

# Deep Learning-Based Garment Warping and Texture-Preserving Try-On System

<sup>1</sup> Dr. Budhewar Anupama Shankarrao, <sup>2</sup> Dipali Gautam Kakade

<sup>1</sup> Professor, Department of Computer Science & Engineering, JSPM University, Pune.

<sup>2</sup> MTech Student (DSAI), Department of Computer Science & Engineering, JSPM University, Pune.

**Abstract**— With the rapid growth of online fashion platforms, there is an increasing demand for intelligent systems that help customers assess clothing before making a purchase. Virtual try-on technology addresses this issue by digitally simulating how garments would appear on an individual's body. This research introduces a deep learning-driven framework that produces realistic clothing visualizations while preserving both garment structure and body proportions. The proposed system integrates multiple components: semantic segmentation, pose estimation, garment deformation, and generative image synthesis. DeepLabV3+ is employed to separate the user from the background, while Media Pipe extracts body landmarks. Thin Plate Spline transformation adapts garments to the detected pose, and texture consistency is ensured through neural style transfer. Finally, a StyleGAN-based refinement module enhances realism in the synthesized output. Experimental evaluation demonstrates that the system generates convincing results suitable for online shopping environments.

**Keywords**— Deep Learning, Convolutional Neural Networks, Computer Vision, Virtual Try-On (VTON), Generative Adversarial Networks (GAN), Spatial Transformer Networks (STN)

## I. INTRODUCTION

The fashion industry has witnessed significant growth in online shopping, driven by convenience and the wide range of products available. Despite these advantages, one of the biggest drawbacks of digital apparel shopping is the inability to physically try on garments before purchase. This limitation often results in uncertainty regarding size, fit, and overall appearance, which contributes to high product return rates. Such returns not only increase operational costs for retailers but also create dissatisfaction among customers.

Although traditional in-store shopping allows individuals to try clothes before buying, it comes with its own challenges. Visiting stores requires time and effort, and shoppers may encounter long queues, limited stock availability, or hygiene concerns when trying on items previously handled by others.

To overcome these issues, virtual try-on technologies have emerged as a promising solution. These systems allow users to preview clothing on their own images by leveraging computer vision and artificial intelligence. By integrating advanced algorithms, they can deliver a realistic and interactive fitting experience. In this study, DeepLabV3+ is utilized for semantic segmentation to isolate the user from the background, while Media Pipe extracts body landmarks for accurate pose estimation. Thin Plate Spline (TPS) transformation is applied to align garments with the user's body shape and posture. Neural Style Transfer ensures fabric texture and lighting consistency, and StyleGAN refines the

final output to produce photorealistic results where the garment blends seamlessly with the user's body.

### 1.1 Motivation

The topic of Image-based Virtual Try-On (VTO) for apparel were chosen because of its revolutionizing potential at the crossroads of fashion and technology. Even though e-commerce are now a major part of the retail experience, the conventional obstacle to online clothing shopping—like sizing uncertainty and the inability to try on clothes in person—remains. A technological answer to these problem is the development of rooms (VTO), which provide users with an experiencing that goes beside simple visual overlay. This topic is driven from the realization that image-based VTO has the power dynamics of the fashion business in addition to providing online buyers with increasing convenience and confidingent. The combination of artificial intelligence and computer vision.

## II. LITERATURE SURVEY

1.Yuan Chang, Tao Peng, "DP-VTON: Toward Detail-Preserving Image-Based Virtual Try-On Network" [1] Image-based virtual try-on systems aim to overlay target garments onto a person's image. A major challenge is producing realistic results while maintaining non-clothing details. To address this, the authors propose DP-VTON, which introduces a clothing warping module that merges pixel-level and feature-level transformations. This design improves garment adaptation while preserving fine details in the final try-on image.

2,Thai Thanh Tuan, Matiuur Rahman Minar, Heejune Ahn, "CloTH-VTON+: Clothing Three-Dimensional Reconstruction for Hybrid Image-Based Virtual Try-On" [2] Deep learning-based VTON systems often struggle with complex poses due to limited geometry handling. CloTH-VTON+ integrates 3D reconstruction with image-based methods. The pipeline automatically builds a 3D clothing model aligned to a reference human body, enabling natural pose and shape transfer. A refinement network corrects misalignments, while generative models fill occluded regions. Experiments show CloTH-VTON+ surpasses prior systems and can extend to video and multi-pose try-on.

3. Debapriya Roy, Sanchayan Santra, "Incorporating Human Body Shape Guidance for Cloth Warping in Model-to-Person Virtual Try-On Problems" [3] Most VTON methods require separate garment images, which limits data availability. Roy

and Santra propose a **model-to-person warping approach**, where clothing is segmented directly from model photos and aligned to the target person. This reduces dependency on isolated product images and improves accuracy for complex fabric patterns. Tests on public datasets confirm better performance compared to benchmark methods.

4. Xiaoyang Lv, Bo Zhang, Jie Li, Yangjie Cao, Cong Yang, “Multi-Scene Virtual Try-On Network Guided by Attributes” [4] Conventional VTON systems often need personal photos, raising privacy concerns. **MS-VTON** avoids this by generating try-on images using only garment inputs and descriptive attributes (e.g., skin tone, viewing angle). The framework has two stages: a Scene Learning Network creates coarse outputs, and a Content Learning Network refines textures and details. Results show strong realism, with MS-VTON achieving a competitive FID score of 9.8.

5. Matteo Fincato, “VITON-GT: An Image-Based Virtual Try-On Model with Geometric Transformations” [5] Fincato introduces **VITON-GT**, which applies multiple geometric transformations to garments before synthesizing the final try-on image. The two-stage transformation module ensures accurate garment projection, while the synthesis module produces realistic outputs. Validated on t-shirt datasets and extended to diverse clothing categories, VITON-GT demonstrates strong generalization and photorealism compared to earlier baselines.

6. Agnè Lagè, Kristina Accutane, “Virtual Try-On Technologies in the Clothing Industry: Basic Block Pattern Modification” [6] This work connects VTON technology with traditional garment design. The authors emphasize the role of **basic block patterns** and ease allowances in clothing construction. These foundational templates define garment shapes and comfort levels, and updating them is essential for integrating 2D/3D design into virtual try-on systems.

7. Prof. Suvarna Bahir, Shivani Shedage, “Image-Based Virtual Try-On Clothes” [7] Bahir and Shedage present a practical VTON system for e-commerce. Their approach adjusts garments according to the wearer’s pose, generates segmentation maps, and produces photo-realistic try-on images. The goal is to enhance customer satisfaction by improving fit visualization in online shopping environments.

### III OBJECTIVE

To create a more engaging and realistic online shopping experience, a virtual try-on system can be developed that integrates several advanced features. The system should focus on enhancing personalization by allowing customers to interact with garments in a way that feels tailored to their preferences. Cutting-edge technologies such as computer vision, artificial intelligence, and augmented reality can be employed to simulate clothing with high precision. A key component is accurate body mapping, where algorithms capture the user’s body shape and dimensions to ensure garments fit naturally in the virtual environment. Beyond fit, the system should also replicate real fabric behaviors, including draping, stiffness, and how materials respond under

different lighting conditions, to provide a lifelike representation of clothing. Finally, interactive tools such as zoom, rotation, and garment adjustments can be incorporated to increase user engagement, giving shoppers the ability to explore their virtual appearance from multiple perspectives and achieve a more immersive experience.

### IV. SYSTEM ARCHITECTURE

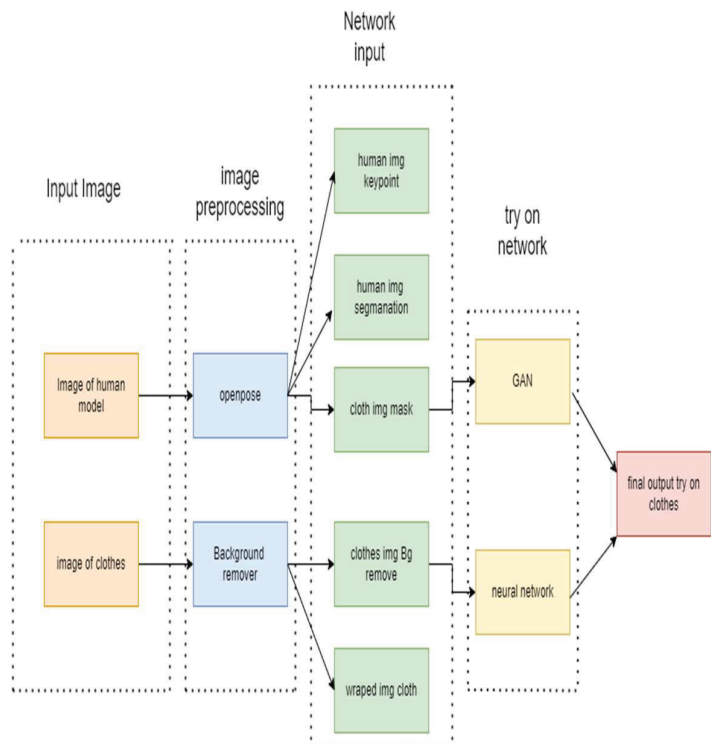


Fig 4.1 System Architecture

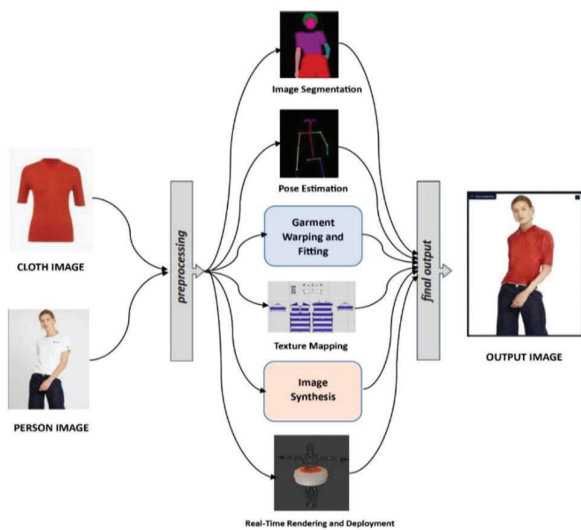
### V. METHODOLOGY

The proposed framework combines multiple deep learning and computer vision methods—such as DeepLabV3+, Media Pipe, Thin Plate Spline (TPS), Neural Style Transfer, and StyleGAN—to enable virtual try-on functionality. The system takes two inputs: a photo of the user and an image of the garment. These inputs are processed through a series of interconnected modules, ultimately producing a synthesized output that shows the user wearing the selected clothing item. During preprocessing, several backend operations are carried out to prepare the inputs and ensure the system produces the intended virtual try-on output.

#### A. Image Segmentation

During preprocessing, image segmentation is employed to separate distinct regions within the user’s photo. This step isolates the person from the background and highlights the body areas where the garment will be applied. For semantic

segmentation, the DeepLabV3+ architecture is utilized because it effectively captures contextual information at multiple scales through aroous spatial pyramid pooling. The input image is passed through the segmentation network to produce a mask that differentiates between body parts, clothing, and background. To enhance the quality of this mask, refinement techniques such as morphological filtering and conditional random fields are applied, which help smooth boundaries and eliminate irregularities. Achieving precise segmentation is critical, as inaccuracies at this stage can compromise the realism and visual quality of the final virtual try-on output.



### B. Pose Estimation

Pose estimation is used to determine the spatial configuration of the user's body joints, providing insight into posture and orientation. Media Pipe Pose is employed to detect key landmarks such as shoulders, elbows, wrists, hips, and knees, with the capability of identifying up to 33 distinct points that describe the human body's structure.

The extracted landmarks are compiled into a pose map, which serves as a guide for aligning the garment image with the user's body. By ensuring that the clothing follows the detected posture and proportions, accurate pose estimation enhances the realism and natural appearance of the virtual try-on output.

### C. Garment Fitting and Warping

After identifying the user's body structure, the clothing image must be reshaped to match the detected pose. To achieve this, Thin Plate Spline (TPS) transformation is applied, which enables smooth geometric deformation while maintaining the garment's overall integrity.

By establishing correspondences between control points on the garment and the user's body landmarks, TPS adjusts

the clothing image so that it naturally conforms to the wearer's shape. This process ensures that critical garment features—such as sleeves, neckline, and waistline—are properly aligned with the user's body, resulting in a realistic and visually coherent virtual try-on experience.

### D. Texture Mapping

During the garment warping process, visual inconsistencies may arise due to variations in lighting or texture. To address this, Neural Style Transfer is applied, ensuring that the original fabric's patterns, colors, and texture details are preserved. This technique extracts stylistic features from the garment image and overlays them onto the synthesized output while maintaining the structural integrity of the user's photo. Leveraging deep convolutional networks, the method produces results that are visually coherent and realistic, enhancing the overall fidelity of the virtual try-on system.

### E. Image Synthesis

The final stage focuses on refining the virtual try-on image to improve detail quality and achieve seamless blending between the garment and the user. StyleGAN is employed as a refinement module, enhancing textures and ensuring photorealistic integration.

StyleGAN operates within a generative adversarial framework: the generator creates synthetic images from latent representations, while the discriminator evaluates their realism against actual images. Through iterative adversarial training, the generator progressively improves, producing outputs that appear natural and visually consistent. As a result, the rendered garment aligns convincingly with the user's body and environment, delivering a polished and lifelike try-on experience.

## VI. Result

The virtual try-on framework integrates multiple deep learning modules to generate realistic clothing visualizations. Experimental evaluations demonstrate that the system successfully preserves garment textures, maintains accurate alignment between the body and clothing, and ensures consistent structural placement across varying user poses. Moreover, the framework is optimized to deliver high-quality outputs across different devices, highlighting its potential for deployment in real-world online fashion applications. This adaptability makes it a promising solution for enhancing digital shopping experiences by providing users with lifelike previews of garments before purchase.

```
[ ] output_grid = np.concatenate([np.array(Image.open('Data_preprocessing/test_img/000001_0.p  
np.array(Image.open('Data_preprocessing/test_color/000001_1.png')),  
np.array(Image.open('results/test/try-on/000001_0.png'))], axis=1)  
image_grid = Image.fromarray(output_grid)  
image_grid
```



## VII. Conclusion

This study introduces an efficient, user-centric virtual try-on framework. By integrating modules for image segmentation, pose estimation, garment warping, texture preservation, and image synthesis, the system is capable of producing realistic visualizations of clothing on user images. Such functionality has the potential to reduce product return rates and strengthen customer trust in online shopping platforms.

Despite its effectiveness, challenges such as occlusion management and achieving real-time performance remain open research problems. Continued advancements in model architectures and computational resources are expected to further improve both the accuracy and realism of future virtual try-on systems, making them increasingly viable for large-scale commercial deployment.

## REFERENCES

- [1] .D. Roy, S. Santra, and B. Chanda, "LGVTON: A Landmark Guided Approach to Virtual TryOn," 2020, arXiv. doi: 10.48550/ARXIV.2004.00562.
- [2] D. Song et al., "Image-Based Virtual Try-On: A Survey," 2023, arXiv. doi: 10.48550/ARXIV.2311.04811.
- [3] S. Huber, R. Poranne, and S. Coros, "Designing actuation systems for animatronic figures via globally optimal discrete search," *ACM Trans. Graph.*, vol. 40, no. 4, pp. 1–10, Aug. 2021, doi: 10.1145/3450626.3459867.
- [4] S. He, Y.-Z. Song, and T. Xiang, "Style-Based Global Appearance Flow for Virtual Try-On," 2022, arXiv. doi: 10.48550/ARXIV.2204.01046.
- [5] Y. Liu, M. Zhao, Z. Zhang, H. Zhang, and S. Yan, "Arbitrary Virtual Try-On Network: Characteristics Preservation and Trade-off between Body and Clothing," 2021, arXiv. doi: 10.48550/ARXIV.2111.12346.
- [6] T. Islam, A. Miron, X. Liu, and Y. Li, "Deep Learning in Virtual Try-On: A Comprehensive Survey," *IEEE Access*, vol. 12, pp. 29475–29502, 2024, doi: 10.1109/ACCESS.2024.3368612.
- [7] B. A. S. N, A. S, and J. V, "Virtual Dressing Room Application Using GANs," in 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India: IEEE, Mar. 2023, pp. 112–116. doi: 10.1109/ICACCS57279.2023.10113074.
- [8] S. Pandey, Y. Srivastava, Y. Meena, and R. K. Dewang, "CLOTON: A GAN based approach for Clothing Try-On," in 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India: IEEE, Aug. 2021, pp. 595–601. doi: 10.1109/SPIN52536.2021.9565973.
- [9] S. Ishikawa and T. Ikenaga, "Image-based virtual try-on system with clothing extraction module that adapts to any posture," *Computers & Graphics*, vol. 106, pp. 161–173, Aug. 2022, doi: 10.1016/j.cag.2022.06.007.
- [10] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, pp. 834–848, Apr. 2018, doi: 10.1109/TPAMI.2017.2699184.