India's No.1 E-School Challenge is back!

Flipkart

GRiD 6.0

SHAPING INDIA'S TECHSCAPE

Team
Robocon
IIT Roorkee

Team Name: **Team Robocon Gravity 3**

Project title:

SMART VISION QUALITY CONTROL SYSTEM

TEAM MEMBERS:

JAY MEENA
NITHISH RAVIKKUMAR
SHYAM GANESH B
PIYUSH RAJ

Executive Summary

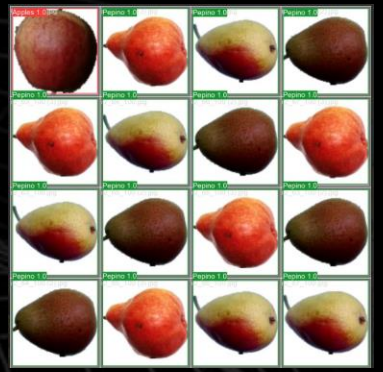
System Overview:

This smart vision quality control system is designed to

a) Process various types of input images

b) Extract relevant information

c) Perform quality assessments on products.




```
</output/>
{
    "Customer_Info": {
        "Name": "Nithish Ravikkumar",
        "Shipping Address": "Cs — 25, Jawahar Bhawan, Iit Roorkee IIT
ROORKEE Roorkee 247667
UTTARAKHAND",
        "Phone": ""
    },
    "Product": {
        "Product_Name": "",
        "Price": ""
    }
}
```


Key Components :

1. ***Text Detection and OCR*** - To extract text information from scanned images/Video input of Bill receipts, and extracting Order and Customer information.
2. ***Barcode Processing*** - To retrieve product information for unboxed/unpackaged items.
3. ***Object Detection Model-*** For unboxed items without barcodes. It detects objects from classes of everyday items and classifies them
4. ***Perishability and Freshness Assessment*** - To check for perishability, and estimate remaining shelf life.

Output :

The system provides **Customer and Order information**, and complete **Product information** as Output.

By combining multiple computer vision technologies, this system offers a comprehensive solution for product identification, quality assessment, and customer data processing in a retail environment.


```
"merchant": "Newegg.com",
"domain": "newegg.com",
"title": "Apple iPhone 6 64GB T-Mobile
"currency": "string",
"list_price": 0,
"price": 1200,
"shipping": "Free Shipping",
"condition": "New",
"availability": "Out of Stock",
"link": "https://www.upcitemdb.com/noro
```

Technical Approach

1. Introduction

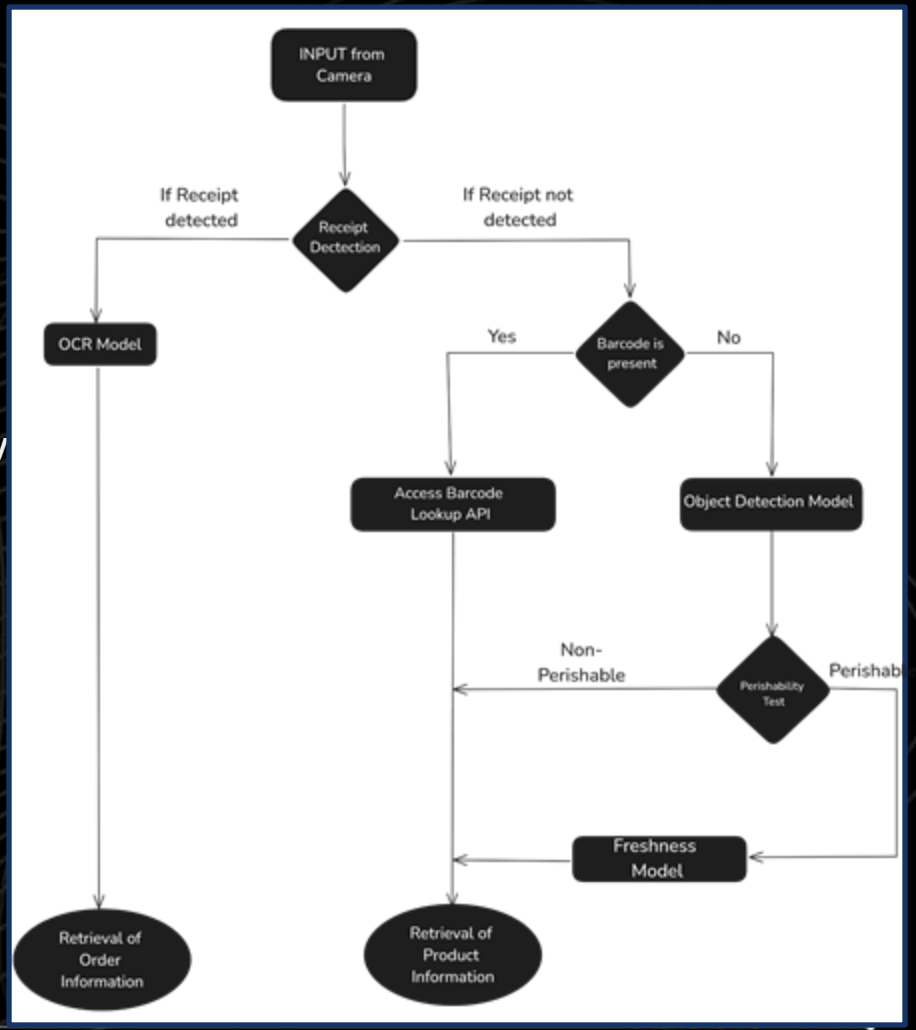
The sheer volume of products in E-commerce and the diversity of items creates the need for rapid processing, demanding innovative solutions that go beyond traditional manual inspection methods.

To address this demand we have created the **Smart Vision Quality Control System**. We have used the following working pipeline to automate and optimise the quality assessment process, from initial receipt of goods to final shipping, ensuring accuracy, consistency, and speed.

1.1. Use Case:

This system has multiple use cases in large scale E-commerce, retail sales and quality control.

Eg: Scenario- A major online grocery retailer needs to efficiently process incoming stock, manage inventory, and ensure product quality before shipping to customers.



1.2. Process Flow:

Bill Processing: As a truckload of mixed groceries arrives at the warehouse:
- System scans and processes the delivery invoice using OCR.
- Automatically updates inventory records and cross-references with the order placed.
- Flags any discrepancies between the order and the delivered items.

Barcode Scanning: For packaged goods (e.g., cereal boxes, canned goods):
- Warehouse staff quickly scan barcodes of multiple items.
- System instantly retrieves product information, updates inventory, and verifies pricing.
- Any mismatched or unknown barcodes are flagged for manual review.

Object Detection: For loose produce and bakery items:
- Items are placed on a conveyor belt passing under a camera.
- System identifies and counts each type of fruit, vegetable, or baked good.
- Automatically updates inventory and sorts items into appropriate storage areas.

Freshness Assessment: For perishable items:
- System analyzes images of produce, meats, and dairy products.
- Provides a freshness score and estimated shelf life for each item.
- Optimizes inventory by suggesting which items to ship first based on freshness.

Outcome: The retailer significantly reduces processing time, minimizes errors in inventory management, and ensures that customers receive fresh, high-quality products.

2. System Architecture

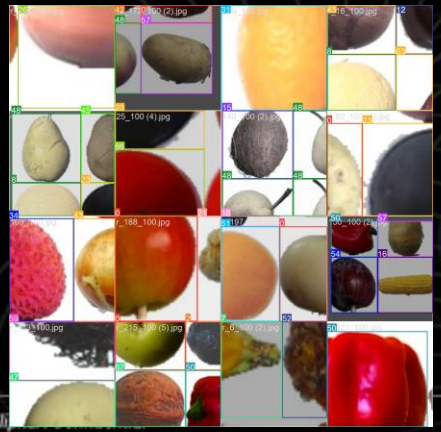
The system consists of five main components which have been incorporated into an integrated pipeline:

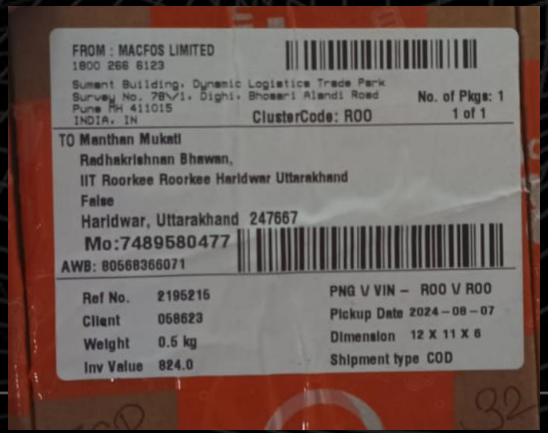
1. Image Acquisition
2. Text Detection and OCR
3. Barcode Processing
4. Object Detection
5. Perishability and Freshness Assessment Model

2.1 Image Acquisition

Input: Frame by frame analysis is applied on Camera video feed to determine which image to use. The High-resolution images of bills, packaged items, and unboxed products are passed on to the pipeline.

Hardware: High-resolution cameras capable of capturing detailed images in various lighting conditions





2.2 Text Detection and OCR

2.2.1. Surya Model

It is an Open-Source document OCR toolkit based on **ViT (Vision Transformer Architecture**). It can be locally run on any device with **minimal GPU** requirements. Hence it is versatile and optimal for our purpose, i.e, to extract text from scanned photos of printed bill materials.

- Performance: Over **90% accuracy** on test dataset.

- **Output**: JSON file containing **extracted text components**

2.2.2. NuExtract Model

NuExtract is a foundation model **from NuMind** designed for structured data extraction from text. We use this to extract key data about customer and product from the **unstructured textual output** of OCR

- Performance: The initial accuracy of the pre trained model on OCR was **39.8% in zero-shot setting**. After finetuning the model on dataset of 200 bill receipt transcripts for 40 epochs, the **accuracy increased of 81.1%.**

- Output format: Customer information (name, shipping address, phone number if present), Order number, **Product information** (Name, cost)

```
</output/>
{
    "Customer Info": {
        "Name": "Manthan Mukati",
        "Shipping Address": "Radhakrishnan Bhawan,
IIT Roorkee Roorkee Harldwar Uttarakhand
False Haridwar, Uttarakhand 247667",
        "Phone": "7489580477"
    },
    "Product": {
        "Product_Name": "",
        "Price": "824.0"
    }
}
```

2.3. Barcoded product analysis

2.3.1 Barcode Detection

If the package detected by camera does not have a bill and/or is unpacked , it is checked for a **barcode**. As all registered consumer products have a barcode, this will be an universal solution to extract product information

For this purpose, we will use **Pyzbar** package to detect and decode barcode directly from the video feed. It supports multiple barcode formats **like EAN, UPC, ISBN**, and hence is versatile and optimum for detecting products of various categories.

2.3.2 Product Information Retrieval

Once we have the barcode data, we will look up the product information via a **barcode database API**. Here, we will be using **UPCItemDB**, a vast collection of product data associated with barcodes (UPC, EAN, etc.) which **does not require any API key.**

- Request structure:

- Endpoint: https://api.upcitemdb.com/prod/trial/lookup?upc={upc_code}

- Method: HTTP GET

- Parameters: UPC, ISBN

- Output: Complete product information (Title, description, Brand, Product link, Price, Product class and Category) in JSON format.

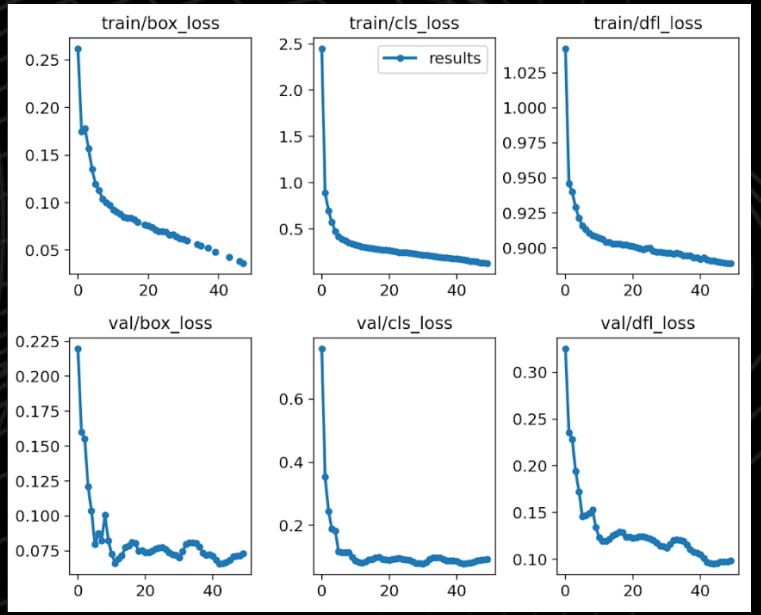
2.4. Object Detection

2.4.1 Model Architecture

If the package detected by camera does not have a bill or barcode, then it is passed through our Custom Object detection model It uses **YOLOv8** as the base model, and has been **fine-tuned on 5,000 images** across **15 classes** of common **Indian fruits and vegetables**. It also performs **instance segmentation** for **multiple objects** in a single image in **real time**, using **live video feed.**

2.4.2 Performance Metrics:

- **Base model has insignificant accuracy** in detection of live object detection. However, after fine- tuning on custom dataset, the **general accuracy reached 94.4%**
- Inference time of the model varies depending on image detail, but is suitable for high speed applications.
- Precision, recall **and major metrics score over 90%,** establishing our custom model as well suitable for usage in real lighting conditions.



2.4.3 Output:

Classification into type of fruit or vegetable

Segmentation of multiple objects, if present, in real- time

2.5 Perishability and Freshness Assessment

2.5.1 Model Architecture:

After object detection and classification, it is passed through our Custom Freshness Assessment model. The base model uses **ResNet-50** (pre-trained), which has been fine-tuned on 12,000 images across 16 classes, consisting of 8 types of fruits, each divided into **fresh** and **rotten** categories. This architecture allows the model to assess perishability in a real-time setting.

2.5.2 Freshness Index:

The Freshness Index aims to estimate the **approximate shelf life** of the detected fruit. The methodology relies on **visual defect analysis** performed using OpenCV, where several factors are considered:

- **Color changes**
- **Texture alterations**
- **Presence of blemishes or mold**

This enables the model to provide accurate shelf-life estimates by evaluating these key visual markers that correlate with fruit freshness.

2.5.3 Performance Metrics:

The model achieved an overall accuracy of **88.3%**, demonstrating reliable performance in detecting fresh and rotten fruits under bright lighting conditions. Precision, recall, and other relevant metrics closely match this accuracy, making the model well-suited for practical applications involving perishability assessment.

2.5.4 Output:

- Classification: Fresh or rotten
- Estimated remaining shelf life (in days)

3. Conclusion

This Smart Vision Quality Control System demonstrates the effective integration of multiple computer vision and machine learning technologies to address various aspects of product identification and quality assessment in retail and e-commerce settings. The combination of OCR, barcode processing, object detection, and freshness assessment provides a comprehensive solution for automated quality control and inventory management.

4. Future Improvements

- Expand object detection to more product categories
- Implement transfer learning for adapting to new products quickly
-Improve video analysis for OCR detection to reduce lag.
-Implement with TPU hardware to increase speed and performance.
- Integrate with inventory management systems

Documentation

Datasets:

Bill receipts dataset used for finetuning Text extraction Model:
Dataset of Fruits and Vegetables used for Object Detection Model:
Fresh and Rotten fruits dataset used for Freshness Estimate Model:

Code:

Integrated Pipeline:

Surya OCR: https://drive.google.com/drive/folders/1PTQDDUhE9cnHZzqh5P7fi6q_FWE6M-qf
Reference: https://github.com/VikParuchuri/surya

Nuextract : https://drive.google.com/drive/folders/1PTQDDUhE9cnHZzqh5P7fi6q_FWE6M-qf
Reference: https://huggingface.co/numind/NuExtract
Dataset: https://www.kaggle.com/datasets/trainingdatapro/ocr-receipts-text-detection

Barcode Reader and API: https://drive.google.com/drive/folders/1WrRbRJ_VxCcpH2ZORUHK9GbLnhY5Pr6g
Reference: https://www.upcitemdb.com/

Object Detection Model: https://drive.google.com/drive/folders/1cPMU9_L4Qk73-sLDGzfphOE5QwrfXfMS
Dataset: https://github.com/marcusklasson/GroceryStoreDataset

Freshness Detection Model: https://drive.google.com/drive/folders/12VJygH36WV05rgE6WC4LyMU5hOj-irXD
Dataset: https://data.mendeley.com/datasets/bdd69gvhv8/1


Flipkart
GRiD
6.0
Thank you!!