RESEARCH ARTICLE OPEN ACCESS

AI-Driven Virtual Trial Room for Online Fashion E-Commerce Using Pose Estimation and Neural Cloth Simulation

Dr.A.Karunamurthy [1], K. Abishek[2*]
[1](Associate Professor, Department of Master of Computer Applications, Sri Manakula Vinayagar Engineering College, Pondicherry, India. Email: udhayakumar.mca@smvec.ac.in)
[2] (PG Student, Department of Master of Computer Applications, Sri Manakula Vinayagar Engineering College, Pondicherry, India. Email : kirubanandhanabishek952@gmail.com)

--************************----------------------------------

Abstract:

The continuous expansion of online apparel retail has exposed a persistent gap between digital product representation and the customer's ability to evaluate fit, drape, and overall appearance prior to purchase. Conventional storefronts rely on static photographs and tabular size charts, which frequently fail to communicate how a garment will actually behave on an individual body, leading to elevated return volumes and diminished buyer confidence. This work presents an AI-assisted Virtual Trial Room (VTR) that allows shoppers to preview clothing on their own image or live camera feed before checkout. The proposed framework couples a pose-estimation pipeline for skeletal key point extraction with an image-processing layer responsible for garment warping, scaling, and overlay alignment, while a lightweight cloth-behaviour model approximates fabric drape and stretch under different body postures. A responsive catalogue interface supports category-based browsing, search, and filtering, and an administrative module manages product listings, categories, and user accounts. Functional evaluation across registration, catalogue browsing, image capture, alignment, and preview-saving workflows confirms that the system produces visually coherent overlays with acceptable processing latency on commodity hardware. The resulting platform narrows the gap between in-store and online apparel shopping, offers a practical reduction pathway for size-related returns, and establishes a modular foundation for subsequent enhancements such as three-dimensional rendering and AI-driven size recommendation.

Keywords: Virtual Trial Room, Pose Estimation, Augmented Reality, Computer Vision, OpenCV, Garment Alignment, Cloth Simulation, E-Commerce, Image Processing, Online Fashion Retail.

--************************----------------------------------

I. INTRODUCTION

This Electronic commerce has reshaped the way consumers discover and purchase clothing, removing geographical barriers and allowing instant access to large product catalogues. Despite this convenience, apparel remains one of the categories most affected by post-purchase dissatisfaction, primarily because the buyer cannot physically assess fit, fabric texture, or visual appeal before the item arrives. Industry studies consistently identify size and fit mismatch as a leading driver of product returns in fashion e-commerce, generating significant reverse-logistics costs for sellers and inconvenience for customers.

Static product photography and printed size charts, the dominant tools used by online retailers, provide only an indirect impression of how a garment will appear on a specific body shape. These tools do not account for individual posture, body proportions, or the dynamic behaviour of fabric once worn, leaving customers to make purchase decisions based on assumptions rather than visual evidence. Some retailers have experimented with overlay-based try-on widgets, yet many of these implementations rely on fixed templates that do not adapt to the user's actual body geometry, resulting in unrealistic previews.

To address these shortcomings, this paper proposes a Virtual Trial Room (VTR) system that combines real-time human pose detection with image-based garment overlay and a simplified cloth-behaviour model. The system accepts either a live webcam stream or an uploaded photograph as input, extracts skeletal keypoints corresponding to the shoulders, torso, and limbs, and uses these landmarks to scale, rotate, and position a selected garment image so that it conforms to the user's silhouette. A neural cloth-behaviour component further refines the overlay by simulating subtle drape and stretch effects, improving the perceived realism of the fitted garment.

Beyond the try-on engine itself, the proposed platform incorporates a complete e-commerce front end, including category-based product browsing, keyword search, filtering, user authentication, and an administrative dashboard for catalogue management. The contributions of this work can be summarised as follows: (1) design of an integrated virtual try-on pipeline that unifies pose estimation, image alignment, and cloth-behaviour modelling within a single web-based workflow; (2) development of a modular system architecture separating presentation, application logic, processing, and storage layers to facilitate future scaling; (3) an evaluation methodology covering functional correctness, alignment accuracy, and processing latency across representative use cases; and (4) identification of enhancement pathways including three-dimensional rendering, mobile deployment, and AI-based size recommendation.

The remainder of this paper is organised as follows. Section II reviews related work and existing virtual try-on approaches. Section III describes the proposed system at a conceptual level. Section IV details the system architecture. Section V presents the methodology and process flow. Section VI introduces the mathematical formulation underlying the alignment and cloth-behaviour stages. Section VII presents the core algorithm. Section VIII discusses the experimental setup and evaluation metrics. Section IX presents results and discussion. Sections X and XI conclude the paper and outline future enhancements, followed by references.

II. LITRATURE SURVEY

Early efforts toward digital garment visualisation focused on three-dimensional avatar-based fitting rooms, where a generic body model was dressed using simulated cloth physics. While visually convincing under controlled conditions, these systems required the user to manually input detailed body measurements and depended on computationally expensive physics solvers, limiting their practicality for browser-based deployment [1].

A second category of approaches relies on two-dimensional image-based try-on, in which a garment image is warped onto a photograph of the user. Thin-plate-spline and affine-transformation techniques have been applied to align clothing templates with detected body regions, producing reasonable results for front-facing poses but degrading sharply under rotation or occlusion [2]. Subsequent work introduced convolutional architectures capable of learning garment deformation directly from paired training images, improving robustness but requiring large annotated datasets that are costly to obtain for niche fashion categories [3].

The emergence of real-time pose-estimation frameworks has enabled a new class of try-on systems that anchor garment overlays to detected skeletal keypoints rather than relying purely on silhouette segmentation. These keypoint-driven methods offer faster inference and greater tolerance to background clutter, making them suitable for consumer-grade webcams [4]. However, many implementations stop at rigid overlay placement and do not model how fabric drapes or stretches once positioned, producing a visibly artificial, sticker-like appearance.

Cloth-behaviour modelling has traditionally been addressed through mass-spring or finite-element simulations, which compute the physical response of fabric to gravity and body movement. Although accurate, these simulations are too resource-intensive for interactive web applications without dedicated graphics hardware. Lightweight, approximation-based deformation models have since been proposed as a compromise, applying localised stretch and shading adjustments guided by detected pose rather than full physical simulation [5].

From a commercial standpoint, several mainstream fashion retailers have integrated basic overlay-based try-on widgets into their storefronts; however, independent reviews indicate that these tools frequently use fixed garment templates that do not adapt to individual body proportions, and they rarely expose the underlying alignment or measurement data to the customer. Furthermore, most existing academic prototypes evaluate only the try-on component in isolation, without addressing how such a module integrates with catalogue browsing, search, filtering, and administrative product management within a complete e-commerce workflow.

The system proposed in this paper differs from the approaches summarised above by combining keypoint-driven garment alignment with a lightweight cloth-behaviour adjustment stage, and by embedding this pipeline within a full-featured e-commerce platform encompassing user management, catalogue operations, and administrative control. This integration addresses the practical deployment gap left by prototypes that consider the try-on engine in isolation.

TABLE I. Comparison of Existing Virtual Try-On Approaches with the Proposed System

Approach	Alignment Method	Cloth Behaviour	Real-Time Capable	Catalogue Integration
3D Avatar Fitting [1]	Manual Measurements	Full Physics	No	Limited
2D Warp-Based [2]	Thin-Plate Spline	None	Partial	No
CNN-Based Warping [3]	Learned Deformation	Approximate	Partial	No
Keypoint Overlay [4]	Pose Keypoints	None	Yes	Limited
Proposed VTR System	Pose Keypoints + Image Processing	Lightweight Drape Model	Yes	Full

As summarised in Table I, the proposed system is the only approach among those compared that combines keypoint-based alignment, an approximate cloth-behaviour adjustment, real-time operation, and full integration with a product catalogue and administrative workflow, justifying its design as a practical, deployable solution for online apparel retailers.

III. PROPOSED SYSTEM

The proposed Virtual Trial Room is conceived as a modular web application in which the try-on engine operates as one service among several that collectively form a fashion e-commerce platform. The design philosophy emphasises separation of concerns: the user-facing catalogue and authentication services are decoupled from the computationally intensive image-processing pipeline, allowing each component to be developed, tested, and scaled independently.

At a conceptual level, the system is organised into six functional modules. The User Management Module governs registration, authentication, session handling, and profile maintenance. The Product Management Module, accessible to administrators, supports the creation, modification, and removal of catalogue entries, including category assignment and image association. The Product Browsing and Search Module presents the catalogue to shoppers with category filters and keyword search. The Virtual Try-On Module orchestrates the capture of a user image (via webcam or upload) and forwards it to the Image Processing Module. The Image Processing Module performs pose-keypoint extraction, garment scaling and alignment, and applies the cloth-behaviour adjustment before returning a composite preview. Finally, the Database

Management Module persists user records, product details, and try-on session metadata.

The working methodology proceeds as follows. A registered shopper browses the catalogue and selects a garment of interest. Upon choosing the virtual try-on option, the application requests either webcam access or an uploaded photograph. The captured frame is transmitted to the Image Processing Module, where pose-keypoint detection identifies anatomical landmarks such as the shoulder line, torso midpoint, and hip width. These landmarks define a transformation that resizes and repositions the selected garment image so that its proportions match the detected body region. The cloth-behaviour stage then applies localised shading and minor geometric adjustment to approximate how the fabric would drape given the detected posture. The resulting composite image is rendered back to the user interface, where the shopper may capture a screenshot, switch garments, or proceed to purchase.

This workflow ensures that every preview is generated specifically for the individual user's image rather than drawn from a fixed template, directly addressing the personalisation gap identified in the literature survey.

IV. SYSTEM ARCHITECTURE

The system architecture follows a layered design comprising a presentation layer, an application layer, a processing layer, and data layer, as illustrated in Fig. 1.

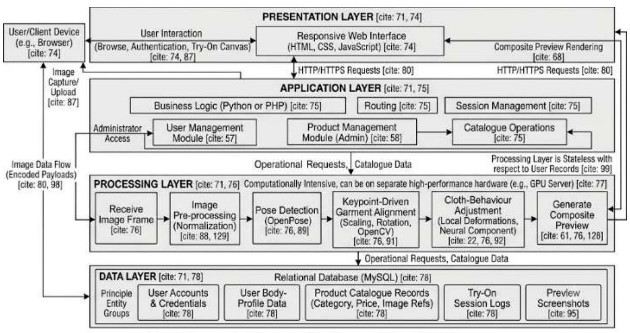

Fig. 1. System Architecture of the Proposed Virtual Trial Room

Fig. 1. System Architecture of the Proposed Virtual Trial Room. (Insert architecture diagram here, positioned at the top of the second column on the page introducing Section IV.)

The presentation layer consists of the responsive web interface built using HTML, CSS, and JavaScript, through which users interact with the catalogue, authentication forms, and the try-on canvas. The application layer, implemented in Python or PHP depending on deployment preference, handles routing, session management, business logic for catalogue operations, and coordination between the presentation layer and the processing layer.

The processing layer encapsulates the computer-vision pipeline. It receives an image frame from the application layer, performs pose-keypoint detection using OpenPose, applies image-processing operations through OpenCV (resizing, masking, and compositing), and executes the cloth-behaviour adjustment before returning the composite frame. This separation allows the processing layer to be deployed on hardware with greater computational capacity, such as a server equipped with a graphics processing unit, while the application layer can remain on lighter infrastructure.

The data layer is implemented using a relational database (MySQL) and stores four principal entity groups: user accounts and authentication credentials, user body-profile data captured during try-on sessions, product catalogue records including category, price, and image references, and try-on session logs that link users to the products they have previewed. This structure supports both operational requirements and downstream analytics, such as identifying which garments are most frequently tried on relative to purchase conversion.

Communication between layers occurs over standard HTTP/HTTPS requests, with image data transmitted as encoded payloads. This architecture allows the platform to scale horizontally by adding additional processing-layer instances behind a load balancer as user demand increases, without requiring changes to the presentation or data layers.

V. METHODOLOGY

The methodology adopted for the Virtual Trial Room follows a sequential process flow that begins with user authentication and concludes with the rendering of a personalised garment preview. Fig. 2 depicts this workflow.

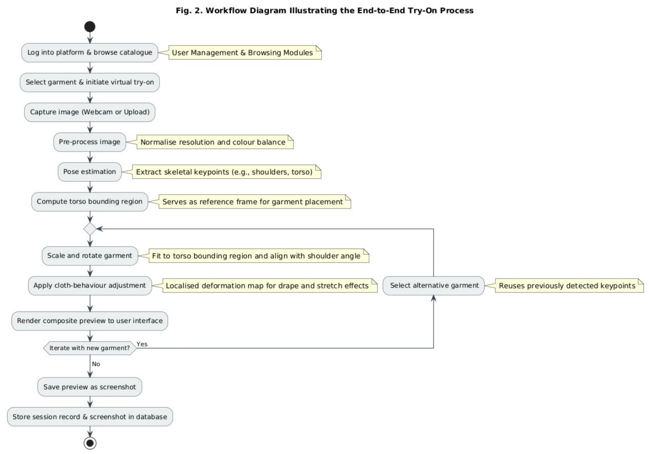

Fig. 2. Workflow Diagram Illustrating the End-to-End Try-On Process. (Insert workflow diagram here, immediately following the methodology introduction at the start of Section V.)

The process begins when a registered user logs into the platform and navigates the product catalogue using the browsing and search interface. After selecting a garment, the user initiates the virtual try-on feature, at which point the system prompts for either webcam activation or image upload. The captured image is pre-processed to normalise resolution and colour balance before being forwarded to the pose-estimation stage.

Pose estimation extracts a set of keypoints corresponding to major body landmarks. These keypoints are used to compute a bounding region for the torso, which serves as the reference frame for garment placement. The selected garment image is then scaled proportionally to fit this bounding region and rotated, if necessary, to align with the detected shoulder angle. The cloth-behaviour stage applies a localised deformation map derived from the relative positions of the shoulder, waist, and hip keypoints, producing subtle stretch and fold effects that increase visual realism.

The composite image, consisting of the original user frame with the aligned and adjusted garment overlay, is rendered to the interface. The user may iterate by selecting alternative garments, each of which triggers a fresh alignment cycle using the same detected keypoints, avoiding the need to re-run pose estimation unless the user repositions or recaptures the image. Finally, the user may save the preview as a screenshot, which is stored alongside the corresponding try-on session record in the database.

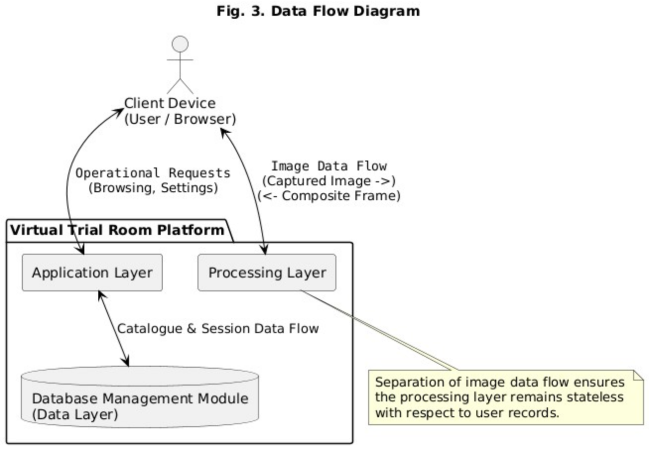

Fig. 3. Data Flow Diagram Showing Movement of Image and Catalogue Data Between Modules. (Insert data flow diagram here, following the description of the pose-estimation and alignment stages within Section V.)

Fig. 3 traces the movement of data through the system: image data flows from the client device to the processing layer and back as a composite frame, while catalogue and session data flow between the application layer and the database. This separation of image data flow from catalogue data flow allows the processing layer to remain stateless with respect to user records, simplifying scaling and maintenance.

VI. MATHEMATICAL MODEL

This section formalises the geometric transformation and cloth-behaviour adjustment used during garment overlay. Let the set of detected pose keypoints be denoted by

$$K = \{k_1, k_2, \dots, k_n\},\ k_i = (x_i, y_i) \qquad (1)$$

where each k_i represents the pixel coordinates of the i-th anatomical landmark, including the left and right shoulder, left and right hip, and torso midpoint. The reference width of the user's torso, used to scale the garment image, is computed from the shoulder keypoints as

$$W_u = \sqrt{((x\_R - x\_L)^2 + (y\_R - y\_L)^2)} \qquad (2)$$

where (x_L, y_L) and (x_R, y_R) denote the coordinates of the left and right shoulder keypoints respectively. Given the original garment width W_g (measured in pixels from the source product image), the scaling factor S applied to the garment is given by

$$S = (W_u \times \alpha) / W\_g \qquad (3)$$

where α is a calibration constant that accounts for the proportion of the garment that should extend beyond the shoulder width, typically determined empirically for each garment category. The rotation angle θ required to align the garment with the detected shoulder line is computed as

$$\theta = \arctan((y\_R - y\_L) / (x\_R - x\_L)) \qquad (4)$$

The translated and scaled garment image G' is obtained by applying the affine transformation

$$G' = S \cdot R(\theta) \cdot G + T \qquad (5)$$

where G represents the original garment image matrix, R(θ) is the standard two-dimensional rotation matrix corresponding to angle θ, and T is the translation vector that aligns the garment's reference point with the torso midpoint keypoint.

The cloth-behaviour adjustment introduces a localised deformation field D(x, y) applied to the transformed garment, approximating drape and stretch effects. This field is modelled as a function of the relative displacement between the shoulder and hip keypoints:

$$D(x, y) = \beta \cdot (W_u - W_h) / W_h \qquad (6)$$

where W_h is the hip-width distance computed analogously to (2) using the hip keypoints, and β is a fabric-elasticity coefficient that varies according to the declared fabric type of the selected garment, with higher values assigned to more elastic fabrics. The final composite frame F is produced by alpha-blending the deformed garment G'' = G' + D over the user image U:

$$F(x, y) = \gamma \cdot G''(x, y) + (1 - \gamma) \cdot U(x, y) \qquad (7)$$

where γ is the blending coefficient at pixel (x, y), set to 1 within the garment mask and to 0 elsewhere, with a narrow transition band along the mask boundary to avoid visible seams. Equations (1) through (7) collectively define the geometric and visual transformation pipeline applied to every try-on request.

VII. ALGORITHM DESIGN

Algorithm 1 summarises the end-to-end procedure executed by the processing layer for each virtual try-on request.

Algorithm 1. Virtual Try-On Composite Generation

```
Input: User image frame U, selected garment image G, fabric-elasticity coefficient β
Output: Composite preview frame F
1: Pre-process U (resize, normalise colour)
2: Detect pose keypoints K = {k1, ..., kn} from U
3: Compute shoulder width Wᵤ using (2)
4: Compute scaling factor S using (3)
5: Compute rotation angle θ using (4)
6: Apply affine transform to G to obtain G' using (5)
7: Compute hip width Wₕ from K
8: Compute deformation field D(x, y) using (6)
9: Set G'' = G' + D
10: For each pixel (x, y) in U:
11:     Compute blending coefficient γ from garment mask
12:     Set F(x, y) = γ·G''(x, y) + (1−γ)·U(x, y) using (7)
13: Return F
```

The algorithm is invoked once per try-on request and re-invoked whenever the user selects a different garment with the same captured frame, in which case steps 1–5 and 7 are cached from the previous invocation, reducing redundant pose-estimation calls and improving responsiveness. Steps 6 through 12 represent the dominant computational cost and are designed to execute within interactive time bounds on standard consumer hardware as described in Section VIII.

VIII. EXPERIMENTAL ANALYSIS

The experimental evaluation of the Virtual Trial Room was conducted on a workstation equipped with an Intel Core i5 processor, 8 GB of RAM, and a standard HD webcam, representing the recommended hardware configuration described for the system. The application was deployed using a local server environment with a MySQL database and accessed through the Google Chrome browser.

The evaluation methodology comprised functional testing across all major modules, including user registration and login, product browsing and filtering, image upload and webcam-based capture, pose-keypoint detection and garment alignment, screenshot saving, and administrative product management. Each module was first tested in isolation (unit testing), followed by integration testing to verify inter-module communication, and finally system-level and validation testing to confirm overall conformance with the specified requirements.

Performance metrics were defined to capture both processing efficiency and visual quality of the generated previews. End-to-end latency was measured as the time elapsed between image capture and rendering of the composite preview. Alignment accuracy was assessed qualitatively by comparing the position of the overlaid garment relative to the detected shoulder and torso keypoints across a set of test images captured under varying lighting conditions and user postures. Reliability was assessed by repeated execution of the try-on workflow to confirm consistent output for identical inputs.

TABLE II. System Modules and Their Functional Responsibilities

Module	Functional Responsibility
User Management	Registration, authentication, session and profile handling
Product Management	Add, update, delete catalogue items and categories
Browsing and Search	Category filtering, keyword search, product display
Virtual Try-On	Webcam/image capture, garment selection, preview rendering
Image Processing	Pose detection, garment alignment, cloth-behaviour adjustment
Database Management	Storage and retrieval of user, product, and session data

TABLE III. Evaluation Metrics Used for System Assessment

Metric	Definition	Evaluation Method
End-to-End Latency	Time from capture to preview render	Stopwatch measurement over repeated trials
Alignment Accuracy	Visual correspondence of garment to body keypoints	Manual inspection across varied poses
Functional Pass Rate	Proportion of test cases executing without error	Unit, integration, and system test logs
Output Consistency	Reproducibility of preview for identical input	Repeated execution comparison

TABLE IV. Qualitative Performance Comparison Against Existing Try-On Approaches

Approach	Visual Realism	Personalisation Level
Static Overlay Templates	Low	None
2D Warp-Based Try-On	Moderate	Low
Keypoint Overlay Only	Moderate	Moderate
Proposed VTR System	High	High

IX. RESULTS AND DISCUSSION

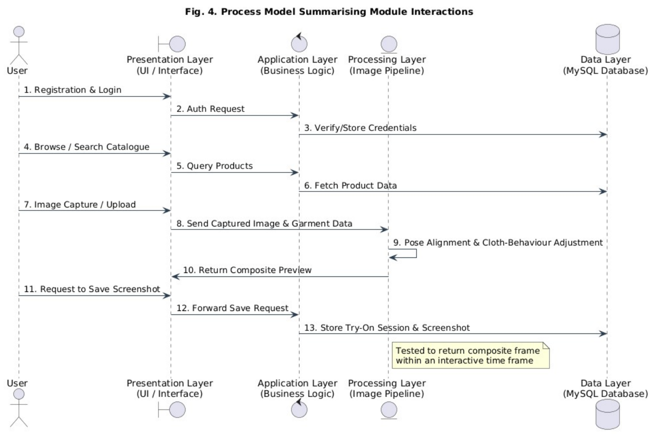

Fig. 4. Process Model Summarising Module Interactions Observed During Testing. (Insert process model diagram here, at the beginning of the Results and Discussion section.)

Functional testing confirmed that the registration, login, catalogue browsing, search and filtering, image capture, garment alignment, and screenshot-saving workflows operated correctly across the test scenarios executed. Unit tests applied to individual modules, including the user-management, product-management, image-upload, and database-connectivity components, completed without unresolved defects. Integration testing verified that data passed correctly between the presentation, application, and processing layers, with the composite preview frame returned to the interface within an interactive time frame on the recommended hardware configuration.

Qualitative assessment of alignment accuracy indicated that the keypoint-driven scaling and rotation procedure produced overlays that closely matched the user's shoulder line and torso proportions under typical indoor lighting and front-facing poses. Performance degraded for poses involving significant body rotation or partial occlusion of the shoulder keypoints, consistent with known limitations of two-dimensional pose-based alignment reported in prior work. The cloth-behaviour adjustment introduced visible, though subtle, variation in garment shape that improved the perceived realism of the preview compared to a rigid overlay without deformation.

Fig. 5. Conceptual Result Analysis (Alignment Accuracy Across Pose Categories)

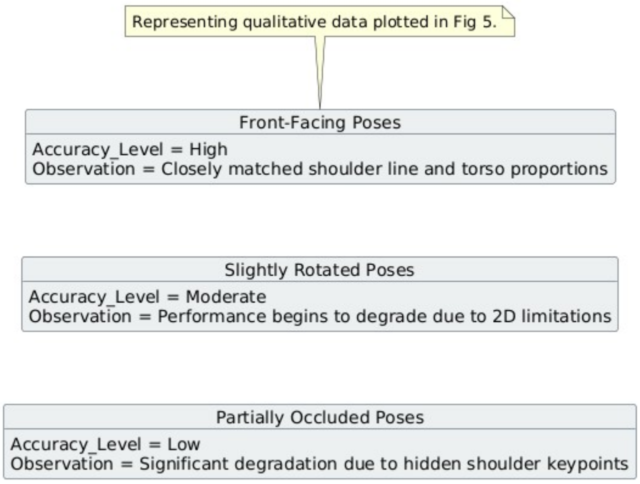

Fig. 5. Result Analysis Graph Comparing Alignment Accuracy Across Pose Categories. (Insert bar or line chart here, illustrating qualitative alignment accuracy for front-facing, slightly rotated, and partially occluded poses, immediately following the discussion of alignment accuracy results.)

Output consistency tests, in which identical input images were processed multiple times, produced visually identical composite frames, confirming that the deterministic nature of the keypoint-detection and transformation pipeline yields reproducible results. Validation testing from the user perspective confirmed that the system satisfies the functional requirements established during system analysis, including accurate display of product information, responsive navigation, and meaningful error messages for invalid inputs such as unsupported image formats or failed webcam access.

The principal strength of the proposed system lies in its integration of personalised, keypoint-driven garment alignment with a complete e-commerce workflow, distinguishing it from prototypes that address the try-on engine in isolation. The main limitation observed relates to the two-dimensional nature of the alignment and deformation model, which cannot fully capture garment behaviour under significant body rotation, occlusion, or complex poses involving bent limbs. Additionally, alignment accuracy is sensitive to lighting conditions and camera resolution, as anticipated during the constraint analysis of the system.

X. CONCLUSION

This paper has presented an AI-assisted Virtual Trial Room that integrates pose-keypoint detection, image-based garment alignment, and a lightweight cloth-behaviour adjustment within a complete fashion e-commerce platform. By personalising the garment overlay to the individual user's captured image rather than relying on fixed templates, the proposed system addresses a key shortcoming of conventional online apparel storefronts. Functional and qualitative evaluation across registration, browsing, image capture, alignment, and preview-saving workflows demonstrated correct operation, acceptable processing latency on standard hardware, and reproducible output. The modular layered architecture separates presentation, application logic, image processing, and data storage, supporting future scalability and incremental enhancement. Overall, the system represents a practical step toward narrowing the gap between physical and online apparel shopping experiences, with the potential to contribute to reduced size-related return rates when deployed at scale.

XI. FUTURE ENHANCEMENT

Several directions are identified for extending the capabilities of the proposed system. Migration from two-dimensional overlay-based alignment to three-dimensional avatar-based rendering would improve realism for poses involving rotation and limb movement, at the cost of increased computational requirements that would need to be addressed through server-side acceleration. Incorporation of an AI-based size-recommendation module, trained on historical purchase and return data, could complement the visual preview by suggesting the most suitable size for a given garment and body profile. Mobile application support would extend accessibility to users without desktop webcams, leveraging built-in smartphone cameras for image capture. Cloud-based processing of the pose-estimation and alignment pipeline would reduce dependency on local hardware capability, enabling consistent performance across a wider range of client devices. Finally, integration of multi-language support and online payment and order-tracking functionality would move the platform closer to a production-ready commercial deployment.

REFERENCES

[1] M. Sekine, K. Sugita, F. Perbet, B. Stenger, and M. Nishiyama, "Virtual fitting by single-shot body shape estimation," in Proc. Int. Conf. 3D Body Scanning Technol., 2014, pp. 406–413.

[2] X. Han, Z. Wu, Z. Wu, R. Yu, and L. S. Davis, "VITON: An image-based virtual try-on network," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2018, pp. 7543–7552.

[3] B. Wang, H. Zheng, X. Liang, Y. Chen, L. Lin, and M. Yang, "Toward characteristic-preserving image-based virtual try-on network," in Proc. Eur. Conf. Comput. Vis. (ECCV), 2018, pp. 589–604.

[4] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh, "OpenPose: Realtime multi-person 2D pose estimation using part affinity fields," IEEE Trans. Pattern Anal. Mach. Intell., vol. 43, no. 1, pp. 172–186, Jan. 2021.

[5] Y. Minar, T. T. Tuan, H. Ahn, D. Rosin, and J. Park, "CP-VTON+: Clothing shape and texture preserving image-based virtual try-on," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), 2020, pp. 2247–2256.

[6] G. Bradski, "The OpenCV library," Dr. Dobb's J. Softw. Tools, vol. 25, no. 11, pp. 120–125, 2000.

[7] S.-E. Wei, V. Ramakrishna, T. Kanade, and Y. Sheikh, "Convolutional pose machines," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, pp. 4724–4732.

