iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 13 **Issue:** XI **Month of publication:** November 2025

DOI: https://doi.org/10.22214/ijraset.2025.74662

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XI Nov 2025- Available at www.ijraset.com

StyleSense: A Multimodal AI Framework for Mitigating E-commerce Fit Uncertainty via Hyper-Realistic Virtual Try-On and Agentic Styling

Dr. C. Udhaya Shankar[1], Elamurugan R[2], Lochana T[3]
[1]*Professor, Department of Artificial Intelligence and Data Science, SNS College of Engineering, Coimbatore, Tamil Nadu, India*
[2, 3]*Bachelor of Technology, Department of Artificial Intelligence and Data Science, SNS College of Engineering, Coimbatore, Tamil Nadu, India*

Abstract: ***The rapid expansion of online fashion commerce has introduced critical challenges, primarily high fit uncertainty, resulting in online apparel return rates that often exceed 24% and translate into billions in annual costs. This high-cost issue is compounded by inefficient, generic search mechanisms and a significant lack of personalized, in-store-like guidance, collectively driving customer dissatisfaction and high logistical burdens.***

The StyleSense project presents a next-generation, AI-powered e-commerce platform designed as a comprehensive solution to these fundamental industry deficiencies. The core innovation is the integration of a three-pillar, multimodal system:

1) ***A hyper-realistic virtual try-on system that utilizes customizable 3D avatars and deep learning for advanced garment draping simulation, thereby substantially boosting purchase confidence and mitigating fit anxiety.***
2) ***An AI-powered style assistant based on GPT/Gemini conversational AI, which curates personalized fashion advice, ensemble suggestions, and sophisticated mix-and-match guidance, effectively mimicking a human consultant.***
3) ***Visual style search, powered by CLIP (Contrastive Language–Image Pretraining) and vector databases***

Architected on a robust full-stack foundation (React, Django/FastAPI), the system further incorporates AI-summarized product reviews and dynamic pricing models to ensure informed decision-making and market competitiveness. By unifying these advanced technologies, StyleSense effectively bridges the experiential gap between digital convenience and physical engagement. The result is a highly reliable and customer-centric fashion ecosystem engineered to significantly reduce returns, lower operational costs, and establish a scalable and sustainable marketplace.

Keywords: StyleSense, Virtual Try-On, 3D Avatars, Gemini AI, Conversational AI, CLIP, Visual Search, E-commerce Returns.

I. INTRODUCTION

The profound shift towards fashion e-commerce as the dominant consumer channel has inadvertently exposed fundamental flaws in the digital shopping experience. The most acute of these deficiencies is fit uncertainty, which manifests as a high product return rate, severely hindering operational efficiency and generating a significant environmental footprint. Quantitative data underscores the severity of this issue.

While the average e-commerce return rate typically hovers around 18.1% , categories requiring specific physical fit, such as clothing and footwear, see these rates surge dramatically. Estimates indicate that online apparel and footwear returns in the United States account for approximately 24.4% of all orders.

The economic consequences of this return volume are substantial. It is estimated that a return rate of 24.4% translates into $38 billion in returned merchandise and an estimated $25.1 billion annually in processing costs across the US apparel and footwear market. This financial burden directly correlates with one of the primary objectives of the StyleSense platform: lowering operational costs through proactive prevention of returns. The underlying cause of this high return frequency is the inherent information asymmetry between physical and digital retail.

Customers lack the crucial tactile and visual data points—such as the feel of the fabric, the true drape, or how an item fits their unique body shape—which they rely upon in physical stores. This deficiency encourages high-risk purchasing behaviors, notably "bracketing," where shoppers buy multiple sizes or variations of an item with the intention of returning what does not fit, further inflating return statistics.

II. EXISTING SYSTEM

The current fashion e-commerce landscape is dominated by major platforms (such as Amazon, Flipkart, Myntra, and Shein), which primarily function as digital catalogs for browsing and purchasing clothing and accessories. While widely adopted, these systems are fragmented and largely lack the integrated, intelligent features necessary to provide a truly efficient, personalized, and confidence-boosting shopping experience.

These conventional methodologies are frequently plagued by inefficiencies, resulting in suboptimal outcomes for both customers and businesses.

A. Key Limitations of Existing E-commerce Systems

1) Static Visualization and Severe Fit Uncertainty: Fixed and Static Solutions: Clothing items are displayed solely through static product images or simple videos. Customers are unable to realistically visualize how garments would drape or fit on their unique body shape, leading to acute fit uncertainty and a lack of confidence during the purchase process.Limited Virtual Try-On: Although some platforms have experimented with basic Augmented Reality (AR) mirrors or simple 2D overlays, these solutions are neither immersive nor fully functional. They lack customizable 3D avatars and the realistic garment draping simulation required to build genuine trust in the product's actual fit and look.

2) Inadequate Personalization and Generic Recommendations: Steep Tool Learning Curves: Existing search and discovery tools (traditional search filters) are too generic and inflexible to accurately capture a user's nuanced preferences, such as specific style, mood, or occasion.

3) Deficiency of Conversational Style Guidance: Reliance on Passive Systems: E-commerce sites may employ simple chatbots, but these are primarily for basic Frequently Asked Questions (FAQs) and order support.

4) Consequences: High Costs and Operational Inefficiency: Due to pervasive size mismatches, fit anxiety, and poor visualization, retailers incur high operational costs from frequent product returns. This phenomenon contributes significantly to environmental waste and logistical strain.

B. Objective

The The primary objective of the StyleSense project is to develop an AI-powered, interactive platform that addresses the core deficiencies of online fashion retail. It aims to combine hyper-realistic virtual try-on, personalized style consultation, and intelligent product discovery to significantly increase purchase confidence and establish a sustainable marketplace. The key objectives for the StyleSense project are as follows:

1) Enhance Shopping Confidence and Reduce Returns: To implement a hyper-realistic virtual try-on system using customizable 3D avatars and advanced garment draping, allowing customers to visualize fit before purchase, thereby mitigating fit uncertainty and hesitation.

2) Provide Personalized Style Guidance: To create an AI-powered style assistant (using GPT/Gemini) that offers expert outfit curation, mix-and-match suggestions, and tailored fashion advice based on user profiles, preferences, and past shopping behavior.

3) Enable Smarter Product Discovery and Search: To develop and deploy a visual style search system (using CLIP and vector databases) where users can upload images to find similar products.

4) Boost Purchase Confidence with Data: To introduce AI-powered review summarization using Natural Language Processing (NLP) to generate concise summaries of customer feedback, extracting key pros and cons to facilitate more informed purchasing decisions.

5) Increase User Engagement and Loyalty: To introduce engaging features such as personalized feeds and dynamic pricing models that adapt to market trends, encouraging customers to spend more time on the platform and increase return frequency.

C. Scope

1) Hyper-Realistic Virtual Try-On (VTO) System: The development and integration of a core VTO module that allows users to generate personalized 3D avatars and employs deep learning for advanced garment draping simulation, directly solving the problem of fit uncertainty in online shopping.

2) Integrated AI Style Consultation: The creation of an expert style assistant utilizing GPT/Gemini conversational AI to deliver personalized fashion advice, curate complete outfits, and provide sophisticated, context-aware styling suggestions based on machine learning.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue XI Nov 2025- Available at www.ijraset.com

3) Advanced Visual Intelligence and Discovery: Implementation of a robust visual style search feature, leveraging CLIP (Contrastive Language–Image Pretraining) and vector databases, enabling users to find products across the catalog using inspiration images.

4) Full-Stack E-commerce Foundation: The establishment of a complete, secure, and functional e-commerce infrastructure, encompassing user account management, a comprehensive product catalog, shopping cart functionality, and secure integrated payment processing.

III. SOFTWARE REQUIRMENTS

A. Programming Language

The Stylesense project leverages Python as its backend programming language. This choice was primarily due to its versatility, ease of use, and convenience for rapid web application development. Python's clarity and extensive range of application development frameworks make it a solid candidate for scalable application development, particularly for applications incorporating AI-driven features such as recommendation systems and automation planning. In front-end development, an interactive and responsive interface is primarily rendered using HTML, CSS, and JavaScript. A clean, customizable, and mobile-responsive UI design is easily achieved with Tailwind CSS. The technologies described above enable the system to provide users with a tailored experience for planning, customizing, and managing smart home setups.

B. System Requirements

RAM: 4GB or 8GB
Windows 10
Intel Core i5/i7 processor
At least 60 GB of Usable Hard Disk Space

C. Libraries and Frameworks

1) Django

Django serves as the primary backend framework for the Stylesense managing all core functionalities such as user requests, aut hentication, and database interactions. Its modular and M VC-based architecture ensures that the system is scalable, maintainable, and easy to extend with additional features. Django also facilitates seamless integration with AI modules , APIs, and the cost estimation engine. The framework supports secure data handling and rapid development of web-based applications, making it ideal for handling the multiple interconnected modules of this project.

2) HTML, CSS, and JavaScript

HTML, CSS, and JavaScript form the backbone of the Stylesense's frontend architecture. HTML is used to structure the content and define the layout of web pages, ensuring that every element, from device selection menus to room layout designs, is properly organi zed. CSS provides styling and visual presentation, allowing the application to deliver an engaging and user-friendly experience. By c ontrolling fonts, colors, layouts, and transitions, CSS ensure the platform is both aesthetically pleasing and intuit JavaScript, on the other hand, brings interactivity to the application. It enables real-time updates without requiring a full page reload, such as dynamica lly updating device recommendations, validating user input instantly, and managing interactive features like drag-and-drop room.

3) FastAPI

Asynchronous API calls are handled by FastAPI, which guarantees an effective frontend, Gemini AI, and Blender script connection. It offers high-performance endpoints that can handle several requests simultaneously, which is essential for AI-based layout recommendations made in real time. FastAPI is an excellent option for creating dependable and responsive backend services because of its support for automatic data validation, documentation, and concurrency. Because it is quick and lightweight, it works well with Django's framework by handling AI and 3D rendering requests asynchronously.

4) Gemini

The Gemini powers the AI Style Assistant, functioning as an expert virtual consultant. It leverages its sophisticated reasoning to provide personalized fashion advice, curate multi-item ensemble suggestions, and offer sophisticated mix-and-match guidance. The model is engineered for multi-turn reasoning and agentic planning, maintaining context to iteratively refine recommendations based on user-defined constraints like budget or occasion

5) SQLite

SQLite is a lightweight relational database used to store all user-related data, including input preferences, AI-generated designs, 3D model references, and detailed cost breakdowns. It is appropriate for projects requiring less database maintenance overhead because of its simplicity, setup convenience, and quick read/write operations. SQLite facilitates concurrent access, guarantees data persistence, and enables efficient retrieval of historical information. This allows the Stylesense to provide individualized, dependable, and consistent user experiences by maintaining an orderly and structured database.

D. System Integration and Testing

The StyleSense platform, integrating advanced multimodal AI, 3D rendering, and e-commerce services, requires rigorous system integration and testing to ensure a seamless, reliable, and confidence-boosting user experience. The testing regimen is designed to validate the precision of AI output, the performance of complex 3D visualizations, and the cohesion of all component modules.

1) Unit Testing

Each critical module is tested independently to verify stability, accuracy, and correctness, particularly focusing on the highly technical AI components.

- Virtual Try-On (VTO) Testing: Unit tests validate the precise generation of the 3D avatar from user input (e.g., measurements or single image), and the stability of the deep learning garment draping simulation across various fabric types (e.g., stretch, silk, denim) and complex poses.
- AI Style Assistant Testing: Tests confirm the response quality and conversational flow of the Gemini AI, verifying that the NLP review summarization correctly extracts key sentiment, fit notes, and pros/cons from unstructured text.
- Visual Search Testing: Individual tests validate the reliable encoding of user-uploaded images by the CLIP model into high-dimensional vectors, and verify the speed and accuracy of the k-Nearest Neighbor (KNN) search within the vector database.

2) Integration Testing

This phase ensures that the distinct AI modules communicate flawlessly and that data outputs are correctly interpreted across the platform.

- AI-to-VTO Data Flow: Verification that the product selected via the AI Style Assistant is accurately loaded and rendered in the VTO 3D environment with the correct material properties and color.
- RAG Pipeline Integrity: Testing ensures that the Retrieval-Augmented Generation (RAG) pipeline dynamically retrieves current product attributes, inventory data, and AI-summarized reviews to properly ground the Gemini response, ensuring factually accurate advice.

3) Performance Testing

The system's performance is benchmarked under simulated high-load scenarios to ensure a low-latency user experience for computationally intensive tasks.

- VTO Rendering Latency: Performance tests measure the time required for 3D avatar generation and the completion of garment draping simulation, ensuring that the hyper-realistic visualization does not introduce unacceptable delays, leveraging FastAPI's asynchronous capabilities.
- AI Response Time: Benchmarking evaluates the response time of the Gemini AI for complex queries (e.g., "Style an outfit for a business casual event in rainy weather") under peak user load, ensuring the RAG lookup and generation process is optimized.

4) Validation and Verification

Beyond technical functionality, this critical stage assesses the platform's core value proposition: customer confidence and business impact.

- Fit Accuracy Validation: VTO results are rigorously verified against quantitative technical metrics, including Measurement Correlation (ensuring digital measurements match real ones, ideally within 1–2 cm) and Semantic Region Distance (SDR), which assesses the structural integrity of the draped garment.
- Recommendation Efficacy: The quality of the AI Style Assistant's recommendations is measured using standard ranking metrics such as NDCG@N (Normalized Discounted Cumulative Gain) and MRR (Mean Reciprocal Rank), confirming the relevance and predictive accuracy of suggested items or outfits.

IV. IDEATE

A. *Proposed System*

1) Requirement Collection and Preprocessing: This phase focuses on gathering both explicit user inputs and raw catalog data, and then converting them into standardized, intelligent representations for subsequent processing. User requirements are gathered via three main channels: key body measurements or a single image for 3D avatar parameterization (SMPL model), natural language queries for the Gemini style assistant's conversational intent, and inspiration image uploads for visual discovery. These raw inputs are then preprocessed: body data is standardized into SMPL parameters, conversational requests are analyzed via NLU into structured intent signals, and images are encoded into high-dimensional query vectors using the CLIP Vision Encoder for rapid search.
2) AI-Powered Design Recommendation: This module uses Gemini AI to decipher user specifications and produce optimum layout recommendations. Gemini AI creates customized floor designs by using contextual reasoning and natural language understanding, guaranteeing ideal lighting, ventilation, space use, and design flow. The methodology allows for interactive design modifications by adjusting recommendations in response to user feedback. Technical architectural planning and user-friendly interactions are connected through this conversational design method.
3) 3D Modelling and Visualization: Blender The core value proposition of StyleSense rests on its ability to transcend the static limitations of existing e-commerce through hyper-realistic 3D Modeling and Visualization. The process starts with generating personalized 3D avatars utilizing the SMPL (Skinned Multi-Person Linear Model), which translates user measurements or a single input image into an accurate digital representation. The critical innovation is the advanced garment draping simulation. This simulation employs deep learning and physics engines to accurately depict complex non-rigid deformation, ensuring realistic visualization of fabric properties like stretch, weight, and folds.
4) Cost Estimation and Budget Planning: To control the variable cost of the Gemini AI Style Assistant, the team should leverage the Retrieval-Augmented Generation (RAG) architecture efficiently. By grounding the LLM with contextually relevant product and review data from the vector database, the response quality is maximized while minimizing the need for larger, more expensive, or longer context windows.
5) Web-Based Interface: The main platform for user interaction is a web interface driven by the Django framework. Homeowners can use this tool to enter their specifications, view 3D models, explore several design iterations, obtain thorough reports, and receive AI-generated plans. Owing to its responsive and accessible design, the interface works seamlessly on various devices.
6) Interactive Chat and Real-Time Refinement: Users can instantly improve their design choices using the system's interactive chat assistant powered by Gemini AI. Users can browse the recommendations promptly and make inquiries. This conversational feature ensures iterative, collaborative, and user-driven planning.
7) Reporting and Documentation: Finally, Documentation includes a Cost Estimation and Budget Planning Report detailing fixed development costs and variable operational expenditures (OpEx) for high-cost services like the Gemini API and VTO GPU compute. The Future Scope section formally outlines the strategic roadmap for next-generation features, including Augmented Reality (AR) live try-on and Emotional AI integration.

B. *Advantages*

The StyleSense platform offers compelling advantages by directly addressing the fundamental deficiencies of conventional digital fashion retail, leading to improved profitability and customer loyalty. The core benefit is the profound reduction in fit uncertainty and high return rates—a multi-billion-dollar problem—achieved through a hyper-realistic Virtual Try-On (VTO) system that utilizes personalized 3D avatars and deep learning for advanced garment draping simulation, substantially boosting purchase confidence. This foundational feature is complemented by the AI-powered style assistant, which leverages GPT/Gemini conversational AI to provide human-level, personalized fashion advice, outfit curation, and sophisticated mix-and-match guidance. Additionally, the platform enhances product discovery through visual style search, which uses CLIP and vector databases, allowing users to find items based on inspirational images rather than generic filters. By integrating these intelligent, interactive tools, StyleSense transforms the passive shopping experience into a reliable, engaging consultation, which concurrently lowers operational costs and establishes a foundation for a more sustainable e-commerce ecosystem. StyleSense provides compounded benefits that secure the platform's long-term commercial viability. Crucially, the substantial reduction in high return rates—which currently cost the apparel and footwear market billions in processing costs—directly lowers operational expenditures and drives retailer profitability. Beyond cost savings, the platform transforms user engagement by offering a sophisticated AI Style Assistant (Gemini-powered) that provides personalized, multi-turn advice grounded in real-time inventory and AI-summarized product reviews, which enhances decision confidence.

V. RESULT AND SCREESNSHOTS

A. Output

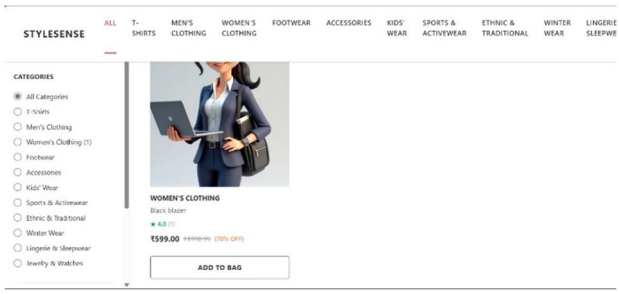


Fig 1: Output Page 1

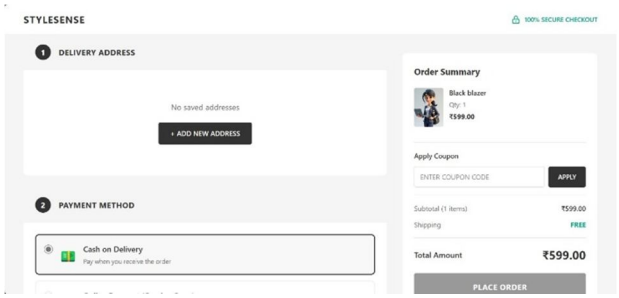


Fig 2: Output Page 2

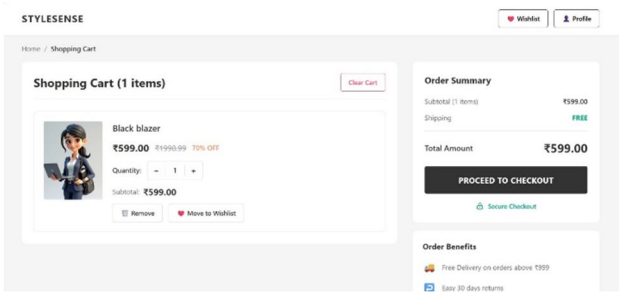


Fig 3: Output Page 3

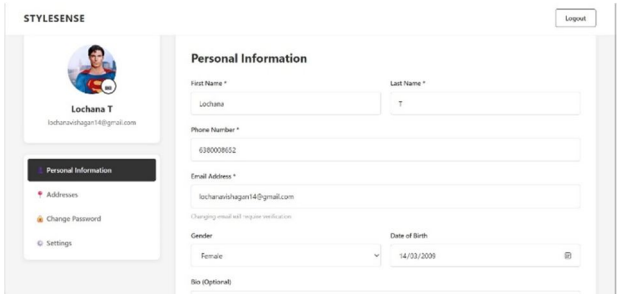


Fig 4: Output Page 4

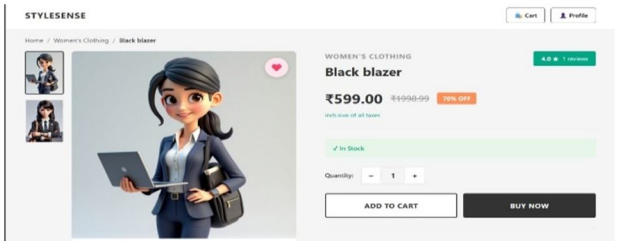


Fig 5: Output Page 5

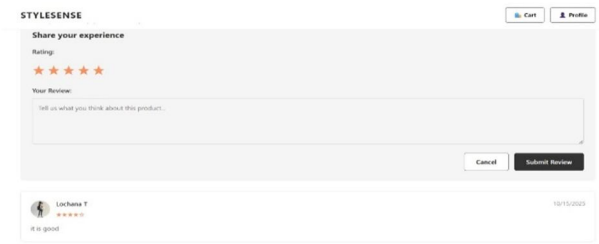


Fig 6: Output Page 6

VI. CONCLUSION

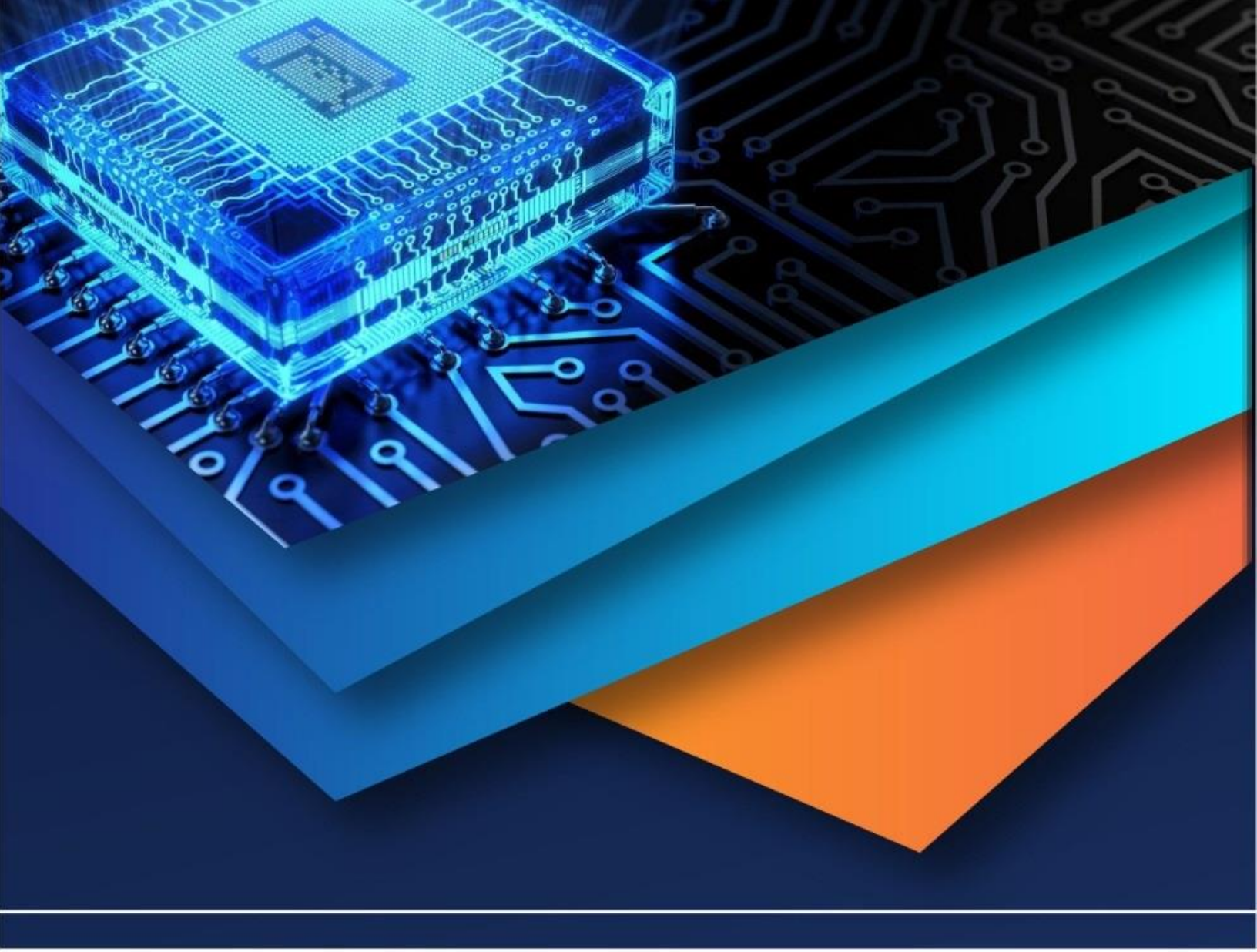
The StyleSense project successfully addresses critical operational and customer challenges in online fashion retail, primarily the multi-billion dollar problem of high return rates driven by fit uncertainty and a lack of personalized guidance. It achieves this by deploying a sophisticated, multimodal AI architecture centered around a hyper-realistic Virtual Try-On (VTO) system using personalized 3D avatars for accurate garment draping, which is seamlessly integrated with a GPT/Gemini conversational AI style assistant that provides expert, grounded outfit curation. Further enhancing the experience, the platform utilizes CLIP-powered visual search and AI-summarized reviews to transform product discovery from passive browsing into an engaging, high-confidence consultation. By unifying visualization accuracy with intelligent, factual guidance, StyleSense not only enhances customer trust and engagement but also establishes a scalable, sustainable e-commerce ecosystem designed to lower operational costs and pave the way for future innovations, such as Augmented Reality (AR) live try-on and advanced emotional intelligence.

VII. FUTURE SCOPE

The future scope for the StyleSense platform is strategically defined by a focus on enhancing digital immersion, deepening personalization, and optimizing global scalability. Building upon the existing 3D framework, future development will introduce Advanced Virtual Realism, primarily through the integration of Augmented Reality (AR) Live Try-On capabilities, allowing users to superimpose garments onto their bodies in real-time using mobile cameras for a live, immersive experience. This realism will be further refined by incorporating Complex Fabric Physics into the simulation, accurately modeling nuanced material properties such as stretch, shine, and transparency. For personalization, the AI Style Assistant will advance to include Emotional Intelligence, enabling the system to infer and respond to user mood or tone during conversational queries, providing suggestions that align with the user's psychological state. Further intelligence integration includes connecting with Wearable Devices to recommend outfits dynamically based on external data like real-time weather or health data. Finally, the project will focus on commercial expansion, scaling the system into a Multi-Vendor Marketplace and prioritizing Sustainability by leveraging AI for accurate demand forecasting, thereby reducing overproduction and waste within the fashion supply chain.

REFERENCES

[1] Generative AI and LLMs (Google Gemini API): The official documentation for Google's Gemini API, which provides the generative AI capabilities for the core AI Style Assistant, driving personalized advice, ensemble creation, and conversational guidance.
[2] 3D Avatar Modeling (SMPL): Research papers detailing the Skinned Multi-Person Linear Model (SMPL), which is essential for accurately generating and customizing the user's realistic 3D avatar for the virtual try-on system.
[3] Garment Simulation (Deep Learning): Research on advanced techniques, such as Generative Adversarial Networks (GANs) and Neural Rendering, used to achieve hyper-realistic and accurate garment draping simulation over the 3D avatar.
[4] Cross-Modal Retrieval (CLIP): Technical resources and papers detailing the CLIP (Contrastive Language–Image Pretraining) Framework, which powers the visual style search by matching user-uploaded images to product text and visual descriptions.
[5] Data Grounding (RAG Architecture): Documentation supporting the Retrieval-Augmented Generation (RAG) architecture, which is implemented to ground the Gemini conversational output in real-time, factual data like inventory and AI-summarized reviews.
[6] High-Performance Backend (FastAPI): The official documentation for FastAPI, the Python web framework used to create the scalable, asynchronous backend API essential for managing the high-throughput demands of simultaneous VTO rendering and AI model inference.
[7] Visual Search Infrastructure (Vector Databases): Documentation for vector indexing tools (e.g., Pinecone/Faiss) necessary to optimize the speed and efficiency of the high-dimensional similarity search required for visual queries and RAG lookups.
[8] Client-Side 3D Rendering (Three.js): The documentation for the Three.js JavaScript library, which is used to render the 3D avatars and the virtual try-on environment directly within the user's web browser with high fidelity.
[9] Frontend Interface (React.js): The official documentation for React.js, the JavaScript library used for building the dynamic, responsive, and engaging frontend user interface that hosts the VTO console and conversational assistant.
[10] E-commerce Foundation Standards: Standard engineering practices and frameworks (including Django) used for building the full-stack foundation, covering secure user account management, the product catalog, shopping cart, and integrated payment gateways.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY