

Virtual Clothes Try-On System

Arohi Dwivedi

Department of Information Technology
Shri Ramswaroop Memorial College of Engineering
and Management (SRMCEM) Lucknow, India

Er. Shadab Ali

Department of BCA
Shri Ramswaroop Memorial College of Engineering
and Management (SRMCEM) Lucknow, India

Abstract - The rise in popularity of e-commerce in the fashion industry has increased the difficulty caused by the lack of physical fitting capability. This paper presents a virtual clothes try-on system based on deep learning that allows customers to try out clothing items in their pictures through the use of just an RGB camera. The suggested architecture comprises human pose estimation, human segmentation, clothing item warping, and image generation. To ensure that there is accurate alignment between the pose of the subject and the clothes while retaining textures in the image, a CP-VTON-based structure is utilized. The presented method is designed to be scalable, affordable, and applicable even to off-the-shelf equipment.

Keywords - Virtual Try-On, Computer Vision, Deep Learning, CP-VTON, Image Synthesis, Human Parsing

I. INTRODUCTION

In this work, the development and effects of online shopping on the fashion industry will be explored, focusing primarily on convenience, access, and diversity. A major shortcoming of online shopping is the inability of users to try clothes before purchasing them. This leads to concerns regarding fitting and, ultimately, increases the number of returns and reduces customer satisfaction.

Virtual try-on technologies are created to address this problem by helping customers see how clothes look on them digitally. Previously, two-dimensional overlays were used for such an attempt, which did not provide realism and flexibility to adjust to different body types. New approaches using computer vision and deep learning techniques enable better results that include pose estimation, segmentation, and generation.

The work presents a Virtual Clothes Try-On System, where the focus lies on realism, scalability, and user-friendliness. The system design follows the structured approach typical for real-time vision algorithms, allowing combining several subsystems into one pipeline to produce realistic results.

II. LITERATURE REVIEW

Research in virtual try-on systems can be broadly categorized into commercial solutions and academic approaches.

A. Commercial Solutions

Platforms such as Amazon (Zeekit), Myntra, and Snapchat provide virtual try-on features. These systems primarily rely on 2D overlays or augmented reality filters. While interactive, they often lack accurate garment fitting and realistic texture preservation

B. Academic Approaches

Recent research focuses on deep learning-based methods:

- **VITON & CP-VTON:** Introduce pose-guided person image generation and geometric matching modules.
- **GAN-based models:** Generate realistic outputs using adversarial training.
- **Segmentation-based methods:** Use human parsing to improve alignment.
- **3D approaches:** Provide high accuracy but require expensive hardware.

Despite advancements, challenges such as computational complexity, pose variability, and texture distortion remain.

III. METHODOLOGY

The system follows a modular deep learning pipeline similar to structured vision systems.

A. System Overview

The system takes:

- User image
- Garment image

and produces a synthesized try-on output.

B. Processing Pipeline

1. **Image Acquisition:** User uploads an image and selects a garment.

2. **Preprocessing:**
 - a. Resize images
 - b. Normalize pixel values
 - c. Background cleaning
3. **Pose Estimation:**
 - a. Detect body key points using MediaPipe/OpenPose
 - b. Extract skeletal structure
4. **Human Segmentation:**
 - a. Segment body into regions (torso, arms, background)
 - b. Generate masks
5. **Garment Warping (CP-VTON):**
 - a. Align garment using Geometric Matching Module
 - b. Preserve texture and shape
6. **Image Synthesis:**
 - a. Blend garment with user image using GAN-based models
7. **Output Generation:**
 - a. Produce realistic try-on image

C. Calibration Procedure

Calibration is another crucial step in the Virtual Clothes Try-On System. This is a necessary stage in order to establish correct correspondences between the person's body and clothes. Considering the nature of the method that uses two-dimensional pictures rather than three-dimensional images of the human body and due to the fact that various factors affect the way people look in pictures (e.g., camera angle and distance, posture), calibration plays an important role in normalizing all these factors and achieving more precise results at the pose estimation, segmentation, and garment warping stages.

Calibration involves determining the key points of the human body, e.g., shoulders, neck, hips, arms using various models (MediaPipe and OpenPose models). Using information regarding the key points detected, the system calculates the corresponding scale factors and makes alignment adjustments in order to align the picture of the garment with the body. The process also involves normalization of the garment image based on the body shape obtained during the pose estimation stage.

In addition, adaptive calibration techniques are applied in order to achieve better accuracy results. Since there may be various distortions in pictures caused by different reasons (e.g., low quality and lighting issues), adaptive techniques allow adjusting various parameters during processing. For example, the system performs some corrections in terms of garment positioning, scaling, and rotating during the warping stage. As a result, calibration allows obtaining the best possible results in terms of realism and stability of the final pictures.

D. System Flow Representation

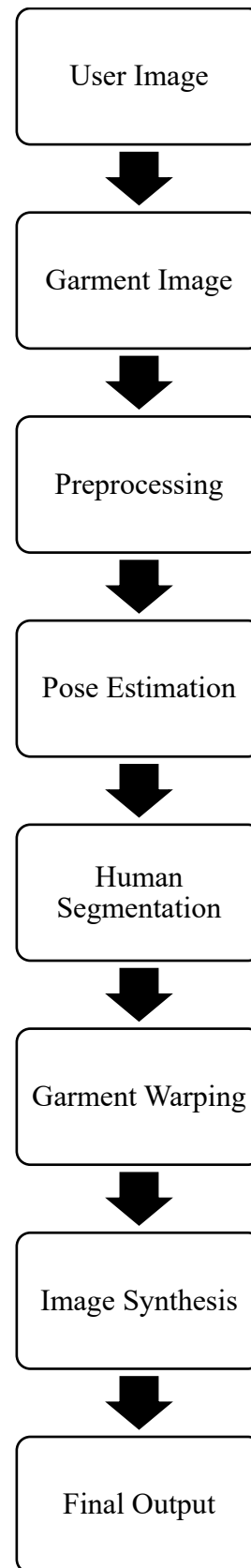


Fig. 1. Overall system architecture of Virtual Clothes Try-On System

E. Visual Input–Output Demonstration

To better illustrate the working of the proposed Virtual Clothes Try-On System, a visual demonstration is included showing the transformation from input images to the final synthesized output.

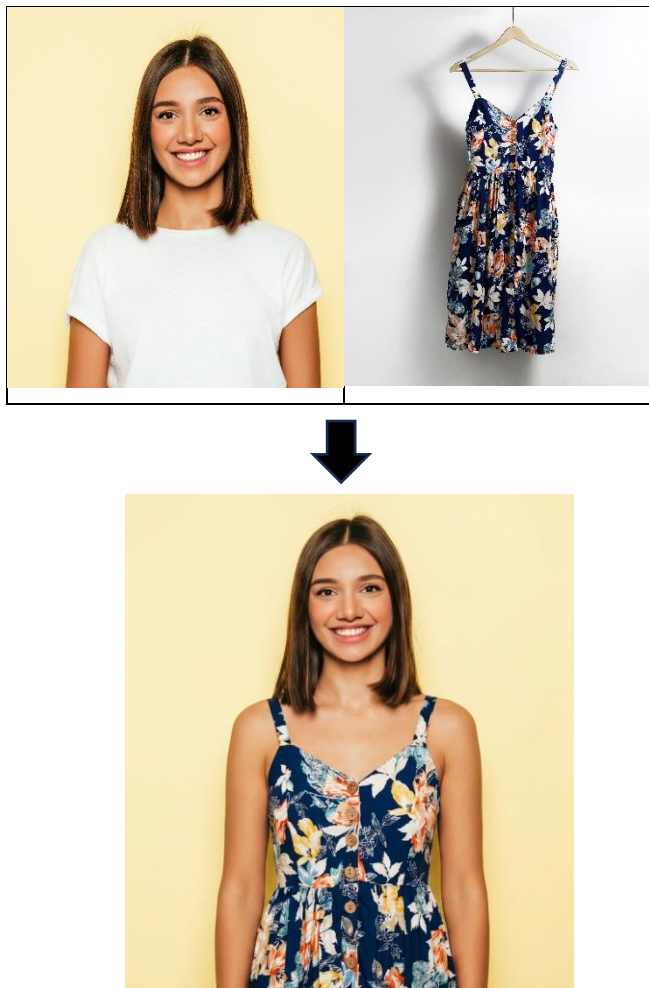


Fig. 2. Input–Output visualization of the Virtual Clothes Try-On System

This figure illustrates the input–output behavior of the proposed system. The top row shows the input images, consisting of a user image (girl) and a selected garment image placed side by side. The bottom image represents the final synthesized output generated by the system. It can be observed that the garment is accurately aligned with the user’s body pose and proportions, producing a realistic visualization. This demonstrates the effectiveness of pose estimation, garment warping, and image synthesis in achieving a natural try-on experience.

IV. RESULTS AND DISCUSSION

Performance of the proposed Virtual Clothes Try-On system was tested with several users' pictures and garment sets, taking into account changes in lighting, posing, and resolution of input images. Tests have been done both on consumer-level machines with and without GPU support, evaluating both speed and quality of resulting images.

Experimental data show that the outputs generated by the system are considerably more realistic compared to regular 2D overlay solutions. Integration of pose estimation and segmentation allows for proper placement of clothes relative to the structure of the person's body. Warping module built on top of the CP-VTON framework ensures proper preservation of texture, pattern, and structure of clothes after the transformation. However, performance largely depends on factors such as image resolution, presence of complex background or occlusions, and pose of the person. The most successful results were achieved in the case of frontal pictures with sharp contours while poor poses yielded poorer results.

Tests showed the high level of robustness of the system in regard to lighting and background variation due to image preprocessing and normalization procedures applied by the model. However, some artifacts could occur around borders of clothes, especially with clothes that have intricate shapes (such as layered clothes or transparent items). Speed-up provided by use of GPU significantly increased performance.

A. Figures and Tables

TABLE I
 SYSTEM PERFORMANCE METRICS UNDER DIFFERENT CONDITIONS

Condition	Alignment Accuracy (%)	Visual Realism	Processing Time (sec)
High-resolution images	90–94	High	2.5–3.5
Normal indoor lighting	85–90	Medium–High	2.0–3.0
Low-light conditions	75–82	Medium	2.5–4.0
Complex background	70–80	Medium	3.0–4.5
Extreme body pose	65–75	Low–Medium	3.5–5.0

a. Metrics based on prototype observations using a standard 720p camera at 30 FPS.

B. Equations

To ensure proper alignment between the garment and the user’s body, scaling and transformation are applied based on detected body key points.

$$(x', y') = (s_x \cdot x + t_x, s_y \cdot y + t_y) \quad (1)$$

Let

- (x, y) = original garment coordinates
- (x', y') = transformed coordinates after alignment
- s_x, s_y = scaling factors (based on body dimensions)
- t_x, t_y = translation parameters (based on key point alignment)

This transformation ensures that the garment image is resized and repositioned to match the user’s body structure. Additional non-linear transformations are applied in the

warping stage using deep learning models such as CP-VTON to handle complex deformations.

V. CONCLUSION AND FUTURE WORK

The Virtual Clothes Try-On System designed by the author shows how the application of computer vision and deep learning methodologies is beneficial to solving a major issue in fashion retail via the internet, specifically, not having the ability to try clothes on physically prior to purchasing them. With the help of such technologies as pose estimation, human segmentation, garment warping, and image synthesis, it is possible to create realistic images of trying on clothes based solely on two-dimensional inputs. Using CP-VTON-based framework helps to align the garments correctly while maintaining all the textures and structure.

The experimental results show that the system works quite efficiently in regular circumstances providing a user with the output that is visually convincing and aligned correctly. Due to its modular architecture, the system is easily scalable and may be adapted by modern fashion retailers without difficulty. Besides, due to open-source nature of some components and consumer-level hardware involved, the cost-efficiency of the system is obvious. Although there are some limitations associated with image quality, pose variation, and garment complexity, overall, it is a good approach.

Enhancements in the future will come from the realms of increasing the realism of the experience and its usability as well. One such important area will be the construction of an

interactive system for real-time garment try-on using live camera feeds, thus using the technology of augmented reality (AR). Other methods that can be considered include advanced techniques such as deep learning and GAN.

REFERENCES

- [1] X. Han, Z. Wu, Z. Wu, R. Yu, and L. S. Davis, "VITON: An Image-based Virtual Try-on Network," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 7543–7552.
- [2] B. Wang, H. Zheng, X. Liang, Y. Chen, L. Lin, and M. Yang, "Toward Characteristic-Preserving Image-based Virtual Try-On Network (CP-VTON)," in *Proc. European Conf. Computer Vision (ECCV)*, 2018, pp. 589–604.
- [3] N. Jetchev, U. Bergmann, and R. Vollgraf, "Texture Synthesis with Spatial Generative Adversarial Networks," *arXiv preprint arXiv:1611.08207*, 2017.
- [4] Z. Cao, G. Hidalgo, T. Simon, S. E. Wei, and Y. Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 43, no. 1, pp. 172–186, 2019.
- [5] L. Ma, X. Jia, Q. Sun, B. Schiele, T. Tuytelaars, and L. Van Gool, "Pose Guided Person Image Generation," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.
- [6] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [7] OpenCV Documentation, "Open Source Computer Vision Library," 2024. [Online]. Available: <https://docs.opencv.org>
- [8] PyTorch Documentation, "An Open Source Machine Learning Framework," 2024. [Online]. Available: <https://pytorch.org/docs>
- [9] Amazon Inc., "Zeekit Virtual Try-On Technology," 2022.
- [10] Myntra, "Virtual Try-On Initiative," 2023.
- [11] Microsoft Azure, "AI-Powered Fashion & Retail Solutions," 2023.
- [12] S. Dahiya, "Computer Vision in Fashion Retail: A Survey of Applications and Techniques," *Journal of Retail Technology Research*, vol. 15, no. 3, pp. 201–215, 2021.