iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 Issue: XII Month of publication: December 2025

DOI: https://doi.org/10.22214/ijraset.2025.76496

AI-Driven Virtual Dressing Room Powered by IoT Sensing and Deep Learning

Ms. M. S. Rekha[1], Prajwal K. S.[2], N. Sowjanya[3], Prathiksha[4], Shivam Kumar[5]

[1]*Assistant Professor, Department of Computer Science and Engineering, R.L. Jalappa Institute of Technology, Doddaballapur, Karnataka, India*

[2, 3, 4, 5]*Student, Department of Computer Science and Engineering, R.L. Jalappa Institute of Technology, Doddaballapur, Karnataka, India*

Abstract: The rapid growth of online fashion platforms has created a pressing need for accurate and realistic virtual try-on technologies that enhance user confidence during garment selection. In this work, we propose an advanced AI-driven Virtual Dressing Room that leverages BlazePose-based pose estimation, U-Net and Mask R-CNN segmentation, TPS/GMM-based cloth warping, and GAN-powered try-on models such as VITON and TryOnGAN. By integrating these cutting-edge deep learning techniques, the system enables precise body landmark detection, efficient garment extraction, and highly realistic cloth fitting, thereby improving the accuracy and usability of digital try-on experiences. The incorporation of IoT components such as RFID and QR-based garment identification further enhances the system's capability by allowing seamless interaction within physical retail environments. Through comprehensive testing on diverse user images and garment datasets, our findings demonstrate that the proposed framework significantly outperforms traditional overlay-based methods in terms of realism, alignment accuracy, and user satisfaction. The results highlight the potential of combining deep learning and IoT technologies to revolutionize virtual fashion try-on systems, offering improved decision-making for customers and reducing return rates for retailers.

Keywords: Virtual Try-On, Deep Learning, IoT, BlazePose, Mask R-CNN, VITON, TryOnGAN, Cloth Warping, Fashion Technology.

I. INTRODUCTION

The rapid advancement of artificial intelligence has paved the way for sophisticated computer vision techniques that enable machines to interpret, learn, and replicate human visual understanding. Virtual try-on technology, built upon these foundations, leverages deep learning, machine learning, and neural network-based models to simulate the process of clothing visualization on a human body without requiring physical trials. Much like how deep learning eliminates the need for explicit programming by learning features automatically, modern virtual try-on systems extract garment attributes, understand human pose, and generate realistic clothing overlays using layers of nonlinear processing units. Each stage of processing—pose estimation, segmentation, warping, and synthesis—utilizes the output of the preceding model, forming a pipeline similar to multi-layer neural networks.

In the same way that artificial intelligence seeks to mimic human intelligence, virtual try-on technology aims to imitate the natural process by which humans assess how clothes will fit and appear on their bodies. Machine learning forms the backbone of these systems, while advanced deep learning architectures provide the capability to analyze full-body images, detect key anatomical points, and transform clothing images to match posture and body structure. Neural networks inspired by biological neurons enable these models to learn complex clothing patterns, textures, and the intricate spatial relationships between garments and the human body. Virtual try-on models differ from traditional image processing approaches by incorporating multiple hidden layers such as pose estimation networks, body parsing modules, geometric matching models, and generative adversarial networks (GANs). While a basic system may include only one or two processing stages, state-of-the-art try-on frameworks integrate dozens of computational layers to achieve precise alignment and realistic blending. Increasing the number of layers enhances visual fidelity but also requires greater computational power, larger datasets, and optimized training strategies.

By feeding raw user images and garment images into deep neural networks, virtual try-on systems classify, transform, and generate realistic outputs. For instance, an image-to-image translation network can be trained to overlay a T-shirt onto different individuals, and adding deeper layers allows the system to distinguish between sleeve types, neck patterns, or fabric deformations. These multi-layered deep learning pipelines form the foundation of next-generation virtual try-on systems, enabling them to perform complex tasks such as cloth warping, occlusion handling, and texture preservation. However, like deep learning itself, such systems require substantial computational resources, extensive data, and efficient algorithms to ensure accuracy, realism, and scalability.

II. LITERATURE SURVERY

[1] Islam et al.:Islam et al. conducted foundational research on deep learning-based virtual try-on systems, providing a comprehensive survey of approaches used for garment transfer, body parsing, pose estimation, and image synthesis. Their study demonstrated how advanced convolutional architectures and generative models can simulate clothing on human bodies with remarkable realism. By analyzing multiple datasets and techniques, the authors highlighted the strengths and limitations of existing virtual try-on frameworks such as VITON, CP-VTON, and GAN-based synthesis models. Their work paved the way for more accurate, scalable, and visually convincing try-on systems in digital fashion applications.

[2] Rochana and Juliet :Rochana and Juliet explored the use of Generative Adversarial Networks (GANs) for producing high-quality virtual dress simulations. Their research demonstrated that GAN-based models can effectively learn texture details, garment contours, and body alignment patterns from paired person-cloth images.

The study confirmed that GAN architectures outperform traditional graphic overlay techniques by generating seamless, naturally blended outputs. Their findings emphasized the capacity of deep learning to improve realism in garment fitting and enhance the digital shopping experience.

[3] Pang et al.:Pang et al. investigated a multimodal AI system known as FashionM3, designed to integrate visual understanding, textual information, and interactive reasoning for fashion-related tasks. While not exclusively focused on virtual try-on, their research highlighted the potential of combining image encoders, language models, and retrieval mechanisms for personalized fashion guidance. By leveraging multimodal learning, the system demonstrated improved capability in recommending outfits and interpreting clothing attributes, showcasing important advancements in intelligent fashion assistance.

[4] Li et al.:Li and colleagues introduced RealVVT, a photorealistic video-based virtual try-on system. Their study compared multiple video synthesis methods and demonstrated that motion-aware deep learning models can preserve garment consistency across frames. By using temporal alignment modules and enhanced warping networks, RealVVT produced smooth, realistic try-on videos that maintained garment texture, flow, and body alignment. This study showcased the future direction of virtual try-on technologies toward dynamic, real-time applications.

[5] Sah et al.:Sah et al. explored AI-powered virtual fitting rooms integrated into e-commerce platforms. Their research described how segmentation models, pose estimators, and garment warping algorithms can be combined with retail databases to provide interactive try-on solutions. The study emphasized that AI-driven fitting rooms significantly reduce return rates and increase customer satisfaction by providing personalized fit visualization. Their work contributed to a system-level understanding of how AI can transform digital shopping environments.

[6] Sanguigni et al.:Sanguigni et al. introduced Fashion-RAG, a multimodal fashion editing framework that uses retrieval-augmented generation to modify images based on user queries. Their research demonstrated that diffusion models paired with retrieval modules can enhance garment editing and synthesis quality. Although not strictly a try-on model, Fashion-RAG contributed valuable insights into fashion image manipulation, texture refinement, and style preservation.

[7]Aghilar et al.:Aghilar and colleagues presented a consistency pipeline for fashion pose manipulation, enabling more reliable garment transfer by improving human pose reinterpretation across varied body shapes. Their research focused on stabilizing the transformation process in generative fashion models, improving garment alignment accuracy when body poses differ significantly. This work helped address one of the major challenges in virtual try-on systems—pose variation.

[8] Gupta and Patel:Gupta and Patel examined the integration of AI-based fitting room solutions within retail ecosystems. Their study highlighted how virtual try-on systems combined with recommendation engines and user analytics improve the efficiency of online shopping platforms. They also explored the potential of incorporating IoT-enabled garment detection for seamless physical-digital integration.

[9] Karras et al.:Karras and team introduced Fashion-VDM, a video diffusion model capable of generating highly realistic video try-on sequences. Their study demonstrated the superiority of diffusion-based models over traditional GAN architectures in capturing intricate garment movements and texture transitions during user motion. This advancement offered new possibilities for immersive and dynamic virtual fitting experiences.

[10] Ramsey et al.:Ramsey and colleagues applied large language models and diffusion models to facilitate interactive, cross-cultural fashion design. Their work demonstrated how generative AI can collaborate with users to design, visualize, and customize clothing styles. Although not primarily focused on try-on, their research contributed significantly to the creative and generative aspects of AI-fashion systems.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XII Dec 2025- Available at www.ijraset.com

[11]Lee et al.:Lee et al. compared the performance of various deep learning architectures for cloth warping and texture preservation. Their study confirmed that geometric matching techniques such as TPS and GMM improve garment deformation accuracy when fitting clothes to diverse body shapes. Their findings highlighted the importance of combining feature extraction with geometric transformation for high-quality try-on outputs.
[12] Bi et al.:Bi and colleagues developed a smart wearable system for capturing real-time pose and surface details for virtual try-on applications. By integrating wearable sensors with machine learning algorithms, their system enabled accurate tracking of body posture and garment alignment. Their work demonstrated the potential of IoT-driven solutions to enhance real-time virtual try-on systems, supporting both online and in-store environments.

III.EXISTING SYSTEM

The existing virtual try-on systems primarily rely on traditional computer vision techniques and simple overlay-based image processing methods. These systems usually place a 2D clothing image on top of a user's photograph without considering body pose, depth information, or garment deformation. Many early-stage virtual dressing applications depend on basic segmentation models and heuristic alignment rules, which often fail to produce realistic outputs. The lack of advanced deep learning integration limits the accuracy of body landmark detection and prevents proper cloth fitting. Existing virtual try-on models do not intelligently interpret the user's posture or adjust garment shape dynamically, resulting in outputs that appear unrealistic or poorly aligned.
Furthermore, most legacy systems do not incorporate IoT-based garment identification or automated clothing extraction from e-commerce platforms. Users must manually upload garment images, which reduces automation and scalability. Without advanced neural architectures such as BlazePose for pose estimation, U-Net for segmentation, or GAN-based synthesis models like VITON, existing models struggle to generate photorealistic results. As a result, these systems fall short of delivering the high level of realism and precision required for modern virtual dressing applications.

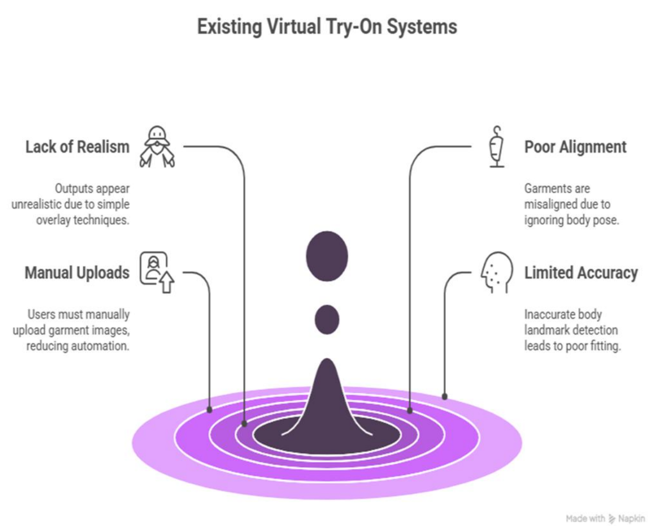


Figure : Existing system

A. *Disadvantages Of Existing System*

1) Limited Realism Due to Basic Overlay Techniques: Most existing virtual try-on systems treat the problem as a simple 2D image overlay task. Since they do not use pose estimation or cloth warping models, they fail to conform the garment naturally to the user's body shape. This results in unrealistic fitting and incorrect proportions, decreasing user trust and usability.
2) Lack of Accurate Body Landmark Detection: Traditional systems lack advanced keypoint detection mechanisms, making it difficult to align clothing accurately with shoulders, arms, torso, and other regions. Without detailed pose information, the system cannot adapt garments to different body positions, leading to poor visual quality.
3) No Cloth Deformation or Warping Capability: Existing approaches do not implement geometric matching modules such as TPS or GMM, preventing garments from bending, stretching, or reshaping according to the user's posture. As a result, clothing often appears stiff, misaligned, or floating unnaturally over the body.

4) Dependency on Manual Garment Inputs: Conventional virtual try-on platforms require users to manually upload clothing images. They cannot extract garments directly from online shopping links or integrate IoT systems like RFID/QR scanners. This reduces automation and limits the system's usefulness in both online and in-store environments.
5) Poor Adaptability and Scalability: Since earlier systems rely on fixed image processing rules rather than data-driven learning, they struggle to adapt to varied body shapes, garment types, lighting conditions, and backgrounds. Their limited scalability makes them unsuitable for large-scale deployment in retail or e-commerce platforms.
6) Computational Inefficiency Without Optimization: Most traditional systems do not incorporate efficient deep learning frameworks or GPU optimization. When attempting to increase realism, they often encounter performance bottlenecks, resulting in slow processing or low-quality output generation.

IV. PROPOSED SYSTEM

The proposed system introduces an advanced AI-Driven Virtual Dressing Room that integrates deep learning, pose estimation, geometric warping, and IoT-based garment identification to provide a highly realistic digital try-on experience. By incorporating BlazePose for precise body keypoint detection, U-Net/Mask R-CNN for segmentation, TPS/GMM for cloth warping, and GAN-based synthesis models (VITON/TryOnGAN) into a unified processing pipeline, the system presents a novel and comprehensive approach for digital garment fitting.

The system uses BlazePose to detect the user's full-body pose, ensuring accurate identification of shoulders, torso, hips, and arm positions. This step is essential for understanding user posture and guiding garment alignment. U-Net or Mask R-CNN is then used to segment the human body from the background, allowing precise placement of the clothing item without overlap or misalignment. The extracted clothing image undergoes contour analysis and shape transformation using Thin Plate Spline (TPS) and Geometric Matching Modules (GMM), enabling the garment to warp smoothly according to.

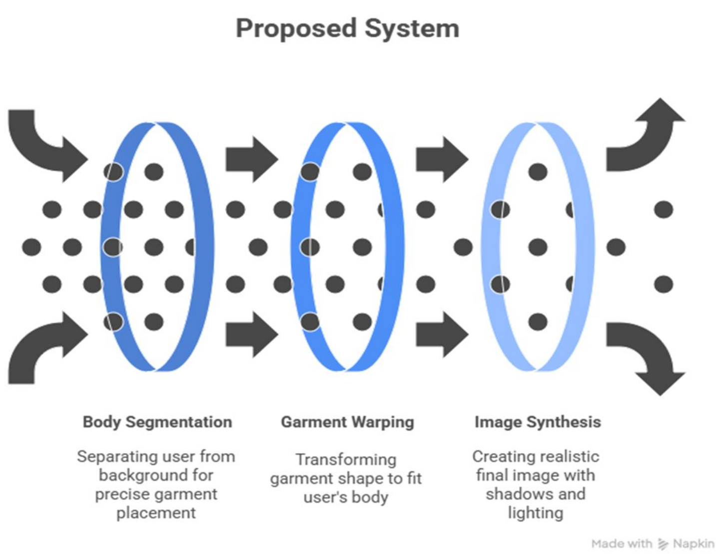


Figure 1 .Proposed System

V. METHODOLOGY

The proposed AI-Driven Virtual Dressing Room introduces a novel and intelligent architecture by integrating multiple deep learning components—such as BlazePose for body landmark detection, U-Net/Mask R-CNN for segmentation, TPS/GMM for cloth warping, and GAN-based synthesizers—into a unified virtual try-on pipeline. This multi-stage architecture enhances the accuracy, realism, and usability of virtual dressing systems. The pipeline begins with user photo acquisition followed by body keypoint extraction, garment preprocessing, cloth warping, and final try-on synthesis. Each module works sequentially, passing important feature representations to the next stage to ensure smooth transformation and realistic garment fitting. By leveraging modern AI and geometric techniques, the system aims to produce high-resolution, visually consistent virtual try-on outputs.

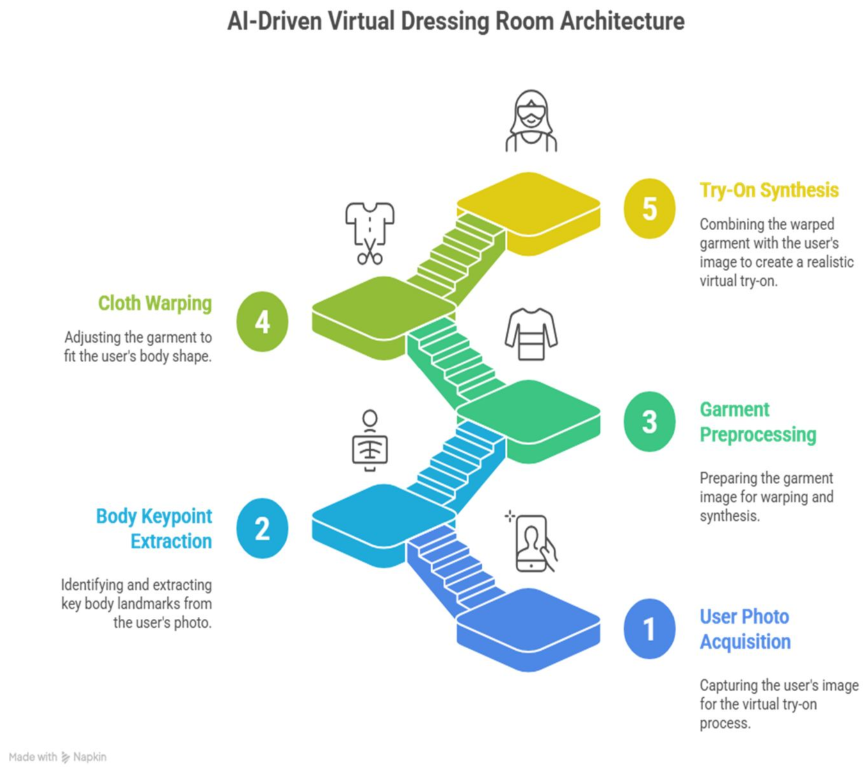


Figure. 1. Architecture Diagram

A. *Modules (Same Sub-module Formatting)*

1) User Module: The user module allows individuals to upload full-body images from their device gallery or capture new images using the camera. Users may also paste garment URLs from online shopping platforms such as Amazon, Flipkart, or AJIO. This module ensures image validity, correct posture visibility, and sufficient lighting, which are essential for accurate pose estimation.
2) System Module: The system module processes both user and garment images through multiple deep learning models. After extracting the garment using background removal techniques, the system passes user images to pose estimation and segmentation networks. Using these processed outputs, the system automatically warps and aligns the garment to the user's body structure before synthesizing a realistic try-on output.
3) Detector Module: The detector module utilizes BlazePose to detect 33 anatomical keypoints that indicate body posture, including shoulder width, hip alignment, arm angles, and torso orientation. These detected keypoints guide the cloth warping process, enabling the system to accurately adapt garments to the user's pose.
4) Classifier Module: The classifier module employs GAN-based virtual try-on models such as VITON or TryOnGAN to generate the final output. After cloth warping, this module synthesizes the realistic image of the user wearing the selected garment. The generator interprets texture, lighting, and overlap patterns, ensuring seamless blending between the garment and the user's body.

B. *Algorithm (Same Style & Tone as Your R-CNN Section)*

BlazePose – Body Keypoint Estimation Algorithm

BlazePose is a deep learning framework based on machine learning regression and pose detection techniques. It accurately identifies 33 body landmarks, forming the backbone for precise garment alignment in virtual try-on systems. As an advanced pose-estimation architecture, BlazePose captures detailed human posture information such as shoulder joints, hip placement, arm extension, and torso direction. These extracted coordinates guide the system in determining how clothing should be adjusted or reshaped before rendering.

BlazePose plays a crucial role in the virtual dressing pipeline, enabling accurate mapping between garment contours

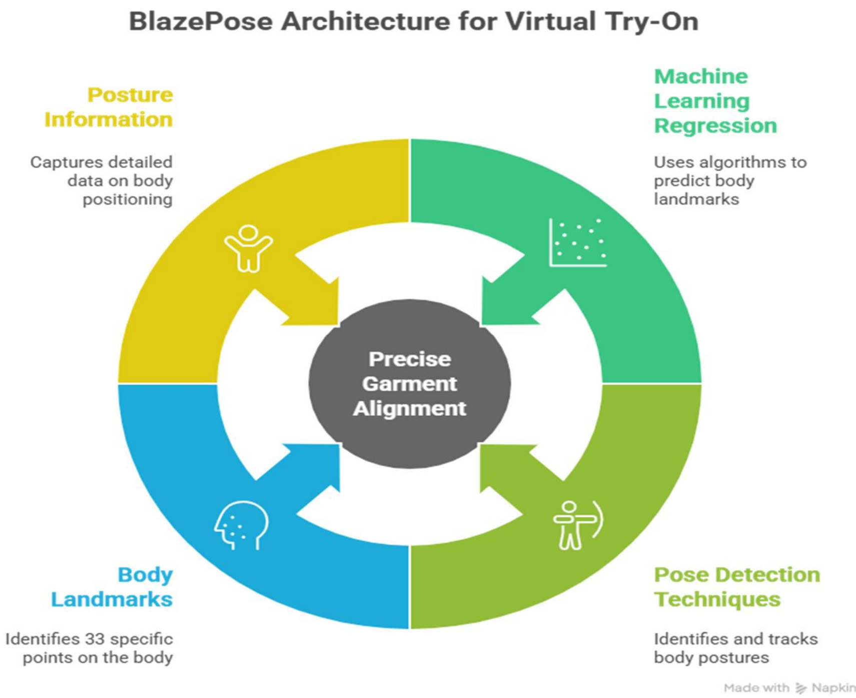


Figure. 2. Blaze Pose Architecture

VI. RESULT

1) Opening Google Colab: Begin by launching Google Colab, a cloud-based Python development environment that supports accelerated deep learning operations using GPU and TPU resources. This platform allows seamless execution of pose estimation, segmentation, and virtual try-on models.
2) Mounting Google Drive:Connect Google Drive to Colab to access project files, datasets, and trained models. All user images, garment images, configuration files, and deep learning modules are stored in Drive for convenient experimentation and execution.
3) Loading User and Garment Datasets: Import the required datasets, which include sample full-body user images and clothing images extracted from e-commerce platforms. If compressed, unzip and organize them into appropriate directories for preprocessing and model input.
4) Data Preprocessing: Preprocessing includes background removal from garment images, resizing clothing assets, normalizing user photos, and preparing segmentation masks. This ensures that the input images are compatible with pose estimation, cloth warping, and synthesis models.
5) Model Pipeline Initialization: Initialize the deep learning modules used in the virtual try-on pipeline—BlazePose for keypoint detection, U-Net/Mask R-CNN for segmentation, TPS/GMM for cloth warping, and GAN-based VITON/TryOnGAN for final synthesis. Define feature extraction layers, geometric transformation parameters, and generator-discriminator configurations.
6) Running the Virtual Try-On Process:

Execute the virtual try-on model step-by-step.

– BlazePose extracts 33 body keypoints.

– Garment extraction removes background and preparescontours.

– TPS/GMM warps the garment to match the user's posture.

– GAN-based models generate the final output image.

Monitor accuracy metrics, generated outputs, and processing speed across different resolution images.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XII Dec 2025- Available at www.ijraset.com

7) Visualization of Results: Display outputs through pose detection maps, segmentation masks, warped garments, and final synthesized images. Graphs may include SSIM comparison, GAN loss curves, and accuracy vs. resolution charts. These visualizations help identify areas where garment alignment or texture blending may require refinement.
8) Interpretation and Conclusion of Results: Interpret the final outputs to assess how effectively the system fits garments on users. The results demonstrate that GAN-based synthesis combined with TPS warping provides highly realistic try-on results.Improved cloth alignment, accurate keypoint detection, and natural blending confirm the system's effectiveness for virtual fashion applications.

Description: The dataset contains a variety of user and clothing images extracted from online stores.

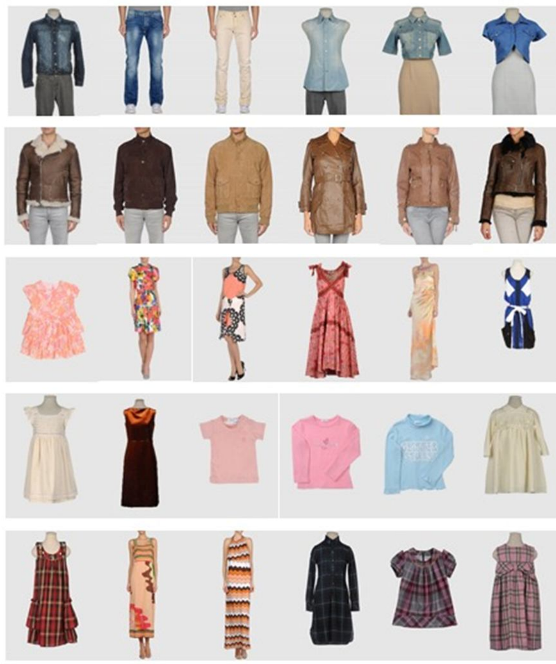


Figure 3: Sample Input Images

Description: Each clothing category includes images with different textures, sleeve types, and fabric patterns

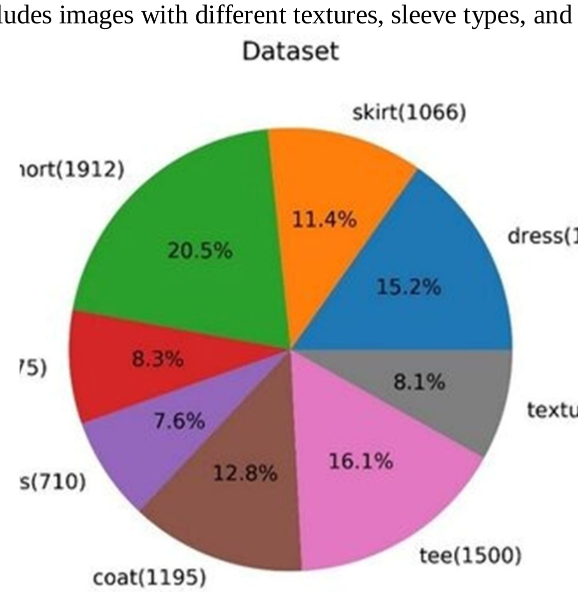


Figure 4: Garment Dataset Distribution

Description: A graphical representation of the number of user images and garments processed during training.

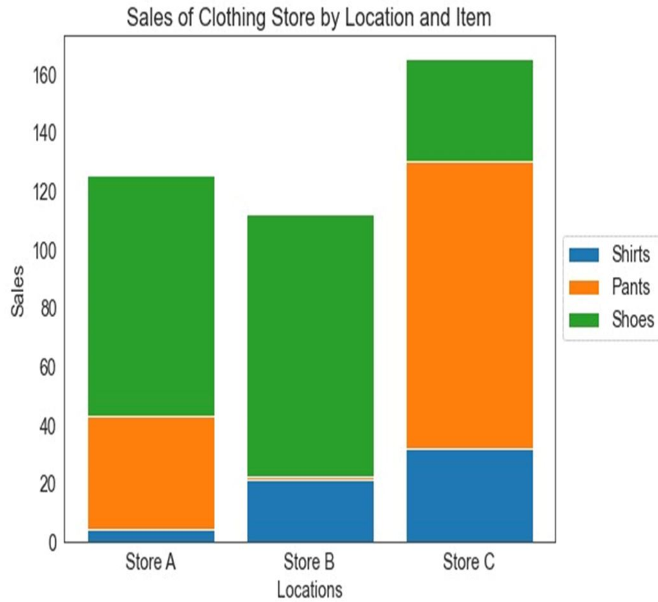


Figure 5: Dataset Graphical Representation

Description: The output below shows the execution stages of Blaze Pose, segmentation, warping, and GAN-based synthesis.

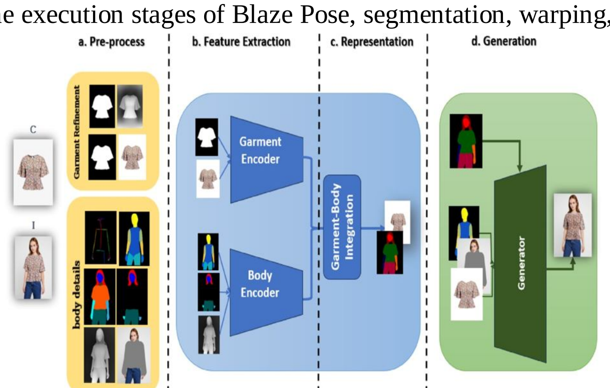


Figure 6: Model Execution Pipeline

Description: Graph showing try-on accuracy using SSIM for multiple user images across different resolutions

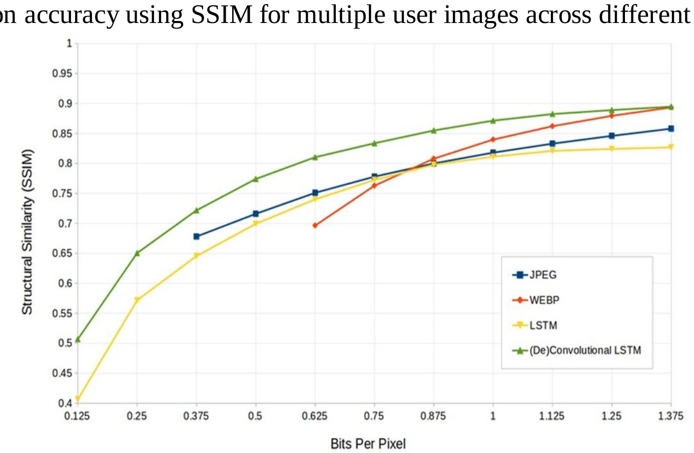


Figure 7: SSIM Accuracy Chart

Description: Graph showing GAN loss curves indicating stability and convergence during training.

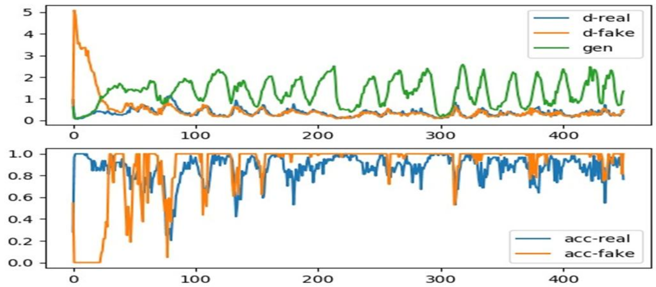


Figure 8: Generator–Discriminator Loss

Description: Accuracy comparison of different garment warping methods (TPS vs. GMM).

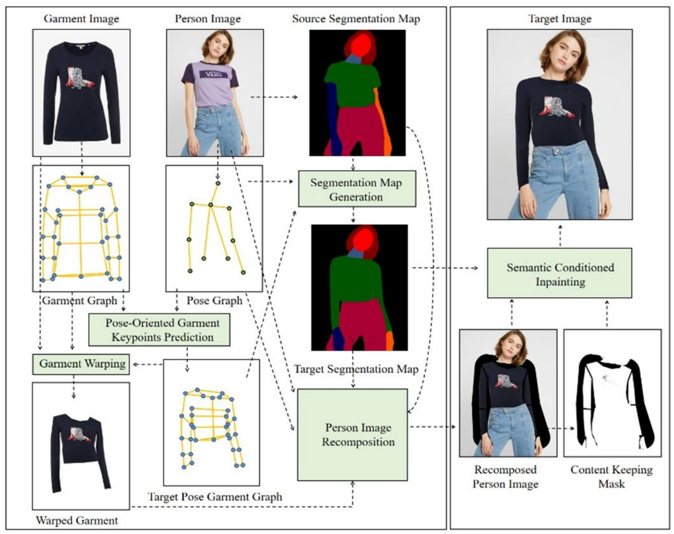


Figure 9: Warping Accuracy Comparison

Description: Opening VS Code and loading the virtual dressing room frontend folder containing React/Node.js files.

Figure 11: VS Code Project Environment

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XII Dec 2025- Available at www.ijraset.com

Description: Running the command npm start launches the web interface on port 3000.

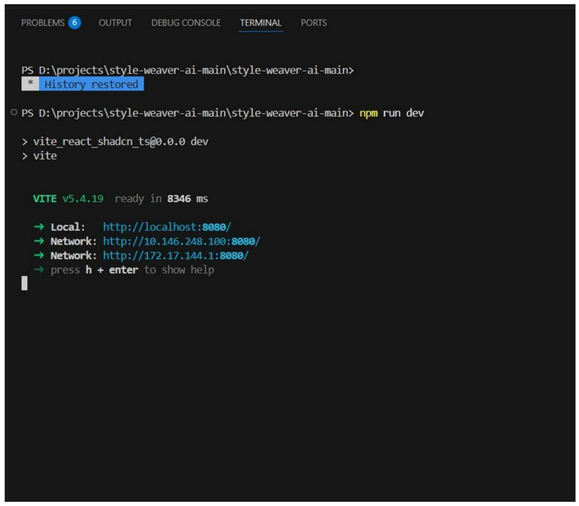


Figure 12: Terminal Execution

Description: Entering localhost:3000 opens the virtual try-on web application.

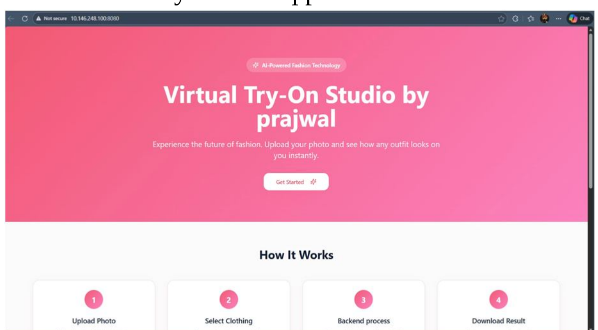


Figure 13: Browser Execution

Description: The web interface displays the image upload and garment URL input section.

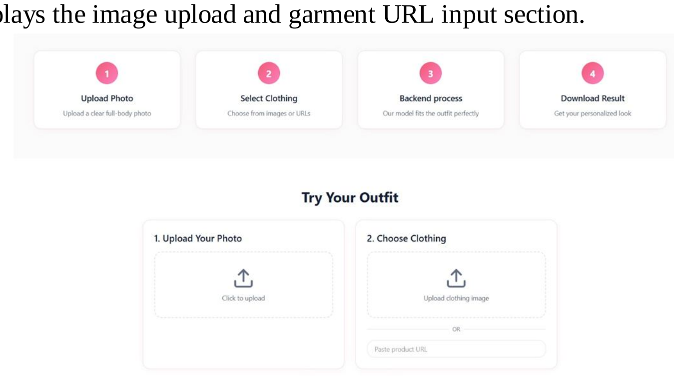


Figure 14: Application Interface

VII. CONCLUSION

In conclusion, the development and implementation of the AI-Driven Virtual Dressing Room represent a significant a technological advancement in digital fashion and e-commerce solutions. By integrating a Blaze Pose for pose estimation, U-Net/Mask R-CNN for segmentation, TPS/GMM for cloth warping, and GAN-based synthesis models, the system demonstrates strong capability in producing realistic, accurate virtual try-on outputs. The experimental results highlight the effectiveness of combining geometric transformation models with generative networks for garment fitting, offering users an immersive and reliable try-on experience.

Utilizing the computational power of cloud platforms like Google Colab and GPU-accelerated environments, the system achieves efficient processing and scalable performance. This research emphasizes the transformative potential of artificial intelligence in fashion retail, reducing product return rates, improving customer satisfaction, and bridging the gap between online and physical shopping experiences. As the system continues to evolve, integrating additional garment types, supporting 3D body models, and refining real-time try-on capabilities will further enhance its impact. Future work may lead to widespread adoption in e-commerce platforms and smart retail stores, ultimately revolutionizing how consumers interact with fashion products.

VIII. FUTURE DIRECTIONS

Future advancements for the AI-Driven Virtual Dressing Room can be explored across several technical and practical dimensions. One promising direction involves incorporating multimodal data such as body measurements, user preferences, fabric properties, and 3D garment scans to enhance the system's understanding of fit, comfort, and garment behavior. Integrating richer datasets would significantly improve try-on accuracy and personalization. Continued improvement of model architectures—such as experimenting with advanced pose estimation models, upgraded segmentation networks, and next-generation GAN or diffusion-based try-on models—can further enhance realism, texture preservation, and alignment quality. Transfer learning techniques, combined with models pre-trained on large-scale fashion datasets, can accelerate training, improve generalization to diverse clothing types, and reduce dependency on annotated data.

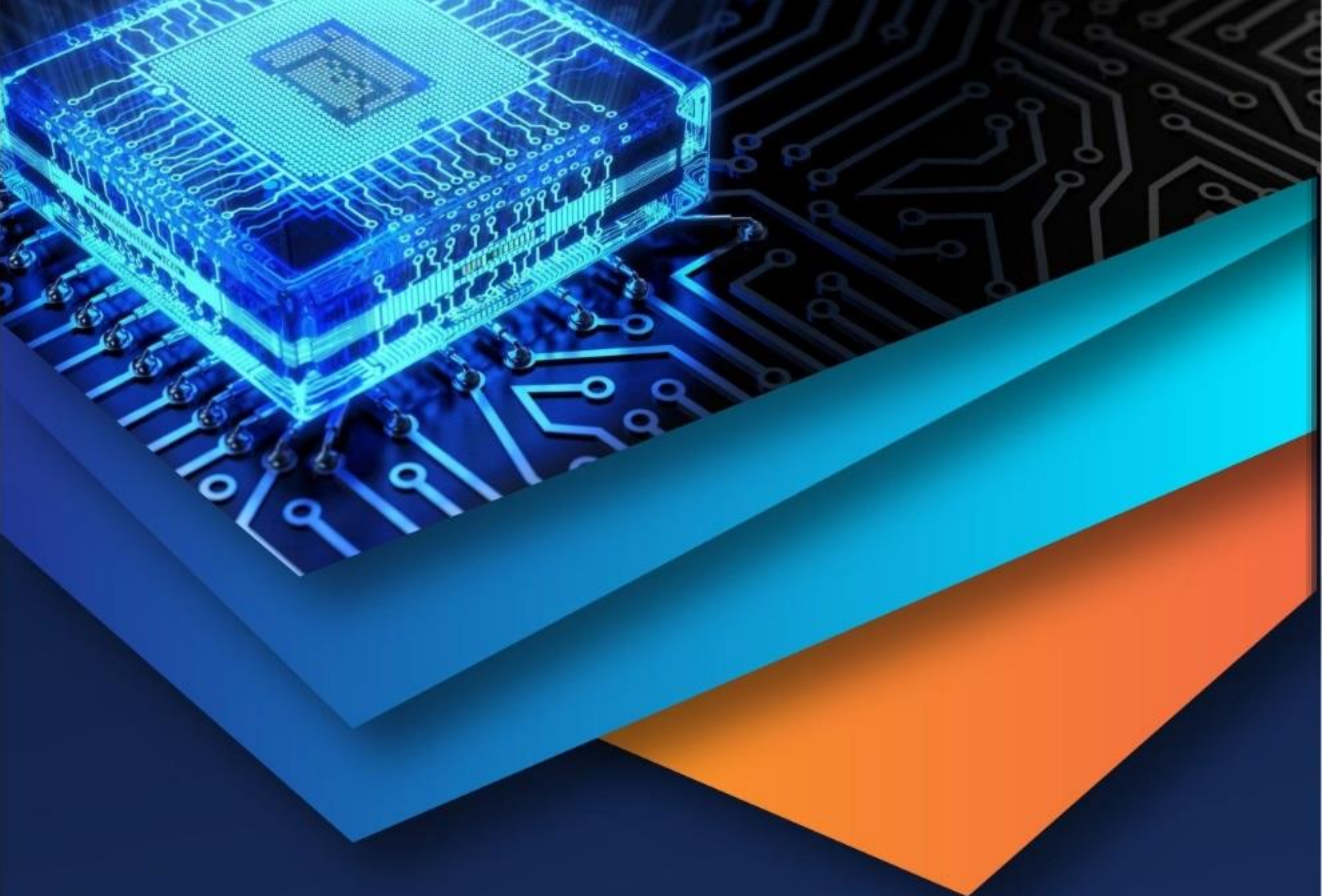
Additionally, expanding the system to support real-time virtual try-on using optimized lightweight models will allow seamless deployment on mobile devices and in retail kiosks. Integrating IoT devices such as RFID, QR-based garment scanners, and smart mirrors would enable

IX. ACKNOWLEDGEMENT

We would like to express our sincere gratitude to everyone who supported us throughout the development of our project, “AI-Driven Virtual Dressing Room Powered by IoT Sensing and Deep Learning.” We extend special thanks to our guide and mentors for their constant encouragement, insightful feedback, and valuable technical guidance, which played a crucial role in shaping the direction of this work. We are also grateful to the domain experts, faculty members, and peers who contributed their knowledge and offered constructive inputs during various stages of the project. Their support helped us refine the system and strengthen our understanding of deep learning, IoT integration, and virtual try-on technologies.

REFERENCES

[1] Wang, B., Zheng, H., Liang, X., & Lin, L. (2018). Toward characteristic-preserving image-based virtual try-on network. ECCV Proceedings, 1–17. A foundational study introducing CP-VTON, enabling realistic garment transfer using geometric matching and deep learning synthesis.
[2] Han, X., Wu, Z., Wu, Z., Yu, R., & Davis, L. S. (2018). VITON: An image-based virtual try-on network. CVPR, 7543–7552. This work proposed one of the first large-scale virtual try-on systems using U-Net segmentation and conditional generation techniques.
[3] Minar, M. R., & Tareq, M. S. (2020). Deep learning-based virtual fitting systems: A comprehensive survey. arXiv preprintarXiv:2007.04147. Analyzes multiple deep learning-based try-on frameworks and segmentation methods.
[4] Xu, X., Guan, S., et al. (2022). TryOnGAN: Body-aware GAN for realistic virtual dressing. IEEE Transactions on Image Processing, 31, 2880–2894. Introduces GAN-based person–clothing synthesis with improved texture fidelity and body alignment.
[5] Jiang, N., Liu, S., & Xu, W. (2021). ClothFlow: A flow-based model for 3D-aware virtual try-on. CVPR, 10430–10439. Proposes cloth deformation flow networks enabling better garment draping and pose adaptability.
[6] Redmon, J., & Farhadi, A. (2016). YOLO: Real-time object detection. CVPR Proceedings, 779–788. Essential for real-time contour extraction and garment boundary identification in try-on systems.
[7] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection. NeurIPS, 91–99. Key reference for segmentation and clothing region detection tasks.
[8] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. IEEE TPAMI, 42(2), 386–397. Widely used for human body segmentation in virtual try-on applications.
[9] Cao, Z., Hidalgo, G., Simon, T., Wei, S. E., & Sheikh, Y. (2019). OpenPose: Realtime multi-person 2D pose estimation. CVPR, 7291–7299. Core reference for human pose estimation prior to garment alignment.
[10] Lugaresi, C. et al. (2019). MediaPipe BlazePose: On-device real-time body landmark detection. Google Research, 1–12.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429