

RAMAKRISHNA MISSION VIVEKANANDA EDUCATIONAL AND RESEARCH INSTITUTE

COMPUTER VISION

PROJECT PROPOSAL

ShopLens: AR Visual Shopping Assistant

Team Name

INSIGHT ENGINEERS

Members

SAYAN DAS

B2430035

RAIHAN UDDIN

B2430070

March 18, 2025

ShopLens is an innovative mobile application that leverages computer vision, augmented reality (AR), and retrieval-augmented generation (RAG) powered by large language models (LLMs) to transform both the seller and buyer experiences in online retail. The app automatically recognizes clothing items, overlays virtual garments, and provides personalized product recommendations by processing real-time data.

I.1 Motivation

In today's fast-paced e-commerce landscape, customers demand accurate product information and personalized recommendations. Inaccurate sizing and poor fit are among the top reasons for high return rates. ShopLens addresses these issues by:

- Automating product recognition to speed up cataloging for sellers.
- Using AR to provide precise body measurements and virtual try-ons.
- Employing RAG to ensure that chatbot recommendations are up-to-date and context-aware.

This project not only improves customer satisfaction and reduces returns but also enhances seller efficiency by minimizing manual data entry and product misclassification.

I.1.1 Workflow

Seller Perspective

1. **Product Image Upload:** The seller uploads product images to the platform.
2. **Automated Attribute Extraction:** The computer vision module automatically extracts key attributes (e.g., garment type, color, fabric, pattern).
3. **Verification and Correction:** The seller reviews and corrects any misidentified details.
4. **Data Integration:** The verified product data is indexed and stored, contributing to model improvements for future recognition.

Buyer Perspective

1. **Product Discovery:** The buyer uploads a reference image or browses the catalog.
2. **Similarity Retrieval:** The system uses the CV module and vector search to retrieve similar products.
3. **AR-based Virtual Try-On:** The AR module overlays virtual garments for an interactive try-on experience.
4. **Chatbot Assistance:** A RAG-powered chatbot offers personalized recommendations and responds to queries (e.g., suggesting alternative products based on price or style).
5. **Purchase Decision:** The buyer evaluates the personalized suggestions and proceeds with the purchase.

SECTION II

TOOLS AND LIBRARIES

- **Computer Vision:** OpenCV, TensorFlow/PyTorch.
- **Augmented Reality:** ARCore (Android).
- **LLM and RAG:** Hugging Face Transformers, LLMs (OpenAI, Gemini, Deepseek), LangChain, vector database.
- **Development Environment:** Git.

SECTION III

METHODOLOGY

The ShopLens system is designed as a modular architecture where each component performs a distinct role while seamlessly integrating with the others. The internal workflow is structured to ensure scalability, accuracy, and real-time responsiveness.

III.1 System Architecture

Our system comprises four main modules:

- **Computer Vision Module:** Automates product recognition and attribute extraction.
- **Augmented Reality Module:** Handles real-time body measurements and virtual try-on experiences.
- **LLM/RAG Module:** Powers a dynamic chatbot to provide personalized recommendations using retrieval-augmented generation.

III.2 Internal Workflow

Computer Vision Module

- **Image Preprocessing:** Uploaded product images are normalized, resized, and augmented to ensure consistency.
- **Feature Extraction:** Deep neural networks (e.g., ResNet or Vision Transformers) extract visual features and generate embeddings.
- **Attribute Classification:** Specialized classifiers analyze the embeddings to identify product type, color, fabric, pattern, and other design elements.
- **Feedback Integration:** Sellers verify and correct attributes, and the corrections are used to periodically retrain and improve the model.

Augmented Reality Module

- **Virtual Try-On:** The module overlays optimized 3D garment models onto the user’s live image stream, aligning them with real-time body tracking data.
- **User Calibration:** On-screen guides help users position themselves correctly, and feedback is used to refine the measurement algorithms.

LLM and Retrieval-Augmented Generation (RAG) Module

- **Query Processing:** User queries and product metadata are encoded into vector representations using advanced embedding models.
- **Semantic Retrieval:** A vector database (e.g., Pinecone) is queried to retrieve the most relevant product data based on semantic similarity.
- **Chatbot Interaction:** The retrieved context is combined with the original query and fed to a large language model (via platforms such as Hugging Face or OpenAI API) to generate personalized and context-aware responses.

- **Continuous Improvement:** Interaction logs and user feedback are utilized to fine-tune both the retrieval process and the LLM prompts.

SECTION IV

WORK PLAN

IV.1 Timeline (4 Weeks)

- **Week 1:** Requirements analysis, system design, and setting up the development environment.
- **Week 2:** Develop and test the computer vision and AR modules.
- **Week 3:** Develop the LLM/RAG-based chatbot and integrate with the product database.
- **Week 4:** Developing application front-end, integration testing, and final deployment.

IV.2 Work Distribution

- **Sayan Das:** Lead on computer vision and AR integration. Responsible for application front-end and deployment.
- **Raihan Uddin:** Responsible for the development of the product detection system and AR virtual try-on. Lead on LLM/RAG chatbot integration.

SECTION V

FUTURE WORK

Beyond the initial 4-week period, the following features could be added:

- **Enhanced Feedback Loop:** Improve model accuracy through continuous seller and user feedback integration.
- **Advanced AR Features:** Implement more sophisticated virtual try-on effects and precise body measurement algorithms.
- **Expanded Product Categories:** Support additional clothing types and accessories.
- **Personalization Enhancements:** Incorporate user data analytics for tailored marketing and product recommendations.
- **Cross-Platform Integration:** Extend support to multiple platforms (iOS, Android, and web).

BIBLIOGRAPHY

- [1] OpenCV: Open Source Computer Vision Library. <https://opencv.org/>
- [2] TensorFlow: An end-to-end open-source platform for machine learning. <https://www.tensorflow.org/>
- [3] Hugging Face: Transformers for Natural Language Processing. <https://huggingface.co/>
- [4] ARCore: Augmented Reality SDK by Google. <https://developers.google.com/ar>
- [5] Lai, Y., et al. ISA: An Intelligent Shopping Assistant. AACL 2020.