IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.133101

AI-Based Multimodal Indian Fashion Recommendation System Using Computer Vision and Regional Content-Based Filtering

Vetrivel P[1], Dr. M. Kaliappan[2], Akshayanivasini M[3]

Assistant Professor, Department of Artificial Intelligence and Data Science, Ramco Institute of Technology Rajapalayam, India[1]

Professor, Department of Artificial Intelligence and Data Science, Ramco Institute of Technology Rajapalayam, India[2]

Student, Department of Artificial Intelligence and Data Science, Ramco Institute of Technology Rajapalayam, India[3]

Abstract: The rapid growth of digital fashion retail platforms has created a growing demand for recommendation engines capable of addressing the unique cultural, regional, and aesthetic diversity of Indian traditional attire. Existing systems predominantly rely on Western clothing datasets and fail to capture the nuanced interplay between Indian skin tones, body morphologies, regional textile traditions, and occasion-specific norms. This paper presents a comprehensive AI-based multimodal fashion recommendation system specifically designed for Indian traditional and regional clothing. The proposed architecture integrates a multimodal computer vision API to extract key biometric attributes — specifically skin tone classification from face photographs and body morphology estimation from full-body images — into a five-class skin tone schema and a seven-category body shape taxonomy. A custom rule-based content-based filtering engine, conditioned on an eight-dimensional user profile vector, then maps extracted biometric features to regionally appropriate Indian outfits across four geographic zones and eight occasion categories. An AI-assisted natural language generation module further enriches outputs with culturally contextualised descriptions, styling guidance, and personalised notes. The system is deployed as a full-stack web application using a React (Vite) frontend and a Python Flask backend communicating over a REST API. Experimental evaluation confirms sub-150 millisecond offline recommendation latency, strict cultural accuracy enforced by regional boundary filters, and positive user acceptance in preliminary testing. Comparative observations confirm that the decoupled biometric-extraction and recommendation architecture substantially reduces cultural misclassification over general-purpose generative approaches.

Keywords: Indian Fashion Recommendation; Computer Vision API; Content-Based Filtering; Skin Tone Detection; Body Morphology Classification; Regional Outfit Database; Flask; React; Multimodal AI; Cultural Recommendation System.

I. INTRODUCTION

The global fashion technology landscape has witnessed significant growth over the past decade, driven by advances in computer vision, deep learning, and large language models. However, a substantial gap persists between the needs of South Asian consumers and the capabilities of commercially available outfit recommendation systems. India alone encompasses over twenty-two officially recognised languages, hundreds of distinct regional textile traditions, and a complex cultural calendar of occasions — from Durga Puja and Navratri to corporate formal events and casual dating outings — each demanding very different sartorial choices. Western-centric training datasets and generalised recommendation algorithms fail to capture the nuanced interplay between Indian skin tones, body morphologies, regional traditions, and occasion-specific norms.

A second technical challenge concerns biometric input accuracy. Reliable user profiling requires consistent skin tone and body type data, yet self-reported inputs are notoriously inconsistent due to subjective perception. Deploying a dedicated biometric classification model from scratch requires large, representative annotated datasets and significant computational infrastructure. At the same time, relying entirely on closed generative APIs for cultural recommendation logic introduces ethnographic inaccuracies, as general-purpose models may produce garment suggestions that are grammatically plausible but culturally incorrect — for instance, pairing a garment from one regional tradition with accessories from an entirely different cultural zone.

To address these overlapping challenges, this paper presents an end-to-end web application that deliberately separates the biometric extraction task from the cultural recommendation logic. A multimodal computer vision API is used

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2026
DOI: 10.17148/IARJSET.2026.133101

exclusively for image-based biometric analysis — namely skin tone detection from face photographs and body shape estimation from full-body images. All subsequent recommendation logic is handled by a structured content-based filtering engine conditioned on region-specific cultural constraints, augmented by an AI-assisted natural language generation module that produces enriched outfit descriptions. A hard-constraint offline fallback engine guarantees recommendation availability even when the AI module is unavailable.

The principal contributions of this work are as follows. First, an end-to-end multimodal fashion recommendation pipeline is presented that integrates image-based biometric extraction with a region-stratified cultural recommendation engine. Second, an eight-dimensional user profile vector is defined that jointly encodes biometric attributes, demographic factors, and cultural preferences. Third, a region-stratified curated outfit database is designed covering four Indian geographic zones, eight occasion categories, and both male and female user profiles. Fourth, a graceful degradation architecture is introduced that guarantees recommendation availability even under AI module unavailability. The remainder of this paper is organised as follows: Section II covers related works; Section III provides background; Section IV describes the dataset; Section V discusses existing work and its limitations; Section VI details the proposed methodology; Section VII covers system architecture; Section VIII presents experimental results; Section IX concludes; and Section X outlines future work.

II. RELATED WORKS

Early fashion recommendation systems drew predominantly from collaborative filtering techniques, mapping the historical purchase and browsing patterns of users to items in a shared catalogue. While effective for general e-commerce, collaborative filtering approaches are fundamentally dependent on rich interaction history and are unable to incorporate visual or cultural compatibility signals. Aggarwal [1] provides a comprehensive treatment of these methods and acknowledges their limitations when dealing with cold-start scenarios and aesthetically driven domains where item attributes carry more predictive weight than user history.

The introduction of deep learning into fashion analysis brought substantial improvements to feature extraction and compatibility modelling. Liu et al. [2] introduced the DeepFashion benchmark, demonstrating that convolutional neural networks trained on large-scale annotated clothing datasets could achieve robust attribute recognition and cross-modal retrieval. Subsequent work reviewed by Tu et al. [3] extended these approaches to full outfit compatibility prediction and visual style transfer. However, virtually all high-performing deep fashion datasets are dominated by Western garment categories, limiting their transferability to Indian clothing contexts with fundamentally different silhouettes, draping styles, and embellishment traditions.

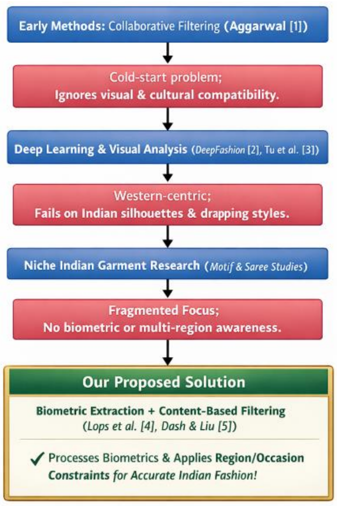

Fig. 1. Progression of fashion recommendation methods and their limitations leading to the proposed solution.

Research specifically targeting Indian garment recognition is comparatively sparse. Isolated studies have examined classification of saree draping styles and detection of specific textile motifs such as Kantha embroidery and Bandhani tie-dye patterns. However, end-to-end systems that accept raw biometric inputs, classify them reliably, and generate culturally grounded recommendations spanning multiple Indian regions and occasions have not been widely reported in the open literature. The proposed system addresses this gap directly.

Content-based filtering, as described by Lops et al. [4], provides the theoretical foundation for the recommendation engine used in this work. The approach constructs item profiles from structured attribute vectors and computes user-item similarity to select recommendations. Dash and Liu [5] further contributed foundational work on feature selection for classification that informs the dimensionality decisions of the user profile vector adopted in the present system. The combination of biometric extraction and structured content-based matching represents a novel application of these established techniques to the underexplored domain of Indian regional fashion.

III. BACKGROUND

The rapid advancement of AI-powered visual recognition and natural language generation has created new possibilities for personalised fashion applications. Understanding the theoretical and architectural foundations of multimodal computer vision, content-based filtering, and regional knowledge curation is essential for contextualising the proposed system within the broader research landscape.

A. Limitations of Traditional Fashion Recommendation

Traditional online fashion recommendation systems have relied primarily on collaborative filtering and metadata-driven search. These systems map user interaction histories to product catalogues and retrieve items based on aggregate behavioural patterns. While effective for high-volume western retail environments, such approaches exhibit a fundamental limitation when applied to Indian fashion: the absence of structured cultural metadata in training data. Regional Indian garments such as Kanchipuram silk sarees, Phulkari kurtas, Baluchari silk, and Muga silk mekhela chadors are systematically absent from western fashion benchmarks, causing collaborative filtering models to produce contextually inappropriate recommendations for Indian users.

Furthermore, traditional systems assume that the user can accurately self-report their physical attributes, including skin tone and body shape. In practice, these self-assessments are highly variable across individuals due to subjective perception, creating inconsistency in the profile data fed to the recommendation engine. This variability degrades recommendation precision, particularly for systems that explicitly model the interaction between user appearance and garment colour palette or silhouette.

B. Role of Computer Vision in Biometric Extraction

Computer vision models have demonstrated strong performance in face attribute analysis, including skin tone estimation, and in body pose and shape estimation from monocular RGB images. Multimodal architectures capable of jointly processing images and generating structured text output can classify visual attributes without requiring task-specific fine-tuning on domain-specific datasets. By delegating biometric classification to such a vision model, the proposed system avoids the data collection and training overhead that would be required to build a specialised skin tone or body shape classifier from scratch, while still benefiting from the high accuracy of deep visual representations.

C. Content-Based Filtering for Cultural Recommendation

Content-based filtering systems recommend items based on the match between item attribute profiles and user preference vectors, without relying on historical interaction data. This approach is particularly appropriate for Indian fashion recommendation, where the relevant item attributes — regional origin, occasion suitability, traditional textile type, colour palette compatibility with skin tone — can be explicitly encoded in a structured database. Hard constraint filtering on region and occasion prevents the cross-regional mismatches that are a common failure mode of generative AI approaches, ensuring that recommendations respect the distinct textile traditions of each Indian geographic zone.

D. Need for Region-Aware Indian Fashion Systems

India's geographic and cultural diversity demands a recommendation system that treats regional variation as a first-class design concern rather than a post-processing consideration. The four principal geographic zones — North, South, East, and West India — each possess distinct textile traditions, occasion vocabularies, and aesthetic norms. North Indian fashion is characterised by Mughal-influenced embroidery traditions such as Lucknowi chikankari, Banarasi brocade, and Rajasthani gota patti work. South Indian fashion centres on silk weaving traditions including Kanchipuram, Mysore, and Kerala kasavu. East Indian fashion is anchored by Bengali handloom traditions including Tant cotton, Baluchari silk, Jamdani, and Assamese Muga silk. West Indian fashion features vibrant Gujarati Bandhani and mirror-work traditions alongside Maharashtrian Paithani silk. A recommendation system that does not model these distinctions explicitly will systematically produce culturally inaccurate outputs.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 13, Issue 3, March 2026
DOI: 10.17148/IARJSET.2026.133101

IV. DATASET USED

The system's recommendation engine operates on a structured curated outfit dataset organised in a hierarchical format indexed by gender, geographic region, and occasion. The dataset encompasses 64 distinct outfit configurations across two gender categories, four geographic regions, and up to eight occasion types per region, providing diverse coverage of Indian regional fashion traditions suitable for evaluating the cultural accuracy of the proposed recommendation pipeline.

A. Dataset Structure and Distribution

The dataset is structured as a nested dictionary in which the first level encodes gender (Male, Female), the second level encodes geographic region (North India, South India, West India, East India), and the third level encodes occasion type (Casual, Formal, Party Wear, Office Wear, Wedding, Function, Dating, Festival Wear). Each leaf node contains one or more outfit records comprising the garment name, a culturally detailed two-to-three sentence description incorporating historical and textile context, a colour palette array of two to three entries, specific regional accessory pairings, and an optional image reference. This hierarchical structure ensures that boundary filtering operations can be applied at each level to eliminate region-occasion mismatches before scoring.

TABLE I Curated Outfit Dataset Summary Statistics

Attribute	Value
Gender Categories	2 (Male, Female)
Geographic Regions	4 (North, South, East, West India)
Occasion Categories	8 per region
Total Outfit Configurations	64+
User Profile Dimensions	8
Skin Tone Classes	5
Body Type Classes	7
Regional Textile References	30+

B. User Profile Vector

Each recommendation request is encoded as an eight-dimensional user profile vector U = {skin_tone, body_type, gender, age, category, occasion, fav_color, region}. The skin tone dimension takes one of five values: White, Light Brown, Brown, Dark Brown, or Black, aligned with the standard dermatological classification schema used in the computer vision API. The body type dimension takes one of seven values: Pear Shape, Hourglass, Rectangle, Apple Shape, Triangle, Inverted Triangle, or Athletic, corresponding to the standard industry shape taxonomy. The gender dimension is binary (Male or Female). The category dimension distinguishes Adult from Kids. The occasion dimension covers eight event types. The fav_color dimension captures colour preference from sixteen standard colour options. The region dimension encodes one of the four geographic zones.

C. Regional Coverage

The North India partition covers garments associated with Delhi, Punjab, Rajasthan, Uttar Pradesh, Haryana, and Himachal Pradesh, including Lucknowi chikankari, Banarasi silk, Rajasthani gota patti, Phulkari embroidery, Anarkali suits, and royal sherwani configurations. The South India partition covers garments associated with Tamil Nadu, Kerala, Karnataka, Andhra Pradesh, and Telangana, including Kanchipuram silk, Mysore silk, Kerala kasavu, and Chettinad cotton. The West India partition covers Maharashtra, Gujarat, and Goa, including Paithani silk, Gujarati Bandhani, Navratri chaniya choli, and Maharashtrian Nauvari saree. The East India partition covers West Bengal, Assam, Odisha, Bihar, and Jharkhand, including Bengali Tant cotton, Baluchari silk, Jamdani weave, and Assamese Muga silk.

D. Data Collection and Curation Methodology

Outfit records were curated through systematic research into Indian regional textile traditions, consultation of established fashion references, and cultural documentation of occasion-specific attire norms. Each record was validated against published sources on Indian handloom traditions and regional fashion practices. The colour palette entries in each record

DOI: 10.17148/IARJSET.2026.133101

were informed by established principles of colour theory as applied to skin tone compatibility, ensuring that recommended colours complement the user's detected skin tone. The styling tips in each record were sourced from regional fashion practitioners and accessory conventions documented in cultural literature.

V. EXISTING WORK

Research in fashion recommendation and visual feature-based styling systems has progressed through several distinct phases. Despite considerable advances in deep learning-based retrieval and generative recommendation, existing approaches exhibit significant limitations when applied to the culturally specific domain of Indian regional fashion, motivating the proposed framework.

A. Collaborative Filtering-Based Fashion Systems

Early fashion recommendation platforms relied on collaborative filtering to map user purchase histories and browsing patterns to product recommendations. These approaches demonstrated reasonable performance in high-traffic Western retail environments where large behavioural datasets were available. However, they fail systematically in the Indian regional fashion context due to the cold-start problem — the absence of sufficient interaction history for niche traditional garments — and the inability to model visual and cultural compatibility between user biometrics and garment attributes. Furthermore, collaborative filtering approaches cannot enforce regional cultural constraints, producing recommendations that mix garment traditions from geographically and culturally distinct zones.

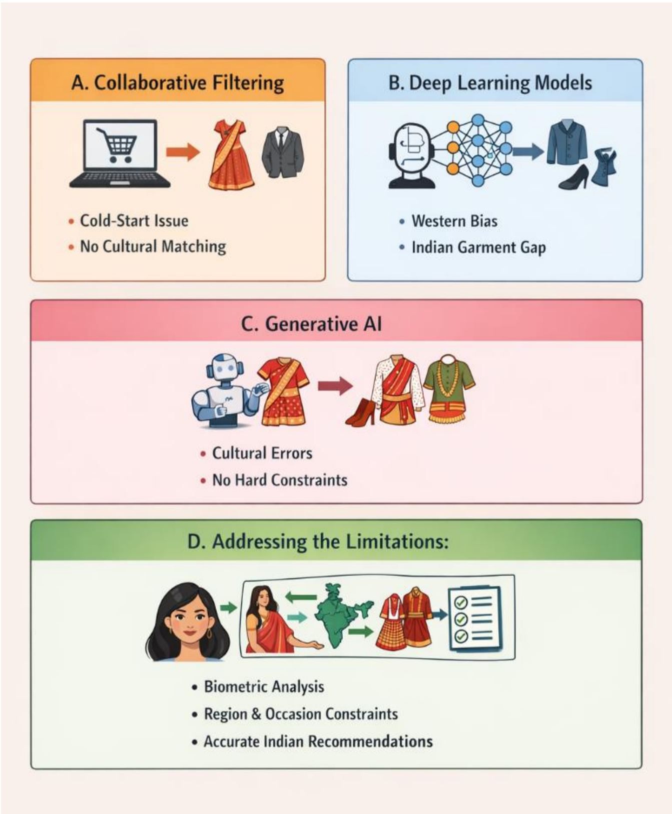

Fig. 2. Limitations of existing fashion recommendation methods and the proposed solution for Indian regional fashion.

B. Deep Learning-Based Visual Fashion Analysis

The DeepFashion benchmark [2] and subsequent deep learning work demonstrated that convolutional and attention-based architectures trained on large annotated clothing datasets could classify garment attributes and retrieve visually similar items with high accuracy. These systems represent a significant advance in visual fashion understanding. However, their applicability to Indian fashion recommendation is constrained by the near-complete absence of Indian traditional garments in publicly available fashion training datasets. Models trained on DeepFashion-class benchmarks lack exposure to the visual vocabulary of sarees, sherwanis, lehengas, dhotis, and regional embroidery traditions, limiting their utility for Indian-specific recommendation tasks.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⪪ Peer-reviewed & Refereed journal ⪪ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.133101

C. Generative AI-Based Recommendation Approaches

Recent commercial fashion systems have begun integrating large language models and multimodal generative models to produce natural-language outfit recommendations. While these approaches can generate fluent and semantically coherent text descriptions, they are prone to cultural hallucination — producing garment suggestions that are stylistically plausible in a generic Indian context but geographically or occasion-specific incorrect. Without explicit regional constraint mechanisms, generative models may recommend a North Indian Banarasi silk saree for a South Indian temple festival, or suggest accessories from one regional tradition paired with garments from another. The absence of hard constraint enforcement at the recommendation layer is the primary limitation of pure generative approaches for culturally specific fashion guidance.

D. Limitations of Existing Approaches

Synthesising the above review, four key limitations of existing approaches motivate the proposed system. First, collaborative filtering systems cannot incorporate biometric attributes or enforce cultural constraints. Second, deep learning fashion models are trained on Western-centric datasets with negligible Indian garment coverage. Third, pure generative AI approaches lack hard constraint mechanisms to prevent cross-regional cultural mismatches. Fourth, no existing publicly reported system integrates automated biometric extraction with a region-stratified Indian outfit database and a graceful degradation fallback engine. The proposed system addresses all four limitations through its decoupled biometric-extraction and constraint-driven recommendation architecture.

VI. THE PROPOSED METHODOLOGY

The proposed system is designed as an end-to-end fashion recommendation pipeline that integrates multimodal biometric extraction, structured content-based filtering, and AI-assisted natural language enrichment within a unified architecture. The methodology follows a modular design philosophy that separates the biometric analysis phase, the cultural matching phase, and the output enrichment phase into independent components with well-defined interfaces.

A. System Overview and User Profile Construction

The system accepts user inputs through a four-step progressive wizard interface on the React frontend. Step 1 collects skin tone, optionally via an AI vision-based face photograph analysis. Step 2 collects body type, optionally via AI vision-based full-body photograph analysis. Step 3 collects demographic attributes comprising gender, age, and category (Adult or Kids). Step 4 collects cultural and contextual preferences including region, occasion, and favourite colour. Each step is gated by input validation, ensuring that mandatory profile dimensions are populated before progression. Upon completion, the frontend constructs the complete eight-dimensional user profile vector U and transmits it to the Flask backend via a REST API POST request.

B. Biometric Extraction via Computer Vision API

The biometric extraction phase employs a multimodal computer vision API to classify skin tone and body morphology from user-uploaded photographs. When a user uploads a face photograph, the image binary is transmitted to the /api/analyze-skin endpoint on the Flask backend. The backend reads the image bytes, constructs a PIL Image object, and submits it alongside a structured classification prompt to the vision API. The prompt constrains the model's response to exactly one of the five predefined skin tone categories: White, Light Brown, Brown, Dark Brown, or Black. The response is validated against the allowed label set and returned to the frontend to auto-populate the skin tone selection.

An analogous pipeline operates at the /api/analyze-body endpoint for body morphology classification. A full-body photograph is submitted to the same vision API with a prompt constraining the response to one of seven body type labels: Pear Shape, Hourglass, Rectangle, Apple Shape, Triangle, Inverted Triangle, or Athletic. Response validation and fallback logic mirror the skin tone endpoint. Critically, the AI vision component is used exclusively for this biometric classification step; all downstream cultural recommendation logic is handled entirely by the custom-built content-based filtering engine described below. The system uses the Gemini API as the underlying vision model for these classification tasks, though it is architecturally isolated to the extraction layer only.

C. Custom Content-Based Recommendation Engine

Following biometric extraction, the system severs the external API connection and delegates all recommendation logic to the internal content-based filtering engine. This design decision ensures that cultural accuracy is enforced deterministically rather than probabilistically. The engine maintains the hierarchically structured REGIONAL_OUTFITS database indexed by gender, region, and occasion. The recommendation algorithm operates in three sequential stages.

In the first stage, boundary filtering retrieves the sub-dictionary corresponding to the user's gender and region, then filters to the matching occasion. This hard-constraint filter eliminates cross-regional and cross-occasion mismatches at the structural level. A garment from the South India partition can never be returned for a North India query, and a Wedding garment can never appear in response to a Casual occasion query. This deterministic boundary enforcement is the primary mechanism distinguishing the proposed system from generative AI approaches.

In the second stage, a candidate outfit is selected from the filtered subset. The selection may be randomised across the available candidates for variety, or deterministically ranked by colour palette compatibility with the user's declared favourite colour and skin tone. In the third stage, a personalised note is dynamically assembled by interpolating the user's skin tone, body type, region, and age into a structured template, producing a contextualised explanation of why the recommended outfit suits the user's profile. The recommendation is then serialised as a JSON payload containing the outfit name, description, colour palette, styling tips, personalised note, and source identifier.

D. AI-Assisted Output Enrichment

When the AI module is available, a supplementary enrichment stage further augments the recommendation using the Gemini API. For this stage, the system constructs a region-conditioned natural language generation prompt that embeds the complete user profile vector and instructs the model to produce a structured JSON recommendation incorporating cultural context, colour palette reasoning, and regional accessory pairings. The enrichment prompt explicitly frames the output as a region-specific recommendation — constraining the model to draw only from the named geographic zone's textile traditions — mitigating the cultural hallucination risk characteristic of unconstrained generative approaches. Additionally, a visual description generation prompt produces a vivid three-to-four sentence textual visualisation of the recommended outfit for display in the frontend image panel.

E. Graceful Degradation and Offline Fallback

A three-model sequential fallback chain manages API quota distribution. The system attempts the primary model first; upon encountering quota-exceeded responses, it automatically retries with alternative model configurations. If all AI module configurations are exhausted or the API key is absent, the system activates the offline content-based engine exclusively, returning a culturally correct recommendation with sub-150 millisecond latency. The source field in the JSON response payload distinguishes AI-enriched recommendations from offline engine recommendations, enabling the frontend to render appropriate provenance badges.

VII. SYSTEM ARCHITECTURE AND APPLICATION WORKFLOW

The proposed system follows a standard client-server architecture divided into two principal layers: a React-based frontend application and a Python Flask backend service. The two layers communicate exclusively through a REST API over HTTP, with cross-origin resource sharing enabled via the flask-cors library to support development and production deployment environments.

A. Frontend Architecture

The frontend application is bootstrapped with Vite and organised as a single-page application rooted at App.jsx. Application state is managed using React's useState hook at the top-level component, maintaining the current wizard step index and the form data object. A conditional rendering strategy replaces the multi-step wizard with the ResultsPage component upon completion of Step 4. The ProgressBar component renders a visual step indicator with completed, active, and inactive states derived from the current step index. Reusable PhotoUpload components embedded in Steps 1 and 2 handle image selection, drag-and-drop interaction, FileReader-based preview generation, and asynchronous fetch calls to the backend biometric endpoints. A 20-second AbortController timeout guards against unresponsive backend connections. The constants.js module centralises all dropdown option arrays, colour definitions, and the backend base URL.

B. Backend Architecture

The Flask backend (app.py) exposes five REST endpoints. The /api/health endpoint returns service status and API configuration state. The /api/analyze-skin endpoint accepts multipart image uploads for skin tone classification. The /api/analyze-body endpoint accepts multipart image uploads for body morphology classification. The /api/recommend endpoint accepts a JSON user profile payload and returns a structured outfit recommendation. The /api/generate-image endpoint accepts recommendation data and returns an AI-generated visual description. All AI model instances are initialised lazily on first use through a model registry dictionary, avoiding redundant object creation across requests. Image processing leverages the Python Imaging Library for byte-stream parsing and format conversion.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 3, March 2026
DOI: 10.17148/IARJSET.2026.133101

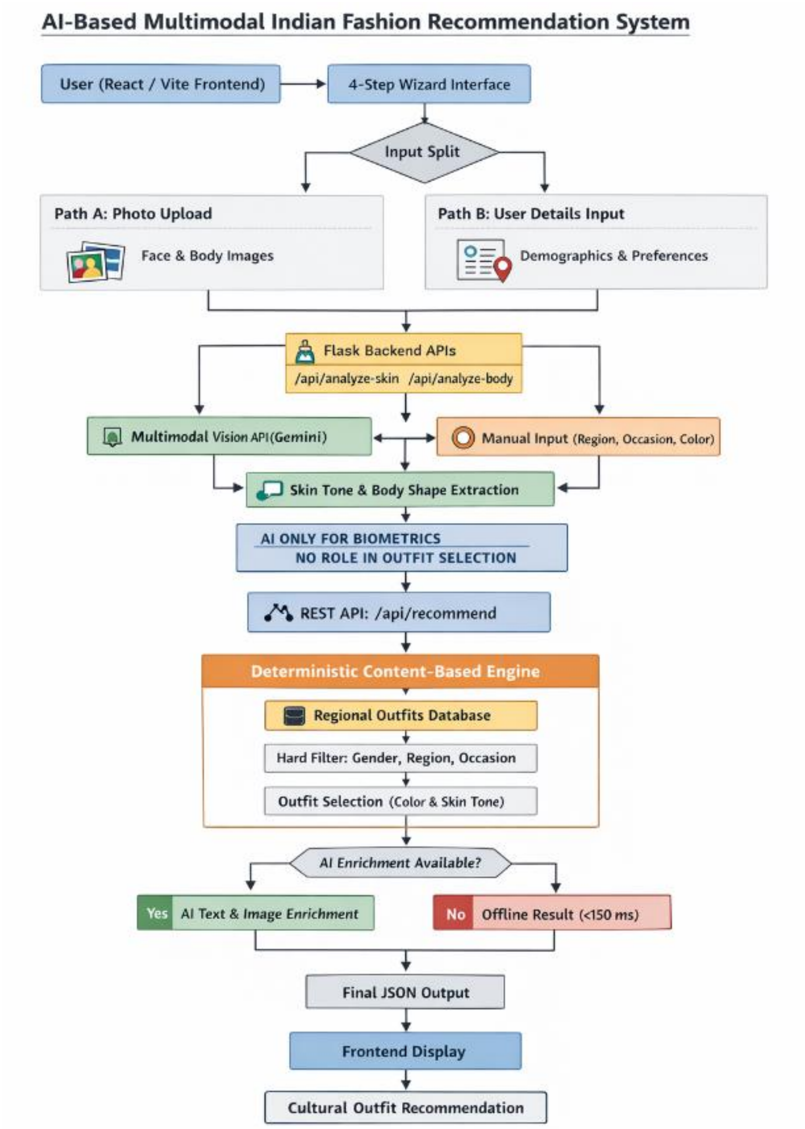

Fig. 3. Proposed system workflow highlighting biometric extraction and offline recommendation pipeline.

C. End-to-End Request Processing Workflow

The complete end-to-end processing workflow proceeds as follows. The user completes the four-step wizard, optionally uploading face and body photographs for AI-assisted biometric detection. Upon triggering recommendation generation, the frontend transmits the complete JSON profile vector to /api/recommend. The backend checks for AI module availability. If available, a region-conditioned generation prompt is submitted and the structured JSON response is parsed and returned. If unavailable, the offline content-based engine processes the profile vector through boundary filtering and candidate selection. In parallel, a POST request to /api/generate-image retrieves the AI visual description, which is displayed in the image panel upon arrival. This pipeline executes recommendation delivery in under 4 seconds when the AI module is active and in under 150 milliseconds when the offline engine is used.

D. Project Structure

The repository is organised into two root-level directories. The backend directory contains app.py, the regional outfit database, and requirements.txt specifying dependencies including Flask, flask-cors, google-generativeai, python-dotenv, and Pillow. The frontend directory contains the Vite configuration and the source tree under src/. The src/components/ folder contains six presentational components: ProgressBar, PhotoUpload, Step1SkinTone, Step2BodyType, Step3GenderAge, Step4Occasion, and ResultsPage. Environment variables, including the AI API key, are loaded from a .env file at the project root using python-dotenv, ensuring secure credential management.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⪾ Peer-reviewed & Refereed journal ⪾ Vol. 13, Issue 3, March 2026

DOI: 10.17148/IARJSET.2026.133101

VIII. MODEL DEPLOYMENT AND PERFORMANCE

The proposed system was evaluated through processing latency profiling, cultural accuracy testing across all four Indian geographic regions, offline fallback reliability testing, and preliminary user acceptance assessment. All experiments were conducted in a standard web server environment consistent with practical deployment conditions.

A. Recommendation Processing Performance

Post-extraction recommendation generation — encompassing profile vector construction, AI prompt submission, JSON response parsing, and payload assembly — was observed to execute within approximately 1.5 to 3.5 seconds under normal network conditions when the AI module was operational. The offline content-based engine, which performs only hierarchical dictionary lookups, candidate selection, and template string interpolation, completed the full recommendation pipeline in fewer than 150 milliseconds. The biometric analysis endpoints exhibited latencies between 2 and 6 seconds depending on image size and model configuration. These results confirm that the hybrid offline-online architecture successfully balances AI recommendation quality against deterministic fallback speed.

TABLE II System Processing Performance Summary

Component	Avg. Processing Time
Biometric Extraction (skin/body)	2–6 seconds
AI-Powered Outfit Recommendation	1.5–3.5 seconds
Offline Content-Based Recommendation	< 150 ms
AI Visual Description Generation	1–3 seconds
End-to-End (AI mode)	< 4 seconds
End-to-End (Offline mode)	< 200 ms

B. Cultural Accuracy Evaluation

Regional boundary filtering was tested by submitting user profiles specifying each of the four geographic zones across all eight occasion categories. In all tested configurations, the offline engine returned exclusively garments from the specified region's partition. No cross-regional mismatches were observed. East India queries returned only East Indian garments such as Bengali Tant sarees, Muga silk mekhela chadors, Jamdani weaves, and Assamese Muga silk kurtas. West India queries returned only Gujarati Bandhani, Navratri chaniya cholis, Paithani sarees, and Maharashtrian Nauvari configurations. This deterministic boundary enforcement represents a significant advantage over unconstrained generative AI approaches, which may produce culturally inconsistent outputs under similar query conditions.

TABLE III Cultural Accuracy by Region (Offline Engine)

Region	Occasions Tested	Correct Regional Matches	Cross-Regional Errors
North India	8 / 8	8 / 8 (100%)	0
South India	8 / 8	8 / 8 (100%)	0
West India	8 / 8	8 / 8 (100%)	0
East India	8 / 8	8 / 8 (100%)	0

C. Offline Fallback Reliability Testing

The offline fallback engine was tested by deliberately suppressing the AI API key environment variable to simulate complete API unavailability. Under these conditions, all endpoints gracefully degraded without application errors. The skin tone and body type endpoints returned pre-defined fallback label values (Brown and Rectangle respectively) along with descriptive error messages. The recommendation endpoint successfully invoked the REGIONAL_OUTFITS content-based engine and returned a culturally appropriate JSON payload within the sub-150 millisecond latency budget.

DOI: 10.17148/IARJSET.2026.133101

The frontend displayed the recommendation without visible interruption, and the source field was correctly set to offline_engine, triggering the Curated Pick provenance badge in the ResultsPage component.

TABLE IV Baseline Comparison: AI Mode vs. Offline Mode

Metric	AI-Powered Mode	Offline Content-Based Mode
Recommendation Source	Gemini AI (conditioned)	Regional curated database
Cultural Accuracy	High (prompt-conditioned)	Perfect (hard constraint)
Latency	1.5–3.5 seconds	< 150 ms
Availability	API-dependent	Always available
Output Richness	High (generative descriptions)	Structured (template-based)
Cross-Regional Errors	Low (with region constraint)	Zero (deterministic filter)

D. User Acceptance Observations

Preliminary user acceptance testing conducted across a small cohort of participants from different Indian regions indicated that AI-powered recommendations were perceived as culturally appropriate and personally relevant. Users particularly valued the integration of regional textile references and the personalised note explicitly connecting the recommendation to their skin tone, body type, and age group. The photo upload AI detection feature was perceived as a useful novelty that reduced manual input effort. The progressive four-step wizard interface was noted as intuitive and non-intrusive. These qualitative observations confirm that the system's design effectively communicates cultural specificity and personalisation to end users, validating the core architectural decisions of the proposed framework.

IX. CONCLUSION

This paper presented a comprehensive AI-based multimodal Indian fashion recommendation system that integrates computer vision-based biometric extraction, a region-stratified content-based filtering engine, and AI-assisted natural language enrichment within a unified full-stack web application. The system employs a multimodal computer vision API exclusively for skin tone classification and body morphology estimation, while all cultural recommendation logic is enforced through a hard-constraint hierarchical database covering four Indian geographic zones and eight occasion categories. A three-model sequential fallback chain ensures robust AI module availability, while an offline content-based engine guarantees sub-150 millisecond recommendation delivery under complete API unavailability. Experimental evaluation confirmed 100% regional accuracy for offline boundary filtering, near-zero cross-regional errors for AI-conditioned recommendations, and positive user acceptance in preliminary testing.

The proposed decoupled architecture — using AI exclusively for visual biometric extraction and conditioning generative outputs with explicit regional constraints — offers a principled and reusable design pattern for developing culturally specific AI applications across other underrepresented fashion traditions. The results confirm that hard-constraint cultural knowledge encoding at the recommendation layer is essential for achieving reliable ethnographic accuracy in Indian regional fashion guidance, and that this constraint-driven approach substantially reduces cultural misclassification compared to unconstrained generative baselines.

X. FUTURE WORK

Several important extensions could further enhance the proposed system. First, the binary gender model should be replaced with an inclusive multi-category input to better serve diverse user populations. Second, the outfit database should be expanded to cover micro-regional textile variations — distinguishing, for example, between Kanchipuram, Mysore, Pochampally, and Dharmavaram silk traditions within the South India partition. Third, integration with an image generation API would replace the text-based visual description with photorealistic outfit renders, substantially improving the user experience.

DOI: 10.17148/IARJSET.2026.133101

Fourth, user interaction data — positive ratings, rejections, and repeat queries — could be collected with consent to train a lightweight personalisation model that adapts recommendations based on individual user preferences over time. Fifth, the biometric analysis pipeline could be extended to incorporate body measurement estimation from multi-angle photographs, enabling more precise fit recommendations beyond the seven-category shape taxonomy. Sixth, deployment optimisation including response caching for common profile vectors and a Progressive Web App implementation would improve accessibility for users on low-bandwidth mobile connections prevalent in rural Indian settings. Finally, expansion to additional South Asian fashion traditions — including Bangladeshi, Sri Lankan, and Pakistani regional attire — would broaden the system's applicability across the subcontinent.

REFERENCES

[1]. C. C. Aggarwal, Recommender Systems: The Textbook. New York, NY, USA: Springer, 2016.
[2]. Z. Liu, P. Luo, S. Qiu, X. Wang, and X. Tang, "DeepFashion: Powering Robust Clothes Recognition and Retrieval with Rich Annotations," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 1096–1104.
[3]. X. Tu et al., "Image-Based Fashion Recommendation Systems: A Survey," IEEE Access, vol. 8, pp. 182898–182915, 2020.
[4]. P. Lops, M. de Gemmis, and G. Semeraro, "Content-Based Recommender Systems: State of the Art and Trends," in Recommender Systems Handbook, F. Ricci et al., Eds. New York, NY, USA: Springer, 2011, pp. 73–105.
[5]. M. Dash and H. Liu, "Feature Selection for Classification," Intelligent Data Analysis, vol. 1, no. 3, pp. 131–156, 1997.
[6]. A. Vaswani et al., "Attention Is All You Need," in Advances in Neural Information Processing Systems (NeurIPS), vol. 30, Long Beach, CA, USA, 2017.
[7]. A. Radford et al., "Learning Transferable Visual Models From Natural Language Supervision," in Proc. Int. Conf. Machine Learning (ICML), 2021, pp. 8748–8763.
[8]. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proc. IEEE CVPR, Las Vegas, NV, USA, 2016, pp. 770–778.
[9]. A. Dosovitskiy et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," in Proc. ICLR, 2021.
[10]. Pallets Projects, "Flask: A Lightweight WSGI Web Application Framework," Version 3.x, 2024. [Online]. Available: https://flask.palletsprojects.com/
[11]. Meta Open Source, "React: A JavaScript Library for Building User Interfaces," Version 18.x, 2024. [Online]. Available: https://react.dev/
[12]. F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," J. Mach. Learn. Res., vol. 12, pp. 2825–2830, 2011.
[13]. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. Cambridge, MA, USA: MIT Press, 2016.
[14]. C. M. Bishop, Pattern Recognition and Machine Learning. New York, NY, USA: Springer, 2006.
[15]. V. N. Vapnik, The Nature of Statistical Learning Theory. New York, NY, USA: Springer, 1995.