SSD Project: Weigh My Plate

Team No : 21

Project Title : Weigh My Plate

Project No : 5

Mentor : Prof. Charu Sharma

GitHub URL : https://github.com/viraj27shah/21_WeighMyPlate

Members :

1. Khule Vaishnavi Dhananjay - 2023201029
2. Nerkar Vedant Narendra - 2023202029
3. Vishnu Ranjith - 2023202004
4. Shah Viraj Utpalbhai - 2023201011

1. Introduction:

a. Overview:

Developed a web-based application using the MERN stack that allows users to upload images containing food items. The application employs trained object detection models, including RCNN and YOLOv5, to recognize, classify and measure the calorie count of 22 specific food classes (12 within the custom dataset and 10 in the RCNN dataset).

b. User Authentication:

- For non-logged-in users, image details are not stored, and the history tab is not visible.
- Logged-in users can access a history tab displaying uploaded images and their corresponding calorie counts.

2. Model Training:

a. Custom Dataset:

Created a meticulous dataset with 40-50 labelled images for each class using RoboFlow. Employed advanced annotation tools such as LabelImg and MakeSense.ai for precise and efficient labelling.

b. Trained Models:

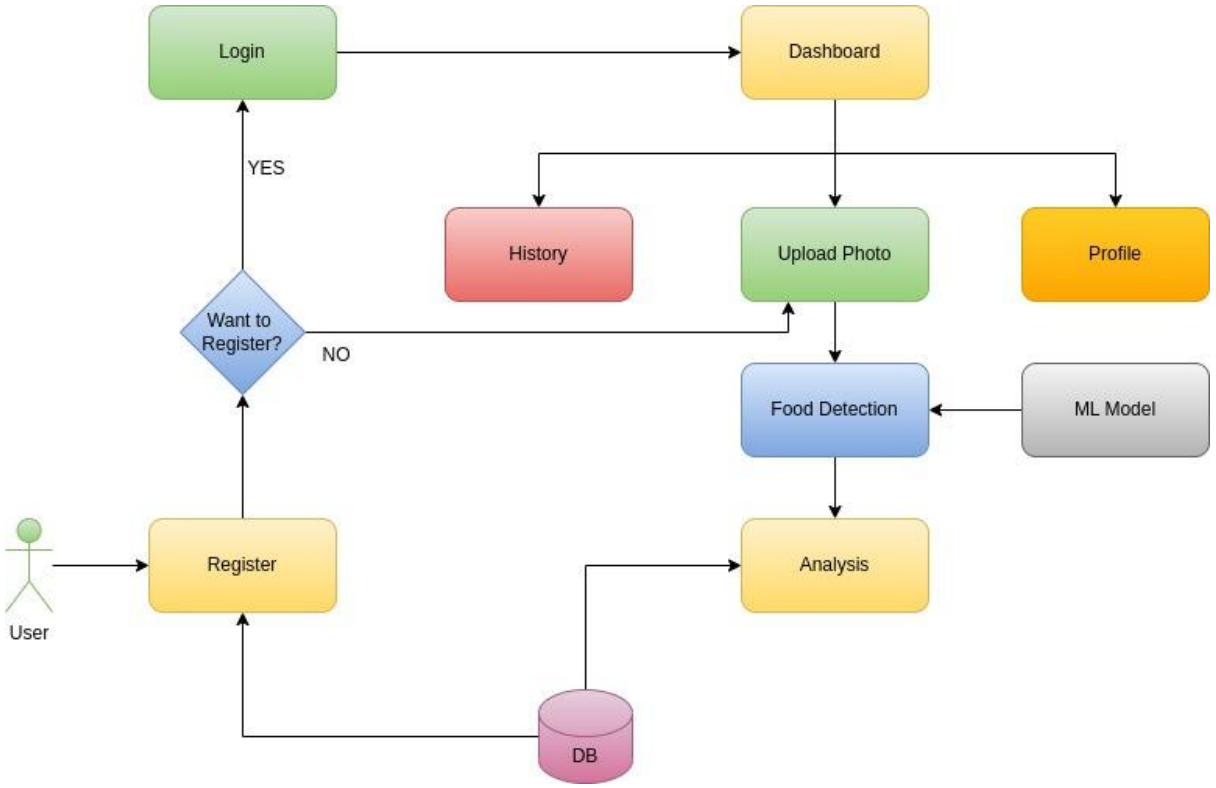
Used RCNN and YOLOv5 for object detection. Achieved a precision of 78%, indicating 78% of predicted positive instances are correct. Attained a mean average precision (mAP) of 76%, demonstrating high precision across different classes. Secured a recall of 69.2%, signifying the model captures 69.2% of all actual positive instances.

c. Training Process:

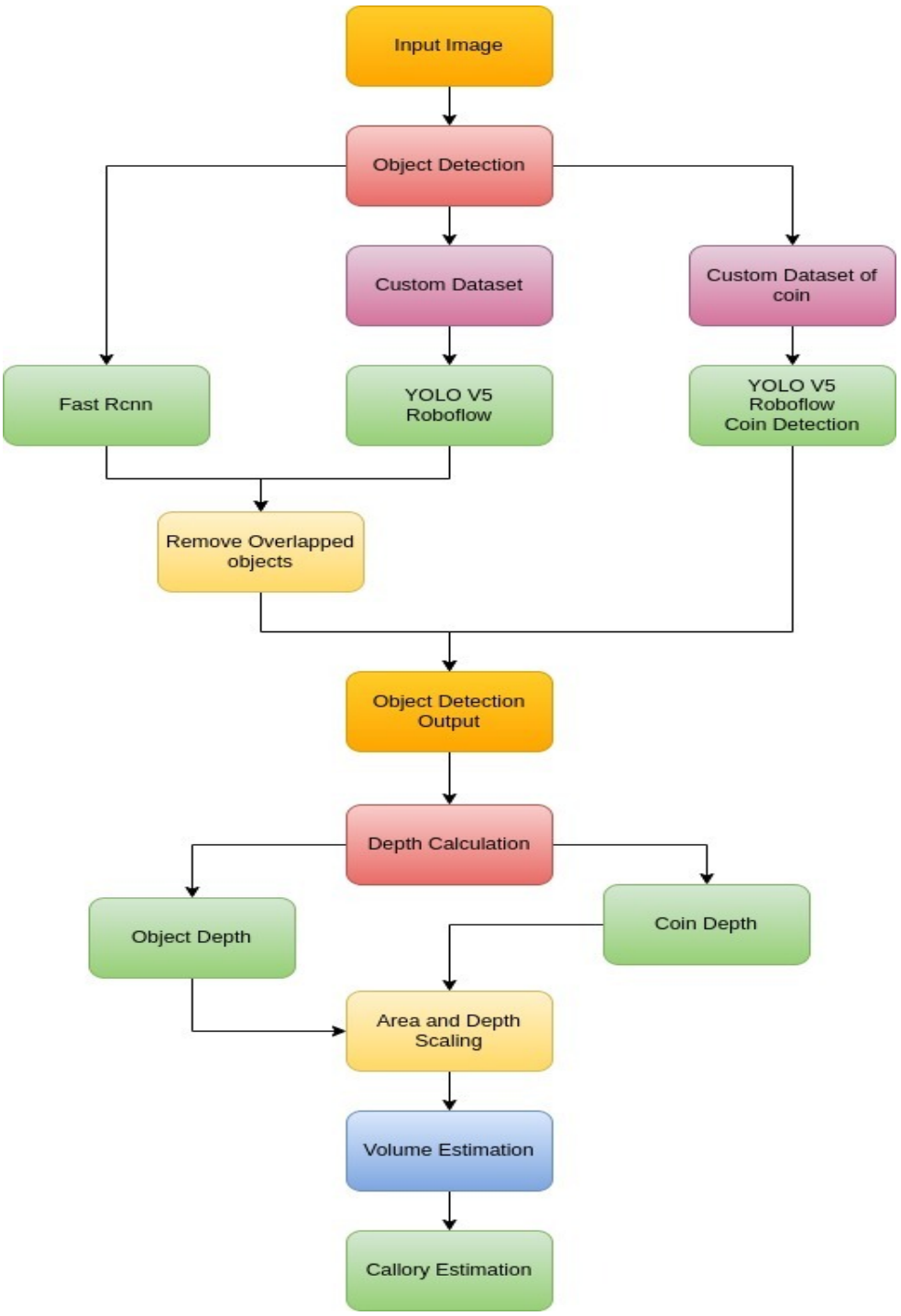
There are overall 3 models that work together to generate the output.

- Faster R-CNN:
    - Faster R-CNN (Faster Region-based Convolutional Neural Network) is a pre-trained object detection model. It detects and generates appropriate bounding boxes on various classes of which only the food-related items are used.
- YOLO v5 for food detection:
    - Faster R-CNN (Faster Region-based Convolutional Neural Network) is another pre-trained object detection model. The difference between this model and Faster R-CNN is that the concept of Transfer learning is implemented on this model. The base YOLO v5 model is used and trained on with the custom dataset. This makes the training process much more efficient as it converges quickly.
- YOLO v5 for coin detection:
    - The same concept as above but implemented for a single class which is ‘coin’.
- MiDaS for depth estimation:
    - MIDAS (Monocular Depth Estimation in Real-Time with Adaptive Sampling) is a deep learning-based method for monocular depth estimation, meaning it estimates the depth or distance of objects in a scene using a single image or video frame.

3. Implementation:



Flow of System



Logic

The overall workflow of the project is as shown in the figure above.

- First the image is uploaded by the user
- Then the image passed through the 3 models
    - Faster R-CNN and the Custom Trained model detect mutually exclusive classes. In case two bounding boxes overlap then the one with the lesser confidence is removed

    - Coin Detection gives the apparent size and the apparent depth of the coin which is placed on a level plane alongside the food plate.
- Using the apparent and the real life dimensions of the coin we can scale the rest of the dimensions of food items to real life values (cm)
- After this, the volume is calculated for each detected object which is then mapped to its appropriate density and calorie values.

3. Web Application:

a. Functionality :

Users can upload images to obtain calorie counts of food items using trained models. Non-logged-in users have no image history, promoting privacy. Logged-in users can access a history tab displaying past uploads and calorie details.

b. Privacy Measures:

Non-logged-in users: No image details stored, history tab not visible. For logged-in users history will be stored securely, providing a personalized experience.

4. Annotation Tools:

a. Labelling Tools:

Employed LabelImg and MakeSense.ai for efficient and accurate annotation. Optimized workflow for precise labeling of objects within images.

Use Cases :

1. Dietary Monitoring
2. Nutritional Education
3. Weight Loss Support
4. Allergen Detection
5. Restaurant Menu Analysis