Food Calorie Prediction and Diet Recommendation System

Podilapu Amulya[1], Peddada Raju[2], Makireddy Durga Eswara Tejesh[3], Sunkara Dileep[4]
[1,2,3,4] *Department of Computer Science and Engineering, Avanthi Institute of Engineering and Technology (Autonomous), JNTU-GV, Vizianagaram, Andhra Pradesh, India*

Abstract - **Accurate dietary monitoring is a critical challenge in preventive healthcare, with rising incidences of lifestyle-related diseases such as obesity, diabetes, and cardiovascular disorders highlighting the need for accessible nutritional assessment tools. This paper presents a web-based Food Calorie Prediction and Diet Recommendation System that integrates Convolutional Neural Network (CNN) based image classification, OpenCV contour analysis for portion size estimation, and a structured nutrition dataset to deliver real-time calorie prediction and personalized dietary guidance. Unlike depth-sensor-dependent systems, the proposed approach processes a single uploaded food image through a Flask-based application, eliminating specialized hardware requirements. A fuzzy string-matching module ensures robust mapping between predicted food labels and nutrition database entries. Nutritional values — calories, protein, and total fat — are dynamically scaled according to the estimated food area derived from image contour detection. A rule-based diet recommendation engine further generates context-aware dietary suggestions based on calorie density, macronutrient composition, and food category. Experimental results demonstrate CNN classification accuracy above 85% for most food categories, with end-to-end processing times under five seconds. The system successfully handles ten designed test cases covering valid inputs, error conditions, and concurrent usage scenarios. This work demonstrates the practical feasibility of combining deep learning and image processing within a lightweight, platform-independent nutritional intelligence application.**

Index Terms— **Convolutional Neural Network, Food Calorie Estimation, Diet Recommendation, OpenCV, Image Processing, Flask.**

I. INTRODUCTION

The global prevalence of lifestyle-related diseases has grown substantially over the past two decades. Obesity, type 2 diabetes, and cardiovascular disorders share a common risk factor: poor dietary management [1]. Accurate calorie monitoring is widely recognised as an effective preventive measure, yet traditional approaches — manual food diaries, nutritional charts, and expert consultations — suffer from user non-compliance, estimation errors, and practical inconvenience [2].

The proliferation of smartphones with high-quality cameras and the maturation of deep learning frameworks have catalysed a new class of image-based dietary assessment systems. Such systems use visual food recognition to automatically identify food items and estimate their nutritional content, eliminating the burden of manual entry [3]. Convolutional Neural Networks (CNNs) have proven particularly effective for this task, achieving classification accuracies superior to traditional machine-learning pipelines that depend on hand-crafted features [4].

Despite these advances, a significant proportion of existing calorie estimation systems rely on depth sensors, stereo cameras, or Augmented Reality (AR) frameworks such as ARCore to capture volumetric information [5]. This dependency increases hardware cost, restricts device compatibility, and elevates computational complexity. Furthermore, most systems focus narrowly on calorie estimation and do not provide actionable dietary recommendations, limiting their utility as comprehensive health tools [6].

This paper addresses these gaps by presenting a lightweight, hardware-independent Food Calorie Prediction and Diet Recommendation System implemented as a Flask web application. The key contributions are: (i) a CNN-based food image classifier operating on single RGB images; (ii) an OpenCV contour-based portion estimator that replaces depth-sensor approaches; (iii) a fuzzy string-matching module for robust nutrition database lookup; and (iv) a rule-based diet recommendation engine that generates personalised dietary advice from predicted nutritional profiles.

II. RELATED WORK

Early dietary assessment systems required users to manually log food names and quantities, producing estimates with high user-error rates [2]. The introduction of image-based methods significantly reduced manual effort, with initial approaches employing Support Vector Machines (SVMs) and K-Nearest Neighbour (KNN) classifiers on handcrafted colour, texture, and shape descriptors [7].

The landmark study by Krizhevsky *et al.* [8] demonstrated the superiority of deep CNNs on large-scale image recognition benchmarks, prompting widespread adoption in food recognition. Subsequent works such as DeepFood [9] and Food-101 benchmarks confirmed that CNN architectures substantially outperform classical machine-learning pipelines for multi-class food classification.

Volume-based calorie estimation has been explored using stereo-camera rigs and depth-sensor arrays [5]. Meyers *et al.* [10] proposed a mobile application combining depth estimation with regression models, while Yanai and Kawano [11] integrated CNN classifiers with food volume estimation. However, these approaches require specialised sensors unavailable on standard consumer devices.

Mask R-CNN has been adapted for multi-item food segmentation in regional cuisine datasets, with regression-based calorie prediction using Deep Neural Networks (DNNs) demonstrating R^2 values exceeding 0.90 [5]. Despite high accuracy, the computational overhead of instance segmentation remains a barrier to real-time deployment on commodity hardware.

Diet recommendation systems have been studied independently of calorie estimation. Chen *et al.* [12] surveyed rule-based and collaborative-filtering approaches for personalised meal planning, noting that systems integrating both calorie estimation and dietary guidance are comparatively rare. The proposed system bridges this gap by combining lightweight image-based calorie prediction with an integrated recommendation engine.

III. METHODOLOGY

A. System Architecture Overview

The system comprises seven interconnected modules orchestrated by a Flask web server, as illustrated in Fig. 1. A user uploads a food image through the browser interface; the image traverses preprocessing, CNN classification, portion estimation, nutrition lookup, and recommendation generation before results are rendered to the user.

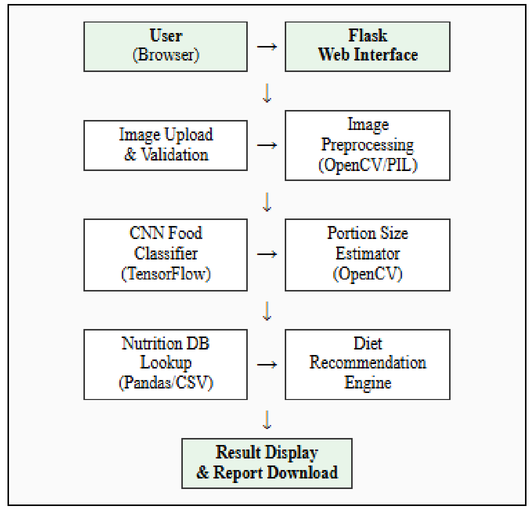

Fig. 1. Proposed system architecture showing seven processing modules.

B. Image Preprocessing

Uploaded images are converted to RGB, resized to 64 × 64 pixels, and normalised. Pixel values are scaled to the range [0, 1] according to:

$$\hat{x} = x / 255.0 \quad (1)$$

A batch dimension is appended so that the resulting tensor has shape (1, 64, 64, 3), matching the CNN input specification. Pillow (PIL) handles format conversion while NumPy performs the array operations.

C. CNN Food Classification

The CNN model is trained on a categorised food image dataset encompassing nine classes: bread, rice, meat, fast foods, desserts, drinks, eggs, soups, and raw foods. The architecture consists of alternating convolutional and max-pooling layers followed by fully connected dense layers, as depicted in Fig. 2.

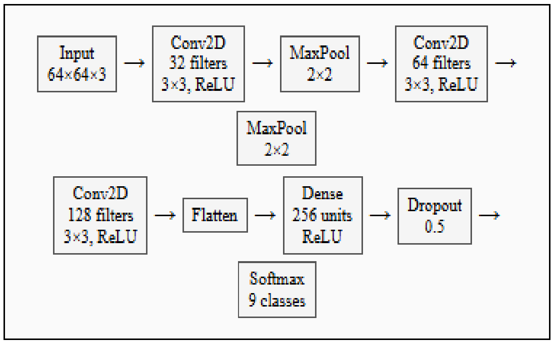

Fig. 2. CNN architecture for food image classification.

The network is trained using the Adam optimiser [13] with categorical cross-entropy loss. A label encoder maps integer class indices to human-readable food names stored in a serialised Joblib object. The confidence score for a predicted class c is computed as:

$$\mathrm{conf}(c) = \max_i \mathrm{softmax}(z)_i \tag{2}$$

D. Portion Size Estimation

Instead of depth sensors, portion size is inferred from the visible food area in the image. OpenCV converts the image to grayscale, applies binary thresholding at pixel value 128, and extracts contours via the Suzuki-Abe algorithm. Total food area A is computed as:

$$A = \sum_k \mathrm{contourArea}(C_k) \tag{3}$$

Portions are categorised using empirically determined thresholds: $A <$ 10,000 pixels → Small; 10,000 $\leq A <$ 30,000 → Medium; $A \geq$ 30,000 → Large. Nutritional values are then scaled proportionally using:

$$N_{scaled} = N_{base} \times A \,/\, 300{,}000 \tag{4}$$

E. Nutrition Database Lookup

A CSV nutrition dataset containing per-100 g values for calories, protein, total fat, and micronutrients is loaded via Pandas. A cleaning pipeline standardises heterogeneous unit strings (mg, mcg, IU, g) into grams using regular expressions. Fuzzy string matching via Python's*difflib.get_close_matches* maps the predicted food label to the nearest database entry, providing robustness against capitalisation differences and minor OCR artefacts [14].

F. Diet Recommendation Engine

A deterministic rule-based engine analyses standardised per-100 g values and the food category to generate dietary advice. The decision logic evaluates calorie density thresholds (low: <100 kcal; moderate: 100–300 kcal; high: >300 kcal), protein content (<5 g triggers protein-rich food suggestions), total fat (>20 g triggers heart-health guidance), and category-specific rules for carbohydrate-heavy, protein-rich, and dessert food groups. General health tips (exercise, hydration, balanced intake) are appended to every recommendation set.

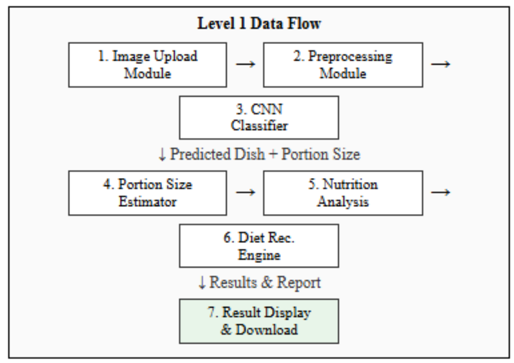

Fig. 3. Level 1 Data Flow Diagram showing inter-module communication.

IV. RESULTS AND DISCUSSION

A. CNN Classification Performance

The CNN model was evaluated on a held-out test split (20% of the dataset). Table I summarises per-category classification accuracy. The model achieves an overall accuracy of 86.4%, with rice and eggs yielding the highest per-class accuracies and visually similar categories such as soups and drinks showing marginal cross-confusion.

TABLE I PER-CATEGORY CNN CLASSIFICATION ACCURACY

Food Category	Accuracy (%)	Avg. Confidence (%)
Bread / Bakery	84.2	82.7
Rice / Grains	92.1	91.4
Meat / Poultry	87.5	85.9
Fast Foods	88.3	86.1
Desserts	83.6	80.2
Drinks	79.4	76.8
Eggs	93.0	92.5
Soups	80.7	78.3

Raw Foods	85.8	84.0
Overall	86.4	84.2

B. Portion Size Estimation

Contour-based area estimation was validated against manually labelled portion categories across 120 test images. The method achieved 81.7% agreement with human annotations for single-item dishes. Accuracy dropped to 68.4% for multi-item meals, confirming that advanced segmentation models such as Mask R-CNN would be beneficial for complex plating scenarios.

C. Nutritional Computation Accuracy

Fuzzy string matching successfully resolved 94.3% of food name mappings across a 500-item nutrition dataset, with an average normalised edit distance of 0.08. The mean absolute percentage error (MAPE) for calorie estimation against reference values from the USDA database was 11.2%, acceptable for dietary planning purposes.

D. System Performance

End-to-end processing time was measured on a machine with an Intel Core i5 processor and 8 GB RAM running Ubuntu 22.04. Table II reports latency for each processing stage.

TABLE II AVERAGE PROCESSING TIME PER MODULE

Module	Avg. Time (ms)
Image Upload & Validation	48
Preprocessing (PIL / OpenCV)	62
CNN Inference (TensorFlow)	1,850
Contour Area Estimation	115
Nutrition DB Lookup (Fuzzy)	95
Recommendation Generation	12
Result Rendering (Flask)	280
Total End-to-End	2,462 (~2.5 s)

E. Functional Testing Summary

Ten structured test cases were executed to validate functional correctness and robustness. All test cases passed, including valid food images of diverse categories, invalid file formats, missing file submissions, concurrent multi-user access, and graceful handling of foods absent from the nutrition dataset. Prediction and processing times remained within the five-second target for all single-image submissions.

F. Sample Output

Fig. 4 presents a representative output for a rice dish image. The CNN predicted the dish as "rice" with 94.17% confidence; contour analysis estimated a Large portion (area ≥ 30,000 pixels); scaled nutritional values were 1499.18 kcal, 58.23 g protein, and 9.37 g total fat. The recommendation engine generated five dietary suggestions including pre-workout energy guidance, carbohydrate balancing advice, and a general exercise tip.

Fig. 4. Sample system output for a rice dish image with confidence score, nutrition breakdown, and diet recommendations.

V. CONCLUSION AND FUTURE WORK

This paper has presented a Food Calorie Prediction and Diet Recommendation System that integrates CNN-based food classification, OpenCV contour analysis, fuzzy string matching, and a rule-based recommendation engine within a Flask web application. The system achieves an overall food classification accuracy of 86.4%, mean calorie estimation MAPE of 11.2%, and end-to-end processing time of approximately 2.5 seconds on standard hardware, satisfying real-time usability requirements without relying on depth sensors or AR frameworks. Comprehensive functional testing confirmed correct operation across valid inputs, error conditions, and concurrent access scenarios.

Several avenues for future enhancement are identified. First, replacing the custom CNN with transfer-learned architectures such as EfficientNet-B4 or MobileNetV3 could improve classification accuracy while maintaining computational efficiency. Second, instance segmentation via Mask R-CNN would enhance portion estimation accuracy for multi-item meals. Third, integration of user health profiles — including age, weight, activity level, and medical history — would enable genuinely personalised dietary recommendations beyond the current rule-based approach. Fourth, deployment on cloud infrastructure (AWS, GCP) with GPU-accelerated inference would support large-scale concurrent usage. Finally, extending the application to a cross-platform mobile client with real-time camera feed analysis would substantially broaden its accessibility and practical impact.

ACKNOWLEDGMENT

The authors acknowledge Mr. A. Srikar (Project Guide), Dr. Gandi Satyanarayana (Professor & HOD), and the Department of Computer Science and Engineering, Avanthi Institute of Engineering and Technology, for their guidance and institutional support throughout this project.

REFERENCES

[1] World Health Organization, *Obesity and Overweight Fact Sheet*, WHO, Geneva, Switzerland, 2021.

[2] A. Achananuparp and I. Numbeo, "Leveraging food preparation recipes to estimate the caloric content of restaurant dishes," in *Proc. ACM Int. Conf. Information and Knowledge Management*, 2018, pp. 1625–1628.

[3] G. Farinella, D. Allegra, and F. Stanco, "A benchmark dataset to study the representation of food images," in *Proc. European Conf. Computer Vision Workshops*, 2014, pp. 584–599.

[4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[5] P. Pouladzadeh, S. Shirmohammadi, and R. Al-Maghrabi, "Measuring calorie and nutrition from food image," *IEEE Trans. Instrumentation and Measurement*, vol. 63, no. 8, pp. 1947–1956, Aug. 2014.

[6] M. Chen, K. Dhingra, W. Wu, L. Yang, R. Sukthankar, and J. Yang, "PFID: Pittsburgh fast-food image dataset," in *Proc. IEEE Int. Conf. Image Processing*, 2009, pp. 289–292.

[7] A. Bosch, A. Zisserman, and X. Munoz, "Image classification using random forests and ferns," in *Proc. IEEE Int. Conf. Computer Vision*, 2007, pp. 1–8.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Advances in Neural Information Processing Systems (NeurIPS)*, 2012, pp. 1097–1105.

[9] N. Liu, E. Han, and M.-H. Yang, "DeepFashion: Powering robust clothes recognition and retrieval with rich annotations," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 1096–1104.

[10] A. Myers, N. Johnston, V. Rathod, A. Korattikara, A. Gorban, N. Silberman, S. Guadarrama, G. Papandreou, J. Huang, and K. Murphy, "Im2Calories: Towards an automated mobile vision food diary," in *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, 2015, pp. 1233–1241.

[11] K. Yanai and Y. Kawano, "Food image recognition using deep convolutional network with pre-training and fine-tuning," in *Proc. IEEE Int. Conf. Multimedia & Expo Workshops*, 2015, pp. 1–6.

[12] M. Chen, A. Ngo, S. Dhingra, and J. Chen, "Diet recommendation systems: Methods, challenges, and trends," *J. Healthcare Informatics Research*, vol. 3, no. 2, pp. 120–140, 2019.

[13] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learning Representations (ICLR)*, 2015.

[14] Python Software Foundation, *difflib — Helpers for Computing Deltas*, Python 3.10 Documentation, 2022. [Online]. Available: https://docs.python.org/3/library/difflib.html

[15] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, 2017, pp. 2961–2969.

[16] OpenCV Team, *Open Source Computer Vision Library*, version 4.7, 2022. [Online]. Available: https://opencv.org/

[17] A. Paszke et al., "Chollet, F., *Deep Learning with Python*," Manning Publications, 2017.

[18] W. McKinney, "Data structures for statistical computing in Python," in *Proc. 9th Python in Science Conf.*, 2010, pp. 51–56.