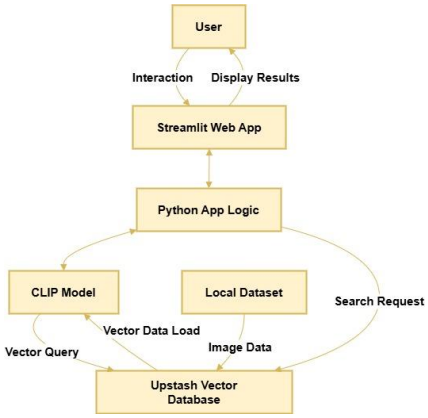
Intelligent Image & Text Retrieval System

1 System Architecture Design

1.1 Overall Architecture

This system is built on a Streamlit-based interface designed to deliver an intuitive and efficient content discovery experience through image and text queries. By integrating the CLIP machine learning model for feature extraction and a powerful vector database, the architecture ensures a seamless and intelligent search process, with particular emphasis on optimizing user interaction and interface design.



1.2 Technology Stacck

• **Frontend Framework**: Streamlit 1.44.1 • **Python Version**: 3.8+ • **Core Dependencies**: clip

2 Dataset Introduction (GroceryStoreDataset)

2.1 Source

The dataset is cloned via Git from: https://github.com/marcusklasson/GroceryStoreDataset.git

2.2 Contents

The dataset contains images of various grocery items.

• The classes.csv file provides category information, including: Class Name, Class ID, Coarse Class Name, Coarse Class ID, Product Description Path, Iconic Image Path

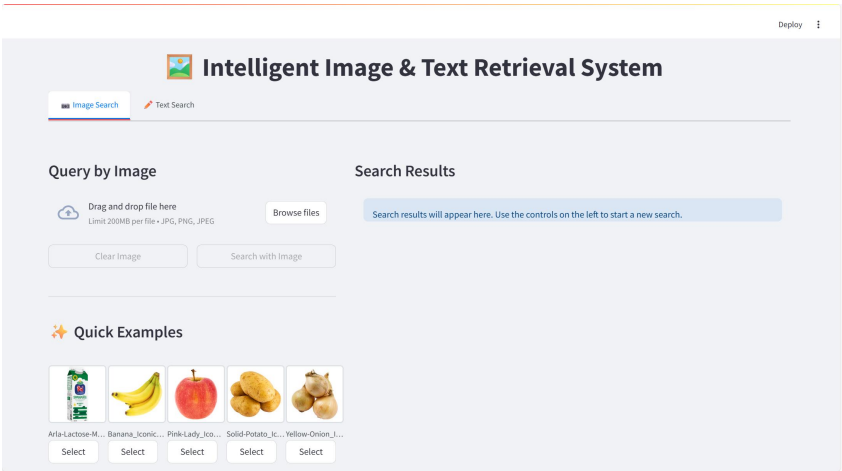
• The train.txt and val.txt files contain the image paths along with their corresponding Class IDs and Coarse Class IDs, used to associate images with their metadata.

2.3 System Integration

The load_data.py script handles the processing of these files. It reads image data and combines it with information from classes.csv (such as class names and description paths) to generate feature vectors. These vectors, along with key metadata (image path, text description path, class name, class ID, etc.), are then stored in the **Upstash vector database**.

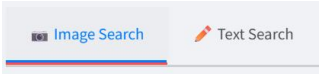
3 User Interface Design

The system's interactive interface aims to deliver a clear and intuitive image-text retrieval experience. The interface is titled "**Intelligent Image & Text Retrieval System**".



3.1. Function Navigation (Tabs)

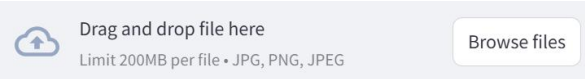
At the top of the interface, there are two main tabs: **Image Search & Text Search**. Users can switch between these modes by clicking the respective tab.



3.2 "Image Search" Tab

• **Query by Image Section**

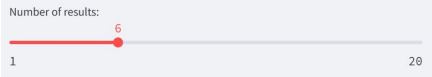
File Upload Area: Users can drag and drop an image file into the designated area or click "Browse files" to upload an image (supports JPG, PNG, JPEG).



Preview: The uploaded image is clearly displayed below the upload area.

Action Buttons: "Clear Image": Removes the currently selected image. "Search with Image": Initiates the image-based search.

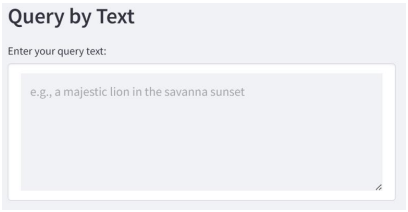
• **Result Count Slider:** Labeled "Number of results:", this slider allows users to specify the number of results to return (1–20, default is 6).



3.3 "Text Search" Tab

• **Query by Text Section**

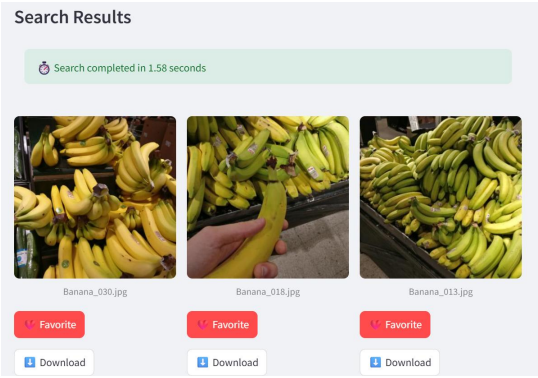
Text Input Box: A multiline text box prompting users to "Enter your query text:", with an example such as ″e.g., a majestic lion in the savanna sunset″.



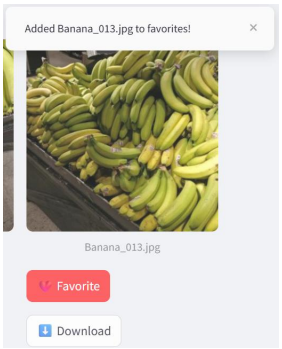
Action Button: "Search with Text": Initiates text-based search. The button is typically disabled when no text is entered.

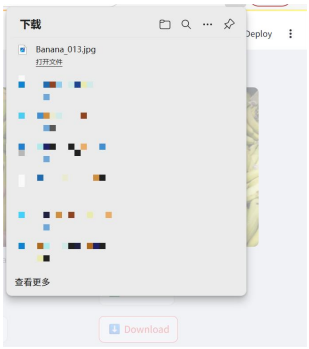
• **Result Count Slider:** Same functionality as in the image search tab.

3.4 Search Results Section



- **Search Feedback:** Displays messages such as "Search completed in 1.58 seconds".
- **Image Grid:** Results are shown in a grid (e.g., 3 columns), each image accompanied by its filename.
- **Per-Image Actions:** Each result includes: "Favorite" / "Favorited" button & "Download" button





4 Additional Features Enhancing User Experience

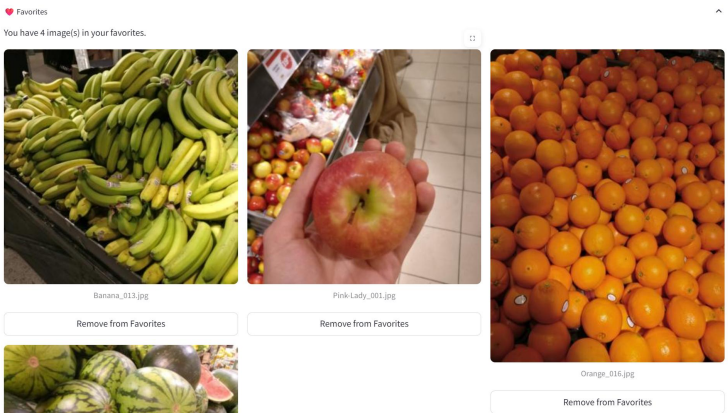
4.1 "Favorites" Function

- **Interface Integration:**

Each result image includes a "Favorite" button. Favorited images are shown in the expandable "Favorites" section at the bottom.

- **Experience Improvement:**

Users can conveniently mark and save images without downloading them immediately. This supports later review and management.



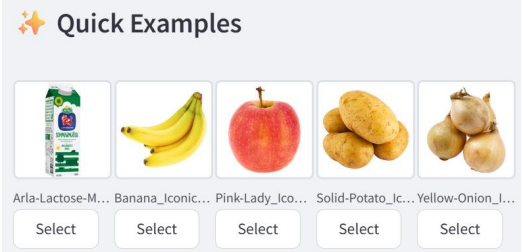
4.2 “Quick Examples” Feature

• **Interface Integration:**

Located in the "Image Search" panel, displays sample images (e.g., milk, banana) with "Select" buttons.

• **Experience Improvement:**

Great for first-time users or those without specific query images. Enables quick exploration of system functionality.

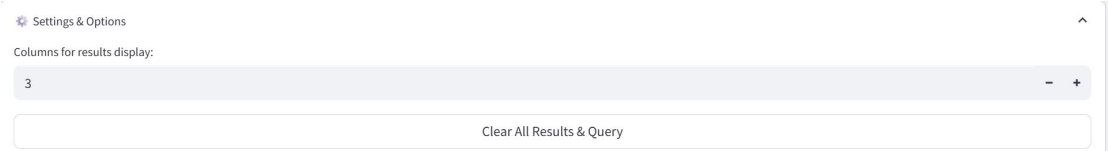


4.3 Customization via “Settings & Options”

• **Interface Integration:**

"Columns for results display:" numeric input.

"Clear All Results & Query" button.



• **Experience Improvement:**

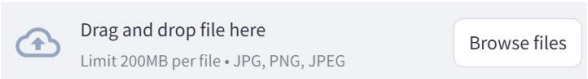
Custom Columns: Users can tailor result layout based on their screen or preference.

One-Click Reset: Enables fresh searches without clutter.

4.4 Visual Feedback and Status Indicators

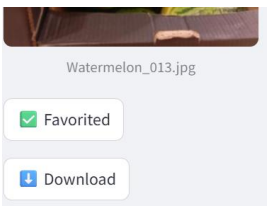
Search completion time displayed (e.g., "Search completed in 1.58 seconds")

Upload area clearly labeled with "Drag and drop file here"



Text input box includes placeholder examples

Favorite buttons change icon and text upon selection



5. User Operations and Friendliness Across Input Modes

The interface distinguishes between image-based and text-based search via separate tabs.

5.1 Impact of Input Modes on User Workflow

5.1.1 Image Input Flow (“Image Search” Tab)

- **Use Case:** Users have an image or want to try examples.
- **Steps:**

  Click "Image Search" tab.

  Upload a file or click "Select" on a quick example.

  View the image preview.

  Adjust the "Number of results" slider.

  Click "Search with Image".
- **Nature:** A visually guided, direct interaction process.

5.1.2 Text Input Flow (“Text Search” Tab)

- **Use Case:** Users want to describe what they’re looking for.
- **Steps:**

  Click "Text Search" tab.

  Enter descriptive text in the input box.

  Adjust the "Number of results" slider.

  Click "Search with Text".
- **Nature:** A concept-driven, semantic interaction process.

5.2 Making Both Input Methods Equally User-Friendly

5.2.1 Clear Functional Separation

Tabs for “Image Search” and “Text Search” ensure users easily identify and switch between search modes.

5.2.2 Unified Layout and Visual Style

- **Similar Structure:** Both modes use a left control panel + right results panel layout.
- **Consistent Display:** Results are shown in the same format for both modes.

5.2.3 Simplified Input and Actions

- **Image Mode:**

Drag-and-drop area and "Browse files" button

Clickable Quick Examples

Preview before search

- **Text Mode:**

Placeholder guides in text box

Button is only enabled when valid input is present

5.2.4 Clear Feedback Mechanisms

- Search time shown after every search
- Sliders and options affect both modes equally
- Real-time feedback on actions (previews, button state)

These design strategies ensure that whether users prefer image-based or text-based queries, they enjoy a smooth, intuitive, and consistent user experience.