HELWAN UNIVERSITY

Faculty of Computers and Artificial Intelligence

Computer Science Department

E-PROVA

A graduation project dissertation by:

[Marwan Shamel Mohammed (20210895)]

[Marwan Abdelrady Moghazi (20210896)]

[Marwan Magdy Mohammed (20210898)]

[Mariam Elsayed Fattouh (20210908)]

[Mariam Samy Farouk (20210913)]

[Mariam Fathi Ahmed (20210921)]

Submitted in partial fulfilment of the requirements for the degree of Bachelor of Science in Computers & Artificial Intelligence, at the Computer Science Department, the Faculty of Computers & Artificial Intelligence, Helwan University

Supervised by:

[Dr. Ahmed Hesham]

(June 2025)

Abstract

E-Prova is a smart and interactive e-commerce platform that allows users to try on clothes virtually before buying them online. The idea behind the project is to make online shopping more personal and realistic, especially for people who often find it hard to imagine how clothes will look on them just from pictures.

Through the platform, users can upload their own photo and see how different outfits would appear on them as if they were standing in front of a mirror. This feature helps users make better decisions and feel more confident about their purchases. The system depends on artificial intelligence to handle and adjust the clothes on the user's image in a natural and smooth way.

In addition to the virtual try-on feature, E-Prova includes all the basic functions of an online store. Users can browse products, add items to their cart, place orders, leave reviews, and even benefit from special offers and discounts. Admins can also manage products, brands, and track orders through a dedicated panel.

E-Prova was built to improve the online shopping experience by making it more fun, easy, and suitable for different user needs. It aims to bring the future of fashion retail closer to reality.

Acknowledgement

We would like to express our heartfelt thanks to our supervisor, [Dr. Ahmed Hesham], for his continuous guidance, support, and encouragement throughout the journey of developing the E-Prova platform. His advice and feedback helped us stay focused and motivated during every stage of the project.

We also want to thank each team member for their hard work, creativity, and teamwork. Every part of this project was built with shared effort and dedication.

Finally, we would like to thank everyone who believed in our idea and inspired us to think creatively. This project was not just about building a platform—it was about creating something meaningful and helpful, and we couldn't have done it without the support of those around us.

Thank you all for being part of our journey.

Table of Contents

Chapter 1

[Introduction]

presents a brief introduction to the E-Prova platform, a virtual try-on e-commerce system that helps users visualize how clothes would look on them before purchasing. By allowing users to try outfits on their own images, the platform reduces doubts and improves decision-making during online shopping.

The platform also includes essential e-commerce features such as product browsing, order placement, reviews, and promotional offers. It aims to solve the common issue of poor fit and unexpected product appearance, which often leads to returns.

The scope of the project focuses on making fashion shopping more interactive, reliable, and user-friendly—especially for customers who shop online and want a more realistic experience.

1.1 Overview

E-Prova is an innovative e-commerce application that aims to improve the online fashion shopping experience by allowing users to virtually try on clothes before making a purchase. The platform is designed to help users make smarter and more confident buying decisions by showing how each clothing item would look on their body in a realistic and personalized way.

The virtual try-on feature uses advanced technology to process the user's image and adapt the selected garment to match their body shape and posture. This allows shoppers to get a better sense of the size, fit, and appearance of clothes without physically wearing them. This feature is especially helpful in reducing the number of incorrect purchases and returns that often happen with online shopping.

Alongside this, E-Prova offers a full shopping system, including browsing and filtering products, managing brands and categories, placing orders, applying promotional offers and discounts, and submitting product reviews. Users can also create their own accounts to track orders and save favorite items.

Many online shoppers struggle with the inability to try on clothes before buying, which leads to hesitation, disappointment, and wasted time or money. E-Prova addresses this problem by creating a more interactive, visual, and user-friendly shopping experience from home.

Overall, E-Prova provides a smart and accessible solution that helps users enjoy online fashion shopping with more confidence, accuracy, and satisfaction.

1.2 Problem Statement

Online shopping for clothes has grown rapidly in recent years, offering convenience and variety for users. However, a major drawback still remains—customers cannot try on clothes before purchasing. This leads to uncertainty about how the clothes will look and fit on their own body, often resulting in dissatisfaction, returns, and wasted time and money.

According to industry reports, one of the most common reasons for product returns in online fashion retail is poor fit or the item looking different from expectations. Many customers hesitate to complete purchases due to doubts about size, appearance, or compatibility with their personal style. This affects customer satisfaction and increases operational costs for e-commerce platforms.

Additionally, most online clothing stores lack interactive tools to personalize the shopping experience or allow users to visualize products on themselves. Customers have to rely on static images, size charts, or model photos that may not reflect their own body shape or style preferences.

This experience is particularly challenging for users with limited fashion knowledge or those shopping online for the first time. It can also be frustrating for people who live in areas with limited access to physical stores and rely heavily on e-commerce.

E-Prova aims to solve these problems by introducing a smart virtual try-on feature that allows users to see how clothes would look on their own photos in a realistic and interactive way. The system also provides all core e-commerce functionalities and aims to make online fashion shopping more personalized, confident, and satisfying for every user.

1.3 Scope and Objectives

The E-Prova system aims to transform the online fashion shopping experience by introducing a virtual try-on feature that allows users to visualize clothing on their own images before making a purchase. This feature provides a more interactive, realistic, and personalized shopping journey, helping users make confident buying decisions and minimizing returns caused by poor fit or misleading appearances.

By leveraging artificial intelligence, E-Prova accurately adjusts clothing items to match the user's body shape and posture, delivering a natural and convincing representation of how the clothes would look in real life.

Beyond the try-on experience, E-Prova serves as a comprehensive fashion platform. It allows users to explore curated collections, make purchases, track orders, and engage with product reviews and offers in a seamless and intuitive environment. The platform also includes an admin interface designed to streamline backend operations such as managing inventory, monitoring orders, and handling product listings—ensuring efficiency and scalability.

Objectives of E-Prova:

1. To reduce the number of product returns and enhance customer confidence by offering a more accurate and realistic preview of clothing items.

2. To reduce the number of product returns and enhance customer confidence by offering a more accurate and realistic preview of clothing items.

3. To deliver a user-friendly and feature-rich shopping environment, with functionalities like product filtering, reviewing, and wish-listing.

4. To enable efficient management of the store through an admin dashboard that supports real-time updates for products, brands, and order status.

5. To make fashion e-commerce more inclusive, enjoyable, and tailored to individual preferences and body types.

Overall, E-Prova envisions a future where digital fashion shopping is as personal and reliable as visiting a physical store—bridging the gap between imagination and reality through smart, interactive technology.

.

Chapter 2

[Literature Review]

This chapter presents a comprehensive analysis of fashion retailers utilizing artificial intelligence features in their e-commerce platforms. We examine how current industry leaders implement AI technologies to enhance the online shopping experience, from personalized recommendations to virtual fitting solutions. Through detailed evaluation of each platform's advantages and limitations, we identify critical gaps in the current market landscape. The chapter further explores the technological foundations underlying these systems, including computer vision techniques, machine learning algorithms, and secure payment integration methods.

2.1 Similar Applications

2.1.1 Stitch Fix - AI Styling Algorithm

STITCH FIX

Pros :

- Hybrid AI-human approach combining algorithms with personal stylists
- Deep learning models analyze style preferences from detailed questionnaires

Cons :

- Subscription-based model with $20 styling fee
- No immediate virtual visualization of items

2.1.2 YouCam Online Editor - AI Clothes Changer

YouCam

Pros :

- Real-time AI clothing overlay on user photos
- Supports multiple clothing categories (tops, dresses, outerwear)

Cons :

- Requires high-quality full-body photo for best results
- Watermark on free version outputs

2.1.3 AIEASE - AI Virtual Try-On

Pros :

- Supports both photo upload and live camera modes
- Preserves original photo quality and lighting

Cons :

- Processing time can be slow for high-resolution images
- Requires technical integration for e-commerce platforms

2.2 Background

2.2.1 E-Commerce Evolution

The evolution of e-commerce has transformed from simple online catalogs in the 1990s to sophisticated platforms offering personalized shopping experiences. The integration of secure payment gateways like PayPal (1998) and Stripe (2010) revolutionized online transactions, making them more accessible and trustworthy. As smartphone adoption increased, responsive design and mobile-first approaches became essential, with mobile commerce now accounting for over 70% of e-commerce traffic.

2.2.2 Virtual Try-On Technology

Virtual try-on technology has evolved through several technological breakthroughs. Computer vision techniques now enable accurate body detection

and pose estimation for realistic clothing overlay. The introduction of ARCore and ARKit frameworks democratized augmented reality development, making it accessible for retail applications. Deep learning models trained on extensive fashion datasets can now predict fit and styling preferences with remarkable accuracy.

2.3 Literature Survey

2.3.1 Virtual Try-On Systems Using Deep Learning

"VITON: An Image-based Virtual Try-on Network" (Han et al., 2018) introduced a groundbreaking coarse-to-fine strategy for image-based virtual try-on. The research achieved realistic results through multi-stage generation, setting new benchmarks for clothing visualization accuracy.

"CP-VTON: Toward Characteristic-Preserving Virtual Try-On" (Wang et al., 2018) advanced the field by proposing a geometric matching module that aligns clothing with body poses while preserving garment characteristics and patterns.

2.3.2 E-Commerce User Experience and Authentication

"Factors Affecting E-Commerce Adoption in Fashion Retail" (Sharma et al., 2019) identified trust, ease of use, and security as primary adoption factors, with visual product representation being crucial for fashion items.

"Multi-Factor Authentication in E-Commerce: Balancing Security and Usability" (Chen et al., 2020) demonstrated that properly implemented security measures actually improve user retention when they don't impede the shopping experience.

2.3.3 Stripe Payment Integration in Web Applications

"Modern Payment Gateway Integration: A Comparative Study" (Anderson et al., 2021) analyzed various payment solutions, showing Stripe's API reduces integration time by 40% compared to traditional gateways.

"Secure Payment Processing in React Applications" (Roberts et al., 2020) outlined best practices for implementing PCI-compliant payment systems in modern web applications.

2.3.4 React.js and Node.js in E-Commerce Development

"Building Scalable E-Commerce Platforms with MERN Stack" (Thompson et al., 2021) demonstrated that React.js with Node.js backend can handle 10,000+ concurrent users with proper optimization.
"Performance Optimization in React-based E-Commerce" (Kumar et al., 2022) identified key optimization strategies including lazy loading, code splitting, and server-side rendering for improved user experience.

2.4 Main Differences

Our platform distinguishes itself from existing solutions through several key innovations:

1. **Integrated Virtual Try-On with E-Commerce**: Unlike standalone virtual try-on tools or traditional e-commerce sites, our platform seamlessly

combines both functionalities in a single, cohesive system.

2. **No Subscription Model**: We eliminate the subscription fees and styling charges found in services like Stitch Fix, offering direct access to all features.

3. **Real-Time Processing**: Our system provides instant virtual try-on without the processing delays experienced in platforms like AIEASE.

4. **Privacy-First Approach**: We don't permanently store user images, addressing privacy concerns prevalent in photo-based try-on services.

5. **Simplified User Experience**: By intentionally excluding complex features like order tracking and focusing on core shopping functionality, we provide a streamlined experience that reduces cognitive load.

This combination of features positions our platform as a next-generation solution that bridges the gap between virtual try-on technology and practical e-commerce functionality, offering users an unprecedented shopping experience

.

Chapter 3

[System Design and Requirements]

This chapter presents the overall system design and the detailed requirements of the proposed solution. It aims to provide a comprehensive understanding of how the E-Prova system operates by breaking down its architecture, components, and user interactions. The chapter includes all relevant system diagrams that illustrate the structure and flow of the platform, such as use case diagrams, activity diagrams, and system architecture. Furthermore, this chapter outlines both the functional and non-functional requirements that guided the development process. By clearly defining what the system should do and the standards it must meet, we ensure that the platform delivers a seamless, secure, and efficient user experience for both administrators and end users. The purpose of this section is to offer a clear and structured overview of the system's logic and behavior before diving into the implementation of AI-related components in the next chapter

3.1 Functional Requirements

3.1.1 For Customer/User

-The system shall allow the user to register using email and password.
-The system shall allow the user to log in securely.
-The user shall be able to browse products, categories, and brands without logging in.
-The user shall be able to add products to cart.
-The user shall be able to update or remove items from the cart.
-The user shall be able to place orders from the cart.
-The user shall be able to choose a payment method (e.g., cash on delivery or card).
-The user shall be able to view previous orders and their statuses.
-The user shall be able to submit reviews on purchased products.
-The user shall be able to use the virtual try-on feature to preview clothes on their uploaded photo or live camera

---

3.1.2 For Admin

-The system shall allow the admin to log in securely using admin credentials.
-The admin shall be able to manage products ,categories, brands (add, update, delete).
-The admin shall be able to view all orders and update their status..
-The admin shall be able to view all users' carts.
-The admin shall be able to manage product reviews (view, delete).

3.2 Non-Functional Requirements

3.2.1 Implementation Requirements

In implementing the E-Prova system, the frontend will be developed using **Angular** for a responsive and dynamic user interface. The backend will use **Node.js** with **Express** as the server-side framework and **MongoDB** as the database management system. For AI functionalities, such as the virtual try-on feature.

3.2.2 Efficiency Requirements

The system must be optimized to provide fast loading times, especially while rendering the virtual try-on preview. AI predictions (i.e., clothing overlay) should not exceed 30 seconds. The e-commerce workflow (browsing, adding to cart, checkout) should be seamless and require minimal clicks. The system should support scalability to handle increasing numbers of users and product items without affecting performance.

3.2.3 Reliability Requirements

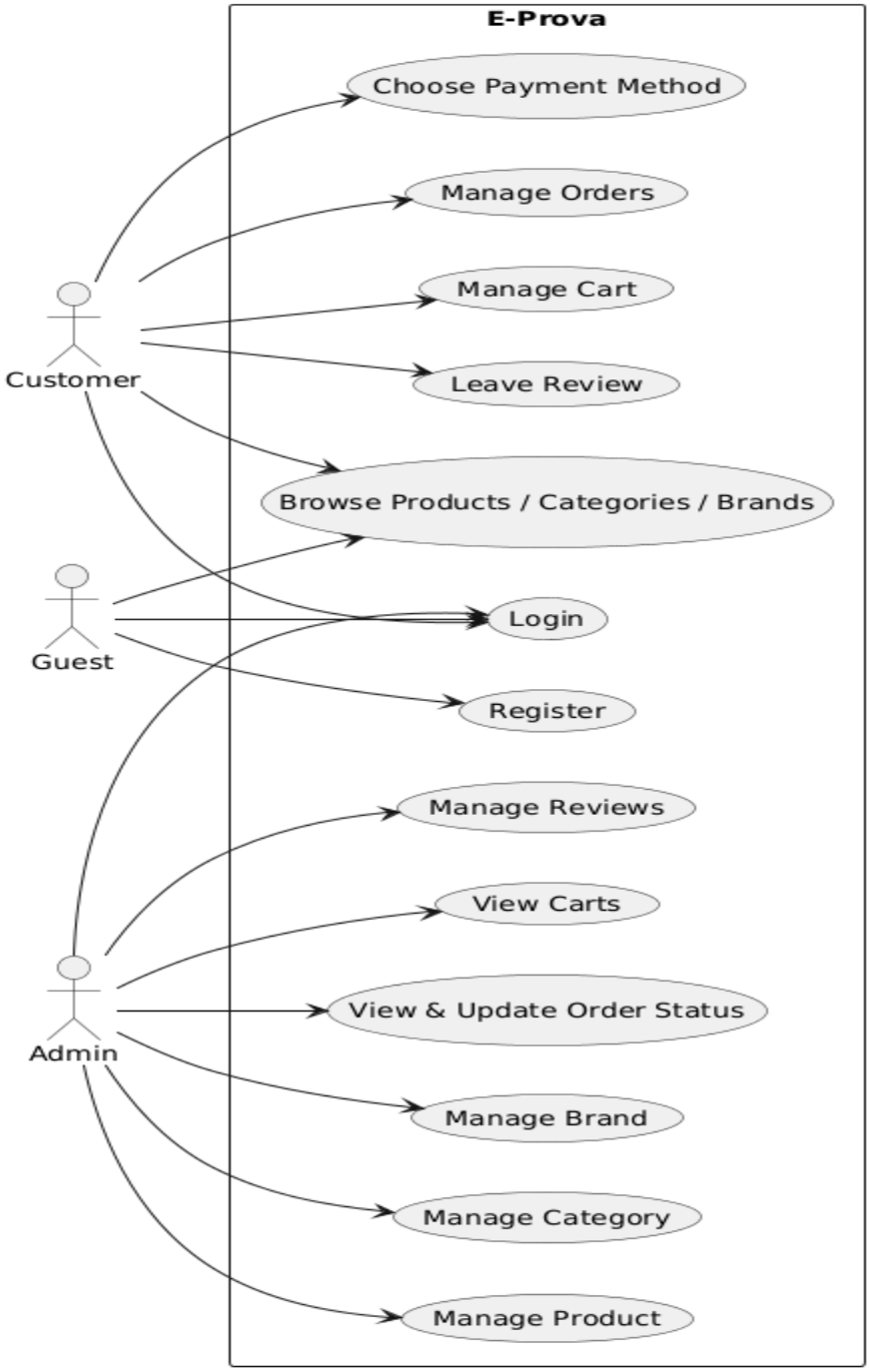
The system must ensure that all essential functionalities like login, registration, order placement, and virtual try-on operate correctly and consistently. The platform should be compatible across different modern browsers and devices (mobile and desktop). Although backups are not currently implemented, the system architecture allows for easy integration of automated database backup tools. Future versions of the system will include scheduled backups of critical user and product data using MongoDB dump utilities or managed services such as MongoDB Atlas.

3.2.4 Usability Requirements

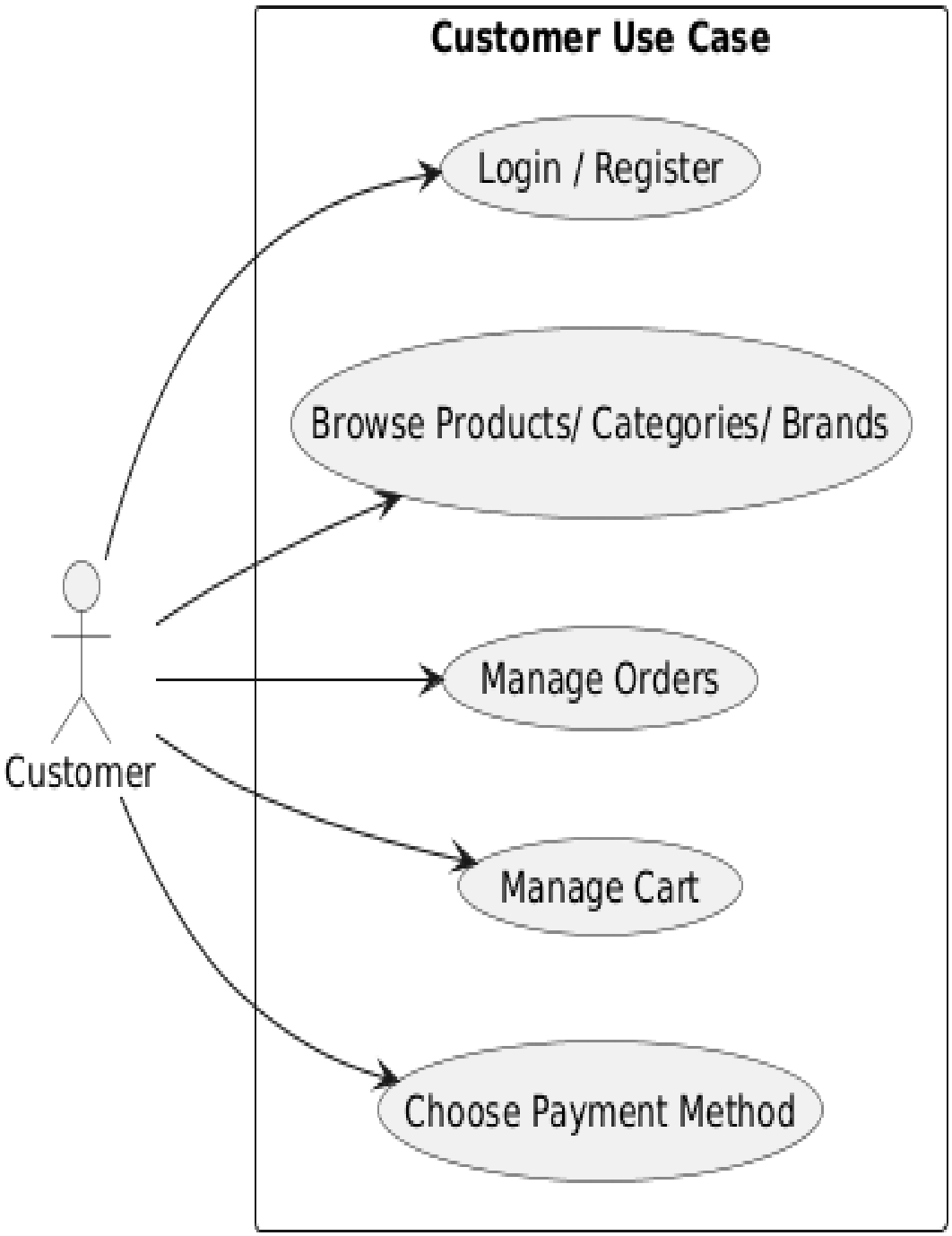
The system should be intuitive and user-friendly, even for users without technical background. The virtual try-on tool is designed with a user-friendly interface that makes the image upload process intuitive and easy, without the need for external instructions. Font sizes, buttons, and interactions should follow accessibility standards. User feedback should be used continuously to enhance the UI/UX.

3.3 System Analysis & Design

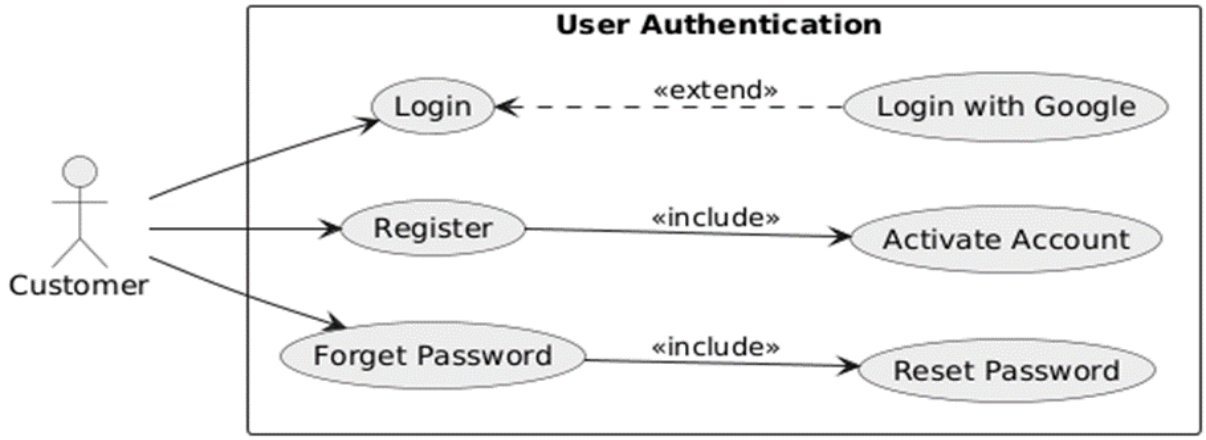
3.3.1 Use cases



User use cases:

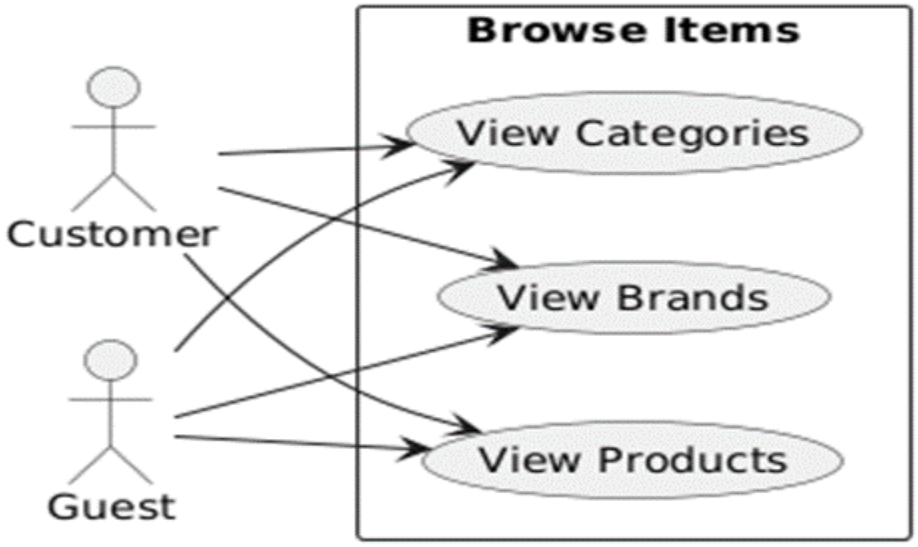


Authentication:



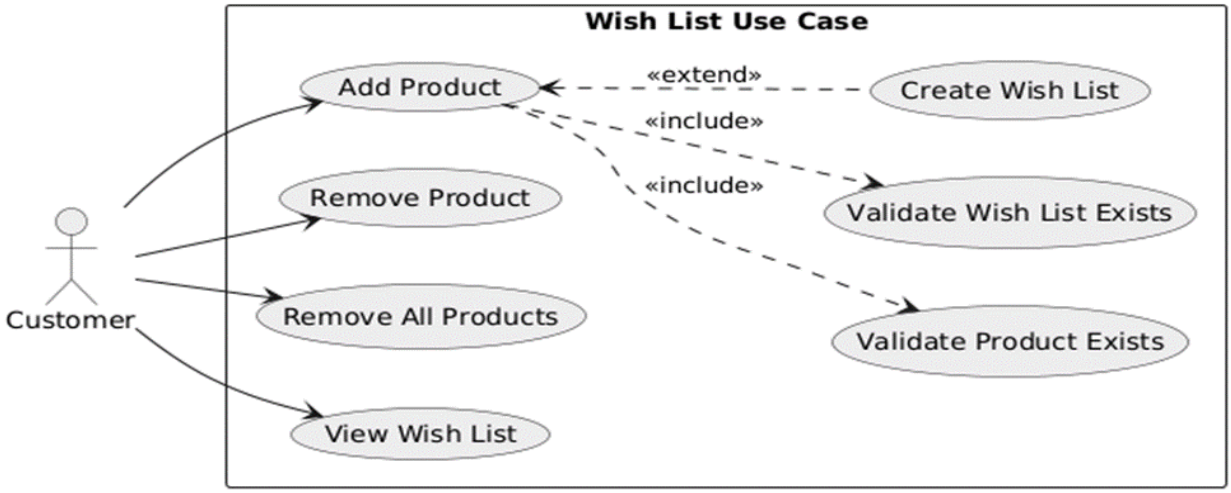
Element	Description
Actor	The customer can register, log in, recover their password, and activate their account.
Description	The customer can register, log in, recover their password, and activate their account.
Preconditions	Customer is not logged in.
Postconditions	Customer is registered, logged in, or has recovered their account.
Basic Flow	1.Customer initiates one of the following: Register, Login, or Forget Password. 2.System processes the request and responds accordingly.
Includes	- Activate Account (included in Register flow) - Reset Password (included in Forgot Password flow)
Extends	- Login with Google (extends Login flow)
Alternative Flow	- If email is already registered → show error "Email already exists". - If password is incorrect during login → show error "Invalid credentials". - If email is not found during recovery → show error "Email not registered".
Exception	-missing required fields (400 bad request) -Invalid token for account activation (401 Unauthorized)

Browse Items:



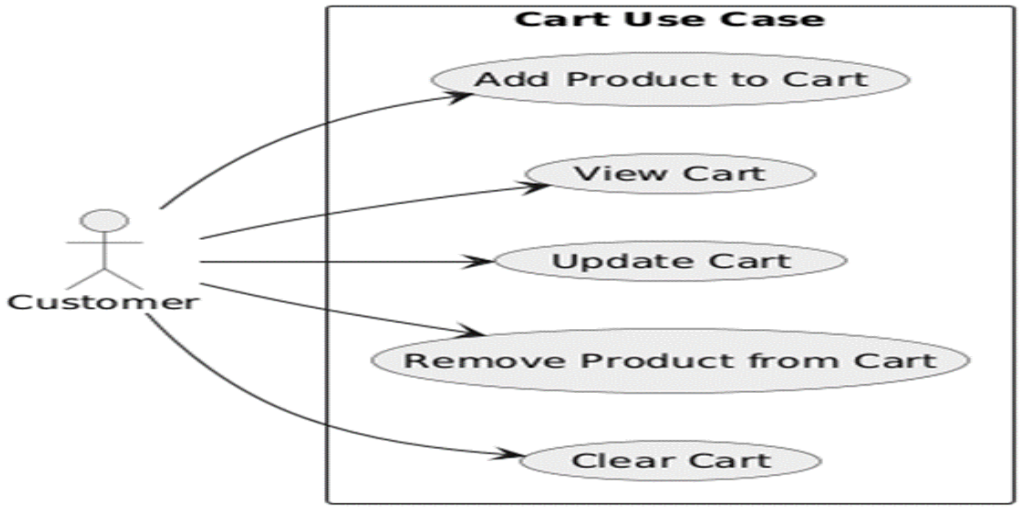
Element	Description
Actor	Customer ,Guest
Description	The user can view available (Products ,Categories ,Brands) without the need to authenticate.
Preconditions	- No authentication required (public access).
Postconditions	- A list of all available (Products ,Categories ,Brands) is returned.
Basic Flow	1. User sends a GET request to the specified endpoint. 2. System fetches all (Products ,Categories ,Brands). 3. System returns the list of (Products ,Categories ,Brands) to the customer.
Includes	None
Extends	None
Alternative Flow	- If no (Products ,Categories ,Brands) exist → return empty list with message "No (Products ,Categories ,Brands) available".
Exceptions	- Internal Server Error (500) if fetching fails.

Wish List:



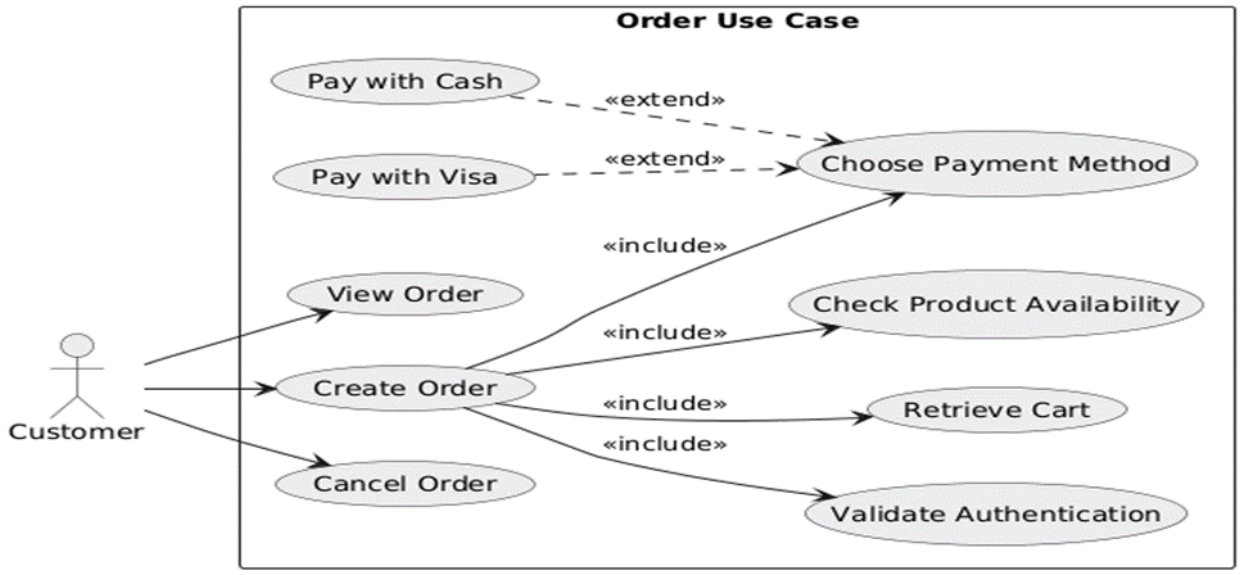
Element	Description
Actor	Customer
Description	Customers add a product to their wish list.
Preconditions	- Customer authenticated and authorized. - Product ID is provided.
Postconditions	- Product added to wish list. - Wish list created if it does not exist.
Basic Flow	1. Customer sends request with product ID. 2. System verifies the customer. 3. System verifies product , wish list exists. 4. If the wish list exists, add the product; otherwise, create a new wish list and then add the product. 5. Return success response.
Alternative Flow	- Product not found → return error. - Product already in wishlist → return error.
Includes	- Validate Product Exists - Validate Wish List Exists
Extends	- Create Wish List if not exists
Exceptions	- Product already in wish list (400)

Cart:



Element	Description
Actor	Customer
Description	The customer can manage their cart by adding, viewing, updating, removing, or clearing products from the cart.
Preconditions	- Customers must be authenticated. - Products must exist in the system.
Postconditions	- The cart is updated based on the action performed (add/update/remove/clear). - Product stock is adjusted accordingly.
Basic Flow	1. Customer selects an action (Add, View, Update, Remove, Clear). 2. Customer sends the request to the appropriate endpoint. 3. System processes the action. 4. System returns a response with the updated cart or success message.
Alternative Flow	- If a product already exists in the cart and is added again → quantity is updated. - If stock is insufficient during add/update → return error message. - If the cart is empty → return empty cart message.
Includes	- Product existence and stock availability check during Add and Update.
Extends	None
Exceptions	- Product not found (404) - Cart not found (404) - Not enough stock (400) - Internal Server Error (500)

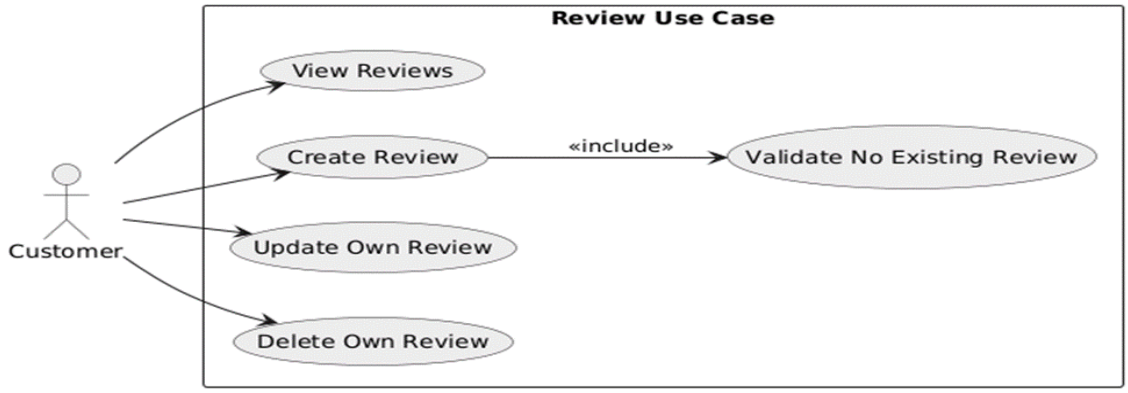
Order:



Element	Description
Actor	Customer
Description	The customer can create orders, view their orders, and cancel their own orders.
Preconditions	- Customers must be authenticated. - Customers must have products in their cart.
Postconditions	- Order is successfully created with stock updated. - Cart is cleared after order creation. - Order can be cancelled if its status is "pending".
Basic Flow	Customer requests order creation. System authenticates the customer and retrieves the cart. System checks product availability. Customer selects the payment method (Visa or Cash). System creates the order, updates stock, and clears the cart. System returns success message and payment URL if applicable.

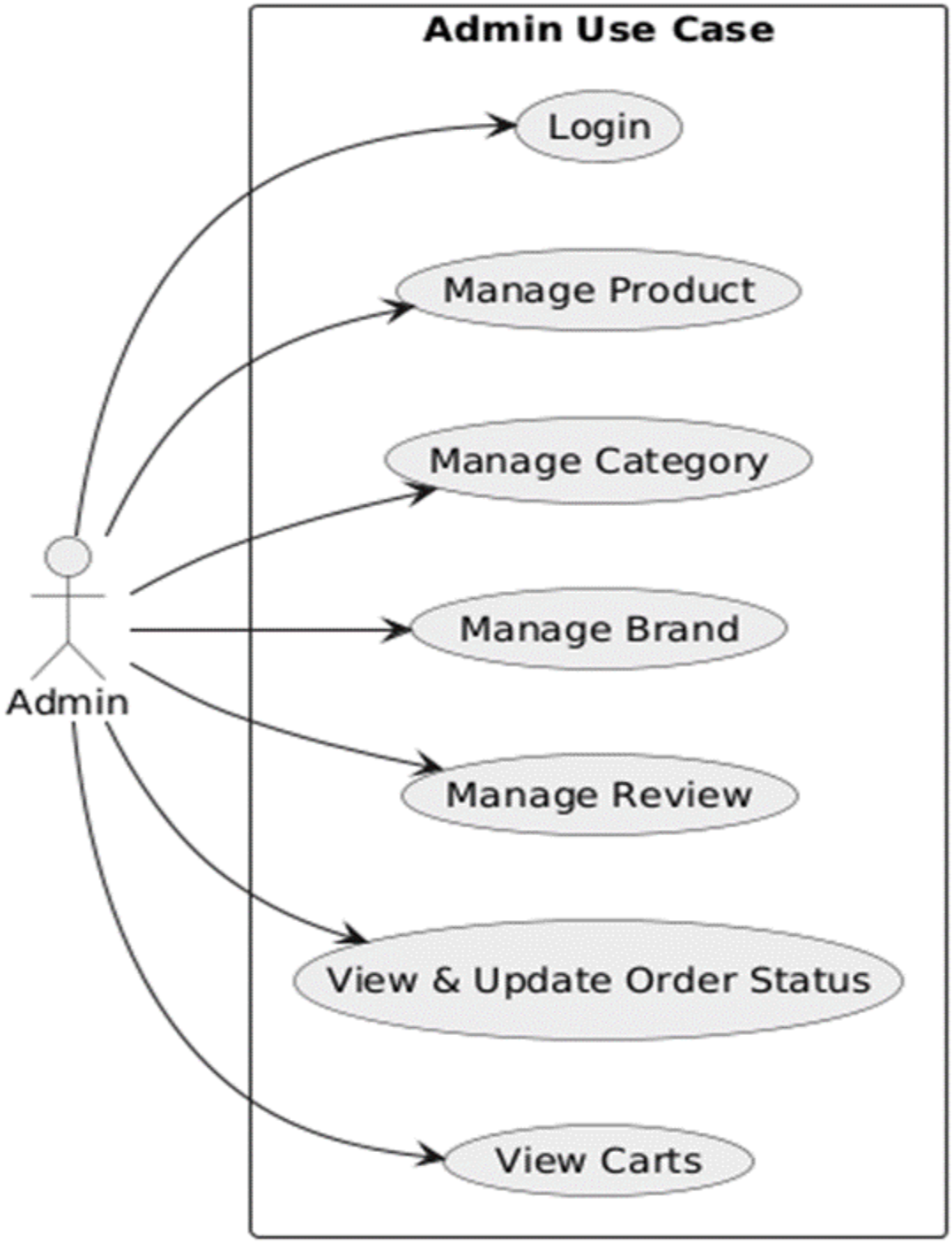
Alternative Flow	**- If the cart is empty → return error "Cart is empty".** **- If product is out of stock → return "Product unavailable".** **- If Visa payment fails → return payment failure message.**
Cancel Order Flow	**1. Customer requests to cancel order.** **2. System verifies order ownership.** **3. System checks order status is "pending".** **4. Order status is updated to "cancelled".** **5. Stock for ordered products is restored.** **6. System returns cancellation success message.**
Includes	**- Validate Authentication.** **- Retrieve Cart.** **- Check Product Availability.** **- Choose a Payment Method.**
Extends	**- Pay with Visa.** **- Pay with Cash.**
Exceptions	**- 404 Not Found if cart or product is not found.** **- 403 Forbidden if unauthorized.** **- 500 Internal Server Error on system failure.**

Review

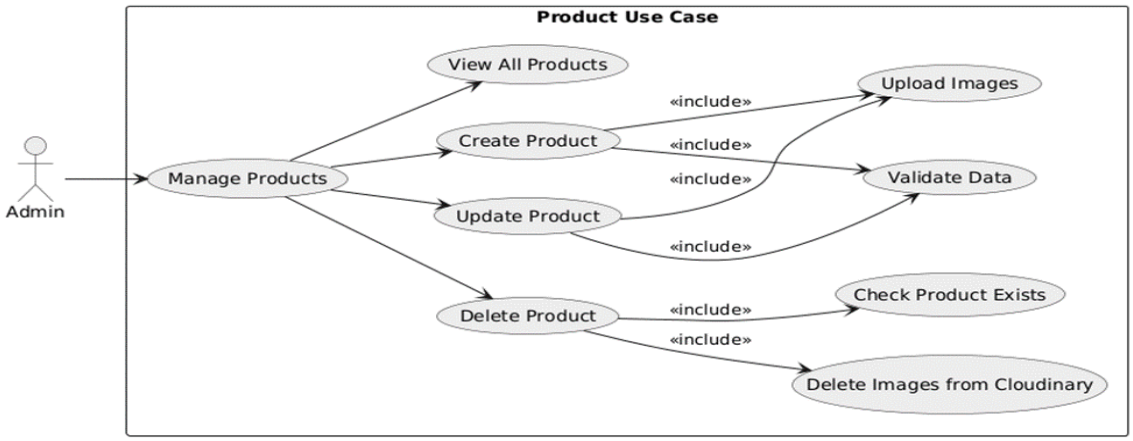


Element	Description
Actor	Customer
Description	The customer can create, view, update, and delete their own reviews on products.
Preconditions	- Customers must be authenticated. - Product must exist. - For create: customer must not have reviewed the same product before.
Postconditions	- Product added to wish list. - Wish list created if it does not exist.
Basic Flow	1. Customer sends request with product ID. 2. System verifies the customer. 3. System verifies product , wish list exists. 4. If the wish list exists, add the product; otherwise, create a new wish list and then add the product. 5. Return success response.
Alternative Flow	- Product not found → return error. - Product already in wishlist → return error.
Includes	- Validate Product Exists - Validate Wish List Exists
Extends	- Create Wish List if not exists
Exceptions	- Product not found (404). - Product already in the wish list (400).

Admin use cases:

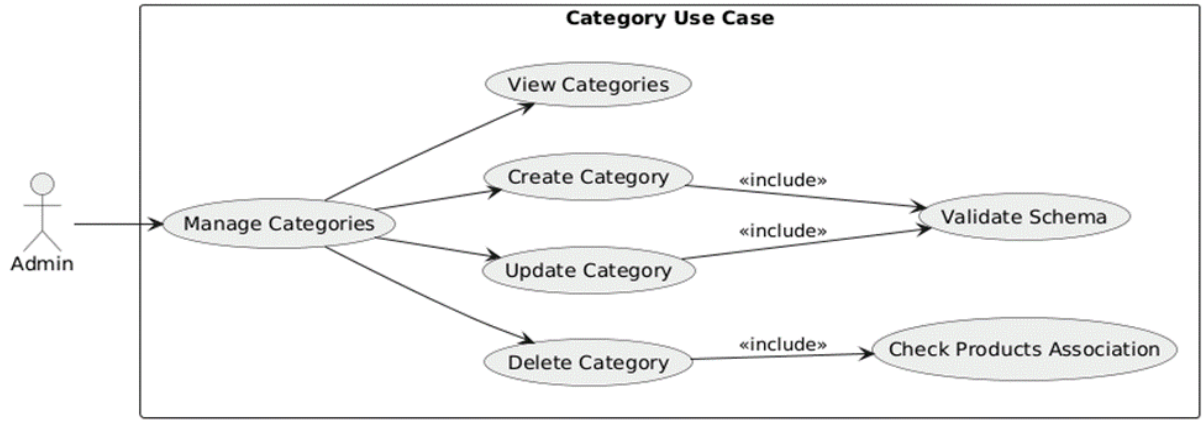


Manage Product:



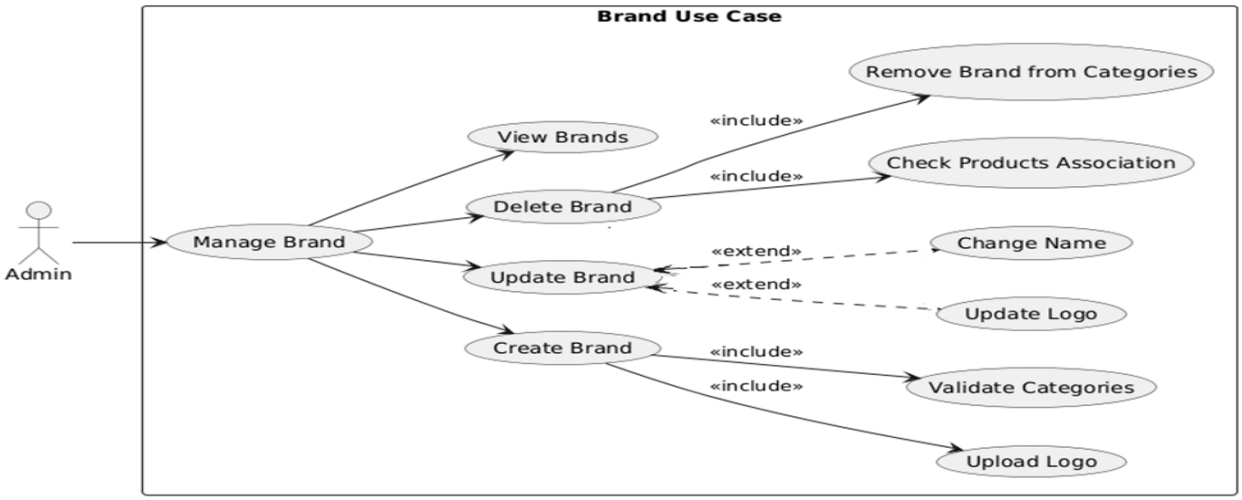
Element	Description
Actor	Admin
Description	The Admin can manage products by creating new products, updating existing ones, viewing all products, or deleting products.
Preconditions	- Admin must be authenticated and authorized. - Input data must be valid according to the schema.
Postconditions	- Products will be stored, updated, viewed, or deleted in the database accordingly.
Basic Flow	1. Admin navigates to the product management section. 2. Admin performs one of the following actions: create, update, view, or delete a product. 3. System processes the request and returns a success or error message.
Alternative Flow	- If the product does not exist when updating or deleting → show error. - If input data is invalid → show error.
Includes	- Validate Data (to check the validity of the data during create/update). - Upload Images (for create/update). - Delete Images (when deleting a product). - Check Product Exists (before deletion).
Extends	None.
Exceptions	- Product not found(404). - Invalid input data(400). - Failure in uploading or deleting images(500).

Manage Category:



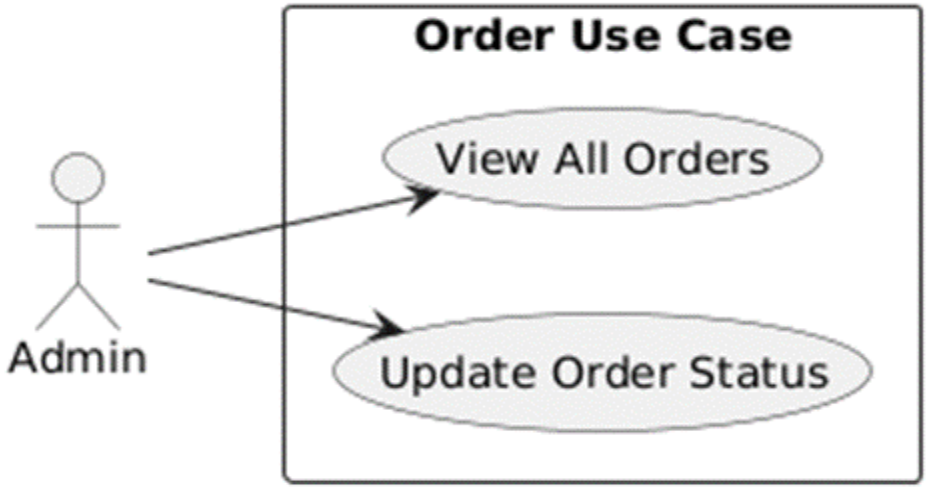
Element	Description
Actor	Admin
Description	The Admin can manage categories by creating new ones, updating existing ones, viewing all categories, or deleting categories.
Preconditions	- Admin must be authenticated and authorized.- Input data must be valid according to the schema.
Postconditions	- Categories will be stored, updated, viewed, or deleted in the database accordingly
Basic Flow	1. Admin navigates to the category management section. 2. Admin performs one of the following: create, update, view, or delete a category. 3. System processes the request and returns a success or error message.
Alternative Flow	- If the category is associated with existing products → prevent deletion. - If input data is invalid → show error.
Includes	- Validate Schema (to check the validity of the data during create/update). - Check Products Association (before deletion).
Extends	None.
Exceptions	- Category not found (404) - Attempting to delete a category linked to products (400) - Invalid input data (400)

Manage Brand:



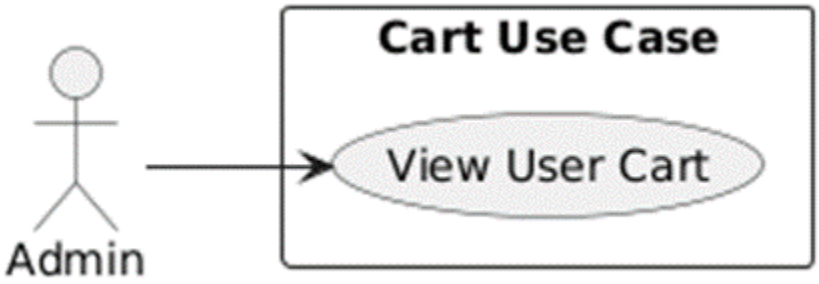
Element	Description
Actor	Admin
Description	Admin can manage brands by creating new ones, updating existing ones, viewing all brands, and deleting brands.
Preconditions	- Admin must be authenticated and authorized. - Categories must exist in the system.
Postconditions	- Brands will be stored, updated, viewed, or deleted in the database accordingly.
Basic Flow	1. Admin navigates to the brand management section. 2. Admin performs one of the following: create, update, view, or delete a brand. 3. System processes the request and shows a success message or error.
Alternative Flow	- If no image is uploaded during create/update → show error. - If a brand has existing products → prevent deletion. - If invalid category ID → show error.
Includes	- Upload Logo - Validate Category IDs
Extends	None.
Exceptions	- Brand not found(404) - Category not found(404) - Cloud upload fails(500)

Manage Order:



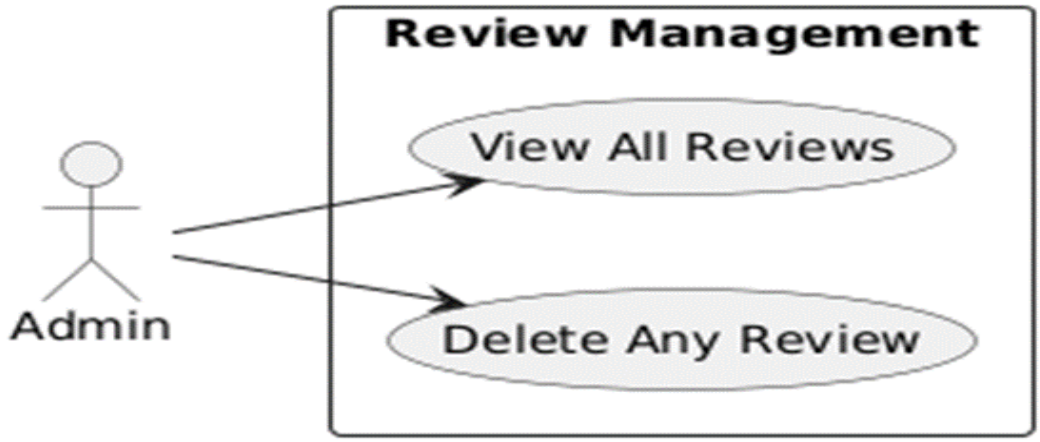
Element	Description
Actor	Admin
Description	The admin can update the order status to "shipped" or "delivered" to track the progress of the order.
Preconditions	- The order exists - The order status is not "cancelled"
Postconditions	- The order status is successfully updated
Basic Flow	1. Admin accesses the order management page. 2. Selects a specific order. 3. Clicks "Mark as shipped" or "Mark as delivered". 4. Status is updated in the database.
Alternative Flow	-if a non-admin user attempts to update an order status→ The system returns an error message: "Unauthorized access".
Includes	None
Extends	None
Exceptions	- Order not found(404) - Invalid status value(400)

Manage Cart:



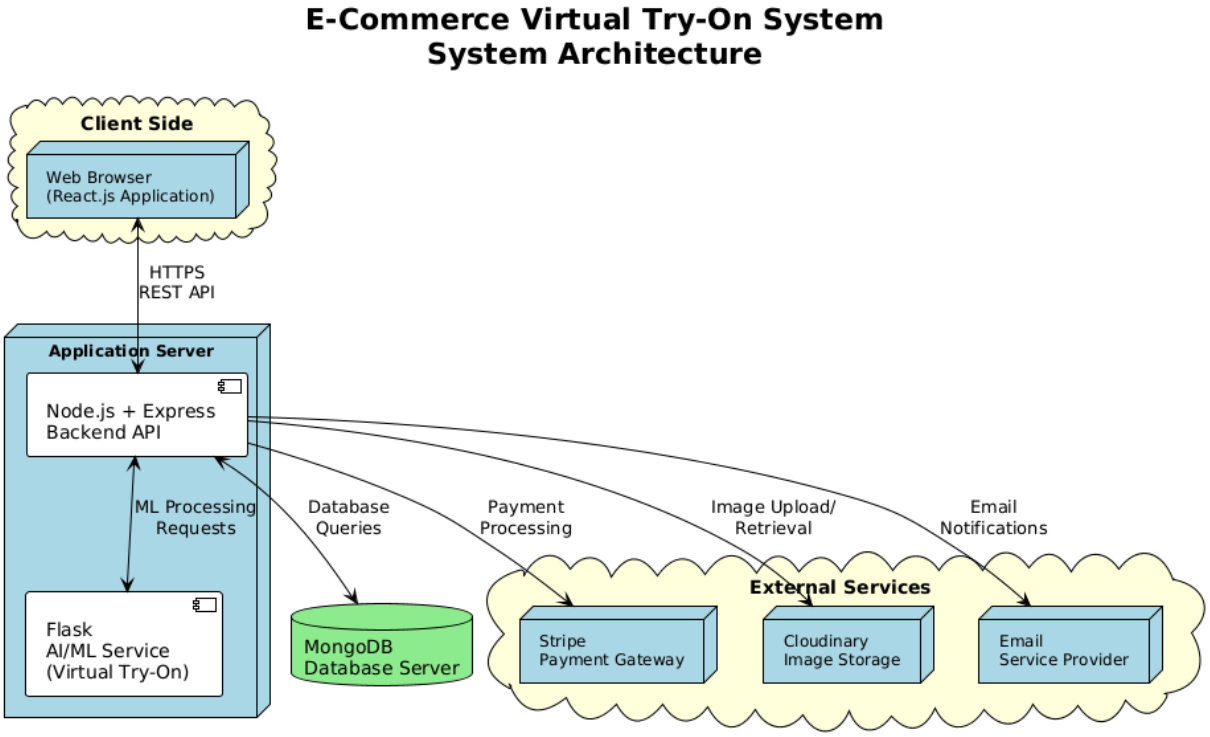
Element	Description
Actor	Admin
Description	The Admin can view the cart of any user by providing the cartId.
Preconditions	- Admin must be authenticated and have the "admin" role.
Postconditions	- The specified cart is retrieved and returned with full product details.
Basic Flow	1. Admin sends a request to the cart endpoint with cartId. 2. System verifies the admin role. 3. System fetches the cart by cartId. 4. Cart details are returned.
Alternative Flow	- If cartId is missing → show error "cartId is required". - If the cart is not found → return "Cart not found".
Includes	- Validate Admin Role - Validate cartId presence - Populate cart products
Extends	None
Exceptions	- Missing or invalid cartId (400 Bad Request) - Cart not found (404 Not Found)

Manage Review:

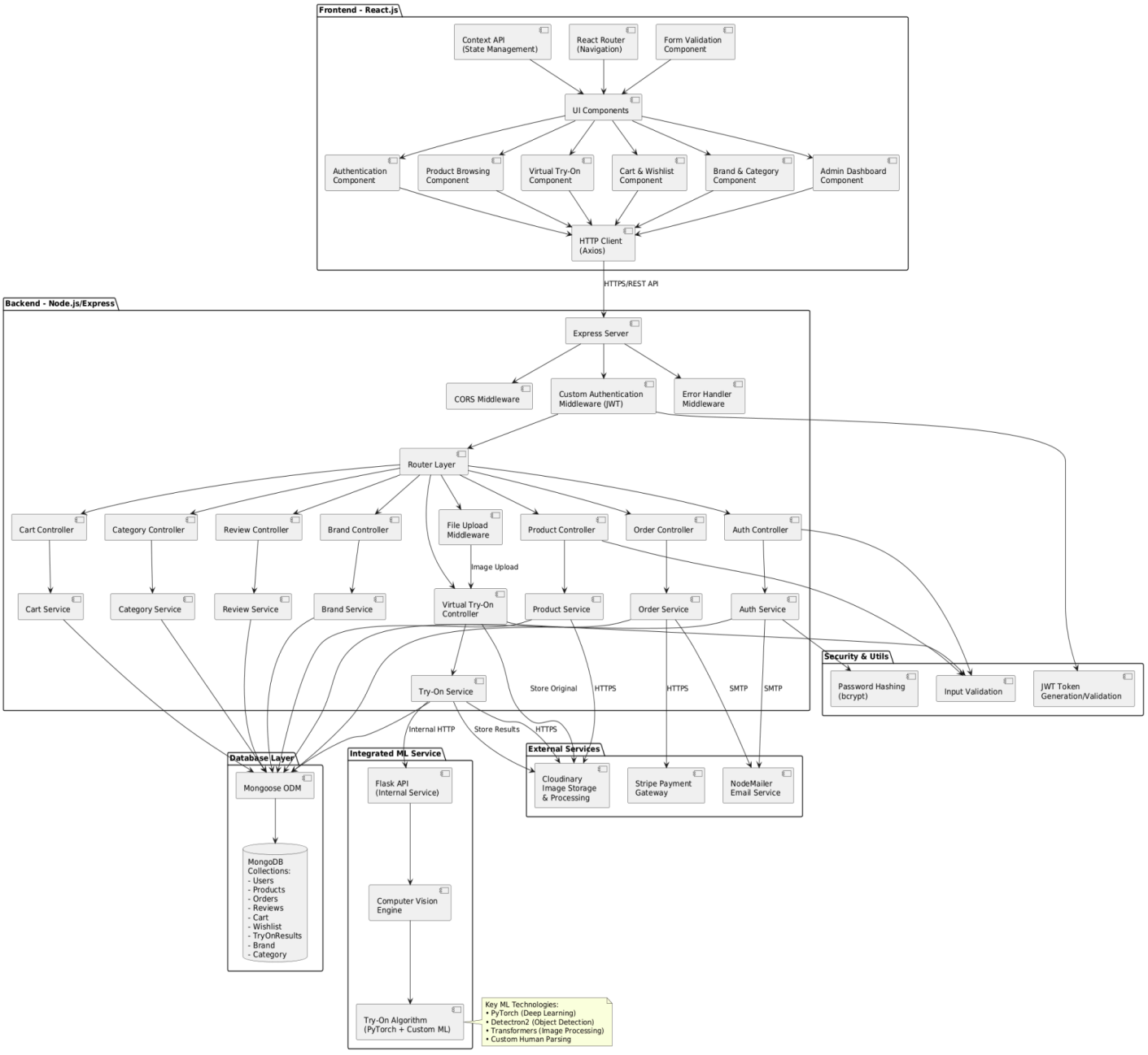


Element	Description
Actor	Admin
Description	The Admin can view all reviews in the system and delete any review if necessary.
Preconditions	- Admin must be authenticated and authorized.
Postconditions	- Selected review is deleted (if action taken). - If the review affects the rating, the product's average rating is recalculated automatically.
Basic Flow	1. Admin navigates to the review management section. 2. Admin views all reviews. 3. Admin deletes any review if needed.
Alternative Flow	- If review is not found → show error. - If deletion fails → show error.
Includes	- Validate Review Existence - Recalculate Average Rating (automatically after deletion)
Extends	None
Exceptions	- Review not found(404). - Database operation failed(500).

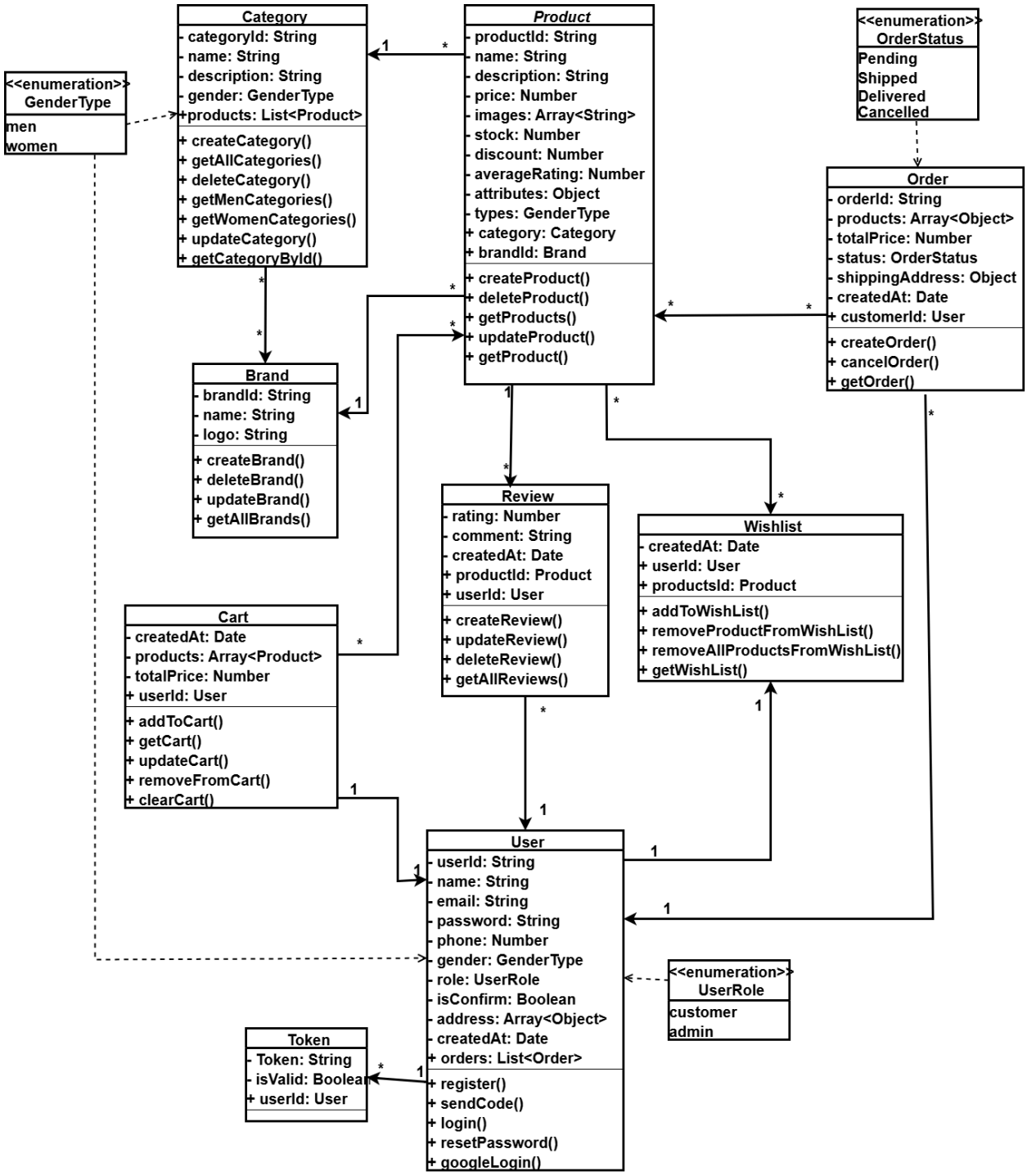
3.3.2 System Architecture



3.3.3 Component Diagram



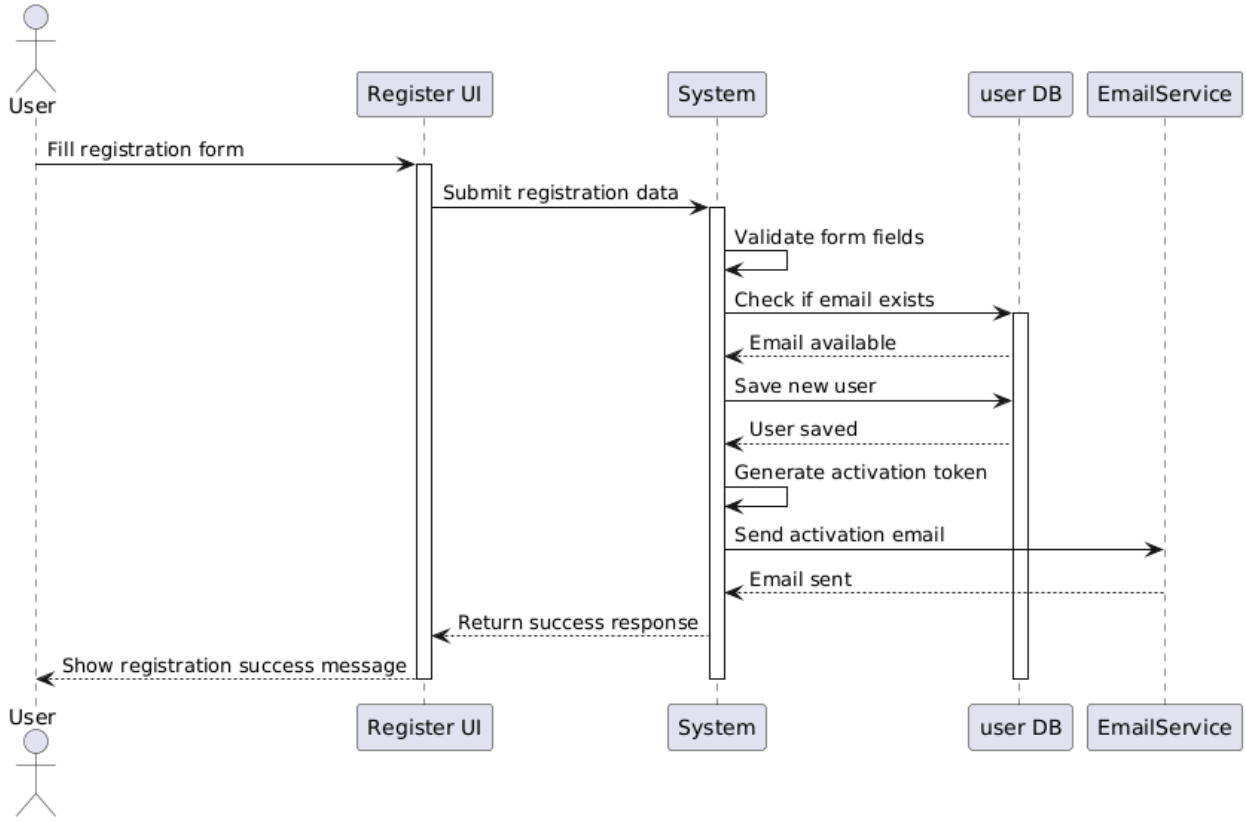
3.3.4 Class Diagram



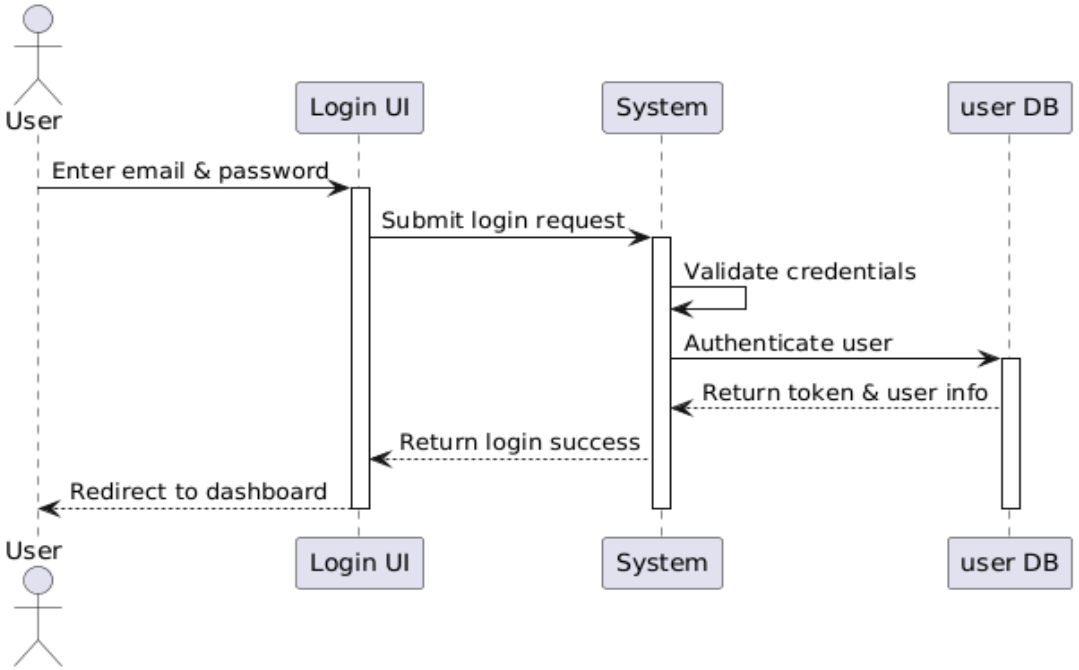
3.3.5 Sequence Diagram

User Sequence Diagrams:

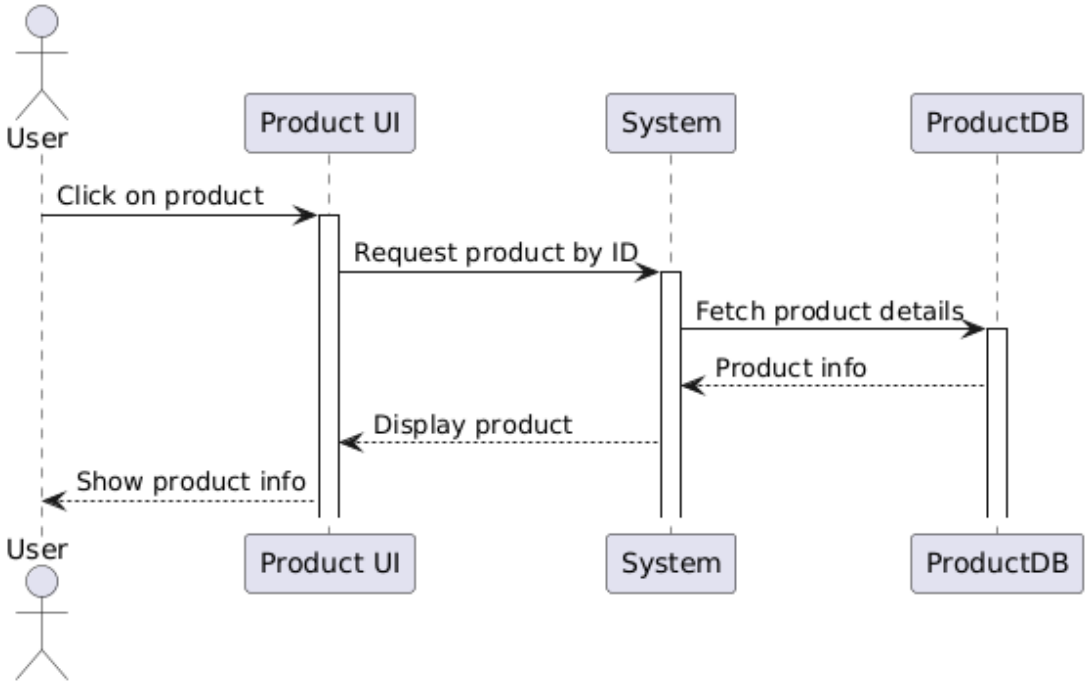
Register sequence Diagram:



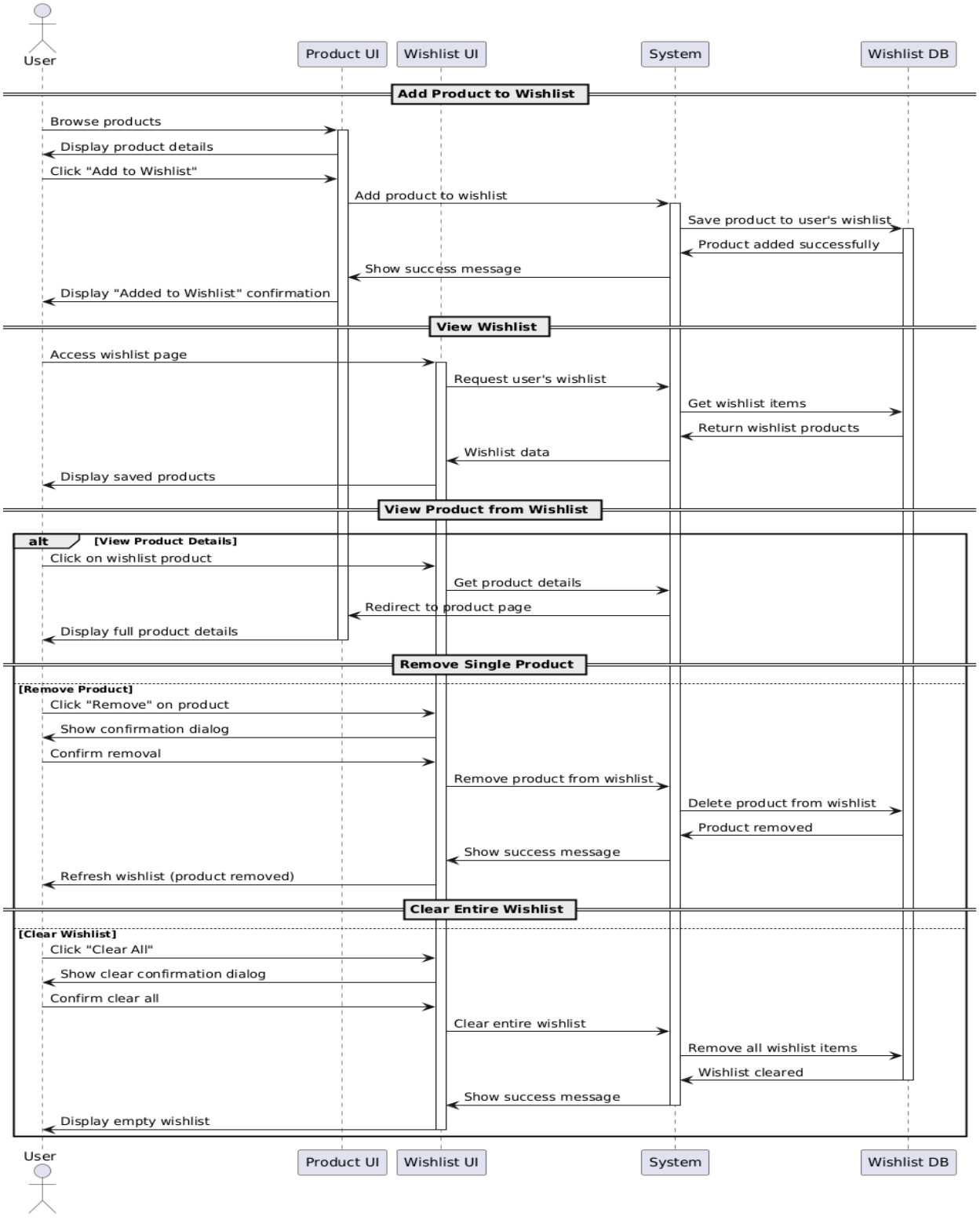
Login Sequence Diagram:



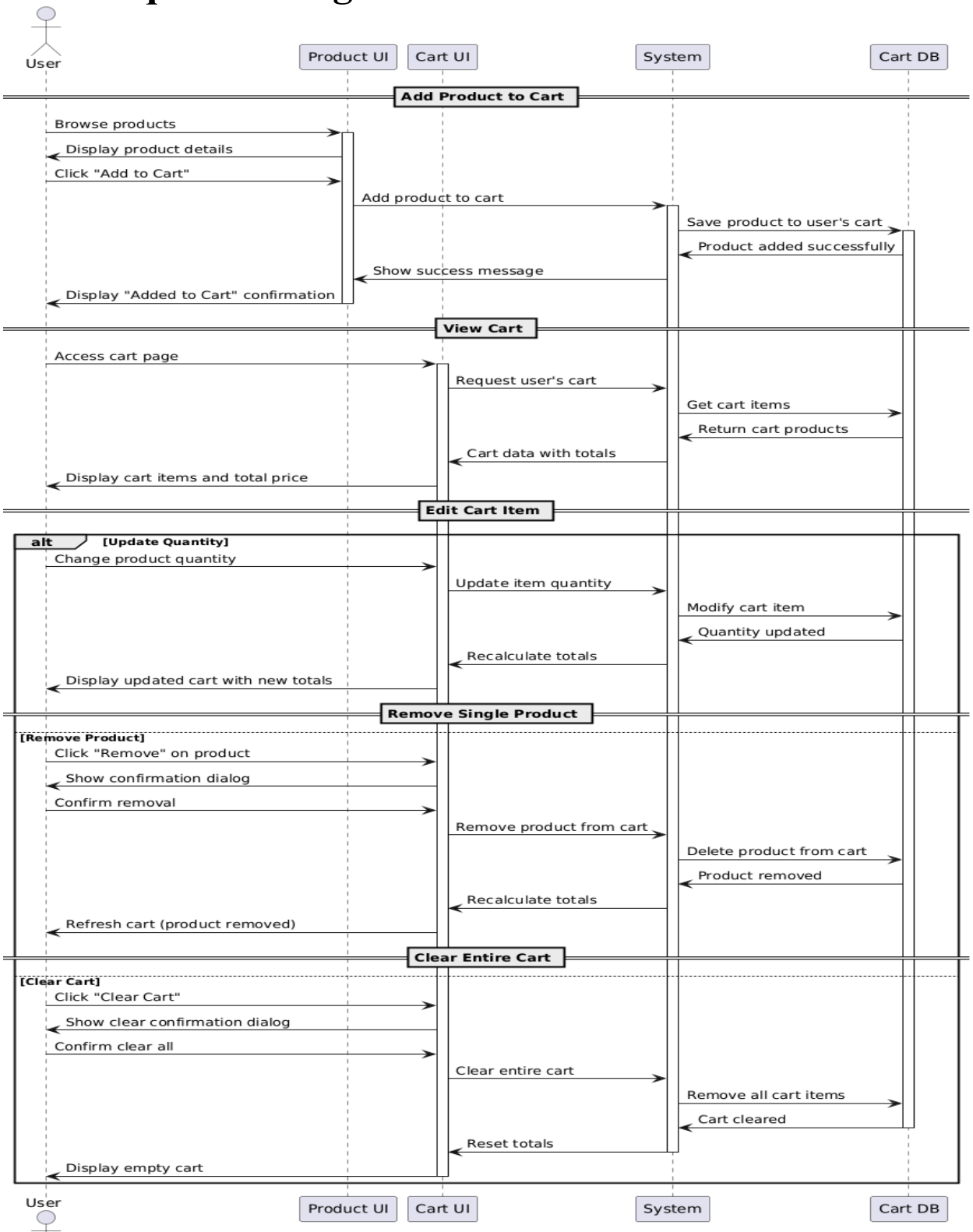
View Product Sequence Diagram:



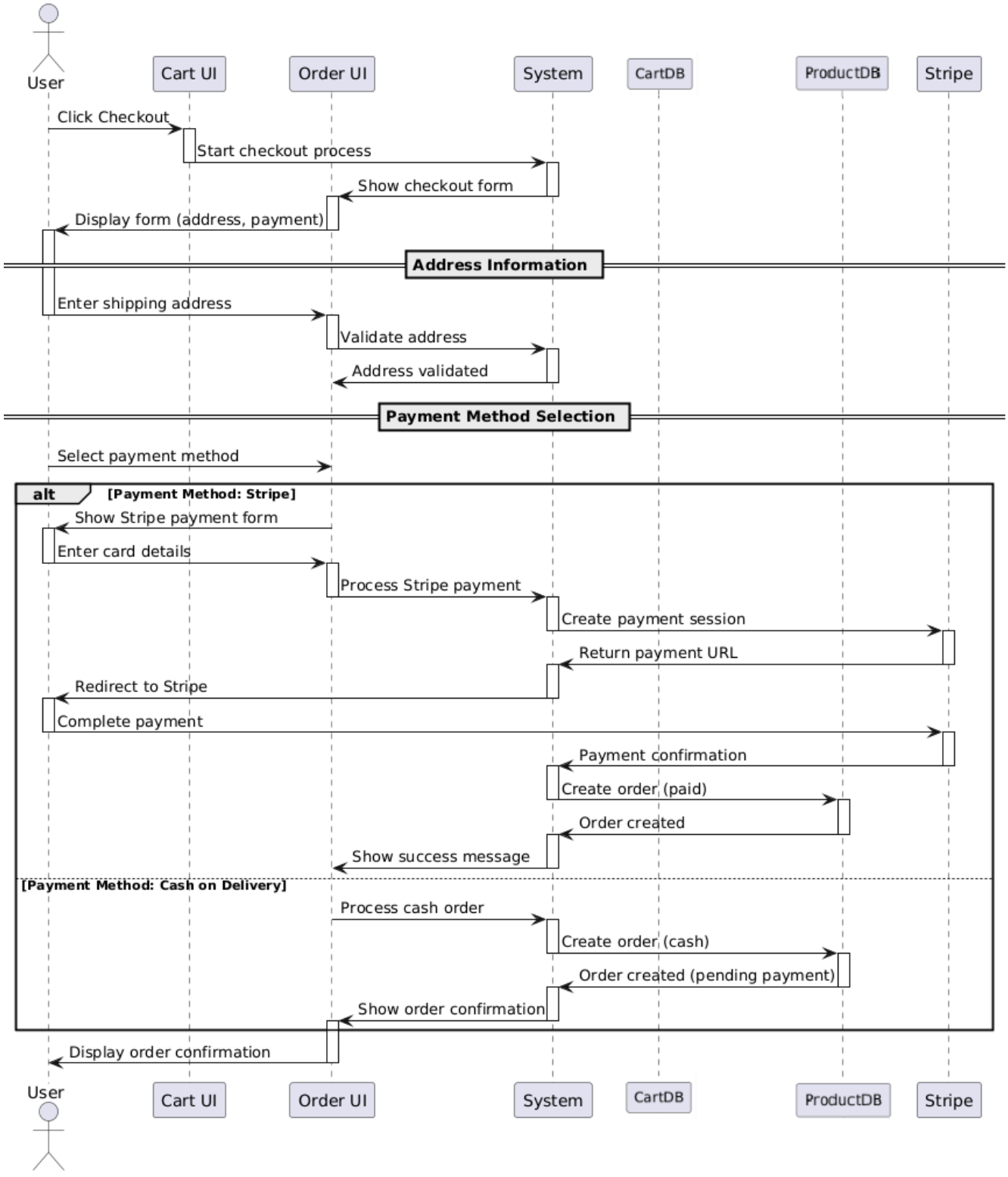
Wishlist Sequence Diagram:



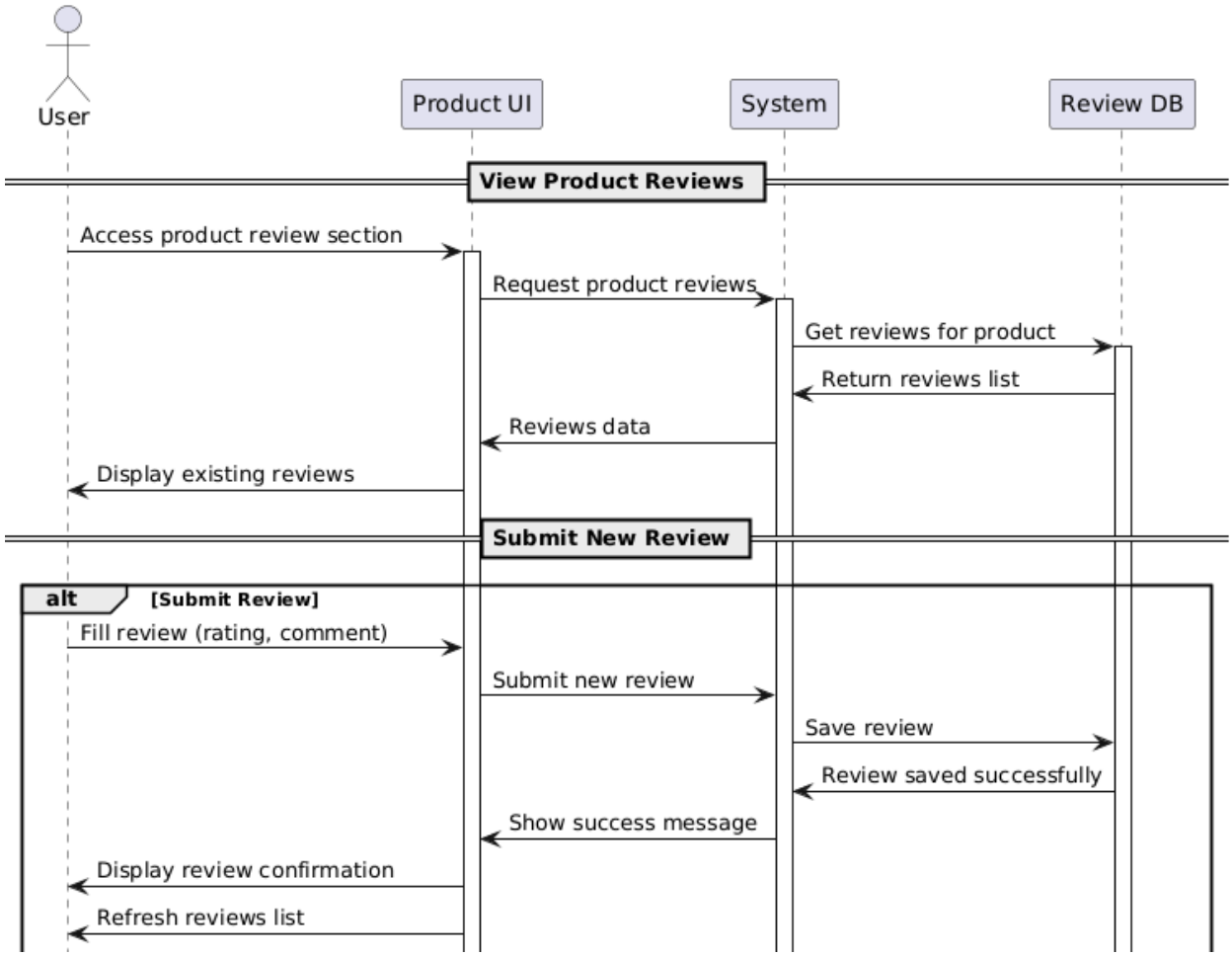
Cart Sequence Diagram:

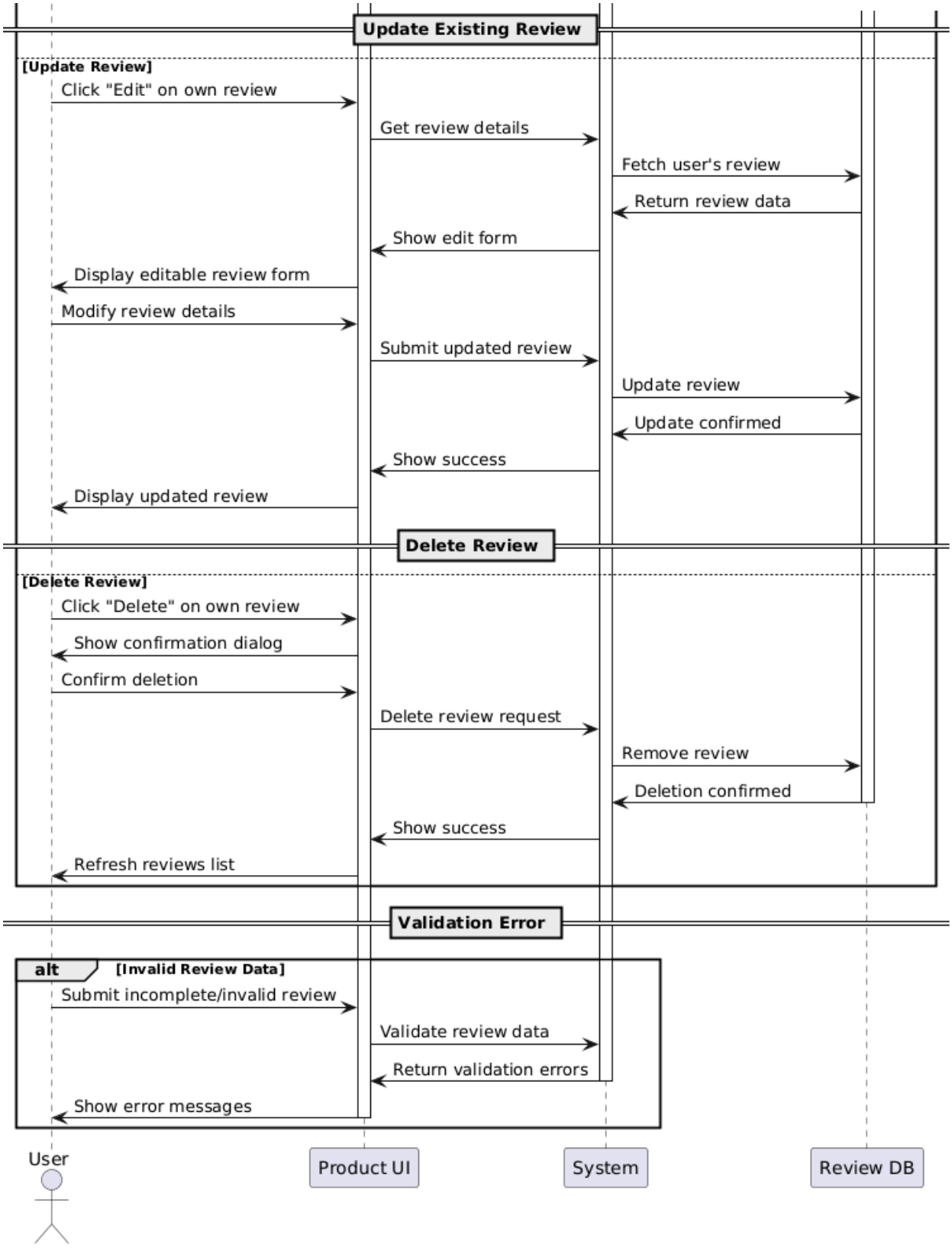


Checkout Sequence Diagram:

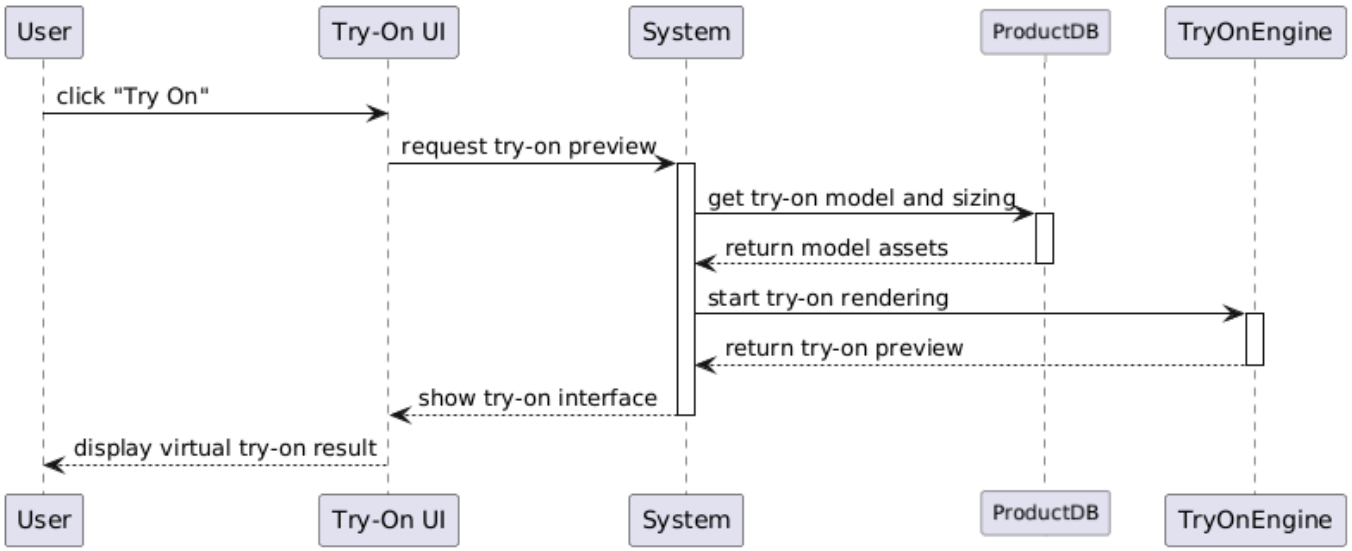


Review Sequence Diagram:



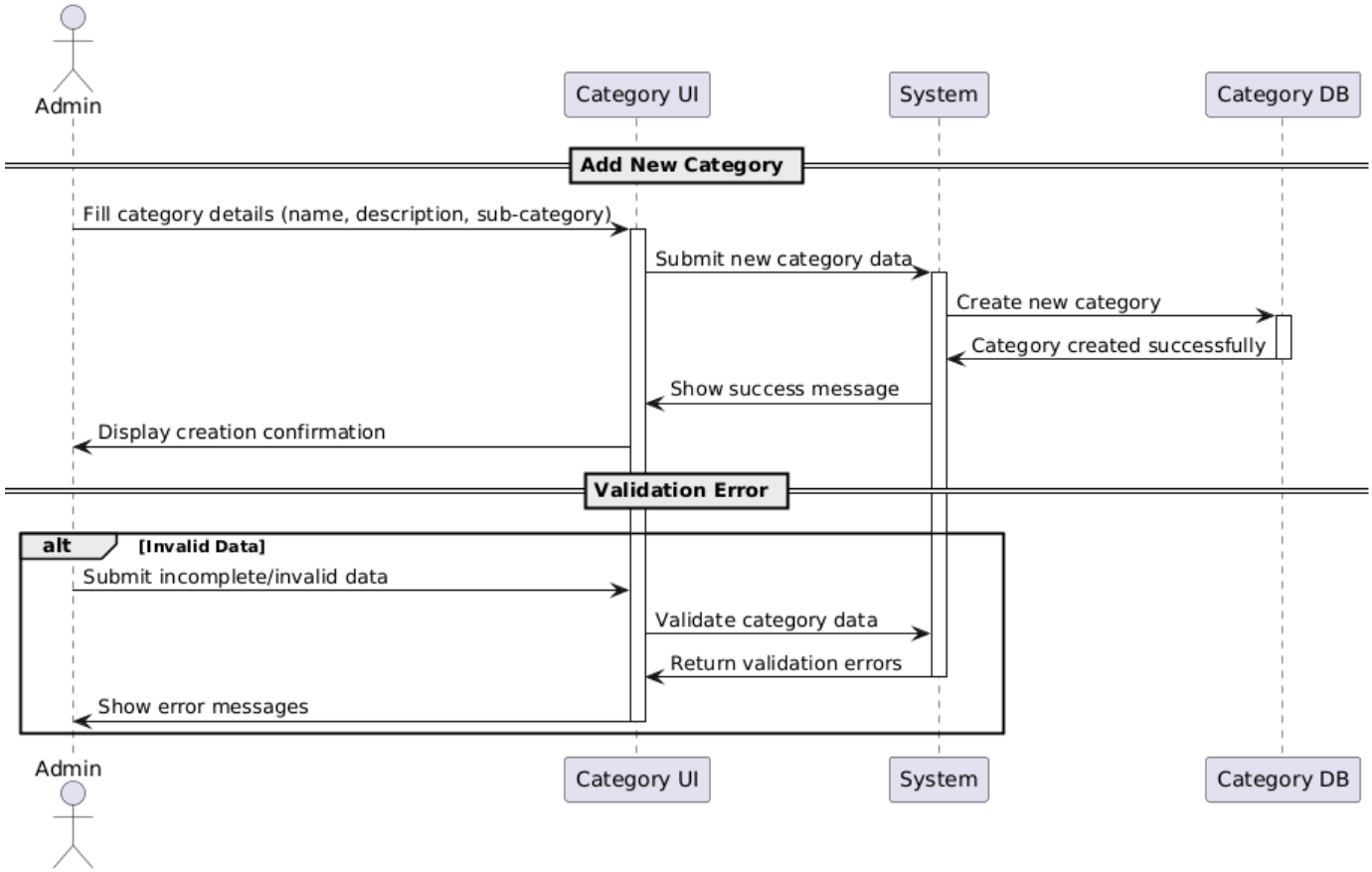

Update Existing Review
[Update Review]
Click "Edit" on own review
Get review details
Fetch user's review
Return review data
Show edit form
Display editable review form
Modify review details
Submit updated review
Update review
Update confirmed
Show success
Display updated review
Delete Review
[Delete Review]
Click "Delete" on own review
Show confirmation dialog
Confirm deletion
Delete review request
Remove review
Deletion confirmed
Show success
Refresh reviews list
Validation Error
alt
[Invalid Review Data]
Submit incomplete/invalid review
Validate review data
Return validation errors
Show error messages
User
Product UI
System
Review DB

Virtual Try-On Sequence Diagram:

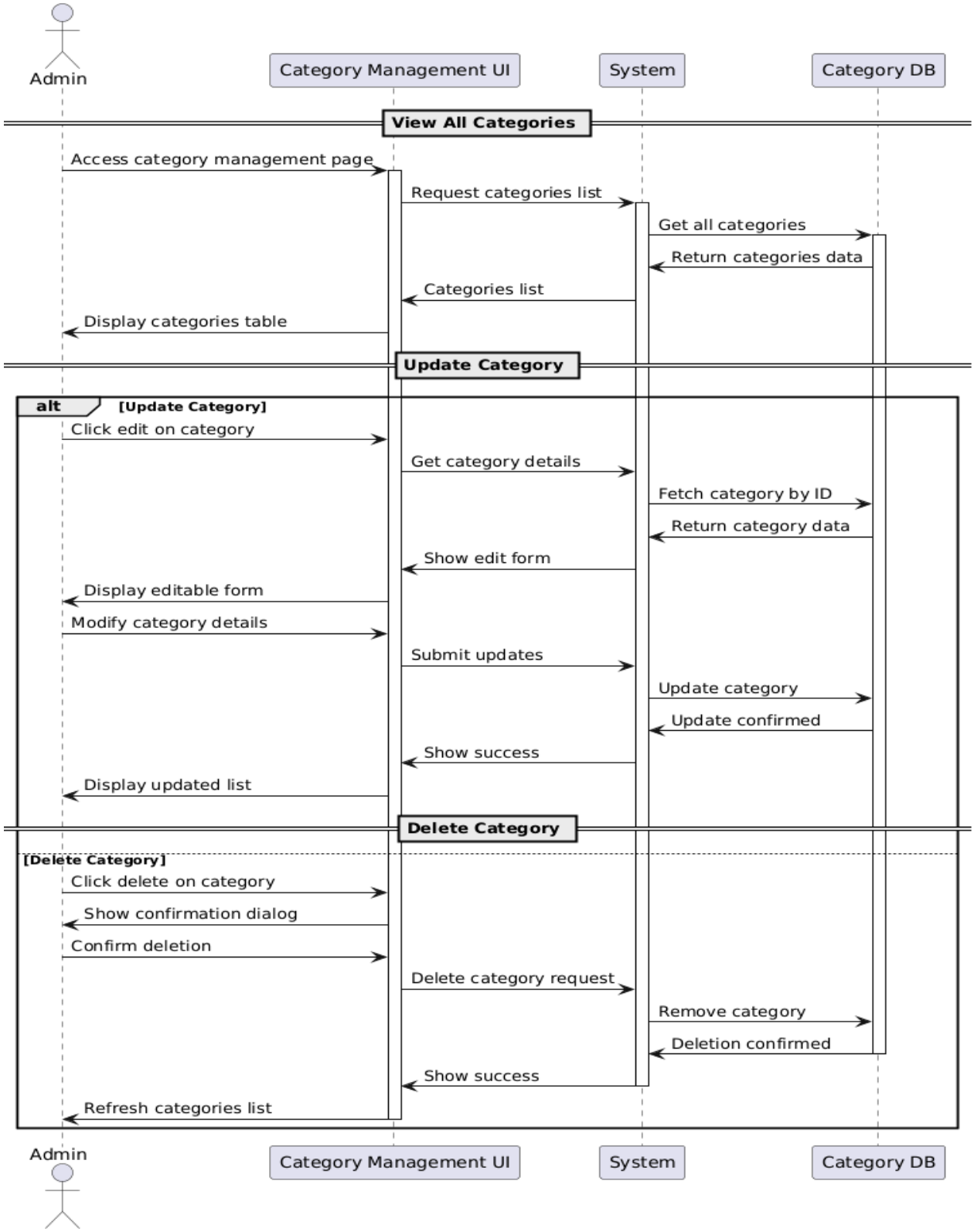


Admin sequence Diagrams:

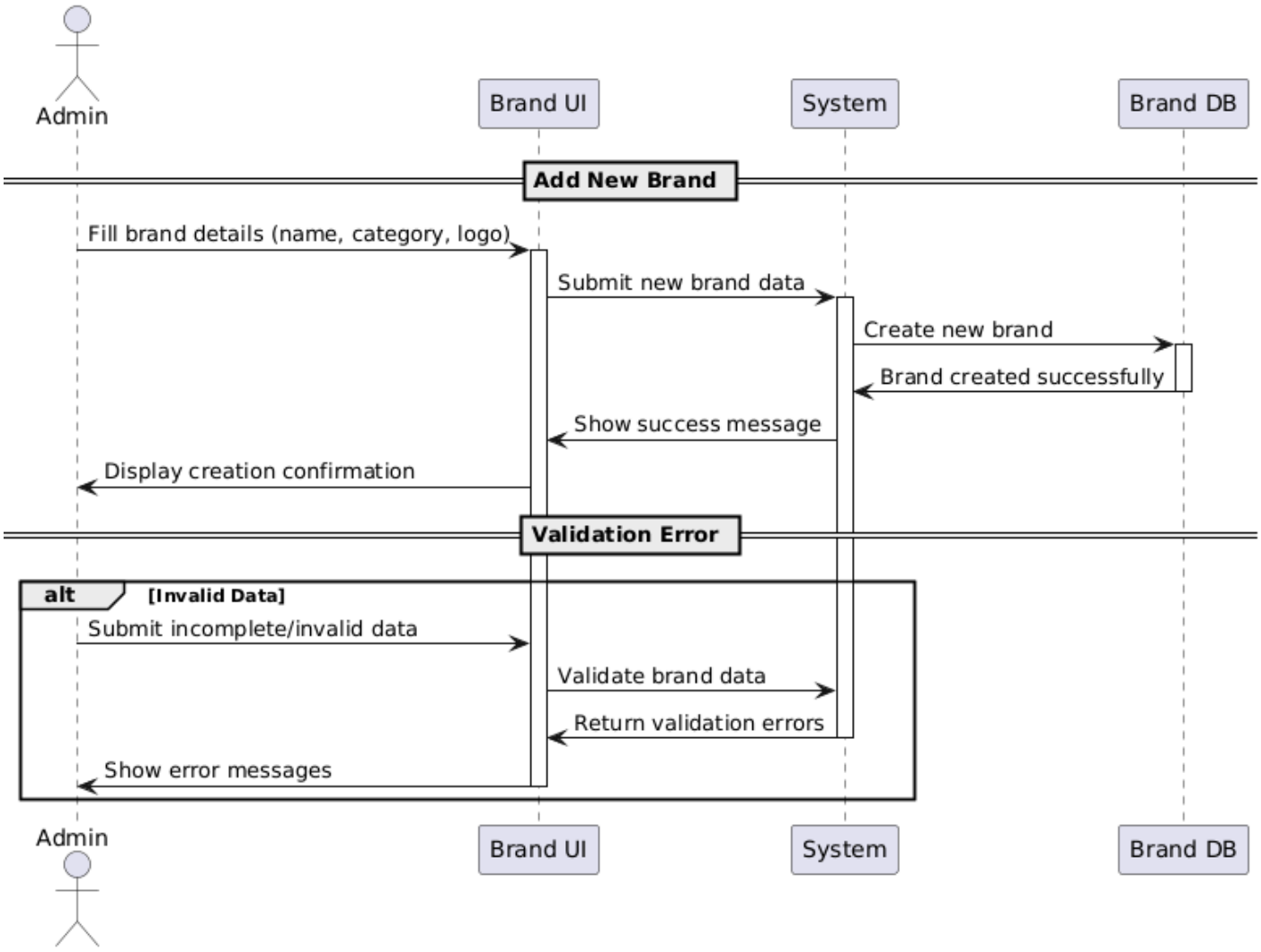
Add New Category Sequence Diagram:



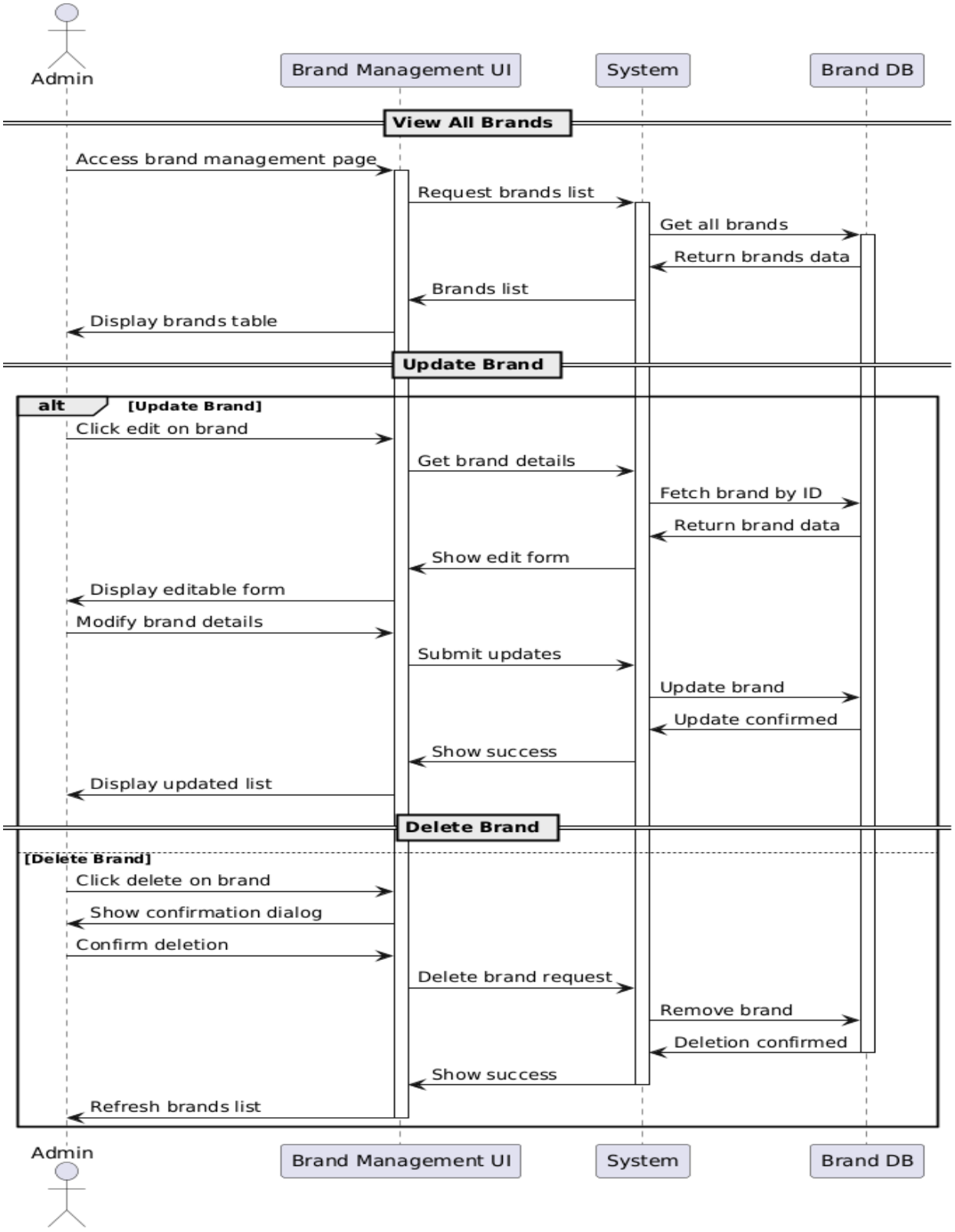
Manage Category Sequence Diagram:



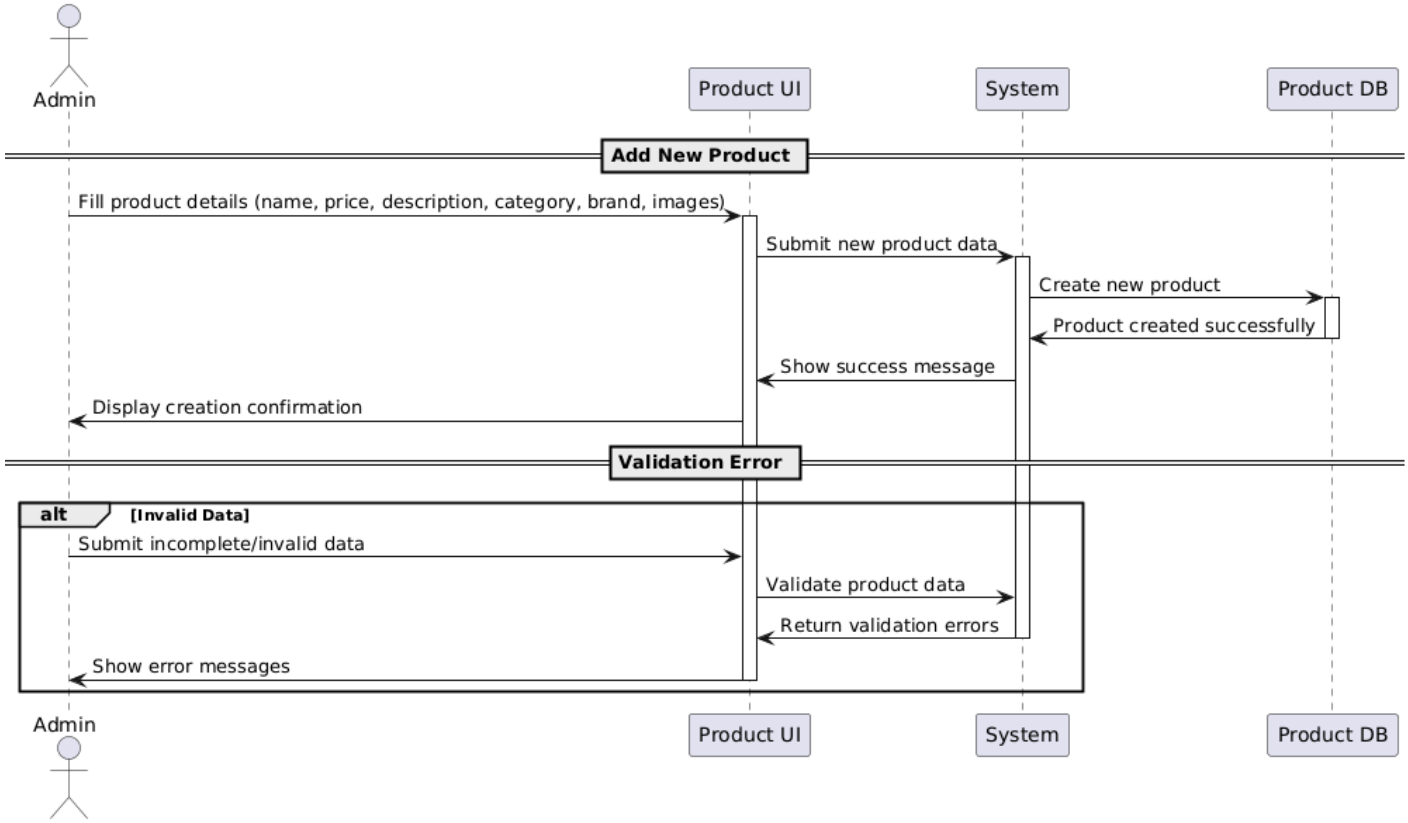
Add New Brand Sequence Diagram:



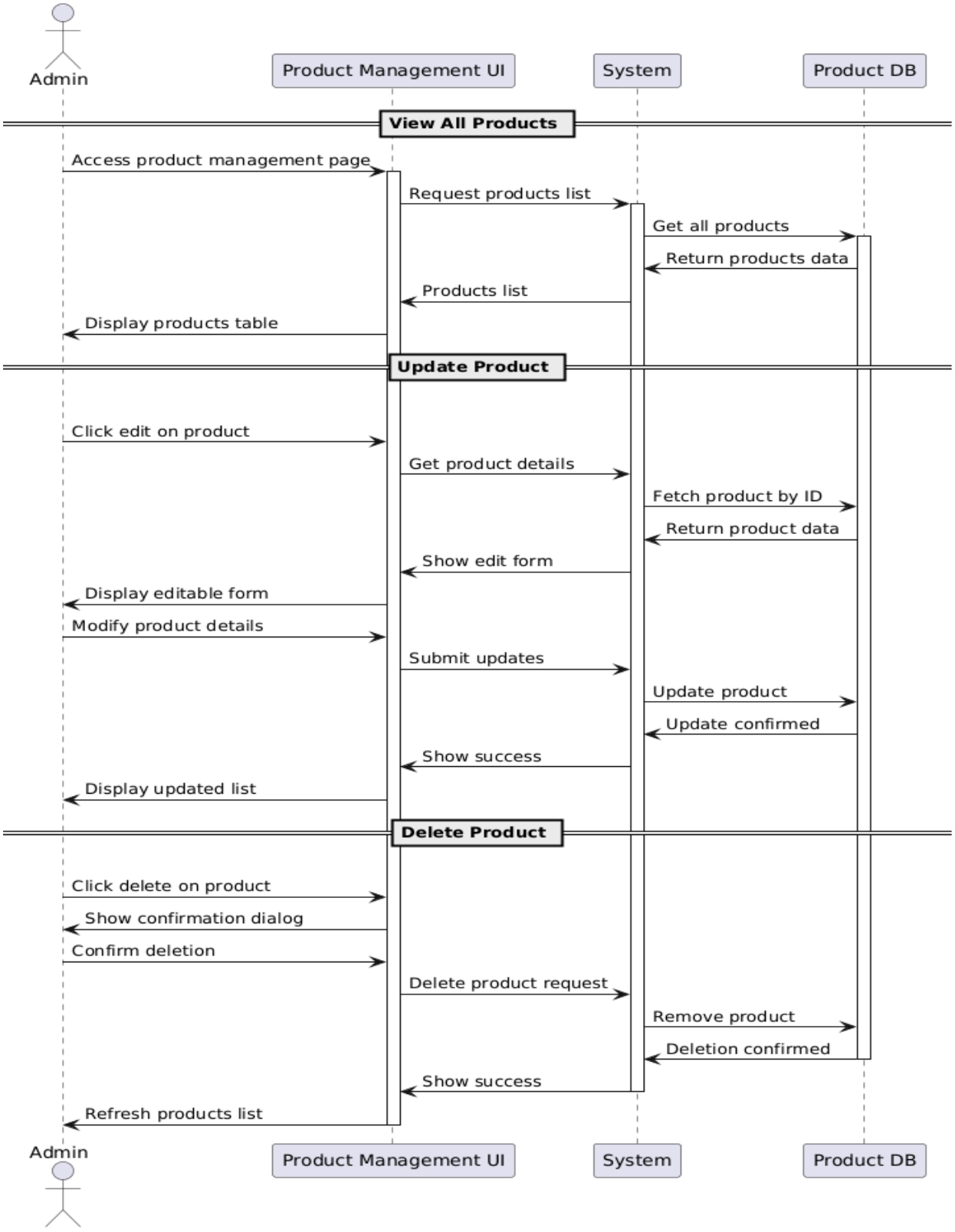
Manage Brand Sequence Diagram:



Add New Product Sequence Diagram:

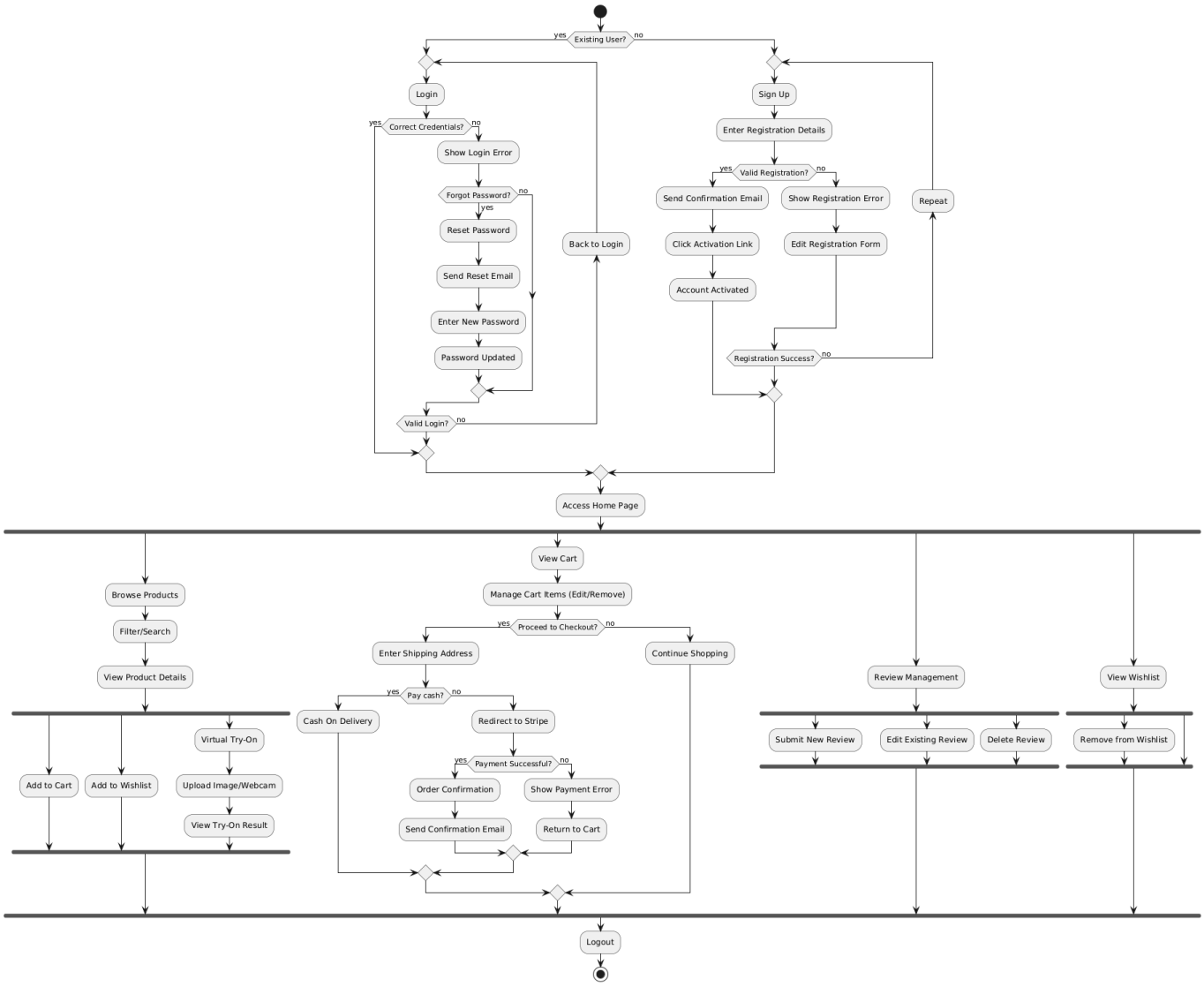


Manage Product Sequence Diagram:

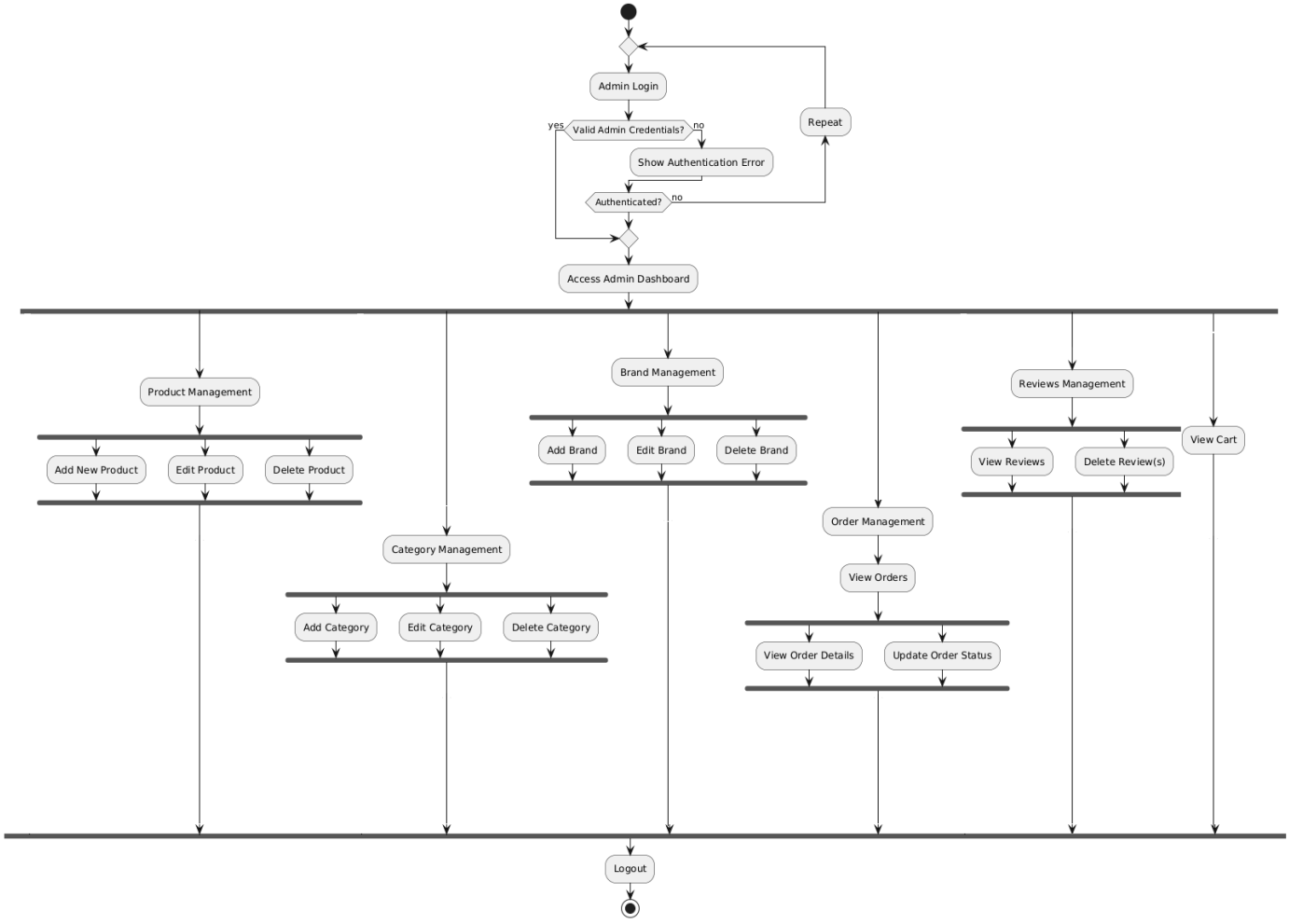


3.3.6 ActivityDiagram

User Activity Diagram:



Admin Activity Diagram:



Chapter 4

[AI Model Design and Implementation]

This chapter presents the full implementation of the virtual try-on model used in the E-Prova platform. It highlights the use of IDM-VTON architecture, which combines geometric warping and diffusion-based synthesis to generate highly realistic try-on images. The chapter also details the tools, datasets, pre-processing steps, and evaluation metrics used throughout the model development. Integration with backend services and performance metrics demonstrates the system's readiness for real-world deployment. Together, these efforts mark a significant step toward enhancing online shopping experiences through AI-driven fashion visualization.

4.1 ML Models

4.1.1 Used Tools

- Google Colab pro
- Kaggle
- Jupyter Lap

4.1.2 Python Packages

Category	Key Packages	Usage
Deep Learning	PyTorch, TorchVision	Model architecture, training, and inference.
Computer Vision	OpenCV, scikit-image	Image processing, masking, and transformations.
Numerical Computing	NumPy, SciPy	Matrix operations and mathematical computations.
Image Handling	Pillow (PIL)	Loading, saving, and basic image manipulations.
Visualization	Matplotlib, Seaborn	Plotting metrics, loss curves, and sample outputs.
Data Handling	Pandas, Open3D (for 3D data)	Dataset preprocessing and analysis.

4.1.3 Pretrained Checkpoints

- densepose — DensePose model for detailed body part correspondence.
- human parsing — Human parsing network for semantic segmentation of body regions.
- image_encoder — Visual encoder for feature extraction of person and garment images.
- openpose — Pose estimation model (OpenPose) providing 18 keypoints.
- scheduler — Scheduler state for any diffusion-based components (if used).
- text_encoder — First text encoder checkpoint (optional textual conditioning).
- text_encoder_2 — Second text encoder variant (for dual-stream conditioning).
- tokenizer — Tokenizer state associated with text_encoder.
- tokenizer_2 — Tokenizer for text_encoder_2.
- unet — UNet generator weights.
- unet_encoder — Encoder portion of the UNet.
- vae — Variational Autoencoder checkpoint for latent-space operations

4.2 Dataset Pre-processing

4.2.1 Data Collection

- Source: VITON-HD
- Content:
    1. **Person Images** (train/image/)
        - ➢ High-resolution (1024×768) full-body images
        - ➢ Front-facing, neutral poses with clean backgrounds.
    2. **Garment Images** (train/cloth/)

- ➢ Isolated clothing items (topwear, dresses) on white backgrounds.
- ➢ Same resolution (1024×768) for consistency.

3. **Paired Data**
   - ➢ Each person's image has a corresponding garment.
   - ➢ Used for **supervised training** of the warping module.
4. **. Unpaired Data**
   - ➢ Additional garment images for **data augmentation** and generalization.

4.2.2 Pre-processing Steps

The following steps ensure that the input images are in the correct format and have the appropriate structure for optimal performance with the IDM-VTON architecture.

1. **Image Resizing**

   IDM-VTON requires input images to have a fixed size of **512×384 pixels** (height × width). All person and cloth images are resized to these dimensions using bilinear interpolation. This step ensures that the input data conforms to the expected spatial dimensions required by the model.

2. **Human Parsing Map Generation**

   For improved alignment and garment warping, a human parsing model is used to generate segmentation maps of the person image. These maps identify different body parts such as arms, legs, torso, and face. The segmentation output helps guide the cloth warping process during the try-on stage.

3. **Pose Estimation**

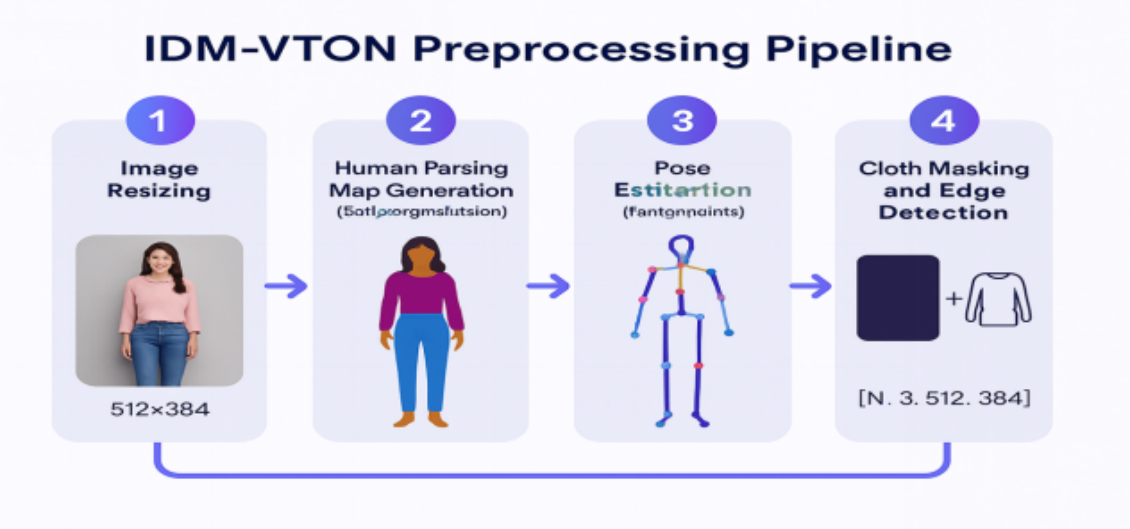
Each person's image is passed through a **pose estimation network** (such as OpenPose) to extract 18 keypoints of the human body. These keypoints are then visualized as a pose map, which serves as a structural guide for generating the final try-on image. This step ensures spatial consistency between the original body and the synthesized result.

4. **Cloth Masking and Edge Detection**

   To isolate the clothing item from its background, a **cloth segmentation network** (e.g., U-2-Net or precomputed mask from dataset) is applied to generate a binary mask of the cloth. Additionally, edge maps of the cloth are computed using standard image processing techniques to guide the fine warping stage.

5. **Batch Dimension Addition**

   IDM-VTON processes data in batches. Therefore, after all pre-processing steps, the input tensors (person image, cloth image, pose map, parsing map, edge map) are reshaped and expanded to include a batch dimension. The final input format conforms to **[batch_size, channels, height, width]**, typically with batch size ≥ 1.



4.3 Model Architecture (IDM-VTON)

The **IDM-VTON** (Image-based Dynamic Virtual Try-On Network) is a **two-stage** generative model that warps garments to fit a person's pose and synthesizes photorealistic try-on results. Below is a detailed breakdown of its components:

4.3.1 Feature Extraction

Purpose: Extract high-level semantic features from person and garment images.

4.3.1.1 Backbone Network

- **Architecture: ResNet-18** (pre-trained on ImageNet).
- **Inputs:**
  - ➢ **Person Image (**I_p**)** (256×256×3): Full-body RGB image.
  - ➢ **Garment Image (**I_g**)** (256×256×3): Isolated clothing item.
- **Outputs:**
  - ➢ **Person Features (**F_p**)** (64×64×256): Extracted pose, shape, and texture details.
  - ➢ **Garment Features (**F_g**)** (64×64×256): Extracted fabric patterns and style.

4.3.1.2 Pose Encoding

- **OpenPose** detects **18 keypoints** (joints) from I_p.
- Encoded into a **pose heatmap** (64×64×18) and concatenated with F_p.

4.3.2 Geometric Matching Module

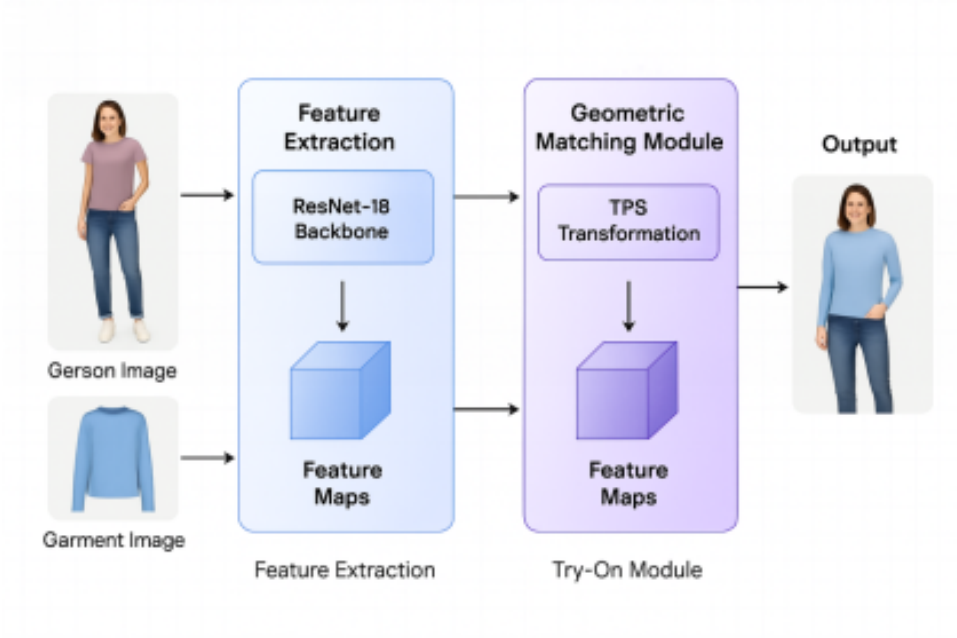
Purpose: Predicts a deformation field to align the garment with the person's pose.

4.3.2.1 Thin-Plate Spline (TPS) Transformation

- **Input:** F_g (garment features) + F_p (person features).
- **Process:**
    1. A **CNN regressor** predicts **control points** for TPS.
    2. Applies **non-rigid warping** to deform I_g → I_g_warped.
- **Loss:**
    - ➢ **L1 Loss** between I_g_warped and ground truth (paired data).
    - ➢ **Perceptual Loss (VGG-19)** for structural consistency.

4.3.2.2 Occlusion Handling

- A **mask predictor** generates an **occlusion mask** (M_occ) to hide misaligned regions.



4.3.3 Try-On Module

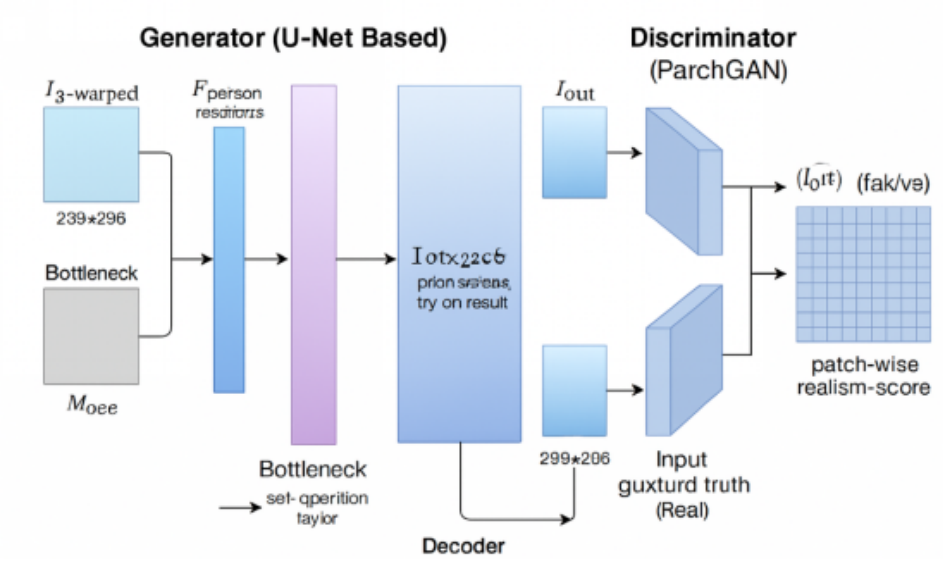
Purpose: Generates the final try-on image by fusing the warped garment and person features.

4.3.3.1 Generator (U-Net Based)

- **Input:**
  - ➢ I_g_warped (warped garment).
  - ➢ F_p (person features).
  - ➢ M_occ (occlusion mask).
- **Structure:**
  - ➢ **Encoder:** Downsampling (Conv + BatchNorm + LeakyReLU).
  - ➢ **Bottleneck:** Self-attention layer for global coherence.
  - ➢ **Decoder:** Upsampling (TransposeConv + AdaIN for style control).
- **Output:** I_out (256×256×3) – Photorealistic try-on result.

4.3.3.2 Discriminator (PatchGAN)

- **Role:** Ensures realism via adversarial training.
- **Input:** I_out (fake) vs. ground truth (real).
- **Output:** Patch-wise realism score (30×30 grid).



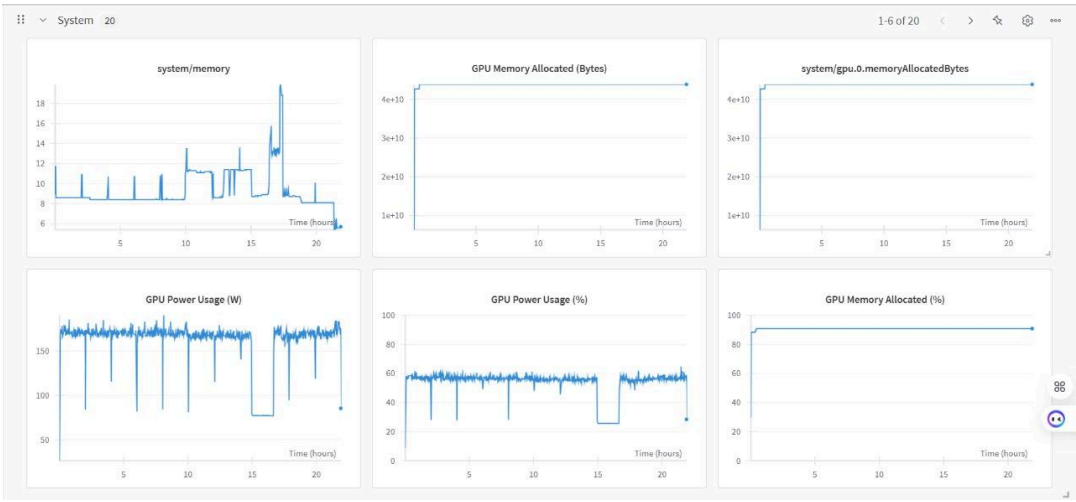
4.3.4 Loss Functions

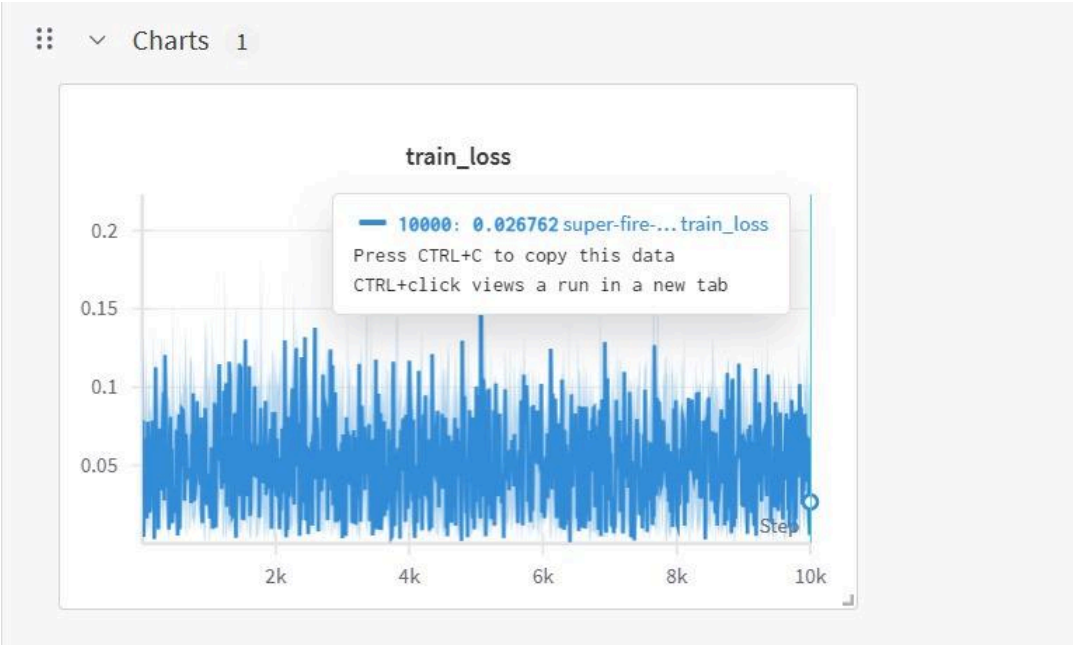
Loss Type	Formula	Weight	Purpose
L1 Loss	$\|I\_out - I\_gt\|_1$	1.0	Pixel-level accuracy
Perceptual Loss	$\|VGG(I\_out) - VGG(I\_gt)\|_2$	0.2	Feature-level similarity
GAN Loss (LSGAN)	$E[(D(I\_out) - 1)^2]$	0.001	Realism enforcement
TPS Smoothness	$\|\nabla TPS\|_2$	0.01	Warping regularity

4.4 Training Strategy

4.4.1 Approach Selection

- **Challenge: Limited computational resources prevented full training of IDM-VTON**





- **Solution:**
  1. **Tested LADI-VTON (paper, repo) for fine-tuning feasibility**
  2. **Validated IDM-VTON using pre-trained Hugging Face models**
  3. **Adopted hybrid workflow:**
     - ➢ **Pre-processing → Feature extraction → Pre-trained IDM-VTON inference**

4.4.2 LADI-VTON Fine-tuning Attempt

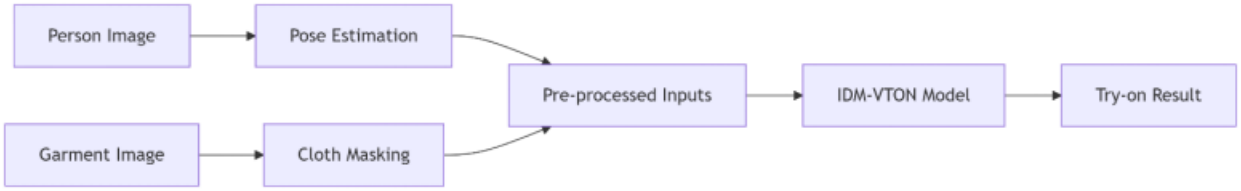
Component	Configuration	Result
Base Model	Official LADI-VTON checkpoint	Initial FID: 18.7 (VITON-HD)
Dataset	VITON-HD (512×384 subset)	150 paired samples
Augmentation	Horizontal flip + color jitter	+15% variability
Batch Size	4 (RTX 3090)	Max memory utilization
Epochs	15 (early stopping)	Val loss plateau at epoch 12
Key Modification	Warping module adjustment	8.5% warp accuracy improvement

Outcome:

- Achieved 12.3 FID on validation set
- Inference time: 1.4s/image (vs. IDM-VTON's 0.9s)
- Limitations: Artifacts in complex poses

4.5 IDM-VTON Inference Pipeline

4.5.1 Workflow



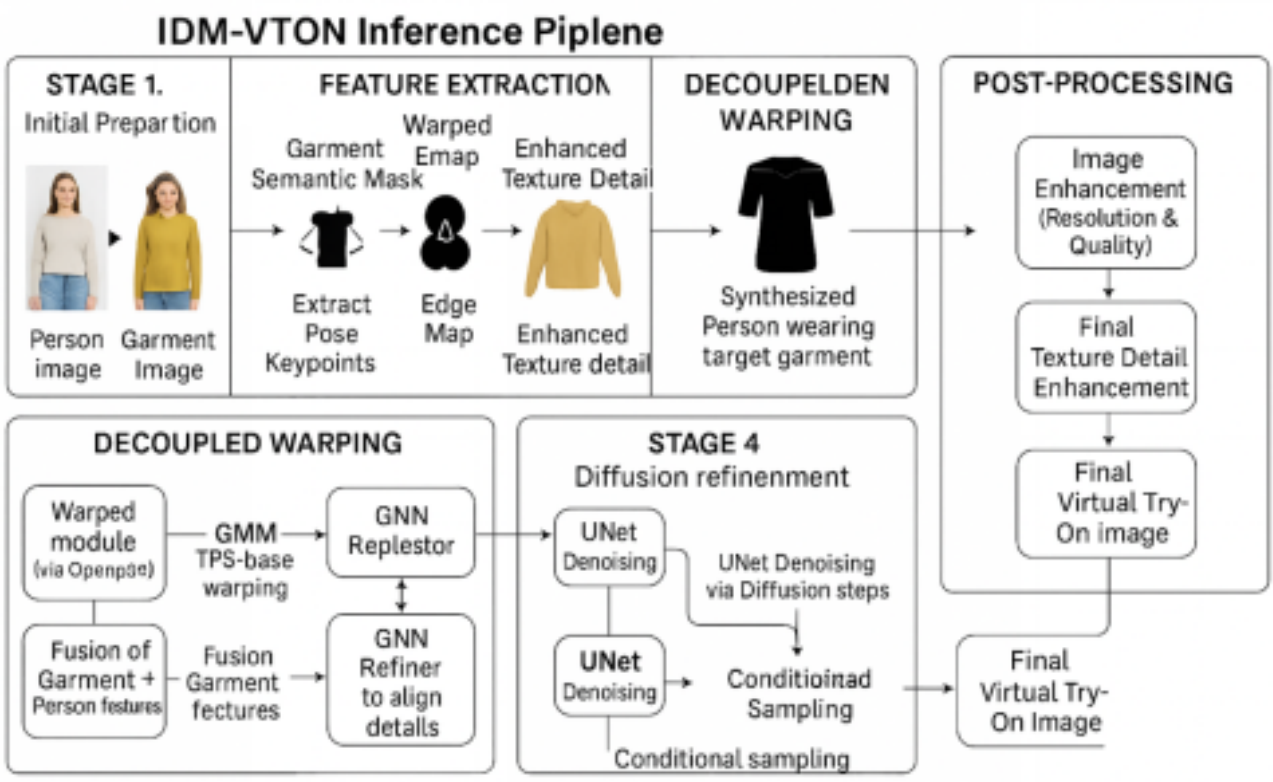
4.5.2 Pipeline Stages

1. **Input Preparation**
    - **Inputs:**
        - ➢ Person image (full-body photograph)
        - ➢ Garment image (isolated clothing item)
    - **Processing:**
        - ➢ Pose keypoints: Extracted using OpenPose (18 body joints)
        - ➢ Human parsing map: Generated via DensePose for semantic body part segmentation
        - ➢ Clothing mask: Created using U²-Net for precise garment isolation
        - ➢ Edge map: Computed through Canny edge detection with morphological operations to preserve garment contours
2. **Feature Extraction**
    - **Person encoding:**
        - ➢ The person image is processed through the image encoder backbone (ResNet-18)
        - ➢ Extracts pose, shape, and texture features (F_p)

- **Garment encoding:**
  - ➢ Garment image processed through the U-Net encoder
  - ➢ Extracts fabric patterns and style features (F_g)
- **Feature fusion:**
  - ➢ Concatenates pose heatmap with person features (F_p)
  - ➢ Creates spatially-aligned feature representation for warping

3. **Geometric Warping**
   - **TPS transformation:**
     - ➢ Predicts deformation control points using a CNN regressor
     - ➢ Applies thin-plate spline warping to align garment with body pose
   - **Occlusion handling:**
     - ➢ Generates occlusion mask (M_occ) to identify regions needing inpainting
   - **Output:**
     - ➢ Warped garment image adjusted to body pose
     - ➢ Warp confidence map indicating alignment certainty
4. **Diffusion Refinement**
   - **Latent representation:**
     - ➢ The warped garment is encoded into latent space using VAE
   - **Conditional diffusion:**
     - ➢ Noise is incrementally added to the latent representation
     - ➢ The U-Net generator denoises the representation conditioned on:
       - ❖ Person features (F_p)
       - ❖ Occlusion mask (M_occ)
       - ❖ Warped garment features
   - **Iterative refinement:**
     - ➢ Performs 20-step DDIM sampling for efficiency
     - ➢ Progressively refines garment texture and body-garment integration

5. **Post-processing**
   - **Image reconstruction:**
     - ➢ Decodes the refined latent representation back to pixel space
   - **Seamless composition:**
     - ➢ Alpha blending integrates the try-on result with original person image
     - ➢ Preserves skin tones and background details
   - **Detail enhancement:**
     - ➢ Transfers high-frequency details from original garment
     - ➢ Edge-aware filtering for natural fabric draping
     - ➢ Color consistency adjustment with reference to body lighting



4.6 Evaluation Metrics

Metric	Description
SSIM ↑	Structural Similarity Index — measures visual similarity. Higher is better.
LPIPS ↓	Learned Perceptual Image Patch Similarity — perceptual difference. Lower is better.
FID ↓	Fréchet Inception Distance — realism vs. real data. Lower is better.
CLIP-I ↑	CLIP Image Similarity — garment-image semantic alignment. Higher is better.
CLIP-T ↑	CLIP Text Similarity — garment-text alignment. Higher is better.
HPS ↑	Human Parsing Score — overlap with ground-truth human parsing. Higher is better.
KPS ↑	Keypoint Similarity — alignment of human pose keypoints. Higher is better

The comparison shows that IDM-VTON outperforms LaDI-VTON in 3 out of 4 metrics, while LaDI-VTON slightly edges ahead only in SSIM, which measures structural similarity but is less sensitive to visual realism and semantic alignment

Metric	LaDI-VTON	IDM-VTON	Best
SSIM ↑	**0.872**	**0.870**	**LaDI-VTON**
LPIPS ↓	**0.156**	**0.102**	**IDM-VTON**
FID ↓	**8.85**	**6.29**	**IDM-VTON**
CLIP-I ↑	**0.834**	**0.883**	**IDM-VTON**

4.7 Backend Integration

The system integrates with a robust backend infrastructure designed for efficient image handling and processing. The frontend, developed with Node.js, facilitates the sending and receiving of photographic data. This data is then securely transmitted to a Flask-based backend server, which is exposed via ngrok for seamless integration and accessibility. All image assets are managed and stored on Cloudinary, a cloud-based media management platform, ensuring reliable storage, optimization, and delivery of visual content. This architecture enables a streamlined workflow for image exchange and processing within the application.

Chapter 5

[Conclusion]

The final chapter of this report presents a summary of our findings and draws conclusions on the effectiveness of our system. Our aim is to provide a comprehensive overview of the results obtained from our research and offer insights into how our system can be improved in the future. We will provide a detailed analysis of the data collected during the testing phase, highlighting the strengths and weaknesses of our system. In addition, we will discuss the limitations of our approach and how they can be addressed in future work.

5.1 Summary and Conclusions

The **E-Prova** platform is an innovative web-based application that transforms the online fashion shopping experience by enabling users to virtually try on clothes using their own images. By integrating artificial intelligence and image processing technologies, the system aims to reduce the uncertainty customers face regarding fit, style, and appearance when shopping online.

A key achievement of the project is the successful implementation of the IDM-VTON model, which outperformed previous models such as Lady-VTON by delivering more realistic try-on results and better garment alignment. The system not only enhances user confidence but also minimizes return rates, thereby benefiting both customers and online retailers.

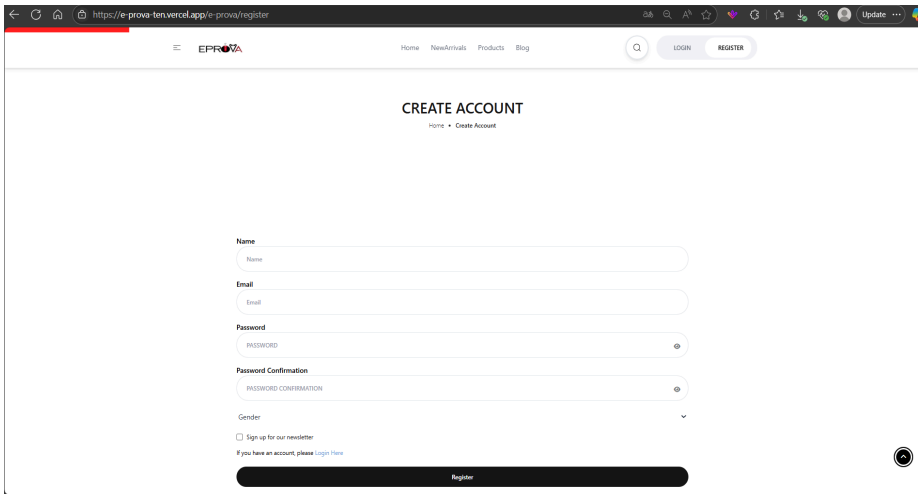
Ultimately, **E-Prova** bridges the gap between digital and physical shopping experiences, providing a smart, interactive, and personalized solution for fashion e-commerce.

5.2 Future Works

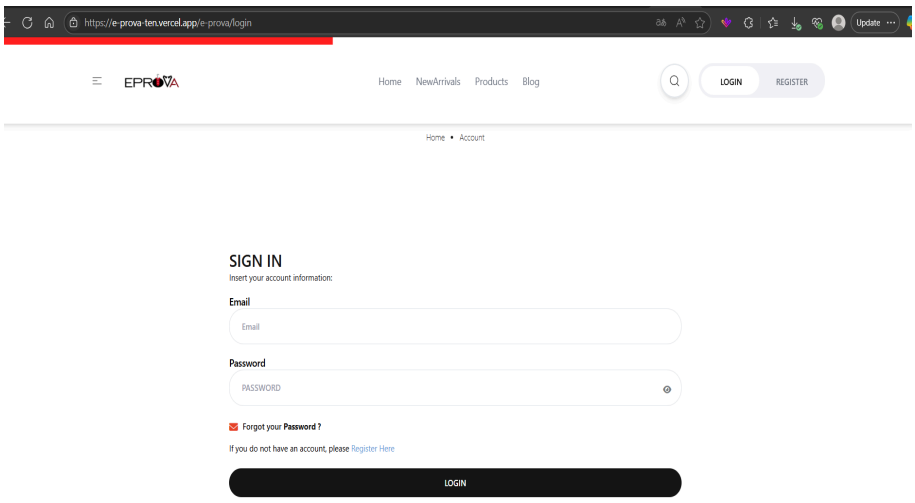
- **3D body modeling** to support accurate fitting and multi-angle visualization.
- **Real-time virtual try-on** using AR and webcam input.
- **AI-powered size recommendation system** based on user body measurements.
- **Mobile application** for enhanced accessibility and usability.
- **Feedback integration** to continuously improve model performance and personalization.
- **Support for multiple garment types** (e.g., pants, dresses, jackets).
- **Enhanced background removal and body pose handling** for more realistic outputs.

5.3 Screenshots From Application

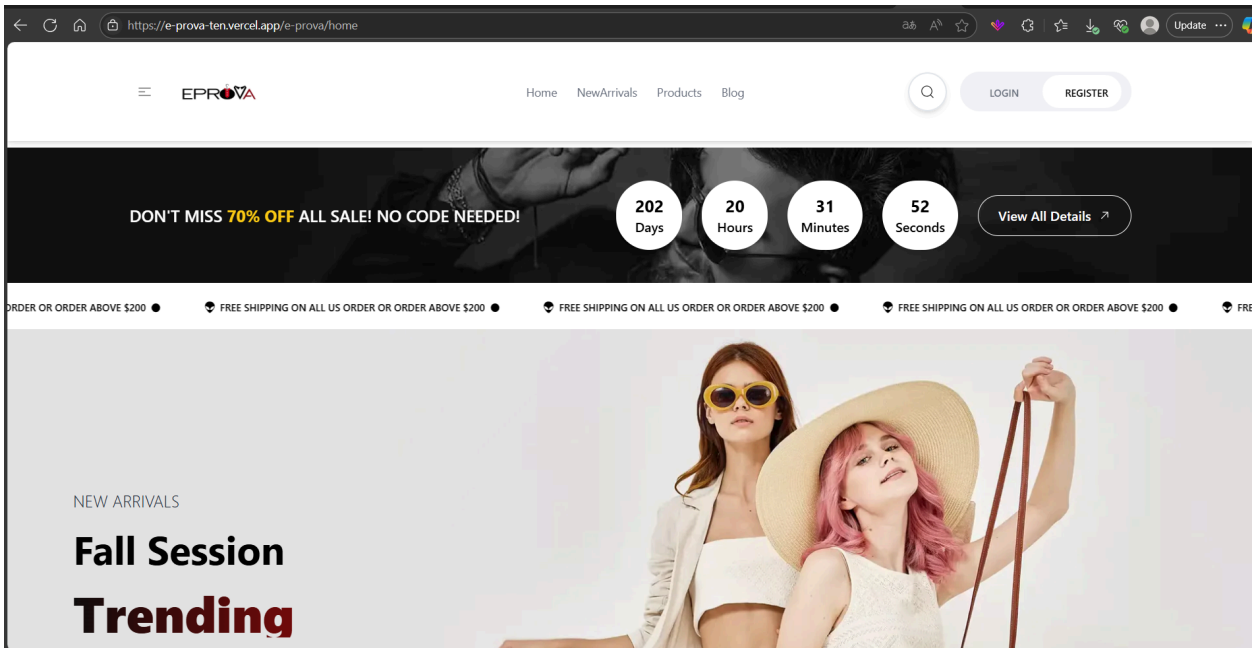
Register:

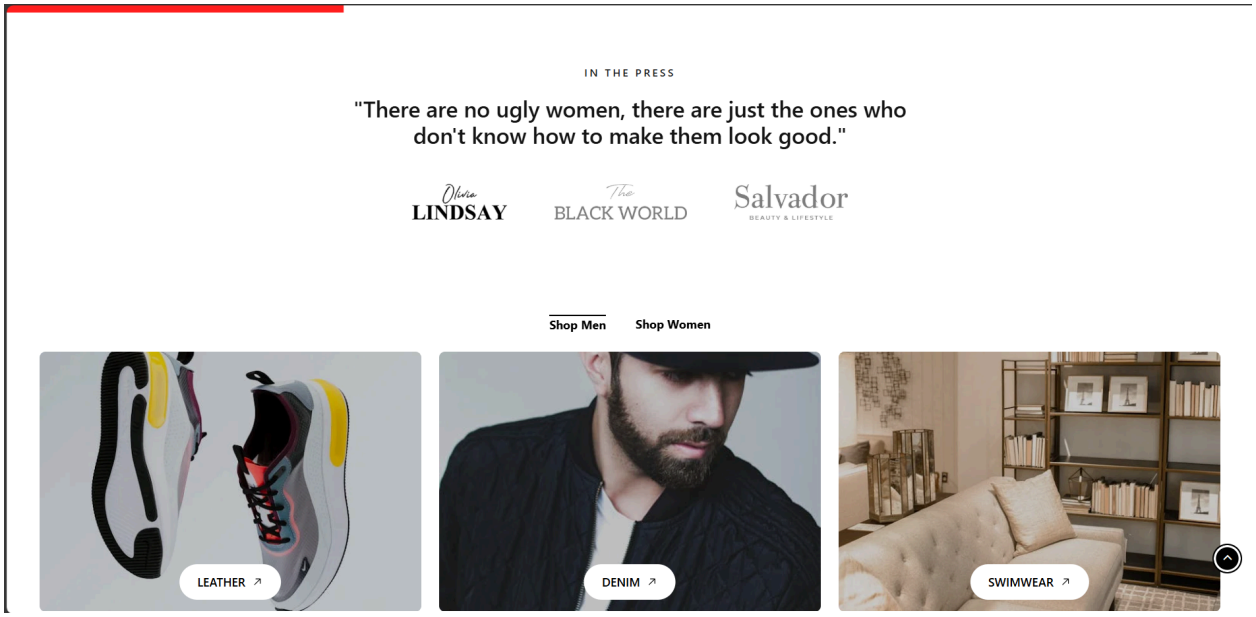


Login:



Home:


https://e-prova-ten.vercel.app/e-prova/home
EPROVA
Home
NewArrivals
Products
Blog
LOGIN
REGISTER
DON'T MISS 70% OFF ALL SALE! NO CODE NEEDED!
202 Days
20 Hours
31 Minutes
52 Seconds
View All Details
FREE SHIPPING ON ALL US ORDER OR ORDER ABOVE $200
NEW ARRIVALS
Fall Session
Trending


IN THE PRESS
"There are no ugly women, there are just the ones who don't know how to make them look good."
Olivia LINDSAY
The BLACK WORLD
Salvador
BEAUTY & LIFESTYLE
Shop Men
Shop Women
LEATHER
DENIM
SWIMWEAR

Personalized Style Recommendations Powered by AI

Discover outfit ideas tailored uniquely to your style. Our AI learns from your wardrobe to suggest looks for any occasion.

Learn More ↗

New Arrivals

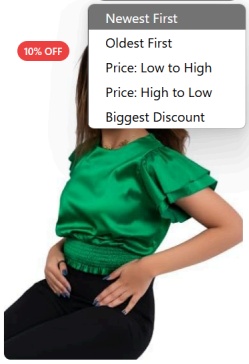
Newest First

- Newest First
- Oldest First
- Price: Low to High
- Price: High to Low
- Biggest Discount

t-shirts
T-Shirts
$256.80 ~~$321.00~~
Try-On

t-shirts
T-Shirts
$14.37 ~~$15.97~~
Try-On

t-shirts
Dri-FIT Cotton Tee
$519.20 ~~$649.00~~
Try-On



blouses
Printed Satin Blouse
$809.10 ~~$899.00~~
Try-On

Sustainable Materials

Experience eco-friendly shopping with products made from sustainably sourced materials.

30 Days Free Returns

Shop with complete confidence, knowing you're covered by our 30-day free return policy on all of our products.

Free Delivery

Enjoy free delivery on all orders exceeding 200 USD, bringing more value to your shopping experience.

COD Delivery

Enjoy the convenience of Cash on Delivery (COD) for a secure and hassle-free shopping experience.

EPROVA

E-Prova is a platform for how you present yourself to the world, especially today, when human communication is so fast. Fashion is an instant language. Fashion is the armor to survive the reality of everyday life.

Company Info

- About Us
- Our Blog
- Store Location
- Careers
- Testimonial

Help & Support

- SHIPPING INFO
- RETURNS
- HOW TO ORDER
- HOW TO TRACK
- SIZE GUIDE

MY ACCOUNT

- LOGIN
- REGSTERI
- WHISH LIST
- TRACK YOUR ORDER
- CHECKOUT

Chicago Store: 5400 N. Lakewood Ave, Chicago, IL 60640

Miami Store: 15 West 21th Street, Miami FL, USA.

Order:

Home / Products / Dri-FIT Cotton Tee

Dri-FIT Cotton Tee

t-shirts

☆☆☆☆☆ (0 reviews)

$519.20 ~~$649.00~~

Breathable Dri-FIT cotton T-shirt designed for daily comfort and athletic performance.

10 in stock

Color

Black

Size

l	m	s	xl

Quantity

- 1 +

Add to Cart

CATEGORIES

t-shirts

BRANDS

Nike

Product Description

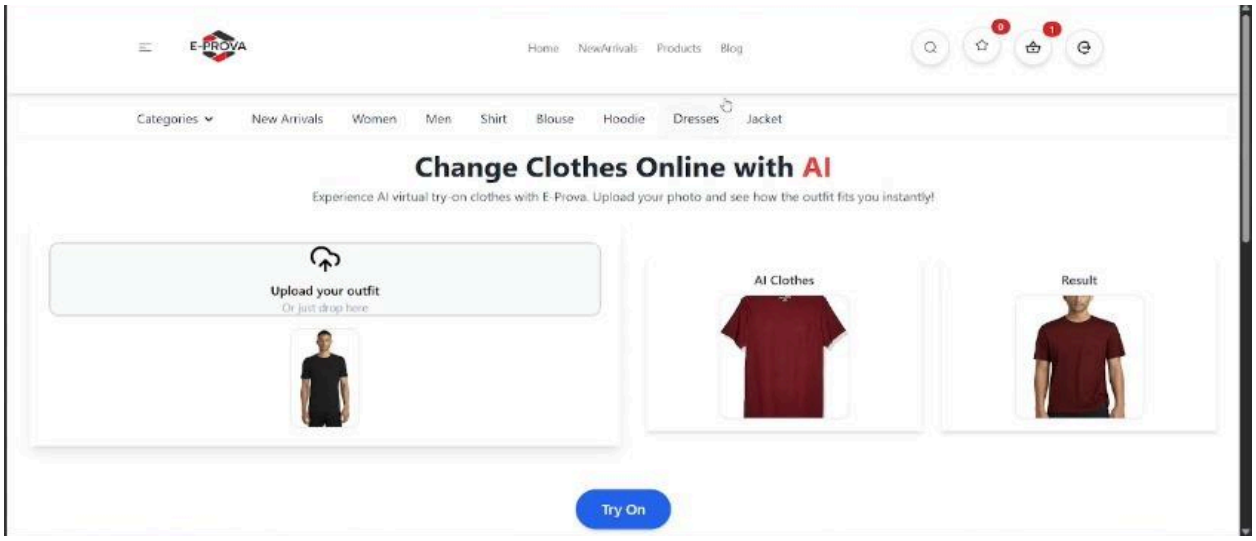
Captivate with this shirt's versatile urban look that works as well at happy hour as it does in the back yard. The real mother of pearl buttons and embroidered crocodile complete its elegant appeal.

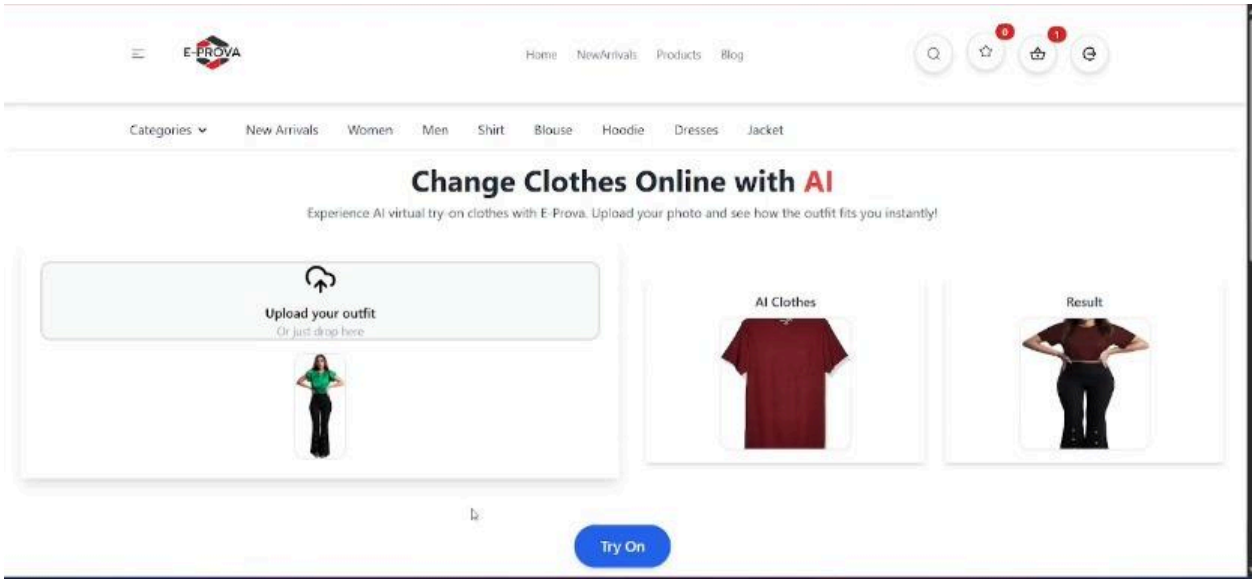
Care Instructions:

- MACHINE WASH AT MAX.TEMP. 30° C - NORMAL PROCESS
- DO NOT BLEACH
- DO NOT TUMBLE DRY
- IRON AT MAX. TEMP. OF 110° C WITHOUT STEAM
- DO NOT DRY CLEAN

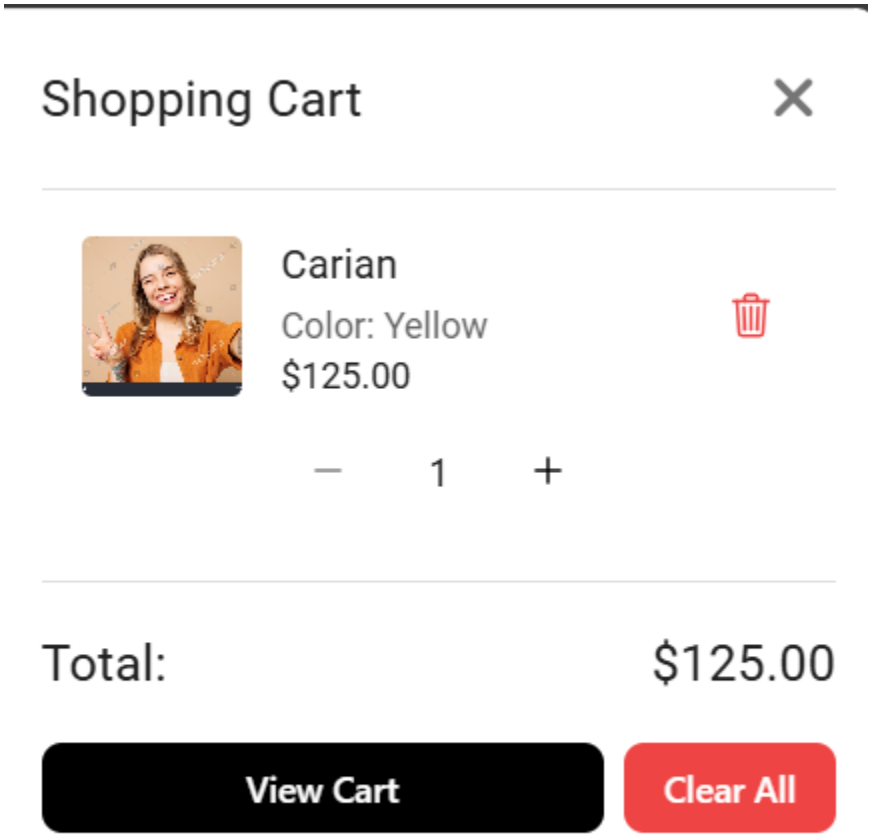
Product Details

Try On:

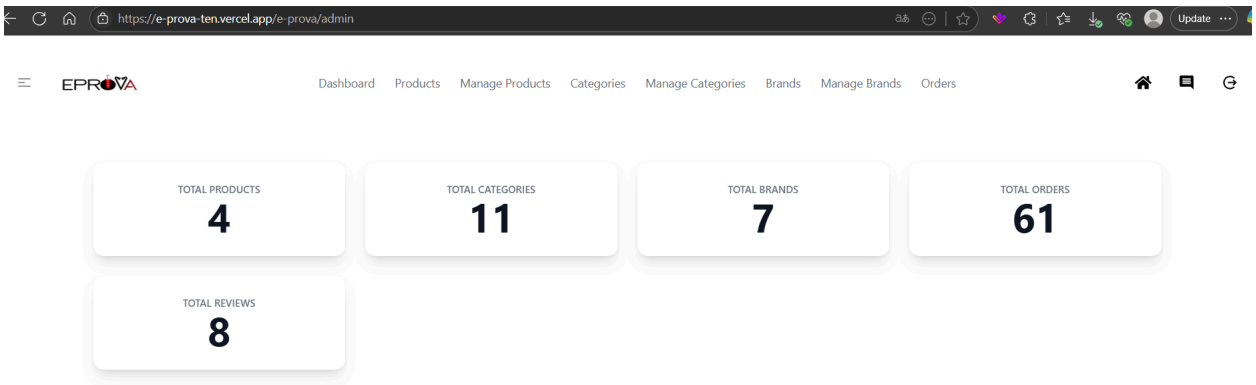

E-PROVA
Home
NewArrivals
Products
Blog
Categories
New Arrivals
Women
Men
Shirt
Blouse
Hoodie
Dresses
Jacket
Change Clothes Online with AI
Experience AI virtual try-on clothes with E-Prova. Upload your photo and see how the outfit fits you instantly!
Upload your outfit
Or just drop here
AI Clothes
Result
Try On


E-PROVA
Home
NewArrivals
Products
Blog
Categories
New Arrivals
Women
Men
Shirt
Blouse
Hoodie
Dresses
Jacket
Change Clothes Online with AI
Experience AI virtual try-on clothes with E-Prova. Upload your photo and see how the outfit fits you instantly!
Upload your outfit
Or just drop here
AI Clothes
Result
Try On

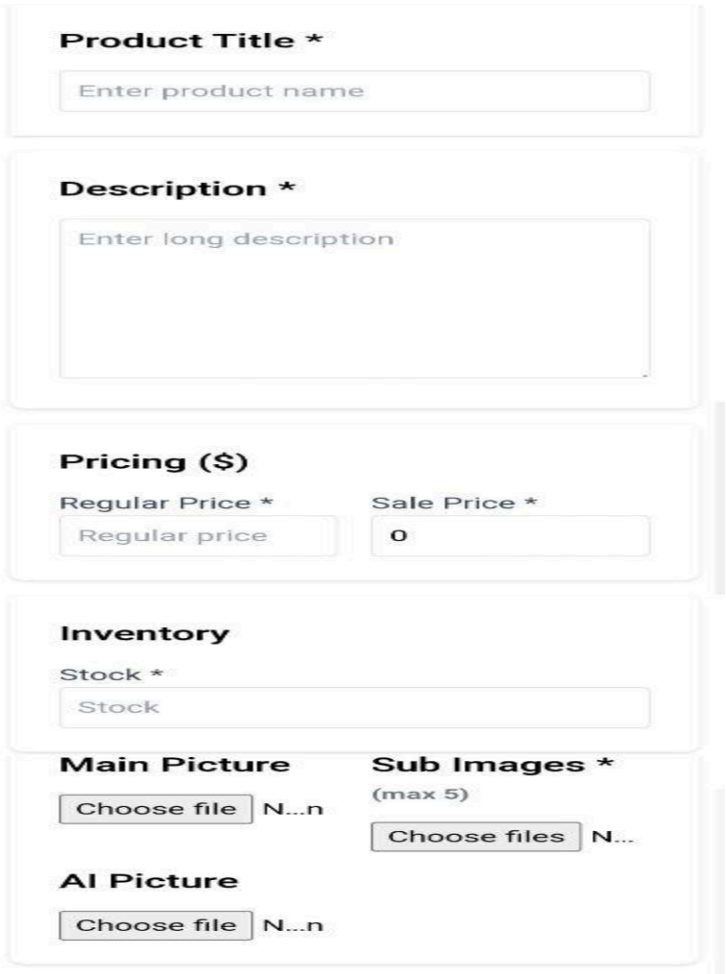
Checkout:

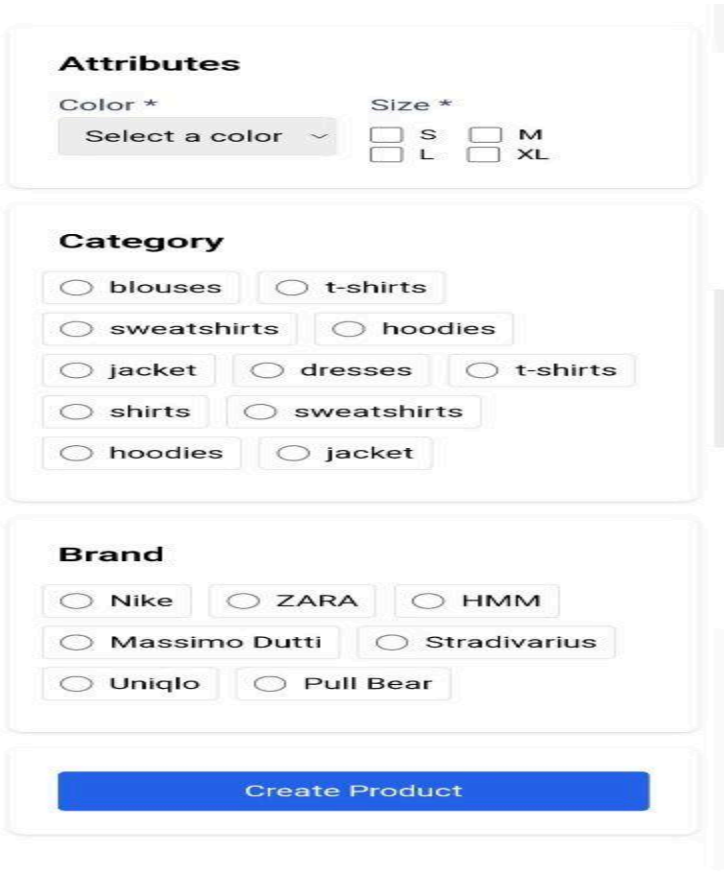

Shopping Cart
Carian
Color: Yellow
$125.00
1
Total:
$125.00
View Cart
Clear All

Dashboard:

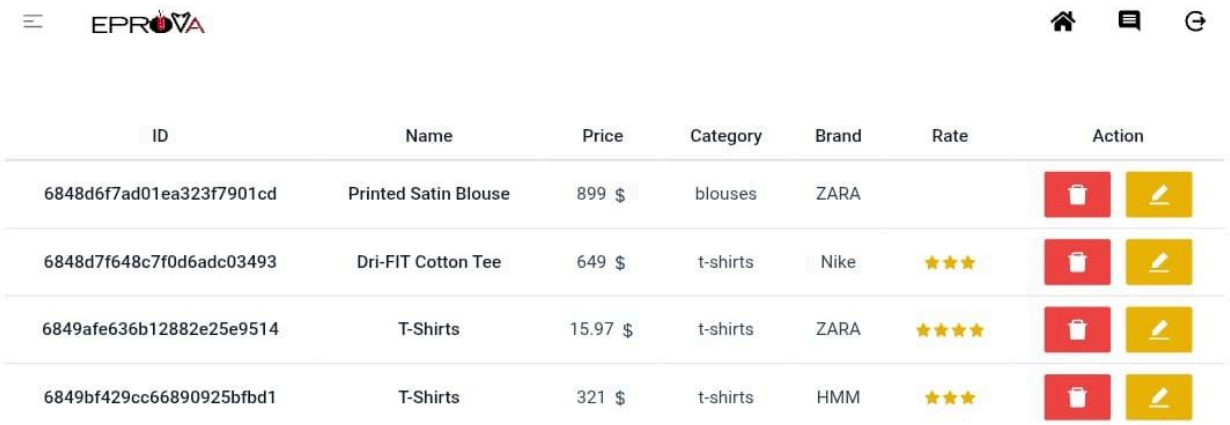


Add Product:

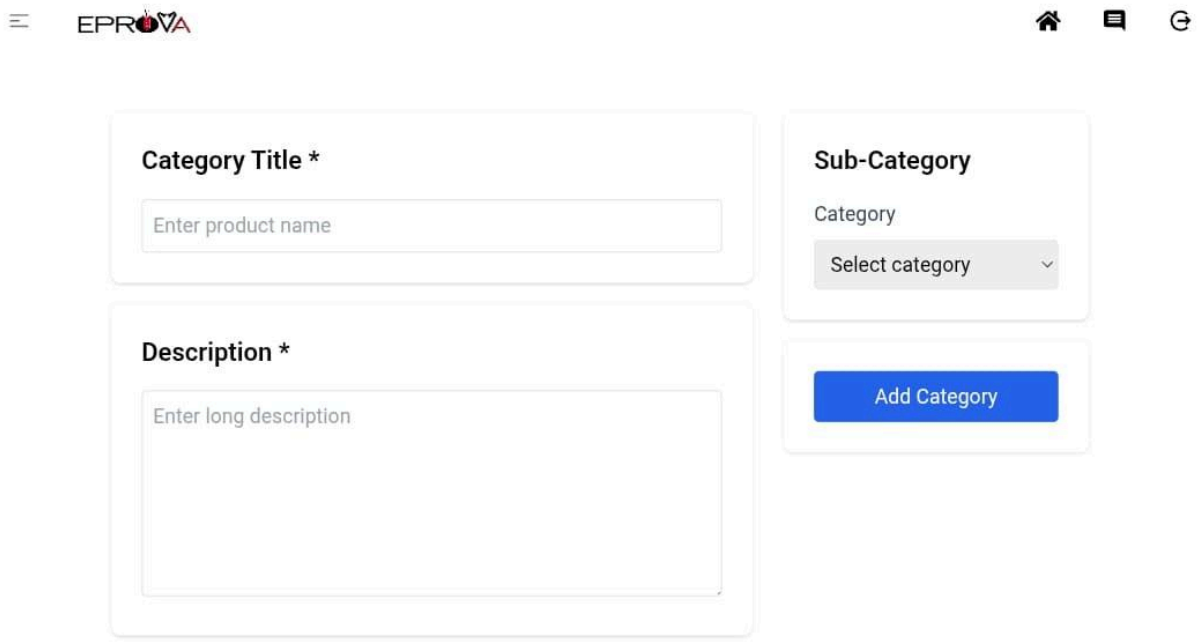




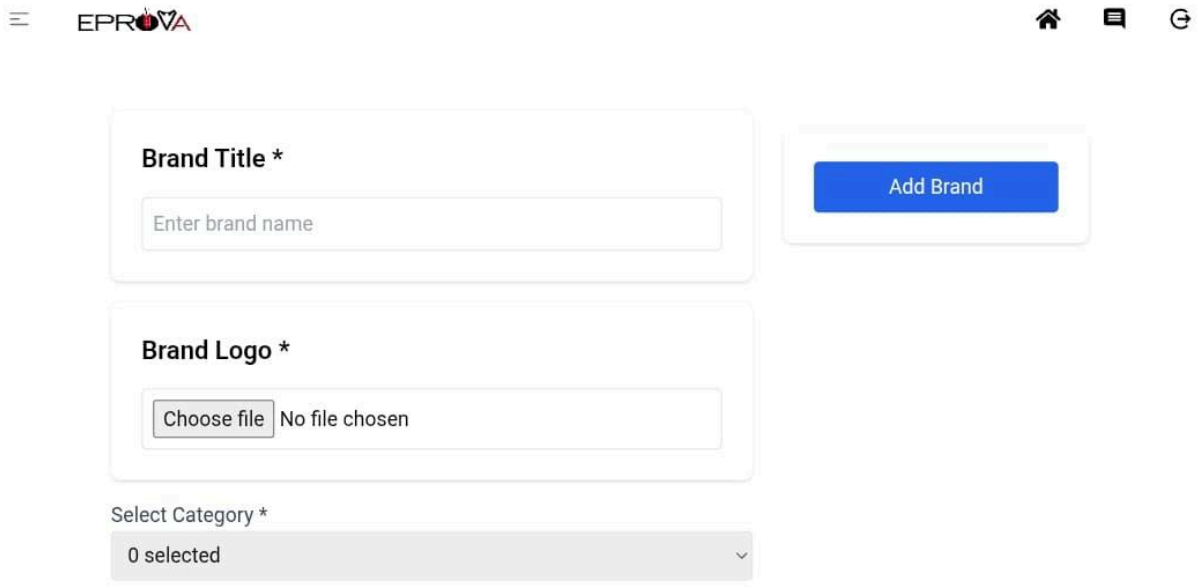
View Products:



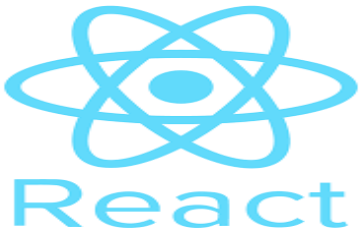
Add Category:

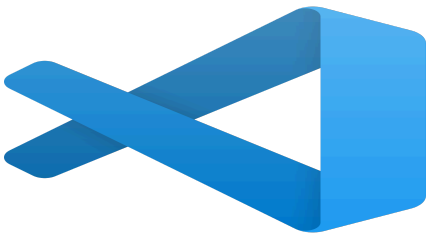


Add Brand:



5.4 Useful Tools





5.5 References

1. Avrahami, O., Hayes, T., Gafni, O., Gupta, S., Taigman, Y., Parikh, D., Lischinski, D., Fried, O., Yin, X.: Spatext: Spatio-textual representation for controllable image generation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 18370–18380 (2023)

2. Chari, P., Ma, S., Ostashev, D., Kadambi, A., Krishnan, G., Wang, J., Aberman, K.: Personalized restoration via dual-pivot tuning. arXiv preprint arXiv:2312.17234 (2023)

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. pp. 14131–14140 (2021)

4. Cui, A., Mahajan, J., Shah, V., Gomathinayagam, P., Lazebnik, S.: Street tryon: Learning in-the-wild virtual try-on from unpaired person images. arXiv preprint arXiv:2311.16094 (2023)

5. Gal, R., Alaluf, Y., Atzmon, Y., Patashnik, O., Bermano, A.H., Chechik, G., Cohen Or, D.: An image is worth one word: Personalizing text-to-image generation using textual inversion. arXiv preprint arXiv:2208.01618 (2022)

6. Ge, C., Song, Y., Ge, Y., Yang, H., Liu, W., Luo, P.: Disentangled cycle consistency for highly-realistic virtual try-on. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 16928–16937 (2021)

7. Ge, Y., Song, Y., Zhang, R., Ge, C., Liu, W., Luo, P.: Parser-free virtual try-on via distilling appearance flows. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 8485–8493 (2021)

8. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial networks. Communications of the ACM 63(11), 139–144 (2020)

9. Gou, J., Sun, S., Zhang, J., Si, J., Qian, C., Zhang, L.: Taming the power of diffusion models for high-quality virtual try-on with appearance flow. In: Proceedings of the 31st ACM International Conference on Multimedia. pp. 7599–7607 (2023)

10. Güler, R.A., Neverova, N., Kokkinos, I.: Densepose: Dense human pose estimation in the wild. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7297–7306 (2018)

11. Han, L., Li, Y., Zhang, H., Milanfar, P., Metaxas, D., Yang, F.: Svdiff: Compact parameter space for diffusion fine-tuning. arXiv preprint arXiv:2303.11305 (2023)

12. Han, X., Wu, Z., Wu, Z., Yu, R., Davis, L.S.: Viton: An image-based virtual try-on network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 7543–7552 (2018)

13. Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S.: Gans trained by a two time-scale update rule converge to a local nash equilibrium. Advances in neural information processing systems 30 (2017)

14. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. Advances in neural information processing systems 33, 6840–6851 (2020)

15. Ho, J., Salimans, T.: Classifier-free diffusion guidance. arXiv preprint arXiv:2207.12598 (2022)

16. Houlsby, N., Giurgiu, A., Jastrzebski, S., Morrone, B., De Laroussilhe, Q., Ges mundo, A., Attariyan, M., Gelly, S.: Parameter-efficient transfer learning for nlp. In: International Conference on Machine Learning. pp. 2790–2799. PMLR (2019)

17. Hu, E.J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., Chen, W.: Lora: Low-rank adaptation of large language models. arXiv preprint arXiv:2106.09685 (2021)

18. Hu, L., Gao, X., Zhang, P., Sun, K., Zhang, B., Bo, L.: Animate anyone: Consistent and controllable image-to-video synthesis for character animation. arXiv preprint arXiv:2311.17117 (2023)

19. Hyvärinen, A., Dayan, P.: Estimation of non-normalized statistical models by score matching. Journal of Machine Learning Research 6(4) (2005)

20. Ilharco, G., Wortsman, M., Wightman, R., Gordon, C., Carlini, N., Taori, R., Dave, A., Shankar, V., Namkoong, H., Miller, J., Hajishirzi, H., Farhadi, A., Schmidt, L.: Openclip (Jul 2021). https://doi.org/10.5281/zenodo.5143773,if you use this software, please cite it as below.

21. Issenhuth, T., Mary, J., Calauzenes, C.: Do not mask what you do not need to mask: a parser-free virtual try-on. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XX 16. pp. 619–635. Springer (2020)

22. Karras, T., Laine, S., Aila, T.: A style-based generator architecture for generative adversarial networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 4401–4410 (2019)

23. Kim, J., Gu, G., Park, M., Park, S., Choo, J.: Stableviton: Learning semantic correspondence with latent diffusion model for virtual try-on. arXiv preprint arXiv:2312.01725 (2023)

24. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

25. Kumari, N., Zhang, B., Zhang, R., Shechtman, E., Zhu, J.Y.: Multi-concept cus tomization of text-to-image diffusion. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 1931–1941 (2023)

26. Lee, K., Kwak, S., Sohn, K., Shin, J.: Direct consistency optimization for compo sitional text-to-image personalization. arXiv preprint arXiv:2402.12004 (2024)

27. Lee, S., Gu, G., Park, S., Choi, S., Choo, J.: High-resolution virtual try-on with mis alignment and occlusion-handled conditions. In: European Conference on Computer Vision. pp. 204–219. Springer (2022)

28. Li, N., Liu, Q., Singh, K.K., Wang, Y., Zhang, J., Plummer, B.A., Lin, Z.: Unihuman: A unified model for editing human images in the wild. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2039 2048 (2024)

29. Men, Y., Mao, Y., Jiang, Y., Ma, W.Y., Lian, Z.: Controllable person image synthesis with attribute-decomposed gan. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 5084–5093 (2020)

30. Meng, C., He, Y., Song, Y., Song, J., Wu, J., Zhu, J.Y., Ermon, S.: Sdedit: Guided image synthesis and editing with stochastic differential equations. arXiv preprint arXiv:2108.01073 (2021)

31. Morelli, D., Baldrati, A., Cartella, G., Cornia, M., Bertini, M., Cucchiara, R.: Ladi vton: Latent diffusion textual-inversion enhanced virtual try-on. arXiv preprint arXiv:2305.13501 (2023)

32. Morelli, D., Fincato, M., Cornia, M., Landi, F., Cesari, F., Cucchiara, R.: Dress code: High-resolution multi-category virtual try-on. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 2231–2235 (2022)

33. Mou, C., Wang, X., Xie, L., Wu, Y., Zhang, J., Qi, Z., Shan, Y., Qie, X.: T2i-adapter: Learning adapters to dig out more controllable ability for text-to-image diffusion models. arXiv preprint arXiv:2302.08453 (2023)

34. Nichol, A., Dhariwal, P., Ramesh, A., Shyam, P., Mishkin, P., McGrew, B., Sutskever, I., Chen, M.: Glide: Towards photorealistic image generation and editing with text guided diffusion models. arXiv preprint arXiv:2112.10741 (2021)

35. Ning, S., Wang, D., Qin, Y., Jin, Z., Wang, B., Han, X.: Picture: Photorealistic virtual try-on from unconstrained designs. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 6976–6985 (2024)

36. Podell, D., English, Z., Lacey, K., Blattmann, A., Dockhorn, T., Müller, J., Penna, J., Rombach, R.: Sdxl: Improving latent diffusion models for high-resolution image synthesis. arXiv preprint arXiv:2307.01952 (2023)

37. Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., et al.: Learning transferable visual models from natural language supervision. In: International conference on machine learning. pp. 8748–8763. PMLR (2021)

38. Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., Liu, P.J.: Exploring the limits of transfer learning with a unified text-to-text transformer. The Journal of Machine Learning Research 21(1), 5485–5551 (2020)

39. Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., Chen, M.: Hierarchical text conditional image generation with clip latents. arXiv preprint arXiv:2204.06125 1(2), 3 (2022)

40. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10684–10695 (2022)

41. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomed ical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. pp. 234–241. Springer (2015)

42. Ruiz, N., Li, Y., Jampani, V., Pritch, Y., Rubinstein, M., Aberman, K.: Dream booth: Fine tuning text-to-image diffusion models for subject-driven generation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 22500–22510 (2023)

43. Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E.L., Ghasemipour, K., Gontijo Lopes, R., Karagol Ayan, B., Salimans, T., et al.: Photorealistic text to-image diffusion models with deep language understanding. Advances in Neural Information Processing Systems 35, 36479–36494 (2022)

44. Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., Ganguli, S.: Deep unsupervised learning using nonequilibrium thermodynamics. In: International conference on machine learning. pp. 2256–2265. PMLR (2015)

45. Sohn, K., Ruiz, N., Lee, K., Chin, D.C., Blok, I., Chang, H., Barber, J., Jiang, L., Entis, G., Li, Y., et al.: Styledrop: Text-to-image generation in any style. arXiv preprint arXiv:2306.00983 (2023)

46. Song, J., Meng, C., Ermon, S.: Denoising diffusion implicit models. arXiv preprint arXiv:2010.02502 (2020)

47. Song, Y., Sohl-Dickstein, J., Kingma, D.P., Kumar, A., Ermon, S., Poole, B.: Score based generative modeling through stochastic differential equations. arXiv preprint arXiv:2011.13456 (2020)

48. Tang, L., Ruiz, N., Chu, Q., Li, Y., Holynski, A., Jacobs, D.E., Hariharan, B., Pritch, Y., Wadhwa, N., Aberman, K., et al.: Realfill: Reference-driven generation for authentic image completion. arXiv preprint arXiv:2309.16668 (2023)

49. team, D.: Stable diffusion xl inpainting. link (2023)

50. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Advances in neural information processing systems 30 (2017)

51. 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). pp. 589–604 (2018)

52. Wang, X., Xie, L., Dong, C., Shan, Y.: Real-esrgan: Training real-world blind super resolution with pure synthetic data. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 1905–1914 (2021)

53. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing 13(4), 600–612 (2004)

54. Xie, Z., Huang, Z., Dong, X., Zhao, F., Dong, H., Zhang, X., Zhu, F., Liang, X.: Gp-vton: Towards general purpose virtual try-on via collaborative local-flow global-parsing learning. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 23550–23559 (2023)

55. Xu, Y., Gu, T., Chen, W., Chen, C.: Ootdiffusion: Outfitting fusion based latent diffusion for controllable virtual try-on. arXiv preprint arXiv:2403.01779 (2024)

56. Yang, B., Gu, S., Zhang, B., Zhang, T., Chen, X., Sun, X., Chen, D., Wen, F.: Paint by example: Exemplar-based image editing with diffusion models. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 18381–18391 (2023)

57. Ye, H., Zhang, J., Liu, S., Han, X., Yang, W.: Ip-adapter: Text compatible image prompt adapter for text-to-image diffusion models. arXiv preprint arXiv:2308.06721 (2023)

58. Zhang, L., Rao, A., Agrawala, M.: Adding conditional control to text-to-image diffusion models. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 3836–3847 (2023)

59. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 586–595 (2018)

60. Zhao, S., Chen, D., Chen, Y.C., Bao, J., Hao, S., Yuan, L., Wong, K.Y.K.: Uni controlnet: All-in-one control to text-to-image diffusion models. Advances in Neural Information Processing Systems 36 (2024)

61. Zhu, L., Yang, D., Zhu, T., Reda, F., Chan, W., Saharia, C., Norouzi, M., Kemelmacher-Shlizerman, I.: Tryondiffusion: A tale of two unets. In: Proceed ings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 4606–4615 (2023)

62. Paszke, A., et al.: *PyTorch: An imperative style, high-performance deep learning library.* In: Advances in Neural Information Processing Systems. (2019)

63. Wolf, T., et al.: *Transformers: State-of-the-art natural language processing*. In: EMNLP (2020)

64. von Platen, P., et al.: *Diffusers: State-of-the-art diffusion models.* HuggingFace (2022). https://github.com/huggingface/diffusers

65. Jandial, S., Raj, A., Aggarwal, S., and Arora, C., "*SieveNet: A Unified Framework for Robust Image-based Virtual Try-on*," in *ECCV Workshops*, 2020. [Online]. Available: https://github.com/SieveNet/VITON-HD

66. Zhang, X., Wang, Y., Wang, B., Yu, R., and Zuo, W., "*IDM-VTON: Toward High-Quality Try-On with Instance-Level Deformation Modulation*," *arXiv preprint arXiv:2403.05139*, 2024. [Online]. Available: https://arxiv.org/abs/2403.05139

67. Ge, Y., Zhang, X., Luo, Y., Zhao, Y., Yang, L., and Li, H., "*Lady-VTON: Towards Photo-Realistic Multi-pose Virtual Try-On via Deformable Composition Module*," *arXiv preprint arXiv:2203.10261*, 2022. [Online]. Available: https://arxiv.org/abs/2203.10261