IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, November 2026
Impact Factor: 8.2

A Deep Learning-Based Multi-Food Detection and Portion Estimation Framework for Automated Nutritional Analysis

Jaydip Jitendra Patil[1], Srushti Deepak Patil[2], Prof. Vikram Popat Deokate[3]
[1],[2] Students, Department of Computer Engineering
[3]Guide, Department of Computer Engineering
Al Ameen Educational & Medical Foundation's College of Engineering & Management Studies, Pune, Maharashtra,

Abstract: *Accurate food recognition and nutritional analysis are important components of digital health and dietary monitoring applications. Existing image-based food recognition systems commonly focus on identifying a single food category from an uploaded image, which limits their effectiveness when analyzing meals containing multiple food items. Furthermore, recognizing a food category alone does not provide sufficient information to estimate the actual quantity consumed or its nutritional composition.*

This research proposes an enhanced deep learning-based framework for multi-food detection and portion-aware nutritional estimation. The proposed system integrates object detection, image segmentation, food classification, portion-size estimation, and nutritional information retrieval into a unified application. A deep learning-based object detection model is used to identify multiple food items within a single meal image. Individual food regions are subsequently processed to estimate their portions using calibrated image geometry and, where available, depth information. The estimated quantities are combined with nutritional information to calculate the calories, carbohydrates, proteins, and fats associated with each food item.

The proposed framework aims to overcome the limitations of conventional single-food classification systems by providing item-wise and meal-level nutritional analysis. Its performance will be evaluated through detection accuracy, segmentation quality, portion-estimation error, and nutritional estimation error.

Keywords: Multi-Food Detection, Deep Learning, YOLO, Image Segmentation, Portion Estimation, Calorie Estimation, Nutritional Analysis, Computer Vision.

I. INTRODUCTION

Maintaining a healthy diet requires continuous monitoring of food consumption and nutritional intake. Traditional calorie-tracking applications depend heavily on manual food selection and portion entry, which can be time-consuming and may introduce inaccuracies.

Deep learning and computer vision provide opportunities to automate food recognition through image processing. Convolutional Neural Networks (CNNs) have demonstrated their usefulness in identifying food categories from images. However, conventional food classification approaches generally predict a single dominant food category and may not effectively analyze meals containing multiple food items.

A typical meal may contain rice, vegetables, chapati, lentils, salad, and other food components. Identifying only one category does not provide a complete representation of the meal.

Another important limitation is portion estimation. Nutritional values depend not only on food identity but also on the quantity consumed. Therefore, an automated dietary monitoring system should identify individual food items and estimate their respective portions before calculating nutritional values.

This research proposes a multi-food detection and portion-estimation framework that combines deep learning-based object detection, image segmentation, geometric analysis, and nutritional information retrieval.

Copyright to IJARSCT
www.ijarsct.co.in

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, November 2026
Impact Factor: 8.2

A. Motivation

The motivation behind this research is to improve automated dietary monitoring by addressing two major challenges: (1) identification of multiple food items from a single meal image, and (2) estimation of individual food quantities for more meaningful nutritional calculations.

B. Research Objectives

- To develop a deep learning-based multi-food detection system.
- To identify individual food categories present in a meal.
- To separate food regions using image segmentation.
- To estimate food portions using a calibrated estimation approach.
- To calculate item-wise nutritional values.
- To generate complete nutritional summaries for meals.

C. Research Contributions

- A multi-food detection pipeline for identifying different food items within a single image.
- Integration of food-region segmentation for individual item analysis.
- A calibrated portion-estimation approach for estimating food quantities.
- Integration of estimated portions with nutritional databases.
- A unified nutritional analysis framework providing item-wise and total meal information.

II. RELATED WORK

Image-based food recognition has been studied using conventional machine learning and deep learning techniques. CNN-based approaches extract visual features such as color, texture, and shape to classify food categories.

The previously published research, "AI-Powered Food and Fitness Guide using CNN," introduced a food recognition and fitness management application using the Food-101 dataset, CNN-based classification, and nutritional information retrieval through the Open Food Facts API [1].

The existing approach provides a foundation for automated food recognition and nutritional monitoring. However, the present research focuses on extending food analysis from single-category recognition toward multi-item detection and portion-aware nutritional estimation.

Object detection techniques provide the capability to identify multiple objects within the same image. Image segmentation methods further support the identification of individual object regions. Combining these techniques with calibrated geometric estimation can support the development of more comprehensive food analysis systems.

Further literature review should examine recent research on: (1) deep learning-based food object detection, (2) food image segmentation, (3) single-view and depth-based food volume estimation, (4) automated dietary assessment, and (5) nutritional estimation using food databases.

III. PROPOSED METHODOLOGY

The proposed framework consists of five major processing stages: food image acquisition and preprocessing, multi-food object detection, food image segmentation and classification, portion-size estimation, and nutritional analysis and reporting.

A. Food Image Acquisition and Preprocessing

The system accepts a food image through a web-based application. The uploaded image undergoes preprocessing operations, including image resizing, pixel normalization, image quality checking, noise reduction where necessary, and preparation for the object detection model. The processed image is passed to the multi-food detection model.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, November 2026
Impact Factor: 8.2

B. Multi-Food Detection Using YOLO

A YOLO-based object detection architecture is proposed for identifying multiple food items within a single image. Unlike a conventional image classification model that predicts a dominant food category, the object detection model identifies individual food regions and generates bounding boxes with corresponding class predictions and confidence scores.

The detection process includes input image processing, feature extraction, object detection, food category prediction, bounding-box generation, and confidence-score calculation. The model will be trained and evaluated using an appropriately annotated food detection dataset.

Each detected food item is represented by its predicted category, bounding-box coordinates, and confidence score.

C. Food Image Segmentation

After object detection, individual food regions are processed using an appropriate image segmentation technique. Segmentation helps separate food regions from the surrounding plate and background.

The process involves extracting detected food regions, generating segmentation masks, identifying visible food areas, calculating region properties, and preparing segmented regions for portion estimation. An appropriate segmentation model will be selected based on dataset availability, computational requirements, and evaluation results.

D. Portion-Size Estimation

Portion estimation is an important component of the proposed framework. The system estimates the quantity of each detected food item using calibrated geometric information. A reference object or depth sensor may be used to establish the physical scale of the food image.

The proposed estimation process includes identifying the food region, extracting its geometric dimensions, estimating physical scale using calibration information, estimating food volume using an appropriate geometric or depth-based method, and converting estimated volume into mass using food-specific density information.

Mass = Estimated Volume × Food Density

Here, mass represents the estimated food quantity in grams, estimated volume represents the calculated food volume, and food density represents the corresponding food-specific density. The estimation method must be validated against actual measured food weights.

A single uncalibrated RGB image cannot reliably establish actual food mass. Therefore, calibration and experimental validation are essential.

E. Nutritional Information Calculation

After estimating the food quantity, the system retrieves nutritional information from a nutritional database. For each food item, the system obtains nutritional values per 100 grams.

Nutrient Amount = (Estimated Food Mass / 100) × Nutrient Value per 100 g

The system calculates calories, carbohydrates, proteins, and fats. The total nutritional value of the meal is calculated by adding the corresponding values of all detected food items.

F. Meal-Level Nutritional Analysis

The final stage combines the results obtained from the detection, segmentation, portion-estimation, and nutritional calculation modules. The system generates detected food categories, individual food quantities, item-wise calorie estimates, macronutrient information, total meal calories, and total meal macronutrient summaries. The results are displayed through an interactive nutritional dashboard.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, November 2026
Impact Factor: 8.2

IV. PROPOSED SYSTEM ARCHITECTURE

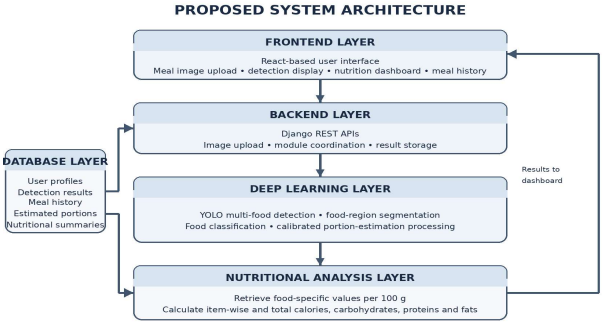


Fig. 2. Proposed system architecture

A. Frontend Layer

The frontend provides an interface for image uploading, displaying detection results, and presenting nutritional information. The existing React-based interface may be extended to include meal image upload, multiple food detection visualization, individual food quantity display, nutritional information dashboard, and meal history.

B. Backend Layer

The backend handles application requests and coordinates the processing modules. Django REST APIs can be used for image upload, object detection requests, segmentation processing, portion-estimation requests, nutritional data retrieval, and result storage.

C. Deep Learning Layer

This layer contains the machine learning components: YOLO-based multi-food detection, food image segmentation, food classification, and portion-estimation processing.

D. Nutritional Analysis Layer

This component retrieves food-specific nutritional information and calculates nutritional quantities using estimated food mass.

E. Database Layer

The database stores user profiles, food detection results, meal history, estimated portions, and nutritional summaries.

V. SYSTEM WORKFLOW

- The user uploads a meal image.
- The backend preprocesses the uploaded image.
- The YOLO model detects multiple food items.
- Individual detected regions are passed to the segmentation module.
- The system estimates food portions using calibrated image information.
- Estimated food masses are converted into nutritional quantities.
- The nutritional database provides food-specific information.
- The system calculates individual and total meal nutrition.
- Results are displayed on the user dashboard.
- Meal information is stored for future dietary monitoring.

VI. EXPERIMENTAL EVALUATION

The proposed system will be evaluated using separate experiments for detection, segmentation, portion estimation, and nutritional calculation.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, November 2026
Impact Factor: 8.2

A. Object Detection Evaluation

The detection model will be evaluated using precision, recall, mean Average Precision (mAP), and Intersection over Union (IoU).

B. Segmentation Evaluation

The segmentation module will be evaluated using Intersection over Union, Dice Similarity Coefficient, and pixel-level segmentation accuracy where appropriate.

C. Portion-Estimation Evaluation

The estimated food quantities will be compared against actual measured food weights. Mean Absolute Error (MAE) can be calculated as follows:

$$MAE = (1/N) \times \Sigma \, |\text{Actual Quantity} - \text{Estimated Quantity}|$$

Mean Absolute Percentage Error (MAPE) may also be reported where actual quantities are nonzero.

D. Nutritional Estimation Evaluation

The calculated nutritional values will be compared with reference nutritional values derived from measured portions and appropriate nutritional records. The evaluation will include calorie estimation error, protein estimation error, carbohydrate estimation error, and fat estimation error.

E. Comparative Analysis

The proposed framework will be compared against the original single-food classification baseline. The comparison will investigate differences in food identification capability, multiple-item detection, portion-estimation capability, nutritional estimation error, and processing time.

VII. RESULTS AND DISCUSSION

This section will present the experimental results obtained from the implemented framework. The following results should be reported: multi-food detection performance, segmentation performance, actual versus estimated food quantities, actual versus estimated nutritional values, comparison with the existing classification baseline, and computational performance.

The analysis should discuss the strengths and limitations of the proposed framework, including difficulties involving overlapping food items, mixed dishes, occlusion, lighting variations, and portion estimation. No numerical performance claims should be made until the corresponding experiments have been completed.

VIII. LIMITATIONS

Overlapping food items may reduce detection performance.
Similar-looking food categories may cause classification errors.
Portion estimation may be affected by image perspective.
Food density varies according to preparation methods.
Nutritional database entries may not perfectly represent the actual meal.
Single-view imaging may introduce uncertainty in volume estimation.
These limitations should be considered during system development and evaluation.

IX. CONCLUSION AND FUTURE SCOPE

This research proposes an enhanced deep learning-based framework for multi-food detection and portion-aware nutritional estimation. The proposed approach extends conventional food image classification by identifying multiple food items within a single meal image and estimating individual food portions. Through the integration of object detection, image segmentation, calibrated portion estimation, and nutritional information retrieval, the system aims to provide more comprehensive nutritional analysis.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 3, November 2026
Impact Factor: 8.2

The framework is intended to support automated dietary monitoring by calculating item-wise and total meal nutritional values. Future work may explore improved three-dimensional food reconstruction, depth-camera integration, real-time mobile food detection, and more accurate food-volume estimation techniques.

REFERENCES

[1] S. Shinde, S. Salve, S. Jadhav, P. Patel, and P. Shinde, "AI-Powered Food and Fitness Guide using CNN," International Journal of Advanced Research in Computer Science & Technology (IJARCST), vol. 8, no. 5, September–October 2025, DOI: 10.15662/IJARCST.2025.0805022.
[2] Add the original YOLO architecture research paper corresponding to the model version actually used.
[3] Add relevant research on food image segmentation.
[4] Add research on calibrated food volume and portion estimation.
[5] Add research on image-based dietary assessment and nutritional estimation.
[6] Add the official documentation and citation for the dataset and nutritional database actually used.

Fig. 1. Proposed end-to-end processing workflow

Meal Image Input	YOLO-Based Detection	Food Region Segmentation	Calibrated Portion Estimation	Nutrition Calculation