International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Applying Pose-Guided Deep Learning for Real-Time Virtual Try-On

Dr. Chandrika J[1], Khushi K[2], Khalandar Bibi[3], Khushi H M[4], Moulya K A[5]

Department of Computer Science and Engineering
Malnad College of Engineering, Hassan, India

Abstract

The quick evolution of computer vision and ma- chine learning technology has created new prospects for virtual shopping experience improvements. In this paper, an innovative system named "Live Virtual Try-On and Size Prediction System Using Real-Time Pose Detection and 3D Modeling" is presented. The system seeks to emulate the physical trial room experience in digital form by providing users with the capability to view garments on their live 3D body model in real-time, forecast clothing sizes from real body dimensions, and have a digital wardrobe of favored outfitting.

The proposed system includes an end-to-end pipeline that begins with user authentication, driven by a legally safeguarded registration and sign-in system based on MySQL. The system's centerpiece utilizes OpenCV and MediaPipe for real-time camera capture and pose keypoint detection, which form the basis of estimating essential body measurements like shoulder width, waist, chest, and hip ratios. These parameters are then run through a size classification model trained using TensorFlow to predict standard clothing sizes (S, M, L, XL, etc.).

A Generative Adversarial Network (GAN), in this case PI- FuHD, is used to construct a realistic 3D avatar of the user. A Virtual Try-On Network (VITON) is then placed on top, warping and rendering chosen clothes onto the 3D model, giving the user the illusion that they are standing in front of a mirror wearing the chosen garment. Additionally, a user interface that is feature-rich and designed similar to a "Flipkart Lite" e-commerce platform is created with Flask and HTML/CSS to enable browsing of clothes, try-on previews, taking screenshots of desired styles, and saving them in a wishlist database.

This combined pipeline not only improves the user experience but also addresses serious challenges such as size mismatch, absence of tactile feedback, and trust in virtual fitting of clothes. The system maintains legal compliance with user data safety, GDPR-like regulations, and data retention laws.

Keywords: Virtual Try-On; Real-Time Pose Estimation; 3D Human Modeling; Size Prediction; GAN (PIFuHD); MediaPipe; VITON; Augmented Reality; Human-Computer Interaction; Fashion Technology

I. INTRODUCTION

The internet revolution has changed most sectors, and fash- ion was no exception. Conventionally, the process of buying apparel is to try them out physically to ensure fit and com- fort. However, as e-commerce gains momentum, customers have the dilemma of selecting apparel without this physical interaction. This usually results in doubt, bad fit selection, and high return rates, which are expensive for retailers and irritant to consumers. Try-on software is a clever solution to allow customers to virtually try clothing on. Rather than using fixed images or size charts, consumers can observe how the clothing will fit their body type from digital avatars or live camera images. The interactive process makes the buyer more confident in their decision, minimizing guesswork and making online purchasing more enjoyable. Even though it's promising, developing good3D virtual try-on systems is challenging.

Initial implementations primarily relied on straightforward 2D overlays that didn't consider body shape or movement, leading to awkward fits. More sophisticated3D models usu- ally demanded costly hardware such as multiple sensors or cameras, making them unaffordable for many consumers. Size estimation techniques also often depend on manual entry or average charts, which don't accommodate individuals' varying body shape. The latest advances in AI and computer vision make these challenges easier to address. Technologies such as MediaPipe allow realtime body keypoints detection with any standard camera, making pose tracking readily available and inexpensive. Meanwhile, GANs such as PIFuHD can build in- tricate 3D avatars from a single image, eliminating the require- ment for elaborate setups. Such 3D models are more realistic and can be used in conjunction with Virtual Try-On Networks (VITON) to project clothing realistically, maintaining textures and creases. When combined with augmented reality, this technology enables users to engage with virtual apparel in real time—turning and zooming clothes just as if they were physical. It provides a rich, immersive shopping experience that is natural and interesting. One of the biggest challenges is still accurate size prediction from a picture. Pose estimation can detect significant landmarks such as shoulders and hips, but converting that information into accurate clothes sizes involves sensitive algorithms. Automating that step minimizes errors and optimizes fit, which is essential to making virtual try-on really valuable. Aside from visuals and computation, the system has to handle user data safely. Functions such as account sign-up, wishlists, and try-on history tailor the experience and enhance retention. Utilizing databases such as MySQL provides effective data management, while robust privacy measures safeguard user data and establish trust. This project will incorporate these technologies into an end-to-end, user-friendly virtual try-on system. Through the combination of MediaPipe for real-time pose estimation, PIFuHD for 3D avatar generation, VITON for photorealistic clothing overlay, and AR for interaction, the system intends to provide a precise and immersive digital fitting room that anyone can access. In addition to enhancing e-commerce, virtual try-on can decrease waste by reducing the rate of returns and overproducing less. It also facilitates innovative new methods of experiencing fashion, including virtual fashion shows and personal styling that accepts varied body types. In short, the combination of AI-based pose estimation, GAN-based 3D modeling, and AR visualization is opening doors to the next level of virtual try-on experiences. This initiative is committed to providing an accessible, practical, and privacy-aware solution for revolutionizing digital fashion retail.

II. LITERATURE SURVEY

Development of virtual fitting systems includes interdisciplinary research encompassing computer vision,

machine learning, human-computer interaction, and fashion technology. This literature review surveys related work in three main areas. Reviewing existing virtual estimation techniques with 3D modeling using body size and GAN technology provide the basis for designing effective and innovative virtual fitting platforms.

A. Existing Virtual Testing in Systems

Virtual Subject Technology has evolved significantly over the past decade, driven by the desire to replicate the tactile experience of physical clothing in a digital environment. Early systems focused primarily on simple 2D overlays, where clothing photos were superimposed on static photos of users. However, these approaches suffered from poor realism, as they were not responsive to precise transformations of posture, shape, or clothing. For example, static image-based apps and web interfaces offer limited interaction and adaptation.

Researchers began using convolutional neural networks (CNNs) to understand the structure of the human body and the shape of clothing, enabling better preservation of body poses and clothing integrity. Han et al. proposed the Virtual Try-On Network (VITON), which made significant advances by using a generative model to combine photographs of humans and clothing while maintaining texture and body contours. VITON and its successors use encoder-decoder architectures to synthe- size plausible images that enhance visual fidelity. However, many systems still require expensive devices or operate only in controlled environments. Additionally, few systems provide seamless, real-time interaction, limiting applications to virtual try-on experiences.

Despite advances, challenges remain in addressing different body types, clothing styles, and user movements, highlighting the need for an integrated system combining pose estimation, 3D modeling, and realistic rendering.

B. Body Size Approach

Accurate estimation of body size is critical to ensure the fit and comfort of virtual clothing, directly affecting user satisfaction. Traditional size estimates rely on manual mea- surements or self-reported data, which are time-consuming and error-prone. Automated methods using computer vision have been developed to extract anthropometric measurements from images and videos.

Classical image processing techniques were initially used to identify body silhouettes and approximate measurements but were sensitive to lighting conditions, clothing, and camera angles. The emergence of deep learning, particularly pose estimation models, provided a more robust framework. Models such as OpenPose and MediaPipe 2D Keypoints represent anatomical landmarks (shoulders, elbows, hips, knees), which form the basis for calculating body dimensions like shoulder width, waist size, and leg length.

An additional approach integrates statistical body models such as SMPL (Skinned Multi-Person Linear Model) to infer 3D shape parameters capturing body variability. These meth- ods improve accuracy but often require controlled input images or specialized hardware. Furthermore, converting exact body size categories (e.g., S, M, L) remains challenging due to in- terpersonal variations and inconsistent size standards. Ongoing work focuses on combining machine learning classifiers with user feedback loops to improve prediction reliability.

III. DISCUSSION

An integrated, strong framework is suggested to strengthen the virtual try-on experience and the accuracy of size estimation through the mixing of real-time pose detection, intelligent size estimation algorithms, and dynamic adjustment of garment fitting. Each part targets important issues in virtual try-on systems: fitting visualization with realistic appearance, precision in size recommendation, and fluid user interaction.

A. Key Findings

Real-time pose estimation allows accurate tracking of the user's body keypoints and enables virtual clothing to dynamically accommodate various postures and movements. This results in a highly immersive and interactive try-on experience, which is better than static or pre-recorded overlays that fail to respond. Technical improvements in pose estimation models like *MediaPipe* and *OpenPose* prove the viability of exact body tracking through ubiquitous cameras without the necessity of high-cost hardware.

Smart size estimation algorithms make use of anthropometric information from pose keypoints in order to produce personalized size estimates. This minimizes the need for manual measurements, increasing convenience and fitting accuracy. Blending statistical body models such as SMPL with machine learning classifiers enhances the robustness of size prediction across a wide variety of body shapes.

Dynamic garment fitting methods change virtual clothing in real time so that virtual tryons look natural and closely resemble physical fit. This kind of flexibility is critical for visual realism and user confidence, particularly during movement of the user or pose change.

B. Research Gaps

Even with these advancements, current systems suffer from a number of shortcomings. Real-time pose estimation may suffer under changing lighting or occlusion conditions that impact measurement accuracy. Size estimation models tend to rely on controlled input environments or need user calibration, which compromises real-world usability. Moreover, most virtual try-on systems do not have fully integrated pipelines that feature pose detection, size estimation, and realistic 3D garment rendering, creating disjointed or inconsistent user experiences.

Privacy issues are another challenge in that most systems send out sensitive body data to remote servers, which raise data security and trust-related concerns for users. Most solutions to date have to do with compromise between on-device computational capacity and model accuracy. The computational requirement of 3D modeling and real-time rendering also im- pacts system responsiveness, particularly on consumer devices.

C. Implications of the Research

This research leads toward a future when virtual try-on technology offers highly tailored, privacy-respecting, and interactive fit experiences accessible over mainstream consumer hardware. By combining real-time pose estimation with clever size estimation and dynamic garment adjustment, the platform has the potential to substantially decrease the incidence of fit- related returns in e-commerce and increase customer satisfaction.

Additionally, the focus on local processing and secure data management is in accordance with increasing privacy regulations and consumer demand. The study also leaves room for other innovations in digital fashion, such as the ability to support more advanced garment types, style advice via AI, and virtual fitting rooms that are accessible by multiple users. Future research can be directed towards

maximizing computational efficiency, increased compatibility with varied clothing styles, and improved robustness under different environmental conditions.

IV. CLOSING THE RESEARCH GAP

In response to the limitations we've identified, we intro- duce a modular framework consisting of three complementary modules—Real-Time Pose Detection, Smart Size Estimation, and Dynamic Garment Fitting Adjustment—combined into an integrated virtual try-on system. The pipeline-based architecture guarantees that each module addresses a vital facet of the virtual try-on process: accurate pose tracking, personalized size suggestion, and realistic garment display.

A. Real-Time Pose Detection Module

This module utilizes state-of-the-art pose estimation pipelines like MediaPipe and OpenPose to monitor anatomical keypoints (e.g., shoulders, elbows, hips, knees) in real time. It serves as the basis for reliable garment alignment and body measurement extraction. Through dynamically mapping body poses, this module surmounts the static limitation of previous 2D overlay systems and increases the sense of realism and interactivity of the virtual try-on experience.

B. Smart Size Estimation Module

A development upon the result of the pose detection module, this part employs anthropometric information to approximate the user's accurate body measurements. Employing statistical body models such as SMPL and machine learning classifiers, it converts live measurements into tailor-made size suggestions (e.g., S, M, L). This makes input by hand or user calibration unnecessary, solving real-world usability issues and increasing size prediction accuracy across varied body types.

C. Dynamic Garment Fitting Adjustment Module

This last module makes virtual clothing look as if it con- forms naturally to the user's pose and motion. It uses 3D modeling methods for fitting and rendering clothes in real- time, for preserving visual realism and for making the shape and drape of the garment respond naturally to posture changes. It closes the gap between static image-based tryons and actual dressing experiences by constantly updating the fit of the virtual garment from live keypoint data.

All three modules are embedded within a real-time, feedback-based loop. Pose detection provides the size estimation system with precise keypoints, which suggests garment sizes that are visualized and refined immediately by the dynamic garment fitting module. The design facilitates scalable deployment, creates an immersive experience, and guarantees that the virtual try-on system accommodates different users and garment types easily, as well as solving practical issues such as lighting variation, occlusion, and body shape variations.

V. CONCLUSION AND FUTURE WORK

The fast development of virtual try-on technology is a major leap towards how customers interact with fashion today in the digital world. The current paper has examined major elements critical to creating a viable virtual try-on system, such as real-time body pose estimation, 3D modeling of humans via generative adversarial networks, and realistic garment render- ing via platforms such as VITON. Together, these technologies promise to bridge the gap between traditional physical shopping and online retail by providing a more interactive, personalized, and immersive experience.

From the literature, it is evident that real-time pose estimation frameworks such as MediaPipe enable efficient detection of human body keypoints using only standard RGB cam- eras, making it feasible to track user movements dynamically without expensive hardware. At the same time, 3D human reconstruction methods such as PIFuHD have shown that they can create high-fidelity avatars from a single image without resorting to complicated multi-camera setups or depth sensors. The combination of these avatars with virtual try-on networks enables clothing to be realistically superimposed, maintaining delicate details like folds and texture, thus improving the visual accuracy of digital tryons.

Nonetheless, a number of challenges still lie unsolved. A significant challenge is accurately estimating clothing size from image data alone. Although pose estimation can identify body landmarks, converting them into accurate size recommendations is a task that demands calibration and is frequently subject to inconsistency due to body shape and posture variations. The literature emphasizes the need for creating adaptive algorithms capable of personalizing size forecasts in line with individual anatomical variation instead of using one- size-fits-all sizing charts. This is crucial to enhancing customer satisfaction and minimizing product returns, which still hound e-commerce fashion websites.

Another key concern raised by current studies is the avail- ability and quality of data. Diverse body types, poses, and fashion styles in high-quality datasets are crucial for training robust AI models. But gathering these datasets is both costly and time-consuming, and this usually constrains model generalization. Synthetic data augmentation methods hold promise but need to be used with great care so as not to introduce bias. Moreover, user privacy and data security are also top concerns, as virtual try-on systems handle sensitive personal images. Future research needs to adhere to ethical guidelines and include robust data protection mechanisms to build user trust.

Additionally, user experience is a critical determinant for the success of virtual try-on systems. Smooth human-computer interaction with intuitive interfaces and augmented reality overlays can raise user interaction dramatically. Low latency and efficient visual rendering on consumer-grade hardware are critical for mass use. Theoretical literature recommends modular system designs to enable independent development and upgrade of pose estimation, 3D modeling, and try-on modules to keep pace with changing technology needs.

Overall, the integration of AI-driven pose estimation, GAN- based 3D reconstruction, and realistic virtual try-on systems holds tremendous potential to disrupt online fashion shopping. The findings from this survey highlight the importance of solving size estimation accuracy, data diversity, privacy, and usability issues in order to achieve fully functional and scalable solutions. By overcoming these issues, virtual try-on platforms in the future will be able to give consumers personalized, immersive, and eco-friendly shopping experiences that most closely mirror the advantages of physical tryons.

This review of this paper provides a solid basis for further research and development work in virtual try-on technology. The convergence of cutting-edge computer vision and deep learning techniques and human-centered design values has the potential to revolutionize not just the way individuals go about buying clothing, but the way fashion technology is developed to satisfy the demands of a networked world.

REFERENCES

[1] Z. He, Y. Ning, Y. Qin, W. Wang, S. Yang, L. Lin, and G. Li, “VTON 360: High-Fidelity Virtual Try-On from Any Viewing Direction,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2025.

[2] T. Islam et al., "StyleVTON: A Multi-Pose Virtual Try-On with Identity and Clothing Detail Preservation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

[3] S. Bai, H. Zhou, Z. Li, C. Zhou, and H. Yang, "Single Stage Virtual Try- On via Deformable Attention Flows," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.

[4] B. Fele, A. Lampe, P. Peer, and V. S˘truc, "C-VTON: Context-Driven Image-Based Virtual Try-On Network," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.

[5] Anonymous, "VTON-MP: Multi-Pose Virtual Try-On via Appearance Flow and Feature Filtering," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

[6] J. Karras, Y. Li, N. Liu, L. Zhu, I. Yoo, A. Lugmayr, C. Lee, and I. Kemelmacher-Shlizerman, "Fashion-VDM: Video Diffusion for Virtual Try-On," 2024.

[7] L. Lu, S. Wu, H. Sun, J. Gou, J. Si, C. Qian, J. Zhang, and L. Zhang, "Self-Supervised Vision Transformer for Virtual Try-On," 2024.

[8] Z. Xie, Z. Huang, F. Zhao, H. Dong, M. Kampffmeyer, X. Dong, F. Zhu, and X. Liang, "PASTA-GAN++: A Versatile Framework for High- Resolution Unpaired Virtual Try-On," 2022.

[9] Y. Choi, S. Kwak, K. Lee, H. Choi, and J. Shin, "IDM-VTON: Improving Diffusion Models for Authentic Virtual Try-On in the Wild," 2024.