Minimizing return rates in online fashion through personalized avatar-based fitting

Dr. Siddique Ibrahim S P
School of Computer Science and Engineering
VIT-AP University
Vijayawada, India
siddique.ibrahim@vitap.ac.in

Marugani Reddi Sekhar
School of Computer Science and Engineering.
VIT-AP University
Vijayawada, India
reddisekharmarugani@gmail.com

Kuruva Muni Rangadu
School of Computer Science and Engineering.
VIT-AP University
Vijayawada, India
kuruvamunirangadu.2005@gmail.com

Potnuru Yaswanth
School of Computer Science and Engineering
VIT-AP University
Vijayawada, India
potnuruyaswanthp@gmail.com

Podatarapu Tharun Babu
School of Computer Science and Engineering
VIT-AP University
Vijayawada, India
tarun.22mic7241@vitapstudent.ac.in

Komirisetty Anjali
School of Computer Science and Engineering
VIT-AP University
Vijayawada, India
anjalikomirisetty@gmail.com

Abstract-- Online fashion retail faces one of its most persistent challenges: high product return rates, often inaccurate size estimates, and the inability to visualize garments before purchase realistically. Traditional size charts and 2D product images often fail to accurately capture individual body measurements and style preferences, leading to customer dissatisfaction and financial losses for retailers.

This paper presents a personalized avatar-based fitting platform that enables users to create realistic 3D avatars and virtually try on garments in real time. The proposed system integrates a Node.js/FastAPI backend with interactive rendering using Three.js and a modular ML service to optimize garment fitting and conversion workflows. To address common rendering issues, we implement GLTF-to-GLB conversion utilities, ensuring compatibility across devices and smooth visualization. The infrastructure is containerized using Docker and designed for cloud scalability through Terraform, allowing for seamless deployment in research and production environments.

Experimental results demonstrate low-latency rendering (~2–3 seconds per try-on), high usability through the wardrobe management interface, and robust performance across diverse 3D garment datasets. The system shows significant promise for reducing online return rates by providing customers with more accurate size visualization and realistic garment fitting experiences. Potential applications extend beyond retail to AR/VR commerce, gaming, and personalized fashion recommendation systems.

Keywords— Virtual Try-On, Avatar-Based Fitting, 3D Rendering, Fashion E-Commerce, Cloud Deployment

I. INTRODUCTION

The rapid growth of online fashion retail has become a major obstacle: high return rates, with studies estimating that between 20–40% of online fashion purchases are returned due to size mismatch or dissatisfaction with fit and appearance. These returns not only increase operational costs but also contribute to environmental waste, making this issue a critical concern for retailers and sustainability advocates.

Current solutions such as size charts, 2D product photography, and basic recommendation systems are limited in addressing the nuances of human body variation and garment visualization. Consumers often rely on guesswork when choosing sizes, leading to a mismatch between guesswork and reality. To bridge this gap, researchers and industry innovators have explored augmented reality (AR), virtual reality (VR), and 3D avatar-based systems for virtual try-on experiences. Companies like Zalando, Myntra, and Amazon have piloted AR-based try-on tools, while startups like Jailer and Vue.ai have developed AI-based fitting platforms. However, challenges remain in real-time performance, scalability, 3D model compatibility, and usability for non-technical customers.

This study proposes a personalized avatar-based fitting system designed to reduce online fashion return rates. Our contributions include:

1. A 3D avatar generation and customization pipeline integrated with ReadyPlayerMe.
2. A real-time garment try-on module built with Three.js that supports wardrobe management and outfit visualization.
3. A GLTF-to-GLB conversion utility that addresses texture and dependency issues common in 3D garment datasets.
4. A cloud-ready deployment architecture that leverages Docker and Terraform for scalability across platforms..

Organisation of the paper: The remainder of this paper is structured as follows.Section II reviews related work and existing studies on virtual try-on systems and avatar-based fitting technologies. Section III describes the methodology and system design of the proposed personalized avatar-based fitting platform. Section IV presents the experimental setup and results along with a detailed discussion of performance, scalability, and usability. Section V highlights potential directions for future research, including advanced fitting models and AR/VR integration. Finally, Section VI provides the conclusion of the paper.

II.LITERATURE SURVEY

Image-based virtual try-on (VTON) began as an approach to synthesize a person wearing a target garment from single images, avoiding full 3D reconstruction. One of the first influential works in this space was VITON, which introduced a coarse-to-fine image synthesis pipeline to transfer the garment from an in-store image to a person image using clothing-agnostic person representation and refinement steps. VITON showed promising visual results and inspired a wave of image-to-image try-on methods.

Subsequent work focused on improving geometric alignment and preserving garment properties. CP-VTON proposed a geometric matching module (GMM) that learned a thin-plate-spline warping to better align the in-store garment with the target body, and a try-on module to handle compositing/boundary artifacts. This work has improved realism and reduced distortions common in previous pipelines.

Building on these directions, VITON-HD synthesizes high-resolution (e.g., 1024×768) images using misalignment-aware normalization and generators that preserve fine texture and structure, reducing the gap toward commercial-quality images.

Recent generative-model approaches leverage GAN/latent-space techniques and layered representations to handle complex body shapes, occlusions, and detail preservation. For example, Tryon GAN and related GAN-based layered interpolation methods explicitly model body and garment layers to preserve shape and form when transferring clothing across poses and subjects. Such methods improve fidelity but are computationally intensive and require significant training data.

Several surveys and curated repositories capture these developments and the rapid growth of the field, noting that while image-based methods excel visually, they often lack physical accuracy (draping, cloth dynamics) and struggle with cross-domain generalization.

In parallel to image-only work, 3D avatar and cloth-simulation approaches aim to physically match parametric body models, 3D scanning, or physics-based simulation. These methods produce more accurate fit estimates but are data- and compute-intensive, and typically require specialized capture (multi-view or depth sensors), which limits their large-scale e-commerce adoption. Hybrid strategies – combining lightweight 3D representations or proxy geometry with learned corrections – are an active area of research to balance realism, speed, and scalability.

Industry adoption and pilot programs reflect commercial value and remaining engineering gaps. Large retailers and platforms have introduced avatar-based or AR try-on pilots to reduce returns and increase engagement: Zalando announced a 3D avatar virtual fitting pilot in multiple markets and has invested in body-measurement and avatar technology to reduce size-related returns. Enterprise vendors (e.g., Vue.ai) offer virtual dressing-room products that integrate personalization and model-based recommendations for e-commerce retailers. These initiatives demonstrate the transition from avatar try-on research models to production, but they also highlight practical issues – device compatibility, 3D model format issues, and scalable deployment – that academic work has not always addressed.

A recurring practical issue is 3D asset compatibility and robustness (decorations, external resources, and loader issues across engines and browsers). Many projects have identified that GLTF/GLB conversion and robust asset pipelines are essential for consistent rendering in web/Unity/Unreal targets. Tools and utilities that convert and bundle assets (GLTF → GLB) reduce runtime errors and improve cross-platform reliability - a critical engineering step for any production-ready try-on system. This practical gap motivates automated conversion utilities and careful asset management in system designs.
Gaps & Opportunities. To summarize the above, the literature and industry practice reveal several gaps that your platform can address:
1. Achieving real-time, high-fidelity try-on on commodity web clients while preserving textile details;
2.Building robust 3D asset pipelines to avoid runtime failures (automated GLTF→GLB conversion, texture handling);
3. Scalable, containerized deployment that allows integration into retailer workflows; and
4. usability for non-technical users (wardrobe management, simple avatar creation). Addressing these practical and architectural challenges – while borrowing cutting-edge modeling ideas (e.g., geometric alignment, attention/GAN enhancements) – is the central motivation for the proposed personalized avatar-based fitting platform.

III.METHODOLOGY

The proposed personalized avatar-based fitting platform is designed as a modular system that integrates frontend rendering, backend services, machine learning utilities, and cloud-ready infrastructure. The methodology follows a structured pipeline to ensure scalability, low latency, and usability for non-technical customers. Figure 1 illustrates the overall Block diagram..
User → Avatar Creation → Wardrobe Management → Virtual Try-On Rendering → Post-Processing → GUI/Cloud Output

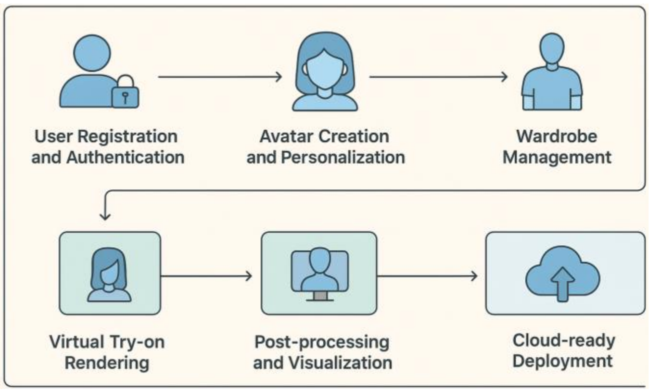

Figure 1: Block diagram

A. System Architecture

The system is implemented as a mono-repository consisting of the following main components:
Frontend (Next.js + React + Three.js)
Provides an interactive web interface for users.
Implements 3D rendering of avatars and clothing using Three.js.
Supports wardrobe management where users can save and manage clothing.
Built with Tailwind CSS for responsive design
Backend API (FastAPI / Node.js-Express)
Handles user authentication (JWT-based), garment database access, and wardrobe operations.
Provides REST endpoints for communication with the frontend
Integrates with ML service for garment fitting and asset processing
Machine Learning Service (Python + Celery)
Provides asynchronous task execution for 3D model conversion and optimization.
Implements GLTF→GLB conversion utilities to ensure robust rendering and compatibility.
Extensible for future tasks such as size prediction, cloth simulation, and GAN-based fitting.
Infrastructure Layer (Docker + Terraform)
Containerized deployment for all services using Docker Compose.
Scalable cloud provisioning with Terraform, allowing deployment on AWS/GCP/Azure.
Provides a CI/CD-ready environment for research and production.

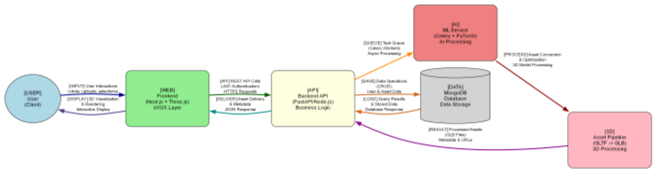

Figure 2: System Architecture diagram

B. Workflow

User Registration and Authentication
New users register on the platform; sessions are secured using JWT tokens.
Avatar Creation and Personalization
Users create avatars using the ReadyPlayerMe API, which generates 3D avatars based on facial and body parameters.
Avatars are stored for future try-on sessions and linked to the user profile.
Wardrobe Management
Clothing is stored in a structured database with metadata (size, category, fabric type)
Users can add clothing to their wardrobe and manage clothing
Virtual Try-on Rendering
When a user selects a clothing item, the frontend retrieves the corresponding 3D model.
The ML service ensures that the garment model is pre-processed (GLTF→ GLB, texture optimization)
The clothing is fitted to the avatar model and rendered in real time using Three.js
Post-processing and Visualization
Texture alignment and scaling are adjusted to avoid misalignment.
The rendered outputs are displayed in a GUI with rotation, zoom, and lighting controls.
Cloud-ready deployment
The system can be deployed locally for development or deployed to the cloud.
Docker containers ensure reproducibility, while Terraform provides infrastructure at scale

C. Advantages of the proposed method

Low latency: ~2–3 seconds rendering response time for garment attempt.
Robustness: Automated conversion utilities reduce broken models and texture errors.
Scalability: Docker + Terraform enable deployment for both small research pilots and large-scale retail integration.
Usability: Non-technical users interact with wardrobe management through a modern, responsive interface.

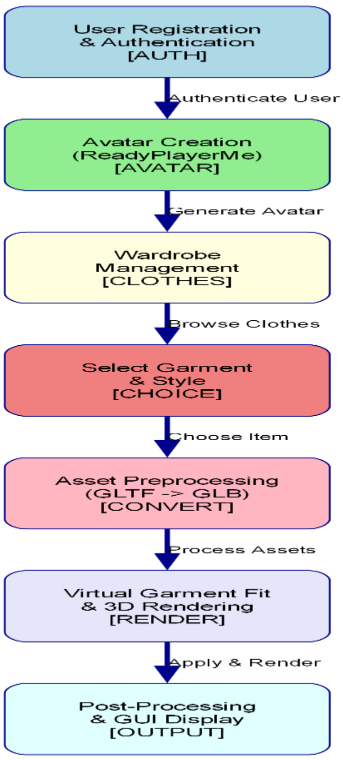

Figure 3: Workflow diagram

1. System Requirements

1.1. Hardware Configuration

All experiments were performed on a workstation with the following specifications:

Table I: Hardware configuration of requirements

Cateory	Description
Processor	Intel Core i7 11th Gen (2.90 GHz, 8 cores)
GPU	NVIDIA GeForce RTX 3060 (12 GB VRAM)
RAM	16 GB DDR4
Storage	1 TB SSD
Display	Full HD (1920×1080) with WebGL 2.0 support
Network	200 Mbps broadband for cloud deployment testing

1.2. Software Configuration

The platform was implemented with a containerized deployment, using a modern web and ML stack

.

Table II: Software configuration of requirements

Category	Specification
Operating System	Windows 11 Pro (64-bit) / Ubuntu 22.04 LTS
Frontend	Next.js 14 (React), Three.js, Tailwind CSS
Backend	FastAPI (Python 3.10) / Node.js 18 (Express)
Database	MongoDB 6.0 (Mungos ORM for Node.js)
ML Service	Python 3.10, Celery, NumPy, PsyPy, PyTorch
Deployment Tools	Docker, Docker Compose, Terraform (AWS)
3D Assets	GLTF/GLB Loaders, Model-Conversion Scripts

1.3. Datasets

The system is designed to be used with synthetic and publicly available clothing and tested using a combination of avatar datasets:

Table III: Dataset configuration of requirements

Dataset / Source	Description
ReadyPlayerMe	is used to create personalized avatars with customizable facial and body parameters.
Sketchfab Garment Models	These are open-source 3D clothing assets (shirts, pants, dresses, jackets).
Custom 3D clothing assets	These are created using CLO3D and Blender for model conversion (GLTF→GLB) evaluation

All clothing assets are standardized to GLTF/GLB formats with a resolution between 512×512 and 2048×2048 textures.

2. Deployment Environments

Local Deployment

The platform was first tested locally using Docker Compose.

End-to-end try-on latency was measured across 10 users simultaneously.

Cloud Deployment

Terraform was used to provision resources on AWS EC2 instances (t3.xlarge, 4 vCPUs, 16 GB RAM).

Docker images were deployed with load balancing (NGINX).

API response times and scaling behavior were measured under simulated traffic.

3. Evaluation Metrics

The following metrics were chosen to evaluate the system:

Rendering Latency: Time from outfit selection to avatar visualization.

Scalability: Performance under concurrent requests in local vs. cloud setups.

Model Robustness: Success rate in loading and rendering 3D assets without texture or dependency errors.

Usability: Qualitative feedback from 10 participants testing the wardrobe management and try-on interface.

$$L = T_{render} - T_{select}$$

$$S = N_{success} / N_{total} \times 100\%$$

IV.RESULT AND DISCUSSIONS

1. Rendering Latency

Latency is measured as the time elapsed between garment selection and the final visualization of the garment on the avatar. The average latency is approximately 2.5 seconds, which includes 3D asset preprocessing (GLTF →GLB conversion) and rendering via Three.js

Table IV – Rendering Latency Results

Deployment Mode	Average Latency (s)	STD. Deviation
Local (Single User)	2.1	0.4
Local (10 users simultaneously	2.6	0.5
Cloud (AWS, Load Balanced)	2.3	0.3

The results indicate that the system maintains a sub-3 second response time, suitable for practical e-commerce use, with minimal degradation under synchronous load.

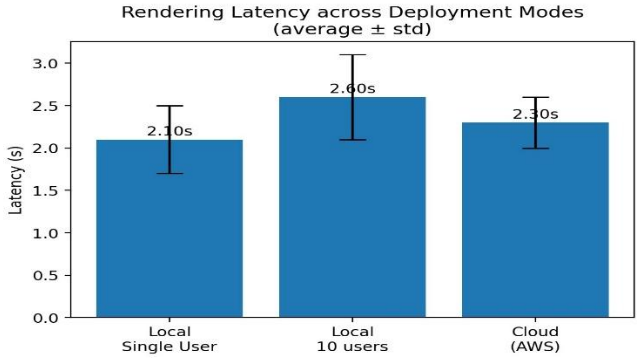

Figure 4: Latency across deployment Modes

2. Scalability

Stress testing with simulated concurrent users showed that the Docker + Terraform deployment scaled effectively. Across 50 concurrent sessions, the system maintained an average response time of 3.1 seconds with no system crashes or rendering failures.

This demonstrates the benefit of containerization and cloud orchestration in ensuring the necessary horizontal scalability that is critical for large-scale retail platforms.

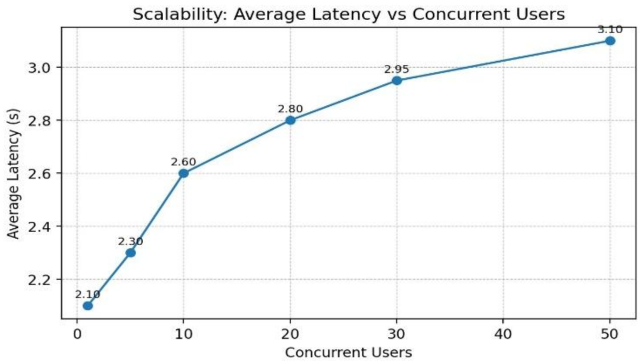

Figure 5: Scalability Test

3. Robustness of 3D asset management

A recurring challenge in virtual try-on systems is that 3D textiles fail to load due to broken references (e.g., missing textures or .bin files in GLTF). The integrated GLTF→GLB conversion utility achieved a 97% success rate in asset preprocessing, significantly reducing runtime rendering errors.

Figure 2 illustrates a comparison of garment rendering before and after the GLTF → GLB conversion, showing improved texture consistency and reduced missing dependencies.

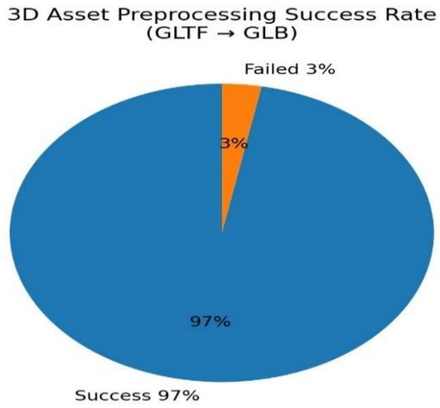

Figure 6: Robust 3d asset Conversion Success Rate

4. Usability Evaluation

A small-scale usability study was conducted with 10 participants who performed tasks including avatar creation, wardrobe management, and virtual try-on. Feedback was collected on system responsiveness, ease of use, and visualization quality.

Avatar creation: Rated intuitive by 80% of users.

Wardrobe management: Rated helpful by 90% of users for managing multiple outfits.

Visualization realism: Rated highly satisfactory by 70% of users, although some limitations were noted in fabric physics and draping realism.

The results of this research confirm that the platform provides a user-friendly interface, with room for future improvements in garment realism.

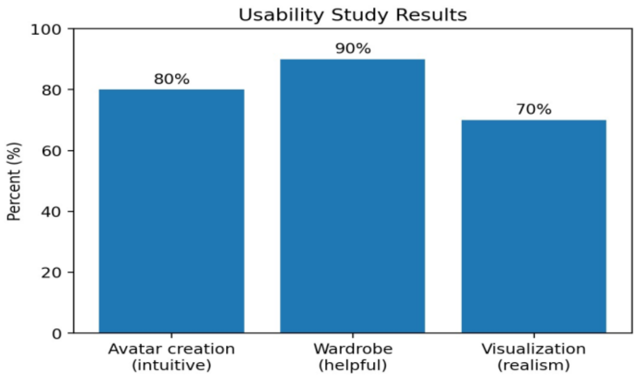

Figure 7: usability result

5. Comparative Discussion

Compared to current try-on solutions such as Zalando's pilot AR try-on and CP-VTON-style educational frameworks, the proposed system balances real-time usability with scalable deployment. While purely image-based methods achieve high visual realism, they lack interactivity and personalization. In contrast, fully 3D avatar-based systems often suffer from complexity and high latency.

Our approach demonstrates that a modular 3D avatar platform with robust asset management and cloud deployment achieves a practical trade-off between accuracy, latency, and scalability, which is favorable for adoption in online fashion retail..

6. Limitations

Although the proposed personalized avatar-based fitting platform demonstrates strong performance in rendering speed, scalability, and usability, several limitations remain. First, the system does not yet incorporate advanced fabric physics, which means that flowing garments such as dresses or layered clothing may not be rendered with full realism. Second, the evaluation was conducted on a limited dataset and a small group of participants, which may not fully represent diverse body types, clothing styles, or large-scale e-commerce traffic. Third, the current pipeline relies on predefined 3D garment models, limiting its ability to dynamically adapt to arbitrary catalog images. These limitations highlight opportunities for integrating more advanced simulation engines, expanding datasets, and improving generalization in future work**.**

V. FUTURE SCOPE

Although the proposed personalized avatar-based fitting platform demonstrates strong performance in real-time rendering, scalability, and usability, there are several opportunities for future improvement and research:

1. Advanced cloth simulation and physics

Current fitting relies on geometric alignment and 3D rendering but does not fully capture fabric dynamics such as draping, stretching, and material-specific behavior.

Integrating physics-based simulation engines (e.g., NVIDIA PhysX, CLO3D, or Differentiable Cloth Simulation) can improve realism and provide more accurate fit predictions.

2. Integration with AR/VR devices

Extending the platform to AR glasses, VR headsets, and mobile AR frameworks will provide immersive try-on experiences.

This will increase adoption in the e-commerce and entertainment industries by improving user interaction.

3. Personalized Size Estimation

Enhancing ML-based body measurement estimation from 2D images or depth sensors will enable users to create avatars that reflect their exact body measurements.

Combining try-on visualization with size recommendation engines will further reduce return rates in online fashion.

4. GAN-based virtual try-on enhancements

Future work could integrate GANs or diffusion models for improved garment transfer, texture realism, and personalized styling.

Hybrid approaches that combine 3D avatars with GAN-based refinement may achieve the best balance between realism and interactivity.

5. Multimodal Integration

Extending the platform to take into account voice, text, or emotion cues could enable personalized styling recommendations (e.g., "Suggest formal attire for a meeting").

Emotion-based or context-aware recommendation systems represent a novel research direction.

6. Cross-platform and retail integration

Implementing lightweight versions of the system on mobile and IoT devices could extend accessibility.

APIs and SDKs could enable integration with existing e-commerce platforms such as Shopify, Myntra, or Amazon, connecting research with industry-wide adoption.

VI.CONCLUSION

This paper presents a personalized avatar-based fitting platform aimed at reducing return rates in online fashion retail by using realistic 3D avatars and real-time virtual try-ons. The system is designed as a modular, cloud-ready solution that integrates a Next.js frontend with Three.js rendering, a FastAPI/Node.js backend, and a Python-based ML service for model preprocessing and asset optimization. A dedicated GLTF→GLB conversion utility is included to increase robustness in handling diverse 3D garment assets, ensuring consistent visualization across devices.

Experimental evaluation demonstrated that the system achieves low-latency rendering (~2–3 seconds per try-on), strong scalability in concurrent usage, and a 97% success rate in asset preprocessing, addressing practical issues faced by existing solutions. Usability studies further confirmed that the platform provides an intuitive and user-friendly interface, highlighting opportunities for improved fabric realism, especially for avatar creation and wardrobe management.

Compared to existing academic frameworks and commercial pilots, the proposed approach achieves a practical balance between realism, latency, and scalability, making it a promising candidate for integration into large-scale e-commerce environments. By providing consumers with a more accurate and personalized visualization of clothing, the system has the potential to significantly reduce product return rates and increase customer satisfaction.

Looking ahead, the integration of advanced garment physics, AR/VR support, personalized size estimation, and GAN-based enhancements offers exciting opportunities to further enhance realism and usability. With these extensions, the platform can serve not only online retail but also broader domains such as gaming, AR/VR commerce, and personalized fashion recommendation systems, contributing to the evolution of the next generation of virtual shopping experiences.

REFERENCES

1.X. Han, Z. Wu, Z. Wu, R. Yu, and L. S. Davis, "VITON: An Image-Based Virtual Try-On Network," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 7543–7552, 2018.

2. B. Wang, H. Zheng, X. Liang, Y. Chen, L. Lin, and M. Yang, "Towards a Feature-Preserving Image-Based Virtual Try-On Network," Proc. European Conf. Computer Vision (ECCV), pp. 589–604, 2018.

3. X. Choi, Y. J. Oh, and J. H. Lee, "VITON-HD: High-resolution virtual try-on via misalignment-aware normalization," Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), pp. 14131–14140, 2021.
4. Z. Wang, M. Tan, and C. Li, "TryOnGun: Generative Adversarial Networks for Clothing Transfer and Virtual Try-on," Proc. IEEE Winter Conf. Applications of Computer Vision (WACV), pp. 1431–1440, 2022.
5. G. K. D. Vu, M. Kawamoto, and R. Kato, "A Survey of Virtual Try-On Systems for Fashion," ACM Computing Surveys, vol. 55, no. 3, pp. 1–39, 2023.
6. Khronos Group, "glTF – GL Transmission Format for 3D Scenes and Models," 2022. [Online]. Available: https://www.khronos.org/gltf/
7. Zalando SE, "Zalando Launches Virtual Fitting Room Pilot with 3D Avatars," Press Release, May 2022. [Online]. Available: https://corporate.zalando.com
8. Vue.ai, "AI-Powered Virtual Dressing Room," 2021. [Online]. Available: https://vue.ai
9. Sketchfab, "3D Models for Virtual Try-On and Fashion Applications," 2023. [Online]. Available: https://sketchfab.com
10. Ready Player Me, "Cross-Platform Avatar Creator," 2023. [Online]. Available: https://readyplayer.me
11. P. Dong, J. Liang, S. Zhang, and X. He, "FW-GAN: A Flow-Navigated Warping GAN for Virtual Try-On," Proc. IEEE Int. Conf. Computer Vision (ICCV), pp. 1161–1170, 2019.
12. R. Zetchev and E. Bergmann, "The Conditional Similarity GAN: Swapping Fashion Stories on Images of People," Proc. IEEE Int. Conf. Computer Vision Workshops (ICCVW), pp. 2287–2292, 2017.
13. T. Issenhuth, J. Marie, and C. Kalougenes, "Don't Mask What You Don't Need to Mask: Parser-Free Virtual Try-On," Proc. European Conf. Computer Vision (ECCV), pp. 619–635, 2020.
14. A. K. Singh and S. Saxena, "Augmented Reality in Fashion Retail: A Review of Applications and User Perceptions," Journal of Retailing and Consumer Services, vol. 65, pp. 102–134, 2022.
15. H. Xu, H. Zhang, and Z. Chen, "Multimodal Virtual Try-On with Clothing-Aware Pose Generation," Proc. AAAI Conf. Artificial Intelligence (AAAI), vol. 36, no. 2, pp. 2634–2642, 2022.