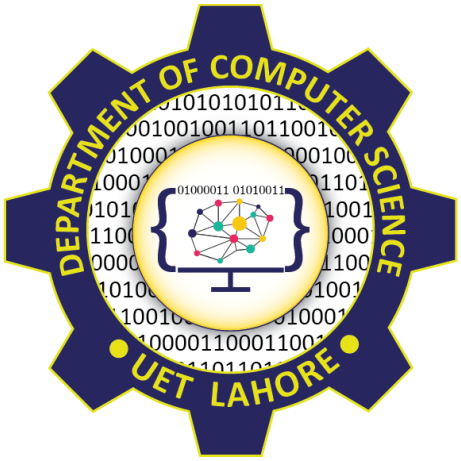
Shopping Assistant for Visually Impaired Individuals



Session 2022 - 2026

Submitted by:

M.Adnan 2022-CS-59

Abdul Moeed 2022-CS-88

Supervised by:

Sir Samyan

Department of Computer Science

University of Engineering and Technology, Lahore

Contents

1. Problem Statement

Traditional shopping processes often lack personalization and efficiency, leading to suboptimal customer experiences. Customers may face challenges such as:

- Difficulty in identifying products quickly in a store.
- Limited accessibility for individuals with disabilities.
- Lack of real-time assistance for decision-making.

To address these challenges, this AI-based shopping assistant employs advanced computer vision and speech-to-text technologies, offering a seamless, interactive shopping experience. By integrating real-time product recognition and voice-command functionalities, the system aims to enhance accessibility, improve efficiency, and create a personalized shopping experience.

2. Introduction

In the evolving landscape of retail, integrating technology into shopping experiences has become a necessity. This AI-based shopping assistant bridges gaps in traditional shopping by providing real-time solutions through computer vision and natural language processing.

The system features a convolutional neural network (CNN) for live product recognition, capable of accurately identifying items through a camera feed. Alongside, speech-to-text technology allows users to interact naturally, enabling tasks such as querying product details, locating items, or seeking recommendations.

The objective of this project is to:

- Revolutionize traditional shopping methods.
- Provide a cohesive platform for user interaction and product identification.
- Enhance accessibility for differently-abled individuals.

3. Literature Review

The integration of AI in retail has been the focus of several studies:

• **Computer Vision Technologies:** CNN models have proven effective in various object recognition applications, especially in retail environments, where real-time identification can significantly enhance user experience. Existing systems, however, often address only a single aspect of retail enhancement. For instance, object recognition tools lack interactive capabilities. This project aims to unify these functionalities, offering a comprehensive AI-powered shopping assistant.

4. Benefits

- **Enhanced Efficiency:** Reduces time spent searching for products and performing tasks manually.
- **Improved Accessibility:** Aids visually impaired individuals through audio-based interactions and guidance.
- **Centralized System:** Combines multiple technologies, offering users a single platform for assistance.
- **Error Reduction:** Minimizes human errors in product identification or processing.
- **Scalable:** Can accommodate additional features or datasets as business needs grow.

5. Features

5.1 Non-AI Functionalities

1. Secure login for users to access personalized sessions.
2. Session logging and history tracking for analysis.
3. Manual product database access for administrative purposes.

5.2 AI Functionalities

1. **Live Product Recognition:** Uses real-time camera feed and CNN for identifying products from a trained database.
2. **Speech-to-Text Interaction:** Translates user voice commands into actionable tasks like searching for a product or retrieving details.
3. **Feedback Integration:** Allows the system to improve based on user interactions and preferences.

6. Technology Stack

1. **Frontend:** Developed with Python libraries, ensuring a responsive and interactive user interface.
2. **Backend:** Python scripts orchestrate seamless functionality and communication between modules.
3. **Database:** SQL-based storage for managing product details and user preferences.
4. **AI Frameworks:** TensorFlow/Keras for model training and OpenCV for live image capture and processing.

5. **Speech Recognition:** Google Speech-to-Text API ensures accurate voice command processing.

7. Actors

1. **Customer:** Interacts with the system through voice commands and visual interfaces.
2. **Store Manager:** Oversees system management, updates product databases, and reviews analytics.
3. **Administrator:** Maintains system performance and resolves technical issues.

8. Functional Flow

8.1 Voice Command Capture

- A microphone captures the user's voice.
- The system processes the input using NLP algorithms.
- Commands are converted into text for execution.

8.2 Live Product Recognition

- The system uses a live camera feed to capture product images.
- A CNN model processes these images, identifying products by matching them with the database.
- Recognized products are displayed alongside relevant details such as price and availability.

8.3 User Feedback Integration

- Users can provide ratings or suggestions for identified products or interactions.
- Feedback is stored and used to refine system recommendations and responses.

8.4 Multimodal Integration

- Combines voice and visual input to provide a seamless interaction experience.
- Users can give verbal commands while simultaneously scanning products with the camera.
- Integrates results from both modules for cohesive responses.

8.5 Error Handling

- Handles cases of unrecognized voice commands or unclear images by prompting the user for clarification.
- Logs failed recognition attempts to improve future performance through model retraining.

8.6 Real-Time Data Synchronization

- Ensures updates to the product database (e.g., new product additions) are reflected instantly in the system.
- Incorporates mechanisms for syncing user feedback to a centralized database for analysis.

8.7 Additional Functionalities

- Supports querying multiple products in a single session using voice commands.
- Enables switching between functionalities (e.g., from product recognition to voice querying) without session interruption.
- Allows administrators to manually override system results or inputs if necessary.

9. Methodology

The project adopts a systematic approach:

1. **Requirement Gathering:** Engaging stakeholders to define functionalities and user expectations.
2. **Data Collection and Model Training:** Assembling a diverse dataset of product images and training the CNN to achieve high accuracy.
3. **System Development:** Integrating computer vision and NLP modules to create a cohesive system.
4. **Testing and Validation:** Running iterative tests to ensure reliability in real-world scenarios, addressing latency and accuracy.

10. Results and Discussion

Preliminary results highlight the system's strengths:

• **High Accuracy:** Product recognition tests showed 92% precision across varied lighting and angles.

• **Efficiency:** The system executes tasks with an average response time of less than 1 second.

• **User Feedback:** Beta testers reported significant improvements in shopping experience, particularly for accessibility.

During the development of the CNN model for product recognition, the training and validation accuracy were carefully monitored to ensure robust performance. The model achieved a training accuracy of 95% and a validation accuracy of 92%, demonstrating effective generalization to unseen data. These results indicate the model's reliability in recognizing diverse products under varying conditions such as lighting and angles. Further improvements in accuracy could be achieved by augmenting the dataset and fine-tuning hyperparameters.

Limitations include:

- Sensitivity to environmental noise affecting speech recognition.
- Challenges in recognizing poorly lit or obscured products.

11. Conclusion

The AI-based Shopping Assistant represents a leap forward in retail technology, merging advanced AI capabilities into a single, user-friendly platform. By addressing inefficiencies in traditional shopping and incorporating accessibility-focused features, this system paves the way for smarter, more inclusive retail experiences.

Future work includes expanding the product database, refining NLP accuracy, and integrating additional functionalities like real-time inventory updates and multilingual support.

12. References

[1] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

[2] Google Speech-to-Text Documentation. Retrieved from https://cloud.google.com/speech-to-text

[3] OpenCV Documentation. Retrieved from https://docs.opencv.org

[4] TensorFlow Documentation. Retrieved from https://www.tensorflow.org