VTON - A 2D Virtual Cloth Tryon

CAPSTONE PROJECT

- Dilshath Shaik

Abstract

- The concept of using images to virtually try on clothes involves transferring a chosen piece of attire onto a person's image, fitting it appropriately to the body.
- Despite significant advancements in this area, the quality of the resulting images often falls short, primarily due to their limited resolution (for instance, 256x192 pixels), which doesn’t meet the expectations of online shoppers.
- The root of this issue lies in a few key challenges: the higher the resolution, the more noticeable the discrepancies between the adjusted clothing and the intended area on the person become.
- To overcome these hurdles, I have come across a new approach to virtual try-ons.
- This technique enables the creation of virtual try-on images at a resolution of 1024x768 pixels.
- The method begins with crafting a segmentation map to direct the synthesis process, followed by an initial adjustment of the clothing item on the model's image.
- These innovations are specifically designed to address alignment mismatches and ensure the preservation of detail in high-resolution inputs.

Background

- The project on high-resolution 2D Virtual Try-On, was developed in response to the limitations of earlier virtual try-on technologies.
- During times like the COVID-19 pandemic, when physical retail experiences were limited and consumers relied heavily on online shopping, a technology like VTON proved invaluable in offering a safe, contactless, and enhanced virtual try-on experience, closely mimicking the in-store fitting room experience from the comfort of one’s home.
- Initially, these technologies could only provide a basic preview of how clothes might look on a person, often with low-quality images that didn't show much detail.
- As online shopping became more popular, there was a clear need for better virtual try-on experiences that could offer more realistic and detailed visualizations.
- This is because high-quality images help customers make better decisions when shopping online, leading to greater satisfaction and potentially fewer returns.
- This project represents a significant step forward in making online shopping more immersive and accurate, providing a better experience for consumers.

Problem Statement

The main challenge that the project, 2D virtual try-on aims to solve is that most online clothes shopping sites use virtual try-on technology that doesn't show clothes clearly or realistically enough. When people shop online, they often get a rough idea of how an outfit might look on them, but the details are blurry and sometimes the fit seems off. This can lead to disappointment when the order arrives—it's just not what they expected. Plus, fitting clothes onto different body types without making them look weird or out of place is really tough with the current technology. The goal is to make a better system that lets shoppers see outfits on them in high quality, just like in real life, making online shopping easier, more satisfying, and hopefully cutting down on the hassle of returns.

Literature Review

The literature on virtual try-on (VTON) technology highlights a journey of continuous innovation aimed at enhancing the online shopping experience. Initial efforts in VTON focused on simple 2D overlays, allowing users to see clothing items on a generic silhouette, but these often lacked realism and failed to account for individual body shapes and garment physics. As the field progressed, the adoption of deep learning marked a significant improvement, enabling more realistic simulations of how clothes fit and look on different body types. Despite these advancements, a persistent challenge has been the production of high-resolution VTON images that can accurately capture and convey the fine details of fabric textures and patterns, which are crucial for a true-to-life try-on experience. Research has increasingly focused on addressing this gap, exploring methods to upscale images without losing quality and integrating 3D body models to overcome the limitations of 2D approaches. However, achieving a balance between high-resolution imagery and computational efficiency remains a technical hurdle. The literature underscores the importance of further innovation in VTON technology, particularly in enhancing image resolution and realism, to meet the evolving expectations of online consumers. This ongoing quest highlights the dynamic nature of the field and the critical role of advanced imaging and modeling techniques in shaping the future of online retail.

Methodology

This project introduces a novel approach to high-resolution virtual try-on, which leverages a clothing-agnostic person representation, incorporating both pose and segmentation maps to effectively eliminate clothing information from input images. The core innovation lies in the ALIAS (ALIgnment-Aware Segment) normalization technique, designed to adeptly handle misalignments between the target clothing item and the person's image. By focusing on the semantic propagation of details through the network, ALIAS normalization ensures the accurate preservation of texture and garment details. The process culminates with the ALIAS generator, which synthesizes the final high-resolution try-on image. This generator is particularly adept at filling in misaligned regions with convincing texture and detail, facilitated by a multi-scale refinement strategy that operates at the feature level. This methodology aims to significantly enhance the realism and quality of virtual try-on images, addressing the long-standing challenge of maintaining high resolution and detail fidelity in digital garment fitting.

Data Collection Plan

- The data collection plan utilizes the VITON dataset, accessible at Kaggle.
- This dataset includes high-resolution images featuring a variety of clothing items and models in diverse poses, tailored for virtual try-on systems.
- The VTON dataset used in the project contains images with a resolution of 256 × 192 pixels.
- It encompasses a broad spectrum of garment styles and body types, vital for developing a system capable of accurately simulating how different clothes look on various body shapes, enhancing the realism and effectiveness of the virtual try-on experience.

Website: https://www.kaggle.com/datasets/seeanbooo/viton-dataset?resource=download

Evaluation Plan

- The evaluation plan for the VITON-HD project encompasses both quantitative and qualitative metrics to assess its performance comprehensively.
- Quantitatively, the plan includes measuring image resolution and detail using metrics like Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), alongside the accuracy of garment alignment with metrics such as Intersection over Union (IoU). Additionally, computational efficiency will be evaluated by recording the processing time and resources required for generating virtual try-on images. Qualitatively, user satisfaction surveys and expert reviews will provide insights into the visual quality, realism, and appeal of the generated images.
- Comparative analysis with existing virtual try-on systems through baseline comparisons and A/B testing will highlight VTON’s advancements in resolution, realism, and user engagement.
- Technical validation will further test the system's robustness across various garments and body types, coupled with an error analysis to identify and improve upon any limitations, ensuring a well-rounded evaluation of the project's success and areas for future enhancements.

Project Plan

The project plan for VITON-HD is designed to develop and evaluate a high-resolution virtual try-on system through a structured approach:

- **Data Preparation:** Collect and preprocess images from the VTON dataset, including segmentation and annotation for accurate model training.
- **System Development:** Implement the VTON framework using advanced computer vision and deep learning techniques, focusing on high-resolution image synthesis and the novel ALIAS normalization for precise garment fitting.
- **Model Training:** Train the VTON model on the prepared dataset, iteratively refining the model based on initial testing to optimize performance.
- **Evaluation:** Assess the system quantitatively and qualitatively, using metrics such as PSNR, SSIM, and IoU for technical evaluation, and conduct user surveys and expert reviews for user experience insights.
- **Optimization:** Analyze evaluation results to identify areas for improvement, optimizing the model for better accuracy, efficiency, and user satisfaction.
- **Integration and Testing:** Prepare the system for integration with online retail platforms, ensuring scalability and user-friendliness, followed by comprehensive testing to ensure robust performance across a wide range of scenarios.
- **Launch and Feedback Collection:** Deploy the VTON system for public use, collecting user feedback for ongoing refinement and updates to enhance the virtual try-on experience.

Conclusion & Discussion

- The VITON-HD project represents a significant leap forward in virtual try-on technology, addressing critical challenges in image resolution and garment alignment that have long impeded the realism and utility of virtual fitting rooms.
- By introducing the ALIAS normalization technique and leveraging advanced deep learning models, this project has successfully demonstrated the ability to produce high-resolution, realistic virtual try-on images that significantly enhance the online shopping experience.
- The system's capability to maintain the texture and detail of clothing items, even at high resolutions, marks a notable advancement over existing solutions.
- This project highlights the critical balance between high-resolution image generation and computational efficiency, underscoring the need for ongoing innovation to enhance realism in virtual try-ons while ensuring system scalability.
- Future efforts will likely focus on optimizing processing techniques and expanding dataset diversity to improve the system's applicability and performance across a broader range of garments and body types.

References

[1] Jetchev, Nikolay, and Urs Bergmann. "The conditional analogy gan: Swapping fashion articles on people images." Proceedings of the IEEE international conference on computer vision workshops. 2017.
[2] Han, Xintong, et al. "Viton: An image-based virtual try-on network." Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
[3] X. Han, Z. Wu, Z. Wu, R. Yu, and L. S. Davis, “VITON: An Image- based Virtual Try-On Network,” in CVPR, 2018.
[4] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In Proc. of the IEEE conference on computer vision and pattern recognition (CVPR), pages 1125–1134, 2017.
[5] Xintong Han, Zuxuan Wu, Zhe Wu, Ruichi Yu, and Larry S Davis. Viton: An image-based virtual try-on network. In Proc. of the IEEE conference on computer vision and pattern recognition (CVPR), pages 7543–7552, 2018.