



# Full-Outlet Virtual Try-On (VTO), Deep Learning for Garment Synthesis, Privacy-Preserving Architectures, Biometric Data and GDPR, Digital Fashion

Muhammad Zubair<sup>1</sup>

## ABSTRACT

Full-outfit virtual try-on (VTO) systems aim to digitally render both upper and lower garments on images of real users while preserving body shape, pose consistency, garment structure, and visual realism. Although significant progress has been achieved in single-garment virtual try-on through GAN-based, warping-based, and diffusion-based architectures, extending these approaches to full-outfit synthesis introduces additional challenges related to multi-garment alignment, occlusion handling, body-garment interaction, and structural continuity across body regions. At the same time, the increasing reliance on real user images raises important legal and privacy considerations, particularly under data protection frameworks such as the General Data Protection Regulation (GDPR) and UK GDPR. This study presents a targeted literature review of representative virtual try-on systems, focusing on architectural design, technical limitations, evaluation practices, and privacy-related considerations. The analysis synthesizes key architectural paradigms, including warping-based networks, flow-based deformation models, diffusion-based generation, and hybrid 2D-3D pipelines, along with their input requirements and reported performance characteristics. The findings highlight recurring technical challenges, such as multi-garment misalignment, occlusion inconsistencies, texture distortion, and identity drift, as well as limitations in current evaluation practices, which often rely on metrics that do not fully capture full-outfit realism or structural coherence. Beyond technical analysis, this study examines the treatment of user data within existing systems, including issues related to data retention, biometric representation, and potential re-identification risk. The review suggests that explicit integration of privacy-by-design principles within VTO architectures remains limited and inconsistently reported. Building on these findings, the primary contribution of this work is a conceptual privacy-preserving architecture for full-outfit virtual try-on systems. The proposed framework separates image acquisition from garment simulation through avatar-based processing, enabling reduced reliance on raw user images and supporting data minimisation strategies. Overall, this study provides a structured synthesis of current approaches and outlines directions for developing technically robust and privacy-aware virtual try-on systems suitable for real-world deployment.

1. University of Hertfordshire,  
United Kingdom

**Keywords:** Full-Outlet Virtual Try-On (VTO), Deep Learning for Garment Synthesis, Privacy-Preserving Architectures, Biometric Data and GDPR, Digital Fashion

## INTRODUCTION

Virtual try-on (VTO) systems aim to digitally render garments on images of real people while preserving body shape, pose, garment structure, and visual realism. Since approximately 2015, advances in deep learning, particularly generative adversarial networks (GANs), spatial transformer networks, pose estimation models, and 3D parametric body representations, have significantly accelerated research in image-based clothing transfer .

Early image-based VTO systems such as VITON introduced a two-stage pipeline combining coarse garment warping and refinement networks to transfer a single garment onto a person image [1]. Subsequent models including CP-VTON [2], VITON-HD [3], and HR-VITON [4] improved high-resolution synthesis, texture preservation, and geometric alignment. Parallel research explored pose-guided person image generation [5, 6] and dense pose mapping [7], further enhancing spatial consistency.

More recently, diffusion-based architectures (e.g., Stable Diffusion adaptations for fashion transfer, 2022-2024) and hybrid 2D-3D pipelines incorporating parametric body models such as SMPL Loper et al. in 2025 [8] have expanded realism and controllability. However, a substantial portion of existing work continues to focus on single-garment replacement, most commonly upper-body garments, often evaluated under controlled datasets such as DeepFashion or Zalando.





### 1.1 From Single-Garment to Full-Outfit Virtual Try-On

Full-outfit virtual try-on, where both upper and lower garments are synthesized simultaneously, introduces substantially greater complexity compared to single-item replacement systems. Unlike approaches that modify only one clothing item, full-outfit systems must maintain geometric consistency across the waist–hip transition to ensure that garments align naturally with the body. They are also required to resolve multi-garment occlusion scenarios, such as a shirt tucked into trousers, where overlapping regions must be rendered realistically. In addition, these systems must accurately model body–fabric interaction across multiple articulation points, capturing how garments respond to movement and body posture. Preserving fine texture details and material properties at garment boundaries is equally critical to maintain visual realism. Finally, they must avoid compounding warping artifacts that can arise when independent modules interact, ensuring a seamless and coherent final output.

When multiple garments are transferred sequentially using single-garment pipelines, cumulative alignment errors frequently arise, particularly at the waist seam, hip curvature, and arm–torso intersections. Fabric draping interactions between garments are only partially captured in most 2D-based systems, which may contribute to artifacts such as texture bleeding, limb occlusion errors, unnatural shading, and silhouette distortion. Despite the commercial relevance of full-outfit try-on for e-commerce and digital retail platforms, systematic academic exploration of fully integrated multi-garment pipelines remains comparatively limited, with much of the literature still emphasizing single-item transfer or partially extensible approaches.

### 1.2 Evaluation Limitations in Existing Literature

Evaluation practices across virtual try-on research remain highly inconsistent across studies. Most existing work relies on commonly used image-generation metrics such as the Structural Similarity Index (SSIM), Fréchet Inception Distance (FID), LPIPS perceptual metrics, and limited-sample human preference studies. However, these evaluation tools were not originally designed to assess garment realism, cross-garment structural continuity, or multi-item alignment in full-outfit synthesis. As a result, they may not fully capture critical aspects of clothing fit, boundary coherence, and fabric interaction across garments.

Moreover, many datasets lack true ground-truth full-outfit pairs, making objective validation challenging. Only a limited number of studies assess robustness across diverse body types or varying poses, which may restrict generalizability. Real-world deployment conditions, such as mobile image capture, uncontrolled lighting environments, and background noise, are also relatively underexplored. Collectively, these observations suggest the absence of widely adopted, standardized full-outfit benchmarks, representing an important methodological gap in the field.

### 1.3 The Underexplored Legal and Biometric Dimension

While technical realism in virtual try-on systems has advanced rapidly, the use of real user images introduces substantial legal and ethical considerations. These systems typically process high-resolution facial imagery, full-body biometric data, body measurements, pose keypoints, and, in many cases, potentially reconstructable 3D body models. Such data types extend beyond ordinary visual information and may fall within the scope of sensitive personal data under modern privacy regulations.

Under regulatory frameworks such as the General Data Protection Regulation (GDPR) (EU, 2016/679) and the UK GDPR, biometric data used for the purpose of uniquely identifying a person is classified as “special category data,” which is subject to stricter processing requirements [9]. Even in cases where a system does not explicitly perform identity recognition, the capture of high-resolution facial imagery and detailed body reconstruction may introduce re-





identification risks. The possibility of reverse engineering identity from stored biometric representations raises additional concerns regarding data minimization, proportionality, and lawful basis for processing.

Despite these regulatory implications, many technical publications in the virtual try-on domain provide limited detail regarding data retention policies, consent mechanisms, storage architectures, or biometric classification considerations. This reflects a broader interdisciplinary gap between advances in computer vision and the implementation of privacy-by-design principles. Bridging this gap is essential to support responsible deployment and user trust.

#### 1.4 Emerging Privacy-Preserving Directions

Recent advances in edge computing, on-device inference, and lightweight body parameter estimation suggest the possibility of alternative system architectures that may reduce legal and privacy risks in virtual try-on (VTO) applications. Conceptually, these approaches include on-device processing with immediate deletion of raw images after inference; retention of de-identified body parameters, such as SMPL coefficients, instead of storing photographs; multi-view reconstruction followed by reuse of an avatar-only representation rather than the original user images; and temporary memory pipelines that avoid persistent server-side storage. Such designs aim to align technical performance with privacy-by-design principles by limiting data persistence and reducing exposure of personally identifiable information.

However, several unresolved questions remain. It is not yet fully clear whether these privacy-preserving pipelines have been systematically evaluated within the VTO literature. Furthermore, even when photographs are discarded, de-identified 3D body models may still qualify as biometric data if they enable unique identification or re-identification, raising regulatory concerns under frameworks such as GDPR. Similarly, uncertainty persists regarding whether storing measurement-only parameters fully satisfies the principle of data minimization, particularly if such parameters can be linked back to individuals. These considerations highlight the need for a structured conceptual perspective that jointly examines technical design choices and their regulatory implications.

#### 1.5 Research Objectives and Research Questions

This study aims to provide a structured conceptual analysis of virtual try-on systems with relevance to full-outfit synthesis, with a particular focus on technical limitations, evaluation practices, and privacy-aware architectural design. The objectives of this study are as follows:

- To synthesize representative literature on virtual try-on architectures relevant to full-outfit scenarios
- To identify key technical challenges associated with multi-garment synthesis
- To examine current evaluation practices and highlight potential limitations
- To explore legal and biometric considerations associated with real-user data processing
- To propose a privacy-preserving conceptual architecture for virtual try-on systems

The study is guided by the following research questions:

- **RQ1:** What architectural approaches are used in virtual try-on systems that are extensible to full-outfit synthesis?
- **RQ2:** What technical challenges arise in multi-garment virtual try-on scenarios?





- **RQ3:** How are existing systems evaluated, and what limitations may exist in current evaluation practices?
- **RQ4:** To what extent do current virtual try-on systems address legal and biometric data considerations?
- **RQ5:** How can privacy-preserving design principles be integrated into virtual try-on architectures?

## 1.6 Scope of the Study

This study adopts a targeted literature review approach, focusing on representative works published between January 2015 and February 2026. The starting year reflects the emergence of parametric body modeling and deep generative methods, while the end date captures recent developments in diffusion-based and hybrid architectures. The aim is not exhaustive coverage, but rather a structured synthesis that supports conceptual and interdisciplinary insights.

## 1.7 Contributions of This Study

This study makes the following contributions:

1. Provides a structured synthesis of representative virtual try-on architectures relevant to full-outfit modeling
2. Identifies recurring technical challenges in multi-garment synthesis, including alignment and occlusion issues
3. Highlights limitations in current evaluation practices for virtual try-on systems
4. Explores legal and biometric considerations associated with real-user image processing
5. Proposes a conceptual privacy-preserving architecture based on avatar-only simulation with reduced data retention
6. Offers directions for future research at the intersection of computer vision and privacy-aware system design

## METHODOLOGY

### 2.1 Review Design and Approach

This study adopts a targeted literature review approach aimed at synthesizing representative research on virtual try-on (VTO) systems, with a particular focus on full-outfit modeling and privacy-aware architectural design. Unlike formal systematic reviews, the objective is not exhaustive coverage but rather a structured and concept-driven synthesis of influential and representative works across technical and regulatory domains.

The review process was guided by predefined research questions, thematic relevance, and conceptual alignment with the study objectives, rather than strict protocol-driven inclusion.

Given the rapid evolution of generative modeling in computer vision, the review covers literature published between January 2015 and February 2026, spanning the introduction of parametric body models (e.g., SMPL) through GAN-based virtual try-on systems and more recent diffusion-based architectures.

The analysis focuses on studies that contribute to understanding architectural design, evaluation practices, and privacy implications in virtual try-on systems, particularly those extensible to full-outfit scenarios.

### 2.2 Information Sources





The literature search was conducted across multiple widely used academic databases, including IEEE Xplore, ACM Digital Library, Scopus, Web of Science, SpringerLink, arXiv, and Google Scholar. These sources were selected to ensure broad coverage of peer-reviewed publications, conference proceedings, and relevant preprints in computer vision, machine learning, and interdisciplinary domains.

### 2.3 Search Strategy

Search queries were constructed using structured Boolean combinations of keywords to capture both technical and regulatory dimensions of the literature. These queries incorporated terms related to virtual try-on, clothing transfer, garment synthesis, full-body synthesis, diffusion models, avatar reconstruction, and 3D human modeling.

In parallel, additional keywords such as privacy, biometric data, and GDPR were included to identify studies addressing legal, ethical, and data protection considerations.

The search strategy was iteratively refined to balance breadth and relevance, with emphasis on studies that are conceptually aligned with full-outfit virtual try-on and privacy-aware system design.

Example core search string:

```
("virtual try-on" OR "clothing transfer" OR "garment synthesis")  
AND ("full outfit" OR "top and bottom" OR "multi-garment" OR "full body")  
AND ("GAN" OR "diffusion" OR "3D model" OR "SMPL" OR "pose transfer")
```

For privacy-related analysis:

```
("virtual try-on" OR "avatar reconstruction")  
AND ("privacy" OR "GDPR" OR "biometric data" OR "data protection")
```

Search strings were adapted to the syntax requirements of each database.

### 2.4 Study Selection Criteria

#### 2.4.1 Inclusion Criteria

Studies were included based on the following criteria:

1. Proposed or evaluated a virtual try-on, garment transfer, or avatar-based clothing synthesis system
2. Used real human images or reconstructed human body representations
3. Included full-body modeling, multi-garment synthesis, or architectures conceptually extensible to full-outfit scenarios
4. Reported architectural, experimental, or evaluation details
5. Were published between January 2015 and February 2026

These criteria were applied with flexibility to prioritize conceptual relevance and representation of key methodological approaches.

#### 2.4.2 Exclusion Criteria

Studies were excluded if they:





1. Relied solely on fully synthetic avatars without connection to real human image input
2. Lacked sufficient architectural or methodological description
3. Were non-English publications
4. Represented editorial notes, posters without technical detail, or marketing-oriented whitepapers

## 2.5 Study Selection Process

The selection process involved a multi-stage screening procedure:

1. Title screening to assess relevance to virtual try-on or avatar reconstruction
2. Abstract screening to evaluate applicability to full-body or multi-garment synthesis
3. Full-text review to extract architectural, evaluative, and privacy-related insights

Duplicates were removed prior to screening.

A total of 17 representative studies were included for in-depth analysis, selected based on their relevance to virtual try-on architectures, multi-garment synthesis, and privacy-related considerations. These studies are intended to provide a representative overview of current approaches rather than an exhaustive coverage of the literature. The selected studies span multiple architectural paradigms, including GAN-based, diffusion-based, and hybrid 2D–3D approaches, enabling a comparative and concept-driven synthesis. The category distribution of included studies is summarized in Table 1.

**Table 1:** Category Distribution of Included Studies (n = 17)

Category	Description	Number of Studies (n)
GAN-based architectures (full-outfit capable or extensible)	Warping-based or adversarial models enabling garment transfer with geometric alignment mechanisms	6
Diffusion-based architectures	Generative models based on denoising processes for high-fidelity garment synthesis and global consistency	4
Hybrid 2D–3D (SMPL-guided) approaches	Systems integrating parametric body models with image-based rendering for improved structural realism	4
Multi-view avatar and full-body	Approaches leveraging multi-view capture and 3D reconstruction for reusable avatar-based try-on	3



---

reconstruction  
pipelines

---

Privacy-aware or edge-based systems Architectures incorporating on-device processing, reduced data retention, or privacy-preserving design elements 2

---

## 2.6 Data Extraction Framework

For each included study, a structured data extraction framework was applied to capture technical, evaluative, and regulatory dimensions. Technical variables included the year of publication and architecture type (e.g., GAN-based, diffusion-based, hybrid 2D-3D). The framework documented whether the system incorporated SMPL or other 3D body models, as well as input requirements such as person images, garment images, pose keypoints, segmentation masks, and depth information. Evaluation variables included reported datasets, quantitative metrics (e.g., FID, SSIM, LPIPS), and whether full-outfit evaluation was explicitly conducted. Privacy-related variables included whether identifiable user images were used, the reported storage architecture (e.g., cloud-based or on-device), and whether consent or data handling practices were described. The extraction process was designed to support comparative and thematic analysis rather than quantitative aggregation.

## 2.7 Assessment of Study Limitations and Bias

Given the diversity of study designs and reporting practices, a qualitative assessment of study limitations and potential biases was conducted instead of applying a formal standardized risk-of-bias tool.

This assessment considered four domains:

- Dataset-related limitations, including restricted body diversity and controlled capture conditions
- Evaluation limitations, such as reliance on synthetic ground-truth or limited human studies
- Reporting limitations, including incomplete disclosure of failure cases or computational requirements
- Privacy-related reporting gaps, particularly regarding data storage, retention, and biometric considerations

These assessments are intended to provide interpretive context rather than definitive quality grading of individual studies.

## 2.8 Data Synthesis Approach

Due to heterogeneity in architectural design, input modalities, datasets, and evaluation strategies, quantitative meta-analysis was not considered appropriate.

Instead, a structured qualitative synthesis was conducted, including:

- Architectural categorization of virtual try-on systems
- Comparative analysis of reported technical challenges and failure modes



- Examination of evaluation practices and their limitations
- Thematic analysis of privacy and legal considerations

This synthesis aims to identify recurring patterns and conceptual insights rather than to produce statistically generalizable conclusions. The findings derived from this synthesis inform the development of the proposed privacy-preserving architectural framework presented in subsequent sections.

## METHODS

### 3.1 Overview of Included Studies

A total of 17 representative studies were included for qualitative analysis, reflecting key developments in virtual try-on systems between 2015 and 2026. These studies illustrate the progression from early 2D warping-based generative adversarial network (GAN) approaches to more recent diffusion-based and hybrid 2D–3D pipelines. Rather than constituting an exhaustive dataset, the selected studies provide a structured overview of dominant architectural trends and emerging design directions relevant to full-outfit virtual try-on. The included works can be grouped into four primary architectural paradigms. The first category consists of warping-based GAN architectures, the second includes flow-based deformation models, the third comprises diffusion-based garment synthesis approaches, and the fourth includes 2D–3D hybrid and avatar-based reconstruction pipelines. This categorization directly addresses RQ1 by identifying and organizing the primary architectural approaches that are relevant or extensible to full-outfit virtual try-on systems.

Across these paradigms, a gradual shift can be observed from localized garment warping toward more holistic and globally consistent synthesis approaches, particularly in diffusion-based and 3D-aware systems. Although early systems focused primarily on upper-body garment replacement, several later models introduced multi-region conditioning, multi-category garment transfer, or full-body modeling frameworks that are conceptually extensible to simultaneous top-and-bottom synthesis.

Table 2 summarizes the key characteristics of the included studies, supporting comparative analysis aligned with RQ1 and RQ4.

**Table 2:** Included Studies

#	Reference	Year	First Author	Publisher	Core Contribution	Relevance to Full-Outfit & Privacy
1	[8]	2015	Loper	ACM TOG	Parametric 3D body model enabling realistic human mesh reconstruction	Foundational for full-body avatar reconstruction; enables measurement-only storage pipelines
2	[5]	2017	Ma	NeurIPS	GAN-based pose-conditioned full-body synthesis	Enables full-body garment transformation; precursor to multi-garment reasoning

3	[6]	2018	Siarohin	CVPR	Deformation-aware pose-based human synthesis	Relevant to garment warping across full body
4	[1]	2018	Han	CVPR	Image-based virtual try-on using warping + refinement	Primarily upper-body; architecture extensible to multi-garment
5	[2]	2018	Wang	ECCV	Geometric matching module for clothing alignment	Improves alignment critical for waist/hip transitions
6	[7]	2018	Güler	CVPR	Dense UV mapping of human body from RGB image	Key for full-body correspondence and garment placement
7	[11]	2019	Dong	ICCV	Optical flow-guided clothing transfer	Improves spatial consistency across body regions
8	[12]	2019	Han X	ICCV	Flow-based clothing deformation modeling	Models cloth deformation across articulated body
9	[3]	2021	Choi	CVPR	High-resolution virtual try-on	Improves texture realism; scalable to multi-garment
10	[4]	2022	Lee	ECCV	High-resolution collaborative parsing + warping	Handles occlusion & structural continuity
11	[13]	2022	Rombach	CVPR	Diffusion-based high-res image synthesis	Backbone for modern fashion try-on diffusion models
12	[14]	2025	Xie	ICCV	Diffusion-guided controllable garment synthesis	Enables higher realism in full-body clothing transfer
13	[15]	2024	He K	ECCV	Dataset enabling upper & lower garment transfer	One of few addressing multi-category garment synthesis



14	[16]	2026	Lee	CVPR	Combines SMPL + neural rendering	Strong relevance to avatar-based full outfit pipelines
15	[17]	2024	Zhu	CVF	Diffusion-based clothing transfer	Capable of multi-region conditioning
16	[18]	2025	He	arXiv	Multi-view 3D try-on pipeline	Supports reusable avatars; privacy-relevant
17	[19]	2024	Blalock	arXiv	Edge-based inference with temporary image storage	Directly addresses privacy-preserving architecture

## 3.2 Architectural Taxonomy of Full-Outfit-Capable Systems

### 3.2.1 Warping-Based GAN Architectures

Foundational systems such as VITON and CP-VTON introduced a two-stage framework consisting of geometric alignment followed by appearance refinement through adversarial learning [1,2]. These architectures typically rely on person images, garment inputs, segmentation maps, and pose keypoints. Their strengths include explicit geometric control and computational efficiency.

However, when extended to full-outfit synthesis, several limitations emerge. Garments are often processed independently, which may lead to error accumulation and reduced structural continuity. Cross-garment interaction is generally not explicitly modeled, which may limit realism in layered scenarios. These observations suggest that while warping-based approaches provide a strong foundation, their direct applicability to full-outfit synthesis remains constrained, thereby partially addressing RQ1.

### 3.2.2 Flow-Based and Deformation Modeling Systems

Flow-based systems such as ClothFlow [12] and FW-GAN [10, 11] introduce dense correspondence modeling to improve garment alignment and articulation. While these methods enhance local realism, interaction between garments and multi-layer occlusion reasoning remain only partially addressed. From the perspective of RQ1 and RQ2, these systems demonstrate improved deformation modeling but still face limitations in representing complex multi-garment interactions required for full-outfit synthesis.

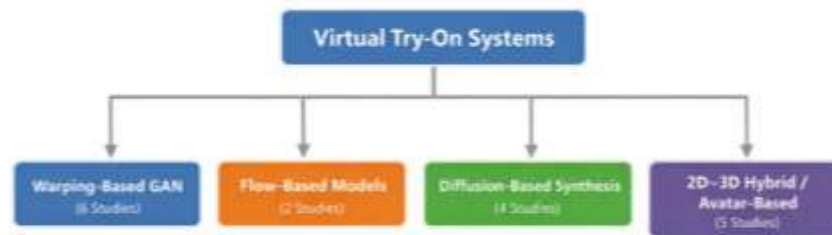
### 3.2.3 Diffusion-Based Garment Synthesis (2022–2026)

Diffusion-based approaches improve global coherence and texture fidelity through probabilistic denoising processes [13]. These models offer advantages such as improved visual consistency and reduced local artifacts. However, explicit modeling of multiple garments remains limited in many implementations. Aligned with RQ1, diffusion-based systems represent a shift toward holistic synthesis, while in relation to RQ2, they highlight ongoing challenges in structured multi-garment reasoning.

### 3.2.4 2D–3D Hybrid and Avatar-Based Pipelines



Hybrid systems integrate parametric body models (e.g., SMPL) with rendering pipelines to improve structural realism. These approaches support better articulation and reusable avatars but introduce increased computational complexity and potential privacy considerations. These findings contribute to RQ1 by identifying hybrid pipelines as a promising direction for full-outfit modeling, while also linking to RQ4 through their implications for biometric data representation. Moreover, taxonomy of architectural paradigms among the 17 included studies full-outfit virtual try-on systems are given below as shown in Figure 1.



**Figure 1:** Taxonomy of Architectural Paradigms Among the 17 Included Studies

### 3.3 Input Requirements Across Studies

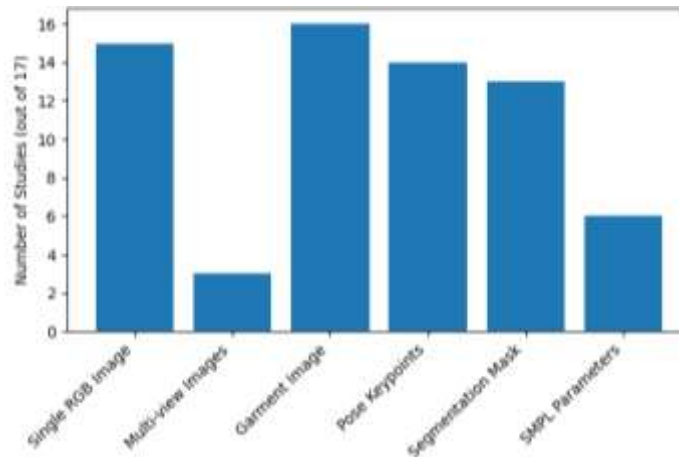
Across the 17 included systems, input requirements vary as shown in Table 3.

**Table 3:** Input Requirements Across Studies

Input Type	Frequency (n=17)
Single-view RGB image	15
Multi-view images	3
Garment image	16
Pose keypoints	14
Segmentation mask	13
SMPL parameters	6

Single-view RGB images and garment images dominate existing pipelines, reflecting the prevalence of image-based approaches. Pose key points and segmentation masks are also widely used. Multi-view inputs and SMPL parameters are less frequent but more common in hybrid systems.

These findings suggest a trade-off between input simplicity and structural accuracy, which is directly relevant to RQ2, as more complex inputs are often required to address multi-garment interaction challenges. Distribution of input modalities are given below in Figure 2.



**Figure 2.** Distribution of input modalities required across included studies

### 3.4 Recurrent Technical Failure Modes

Several recurring technical challenges were identified across the included studies. Waist–hip misalignment remains a commonly observed issue, particularly in systems that process garments independently. Occlusion inconsistencies, fabric detail loss, identity drift, and color mismatch were also reported across multiple studies. These failure modes collectively address RQ2 by highlighting the key technical limitations associated with multi-garment synthesis in current systems. Importantly, the reporting and evaluation of these challenges vary across studies, suggesting that these limitations may not be uniformly assessed.



**Figure 3.** Conceptual illustration of recurrent technical failure modes observed across full-outfit virtual try-on systems

### 3.5 Evaluation Practices Among Included Studies

Evaluation practices vary considerably across the included literature. Common metrics include FID, SSIM, LPIPS, and human preference studies. However, these metrics may not fully capture full-outfit realism, particularly aspects



such as garment interaction and structural continuity. These observations directly address RQ3 by indicating that existing evaluation practices may not be fully aligned with the requirements of full-outfit virtual try-on systems. Evaluation across diverse body types and real-world conditions appears limited, and only a subset of studies incorporates real-user testing. Additionally, explicit evaluation of privacy-related risks is largely absent, further reinforcing limitations in current evaluation frameworks.

### **3.6 Preliminary Privacy Observations**

Analysis of the included studies indicates that identifiable human imagery is frequently used, and cloud-based processing is commonly implied. While some studies acknowledge data handling considerations, detailed reporting on consent, storage, and biometric classification remains limited. A small number of studies explore on-device processing and reduced data retention approaches. These findings address RQ4 by highlighting the current level of engagement with legal and biometric considerations, and RQ5 by identifying emerging directions toward privacy-preserving architectural design. However, systematic evaluation of re-identification risk or regulatory classification of body representations remains limited in the reviewed literature.

## **PROPOSED PRIVACY-PRESERVING ARCHITECTURE FOR FULL-OUTFIT VIRTUAL TRY-ON**

### **4.1 Motivation and Alignment with Findings**

The results presented in Section 3 highlight several recurring patterns across existing virtual try-on systems. In particular, the analysis aligned with RQ2 and RQ3 indicates persistent technical limitations in multi-garment interaction and evaluation practices, while findings related to RQ4 suggest limited integration of privacy and regulatory considerations within current architectures. In response to these observations, this section proposes a conceptual privacy-preserving architecture designed to address both technical and regulatory challenges in full-outfit virtual try-on systems. The proposed framework is informed by the architectural trends identified in RQ1 and aims to operationalize privacy-by-design principles in line with RQ5.

### **4.2 Architectural Design Objectives**

Based on the identified gaps, the proposed architecture is guided by several key design objectives that collectively aim to balance technical performance with regulatory and ethical considerations. First, the architecture prioritizes the minimization of raw image retention, in alignment with RQ4, to reduce risks associated with storing sensitive visual data. In parallel, it emphasizes the reduction of identifiability in stored representations, ensuring that user privacy is preserved even when data is processed or reused. Additionally, the system is designed to support multi-garment structural consistency and interaction modeling, as aligned with RQ2, enabling more accurate and realistic simulations across diverse clothing combinations. The architecture also ensures compatibility with reusable avatar-based simulation pipelines, facilitating scalability and adaptability in virtual try-on applications. Finally, it adheres to privacy-by-design principles, including data minimization and storage limitation, thereby embedding compliance and ethical responsibility directly into the system's core design.

### **4.3 Proposed Four-Stage Pipeline**

The proposed architecture consists of a four-stage pipeline designed to separate personal image processing from garment simulation tasks.

#### **Stage 1: Image Capture and Preprocessing**



The system accepts either a single-view image or a structured multi-view capture (e.g., front, back, and side views). Where feasible, preprocessing is performed locally to reduce unnecessary transmission of raw image data. This stage aligns with RQ4 by minimizing exposure of identifiable data during initial acquisition.

### Stage 2: Parametric Body Model Extraction

A parametric body model (e.g., SMPL) is estimated from the input images, producing shape ( $\beta$ ) and pose ( $\theta$ ) parameters along with derived anthropometric measurements. Instead of storing raw images, the system retains only parametric representations or measurement vectors, thereby reducing identifiability risk. This design choice directly addresses RQ4 and partially responds to the limitations identified.

### Stage 3: Avatar Generation and Data Minimisation

A reusable digital avatar is generated from the extracted parameters. Following avatar generation, raw input images are deleted or discarded, with only de-identified representations retained. Optional facial representations may be included only under explicit user consent. This stage operationalizes privacy-by-design principles and aligns with RQ5 by embedding data minimisation and storage limitation into the system architecture.

### Stage 4: Full-Outfit Simulation Module

Garment simulation is performed exclusively on the avatar representation using diffusion-based rendering, physics-aware modeling, or neural rendering techniques. By decoupling simulation from raw image data, this stage reduces reliance on personal data while enabling reusable and scalable try-on functionality. This component also responds to RQ2 by providing a framework that can support improved multi-garment consistency through unified avatar-based modeling. The complete stages of proposed architecture are given below in Figure 4.

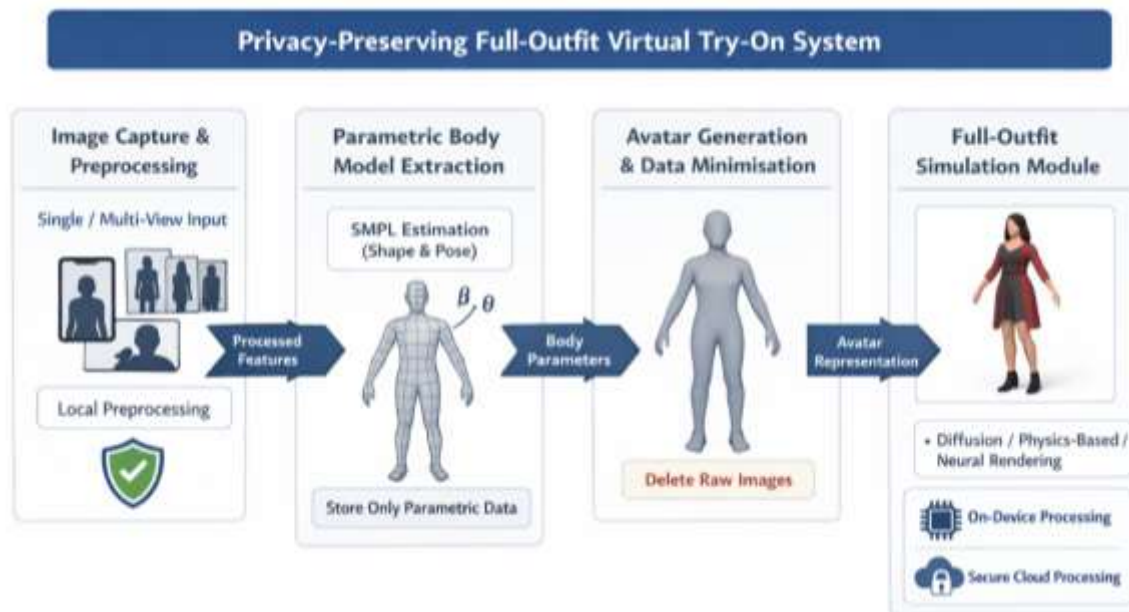


Figure 4: Architecture Diagram



#### 4.4 Edge vs Cloud Deployment Considerations

Deployment architecture plays a critical role in privacy and performance trade-offs. On-device processing reduces data exposure and simplifies compliance but introduces computational constraints. Cloud-based approaches offer scalability but may increase privacy risks if raw data are stored or transmitted without safeguards. A hybrid architecture, combining on-device feature extraction with controlled cloud-based rendering, may offer a balanced solution. This consideration aligns with both RQ4 and RQ5 by addressing practical implementation challenges.

#### 4.5 Open Challenges and Limitations

Despite its conceptual advantages, the proposed architecture introduces several open challenges that warrant careful consideration. One key issue is the trade-off between geometric accuracy and data minimisation, where efforts to reduce stored data may compromise the fidelity of body or garment representation. Additionally, there is a risk of potential reversibility of parametric body representations, raising concerns about whether anonymized models could be reconstructed to reveal identifiable information. The architecture also faces computational constraints for on-device rendering, particularly in resource-limited environments where real-time performance is essential. Furthermore, the limited availability of physics-based multi-garment simulation models restricts the system's ability to accurately capture complex garment interactions. These challenges reflect broader limitations identified in RQ2 and RQ3 and highlight important directions for future investigation.

#### 4.6 Contribution and Positioning

The proposed architecture contributes to the literature by integrating technical design considerations with privacy-aware system principles. Rather than introducing a fully implemented system, it provides a conceptual framework that can guide future development of full-outfit virtual try-on systems under regulatory constraints. In doing so, it directly addresses RQ5 while synthesizing insights derived from RQ1–RQ4.

### DISCUSSION

#### 5.1 Technical Realism and Regulatory Considerations

This study highlights an apparent divergence in the development of virtual try-on (VTO) systems between advancements in visual realism and the integration of privacy and regulatory considerations. Architectural innovation has led to improvements in geometric alignment, texture fidelity, and global coherence, particularly through GAN-based, diffusion-based, and hybrid 2D–3D pipelines. At the same time, the findings suggest that integration of regulatory and privacy-aware design considerations within system architectures remains limited or inconsistently reported. While some studies acknowledge data-related concerns, detailed implementation of data minimisation strategies, legal risk assessment, or verifiable deletion mechanisms is not commonly described. This indicates that technical and regulatory dimensions are often developed in parallel rather than through integrated design approaches. As VTO systems move toward real-world deployment, this separation may introduce challenges related to compliance, transparency, and user trust.

#### 5.2 The Evolving Status of Parametric Body Models

A key issue emerging from this analysis concerns the classification and treatment of parametric body representations, such as SMPL coefficients. From a technical perspective, these parameters encode body shape and pose through structured computational processes and enable detailed reconstruction of human geometry. From a regulatory perspective, their classification depends on context, particularly whether they enable or are used for identification. Current literature provides relatively limited discussion on how such representations should be treated under data protection frameworks, particularly in relation to identifiability and biometric classification. This creates an





interpretive grey area in which technically derived representations may carry latent identification risks without being explicitly addressed within system design or evaluation. Further interdisciplinary investigation may be required to clarify how such representations should be governed within privacy-aware virtual try-on systems.

### **5.3 Full-Outfit Complexity and Privacy Implications**

Full-outfit virtual try-on systems introduce additional layers of complexity compared to single-garment pipelines. Multi-garment synthesis requires more detailed modeling of body geometry, garment interaction, and occlusion handling, often involving richer input data and more expressive representations. These technical requirements may increase the granularity and sensitivity of underlying data representations, particularly in systems involving multi-view capture or 3D reconstruction. As a result, improvements in visual realism may be associated with increased exposure to privacy-related risks, including potential re-identification or inference of sensitive attributes. This suggests a potential trade-off between realism and data minimisation, which remains relatively underexplored in current literature and warrants further investigation.

### **5.4 Limitations of Current Evaluation Practices**

The findings indicate that evaluation practices in virtual try-on research remain largely centered on perceptual similarity metrics such as FID, SSIM, and LPIPS. While these metrics provide useful indicators of visual quality, they may not fully capture properties specific to full-outfit synthesis, including garment interaction, structural continuity, and realistic fit. Evaluation under diverse real-world conditions, such as variations in body types, poses, and environmental settings, also appears comparatively limited. In addition, evaluation frameworks rarely incorporate privacy-related dimensions, such as re-identification risk, data retention transparency, or regulatory compliance considerations. Taken together, these observations suggest that current evaluation approaches may benefit from broader, multi-dimensional frameworks that integrate both technical performance and data governance aspects.

### **5.5 Implications of Privacy-Preserving Architectural Design**

The conceptual architecture proposed in this study illustrates one possible approach to integrating privacy considerations into virtual try-on system design. By separating image acquisition from simulation through avatar-based processing and reducing reliance on raw image storage, the framework demonstrates how privacy-by-design principles may be embedded within system pipelines. However, several practical challenges remain, including potential trade-offs between geometric fidelity and data minimisation, computational constraints associated with on-device processing, and uncertainties regarding the reversibility of parametric representations. These considerations suggest that while privacy-preserving architectures are conceptually viable, their practical implementation requires further empirical validation and technical development.

### **5.6 Toward an Interdisciplinary Research Framework**

Virtual try-on research increasingly spans multiple domains, including computer vision, generative modeling, human-computer interaction, and data protection. However, these domains often remain only loosely integrated, with limited cross-disciplinary evaluation or design frameworks. Future research may benefit from more explicitly interdisciplinary approaches that combine technical system development with legal, ethical, and governance-oriented analysis.

Potential directions for future work include several important areas that can enhance both the robustness and trustworthiness of virtual try-on systems. These include the development of privacy-aware dataset design strategies that ensure sensitive user information is protected from the outset, as well as the creation of evaluation metrics that incorporate not only structural performance but also regulatory and ethical dimensions. Additionally, there is a need for transparent consent frameworks and data governance mechanisms to ensure users retain control over their data,





alongside comprehensive system-level documentation that clearly outlines privacy and compliance considerations. Collectively, these approaches may support the development of more reliable, trustworthy, and deployable virtual try-on systems.

### 5.7 Synthesis of Key Research Gaps

The combined analysis highlights several key areas that require further investigation. There remains a lack of standardization in evaluation metrics for full-outfit synthesis, making it difficult to consistently assess system performance across studies. Ongoing challenges also persist in accurately modeling multi-garment interaction and maintaining structural continuity in complex scenarios. Furthermore, there is uncertainty surrounding the legal classification of parametric body representations, particularly in the context of privacy and data protection regulations. Another critical gap is the limited integration of privacy considerations into system-level design, indicating that privacy is often treated as an afterthought rather than a foundational component. Finally, there is noticeable fragmentation between technical and regulatory research perspectives, which can hinder the development of holistic solutions.

These gaps should be viewed as opportunities for advancement rather than indications of stagnation in the field. Addressing them has the potential to significantly improve both the technical capabilities and regulatory alignment of future virtual try-on systems.

### CONCLUSION

This study provides a structured synthesis of virtual try-on systems, highlighting both technological advancements and persistent limitations across multi-garment modeling, evaluation practices, and privacy integration. While recent approaches demonstrate strong progress in visual realism and generative capabilities, they continue to face limitations in accurately modeling complex garment interactions, maintaining structural consistency across full outfits, and embedding privacy mechanisms at the system level. Additionally, this work is limited by its reliance on existing literature, which itself reflects fragmentation in evaluation standards and regulatory interpretation, potentially constraining the generalizability of the proposed architectural perspective.

Future research should focus on developing standardized evaluation frameworks that incorporate both structural fidelity and regulatory compliance, as well as advancing physics-based multi-garment simulation models to improve realism. There is also a need to investigate the legal and technical implications of parametric body representations, particularly in relation to identifiability and data protection. Moreover, integrating privacy-aware dataset design, transparent consent mechanisms, and system-level documentation can support more trustworthy implementations. Addressing these directions will be critical for enabling scalable, privacy-preserving, and practically deployable virtual try-on systems.

### REFERENCES

1. Han X, Wu Z, Wu Z, Yu R, Davis LS. Viton: An image-based virtual try-on network. In Proceedings of the IEEE conference on computer vision and pattern recognition 2018 (pp. 7543-7552).
2. Wang B, Zheng H, Liang X, Chen Y, Lin L, Yang M. Toward characteristic-preserving image-based virtual try-on network. In Proceedings of the European conference on computer vision (ECCV) 2018 (pp. 589-604).
3. Choi S, Park S, Lee M, Choo J. Viton-hd: High-resolution virtual try-on via misalignment-aware normalization. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition 2021 (pp. 14131-14140).





4. Lee HS, Oh SH, Ha SH. Virtual Fitting System Using Deep Learning Methodology: HR-VITON Based on Weight Sharing, Mixed Precision & Gradient Accumulation. *The Journal of Information Systems*. 2022;31(4):145-60.
5. Ma L, Jia X, Sun Q, Schiele B, Tuytelaars T, Van Gool L. Pose guided person image generation. *Advances in neural information processing systems*. 2017;30.
6. Siarohin A, Sangineto E, Lathuiliere S, Sebe N. Deformable gans for pose-based human image generation. *InProceedings of the IEEE conference on computer vision and pattern recognition 2018* (pp. 3408-3416).
7. Güler RA, Neverova N, Kokkinos I. Densepose: Dense human pose estimation in the wild. *InProceedings of the IEEE conference on computer vision and pattern recognition 2018* (pp. 7297-7306).
8. Loper M, Mahmood N, Romero J, Pons-Moll G, Black MJ. SMPL: A skinned multi-person linear model. *InSeminal Graphics Papers: Pushing the Boundaries, Volume 2 2023 Aug 1* (pp. 851-866).
9. Deac A. Regulation (eu) 2016/679 of the european parliament and of the council on the protection of individuals with regard to the processing of personal data and the free movement of these data. *Perspectives of Law and Public Administration*. 2018;7(2):151-6.
10. Kuppa G, Jong A, Liu X, Liu Z, Moh TS. Shineon: Illuminating design choices for practical video-based virtual clothing try-on. *InProceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision 2021* (pp. 191-200).
11. Dong H, Liang X, Shen X, Wu B, Chen BC, Yin J. Fw-gan: Flow-navigated warping gan for video virtual try-on. *InProceedings of the IEEE/CVF international conference on computer vision 2019* (pp. 1161-1170).
12. Han X, Hu X, Huang W, Scott MR. Clothflow: A flow-based model for clothed person generation. *InProceedings of the IEEE/CVF international conference on computer vision 2019* (pp. 10471-10480).
13. Rombach R, Blattmann A, Lorenz D, Esser P, Ommer B. High-resolution image synthesis with latent diffusion models. *InProceedings of the IEEE/CVF conference on computer vision and pattern recognition 2022* (pp. 10684-10695).
14. Xie Y. Exploring The Practical Applications of Text-Guided Image Stylization: A Case Study of Diffstyler. *InITM Web of Conferences 2025 (Vol. 78, p. 04011)*. EDP Sciences.
15. He K, Yao K, Zhang Q, Yu J, Liu L, Xu L. Dresscode: Autoregressively sewing and generating garments from text guidance. *ACM Transactions on Graphics (TOG)*. 2024 Jul 19;43(4):1-3.
16. Lee J, Jung H, Choi J. 3D-aware virtual try-on using only 2D inputs. *Computer Vision and Image Understanding*. 2026 Jan 21:104661.
17. Zhu L, Yang D, Zhu T, Reda F, Chan W, Saharia C, Norouzi M, Kemelmacher-Shlizerman I. Tryondiffusion: A tale of two unets. *InProceedings of the IEEE/CVF conference on computer vision and pattern recognition 2023* (pp. 4606-4615).
18. He Z, Ning Y, Qin Y, Wang G, Yang S, Lin L, Li G. Vton 360: High-fidelity virtual try-on from any viewing direction. *InProceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2025* (pp. 26388-26398).
19. Blalock J, Munechika D, Karanth H, Helbling A, Mehta P, Lee S, Chau DH. Mobile fitting room: On-device virtual try-on via diffusion models. *arXiv preprint arXiv:2402.01877*. 2024 Feb 2.

