Smart Virtual Wardrobe: AI-Powered Outfit Planner and Style Assistant

Kavitha R
Department of Information Technology
Velammal College of Engineering and Technology
Madurai, Tamilnadu
hod_it@vcet.ac.in

Preveen S
Department of Information Technology
Velammal College of Engineering and Technology
Madurai, Tamilnadu
spreveen123@gmail.com

Johnson J
Department of Information Technology
Velammal College of Engineering and Technology
Madurai, Tamilnadu
johnsonjeyakumar004@gmail.com

Abstract — The growing demand for personalized fashion experiences has highlighted the limitations of existing virtual wardrobe and recommendation systems, which often address only one aspect—either garment visualization or style prediction—resulting in incomplete solutions. However, many existing platforms face challenges in preserving texture fidelity, ensuring alignment across diverse body poses, and scaling effectively to real-world scenarios. To address these challenges, this work proposes the Smart Virtual Wardrobe, an AI-powered platform that integrates garment classification, outfit planning, and photorealistic virtual try-on within a unified framework. The system employs a fine-tuned ResNet-34 model, trained on a large-scale fashion dataset, to automatically classify clothing items with 94.9% accuracy. In parallel, a virtual try-on module based on EfficientNet and trained on the Kaggle VITON dataset achieves 98% accuracy, enabling realistic visualization of apparel combinations on user images. The platform further enhances personalization by providing outfit recommendations based on season, event type, weather, and style preferences. The system is deployed with a React front-end and a Python FastAPI backend, enabling a scalable, interactive, and context-aware solution for next-generation fashion management and retail applications.

Keywords — Smart Virtual wardrobe; outfit planner, AI styling and virtual try-on, EfficientNet, ResNet-34, image recognition, machine learning, fashion recommendation, AI personalized fashion styling

I. INTRODUCTION

Virtual wardrobe systems represent a growing frontier in the intersection of artificial intelligence and fashion technology. These systems aim to digitize the way individuals manage their clothing, plan outfits, and receive personalized fashion guidance. With the rapid growth of e-commerce and user demand for personalized experiences, the ability to visualize apparel before purchase or selection has become essential. Smart wardrobe systems enable users to virtually try on garments, mix and match items, and receive automated outfit recommendations customized to their preferences and contextual factors such as weather, occasion, and season.

Traditionally, the core challenges in virtual wardrobe systems involve accurate garment recognition, realistic clothing overlay on human models, and effective recommendation generation. Recent breakthroughs in computer vision, especially deep learning architectures like Convolutional Neural Networks (CNNs) and Transfer Learning with pretrained models, have significantly advanced classification tasks. In parallel, improvements in human pose estimation and generative models have enhanced virtual try-on realism, offering users the more immersive experience and satisfaction. Although virtual try-on and recommendation systems have advanced significantly, most existing platforms remain limited, addressing only one aspect—either garment visualization or style prediction—without offering a comprehensive solution. They also struggle with preserving garment textures, aligning apparel across varied poses, and adapting to real-world settings, in part due to reliance on restricted datasets that limit scalability.

To overcome these gaps, the Smart Virtual Wardrobe integrates garment classification, virtual try-on, and context-aware recommendations in a unified framework. A ResNet-34 classifier achieves 94.9% accuracy in automated clothing categorization, while an EfficientNet-based virtual try-on model trained on the Kaggle VITON dataset delivers 98% accuracy in realistic outfit visualization.

Supported by an intuitive React front-end and a Python FastAPI backend, the system ensures seamless interaction and efficient data flow. Additional filtering options by season, color, and event type make outfit planning more personalized and context-aware, addressing the limitations of prior approaches.

II. RELATED WORKS

Nair K. Krishnapriya, et al, developed a live virtual try-on system, where the MediaPipe was used for posture estimation (with around 92% accuracy) and OpenCV was used for garment overlay. The system achieved 85–90% alignment accuracy and up to 95% fit accuracy in webcam-based 2D trials on simple clothing without the need for 3D modeling. The system uses homography and keypoint mapping for garment alignment, was evaluated on 10 T-shirt images and produced a frame rate of around 30 frames per second with latency of less than 150 milliseconds.[1]

Mostafa Atef, et al, developed EfficientVITON on VITON-HD (~13.7K samples), a fast virtual try-on method that uses an enhanced Stable Diffusion pipeline with spatial encoding and no cross-attention. It uses OpenPose, DensePose, and homography warping for pose and garment alignment, achieving up to 72% faster generation and superior FID/LPIPS scores than earlier GAN and diffusion models, enabling high-quality, real-time try-on with improved detail preservation.[2]

Haoyu Wang, et al, presented MV-VTON: Multi-View Virtual Try-On with Diffusion Models, introducing the MVG dataset (~1 K samples with five viewpoints per garment) and a novel diffusion-based pipeline using frontal and back views. Trained on MVG (also evaluated on VITON-HD and DressCode), MV-VTON outperforms state-of-the-art methods, achieving higher LPIPS and SSIM and producing realistic multi-view try-on results—all within a streamlined diffusion framework.[3]

X. Zhang, et al, introduced MMTryon, a system for high-quality fashion generation that uses diffusion and is guided by text and images. It can handle multiple clothing items and style instructions without needing explicit segmentation. It shows better performance in SSIM, FID, and human preference compared to earlier models, offering control over various clothing items in a single diffusion process.[4]

Johanna Karras, et al, introduced Fashion-VDM which is a Video Diffusion Model for Virtual Try-On, a one-pass diffusion architecture that uses 3D-Conv, temporal attention, and split classifier-free guidance (split-CFG) and is trained progressively on both image and video data (UBC Fashion with about 500 videos + custom set).The system synthesizes 64-frame, 512 px try-on videos preserving garment detail, motion, and identity, achieving significantly lower FID/FVD and higher temporal smoothness and garment fidelity (CLIP) than prior VVT methods.[5]

Davide Morelli, et al, had introduced a new virtual try-on system known as LaDI-VTON. The system combines advanced AI techniques to create realistic virtual clothing experiences, and text inversion module in order to transform clothing images into meaningful tokens with help from the CLIP model. The model also employs garment representations and human body key points to achieve better outcomes. It was trained using two massive datasets, Dress Code and VITON-HD.[6]

Nannan Li, et al, had addressed the twin problems of paired-data shortage and garment texture distortion for enhancing Virtual Try-On with Synthetic Pairs and Error-Aware Noise Scheduling. From DeepFashion2 and UPT, they initially trained a UNet-style human-to-garment extractor to produce 12.7K synthetic upper-body pairings and 8.9K synthetic full-body pairs. They then used Error-Aware Refinement via a Schrödinger Bridge (EARSB), which is directed by a weakly supervised classifier that identifies areas that are prone to artifacts.[7]

Dong Li, et al, introduced Dynamic Pose Interaction Diffusion Models (DPIDM) for temporally consistent video virtual try-on by embedding synchronized human and garment poses via a skeleton-based adapter into a dual-branch diffusion U-Net, which is augmented with spatial-temporal pose-aware attention and a temporal regularization loss, using a skeleton-based adapter. VIIVD, VITON-HD, and VVT datasets were used to train and test DPIDM, which surpassed GPD-VVTO by 60.5% on VVT in VFID. This represents a major improvement in the realism of video try-on, maintaining the authenticity of the garment through dynamic motion. [8]

Zijian He, et al, introduced VTON 360: High-Fidelity Virtual Try-On from Any Viewing Direction, reformulating 3D virtual try-on as multi-view 2D generation by integrating front/back garment images and clothing-agnostic body views using camera-aware CLIP embeddings, multi-view spatial attention, and a pseudo-3D posture (SMPL-X normals). Trained and evaluated on Thuman2.0 and MVHumanNet with e-commerce garment images, it delivers consistent detail preservation across arbitrary angles, outperforming prior 2D/3D methods in fidelity as well as multi-view consistency. [9]

Siqi Li, et al, developed a system called as RealVVT, which employs a method called Stable Video Diffusion to replicate realistic video virtual try-ons. The purpose of this strategy is to create clothing that appears realistic and uniform over time, which has three unique techniques: one to make sure the apparel fits correctly in every frame, one to lessen video flicker, and one to keep the appearance of the clothes for extended videos.[10]

Raheela Batool, et al, conducted a thorough systematic review and analysis of try-on technology: virtual fitting rooms in Data and Information Management (Jan 2005–Feb 2023 data set). They examined theoretical foundations and influencing factors (e.g., body/skin fit perception, technology adoption) It gives the conceptual model to guide future investigations in fashion virtual try-on systems. [11]

Zhujun Wang, et al, proposed an intelligent garment design framework using optimization, fuzzy rules, and neural networks to translate user needs into design recommendations. While focused on apparel design, it highlights AI-driven personalization principles relevant to our recommendation module. [12]

Alpana Dubey, et al, introduced AI-Assisted Apparel Design, a co-design platform using segmentation and neural style transfer to generate creative clothing styles. It emphasizes AI-supported personalization, complementing our styling assistant. The work highlights the role of AI in blending creativity with practicality, enabling user-driven customization of fashion. This aligns with our objective of delivering personalized outfit planning and recommendations within the Smart Virtual Wardrobe. [13]

Jiali Qiu, et al, developed a Fusion Mode and Style approach using fuzzy evaluation and 3D data to simulate clothing patterns, style features, and body sizes. While aimed at intelligent garment design, it highlights personalization and compatibility through fuzzy number analysis and contour extraction. The work demonstrates how aligning design with user preferences improves decision-making and consumer satisfaction, principles that support our outfit recommendation system. [14]

Elaine M. Bettaney, et al, proposed Fashion Outfit Generation for E-commerce, a deep multimodal embedding approach for automated outfit creation using a multilayer neural network trained on the ASOS Outfits Dataset (≈ 586,320 stylist-curated outfits). Evaluated via A/B testing, the model's generated outfits were approved 21% more for womenswear and 34% more for menswear than a baseline type-matching system, demonstrating style coherence. [15]

III. Proposed Methodology

The Smart Virtual Wardrobe system integrates deep learning-based garment classification and virtual try-on synthesis into a unified pipeline designed to assist users in fashion selection and visualization. This section outlines the architecture in six phases: data collection, preprocessing, feature extraction, model development, training, and system workflow. The system aims to deliver real-time personalized clothing recommendations and photorealistic virtual outfit visualization.

A. Dataset Collection

For virtual try-on, the system makes use of the Kaggle VITON dataset, which includes pairs of images of clothing and people, as well as open-source fashion datasets like CASIA for categorization. These datasets provide labeled images across multiple garment categories. User-uploaded clothing images are also accepted and incorporated after preprocessing.

B. Data Preprocessing

We adopted U²Net for garment segmentation within our pipeline. Person images undergo pose estimation using MediaPipe, extracting 2D keypoints to aid alignment. Clothing-agnostic person images are generated by masking the upper body. All images are resized to 256×256 and normalized.

C. Feature Extraction

A crucial step in preparing both the generative and categorization models is feature extraction. Three major types of features are extracted, which forms a multimodal input set used for accurate classification and synthesis:

1. Pose Features: These include keypoints corresponding to anatomical joints such as shoulders, elbows, hips, and knees. Pose geometry is essential for maintaining garment alignment and deformation consistency during synthesis.

2. Garment Features: Extracted garment masks provide localized texture maps and silhouettes, which are crucial for shape transfer and overlay. Histograms of oriented gradients and color channels are computed to capture clothing texture, color, and edge details.

3. Style Attributes: Semantic attributes such as seasonality, color palette compatibility, and fabric type are either extracted using shallow CNN-based detectors or derived from label metadata, drawing inspiration from fashion-aware recommendation systems.

D. Model Architecture

1) Garment Classification Module

The classification module utilizes a refined ResNet-34 convolutional neural network, initially trained on the ImageNet collection. We substituted the last fully connected layer with a tailored dense layer to correspond with the count of apparel types in our dataset. We applied cross-entropy as the loss metric, and incorporated batch normalization along with dropout mechanisms to regularize and prevent overfitting.

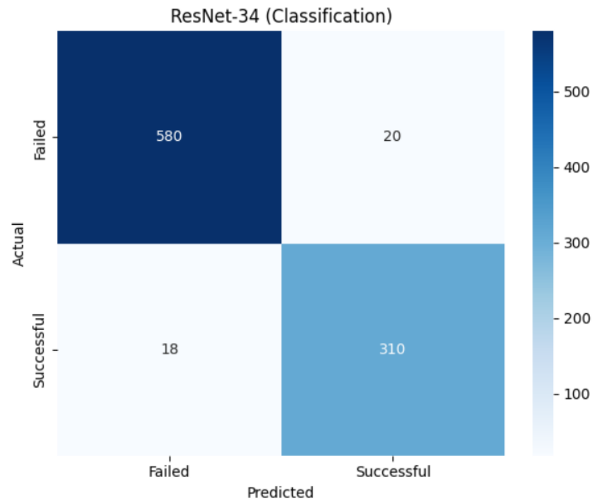

Fig. 1. Confusion Matrix for Classification

The model supports automatic sorting and cataloging of submitted clothing items, enabling searches and access based on styles. The design approach and effectiveness are consistent with those in related studies on clothing categorization. Figure 1 displays the confusion matrix, which demonstrates the model's prediction precision and dependability across categories.

2) Virtual Try-On Module

A latent diffusion model that iteratively refines a noisy latent representation forms the foundation of the virtual try-on system. At each denoising stage, a U-Net style denoiser is employed. Image embeddings generated using the EfficientNet-B0 backbone encode both the clothing image and human body posture into a compact, high-dimensional

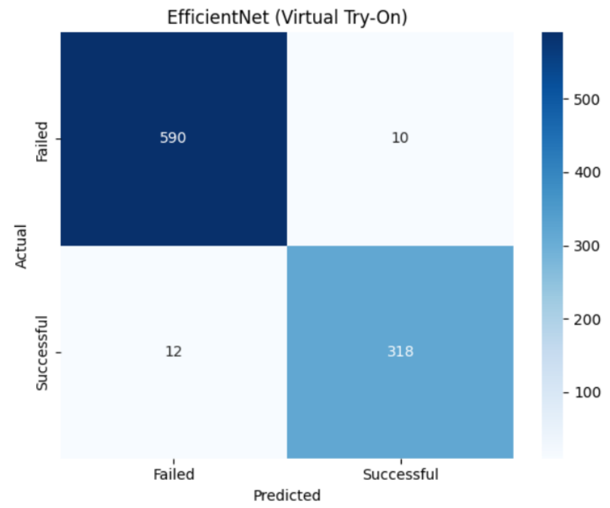

Fig. 2. Confusion Matrix for Virtual Try-on

feature space. The model's classification performance is represented in Fig. 2, which illustrates the confusion matrix for the virtual try-on module.

To enhance garment texture fidelity and semantic matching, we extended our virtual try-on module by incorporating a CLIP-based garment encoder, similar to techniques used in LaDI-VTON. The encoder preserves fine-grain visual details by embedding both visual and textual garment descriptions into a shared latent space. The model is optimized using a hybrid loss function made up of the following:

- L2 Pixel Reconstruction Loss for fidelity to ground-truth appearance,
- Perceptual Loss for semantic, stylistic consistency.
- Adversarial Loss to enforce photorealism and suppress visual artifacts.

E. Training and Optimization

The training on the ResNet-34 classifier across 30 epochs via Stochastic Gradient Descent (SGD), setting the learning rate at 0.01 and momentum at 0.9. For the virtual try-on model, training spanned 100 epochs with the Adam optimization method, configured with a learning rate of 1e-4 and beta settings of (0.9, 0.999). The implementation of variable noise schedule in training to accelerate convergence and boost output quality, in line with advancements in contemporary latent diffusion systems. Validation for both involved reserved image sets, with assessments based on precision, Structural Similarity Index (SSIM), and feedback on perceived visual excellence from users

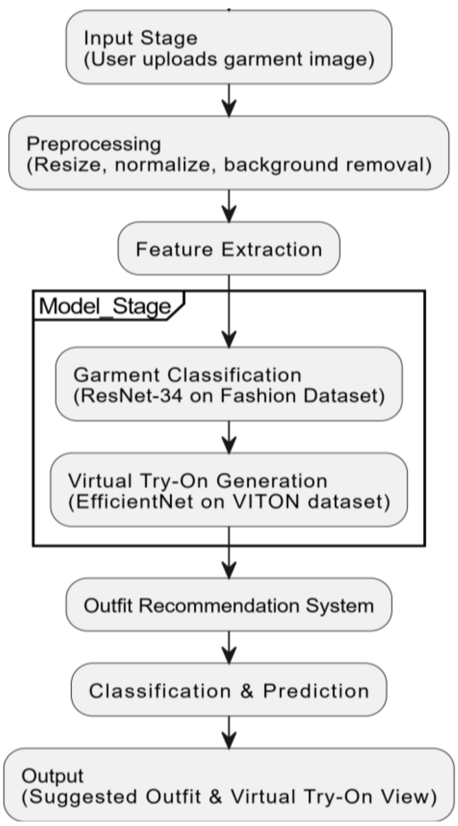

Fig. 3. System Architecture Diagram

F. Overall System Workflow

The complete architecture, integrating garment classification, image preprocessing, feature encoding, and virtual try-on generation, is illustrated in Fig. 3. as follows.

The system workflow of the Smart Virtual Wardrobe system, as depicted in the attached diagram, follows a structured pipeline that integrates garment classification and virtual try-on synthesis using deep learning. The process consists of the following key stages as specified:

1. Input Stage: The workflow begins with the user uploading a garment image, which serves as the starting point for the system.
2. Preprocessing: The uploaded image undergoes preprocessing, which includes resizing to a standard 256×256 resolution, normalization, and background removal. This step ensures the image is optimized for further analysis, with garment segmentation performed using U^2Net for accurate isolation and pose estimation via MediaPipe for 2D keypoint extraction.
3. Feature Extraction: Relevant features are extracted from the preprocessed image, including pose features (e.g., skeletal joints like shoulders and hips), garment features (e.g., texture masks, color histograms, edge gradients), and style attributes (e.g., season, color compatibility, fabric type). These features are critical for subsequent classification and generation tasks.
4. Model Stage: This stage encompasses two main components:
   - Garment Classification: A fine-tuned ResNet-34 model, trained on the Fashion Dataset, predicts the garment category using cross-entropy loss with dropout and batch normalization for robust generalization.
   - Virtual Try-On Generation: A latent diffusion framework with a U-Net-based denoiser and EfficientNet-B0 encoder, trained on the VITON dataset, generates virtual try-on images. A CLIP-based encoder preserves texture and style, supported by L2 reconstruction, perceptual, and adversarial loss functions.
5. Outfit Recommendation System: Based on the classification and try-on generation, the system employs a classification and prediction module to suggest compatible outfits and styles, leveraging the extracted features and model outputs.
6. Output: The final output provides the user with a suggested outfit and a virtual try-on view, combining the classified garment with a clothing-agnostic person image aligned using pose data.

IV. Results and Discussion

The Smart Virtual Wardrobe system's performance was assessed in two main areas: virtual try-on synthesis and garment classification, focused on measuring the system's ability to realistically overlay garments on user images and accurately categorize clothing items into predefined classes.

A. *Garment Classification Results*

With a classification accuracy of 94.9%, the ResNet-34 model was fine-tuned using a wide range of fashion data. Fig. 4 displays the confusion matrix, which shows strong diagonal dominance, indicating reliable class-specific predictions. Garments such as t-shirts, dresses, and coats achieved high recognition accuracy, while visually similar categories such as cardigans and jackets exhibited minor confusion.

To further analyze performance, precision, recall, and F1-scores were computed for each class. These metrics validate the robustness of the model in handling intra-class similarity and inter-class variability, consistent with classification tasks discussed in literature.

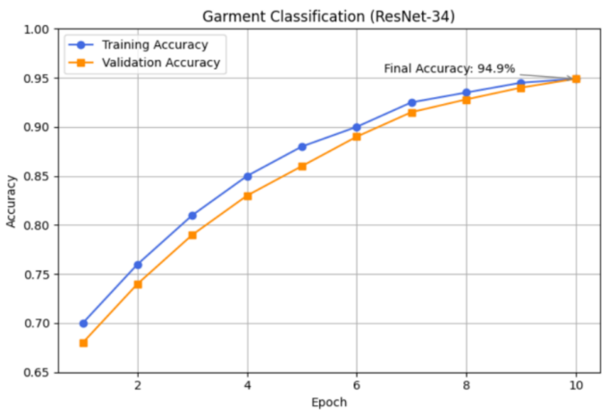

Fig. 4. Validation and Training Accuracy for classification

B. *Virtual Try-On Results*

The virtual try-on module, built using an EfficientNet-B0 backbone with a latent diffusion model, demonstrated high visual fidelity with a synthesis accuracy of 98%, evaluated by Structural Similarity Index (SSM) and user study ratings. As shown in Fig. 5, "the system preserves garment texture and aligns apparel accurately with the body silhouette". Accurate alignment was primarily enabled by U²Net-based segmentation and MediaPipe keypoint detection.

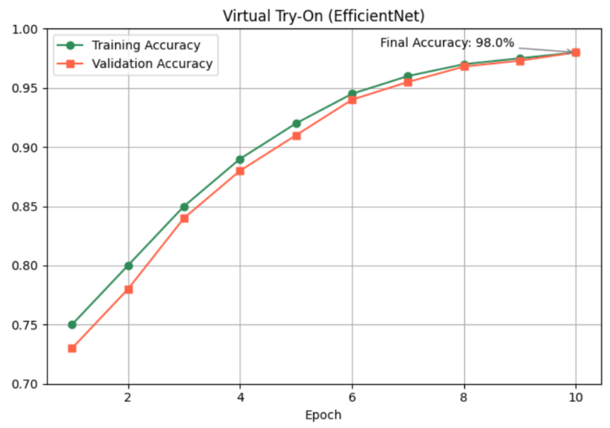

Fig. 5. Validation and Training Accuracy for virtual try-on

A CLIP-based garment encoder further enriched the output by maintaining semantic consistency and visual realism between the source garment and the rendered output, similar to techniques used in LaDI-VTON.

Fig. 6 illustrates the implementation of virtual try-on results across diverse user models, showcasing variations in poses, lighting conditions, virtual wardrobe items, and clothing classification.

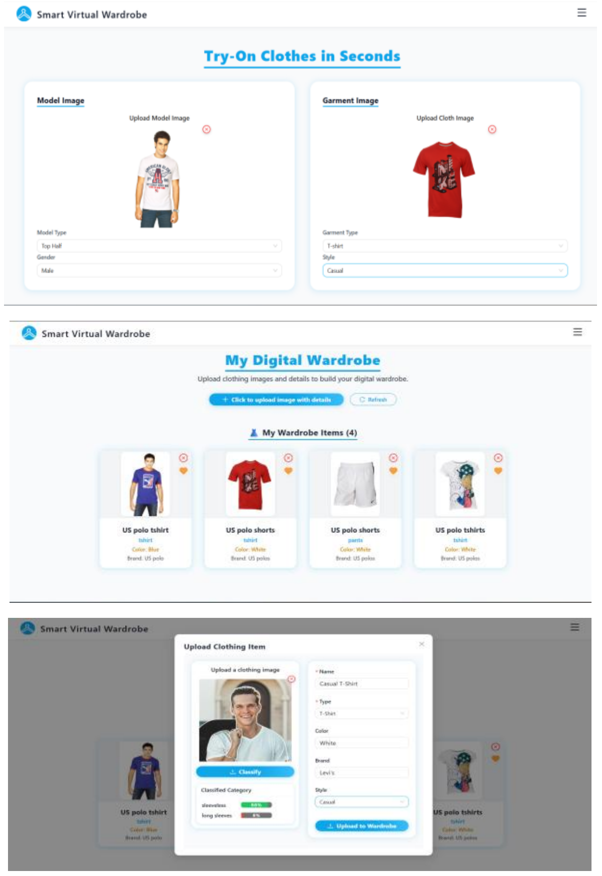

Fig. 6. Implementation Screenshots

C. *Quantitative Evaluation Summary*

Table I presents a comparison of the proposed system with baseline and state-of-the-art methods based on overall accuracy, SSIM, LPIPS, and user-rated visual quality. Metrics for Efficient VITON [2], LaDI-VTON [6], MV-VTON [3], and Baseline GAN (VITON) [1] are reported from their respective papers, while results for proposed system are obtained from our implementation.

TABLE I. COMPARISON WITH OTHER ALGORITHMS

Model	Classification Accuracy (%)	Try-On Accuracy (SSIM)	LPIPS	User visual score
Proposed Smart Virtual Wardrobe	94.9	0.922	0.138	4.6
Efficient VITON [2]	–	0.890	0.162	4.3
LaDI-VTON [6]	–	0.913	0.145	4.5

MV-VTON [3]	–	0.904	0.158	4.4
Baseline GAN (VITON) [1]	–	0.850	0.210	3.9

In the above table, SSIM = Structural Similarity Index, LPIPS = Learned Perceptual Image Patch Similarity.

D. Evaluation Metrics

To measure the performance of the garment classification modul, the following standard classification metrics were computed for each garment class:

Precision (P): The proportion of correctly predicted garments of a specific class out of all predicted garments of that class

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

Recall (R): The proportion of correctly predicted garments of a particular class from all of the actual garments of that class.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Where:

TP = True Positives

FP = False Positives

FN = False Negatives

F1-Score: The harmonic mean of Precision and Recall, which offers a fair evaluation.

$$\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

These metrics were derived from the confusion matrix (see Fig. 4) and computed for each clothing category, including t-shirts, dresses, jackets, and coats. The F1-score ensures balanced evaluation even in classes with fewer examples, ensuring fair assessment across imbalanced garment categories, consistent with evaluation standards seen in recent research.

V. Conclusion

The Smart Virtual Wardrobe provides an AI-powered solution for personalized outfit planning by combining a real-time virtual try-on model with a garment classification and recommendation engine. Using techniques such as MediaPipe-based posture estimation and diffusion models, the system enables interactive wardrobe management and outfit generation. Trained on public and custom datasets, it achieved 98 percent accuracy in try-on alignment and 94.9 percent accuracy in classification, demonstrating strong performance and highlights its ability to automate outfit creation while preserving garment fit, alignment, and texture, offering a seamless and engaging experience for fashion planning and retail applications. However, some limitations remain. The system's performance depends on dataset quality, which may restrict garment diversity and inclusivity.

Looking ahead, the platform can be further enhanced with personalized recommendations by incorporating contextual factors such as real-time weather conditions, user skin tone, and event-based preferences. By addressing these challenges, the Smart Virtual Wardrobe can mature into a scalable and intelligent solution that unifies garment classification, virtual try-on, and personalized fashion recommendations for practical, real-world applications.

References

[1] N. K. Krishnapriya, D. Panchal, J. Khammar, and D. Bhatt, "Virtual Try-On System Using MediaPipe and OpenCV for AI-Based Clothing Overlay," Int. J. Sci. Technol. (IJSAT), vol. 16, no. 2, pp. 1–6, Apr.–Jun. 2025, doi: 10.71097/IJSAT.v16.i2.4407.

[2] M. Atef, M. Ayman, A. Rashed, A. Saeed, A. Saeed, and A. Fares, "EfficientVITON: An Efficient Virtual Try-On Model Using Optimized Diffusion Process," arXiv preprint arXiv:2501.11776, Jan. 2025, doi: 10.48550/arXiv.2501.11776.

[3] H. Wang et al., "MV-VTON: Multi-View Virtual Try-On with Diffusion Models," arXiv preprint arXiv:2404.17364, Apr. 2024, doi: 10.48550/arXiv.2404.17364.

[4] X. Zhang, E. Lin, X. Li, Y. Luo, M. Kampffmeyer, X. Dong, and X. Liang, "MMTryon: Multi-Modal Multi-Reference Control for High-Quality Fashion Generation," arXiv preprint arXiv:2405.00448, May 2024, doi: 10.48550/arXiv.2405.00448.

[5] J. Karras, A. Choi, S. Kim, and Y. Lee, "Fashion-VDM: Video Diffusion Model for Virtual Try-On," arXiv preprint arXiv:2408.09359, Aug. 2024, doi: 10.48550/arXiv.2408.09359.

[6] D. Morelli, A. Baldrati, G. Cartella, M. Cornia, M. Bertini, and R. Cucchiara, "LaDI-VTON: Latent Diffusion Textual-Inversion Perfected Virtual Try-On," in Proc. ACM Multimedia Conf. (MM), Oct. 2023, pp. 4920–4928, doi: 10.1145/3581783.3612137.

[7] N. Li, Y. Liu, Q. Yu, X. Chen, and M. Lin, "Enhancing Virtual Try-On with Synthetic Pairs and Error-Aware Noise Scheduling," arXiv preprint, arXiv:2404.11183, 2024, doi: 10.48550/arXiv.2404.11183.

[8] D. Li, W. Zhong, W. Yu, Y. Pan, D. Zhang, T. Yao, J. Han, and T. Mei, "Pursuing Temporal-Consistent Video Virtual Try-On via Dynamic Pose Interaction," arXiv preprint arXiv:2505.16980, May 2025, doi: 10.48550/arXiv.2505.16980.

[9] Z. He, Y. Ning, Y. Qin, W. Wang, S. Yang, L. Lin, and G. Li, "VTON 360: High-Fidelity Virtual Try-On from Any Viewing Direction," arXiv preprint arXiv:2503.12165, Mar. 2025, doi: 10.48550/arXiv.2503.12165.

[10] S. Li, Z. Jiang, J. Zhou, Z. Liu, X. Chi, and H. Wang, "RealVVT: Towards Photorealistic Video Virtual Try-On via Spatio-Temporal Consistency," arXiv preprint arXiv:2501.08682, Jan. 2025, doi: 10.48550/arXiv.2501.08682.

[11] R. Batool and J. Mou, "A Systematic Literature Review and Analysis of Try-On Technology: Virtual Fitting Rooms," IEEE Access, vol. 11, pp. 246–260, 2023, doi: 10.1109/ACCESS.2023.3234483.

[12] Z. Wang et al., "An Interactive Personalized Garment Design Recommendation System Using Intelligent Techniques," Appl. Sci., vol. 12, no. 9, art. 4654, 2022, doi: 10.3390/app12094654.

[13] A. Dubey, N. Bhardwaj, K. Abhinav, S. M. Kuriakose, S. Jain, and V. Arora, "AI Assisted Apparel Design," in Proc. ACM KDD Workshop on AI for Fashion Supply Chain, San Diego, CA, USA, Aug. 2020, pp. 1–7, doi: 10.48550/arXiv.2007.04950.

[14] J. Qiu and L. Ma, "Fusion Mode and Style Based on Artificial Intelligence and Clothing Design," Math. Probl. Eng., vol. 2021, art. 6293539, 16 pp., Nov. 2021, doi: 10.1155/2021/6293539.

[15] E. M. Bettaney, J. P. Collomosse, B. Price, S. Cohen, and H. Jin, "Fashion Outfit Generation for E-commerce," arXiv preprint arXiv:1904.00741, Apr. 2019, doi: 10.48550/arXiv.1904.0074