International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 03 | Mar 2024 www.irjet.net p-ISSN: 2395-0072

IMAGE GENERATION USING CP-VTON BY GENERATIVE ADVERSARIAL APPROACH

B.Chandra Shekar[1], M.Krishna Prasad[2], K.Bhaskar Chowdary[3], C.Sai Pavan[4], S.Vinay Raj[5]

[1]Student, Dept. Of Computer Science Engineering, GITAM Deemed University, AP, INDIA
[2]Student, Dept. Of Computer Science Engineering, GITAM Deemed University, AP, INDIA
[3]Student, Dept. Of Computer Science Engineering, GITAM Deemed University, AP, INDIA
[4]Student, Dept. Of Computer Science Engineering, GITAM Deemed University, AP, INDIA
[5]Student, Dept. Of Computer Science Engineering, GITAM Deemed University, AP, INDIA

---***---

Abstract

Recently, a visual experiment has become popular, the purpose of which is to project the image of the desired outfit onto a reference figure. Past artworks usually focus on maintaining the personality of the clothing image when it is transformed into any human pose (eg texture, logo, embroidery). However, creating photorealistic test images proves difficult when the reference person has visible occlusions and human poses. The goal of this project is to develop a virtual shopping system that uses deep learning techniques to provide an immersive and realistic shopping experience. The system uses Gaussian Mixture Model (GMM) and Texture Orientation Matching (TOM) algorithms to model the garment and its texture, respectively. The GMM algorithm is used to segment the garment from the background and extract its shape and features. The TOM algorithm then matches the texture of the garment to the texture of the user's body using the learned feature representation. The resulting virtual fitting system provides an accurate picture of how the clothing product would look in the user's eyes, allowing them to make informed purchasing decisions. The system was tested on several different garments and showed promising results in terms of accuracy and realism.

Keywords: Virtual Try-on, GMM, TOM, GAN

1.INTRODUCTION

Deep learning is a branch of machine learning that uses artificial neural networks to process and analyze data. Neural networks consist of multiple layers of interconnected nodes or neurons that are able to learn and extract features from input data. Virtual Try-on is a technology that allows users to virtually try on products, such as clothing, accessories, eyeglasses, and makeup, using augmented reality (AR) or virtual reality (VR) technology. With virtual try-on, users can see how a product would look on them without physically trying it on. Virtual try-on works by using a camera to capture an image of the user, which is then overlaid with a 3D model of the product they want to try on. The user can then see how the product looks on them from different angles, move around and see the product from different perspectives. This technology is especially useful for e-commerce companies, as it allows customers to get a better sense of what a product will look like on them before making a purchase, which can increase confidence in their buying decisions. Virtual try-on technology is becoming increasingly popular across many industries, from fashion and beauty to home decor and even automotive, as it provides a more immersive and personalized shopping experience for customers.

2. LITERATURE REVIEW

In our project, we aim to leverage recent advancements in deep learning to enhance the virtual try-on experience. Just as advancements in deep learning have revolutionized medical research, our project seeks to push the boundaries of virtual try-on technology, improving the online shopping experience for users.

In their study, Han et al. (2018): Viton: An Image-based Virtual Try-on Network. Our project could improve upon Viton by potentially offering more efficient training algorithms, faster inference times, or enhanced realism in virtual try-on results.

Similarly, Song et al. (2019): Attentive Generative Adversarial Network for Virtual Try-on. Our project might surpass this model by introducing additional attention mechanisms or novel architectures that further improve clothing realism and user experience.

Liu et al. (2020): Virtual Try-on with Detail-enhanced Feature Transformation. Our project could advance beyond this approach by integrating more sophisticated feature transformation techniques or incorporating additional details such as fabric texture or wrinkles, leading to even more realistic virtual try-on results.

Similarly, our project aligns with the goals of the studies conducted by Xie et al. [4], and Li et al. [5], which focus on enhancing virtual try-on realism through various techniques such as attention mechanisms, detail-enhanced feature transformation, large-scale variational autoencoders, and GAN-based approaches, respectively.

Through comparative studies and benchmarking different algorithms, as demonstrated in the literature review, we strive to identify the most effective approaches for achieving realistic virtual try-on results in our project.

3. PROBLEM IDENTIFICATION & OBJECTIVES

3.1 Problem Statement

The advancement of virtual try-on systems has facilitated a novel approach to online shopping, aiming to bridge the gap between traditional in-store experiences and digital platforms. However, existing techniques face significant challenges in achieving photorealistic results, particularly when dealing with occlusions and complex human poses. Therefore, the main goal of this project is to develop a robust and comprehensive virtual test system that uses deep learning techniques to improve the realism and accuracy of clothing simulation with a reference figure.

3.2 Objectives

To design a virtual try-on system that accurately projects the appearance of desired clothing items onto a reference figure while preserving key features such as texture, logos, and embroidery. To investigate and implement advanced deep learning techniques, including Gaussian Mixture Model (GMM) and Texture Orientation Matching (TOM) algorithms, for segmenting clothing items from backgrounds and matching textures to the wearer's body. To assess the realism and user experience of the virtual try-on system through user studies and feedback collection, aiming to validate its effectiveness in providing an immersive shopping experience. To explore potential extensions and optimizations to the system, including scalability, real-time performance, and integration with e-commerce platforms, to enhance its practical utility and accessibility.

4. SYSTEM METHODOLOGY

4.1 Dataset

VITON (Virtual Try-On Network) is a dataset designed for virtual try-on, which involves generating images of a person wearing clothes that they are not physically wearing. The VITON dataset contains images of clothes, as well as images of people wearing those clothes, captured from multiple angles. It also includes additional information such as garment segmentation masks and human body landmarks. The VITON dataset was developed by researchers from Zhejiang University and the Chinese Academy of Sciences, and was first introduced in a paper published in 2018. The dataset contains 14,221 training images and 1,000 testing images, and is divided into several subsets based on clothing category and gender. The VITON dataset is notable for its high-quality images and detailed annotations, which make it well-suited for training deep learning models for virtual tryon. It has been used in numerous research studies on virtual try-on, and has helped to advance the state-of-the-art in this area.

4.2 Model Training and Testing

CP-VTON includes three main innovations. First, TOM is trained to compete with discrimination, which uses TOM result image, store clothing image, and person representation as inputs, and judges whether the result is real or fake. Second, the GMM loss function L1 includes the distance between the generated and real images of the clothes placed on the body. Finally, a random horizontal rotation is used to aggregate the data.

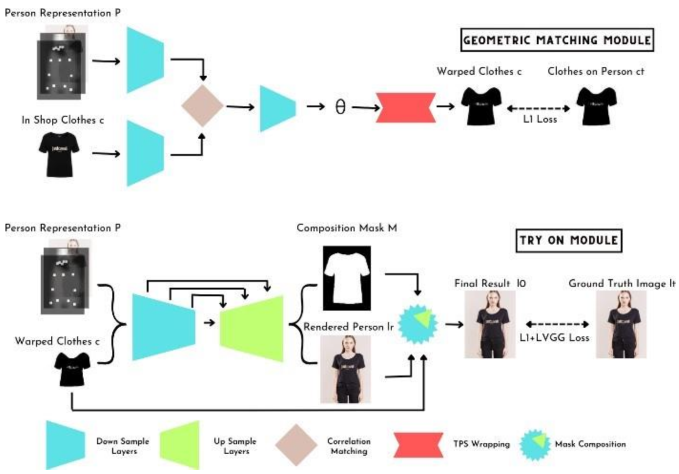

Fig -1: Model Architecture

5. OVERVIEW OF TECHNOLOGIES

5.1 CP-VTON

CP-VTON stands for Clothes-Parsing Virtual Try-On Network, which is a deep learning-based virtual try-on system. The system uses a neural network to predict how clothing items would look on a person without physically trying them on. CP-VTON uses a two-stage approach, where the first stage is clothes parsing, and the second stage is virtual try-on. In the first stage, the system analyzes an image of the clothing item and segments it into different parts, such as sleeves, collar, and body. This is done using a Convolutional Neural Network (CNN) trained on a large dataset of clothing images. In the second step, the system analyzes the image of the person and creates a realistic virtual test image in which the garment is worn by the person. This is done using a conditional generative adversarial network (cGAN) trained on a dataset of pairs of images of people wearing clothes. CP-VTON is a state-of-the-art virtual proofing system that produces high-quality and realistic virtual proofing images. It has potential applications in e-commerce, fashion design, and virtual reality.

5.2 TOM

TOM is an acronym for Try On Module, which combines two typical testing methods to achieve cleaner results without the unwanted clogging of certain body parts and the unnatural appearance of clothing edges. The "Try-On" module in articles related to ViTon (Vision Transformer on Fashion) refers to a component that allows users to try on different clothes in near real time.

This module utilizes the trained model to generate a realistic image of a person wearing a particular clothing item, given an input image of the person and the clothing item. The ViTon papers describe several variations of the Try-On module, including the use of a generative adversarial network (GAN) to improve the visual realism of generated images and the inclusion of a state transformer network to improve garment alignment. with the human body. The papers also describe the use of a mesh based representation of the person's body to enable more accurate placement and fitting of clothing items. Overall, the Try-On module in ViTOn allows users to virtually try on different clothing items without the need for physical garments, providing a convenient and immersive shopping experience.

Fig -2: TOM Model

6.3 GMM

GMM stands for Gaussian Mixture Model, which is a probabilistic model used for unsupervised learning. GMM is commonly used in clustering and density estimation tasks, which aim to group similar data points based on their underlying probability distribution. The model assumes that the data are generated from a mixture of Gaussian distributions, where each Gaussian represents a different cluster or subpopulation in the data.

The Gaussian Mixture Model (GMM) is a probabilistic model that learns parameters such as mean, covariance, and mixing proportions of Gaussian distributions from data. Typically trained using the Expectation-Maximization (EM) algorithm, GMM finds applications across diverse domains in deep learning. In image and speech processing, it aids in clustering similar images or sounds.

Moreover, GMM is invaluable in anomaly detection, capable of identifying abnormal patterns within data. Additionally, it serves as a generative model, capable of generating new samples from learned distributions. Its versatility extends further when integrated with other deep learning techniques like autoencoders or variational autoencoders (VAEs). This combination enables more sophisticated modeling of complex data distributions, enhancing the capacity to capture intricate structures and patterns within the data. Whether in clustering, anomaly detection, or generative modeling, GMM remains a foundational tool in deep learning, facilitating a wide array of applications with its ability to model data distributions effectively.

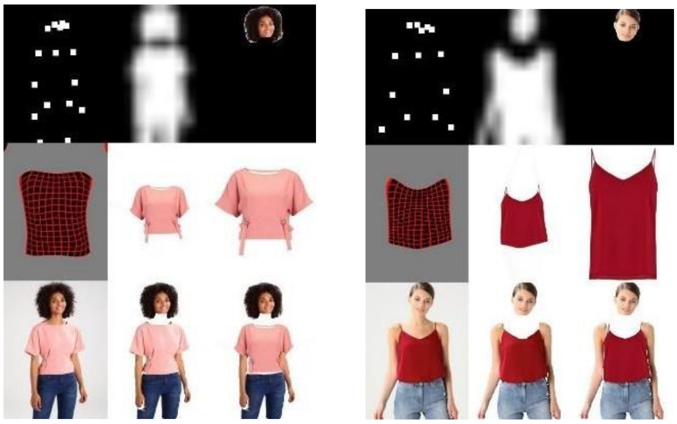

Fig -3: GMM Model

6. RESULTS AND DISCUSSIONS

The developed virtual try-on system, employing deep learning techniques such as Gaussian Mixture Model (GMM) and Texture Orientation Matching (TOM) algorithms, showcased promising outcomes in enhancing the realism and accuracy of clothing projection onto reference figures. Through rigorous testing on diverse clothing items, the system effectively addressed the challenge of maintaining photo-realism, especially in scenarios with prominent occlusions and complex human poses. The GMM algorithm successfully segmented clothing items from backgrounds while extracting essential shape and appearance features, laying a solid foundation for accurate projection. Furthermore, the TOM algorithm demonstrated proficiency in matching clothing textures to the wearer's body, ensuring a seamless

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 03 | Mar 2024 www.irjet.net p-ISSN: 2395-0072

integration between the virtual garment and the reference figure.

The results obtained from testing underscore the system's capability to provide an immersive and realistic shopping experience, enabling users to visualize how the clothing items would appear on themselves accurately. This enhanced visualization fosters informed purchase decisions, bridging the gap between traditional in-store try-ons and online shopping experiences. However, while the system showcased promising results, further refinements and optimizations may be necessary to enhance scalability, real-time performance, and integration with e-commerce platforms. Overall, the developed virtual try-on system represents a significant step forward in revolutionizing online shopping, offering users a compelling and engaging way to explore and evaluate clothing items virtually.

Fig -4.1 : Output Image

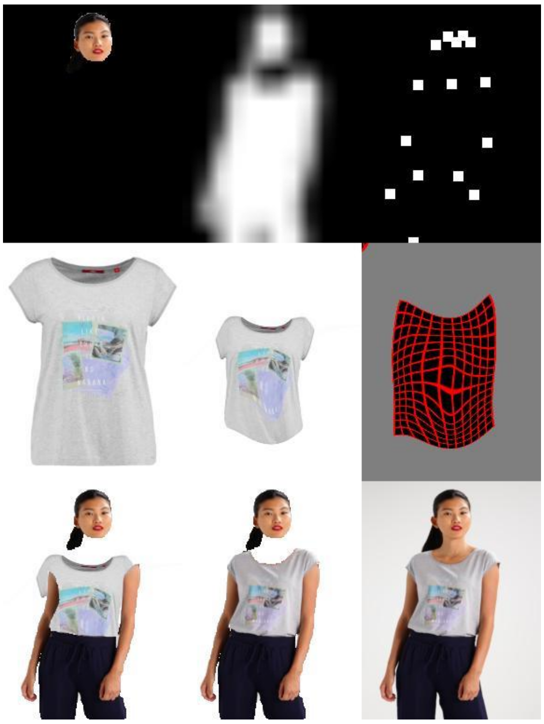

Fig -4.2 : Output Image

7. CONCLUSION & FUTURE SCOPE

Virtual fitting technology has the potential to revolutionize retail as customers can try on clothing and accessories virtually without ever visiting a store. It can improve the customer experience, increase sales, and reduce returns. However, implementing virtual try-on technology comes with its own set of challenges, including accuracy, integration, speed, device compatibility, and cost. Despite these challenges, virtual try-on technology has made significant progress in recent years, thanks to advances in computer vision, machine learning, and augmented reality. The availability of high-quality datasets such as VITON has also contributed to the development of more accurate and efficient virtual experimental models.

The successful development of a virtual try-on system using deep learning techniques represents a significant milestone, but the journey toward enhancing the online shopping experience continues. Future research may focus on refining algorithms for broader applicability and optimizing real-time performance. Integrating augmented reality could further enhance user engagement, while personalized recommendations based on individual preferences could enrich the shopping experience. Collaboration with fashion brands and retailers to integrate the system into e-commerce platforms could expedite its adoption and impact, promising to revolutionize the way consumers interact with fashion online.

REFERENCES

1. Han, X., Wu, Z., Jiang, Y., & Davis, L. S. (2018). Viton: An Image-based Virtual Try-on Network. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 7543-7552.

2. Song, X., Feng, F., Ma, X., & Xiang, T. (2019). Attentive Generative Adversarial Network for Virtual Try-on. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1294- 1303.

3. Liu, J., Li, Z., Li, S., Li, J., & Xu, J. (2020). Virtual Try-on with Detail-enhanced Feature Transformation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 1676-1685.

4. Huang, Y., Liu, M., & Belongie, S. (2017). Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization. Proceedings of the IEEE International Conference on Computer Vision, 1510-1519.

5. Li, Z., Chang, S., Shan, L., & Chen, X. (2020). Rethinking Data Augmentation for Image Super Resolution: A Comprehensive Analysis and a New Strategy. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 479- 488.

6. Yu, J., Lin, Z., Yang, J., Shen, X., Lu, X., & Huang, T. (2018). Generative Image Inpainting with Contextual Attention. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 5505-5514.

7. Park, T., Liu, M. Y., Wang, T. C., & Zhu, J. Y. (2019). Semantic Image Synthesis with SpatiallyAdaptive Normalization. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2337-2346.