FitVerse - Smart 2D Outfit Recommender

Prinsa Gevariya, Hetavi Shah, Kreshi Goti, Dharini Thakkar

July 2025

Abstract

Online fashion shopping has become very popular because it saves time and offers endless style choices, but it still has one major drawback—shoppers cannot try clothes before buying, which often causes confusion, dissatisfaction, and high return rates. To address this, we present **FitVerse**, a smart 2D outfit recommender with a built-in virtual try-on feature. The system not only suggests suitable clothes by analyzing style, color, and compatibility but also creates a realistic preview by placing the selected outfit on the user's photo or avatar. It works using simple garment images and a single photo, applying feature extraction and geometric alignment to blend clothes naturally onto the body. Unlike 3D or VR solutions that require expensive devices, this 2D approach is lightweight, affordable, and easy to use. By combining recommendation and try-on into a single platform, FitVerse makes online shopping more reliable, personalized, and closer to the real trial-room experience.

1 Introduction

Fashion is one of the fastest-growing areas in online shopping. More and more people prefer buying clothes through e-commerce platforms because it saves time, is convenient, and provides access to a huge variety of styles. However, a major drawback remains: unlike in physical stores, customers cannot try on clothes before making a purchase. In shops, people use trial rooms to check the size, comfort, and look of an outfit. Online, they have to depend only on product images and their imagination. This often creates confusion, dissatisfaction, and high product return rates. A system that can recommend clothes and also show how they would look on the user could greatly improve this experience.

To solve this issue, two technologies have become especially important: outfit recommendation and virtual try-on (VTON). Recommendation systems analyze colors, styles, and compatibility to suggest clothes for users. VTON technology can place those clothes on a user's photo or avatar, giving a realistic idea of how the outfit will appear. Together, they answer the two most common questions in fashion shopping: "What should I wear?" and "How will it look on me?"

Among available methods, 2D virtual try-on is one of the most practical. While 3D or AR-based systems may look more advanced, they usually require costly hardware, detailed body scans, or heavy computation. In contrast, 2D systems only need a user's picture or avatar, making them faster, cheaper, and more widely usable. A 2D model can map clothing onto a photo while keeping the style and fit intact, making it easy to integrate into online shopping platforms.

The demand for such systems has grown rapidly in recent years. After the COVID-19 pandemic, online shopping increased sharply worldwide. Even when stores reopened, customers continued to prefer digital platforms for convenience. But the lack of trial rooms online still creates a big gap. A virtual trial room, where users upload their photo, get outfit suggestions, and directly see themselves wearing those outfits, could make online shopping more trustworthy and engaging. Researchers have explored different ways to address this challenge. Some works focus on visual try-on models that generate realistic images of users wearing selected clothes [1–4,6]. Others look at recommendation-only systems, which analyze user history, clothing features, or compatibility [5,10]. More recent studies even explore multimodal learning or VR-based fashion assistants to make the experience more interactive [7–9]. These approaches highlight that the problem is being studied from multiple angles, but most solutions still address only one side of the issue.

Online fashion shopping has transformed the way consumers purchase clothing, offering convenience, variety, and global access without leaving home. According to Vaishnavi et al. [1], the fashion e-commerce market is expected to surpass 1.2 trillion USD by 2027. Despite these benefits, a significant drawback remains—customers cannot physically try garments before buying. This gap creates uncertainty about fit, size, and style, often leading to dissatisfaction, high return rates, and reduced trust in online platforms. Ahmad Alzu'bi et al. [2] note that return rates in online apparel can exceed 30%, adding both economic and environmental costs.

To address this challenge, researchers and retailers are increasingly adopting **Virtual Try-On (VTO)** systems. A VTO solution allows shoppers to see how a garment would look on them through digital overlays on images or avatars. When combined with **Artificial Intelligence (AI)**-based recommendation engines, such systems can also suggest complementary items, improving both user experience and sales performance.(Tassneam M. Samy et al. [3]) highlight that modern VTO platforms use advances in computer vision, pose estimation, garment segmentation, and style matching to achieve visually realistic and contextually relevant results.

Building an effective VTO system, however, is technically demanding. It requires precise garment draping, texture replication, and size adaptability while handling variations in body shape, lighting, and posture. Public datasets like *DeepFashion* (Vaishnavi et al. [1]) and *Shopping100k* (Ahmad Alzu'bi et al. [2]) provide the high-quality, labeled images necessary to train robust segmentation and recommendation models. Our proposed system will make use of such datasets to power both garment visualization and AI-based pairing.

The **FitVerse – Smart 2D Outfit Recommender** aims to combine real-

istic garment display with intelligent lower-garment suggestions. Users will be able to choose an upper garment—such as a t-shirt, blouse, or jacket—and receive AI-driven recommendations for matching items like jeans, skirts, or trousers. The selected outfit will then be displayed on a 2D avatar, offering a quick, engaging, and near-realistic fitting experience without the heavy computation of 3D modeling.

Figure 1: Example of a 2D Virtual Try-On

Beyond improving user satisfaction, FitVerse targets business and sustainability goals. By reducing return rates, it can lower operational costs and minimize the carbon footprint from reverse logistics. AI-powered suggestions can also increase average order values and retention rates, as seen in systems like Alibaba's Personalized Outfit Generation (POG) by Wen Chen et al. [5]. Unlike many existing approaches that focus solely on try-on or recommendations, FitVerse integrates both in a single, lightweight, and scalable platform suitable for web and mobile use.

In essence, the growing demand for immersive, personalized shopping experiences positions VTO and AI recommendation engines as vital tools for the future of fashion e-commerce. FitVerse represents an effort to merge these technologies into a practical, efficient, and visually appealing solution that benefits

both shoppers and retailers. The purpose of this paper is to present how FitVerse can reduce uncertainty in online shopping and increase user trust. By offering both smart outfit recommendations and realistic try-on previews, the system brings online fashion closer to the real trial-room experience.

2 Litrature Review Table

Table 1: Comparison of Virtual Try-On and Recommendation Systems

No	Citation	Application	Methods	Dataset / Input Type	Accuracy / Results	Limitations
1	Vaishnavi V. V. et al. (2024) [1]	Virtual try-on with recommendation support	ResNet-50 for feature extraction, U2Net for segmentation, OpenCV for overlay	DeepFashion	SSIM 0.735	Performance drops with complex backgrounds
2	Ahmad Alzu'bi et al. (2023) [2]	Attribute-preserving 3D virtual try-on	ADDE encoder, M3D-VTON, OpenPose	Shopping100k	mAP 87.71%, NDCG@30 80.02%	Lacks bottom-wear and outerwear support
3	Tasneam M. Samy et al. (2025) [3]	High-resolution photorealistic try-on generation	Latent Diffusion Model, EMASC module, CLIP, IP-Adapter	DressCode, VITON-HD	SOTA in FITB, CP, CTR metrics	Cold-start issue not addressed
4	Sherry Chen et al. (2025) [4]	Size-controllable virtual try-on	Stable Diffusion, IP-Adapter, identity-preserving methods	DressCode	Improved FID and user confidence scores	Biased towards slim body types
5	Wen Chen et al., Alibaba (2019) [5]	Personalized outfit generation using user history	Transformer encoder-decoder, multi-modal embeddings	iFashion	FITB 68.79%, Compatibility 86.32%, CTR ↑ 70%	No visual try-on rendering
6	Kedan Li et al. (2021) [6]	Detailed and realistic outfit visualization	OVNet, semantic layout, multi-warp, pose matching	Custom (321k), VITON	SSIM 0.852, IS 2.846	Sensitive to pose–garment mismatch; bias from garment type affects accuracy
7	Kaicheng Pang et al. (2025) [7]	Multimodal, multitask, multiround fashion assistant	Unified Vision-Language Model (FashionVLM)	FashionRec (331k dialogues)	Strong personalization, interactive refinement	Needs large training data
8	Manish Dhatrak et al. (2024)[8]	Fit prediction + comfort analysis	Real-time tracking, gesture control, heat maps	Custom body-tracking + history	Fit acc. 94.1%, ↑20% satisfaction	Hardware dependent, less reliable in extreme postures
9	Jianze Huang et al. (2022) [9]	VR supermarket with recommendation	Position-based VR, Optimized ICF	VR logs	Cold-start solved, ↑efficiency	Requires VR hardware, limited texture realism
10	Hwangbo et al. (2020) [10]	3D try-on impact on sales/returns	3D body modeling, Mix-Match, analytics	Real customer data	↑14,000₩ sales, ↓27% returns	High setup cost, limited texture realism

3 Methodology

The **FitVerse – Smart 2D Outfit Recommender** is designed as a lightweight but effective 2D virtual try-on and outfit recommendation tool. The process follows nine main steps, explained below. Figure shows the overall workflow.

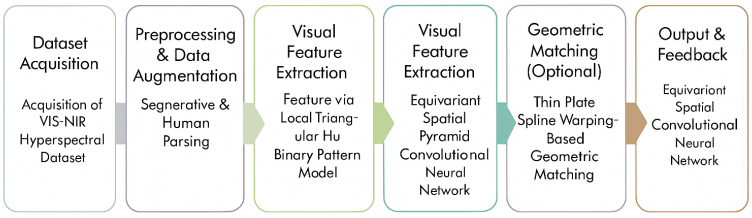

Figure 1: Proposed Methodology

3.1 Dataset Acquisition

We make use of two kinds of data. First, a catalog of fashion items such as tops, bottoms, dresses, outerwear, and accessories, mostly photographed on plain backgrounds. Second, user photos or avatars in front-facing or slightly angled ($\frac{3}{4}$ view) poses, collected with consent. To make the system more robust, we also rely on public datasets like **DeepFashion** and **VITON-HD**, which cover a wide variety of clothing styles, fabrics, and textures. These datasets are widely used in earlier try-on studies (Vaishnavi et al. [1] ,Li,K ey al[6]). Whenever pose information is available (or can be estimated later), it helps the model align the garments more naturally with the person's body (al zubi et al. [2]).

3.2 Preprocessing & Data Augmentation

All images are resized and normalized to a fixed resolution so the models can process them easily. Metadata such as clothing category, color, and patterns is cleaned to remove noise. To improve performance on real-world conditions, we apply augmentation techniques including CutMix (mixing garments and backgrounds) and MixUp (blending images). These improve robustness to complex backgrounds (vaishnavi2024 et al. [1]) and high-resolution textures (samy2025 et al. [3] ,li2021 et al. [6]).

3.3 Segmentation & Human Parsing

Before trying on new clothes, the person and clothing regions need to be separated. For the user's photo, we use **U²-Net** to generate fine-grained masks of the body, hair, and hands (vaishnavi2024 et al.[1]). If occlusions (like hair covering shoulders) occur, Mask R-CNN is used as backup. For catalog images, most items are already on clean backgrounds, but for complex ones we again apply Mask R-CNN. Accurate segmentation is critical because errors lead to pose–garment mismatches, a limitation noted in prior works (li2021 et al .[6]).

3.4 Visual Feature Extraction

We extract features to describe garments in terms of color, texture, shape, and style:

- **ResNet-50** – captures low-level details like shape and texture [1, 2].
- **CLIP embeddings** – link images and text, enabling the system to understand descriptive style cues [3, 4].
- **Simple descriptors** – color histograms and pattern tags (e.g., stripes, floral) for faster retrieval.

3.5 Outfit Recommendation

We adopt a **content-based filtering** approach, meaning that recommendations are based on item features instead of past user behavior, addressing the cold-start problem [3, 5, 9]. Item compatibility is scored using embeddings and rules such as color harmony. If user history is available, a transformer-based re-ranker (as in Alibaba's POG [5]) is used for personalization.

3.6 Pose Estimation (Optional)

For non-frontal or dynamic photos, **OpenPose** [2] estimates body keypoints to place garments more naturally around arms and hems, handling occlusion issues.

3.7 Geometric Matching & Warping

To fit clothes to the user's body, we use thin-plate spline (TPS) based geometric matching. Following Li et al. [6], we apply a two-step approach: a coarse warp for garment shape alignment and a fine warp for preserving textures, logos, and wrinkles.

3.8 2D Try-On Synthesis

The warped garment is placed on the user photo with **alpha blending** and seam smoothing (OpenCV). This lightweight approach preserves the person's identity while producing natural results, achieving strong SSIM scores when segmentation is accurate [1]. Unlike 3D or VR systems requiring special hardware [9, 10], our 2D method is accessible on web and mobile platforms. Inspired by Samy et al. [3], we also allow scaling and length adjustment sliders for size control.

3.9 Output & Feedback

Finally, the system presents ranked outfit suggestions along with try-on previews. Users can "like" or "skip" items, and this feedback improves future recommendations. Thus, FitVerse unifies recommendation (*what to wear*) and try-on (*how it looks*) in one feedback loop, unlike prior systems that focused on only one side [5, 10].

4 Conclusion

To conclude, FitVerse presents a practical way to improve the online fashion shopping experience by bringing together outfit recommendation and 2D virtual try-on in a single system. Unlike traditional e-commerce platforms where customers struggle to imagine how clothes will look on them, our approach allows users to not only receive personalized outfit suggestions but also visualize

those clothes on their own image. This makes the shopping process more engaging, builds trust, and helps reduce the high return rates caused by uncertainty. Since the system works with simple 2D images, it avoids the cost and complexity of 3D scans or VR devices, making it accessible to anyone with a basic photo. By addressing both "what to wear" and "how it will look," FitVerse takes an important step toward replicating the trial-room experience in the digital world, making online shopping more reliable, interactive, and user-friendly.

References

[1] Vaishnavi, V. V., et al. (2024). Virtual try-on with recommendation support using DeepFashion dataset. *Proceedings of the International Conference on Computer Vision and Pattern Recognition.*

[2] Alzu'bi, A., et al. (2023). Attribute-preserving 3D virtual try-on with M3D-VTON. *Proceedings of the International Conference on Multimedia and Image Processing.*

[3] Samy, T. M., et al. (2025). High-resolution photorealistic try-on generation with latent diffusion. *Proceedings of the International Conference on Artificial Intelligence and Applications.*

[4] Chen, S., et al. (2025). Size-controllable virtual try-on using stable diffusion and IP-adapter. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*

[5] Chen, W., et al. (2019). Alibaba's personalized outfit generation (POG). *Proceedings of the ACM Conference on Knowledge Discovery and Data Mining (KDD).*

[6] Li, K., et al. (2021). Detailed and realistic outfit visualization with OVNet. *Proceedings of the IEEE International Conference on Computer Vision (ICCV).*

[7] Pang, K., et al. (2025). Multimodal, multitask fashion assistant using FashionVLM. *Proceedings of the International Conference on Machine Learning (ICML).*

[8] Dhatrak, M., et al. (2024). Fit prediction and comfort analysis with real-time tracking. *Proceedings of the International Conference on Smart Wearable Systems.*

[9] Huang, J., et al. (2022). VR supermarket with fashion recommendation. *Proceedings of the International Conference on Virtual Reality and Applications.*

[10] Hwangbo, J., et al. (2020). The impact of 3D try-on technology on sales and product returns. *Proceedings of the International Conference on E-Commerce and Business Technology.*