International Journal of Engineering Research and Development
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 22, Issue 5 (May 2026), PP 235-239

AI-Powered Fashion Recommendation System Using Computer Vision and Deep Learning

Dr. K N D Malleswararao[1]*, V. Uma Sri[1], U. Harshitha[1], T. Mydhili[1], T. Harika[1], and K. Joseph Samuel[1]
[1]*Department of IT,*
Dhanekula Institute of Engineering and Technology, AP, INDIA

ABSTRACT
In recent years, the widespread presence of social media and online shopping has significantly enhanced the availability of fashion trends and related information. However, in spite of the availability of a number of resources on the internet, users often face difficulties in taking proper fashion-related decisions due to the lack of personalized information based on factors such as body type, skin tone, and occasion. Thus, users may face ambiguity and a lack of confidence in their fashion-related decisions. Keeping the above-mentioned problem in view, the present project aims at developing a Fashion Recommendation System using AI technology. The proposed system is based on the concept of using computer vision and deep learning techniques for developing a personalized fashion recommendation system. The system will make use of various computer vision techniques and machine learning algorithms for developing a personalized fashion recommendation system. The system will make use of various image processing techniques for pre-processing the input images. The facial features will be detected using the Haar Cascade Classifier, while the body will be segmented using the YOLO object detection algorithm. Apart from that, a deep learning algorithm will also be used for predicting various attributes such as body type, skin tone, and fashion sensibility. Grad CAM will also be used for the purpose of visualization. Thus, the system will make personalized fashion recommendations based on the analyses carried out above. The proposed system is expected to provide a smart solution for personalized fashion recommendations with high accuracy.

Date of Submission: 03-05-2026 Date of acceptance: 13-05-2026

I. INTRODUCTION

The contemporary form of fashion is a complex phenomenon that has grown from being an important factor of personal identity & self-expression in the digital era. The explosive growth of social media platforms and online shopping websites made the exchange of fashion trends a global and instantaneous process that can affect millions of users in minutes. While users can find tonnes of information regarding fashion, they still struggle with a suitable outfit that goes well with their body shape, skin tone, personality and occasion. Existing fashion recommendation systems are mostly trend-based or purchase history-based without concerning with the individual traits of each person. For these reasons, the users are confused and less confident which also leads them to wear bad clothes. It creates need for a personalized and intelligent, efficient fashion recommendation system which is also supported by some previous research [1].To break free from these constraints, a cutting-edge system titled "AI-Based Fashion Recommendation System Using Computer Vision and Deep Learning" has been introduced. The suggested system proposes to make a personal fashion recommendation such as body shape, facial structure and skin color by using some of the vital features (user-specific).

The Haar Cascade is used by the system to efficiently detect facial regions and distinguish between primary facial characteristics. Thirdly, YOLO (You Only Look Once) architecture [2], [3] is used to perform the body anlaysis and object detection due to the fact that YOLO has extremely high speed based upon technology including real-time detections. Such structure, positioning and proportion of the body features are key to a successful outfit recommendation and is accurately achieved with YOLO. The derived features are subsequently investigated through a deep learning-based synthesis from which eventual models such as Convolutional Neural Networks (CNNs) and Transformer architectures known for excellent performance in computer vision and image understanding tasks [4], [5], [6] provide the end evaluations.

The Grad-CAM algorithm [8] is integrated into the proposed system to increase prediction transparency and explainability. Grad-CAM is a useful model-agnostic class-discriminative localization technique that identifies the most relevant areas of an image for making prediction, giving users better insight into what the system picks up on from them before deciding how they should be recommended. Depending on the detected

features and deep learning predictions, the system gives appropriate fashion suggestions focusing on clothing styles that suit the user, compatible color combinations, accessories, identifying clothes suitable for different occassions (casual wear, party and formal event).

Importance of Artificial Intelligence (AI) And Computer Vision in Fashion Industry 70+ More AI applications from [11] This section presents the past research works and some more intelligent fashion systems like DeepFashion [9], virtual try on framework VITON[10]. These advances provide the impetus for presenting a solution that leverages an AI-driven mix of deep learning and computer vision to design a smart, user-friendly fashion recommendation system. In short, this system assists the decision-making of a suitable combination in order for the user to dress properly, builds trust while dressing better, and makes online shopping from e-commerce sites with size correct prediction.

Computer Vision, Deep Learning and Explainable AI make the system more robust, interactive and user-friendly. This indicates that presented solution can effectively reshape the future of customized online fashion and smart styling solutions.

II. MATERIAL AND METHODS

2.1 System Architecture

The proposed system will be based on the integration of computer vision, deep learning, and multimodal AI approaches for generating personalized fashion recommendations. It will include several modules related to the processes of image processing, identification of body parts, fashion style analysis, and visualization of outfits. At first, a picture will be uploaded by the user and then processed using typical computer vision methods. Then, this image will be fed into a body parts detection module and YOLO algorithm-based models will be applied for detecting particular body parts [2], [3]. Convolutional neural networks and transformers will be employed to extract more abstract visual features as well as analyze the relationships [4]-[7], [12]. The key principles of fashion recommendation and data sets will lie at the core of fashion recommendations and visualization approaches [1],[9].

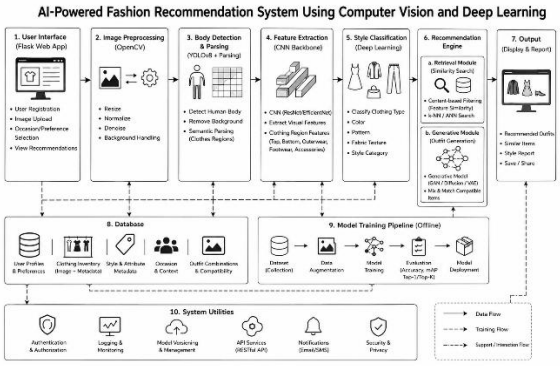

Fig-1.System Architecture

2.2 Input and Processing

The user images can be prepared and enhanced in the suggested module. The input images can also be of different sizes and quality so they need to be inspected and analyzed first before going into any other process. Basic image processing operations like resizing, denoising, and landmarks localization from a face are used to enhance an image. Enhancing an image is essential for deep learning-based computer vision systems [4], [5]. It is processed with YOLO and SSD methods [2], [23] on the enhanced image.

2.3 AI-Based Style Analysis and Recommendation

During this step, the system utilizes high-level style analysis using multimodal AI. EfficientNet and DenseNet, among other deep learning methods, are used for extracting prominent visual information from the images. At the same time, transformers are used to boost contextual awareness in images. Vision-language models are used to understand fashion attributes in a semantic way. Based on the results obtained, the system predicts body types, skin tones, and style categories and generates recommendations for clothes. The recommendations method is based on previous studies on the topic of fashion matching and virtual try-ons. Furthermore, the process utilizes Grad-CAM for explainability and visualization of critical image regions used during prediction

2.4 Visualization and Output Generation

The final step is to make plots for the proposed combinations of dresses. The systems two methods of producing

images ranging from simple to advanced generation techniques such as generative adversarial networks and

diffusion modeling that can be applied in the system [13], [16]. Additionally, they propose a variational auto-encoder to ensure variance and adequate representation in the generated images [17]. We visualize the clothing outfits with color encoding whenever possible using SVG-based visualization, which generates high-quality images. Moreover, they are a perfect supplement to connect text descriptions and actual images.

III. RESULTS AND DISCUSSION

3.1 Style Analysis Phase of the System

The Fig-2 depicts the style analysis phase of the system. In this phase, the image provided by the user is analyzed in order to establish the salient attributes associated with the image of an individual. In this case, the system identifies the body shape of the individual as hourglass, the skin tone as medium, and the style orientation as contemporary chic. Based on the attributes provided in the image, the system predicts the best fashion for the individual, which in this case is casual wear and has a confidence level of 80%, in addition to a number of colors to match the look of the individual. This phase of the system is very important since an image analysis is done in order to come up with fashion attributes used in generating the required outfits.

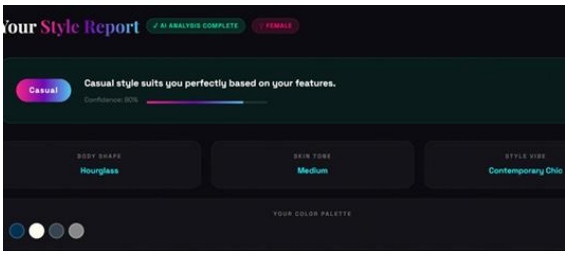

Fig-2.Style Analysis Phase of the System

Attribute	Detected Result	Description
Body Shape	Hourglass	Balanced upper and lower body proportions
Skin Tone	Medium	Moderate warm complexion
Style Orientation	Contemporary Chic	Modern fashion preference
Predicted Fashion Style	Casual Wear	Recommended outfit category
Confidence Score	80%	Prediction confidence
Recommended Colors	Beige, Navy Blue, Olive Green	Suggested color palette

Table 1: Extracted User Fashion Attributes & Prediction Results.

The confidence score for fashion classification is calculated using the Softmax function:

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^{n} e^{z_j}}$$

3.2 AI-Based Outfit Visualization and Analysis Output

In this figure(Fig-3), one can see the key stage of results, in which the process of analysis and visualization of the final outfits occurs at once. The upper part of the screen shows the detection result using YOLOv8 technology with a confidence of 95%, where the proportions of the human body are determined, after which the Grad-CAM analysis is shown with marked areas, highlighting critical factors, such as facial structure, shoulders' width, and overall human body proportions. In the lower part, the system generates images with outfit illustrations created based on artificial intelligence. In total, three different outfits are illustrated using this method. These visuals illustrate an outfit combination that was obtained as a result of the previous analysis, allowing the user to clearly see what clothes would look like in reality.

YOLOv8 object detection performance is evaluated using Intersection over Union:

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union}$$

Grad-CAM computes feature importance using gradients:

$$L^{c}_{Grad-CAM} = ReLU\left(\sum_k \alpha^{c}_{k} A^{k}\right)$$

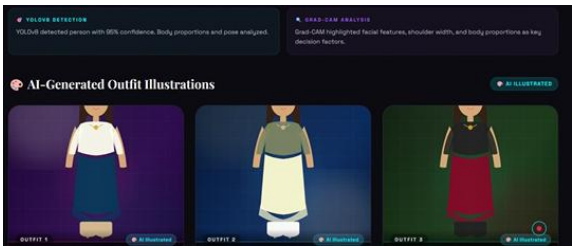

Fig-3.AI Based Outfit Visualization and Analysis Output

3.3 Detailed Outfit Recommendations Output

This diagram (Fig-4) represents the last stage in the system where specific suggestions for outfits are made in an organized manner for the benefit of the user. There are three outfits represented in this system which include the Classic Elegance, Smart Casual, and Evening Ready outfits, each meant for different occasions like workplace, casual meetups, and evenings, respectively. All the constituents of these outfits are mentioned here along with their color combinations so that the user can get a proper idea on how to wear them. The most important point that this section includes is the proper occasion where the outfits should be worn and helpful suggestions like wearing minimum accessories or wearing a blazer when going to work.

Outfit Type	Clothing Components	Recommended Colors	Suitable Occasion	Styling Suggestion
Classic Elegance	Blazer, Shirt, Trousers	Black, White, Grey	Office / Workplace	Use minimal accessories
Smart Casual	T-Shirt, Jeans, Sneakers	Blue, Beige, Olive	Casual Meetups	Add a lightweight jacket
Evening Ready	Dress/Shirt, Formal Shoes	Maroon, Navy Blue	Parties / Evening Events	Use elegant accessories

Table 2. Detailed Outfit Recommendations for Different Occasions

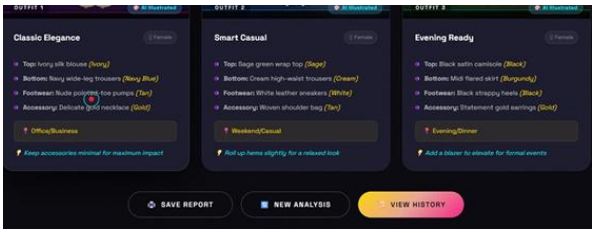

Fig-4.Based Detailed Outfit Recommendations Output

IV. CONCLUSION

The project is a significant advancement in the field of AI-based fashion technology since it shows the potential for the integration of various state-of-the-art methodologies, including computer vision, deep learning, and multimodal analysis, for generating intelligent fashion recommendation systems. An end-to-end system is developed for the integration of various components, including image preprocessing, face/body analysis, intelligent recommendation generation, and visualization. The key contributions of the project are the integration of various AI-based components using a single system, the application of explainability techniques for improved system transparency, and the development of a reliable system for visualization. Furthermore, the project shows the potential for developing intelligent fashion recommendation systems using accessible tools and APIs without the requirement for complex infrastructure. However, the project also highlights the limitations in the system, including the lack of real-time fashion trends, the absence of wardrobe integration, and the application of simple body detection techniques. The system demonstrates the potential for artificial intelligence in developing intelligent fashion recommendation systems for the end-user. Therefore, the project is a significant advancement in the field of AI-based fashion technology. The way people use this thing is better when it gives them ideas for what to wear based on their body shape, skin tone and the style they like. People trust it more. Understand how it makes predictions and suggestions. This is because it uses techniques like Grad-CAM that help explain what it is doing.

REFERENCES

[1] Y. Hu, X. Yi, and L. S. Davis, "Collaborative Fashion Recommendation: A Functional Tensor Factorization Approach," Proc. ACM Multimedia, 2015, pp. 129–138.
[2] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv:1804.02767, 2018.
[3] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," arXiv:2004.10934, 2020.
[4] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv:1409.1556, 2014.
[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Proc. CVPR, 2016, pp. 770–778.
[6] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," Proc. ICML, 2019, pp. 6105–6114.
[7] A. Dosovitskiy et al., "An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale," Proc. ICLR, 2021.
[8] R. R. Selvaraju et al., "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," Proc. ICCV, 2017, pp. 618–626.
[9] Z. Liu et al., "DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations," Proc. CVPR, 2016, pp. 1096–1104.
[10] X. Han et al., "VITON: An Image-Based Virtual Try-On Network," Proc. CVPR, 2018, pp. 7543–7552.
[11] A. Radford et al., "Learning Transferable Visual Models From Natural Language Supervision," arXiv, 2021.
[12] A. Vaswani et al., "Attention Is All You Need," Proc. NeurIPS, 2017, pp. 5998–6008.
[13] J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," Proc. NeurIPS, 2020.
[14] C. Lassner et al., "Unite the People: Closing the Loop Between 3D and 2D Human Representations," Proc. CVPR, 2017.
[15] H. Adam and A. D. Galinsky, "Enclothed Cognition," Journal of Experimental Social Psychology, vol. 48, no. 4, pp. 918–925, 2012.
[16] I. Goodfellow et al., "Generative Adversarial Nets," Proc. NeurIPS, 2014, pp. 2672–2680.
[17] D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes," Proc. ICLR, 2014.
[18] T.-Y. Lin et al., "Microsoft COCO: Common Objects in Context," Proc. ECCV, 2014, pp. 740–755.
[19] M. Everingham et al., "The Pascal Visual Object Classes (VOC) Challenge," IJCV, vol. 88, no. 2, pp. 303–338, 2010.
[20] G. Huang et al., "Densely Connected Convolutional Networks," Proc. CVPR, 2017, pp. 4700–4708.
[21] A. G. Howard et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," arXiv:1704.04861, 2017.
[22] M. Sandler et al., "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Proc. CVPR, 2018, pp. 4510–4520.
[23] W. Liu et al., "SSD: Single Shot MultiBox Detector," Proc. ECCV, 2016, pp. 21–37