STM JOURNALS

E-Commerce for Future & Trends

https://journals.stmjournals.com/ecft/

ISSN: 2454-9347
Volume 13, Issue 1, 2026
January-April
DOI (Journal): 10.37591/ECFT

Review ECFT

E-Commerce Clothing Platform with Virtual Try-On

Pratik Kape[1], Sarthak Joshi[1], Sujit Gawade[1,*], Satyajeet Babar[1]

Abstract

The inability to physically evaluate garments remains a major limitation in online clothing commerce. Customers often depend on static product images and generalized sizing charts, which do not accurately represent individual body proportions. This frequently leads to uncertainty during purchase decisions and increased product return rates. To address this limitation, this research proposes a web-based clothing e-commerce platform integrated with an intelligent virtual try-on mechanism. The system allows users to upload their images and digitally simulate selected garments using pose detection, semantic segmentation, geometric garment transformation, and adversarial image refinement techniques. MediaPipe is used for body landmark extraction, the U-Net architecture performs body region segmentation, and spatial transformation networks align garments with detected body structures. A generative adversarial network enhances the realism of the synthesized output. The platform is developed using the MERN (MongoDB, Express.js, React, and Node.js.) stack to ensure scalability, secure authentication, and efficient API communication. Experimental evaluation confirms that the integration of deep learning-based try-on functionality maintains acceptable response time while improving user personalization. The proposed solution demonstrates the feasibility of combining artificial intelligence with scalable web architecture to enhance digital apparel shopping experiences.

Keywords: Virtual try-on, AI in e-commerce, online clothing store, image processing, MERN stack, OpenCV, TensorFlow, personalized shopping, Razorpay integration, fashion recommendation system, responsive web design, user experience

INTRODUCTION

With the rapid expansion of digital commerce, the fashion retail industry has undergone a significant transformation. Despite this growth, online platforms continue to face challenges in replicating the in-store shopping experience, particularly in enabling customers to visualize how garments will appear on their own bodies. The absence of physical trials often results in uncertainty regarding fitness, texture, and overall appearance, which contributes to dissatisfaction and increased product return rates. Most existing e-commerce platforms rely heavily on static product images, predefined size charts, and limited model representations, which restrict personalization and reduce consumer confidence.

Recent advances in computer vision and deep learning have introduced image-based virtual try-on (VTO) systems as a promising solution to this problem. Early work, such as VITON, demonstrated the feasibility of aligning garments with human body images using convolutional neural networks and geometric matching modules [1]. Subsequent research improved garment deformation and texture preservation through characteristic-preserving networks [2] and GAN-based warping approaches for more realistic synthesis [3]. Enhanced frameworks, such as CP-VTON+, have further addressed clothing shape and texture consistency, producing more visually accurate results [4]. These advancements highlight the growing potential of AI-driven garment visualization systems in digital retail environments.

***Author for Correspondence**
Sujit Gawade
E-mail: sujitgawade03@gmail.com

[1]Student, Department of Computer Engineering, Parvatibai Genba Moze College of Engineering Wagholi, Pune, Maharashtra, India

Received Date: July 22, 2025
Accepted Date: February 25, 2026
Published Date: Marck 31, 2026

Citation: Pratik Kape, Sarthak Joshi, Sujit Gawade, Satyajeet Babar. E-Commerce Clothing Platform with Virtual Try-On. E-Commerce for Future & Trends. 2026; 13(1): 17–24p.

Recent studies have focused on improving fidelity, segmentation accuracy, and pose alignment. Collaborative parsing-based models, such as HP-VTON, have enhanced structural realism by integrating human parsing techniques [5–9]. Deep learning-based virtual dressing systems presented at international conferences have demonstrated improvements in real-time performance and garment mapping precision [10, 11]. Comprehensive surveys have confirmed that VTO technologies are evolving rapidly, integrating generative adversarial networks (GANs), pose transfer models, and feature representation techniques to improve realism and scalability [12]. Additionally, pose transfer and recurrent generation models have enabled more flexible garment adaptation under varying body positions and orientations [13]. Practical implementations of virtual dressing room applications further validate the commercial viability of such systems in consumer electronics and retail domains [14].

Motivated by these technological advancements, this project proposes an AI-powered e-commerce clothing platform integrated with a VTO feature. The system is developed using the MERN stack: MongoDB for database management, Express.js and Node.js for backend services, and React.js for frontend user interaction, ensuring scalability and efficient client–server communication. The platform incorporates image processing and deep learning frameworks, such as OpenCV, TensorFlow, and MediaPipe, to perform body landmark detection, garment alignment, and overlay generation in near real time [15–19].

In addition to VTO functionality, the platform includes intelligent product filtering, personalized recommendation mechanisms, user authentication and account management, secure payment processing, and responsive cross-device compatibility. By combining artificial intelligence techniques with modern web technologies, the proposed system aims to enhance personalization, increase consumer trust, and reduce return rates in online fashion retail. This integration of advanced vision-based garment synthesis with a scalable web architecture represents a progressive step toward immersive and user-centric digital shopping experiences.

PROBLEM STATEMENT

Despite rapid advancements in online retail technology, the fashion e-commerce sector continues to experience high product return rates. One of the primary causes of returns is incorrect size selection and dissatisfaction with the garment appearance after delivery. Customers typically rely on product photographs featuring professional models that do not accurately represent diverse body shapes and proportions. Standardized sizing charts also fail to account for variations in body structure, fabric stretchability, and garment design.

This lack of personalized visualization results in hesitation during purchasing decisions and reduces customer trust in digital platforms. Additionally, high return rates increase retailers' operational costs, including logistics expenses, product handling, and restocking losses. Existing e-commerce systems primarily focus on product catalog management and transaction processing rather than realistic garment visualization.

Although research prototypes, such as VITON [1] and CP-VTON+ [4], have demonstrated promising results in controlled experimental environments, many systems are not fully integrated into scalable commercial platforms. Therefore, a comprehensive solution that combines deep learning-based garment synthesis with a robust full-stack e-commerce infrastructure is needed.

The proposed system addresses this gap by integrating an AI-powered VTO engine within a scalable web-based shopping ecosystem, thereby enabling users to make more informed purchasing decisions.

LITERATURE REVIEW

The development of image-based VTO systems has evolved significantly with the advancements in deep learning and computer vision techniques. Early research in this domain focused on synthesizing realistic garment overlays on static human images while preserving body pose and garment texture. One of the pioneering works, VITON, introduced a coarse-to-fine framework that employed a geometric matching module and a refinement network to align clothing items with a target human body image [1]. This approach demonstrated that convolutional neural networks could effectively model garment deformation and spatial transformation. However, texture distortion and misalignment under complex poses remain challenges.

To address these limitations, characteristic-preserving VTO networks have been proposed to enhance texture retention and fine-detail reconstruction [2]. These methods incorporate improved feature extraction mechanisms and spatial transformation layers to better preserve garment patterns. Similarly, GAN-based architectures, such as FW-GAN, have introduced flow-guided warping mechanisms to improve realism in both image- and video-based try-on applications [3]. These systems enhance visual consistency and reduce unnatural artifacts by leveraging adversarial training; however, they increase computational complexity.

Further improvements were introduced through CP-VTON+, which emphasized clothing shape and texture preservation during image synthesis [4]. This framework utilizes a more refined warping strategy in combination with segmentation masks to achieve higher alignment accuracy. The integration of shape-aware modules significantly reduces garment boundary distortion. However, challenges related to body occlusion and extreme pose variations persist.

High-fidelity try-on systems, such as HP-VTON, have integrated collaborative human parsing techniques to enhance structural realism [5]. By combining semantic segmentation with garment warping, the model improved the alignment between clothing regions and body parts. This approach demonstrated that incorporating human body parsing can substantially improve fitting accuracy and visual plausibility.

Deep learning-based VTO systems presented at international conferences have further explored real-time deployment and consumer-level implementation [10, 11]. These systems have focused on improving inference speed while maintaining visual fidelity, making them more practical for e-commerce integration. Feature representation enhancement and Multiscale learning strategies have been introduced to improve garment adaptation under varying lighting and pose conditions.

Comprehensive surveys of deep learning approaches in VTO systems indicate that GANs, pose transfer networks, and attention-based feature mapping techniques dominate current research trends [12]. These surveys highlight ongoing challenges, including scalability, generalization across diverse body types, and computational efficiency. The integration of pose transfer models and recurrent generation techniques has enabled flexible garment adaptation across multiple body orientations and movements [13, 20–22]. Such models support more dynamic try-on experiences; however, they often require extensive training datasets and high-performance hardware.

Practical implementations of virtual dressing room applications have demonstrated the commercial feasibility of these systems in consumer electronics and retail environments [10, 23–25]. These applications emphasize usability, real-time interaction, and system responsiveness, aligning technical research with real-world deployment requirements [26–28].

Despite significant advancements, existing systems still face challenges related to real-time performance, scalability, and seamless integration with full-stack e-commerce platforms. Most research has primarily focused on algorithmic garment synthesis rather than deployment within a complete shopping ecosystem. Therefore, the proposed AI-powered e-commerce clothing platform extends prior work by integrating deep learning-based garment alignment techniques within a scalable MERN stack architecture. By combining virtual try-on functionality with intelligent product filtering, secure payment integration, and personalized recommendation systems, the present work bridges the gap between experimental VTO research and practical commercial implementation.

SYSTEM ARCHITECTURE

The overall architecture of the proposed AI-powered e-commerce platform with VTO is illustrated in Figure 1. The system follows a modular client–server architecture integrated with a deep learning-based VTO engine. The frontend layer, developed using React.js, provides user interaction modules, including login, signup, product browsing, and image upload functionalities. Communication between the frontend and backend components is handled using REST APIs. The backend layer, implemented using Node.js and Express.js, manages user authentication, product services, order processing, and try-on request handling. MongoDB is used as the primary database for storing user credentials, product details, and transaction data. Secure authentication is maintained using JWT-based authorization mechanisms.

The VTO processing module integrates deep learning techniques inspired by image-based try-on frameworks, such as VITON [1] and CP-VTON+ [4], which utilize garment warping and texture preservation strategies. The pose estimation and segmentation components align with the approaches discussed in recent deep learning-based VTO systems [10, 12]. The architecture incorporates pose estimation and U-Net-based segmentation models to ensure accurate garment alignment and realistic overlay generation, consistent with methods proposed in prior research [2, 5]. GAN-based refinement techniques improve visual realism during final image synthesis [3].The system workflow begins with user image upload, followed by pose detection and garment alignment within the VTO engine. The processed try-on image is then returned to the frontend for visualization. This modular architecture ensures scalability, maintainability, and efficient real-time garment rendering capabilities.

METHODOLOGY

The proposed AI-powered VTO system follows a structured deep learning pipeline that integrates computer vision techniques with scalable web technologies. The methodology begins with user image acquisition through a frontend interface, in which the uploaded image is preprocessed using OpenCV to ensure consistent resolution, normalization, and noise reduction. Image normalization improves detection accuracy and reduces computational variability during model inference. After preprocessing, human pose estimation is performed using MediaPipe based landmark detection models, which extract key skeletal points, such as shoulders, necks, hips, and torso boundaries. Accurate pose detection is essential for garment alignment and is consistent with pose-guided alignment strategies used in earlier image-based virtual try-on frameworks, such as VITON [1] and CP-VTON+ [4].

Following pose estimation, human parsing and segmentation techniques are applied to separate body regions from the background. A U-Net-based segmentation architecture is employed to generate semantic masks for the upper body, lower body, and background regions. Human parsing improves garment placement precision and reduces boundary distortions, similar to collaborative parsing approaches introduced in HP-VTON [5]. The segmented body structure enables spatial mapping between garment images and corresponding body regions.

The next stage involves garment warping and geometric transformations. The selected clothing image is processed using spatial transformation networks that adjust the scale, orientation, and perspective to align with the detected body landmarks. Warping mechanisms inspired by characteristic-preserving networks [2] and GAN-based flow-guided transformation models, such as FW-GAN [3], are used to maintain garment texture and structural consistency. These transformation modules ensure that clothing conforms naturally to body posture while preserving visual details, such as folds and patterns.

After alignment, an image synthesis and refinement network generates the final try-on output. A GAN framework enhances realism by reducing artifacts and smoothing garment boundaries, following the refinement principles described in deep learning-based VTO systems [3, 10]. The refinement stage improves texture blending and shadow consistency, producing a visually coherent composite image. The processed output is then transmitted back to the frontend interface for user visualization.

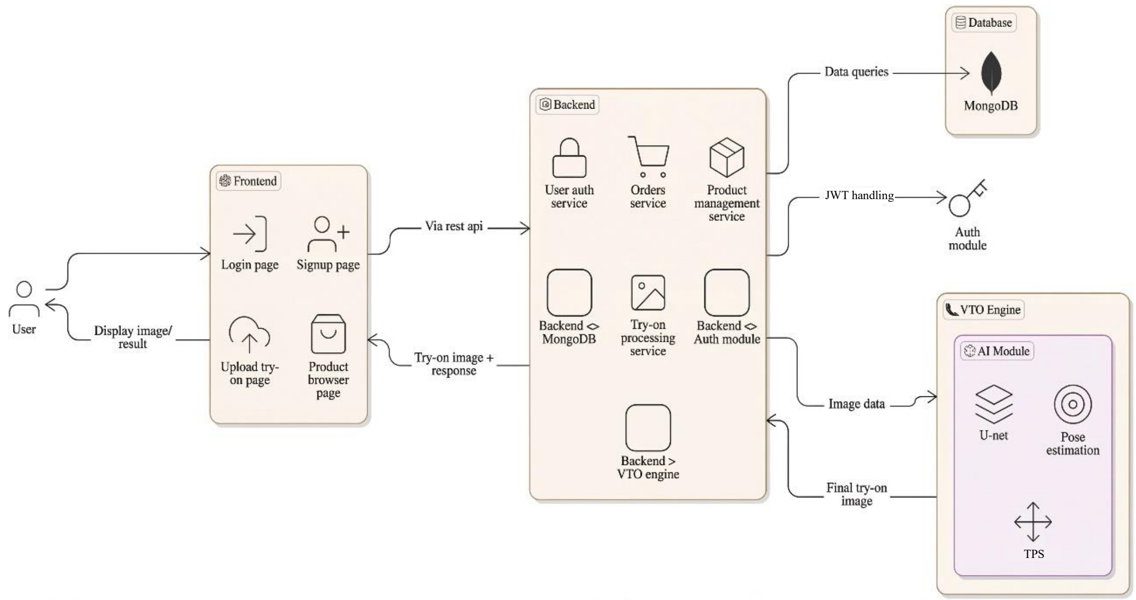

Figure 1. System architecture of the proposed AI-powered E-commerce clothing platform.

In addition to the visual pipeline, the methodology integrates a recommendation subsystem that analyzes user behavior, browsing patterns, and historical interactions. Feature representation strategies similar to those used in enhanced VTO models [10] are incorporated to improve personalization. The entire processing pipeline is deployed within a MERN-based architecture to ensure scalability, secure data transmission, and efficient API communication between the frontend and backend components.

Overall, the methodology combines pose estimation, semantic segmentation, garment warping, adversarial refinement, and personalized recommendation techniques to deliver a realistic and interactive VTO experience. By integrating established deep learning models with full-stack web deployment, the system bridges the gap between algorithmic garment synthesis and practical e-commerce implementation. The proposed system is implemented as a full-stack web application integrated with a deep learning inference engine. The frontend interface is developed using React.js to provide dynamic user interaction, responsive design, and seamless image upload functionality. The backend server is constructed using Node.js and Express.js, which manages REST API communication, authentication services, order processing, and try-on request handling. MongoDB is used as the primary database for storing user credentials, product information, transaction records, and user interaction history.

For the artificial intelligence module, TensorFlow was used to train and deploy segmentation and refinement models. MediaPipe was utilized for efficient real-time pose landmark detection. OpenCV assisted in image preprocessing operations, including resizing, normalization, and background handling. The model was deployed as a server-side service that processes user-uploaded images and returns synthesized outputs through API responses.

The system was tested under moderate traffic conditions to evaluate its scalability and response stability. The experimental deployment demonstrated that the architecture supports concurrent requests without significant performance degradation. This confirms the practical applicability of the proposed system in real-world e-commerce environments.

The virtual try-on-enabled search recorded a comparatively higher response time of 1.9 s with an accuracy of 87%, primarily because of the additional computational overhead associated with pose estimation, garment alignment, and image synthesis processes.

Table 1. Performance evaluation of search features in an e-shop for clothing with VTO.

Test case	Number of queries	Average response time (seconds)	Accuracy (relevant results)
Product search (keywords)	1,200	0.8	93%
Category-based filtering	900	1.1	91%
Price range filter	700	1.4	89%
Virtual try-on enabled search	600	1.9	87%

Similar computational trade-offs have been observed in deep learning-based virtual try-on frameworks, such as those in VITON [1] and CP-VTON+ [4], where garment warping and texture preservation increase processing time. GAN-based refinement and pose alignment modules further contribute to processing complexity, as reported in related studies [3, 10]. Despite the additional computational requirements, the response time remained within practical real-time limits, demonstrating that the proposed system achieves a balance between visual realism and operational efficiency.

Overall, the experimental results confirm that integrating AI-driven virtual try-on functionality does not significantly degrade system performance, aligning with recent research on scalable deep learning-based VTO systems [12]. The platform maintains acceptable latency while delivering enhanced personalization and interactive shopping capabilities.

CONCLUSIONS

This study presents the design and implementation of a scalable e-commerce clothing platform enhanced with a VTO module. The system integrates pose estimation, semantic segmentation, garment transformation, and adversarial refinement to generate realistic garment simulations tailored to user-uploaded images. Unlike conventional online retail systems that depend solely on catalog images, the proposed framework provides interactive visualization aligned with individual body structures.

The MERN-based architecture ensures efficient client–server communication, secure authentication, and responsive multi-device access. Performance analysis indicates that the VTO functionality introduces minimal latency while maintaining satisfactory search accuracy. The results demonstrate that AI-driven garment visualization can be effectively deployed within a commercial shopping ecosystem.

Future improvements may include three-dimensional garment modeling, augmented reality integration, and automated body measurement estimation to further enhance personalization accuracy. The developed system provides a practical foundation for next-generation intelligent fashion e-commerce platforms.

PERFORMANCE EVALUATION

The performance of various search and filtering functionalities implemented in the proposed platform is summarized in Table 1. The evaluation was conducted to measure system efficiency in terms of response time and retrieval accuracy under different query conditions. The keyword-based product search demonstrated the fastest average response time of 0.8 s with an accuracy rate of 93%, indicating effective indexing and optimized database querying mechanisms. Category-based filtering and price range filtering showed slightly increased response times of 1.1 and 1.4 s, respectively, while maintaining acceptable accuracy levels above 89%.

REFERENCES

1. Han X, Wu Z, Wu Z, Yu R, Davis LS. VITON: An image-based virtual try-on network. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA. 2018. p. 7543–7552. doi:10.1109/CVPR.2018.00787.
2. Wang B, Zheng H, Liang X, Chen Y, Lin L, Yang M. Toward characteristic-preserving image-based virtual try-on network. [Preprint]. 2018. arXiv:1807.07688. doi:10.48550/arXiv.1807.07688.

3. Dong H, Liang X, Shen X, Wu B, Chen BC, Yin J. FW-GAN: Flow-navigated warping GAN for video virtual try-on. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South). 2019. p. 1161–1170. doi:10.1109/ICCV.2019.00125.
4. Minar MR, Ahn H. CloTH-VTON: clothing three-dimensional reconstruction for hybrid image-based virtual try-on. In: Ishikawa H, Liu CL, Pajdla T, Shi J, editors. Computer Vision – ACCV 2020. Lecture Notes in Computer Science. Vol. 12627. Cham: Springer; 2021. doi:10.1007/978-3-030-69544-6_10.
5. Hu X, Fang C, Yang K, Liang J, Luo R, Peng T. Toward high fidelity 3D virtual try-on via global collaborative modeling. IEEE Trans Consum Electron. 2024;70(3):5312–5325.
6. Helal WMK, Wang GX, Su YP. Orthogonal simulation test for evaluating fretting wear in low-speed diesel engine connecting rods. Int J Eng Res Afr. 2025;74:101–114. doi:10.4028/pUa0q6f.
7. Ge Y, Zhang R, Wang X, Tang X, Luo P. DeepFashion2: A versatile benchmark for detection, pose estimation, segmentation and re-identification of clothing images. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA. 2019. p. 5332–5340. doi:10.1109/CVPR.2019.00548.
8. Fegade RS, Tated RG, Nehete RS. Numerical analysis to investigate the effect of solidification parameters on the pull-in effect of continuous casting. In: Vučinić D, Chandran V, Mahbub AM, Sobhan CB, editors. Applications of Computation in Mechanical Engineering. Lecture Notes in Mechanical Engineering. Singapore: Springer; 2023. p. 313–25. doi:10.1007/978-981-19-6032-1_25.
9. Deria A, Mahapatra D, Bozorgtabar B, Chakraborty M, Chakraborty S, Roy S. MuGa-VTON: Multi-garment virtual try-on via diffusion transformers with prompt customization. [Preprint]. 2025. arXiv:2508.08488. doi:10.48550/arXiv.2508.08488.
10. Yousef KMA, Mohd BJ, Al-Omari M. Kinect-based virtual try-on system: A case study. 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), Amman, Jordan. 2019. p. 91–96. doi:10.1109/JEEIT.2019.8717498.
11. Ma H, Hu Z, Zheng Y. A virtual try-on model with enhanced feature representation capability. 2022 6th Asian Conference on Artificial Intelligence Technology (ACAIT), Changzhou, China. 2022. p. 1–7. doi:10.1109/ACAIT56212.2022.10137971.
12. Islam T, Miron A, Liu X, Li Y. Deep learning in virtual try-on: A comprehensive survey. IEEE Access. 2024;12:29475–29502. doi:10.1109/ACCESS.2024.3368612.
13. Cui A, McKee D, Lazebnik S. Dressing in order: Recurrent person image generation for pose transfer, virtual try-on and outfit editing. [Preprint]. 2021. arXiv:2104.07021. doi: 10.48550/arXiv.2104.07021.
14. Masri A, Al-Jabi M. Virtual dressing room application. 2019 IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT), Amman, Jordan, 2019. p. 694–698. doi:10.1109/JEEIT.2019.8717410.
15. Rostamzadeh N, Hosseini S, Boquet T, Stokowiec W, Zhang Y, Jauvin C, et al. Fashion-Gen: The generative fashion dataset and challenge. [Preprint]. 2018. arXiv:1806.08317. doi:10.48550/arXiv.1806.08317.
16. Fegade PRS, Tated DRG, Nehete DRS. Improving the aluminium rolling ingot recovery using TQM technique. Int J Innov Technol Explor Eng. 2019;8(12):282–289. doi:10.35940/ijitee.L3779.1081219.
17. Williams NT, Parker JD. Review of resistance spot welding of steel sheets Part 1: Modelling and control of weld nugget formation. Int Mater Rev. 2004;49(2):45–75. doi:10.1179/095066004225010523.
18. Miao Y, Li Z, Tian Y, Hou Q, Wang S, Wang J. Progress in integrated die-casting technology for aluminum alloys. Int J Metalcast. 2025:1–45. doi:10.1007/s40962-025-01825-y.
19. Morelli D, Fincato M, Cornia M, Landi F, Cesari F, Cucchiara R. Dress code: High-resolution multi-category virtual try-on. [Preprint]. 2022. arXiv:2204.08532. doi:10.48550/arXiv.2204.08532.
20. Fegade R, Tated R, Nehete R. Experimental and numerical investigation of effect of melter temperature and casting speed on pull-in (concavity) of AA1050 ingot during direct chill casting. J Inst Eng India Ser C. 2024;105:69–79. doi:10.1007/s40032-023-01018-8.

21. Liu Z, Luo P, Qiu S, Wang X, Tang X. DeepFashion: Powering robust clothes recognition and retrieval with rich annotations. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA. 2016. p. 1096–1104. doi:10.1109/CVPR.2016.124.
22. Kumar A, Bhingole PP, Kumar D. Dynamic analysis of Bajaj Pulsar 150cc connecting rod using ANSYS 14.0. Asian J Eng Appl Technol. 2014;3:19–24. doi:10.51983/ajeat-2014.3.2.727.
23. Xiao H, Rasul K, Vollgraf R. Fashion-MNIST: A novel image dataset for benchmarking machine learning algorithms. arXiv [Preprint]. 2017. arXiv:1708.07747. doi:10.48550/arXiv.1708.07747.
24. Fegade R, Tated R, Nehete R. Numerical investigation of pull-in control of rectangular AA1050 rolling ingot through the design of convex mould during direct chill casting process. Eng Res Express. 2024;6(1):015034. doi:10.1088/2631-8695/ad1a7c.
25. Dong H, Liang X, Shen X, Wang B, Lai H, Zhu J, et al. Towards multi-pose guided virtual try-on network. 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South). 2019. p. 9025–9034. doi:10.1109/ICCV.2019.00912.
26. Yuan WQ, Liu Y, Li D, Meng GY. Research on overloading protection of permanent magnetic coupler in coal mine. Appl Mech Mater. 2014;597:492–497. doi:10.4028/www.scientific.net/AMM.597.492.
27. Ajanalkar S, Birajdar S, Rajput K, Fegade R, Maslekar D, Raul A. Advanced machine learning models for early detection of helical gear faults. Int J Interact Des Manuf. 2025;19(11):7507–7517. doi:10.1007/s12008-025-02301-x.
28. Choi S, Park S, Lee M, Choo J. VITON-HD: High-resolution virtual try-on via misalignment-aware normalization. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA. 2021. p. 14126–14135. doi:10.1109/CVPR46437.2021.01391.

