Cloth segmentation, styling and pose transfer for different body types

Aadhya Raj
aadhya19002@iiitd.ac.in
Indraprastha Institute of Information Technology
Delhi, India

Aparna Jha
aparna19410@iiitd.ac.in
Indraprastha Institute of Information Technology
Delhi, India

Arihant Singh
arihant19298@iiitd.ac.in
Indraprastha Institute of Information Technology
Delhi, India

Harshit Singh
harshit19424@iiitd.ac.in
Indraprastha Institute of Information Technology
Delhi, India

Meet Modi
meet19435@iiitd.ac.in
Indraprastha Institute of Information Technology
Delhi, India

Smiti Chhabra
smiti19112@iiitd.ac.in
Indraprastha Institute of Information Technology
Delhi, India

Abstract

Recent years have witnessed the increasing demands for online shopping for fashion items. Despite the convenience online fashion shopping provides, consumers are concerned about how a particular fashion item would look on them when buying apparel online. Thus, allowing consumers to virtually see how the clothes look on different body types will enhance the shopping experience, transforming the way people shop for clothes. The aim of this project is to take a piece of clothing and virtually put on those clothes on different body types to help with this issue.

1. Introduction

1.1. Motivation

This pandemic saw an increase in online shopping due to mobility restrictions. Ordering something with a click from anywhere and getting it delivered wherever you want is a luxury that saves a lot of time and effort. However, one major challenge is wondering how the product would look on us. This served as the motivation for this project. This project will help us try different clothes on different body types with different poses. We believe this will solve one of the significant issues with online shopping and help increase its reach, comfort, and accessibility.

1.2. Problem Statement

One of the challenges of online shopping remains the mystery of imagining ourselves in the selected clothes. Through our project, we aim to cater to this problem. We aim to develop a system that segments clothes from images, styles them, and applies pose transfer techniques to get different poses in that output on different body types.

1.3. Challenges Involved

- While obtaining the cloth mask from different images from the dataset, the hair could be present on top of the dress. In that case, transferring the dress from one pose to another could lead to mismatches because the algorithm fails to determine what could be present in the location where hair was present on top of the dress.
- Transferring initial poses to complex poses could lead to unclear output poses or being transferred improperly because the poses are rarely present in the dataset.

2. Related Work

- **Paper 1** [1]: This paper by Xintong Han, Zuxuan Wu, Zhe Wu, Ruichi Yu, and Larry S. Davis presents an image-based Virtual Try-On Network (VITON) without using 3D information that claims to transfer the desired clothing item onto the corresponding region of a person using a coarse-to-fine strategy. A coarse sample is first generated using six convolutional layered multi-task encoder-decoders conditioned on a detailed clothing-agnostic person representation. The coarse results are further enhanced with a refinement network that learns the optimal composition. The model fails when applied to rarely-seen poses or when there is a huge mismatch between current and target clothing shapes.

- **Paper 2** [2]: This paper by Aiyu Cui, Daniel McKee, and Svetlana Lazebnik proposes a flexible person generation framework called Dressing in Order (DiOr), which supports 2D pose transfer, virtual try-on, and several fashion editing tasks. They represent each person as a (pose, body,{garments}) tuple. For the

implementation, they ran an off-the-shelf human parser on the source garment to obtain the masked garment segment. To perform pose transfer, they set the body image and the garment set to be those of the source person and render them in the target pose. This method also has some limitations and failures. The complex or rarely seen poses are not always rendered correctly, unusual garment shapes are not preserved, some ghosting artifacts are present, and holes in garments are not always filled in properly.

- **Paper 3** [3] : In this paper by Zhen Zhu, Tengteng Huang, Baoguang Shi, Miao Yu, Bofei Wang, and Xiang Bai on progressive pose attention transfer for person image generation, Pose-Attentional Transfer Blocks were used for transferring poses. They consider the appearance features along with the pose using progressive pose transfer. Further, a Pose-Attentional Transfer Network is proposed to make this process efficient and straightforward that can be extended to non-rigid bodies. Two datasets are used, along with a new metric specifically designed to evaluate the shape consistency.

- **Paper 4** [4]: In the paper Yining Li, Chen Huang, and Chen Change Loy take the source and target image and first use a variant of u-net to encode the image and target pose. Then using the reference and target pose, a 3-d flow map is created. To get an image in target pose, the features of the reference image are warped through a 3-d flow map and then to a visibility map. A visibility map is created to consider the lost pixels due to occlusions. Then concatenating the warped image features and target pose features take place to obtain the pose-transferred image.

- **Paper 5** [5]: In the paper by George. A. Cushen and Mark. S. Nixon, they present a real-time clothing segmentation model for video and single images. It initializes points on the upper body clothing instead of detecting the face of the subject using distance metrics which is advantageous because it helps prevent skin segmentation instead of clothing. It takes advantage of intensity and hue histograms for efficient segmentation.

3. Methodologies

3.1. Techniques/Algorithms

The first step of the problem would be segmenting the clothes from the human body, which we can perform with the help of Adaptive Content Generation and Preservation Network (ACGPN), which keeps clothes and humans intact. We can also use pre-trained U-2-Net models for cloth segmentation and Saliency Map Generation, which accurately segments cloth components based on upper body cloth, lower body cloth, full body cloth, and background. We will use neural networks style transfer for style transfer in the segmented region cases. For cases where the transfer of a single segment of cloth is required, we will use the GAN G1, G2, and G3 to create the pose map, synthesized map, and finally transfer the segmented cloth on the input image. Finally, we will use the saliency full body map and the map with the new cloth/style transferred to perform blending to smooth out the edges S, which will blend the style transferred image with the Saliency map created using U-2-Net. For pose transfer, we plan to use the pose transfer algorithm from DiOr on different body types images obtained from clothes segmentation and wrapping.

3.2. Novelty

While there are several algorithms for pose transfer and cloth segmentation, we couldn't find any for different body types. Through this project, we aim towards inclusivity by accommodating different body types. Also, we couldn't find any single algorithm to perform all the tasks of cloth segmentation, styling, and pose transfer together which will be accomplished through our project.

3.3. Result Analysis and Evaluation

We can evaluate the models based on the clarity of images formed primarily for regions where a portion of the image is replaced, like hair and smoothness for edges. Next, we can also check if the results are better in high contrast, specific body types, and certain poses. We can compare the initial image with the final output to see which and how the features change through the various stages.

4. Potential Contributions

- Aadhya Raj : Work on the backend portion for cloth segmentation, styling, and pose transfer.
- Aparna Jha : Work on the backend portion for cloth segmentation, styling, and pose transfer.
- Arihant Singh : Work on the website's frontend portion and integrate the backend and frontend.
- Harshit Singh : Work on the website's frontend portion and integrate the backend and frontend.
- Meet Modi : Work on the backend portion for cloth segmentation, styling, and pose transfer.
- Smiti Chhabra : Work on the backend portion for cloth segmentation, styling, and pose transfer.

References

[1] Han, Xintong, Zuxuan Wu, Zhe Wu, Ruichi Yu, and Larry S. Davis. "Viton: An image-based virtual try-on network." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7543-7552. 2018.

[2] Cui, Aiyu, Daniel McKee, and Svetlana Lazebnik. "Dressing in order: Recurrent person image generation for pose transfer, virtual try-on and outfit editing." In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 14638-14647. 2021.

[3] Zhu, Zhen, Tengteng Huang, Baoguang Shi, Miao Yu, Bofei Wang, and Xiang Bai. "Progressive pose attention transfer for person image generation." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2347-2356. 2019.

[4] Li, Yining, Chen Huang, and Chen Change Loy. "Dense intrinsic appearance flow for human pose transfer." In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3693-3702. 2019.

[5] Cushen, George A., and Mark S. Nixon. "Real-time semantic clothing segmentation." In *Advances in Visual Computing: 8th International Symposium, ISVC 2012, Rethymnon, Crete, Greece, July 16-18, 2012, Revised Selected Papers, Part I 8*, pp. 272-281. Springer Berlin Heidelberg, 2012.