuation rate, visual consistency score, and energy efficiency coefficient. Visual consistency scores were obtained through a double-blind user study (5-point Likert scale, 30 participants), while the content adaptation rate was calculated as the ratio of successfully adapted multimodal outputs to the total number of evaluated outputs. The energy efficiency factor is defined as the ratio of rendered frames per joule of GPU power consumption; higher values indicate greater resource utilization. The energy efficiency factor values shown in Figure 6 were calculated based on power consumption measured under the same workload conditions and complement the metrics in Table 4. The results are shown in Figure 6:

Figure 6 compares three media environments in disseminating digital clothing content. TikTok achieves the highest content adaptation rate (88.3%) and visual consistency (4.30); Metaverse Show records the highest user engagement (42%) due to its immersive nature; AR Fitting excels in rendering efficiency (90%) but has a lower adaptation rate (58.4%). AR Fitting also shows the smallest propagation delay fluctuation (12 ms), ensuring smooth interaction. These results reflect that embedding material microstructures and physical constraints via PINN optimizes dynamic texture–scene coordination, while the multimodal recommendation algorithm enhances user participation in Metaverse Show. The low latency further demonstrates that mixedprecision computation and multi-level caching support real-time interactivity.

PERFORMANCE CONTENT QUALITY ASSESSMENT

This study verifies the stage adaptability and style transfer effect of the generated costumes through doubleblind tests and quantitative analysis. The experiment invites costume designers to form an evaluation team and uses a 5-point scoring system to independently score the stage performance of the dynamic texture generation, motion response algorithm, and style transfer module. The evaluation dimensions include dynamic effects and movement synchronization (such as the swing amplitude of the skirt when turning around), the matching degree of material gloss and stage lighting, and the fit between artistic style and performance theme. Figure 7 presents the key results of the performance content quality evaluation.

Figure 7(a) shows that AIGC–PINN outperforms GMF and Transformer in dynamic effects (4.72 ± 0.15), gloss matching (4.65 ± 0.18), and artistic style fit (4.80 ± 0.12). In Figure 7(b), the horizontal axis represents evaluation metrics (Dynamic, Material, Artistic), where Material corresponds to designers' perception of gloss matching and other material-related attributes. Figure 7(b) reveals strong correlations, with air permeability positively linked to dynamic effects (0.83), and a strong correlation is observed between the physical parameter gloss and the perceptual evaluation of gloss matching (correlation coefficient = 0.78). These advantages stem from embedding physical parameters into the diffusion process via a material microstructure database and constraint equations, applying mixed loss to suppress fuzzification, and using a dynamic texture module driven by motion capture to couple fabric deformation with elastic modulus.

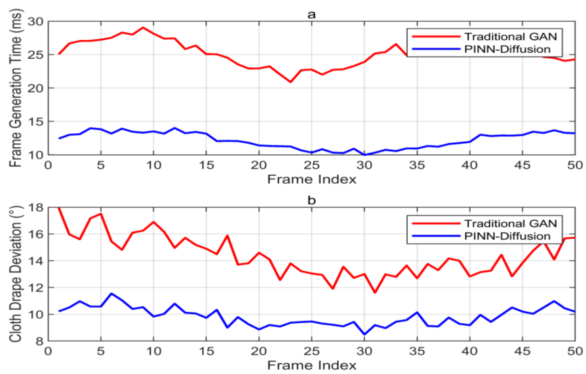

FIGURE 5. Dynamic Comparison of the Performance of Traditional GAN and PINN–Diffusion Models in Virtual Catwalk Scenes:
Figure 5(a) Performance Comparison of Models in Virtual Catwalk Scenes: Single-Frame Generation Time; Figure 5(b) Performance Comparison of Models in Virtual Catwalk Scenes: Fabric Drape Angle Deviation

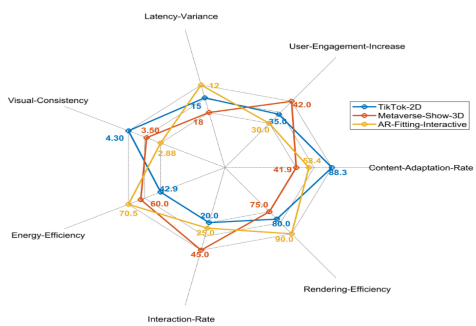

FIGURE 6. Radar Chart Comparing the Performance of Multi-Modal Propagation Paths of Digital Clothing

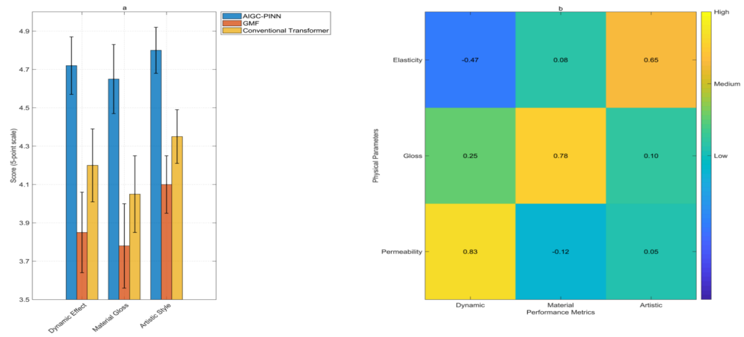

FIGURE 7. Performance Content Quality Assessment: Figure 7(a) Designer Score Comparison Bar Graph; Figure 7(b) Physical Parameter and Score Correlation Heat Map

IV. CONCLUSION

This study explores AIGC-driven digital clothing performance and multimodal communication. By embedding the PINN into the diffusion process, it enables dynamic visualization of regenerated fibers and plant-based leather. Motion capture, style transfer, and motion response algorithms enhance artistic expression and physical authenticity, while a cross-media adaptation mechanism with dynamic texture mapping and lowlatency transmission supports AR fitting, social media, and digital stage broadcasts. Centered on low-carbon materials, the study delivers a full-link solution from generation to dissemination, demonstrating the potential of integrating AIGC with material property modeling for textile visualization and supporting the lowcarbon transformation of the textile industry.

AUTHOR CONTRIBUTIONS

All work in this study was independently completed by Lizhen SHEN.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] W. Z. Wang, H. M. Xiao, and Y. Fang, “Clothing image attribute editing based on generative adversarial network, with reference to an upper garment,” International Journal of Clothing Science and Technology, vol. 36, no. 2, pp. 268-286, 2024, doi: 10.1108/IJCST-09-2023-0129.

[2] Y. Hu, “Research on the Design Method of Traditional Decorative Patterns of Ethnic Minorities under the Trend of AIGC,” Journal of Electronics and Information Science, vol. 8, no. 5, pp. 58-62, 2023, doi: 10.23977/jeis.2023.080509.

[3] M. P. Sikka, A. Sarkar, and S. Garg, “Artificial intelligence (AI) in textile industry operational modernization,” Research Journal of Textile and Apparel, vol. 28, no. 1, pp. 67-83, 2024, doi: 10.1108/RJTA-04-2021-0046.

[4] W. H. Akhtar, C. Watanabe, Y. Tou, and P. Neittaanmäki, “A New Perspective on the Textile and Apparel Industry in the Digital Transformation Era,” Textiles, vol. 2, no. 4, pp. 633-656, 2022, doi: 10.3390/textiles2040037.

[5] S. Kang and J. Chun, “Digitalization of Fashion Shows in the Pandemic Era: A Focus on Fashion Films and Fashion Gamification,” Fashion & Textile Research Journal, vol. 24, no. 1, pp. 29-41, 2022, doi: 10.5805/sfti.2022.24.1.29.

[6] A. S. M. Sayem, “Digital fashion innovations for the real world and metaverse,” International Journal of Fashion Design, Technology and Education, vol. 15, no. 2, pp. 139-141, 2022, doi: 10.1080/17543266.2022.2071139.

[7] S. E. Kim and M. J. Kim, “Analysis of Fashion Brand Cases Using 3D Virtual Clothing Technology - Focusing on Green Design Perspective,” Journal of the Korea Fashion & Costume Design Association, vol. 26, no. 2, pp. 115-127, 2024, doi: 10.30751/KFCDA.2024.26.2.115.

[8] Z. Wang, J. Wang, X. Zeng, S. Sharma, Y. Xing, S. Xu, et al., “Prediction of garment fit level in 3D virtual environment based on artificial neural networks,” Textile Research Journal, vol. 91, no. 15-16, pp. 1713-1731, 2021, doi: 10.1177/0040517520987520.

[9] H. J. Chen, H. H. Shuai, and W. H. Cheng, “A Survey of Artificial Intelligence in Fashion,” IEEE Signal Processing Magazine, vol. 40, no. 3, pp. 64-73, 2023, doi: 10.1109/MSP.2022.3233449.

[10] W. Leal Filho, M. A. P. Dinis, O. Liakh, A. Paço, K. Dennis, F. Shollo, et al., “Reducing the carbon footprint of the textile sector: An overview of impacts and solutions,” Textile Research Journal, vol. 94, no. 15-16, pp. 1798-1814, 2024, doi: 10.1177/00405175241236971.

[11] M. M. Rahman, “Applications of the Digital Technologies in Textile and Fashion Manufacturing Industry,” Technium: Romanian Journal of Applied Sciences and Technology, vol. 3, no. 1, pp. 114-127, 2021, [Online]. Available: https://ssrn.com/abstract=3777742.

[12] A. Gangoda, S. Krasley, and K. Cobb, “AI digitalisation and automation of the apparel industry and human workforce skills,” International Journal of Fashion Design, Technology and Education, vol. 16, no. 3, pp. 319-329, 2023, doi: 10.1080/17543266.2023.2209589.

[13] L. B. Nhelekwa, J. Z. Mollel, and I. W. R. Taifa, “Assessing the digitalisation level of the Tanzanian apparel industry: Industry 4.0 perspectives,” Research Journal of Textile and Apparel, vol. 28, no. 2, pp. 185-205, 2024, doi: 10.1108/RJTA-11-2021-0138.

[14] M. Honauer, “Beyond costume tradition and physical computing: Characterizing the profile of interactive costume creators,” Digital Creativity, vol. 32, no. 2, pp. 99-115, 2021, doi: 10.1080/14626268.2021.1922459.

[15] V. Linfante and C. Pompa, “Space, Time and Catwalks: Fashion Shows as a Multilayered Communication Channel,” ZoneModa Journal, vol. 11, no.1 , pp. 15-42, 2021, doi: 10.6092/issn.2611-0563/13100.

[16] N. Asfand and V. Daukantienė, “A study of the physical properties and bending stiffness of antistatic and antibacterial knitted fabrics,” Textile Research Journal, vol. 92, no. 7-8, pp. 1321-1332, 2022, doi: 10.1177/00405175211055070.

[17] N. Asfand and V. Daukantienė, “Study of the Tensile and Bending Stiffness Behavior of Antistatic and Antibacterial Knitted Fabrics,” Fibres & Textiles in Eastern Europe, vol. 31, no. 3, pp. 37-45, 2023, doi: 10.2478/ftee-2023-0026.

[18] A. Patti and D. Acierno, “Materials, Weaving Parameters, and Tensile Responses of Woven Textiles,” Macromol, vol. 3, no. 3, pp. 665-680, 2023, doi: 10.3390/macromol3030037.

[19] H. Yan, H. Zhang, J. Shi, J. Ma, and X. Xu, “Toward Intelligent Fashion Design: A Texture and Shape Disentangled Generative Adversarial Network,” ACM Transactions on Multimedia Computing, Communications and Applications, vol. 19, no. 3, pp. 1-23, 2023, doi: 10.1145/3567596.

[20] H. Yan, H. Zhang, J. Shi, J. Ma, and X. Xu, “Inspiration Transfer for Intel-

ligent Design: A Generative Adversarial Network with Fashion Attributes Disentanglement," IEEE Transactions on Consumer Electronics, vol. 69, no. 4, pp. 1152-1163, 2023, doi: 10.1109/TCE.2023.3255831.

[21] H. Yan, H. Zhang, J. Shi, and J. Ma, "Texture Brush for Fashion Inspiration Transfer: A Generative Adversarial Network with Heatmap-Guided Semantic Disentanglement," IEEE Transactions on Circuits and Systems for Video Technology, vol. 33, no. 5, pp. 2381-2395, 2023, doi: 10.1109/TCSVT.2022.3224190.

[22] R. Dan and Z. Shi, "Finite element simulation on the relationship between Journal of Clothing Science and Technology, vol. 33, no. 2, pp. 289-301, 2021, doi: 10.1108/ijcst-03-2020-0037.

[23] R. Dan, Z. Liu, and Z. Shi, "Functional relationship among pressure, displacement and angle for the waist of elastic pantyhose using finite element simulation," Textile Research Journal, vol. 93, no. 9-10, pp. 2103-2112, 2022, doi: 10.1177/00405175221135142.

[24] R. Dan and Z. Shi, "Dynamic simulation of the relationship between pressure and displacement for the waist of elastic pantyhose in the walking process using the finite element method," Textile Research Journal, vol. 91, no. 13-14, pp. 1497-1508, 2021, doi: 10.1177/0040517520981741.

[25] R. B. Malabadi, K. P. Kolkar, R. K. Chalannavar, and H. Baijnath, "Plant-based leather production: An update," World Journal of Advanced Engineering Technology and Sciences, vol. 14, no. 01, pp. 031-059, 2025, doi: 10.30574/wjaets.2025.14.1.0648.

[26] S. Mandal and J. Venkatramani, "A review of plant-based natural dyes in leather application with a special focus on color fastness characteristics," Environmental Science and Pollution Research, vol. 30, no. 17, pp. 48769-48777, 2023, doi: 10.1007/s11356-023-26281-1.

[27] Y. Xu, C. Zhi, S. Wang, J. Chen, R. Sun, Z. Dong, et al., "FabricGAN: An enhanced generative adversarial network for data augmentation and improved fabric defect detection," Textile Research Journal, vol. 94, no. 15-16, pp. 1771-1785, 2024, doi: 10.1177/00405175241237479.

[28] V. Werlen, C. Rytka, C. Dransfeld, C. Brauner, and V. Michaud, "A multiscale consolidation model for press molding of hybrid textiles into complex geometries," Polymer Composites, vol. 45, no. 6, pp. 5460-5478, 2024, doi: 10.1002/pc.28139.

[29] W. Li, Z. Wei, Z. Liu, Y. Du, J. Zheng, H. Wang, et al., "Qualitative identification of waste textiles based on nearinfrared spectroscopy and the back propagation artificial neural network," Textile Research Journal, vol. 91, no. 21-22, pp. 2459-2467, 2021, doi: 10.1177/00405175211007516.

[30] A. Zulfiqar, T. Manzoor, M. B. Ijaz, H. H. Nawaz, F. Ahmed, S. Akhtar, et al., "Artificial-Neural-Network-Based Predicted Model for Seam Strength of Five-Pocket Denim Jeans: A Review," Textiles, vol. 4, no. 2, pp. 183-217, 2024, doi: 10.3390/textiles4020012.

[31] L. Xu, F. Li, and S. Chang, "Incorporating convolutions into transformers for textile fiber identification from fabric images," Textile Research Journal, vol. 93, no. 23-24, pp. 5380-5390, 2023, doi: 10.1177/00405175231194797.

[32] Q. Wu, B. Zhu, B. Yong, Y. Wei, X. Jiang, R. Zhou, et al., "ClothGAN: Generation of fashionable Dunhuang clothes using generative adversarial networks," Connection Science, vol. 33, no. 2, pp. 341-358, 2020, doi: 10.1080/09540091.2020.1822780.