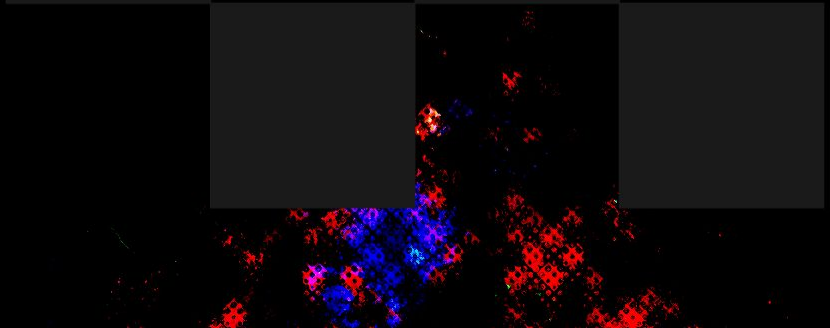
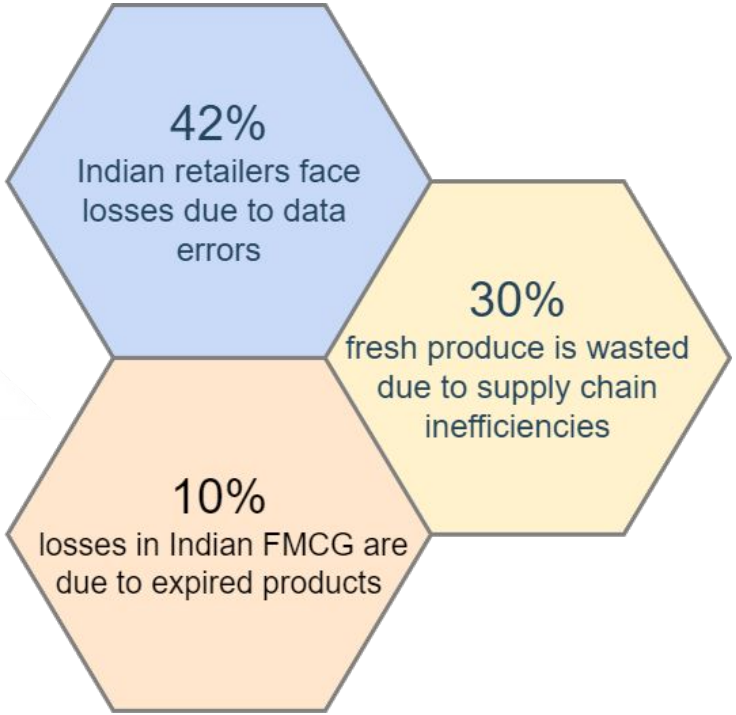
Flipkart GRiD 6.0

AI-POWERED PRODUCT DETAILS EXTRACTION, AND FRESHNESS DETECTION OF FRESH PRODUCE

(Robotics Track)



Problem Overview



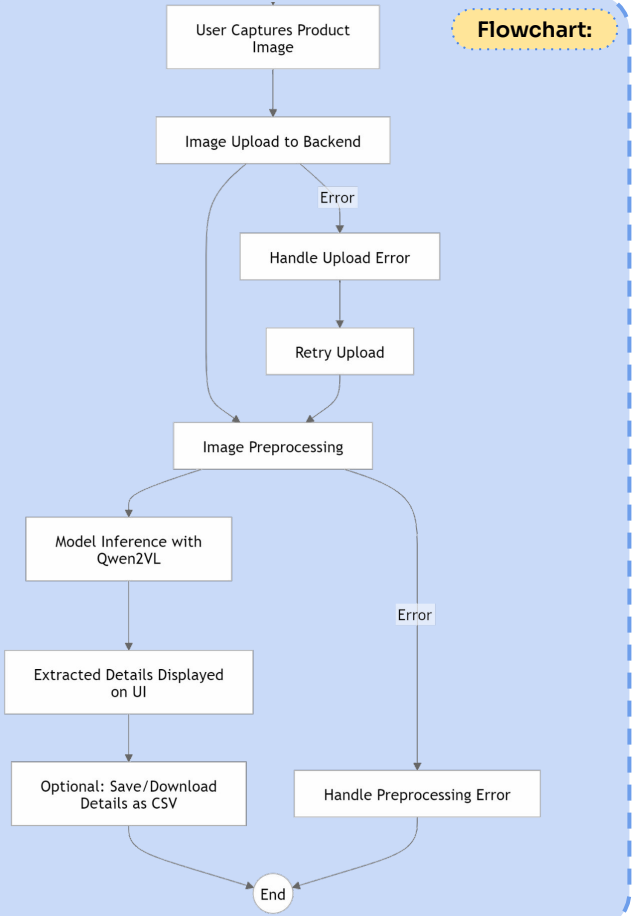
Problem Statements:

To automate the extraction of product details like **Brand**, **MRP**, **Expiry Date**, and **Net Content** from packaging, eliminating the errors and inefficiencies of manual data entry.

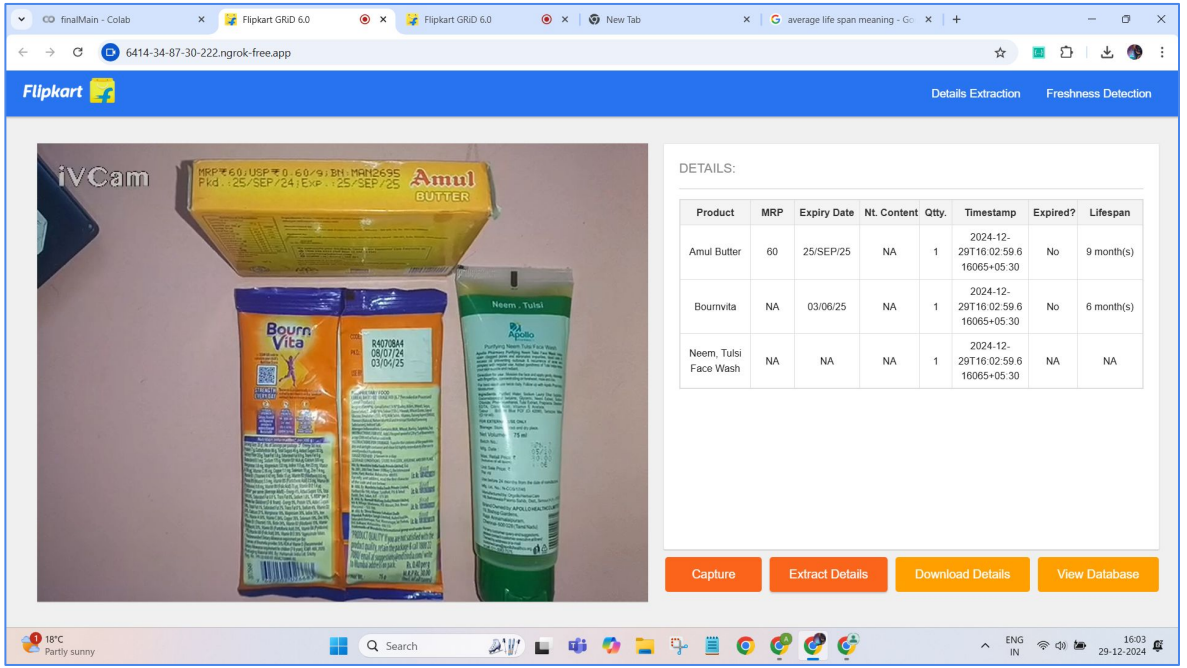
To detect the **freshness** of perishable items like fruits and vegetables enabling retailers to optimize food quality and minimize wastage.

Solution: Extraction of product details

Flowchart:



Application Overview:



Solution: Freshness detection of fresh produce

User captures image of a fruit, vegetable, bread etc.

↓

Clicks on the "Freshness Index" button to process the image.

↓

The system automatically detects the name of the fruit or vegetable.

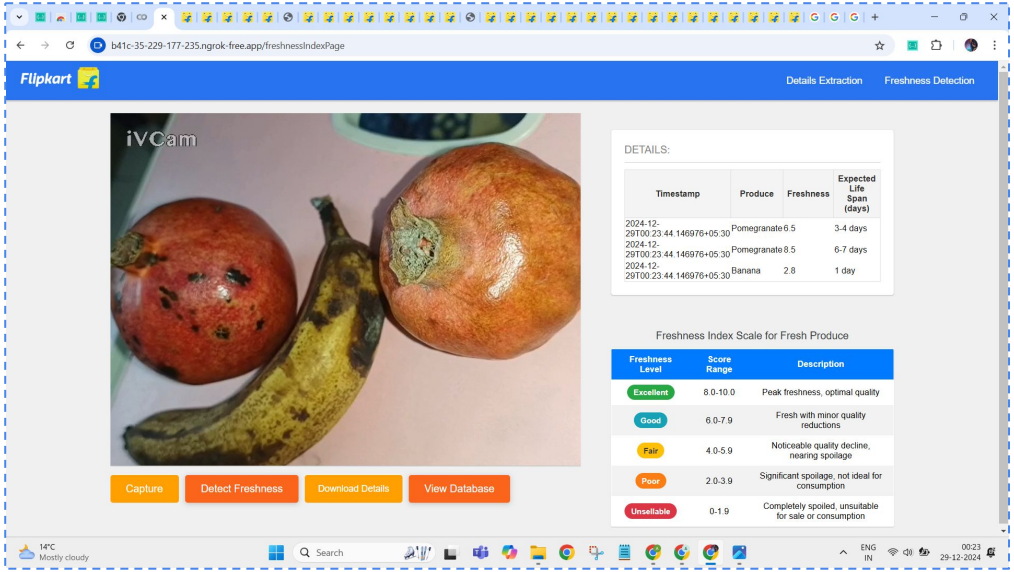
↓

It calculates and displays the Freshness Index.

↓

User can save the details for future reference or integration.

Application Overview:



INPUT	OUTPUT
	Name: tomato Freshness Index: 8.0
	Name: coriander Freshness Index: 6.7
	Name: apple Freshness Index: 7.0
	Name: beet Freshness Index: 7.9
	Name: spinach Freshness Index: 8.5

Technology used

Environment Setup: Python v3.7 or higher

Base Model: Qwen2VL-7B-Instruct

Why Qwen2VL model?

- Vision-Language Alignment
- Dynamic Resolution
- Open Source

Fine-tuning Setup: self curated dataset (600+)

WebApp Setup:

- **Backend:**
  FastAPI: A modern web framework for building APIs with Python.
- **Frontend:**
  HTML, CSS, and JavaScript

Database Setup:

- SQLite3
- Includes: Product details, and freshness details database

Deployment Setup:

- Google Colab (cloud deployment)
- Docker and ngrok

python™

pandas

Google Colaboratory

Qwen2-VL

Image	Entity Name	Entity Value
	Item Name	Frooti
	MRP	150.00
	Expiry Date	25/02/25
	Net Content	500 g

Sample Outputs

Future Scope

Enhanced Model Training with Larger Datasets: Continue fine-tuning the model with diverse datasets to enhance extraction accuracy and speed.

Scalability through GPU Upgrades: Leverage more powerful GPUs to reduce training time and enable fine-tuning with larger models for better accuracy.

Multi-Language Support through Fine-Tuning: Enable text extraction in regional languages for broader accessibility.

Appendices

GitHub Repository:
https://github.com/AnamtaRehman/flipkartGrid

References:
https://viso.ai/deep-learning/vision-language-models/
https://en.wikipedia.org/wiki/Optical_character_recognition

THANK YOU