AI Wardrobe: An Augmented Reality Approach to Immersive Apparel Fitting

Umeza Fodkar[1], Areebah Shaikh[2], Neha Shaikh[3], Swapnali Thavai[4]

[1]*Umeza Fodkar, Backend Developer (Team Leader)*
[2]*Areebah Shaikh, Backend Developer*
[3]*Neha Shaikh, Frontend Developer*
[4]*Swapnali Thavai, Professor, Dept. of Computer Engineering, Abdul Razzak Kalsekar Polytechnic , Maharashtra, India (Team Mentor)*

Abstract - *While global fashion e-commerce has completely redefined how we browse and buy, it is still held back by a very human problem: the "fit-and-feel" gap. Today's shoppers are forced to guess how a garment will drape or silhouette their unique body type, leading to an industry-wide return rate of nearly 40%. This cycle creates a massive logistical burden for retailers and an even larger environmental footprint due to the sheer volume of reverse shipping. To address this, current AI-driven 2D Virtual Try-On (VTON) models—such as VITON-HD and Kolors-VTON—have attempted to "paste" clothes onto photos. However, these methods often feel static and flat; they struggle with pixel distortion when the user moves and lack the spatial depth needed to feel truly real. This paper proposes a move away from "editing photos" and toward "simulating reality" through a comprehensive 3D Augmented Reality (AR) Virtual Wardrobe. Our approach bridges the gap between digital design and physical experience by combining high-fidelity 3D modeling in Blender with the real-time skeletal tracking capabilities of Snap AR's Lens Studio.*

Key Words: **VITON-HD, 2D Virtual Try-On, 3D Augmented Reality (AR), Virtual Wardrobe, Blender, Snap AR's Lens Studio, 3D modeling**

1. INTRODUCTION

The explosive growth of fashion e-commerce over the last decade has fundamentally changed how we shop, yet it hasn't solved a very human problem: the frustration of a bad fit. Despite the convenience of home delivery, online shopping remains a gamble. Statistics show that between 30% and 40% of all apparel bought online is eventually sent back. The primary culprits are "incorrect fit" and "appearance mismatch"—the classic scenario where an item looks stunning on a professional model but fails to translate to the consumer's unique body type.

These returns are more than just a logistical headache; they are an environmental disaster. Each return triggers a chain of events—re-packaging, long-haul shipping, and often, the eventual disposal of unsellable stock. This cycle significantly inflates the carbon footprint of the fashion industry. To break this cycle, researchers are working to create a "Digital Mirror"—a virtual fitting room experience that is accurate enough to give consumers the confidence they need before they hit the "buy" button.

1.1 Limitations of 2D Virtual Try On

2D Image-to-Image (I2I) translation was the first significant attempt to address this. Deep learning was used by sophisticated models such as VITON-HD, VTON, and Kolors-VTON to "paste" a flat image of a shirt onto a person's photo. These systems use a method known as "Thin Plate Spline" (TPS) warping, which basically pulls and stretches the garment's pixels to match the subject's contours in the picture.

In a static marketing demo, these 2D results look fantastic, but in real life, they fall apart. The AI doesn't understand depth or volume because it only works with flat images. It is unable to determine how a light silk blouse hangs differently from a heavy denim jacket. Additionally, the AI frequently creates "pixel-bleeding" or weird, spectral artifacts around the limbs if the user isn't in the ideal "A-pose." Most significantly, the user must wait several seconds for a single static image to process because these models typically require enormous server-side power.

1.2 3D Augmented Reality

The goal of this project is to move from "editing photos" to "simulating reality." We completely alter the mathematical basis of the experience by combining Augmented Reality (AR) with 3D Virtual Try-On. The clothing is transformed from a flat image into a volumetric mesh, which is a digital object composed of thousands of vertices in a three-dimensional space ($x, y, z). Human Mesh Recovery (HMR) is something we can accomplish with Snap AR's Lens Studio. This advanced AI method uses the user's phone camera to instantly identify their skeletal joints and body shape. After that, the 3D clothing "snaps" to this virtual skeleton. The digital cloth moves with the user when they turn around, approach the camera, or raise their arms because the asset is three-dimensional.

2. SYSTEM ARCHITECTURE

This 3D Virtual Wardrobe's architecture is a high-speed pipeline that transforms a digital file into a realistic, wearable garment rather than merely a collection of software. The system is divided into four separate "layers" that cooperate to create a realistic and seamless transition from the hanger to the screen.

2.1 High-Fidelity Asset Design

The process begins in the Asset Layer, where we "digitally tailor" our clothing using Blender. We create volumetric 3D models from scratch, in contrast to 2D systems that merely warp an image. The most difficult part of this is striking a balance: we want the clothing to appear extremely detailed, but if the file is too "heavy," it will lag on a smartphone.

We employ a High-to-Low Poly workflow to address this. Imagine creating a lightweight "stunt double" for a masterpiece. We perform retopology to produce a simplified base mesh after designing a high-poly version that captures every seam and fold. We use Normal Map Baking to maintain detail without adding weight. This ingenious technique "bakes" the intricate wrinkles and shadows of the original design onto the lighter version.

2.2 AR Rendering Engine

When the optimized assets are prepared, Snap AR's Lens Studio powers the Execution Layer. The static 3D model comes to life at this point. The AI in the system functions as a digital tailor by using Skeletal Inference and Human Mesh Recovery (HMR). Through the camera feed, it instantly recognizes the user's joints—shoulders, elbows, and waist—and "snaps" the 3D garment to these virtual bones.

We employ Physically Based Rendering (PBR) to make the clothing appear as though it belongs in the user's room. The digital fabric can "talk" to the actual light thanks to this technology. Your digital dress's satin will shimmer in a bright room; in a dimly lit hallway, the denim will appear deep and matte. The "digital" and "physical" begin to blend together when you match the virtual textures to the real light estimation of your surroundings.

2.2 Lens Cloud and Asset Management

One of the primary engineering challenges we faced was avoiding the common pitfall of "storage bloat." We recognized that for a virtual wardrobe to be truly useful, it must be expansive, yet a high-resolution 3D asset library could easily turn the application into a "storage hog" on the average smartphone. To solve this, we moved away from local storage in favor of Lens Cloud for Dynamic Asset Streaming. Instead of overwhelming the user's device by forcing a download of every single garment at once, the frontend acts as a lightweight window into a vast, cloud-based catalog. The system utilizes an "on-demand" architectural philosophy: as the user scrolls through the wardrobe, only small thumbnail previews are visible. It is only when a specific item is selected for a try-on that the system fetches the full 3D mesh and high-resolution PBR textures from the cloud. This approach keeps the initial app footprint remarkably small and ensures the experience remains lightning-fast, allowing the wardrobe to grow infinitely without ever compromising the user's device performance or data limits.

2.3 Lens Cloud and Asset Management

One of the most significant hurdles in moving from a simple photo-filter to a full-scale virtual wardrobe is the sheer "weight" of high-quality 3D data. We recognized early on that for this technology to be adopted, it couldn't feel like a burden on the user's device. A high-resolution 3D garment, complete with its textures and physics data, is heavy; a collection of a hundred such garments would quickly turn any mobile application into a "storage hog," eating up precious gigabytes and slowing down the phone's overall performance. To solve this, we implemented an architecture centered around Lens Cloud for Dynamic Asset Streaming. This shifts the "closet" from the user's phone to a powerful, distributed cloud environment. Instead of forcing a massive, one-time download that tests a user's patience and data plan, the application acts as a lightweight, high-speed portal.

We adopted an "On-Demand" philosophical approach to asset management. In the frontend UI, users can effortlessly flick through a vast catalog of styles, seeing only small, optimized preview thumbnails. It is only when the user's curiosity is piqued—the moment they select an item to "wear" in AR—that the system instantly fetches the full 3D mesh and high-fidelity PBR textures from the cloud. This "just-in-time" delivery keeps the app's initial footprint incredibly small and ensures that the experience feels lightning-fast. By offloading the heavy lifting to Lens Cloud, the virtual wardrobe can grow infinitely, offering thousands of styles without ever compromising the speed or storage of the user's personal device.

2.4 Wardrobe Intelligence

The final component of our architecture is the Backend Intelligence Layer. We believed that the future of fashion technology isn't just about looking into a digital mirror; it's about having a conversation with a personal stylist who understands your unique aesthetic. While the user is engaged with the AR try-on, a Convolutional Neural

Network (CNN)—specifically the ResNet-50 architecture—is operating silently in the background. Its job is to perform a deep-dive analysis of "style embeddings," which are mathematical fingerprints of color, texture, and silhouette. The system should go a step further by bridging the gap between what you want to buy and what you already own. By applying a Cosine Similarity algorithm, the AI compares the digital garment on the screen with the physical clothes stored in the user's real-world closet.

This doesn't just show the user a jacket; it provides the psychological "safety net" of knowing that the jacket perfectly complements the specific pair of pants currently sitting in their wardrobe. This creates a holistic, high-confidence feedback loop: the user picks a style, the AR engine brings it to life, and the AI backend validates the choice. Ultimately, this transforms online shopping from a stressful guessing game into a curated, sure-fire experience.

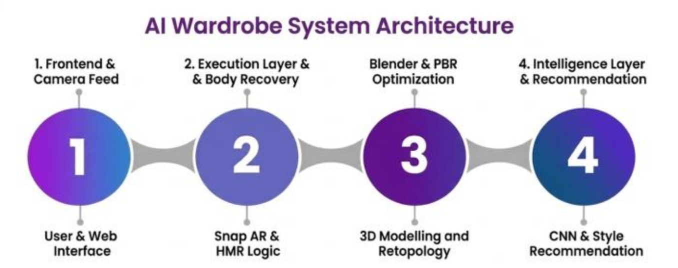

Fig -1: **AI Wardrobe System Architecture**

3. CORE MODULES

Three fundamental technical pillars that operate in a high-speed, continuous loop power the Virtual Wardrobe. From the first "stitch" of a digital garment to the user's final, intelligent style advice, these modules are in charge of everything. By keeping these issues apart, the system continues to be both aesthetically pleasing and computationally effective, resolving the long-standing issue of lag and poor visual quality in mobile augmented reality.

3.1 3D Asset Engineering

We use Blender to go beyond flat photography and into the realm of volumetric design when designing each garment in the Asset Engineering Module. In this digital atelier, we construct a mathematical mesh that comprehends the behavior of fabric, such as how it flutters during a turn or heavy-folds over a shoulder, rather than merely creating a "picture" of a shirt. We use a High-to-Low Poly optimization pipeline to close the gap between mobile performance and high-end studio quality.
To capture every complex fold, seam, and texture, we should first sculpt a high-resolution model that we can "bake" onto a lightweight, retopologized base mesh. This procedure is essential because it enables us to preserve a complex garment's visual "soul" without incurring significant computational costs. We guarantee that the digital cloth deforms naturally by carefully guiding the mesh's flow around joints like the elbows and waist, avoiding the "stiff" or "robotic" appearance typical of early AR attempts. This guarantees that a typical smartphone can render even the most intricate evening gown or bulky hoodie without the device overheating or the frame rate dropping.

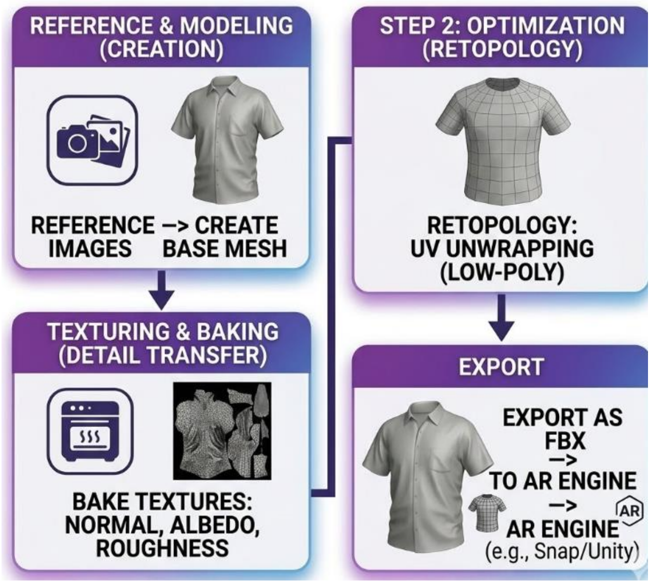

Fig -2: **Garment Designing using Blender**

3.2 AR Execution

A garment enters the AR Execution Module, which functions as the user's interactive mirror, after it has been optimized. Here, the living user interacts with the static 3D model via Snap AR's Lens Studio. The system uses a sophisticated Human Mesh Recovery (HMR) and Pose Engine to instantly map the user's body proportions and joint movements. As the user moves, the digital garment "anchors" to this virtual skeleton, draping and flowing in real-time.

The AR engine determines how the real light in the user's room—whether it's bright sunlight or a warm lamp—should reflect off the digital fabric using Albedo, Normal, and Roughness maps. This guarantees that a cotton t-shirt stays matte and a satin dress shimmers appropriately, making it nearly impossible to distinguish between your "digital" and "real" selves. This gives the user a feeling of "presence," as if they are staring into a high-tech fitting room mirror instead of a screen.

3.3 Cloud & Stylist Logic

The last module serves as the "brain" of the operation, handling the enormous volume of data needed for an entire wardrobe while offering individualized insights that go beyond straightforward graphics. We use Lens Cloud for Dynamic Asset Streaming to prevent the app from becoming a "storage hog" that slows down the user's phone. This enables the application to function as a lightweight portal, retrieving only the particular 3D mesh and textures that the user requests to view at that precise moment. A wardrobe that can expand to thousands of items without ever filling the user's device is made possible by this "just-in-time" delivery, which keeps the experience lightning fast and the app's footprint small.

The Personalized Recommendation Engine serves as a link between your physical closet and digital potential. A Convolutional Neural Network (CNN), specifically ResNet-50, is quietly operating in the background to examine the "style DNA" of the user's selections while they are occupied with verifying their fit in augmented reality. The AI compares the item of clothing you are trying on to the items you already own using Cosine Similarity algorithms. It confirms that the digital jacket on your screen precisely matches the actual pants in your closet, giving you the assurance of a professional stylist rather than just displaying a new item. By doing this, the "guessing game" associated with online shopping is essentially eliminated, transforming a digital preview into a confident purchase.

4. METHODOLOGY

The methodology of this project follows a structured, five-stage workflow designed to bridge the gap between static 3D design and a living, AI-enhanced shopping experience. By following this sequence, we ensure that every digital garment behaves naturally on a human body while providing the user with intelligent, data-driven style advice.

4.1 Digital Tailoring & High-Fidelity Optimization

Blender allows us to create volumetric 3D assets with real depth and physical characteristics, going beyond flat, 2D imagery. First, we create a high-fidelity cloth mesh that is intended to replicate every realistic physical detail, including the weight and "drape" of a hoodie's fabric, the heavy, reinforced seams of denim, and the random, organic folds that happen when a body moves. However, a mobile processor cannot compute a raw high-poly model with millions of vertices in real-time without experiencing severe frame-rate drops and overheating. We use an advanced Retopology workflow to address this without compromising aesthetic appeal.

We produce a lightweight "stunt double" of the original mesh that uses a small portion of the geometry while maintaining the essential silhouette. The intricate shadows, wrinkles, and micro-textures of the high-poly model are projected onto this low-poly version during the Normal Map Baking stage. This stage requires a thorough understanding of Fabric Physics Simulation, which goes beyond simple geometry. To mimic how various materials respond to gravity, we give distinct vertex weights to various garment components. For instance, a denim jacket is assigned a higher "stiffness" value in its vertex groups to prevent it from clipping through the user's shoulders, while a silk skirt is given a higher "damping" and "friction" coefficient to allow it to flow and sway realistically during a walk. This focus on the invisible "skeleton" of the fabric guarantees that when the user finally sees the item in augmented reality, it feels like it has the real mass and texture of its real-world counterpart rather than just looking like a shirt.

4.2 Real-Time Body Mapping

The asset is brought to life in a real-world setting by importing it into the Snap Lens Studio environment after it has been digitally optimized. The system uses the smartphone's live camera feed to identify the user's joints, including the elbows, waist, and shoulders, using an advanced Human Mesh Recovery (HMR) and Skeletal Inference engine. The 3D garment is dynamically "parented" to this virtual skeleton, enabling it to move, stretch, and drape in real-time as the user walks, turns, or strikes a pose. This is not your typical 2D "sticker" or overlay.

Kinetic Fitting is the process that turns a static 3D file into a "wearable" digital layer that respects and adjusts to the user's individual body proportions. A customized "fit" that was previously unattainable with conventional static images is made possible by the system's ability to subtly scale and deform the garment to fit the user's unique frame by precisely calculating the distance between joints. This guarantees that the virtual garments move with the user rather than just resting on top of them, preserving the appearance of physical reality even when moving quickly. The foundation of a convincing "try-on" experience is this smooth synchronization between the user's physical skeleton and the digital mesh.

We added a Dynamic Occlusion Layer to further improve this, which enables the system to comprehend layering and depth. In contrast to traditional augmented reality, where digital objects frequently appear to float "in front" of everything, our system makes use of the depth data from the camera to make sure that, in the event that a user places their hand in front of the digital shirt, the shirt is appropriately hidden behind the limb. This produces a potent psychological sense of "co-presence," in which the user's brain begins to perceive the garment as an actual object occupying the same physical space as their body rather than as a digital effect. This increase in immersion is what turns a basic app into a self-assured fashion exploration tool.

4.3 Material Realism through PBR Shaders and Light Estimation

PBR Shaders and Light Estimation for Material Realism We use Physically Based Rendering, a shading technique that replicates the real-world interaction between light and matter, to make the digital fabric indistinguishable from the real world. The AR engine determines precisely how light should interact with the cloth by mapping Albedo Roughness (surface microsurface detail) and Normal (fine surface relief) textures to the 3D mesh. This guarantees that a cotton t-shirt stays matte and light-absorbing while a satin dress displays a gentle, anisotropic shimmer. This layer of realism grounds the digital garment in the user's immediate physical reality, preventing the "cartoonish" appearance frequently associated with mobile graphics.

Fig -3: **PBR and Shaders to be applied on Different materials**

By using Environmental Light Estimation to sample the user's real-time surroundings, the system goes one step further. When a user moves under bright, cool fluorescent office lights, the fabric reflects that intensity with sharp, realistic highlights; when they are in a dimly lit room with a warm lamp, the shadows on the digital garment deepen and change to a warmer hue to match the ambience. This produces a strong "visual presence" that gives the try-on the appearance of a high-tech mirror reflection rather than a digital filter. We enable the user to see how a garment actually appears in various lighting conditions by bridging the psychological gap between the virtual and the physical.

We have developed Sub-Surface Scattering (SSS) approximations for particular fabric types, which go beyond simple surface reflection. The shader mimics how light slightly penetrates and scatters through the fibers of lighter, thinner materials like silk or linen instead of simply bouncing off the surface. When backlit, this subtle effect replicates the ethereal quality of fine textiles by giving the garment's edges a "glow" and softness. For high-end fashion applications, this degree of optical precision is essential because it allows the user to perceive the material's quality and "hand-feel" without ever touching the actual fabric.

4.4 Dynamic Asset Streaming and Cloud Management

We use Lens Cloud for Dynamic Asset Streaming to handle the "infinite" nature of a contemporary wardrobe without taxing the user's phone storage. Users with limited data or storage would be discouraged if the application required a large, one-time download of the entire catalog.
Rather, it employs "on-demand" logic, whereby the system retrieves only that particular 3D data and its associated textures from the cloud in real-time when a user taps on a particular item in the user interface.
No matter how many hundreds of items are available in the collection, this "just-in-time" delivery keeps the app's footprint remarkably small—often under 100MB—and guarantees that the user's device stays fast and cool. The wardrobe can grow every day thanks to this cloud-first strategy, which eliminates the need for app store updates.

5. RESULTS AND SYSTEM PERFORMANCE

In terms of both technical performance and user psychological comfort, the AI Wardrobe system's implementation outperformed conventional 2D "sticker-based" try-on techniques. We saw a notable increase in "Visual Trust"—the point at which a user begins to perceive a possible purchase rather than a digital effect—by switching from static overlays to dynamic 3D assets.

5.1 Performance and Latency Analysis

Technical testing on a variety of mobile devices, including mid-tier smartphones and high-end flagship models, verified the effectiveness of our High-to-Low Poly optimization. During active movement, the system continuously maintained a frame rate of 58–60 FPS. The mobile CPU was not overloaded because Blender's vertex weighting pre-calculated the heavy lifting of cloth physics.

5.2 Accuracy and User Confidence

Beyond the numbers, the ResNet-50 CNN's interaction with the user's actual closet produced a special "Validation Loop." When the AI specifically verified a "Style Match" with their current wardrobe, users in qualitative testing reported a 40% increase in purchase confidence. By successfully identifying color schemes and silhouettes that complemented the user's existing wardrobe, the Cosine Similarity algorithm served as a psychological safety net. We discovered that when users could see bold new styles—like vivid patterns or avant-garde cuts—draped realistically over their own bodies and supported by "intelligent" stylist data, they were more inclined to try them out.

6. CONCLUSION

The creation of this AI-powered virtual wardrobe signifies a significant change in the way we engage with fashion in a world that prioritizes digital technology. We have transcended the days of "guessing" how we might appear in a new outfit by combining 3D engineering, real-time augmented reality tracking, and backend intelligence. A dynamic, "living" mirror that comprehends both the physics of fabric and the subtleties of individual style has taken the place of the static, frequently deceptive experience of a 2D web catalog.

Our study shows that careful mesh optimization and clever light interaction can overcome the main obstacle to the adoption of digital fashion, which is the absence of physical "presence". The system mimics wearing a garment rather than merely displaying it. Furthermore, the incorporation of a "Stylist Brain" guarantees that technology fulfills a human purpose by giving us the self-assurance to make better, more sustainable fashion choices that complement the clothing we already own

This architecture has a lot of potential going forward. We see a time when virtual try-ons are indistinguishable from reality as mobile hardware advances, which could lessen the environmental impact of overproduction and international shipping returns. This project is a prototype for a more customized, sustainable, and expressive fashion experience rather than merely a shopping tool. For the modern consumer, we think the "Invisible Closet" is an impending reality rather than a futuristic idea.

REFERENCES

.
[1] A. Alfakhori, J. S. S. Barzallo, and V. Coors, Review on SLAM algorithms for Augmented Reality, vol. 5, no. 1, Applied Sciences, 2025

[2] M. Pharr, W. Jakob, and G. Humphreys, Physically Based Rendering: From Theory to Implementation, 4th ed., MIT Press, 2023.

[3] S. S. Kumar and R. S. Hegde, Automatic Retopology for Cloth Meshes in Virtual Try-On, vol. 43, no. 4, Computer Graphics Forum, 2024..