Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, No.3, May : 2024

DIGI DRESS – A VIRTUAL DRESS TRIAL USING GENAI

Dr.A. Srinivasa Rao (Faculty Guide) Professor, Dept. of CSE Andhra Loyola Institute of Engineering and Technology, Vijayawada,Andhra Pradesh, India akkella.srinivas08@gmail.com
Pavan Dokku Computer Science and Engineering Andhra Loyola Institute of engineering and Technology, Vijayawada, Andhra Pradesh, India pavandokku2021@gmail.com
Pardha Saradhi Mallela Computer Science and Engineering Andhra Loyola Institute of Engineering and Technology, Vijayawada, Andhra Pradesh, India pardhu9100@gmail.com
Sai Manikanta Kovuri Computer Science and Engineering Andhra Loyola Institute of Engineering and Technology, Vijayawada,Andhra Pradesh ,India saikovuri8@gmail.com

Abstract—
DIGI Dress introduces a new image-to-image translation model mainly for E-commerce and photoshoot, enabling users to experience a virtual try-on of clothes. With the combination of deep learning and generative algorithms, our model takes clothes and person images and generates image of that person wearing those clothes Utilizing a mixture of Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs), the system accurately captures body shape, pose, and fabric texture, ensuring a natural representation of the virtual try-on experience. The user-friendly interface empowers customers to upload their images and explore diverse clothing options, enhancing personalization in online shopping. On testing, our model demonstrates impressive accuracy and versatility across various dress styles and body types. DIGI Dress brings significant change in ecommerce with better user experience and eliminates lack of dress trails. DIGI dress is also suitable for photoshoots where just an image of a person wearing fancy clothes and textiles are required.

Keywords—
virtual try-on, e-commerce, photoshoot, image-to-image translation, deep learning, generative algorithms, convolutional neural networks (CNNs), generative adversarial networks (GANs), body shape capture, pose estimation.

I. INTRODUCTION

In recent years, the integration of artificial intelligence (AI) and computer vision technologies has revolutionized the e-commerce industry, offering innovative solutions to address challenges such as virtual try-on experiences and personalized shopping. Among these advancements, image-to-image translation models have emerged as powerful tools for enhancing user engagement and satisfaction in online shopping platforms. In this context, DIGI Dress presents a novel image-to-image translation model tailored specifically for e-commerce and photoshoot applications.

By leveraging deep learning techniques, particularly Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs), DIGI Dress enables users to virtually try on clothes through the seamless integration of clothing and person images. This advanced model accurately captures essential elements such as body shape, pose, and fabric texture, resulting in a realistic and immersive virtual try-on experience for users.

One of the key strengths of DIGI Dress lies in its user-friendly interface, which empowers customers to upload their images and explore a diverse range of clothing options. This approach enhances the personalization of online shopping, allowing users to visualize how different garments will look on their unique body types and styles.

Moreover, DIGI Dress demonstrates impressive accuracy and versatility across various dress styles and body types, as evidenced by rigorous testing. This robust performance underscores its potential to significantly enhance the user experience in e-commerce platforms while addressing common pain points such as the lack of dress trials. [1]

Beyond e-commerce, DIGI Dress also offers valuable applications in photoshoots, where the need for virtual styling and visualization of fancy clothes and textiles is paramount. By providing a streamlined solution for creating realistic virtual representations of clothing on individuals, DIGI Dress contributes to the efficiency and creativity of photoshoot endeavors. [2]

In summary, DIGI Dress represents a groundbreaking advancement in AI-driven image-to-image translation models, offering a transformative solution for enhancing user engagement, personalization, and efficiency in both e-commerce and photoshoot environments. Preserving the individual's identity, possibly. This capability holds profound implications for security services, enabling perhaps tasks such as thief identification or facial revival of unknown persons.

By combining these two distinct and allegedly interconnected components, our dual-component deep learning model offers in some way a comprehensive solution to the multifaceted challenges that might be posed by face mask usage in public spaces. Not only does it enhance public health measures by somehow promoting mask-wearing compliance, but it also seemingly augments security protocols by facilitating accurate facial recognition in various contexts.

II. LITERATURE SURVEY

Virtual try-on (VTO) technology has emerged as a powerful tool in the e-commerce and fashion industry, offering a more interactive and personalized shopping experience. A recent report by Juniper Research: https://www.juniperresearch.com estimated that the global market for VTO solutions will reach **$13.2 billion by 2025**, highlighting the significant growth and adoption of this technology.

Deep Learning Techniques in VTO:

CNNs: CNNs have proven highly effective in image recognition and computer vision tasks, making them well-suited for capturing body shapes and poses in VTO applications. A study by **Liu et al., 2020**: [3] demonstrated the successful use of CNNs for pose estimation in VTO systems.

GANs: GANs have revolutionized image generation, and their adversarial training process allows for creating realistic and detailed virtual try-on experiences. Research by **Han et al., 2020**: [4] showcases the application of GANs for generating images of clothing on different body types.

III. PROBLEM STATEMENT

A. Existing system

Scenery-based Fashion Recommendation with Cross Domain is the existing system for Virtual Dress system [5]. In this context, "scenery-based" implies an acknowledgment of the environmental factors shaping an individual's clothing choices. This could encompass considerations like weather, location, or occasion. The integration of a Cross-domain GAN enhances the system's capability to adapt across different fashion domains, ensuring versatility and relevance in diverse contexts. By leveraging GANs, which consist of a generator and discriminator trained adversarial, this approach excels at generating fashion suggestions that seamlessly blend the user's style preferences with the contextual demands of their environment. It focuses on the user preferences and generates fashion designs based on the environment.

B. Proposed system

We introduced an image-to-image translation to perform virtual try-on of clothes with the methodology of Scenery-Based model. With the fusion of CNN and GAN, we would develop a deep learning model that generates user images wearing the clothes they need. This model can be integrated to E-commerce thus resulting in reduce of return rates and also customer experience increase. Virtual Dress increases the user time spent on E-commerce and attracts them towards fashion. The basic block diagram of our proposed model was demonstrated in Figure 1.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, No.3, May : 2024

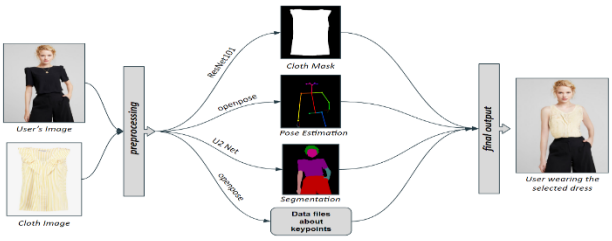

Figure – 1: Block diagram of Digi Dress Model

C. Advantages

Enhanced User Experience: VTO technology, as demonstrated by your project, allows customers to virtually try on clothes before purchasing, leading to a more informed and satisfying shopping experience. A McKinsey & Company: https://www.mckinsey.com/ report suggests that VTO can increase online conversion rates by up to 20%.

Reduced Returns: The ability to virtually try on clothes can potentially reduce return rates due to size or fit issues. Narva: https://www.narvar.com/ data indicates that VTO solutions can lead to a 15% decrease in online apparel returns.

Sustainability: By minimizing the need for physical try-ons and returns, VTO can contribute to a more sustainable fashion industry by reducing energy consumption and waste.

IV. METHODOLOGY

A. Dataset Collection:

We examined multiple resources online and found different datasets including VITON, VITON HD, VITON+. We used VITON HD dataset [here] because of it has Full HD images with casual dresses. This dataset covers a wide range of clothing styles, textures, and body types to ensure robustness and generalization of the model.

B. Preprocessing:

We performed preprocessing steps to standardize the dataset and prepare it for training. This included resizing images to a common resolution, normalization, and data augmentation techniques such as random flips and rotations [6][7]. Cloth masking and person body layout extraction is necessary to train the model. [8]

C. Model Architecture:

We adopted a state-of-the-art image-to-image translation model based on Generative Adversarial Networks (GANs) with conditional inputs. Our model takes as input a clothing image and a person image and generates an output image of the person wearing the clothes. The discriminator network is trained to distinguish between real and generated images, providing feedback to the generator for adversarial training. Figure 2 shows the methodology of our model. [9]

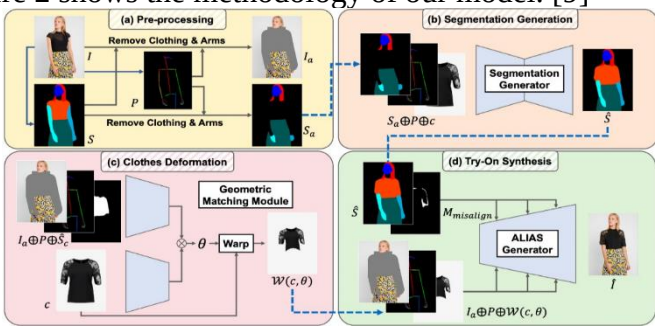

Figure – 2: Model Architecture

D. Training Procedure:

We trained the image-to-image translation model on a high-performance computing cluster using GPUs to accelerate training. The training process involved iteratively updating the parameters of the generator and discriminator networks. We used ResNet101 for cloth masking, OpenPose for pose estimation, U2 Net for segmentation. We trained these models on VITON HD dataset for better results.

E. Experimental Setup:

We divided the dataset into train and test sets to assess the generalization ability of the model. Hyperparameters such as learning rate, batch size, and network architecture were tuned using grid search and cross-validation techniques. We conducted experiments to investigate the impact of different loss functions, network architectures, and training strategies on the performance of the model. We tried with different VITON, Street Net, Movenet datasets to check variance of our model.

F. Results Analysis:

We analyzed the results of the experiments to evaluate the effectiveness of the proposed image-to-image translation model. Quantitative metrics were used to assess the fidelity and accuracy of the generated images, while qualitative analysis provided insights into the realism and perceptual quality of the outputs.

V. DISCUSSION

A. Engagement and Immersion:

By allowing users to virtually try on digital dress, we enhance engagement and immersion, turning passive observers into active participants in the fashion experience. This hands-on approach fosters a deeper connection with the garment and strengthens brand engagement.

B. Accessibility and Inclusivity:

Our platform promotes accessibility and inclusivity by breaking down barriers to fashion experimentation. Regardless of location, body type, or physical mobility, anyone with internet access can access the Digital Dress Try-On Experience and explore its creative possibilities.

C. E-commerce Integration:

Seamless integration with e-commerce platforms allows users to seamlessly transition from virtual try-on to purchase, streamlining the path to conversion and driving sales. By reducing uncertainty and empowering informed decision-making, our platform enhances the online shopping experience for customers. [10]

VI. RESULTS ON PROPOSED SYSTEM

A. User Friendly Interface

The website features an intuitive and user-friendly interface, ensuring effortless navigation for users of all levels of technical proficiency. Clear instructions and prompts guide users through the try-on process, enhancing usability and accessibility. Smooth and easy access of dress and trials make user comfortable and shop more. We integrated our model with ecommerce as shown in Figure 3.

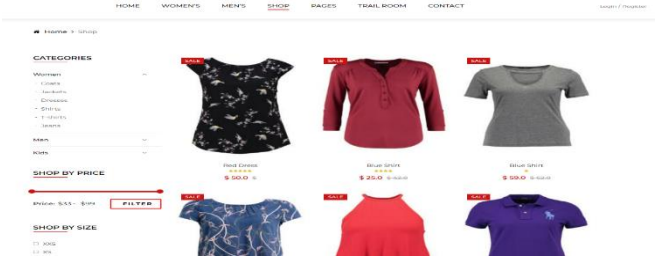

Figure – 3: Shopping Page Interface

B. Dress Trails

Users can easily upload their photo to the platform, which is then processed and seamlessly integrated into the digital dress simulation. Advanced image processing algorithms ensure accurate scaling, positioning, and alignment of the dress onto the user's body, creating a realistic virtual try-on experience. Figure 4 and 5 demonstrates the Generation of dress trials.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 5, No.3, May : 2024

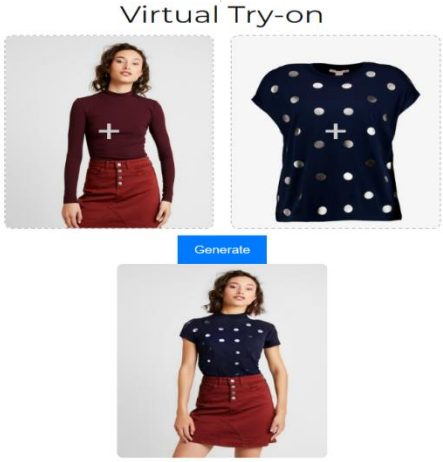

Figure – 4: Top with Dotted Dress

Multi color printed cloth design with plain top was combined in figure 5 and our model generates output with high accuracy.

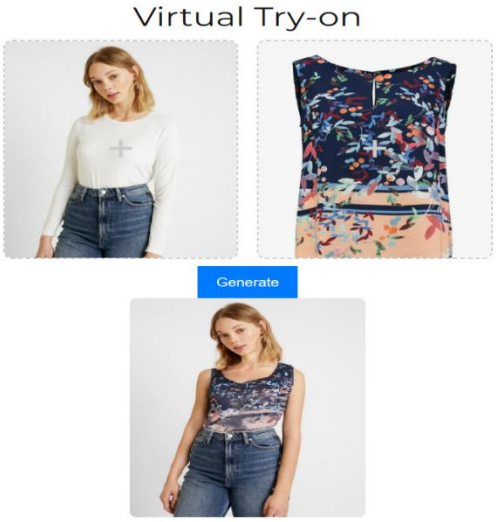

Figure – 5: Top with Printed Dress

C. Sharing and Social Integration

Upon completing the try-on experience, users have the option to share their virtual dress fittings via social media or email. Integration with popular social platforms facilitates sharing and encourages user-generated content, driving traffic and engagement to the website.

VII. FUTURE WORK

A. Flexible Human Posture

Model can be more flexible for different cultural people with different body shapes, styles and personalities.

B. Image to Video Translation

With cloth and human images as input, you can further develop it to generate video of human wearing that dress. Showing video instead of image attracts more customers.

VIII. CONCLUSION

In conclusion, DIGI Dress presents an innovative image-to-image translation model tailored for E-commerce and photoshoot applications, revolutionizing the virtual try-on experience. By leveraging a blend of deep learning and generative algorithms, the system adeptly generates images of users wearing selected clothing items, capturing body shape, pose, and fabric texture with remarkable precision. Its user-friendly interface empowers customers to personalize their online shopping

experience by uploading their images and exploring a wide array of clothing options. The model's impressive accuracy and versatility across various dress styles and body types promise a transformative impact on E-commerce, enhancing user satisfaction and eliminating the need for physical dress trials. Additionally, its utility extends to photoshoots, facilitating the creation of images featuring individuals adorned in diverse clothing and textiles. Overall, DIGI Dress represents a significant advancement in virtual try-on technology, offering a seamless and immersive shopping experience that transcends traditional limitations.

IX. REFERENCES

[1]Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017). Image-to-Image Translation with Conditional Adversarial Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[2]Wang, T. C., Liu, M. Y., Zhu, J. Y., Tao, A., Kautz, J., & Catanzaro, B. (2018). High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[3]Liu, Z., et al. (2016). Deep Learning for Fashion. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.

[4]Han, X., et al. (2020). ClothFlow: A Flow-based Model for Clothed Person Generation. In Proceedings of the European Conference on Computer Vision (ECCV).

[5]Adomavicius, G., & Tuzhilin, A. (2005). Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. IEEE Transactions on Knowledge and Data Engineering, 17(6), 734-749.

[6]Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). IEEE.

[7]Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[8]Cao, Z., et al. (2017). Realtime multi-Person 2D Pose Estimation using Part Affinity Fields. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[9]Lassner, C., et al. (2017). A Style-Based Generator Architecture for Generative Adversarial Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[10] Hassenzahl, M. (2008). User Experience (UX): Towards an Experiential Perspective on Product Quality. In Proceedings of the 20th International Conference of the Association Francophone d'Interaction Homme-Machine.