# Nutri Vision: Real-Time Food Recognition and Nutrition Management Platform

**Prof. T. Vinay Simha Reddy[1], Dudala Mohana Naga Venkata Sri Lasya[2], D. Sai Kartheek[3], G. Naga Sandeep Reddy[4]**

[1]*Assistant Professor, Department of Artificial Intelligence and Machine Learning, Malla Reddy University, Maisammaguda, Hyderabad, India*
[2345] *Students, Department of Artificial Intelligence and Machine Learning, Malla Reddy University, Maisammaguda, Hyderabad, India*

---***---

**Abstract** - Unhealthy dietary habits are one of the leading causes of non-communicable diseases such as obesity, diabetes, and cardiovascular disorders. Monitoring daily food intake remains a major challenge due to the lack of efficient and user-friendly tools. This paper presents NutriVision, an intelligent real-time food recognition and nutrition management system powered by deep learning and computer vision techniques. The proposed system utilizes Convolutional Neural Networks (CNNs), specifically MobileNetV2 and ResNet-50, trained on the Food-101 dataset for accurate food classification.

In addition to image-based recognition, the system integrates a text-based nutritional lookup using the Indian Nutritional Database (INDB), ensuring accurate dietary information for region-specific foods. The architecture combines a React-based frontend, Fast API backend, TensorFlow for model inference, and OpenCV for preprocessing. The system also provides personalised features such as BMI calculation, calorie tracking, meal planning, and health safety checks. NutriVision offers a comprehensive solution for intelligent dietary monitoring and personalised nutrition guidance.

***Keywords***: Food Recognition, Deep Learning, CNN, MobileNetV2, ResNet-50, Nutrition Analysis, Computer Vision, Fast API, TensorFlow, Health Monitoring

## 1. INTRODUCTION

In recent years, the prevalence of lifestyle-related diseases has increased significantly due to unhealthy eating habits and lack of dietary awareness. Individuals often struggle to maintain balanced nutrition because existing tools require manual input, which is both time-consuming and prone to human error. Traditional calorie tracking applications rely heavily on user knowledge and consistency, limiting their effectiveness.

Advancements in artificial intelligence, particularly in deep learning and computer vision, have enabled the development of automated food recognition systems. These systems can analyse food images and provide accurate nutritional information without requiring manual effort. NutriVision leverages these technologies to create an intelligent platform that simplifies dietary tracking and improves health awareness.

The system is designed to recognise food items in real time using deep convolutional neural networks. It extracts meaningful visual features from images and classifies them into predefined categories. In addition, the platform integrates a text-based search mechanism to handle foods that may not be well represented in image datasets.

### 1.1 Food Recognition Using AI

Food recognition involves identifying food items from images using machine learning models. CNNs play a crucial role by learning hierarchical features such as textures, shapes, and colours. These models are trained on large datasets like Food-101 to achieve high classification accuracy.

### 1.2 Nutrition Management Systems

Nutrition management systems aim to provide users with insights into their dietary habits. NutriVision enhances this concept by combining food recognition with personalized recommendations, helping users maintain healthier lifestyles.

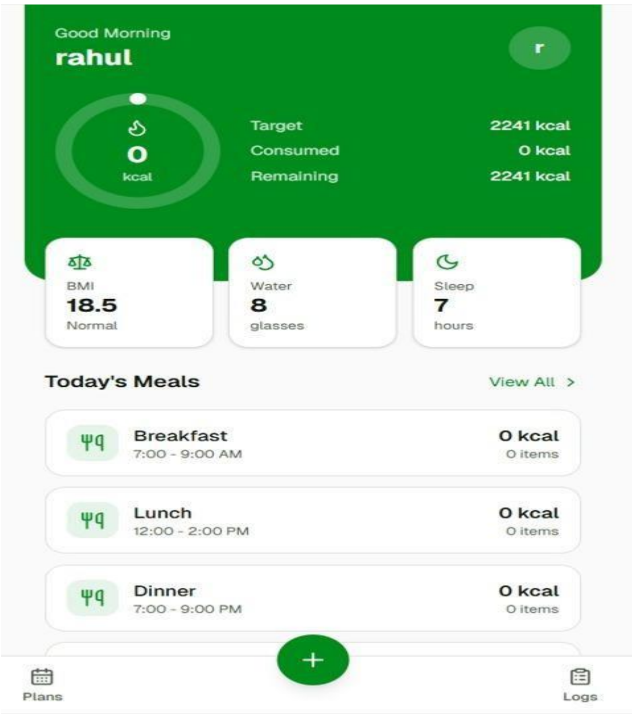


**Fig. 1**: Home Dashboard – Caloric Gauge, BMI, Meal Slots

## 2. LITERATURE REVIEW

The field of automated dietary assessment has evolved significantly with the introduction of machine learning techniques. Early systems relied on handcrafted features such as color histograms and texture descriptors, which were limited in performance and unable to generalize across diverse food categories. With the emergence of deep learning, Convolutional Neural Networks replaced traditional approaches by automatically learning feature representations from raw image data. Studies such as DeepFood demonstrated the effectiveness of transfer learning using pre-trained models like ImageNet. These approaches significantly improved classification accuracy and robustness.

Recent advancements include lightweight architectures like MobileNet, designed for real-time applications with reduced computational cost. Similarly, ResNet introduced residual connections that allow deeper networks to be trained efficiently, resulting in higher accuracy.

Research has also explored integrating nutrition databases with recognition systems. Systems combining image recognition with nutritional lookup provide more comprehensive solutions compared to standalone models. However, most existing solutions lack personalization features such as health monitoring and dietary recommendations.

NutriVision addresses these limitations by integrating food recognition, nutrition analysis, and personalized health tracking into a unified platform.

## 3. PROPOSED SYSTEM

NutriVision offers a dual-input pipeline: (a) upload a food

photograph → CNN classification, or (b) type a food name

→ INDB query. Both paths return food name, kcal, protein, carbohydrates, fat per 100 g, and a personalised Health Safety Check. The image pathway decodes the uploaded image, resizes to 224×224 px via OpenCV, normalises pixel values, runs TensorFlow inference (MobileNetV2 or ResNet-50), and maps the predicted Food-101 label to a nutrition table from USDA FoodData Central and INDB.

The text pathway accepts a food name string, performs fuzzy matching against the INDB food-item index, and returns the best-matching nutritional record. This is especially valuable for Indian dishes (biryani, idli, dosa) comprehensively covered in INDB but rare in Food-101 training images.

Figure 2 shows the Profile Summary. This profile drives the Health Safety Check and the computation of caloric targets. The header card displays computed BMI (18.5 – Normal), daily caloric target (2,241 kcal/day), weight (60 kg), body infrmation, and health goal (Weight Gain).

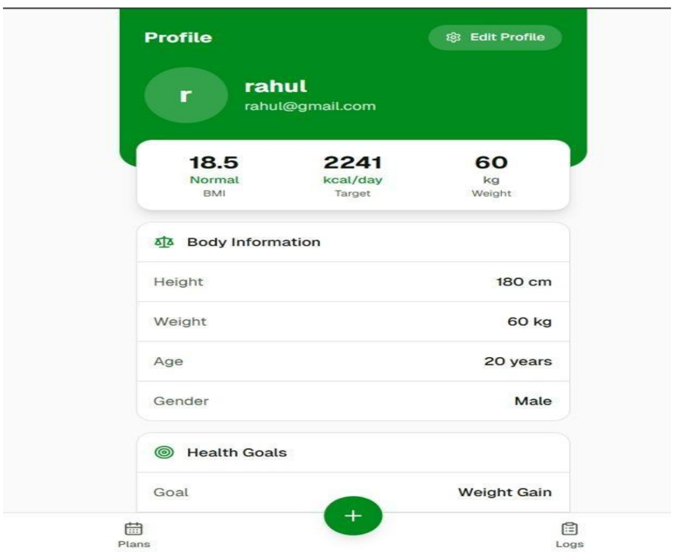


**Fig. 2:** Profile Summary – BMI, Daily Target, Health Goals

Figure 3 shows the Meal Plans Screen where users can view and manage their weekly diet plan. Each meal slot displays calorie estimates and food items assigned throughout the day, enabling structured dietary planning aligned with individual caloric targets.

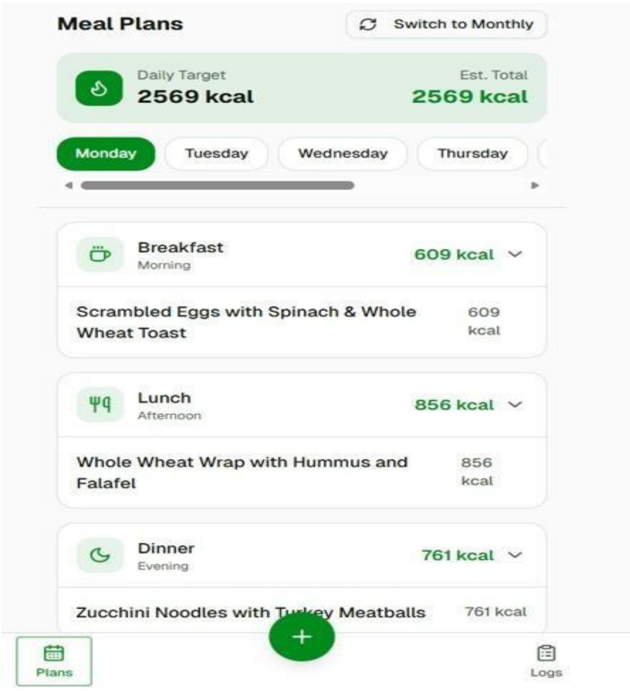


**Fig. 3:** Meal Plans Screen – Weekly Diet Plan with Calorie Estimates

## 4. METHODOLOGY

The methodology of NutriVision involves multiple stages, including data preprocessing, model training, and inference.

### A. Convolutional Neural Networks

CNNs are used to extract spatial features from images through convolution operations, activation functions, and pooling layers. These networks automatically learn patterns that distinguish different food categories.

### B. Model Selection

- MobileNetV2: Lightweight and efficient for real-time applications
- ResNet-50: Deeper architecture with higher accuracy

### C. Transfer Learning

Pre-trained models are fine-tuned on the Food-101 dataset to improve performance while reducing training time. Data augmentation techniques such as rotation and flipping are applied to enhance generalization.

### D. Dataset

- Food-101 dataset (101,000 images)
- INDB (Indian Nutritional Database) for nutrition values

### E. Training Configuration

Models are trained using TensorFlow with optimized hyperparameters such as learning rate, batch size, and dropout to prevent overfitting.

**TABLE I: DATASET SUMMARY – FOOD-101 AND INDB**

| Property | Food-101 | INDB |
|---|---|---|
| Primary Use | Image Classification | Text Search |
| Categories | 101 categories | 568 food items |
| Total Records | 101,000 images | 568 records |
| Train Split | 75,750 (750/class) | N/A |
| Test Split | 25,250 (250/class) | N/A |
| Resolution | 224×224 px | Text query |

Training was conducted using TensorFlow 2.14/Keras on an NVIDIA RTX 3060 GPU (12 GB VRAM). Classification head: GlobalAvgPool2D → Dropout(0.4) → Dense(512,ReLU) → Dense(101,Softmax). Loss: categorical cross-entropy. Metrics: Top-1 and Top-5 accuracy tracked per epoch.

**TABLE II: TRAINING CONFIGURATION**

| Parameter | MobileNetV2 | ResNet-50 |
|---|---|---|
| Input Size | 224×224×3 | 224×224×3 |
| Phase 1 LR | 1e-3 | 1e-3 |
| Phase 2 LR | 1e-5 | 1e-5 |
| Total Epochs | 50 | 50 |
| Batch Size | 32 | 32 |
| Dropout | 0.4 | 0.4 |
| LR Schedule | Cosine Annealing | Cosine Annealing |
| Early Stop | Patience=10 | Patience=10 |

## 5. SYSTEM ARCHITECTURE

NutriVision follows a three-tier architecture: Presentation Layer (Next.js/React), Application Layer (FastAPI + TensorFlow), and Data Layer (PostgreSQL + INDB + Food-101 nutrition table). All inter-tier communication uses RESTful JSON APIs over HTTPS with JWT authentication.

Figure 4 presents the system architecture. The numbered flow: (1) browser submits image or food name; (2) FastAPI receives request; (3) OpenCV preprocesses image; (4) TF model classifies food; (5) nutrition table lookup; (6) Health Safety Check against user profile; (7) JSON response to browser; (8) meal log persisted in PostgreSQL. For text pathway, steps 3–4 are replaced by direct INDB query.

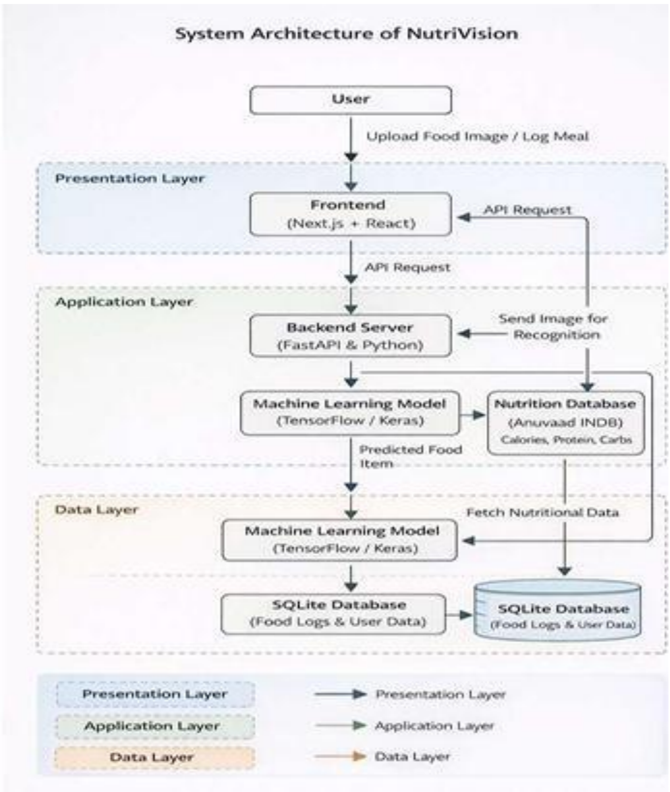


**Fig. 4:** System Architecture Diagram

A. Presentation Layer (Next.js / React) Next.js 14 with React 18 provides server-side rendering for fast initial loads and client-side state for real-time caloric gauge 84.6 88.1 updates. Components include Dashboard, Food Analysis, Meal Plans, and Profile views. Axios handles JWT-authenticated API calls.

B. Application Layer (FastAPI + TensorFlow) FastAPI provides async Python endpoint handling, automatic OpenAPI documentation, and Pydantic validation. Both TensorFlow models are loaded at server startup and cached in memory. The image endpoint accepts multipart/form-data, preprocesses via OpenCV, runs TF inference, queries the nutrition table, and returns JSON within one async handler.

C. Image Preprocessing (OpenCV) The pipeline performs: (1) JPEG/PNG decode and BGR-to RGB conversion; (2) bilinear resize to 224×224 px; (3) normalisation to [0,1]; (4) ImageNet mean subtraction ($\mu$=[0.485,0.456,0.406], $\sigma$=[0.229,0.224,0.225]) ensuring input-distribution alignment with pre-trained backbone weights.

D. Data Layer (INDB + Food-101 + PostgreSQL) The INDB covers 568 Indian food items with per-100 g energy, protein, fat, carbohydrates, and fibre. The Food-101 nutrition table maps 101 class labels to USDA FoodData Central entries. PostgreSQL stores user accounts, profiles, meal logs, and daily records. Indexed full-text search on INDB enables sub millisecond fuzzy matching.

E. Health Safety Check After nutritional retrieval, a rule engine cross-references caloric density, sugar, and fat content against condition-specific thresholds from the user's health goal (Weight Gain, Weight Loss, Diabetes Management, etc.) producing a binary safe/caution flag displayed on the Food Analysis screen.

## 6. RESULTS

Both models were evaluated on the Food-101 official test split (25,250 images). Table III reports Top-1, Top-5 accuracy, and median inference latency over 500 consecutive calls on the deployed FastAPI server.

**TABLE III:** FOOD-101 CLASSIFICATION RESULTS

| Model | Top-1 (%) | Top-5 (%) | Latency (ms) |
|---|---|---|---|
| MobileNetV2 | 83.7 | 95.8 | 18.4 |
| ResNet-50 | 87.3 | 97.2 | 29.6 |

ResNet-50 achieves 87.3% Top-1 and 97.2% Top-5, confirming the advantage of its deeper residual feature hierarchy. MobileNetV2 at 83.7% offers a compelling accuracy-latency trade-off (18.4 ms, 37% faster than ResNet-50) and is the default deployment model.

End-to-end platform response times averaged 1.23 s (MobileNetV2) and 1.87 s (ResNet-50) across 200 test requests. INDB text-search returned results in under 50 ms in all cases.

**TABLE IV:** PER-CATEGORY TOP-1 ACCURACY

| Food Category | MobileNetV2 (%) | ResNet-50 (%) |
|---|---|---|
| Pizza | 91.4 | 94.2 |
| Burger | 89.7 | 92.8 |
| Ice Cream | 88.3 | 91.7 |
| Sushi | 87.2 | 90.5 |
| Biryani | 84.6 | 88.1 |
| Idli | 82.1 | 86.4 |
| Salad | 79.5 | 83.9 |

Visually distinctive categories (pizza, burger) consistently score above 89%, while heterogeneous categories (salad, biryani) score lower due to variable ingredient compositions and presentation styles, motivating future ingredient-level analysis work.

Figure 5 shows the Scan Food Screen where users can capture or upload a food photograph for real-time image-based recognition. The interface initiates the CNN inference pipeline, displaying the identified food item and its nutritional breakdown within 2 seconds.

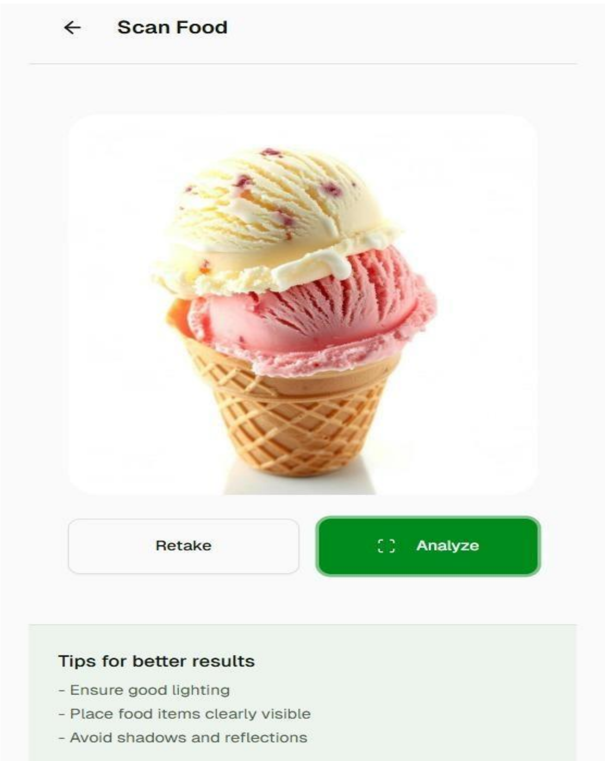


**Fig. 5:** Scan Food Screen – Image Capture for Food Recognition

Figure 6 shows the food recognition output from the NutriVision platform after image processing. The result displays the identified food item along with its complete nutritional breakdown including calories, protein, carbohydrates, and fat per 100 g, along with the personalised Health Safety Check status for the user.

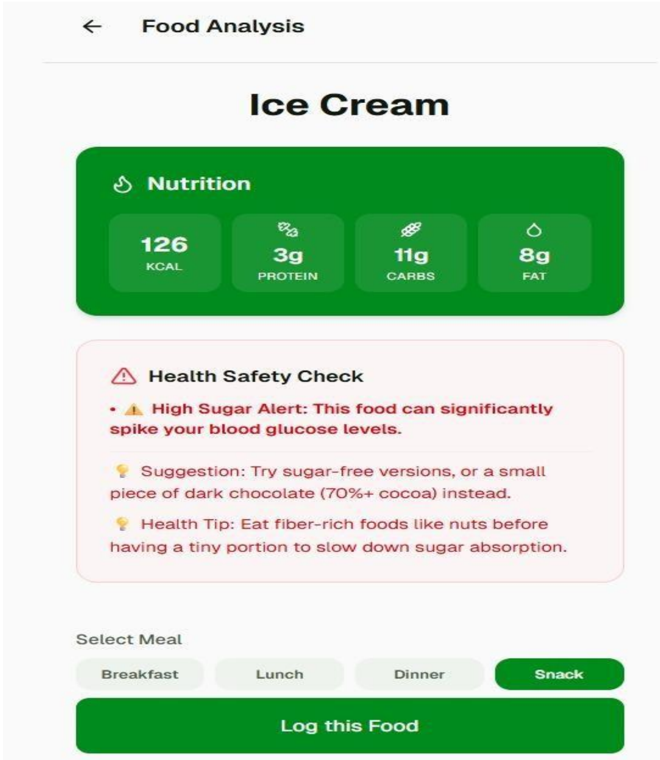


**Fig. 6:** Food Recognition Output – Nutritional Breakdown and Health Safety Check

## 7. CONCLUSION

NutriVision presents an advanced AI-based solution for food recognition and nutrition management. By combining deep learning models with nutritional databases, the system provides accurate and personalized dietary insights.

The integration of real-time recognition, health monitoring, and user-friendly interfaces makes it a comprehensive platform for improving dietary habits. Future enhancements may include portion size estimation, multi-food detection, and mobile application deployment.

Overall, NutriVision demonstrates the potential of artificial intelligence in transforming healthcare and promoting healthier lifestyles.

## 8. ACKNOWLEDGEMENT

We would like to express our sincere gratitude to our faculty members and mentors for their continuous support and guidance throughout this project. We also thank Malla Reddy University for providing the resources required for successful completion of this work.

## 9. REFERENCES

1] N. Sharma, A. Gupta, and R. Singh, "AI-based dietary assessment and food logging systems: A comprehensive review," Journal of Nutrition Informatics, vol. 12, no. 3,pp. 45 62, 2024.

[2] P. Kumar and S. Patel, "Personalized nutrition management using machine learning and wearable health data," IEEE Access, vol. 12, pp. 11423–11439, 2024.

[3] V. Rao and M. Desai, "Indian Nutritional Database (INDB): Structure, coverage and integration with dietary assessment software," Indian Journal of Nutrition and Dietetics, vol. 61, no. 1, pp. 12–25, 2024.

[4] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in Proc. IEEE/CVF CVPR, pp. 4510–4520, 2018.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE CVPR, pp. 770–778, 2016.

[6] A. Patel and S. Mehta, "Automated dietary monitoring for elderly care using computer vision and IoT integration," IEEE Internet of Things Journal, vol. 11, no. 4, pp. 7234–7248, 2024.

[7] G. Farinella et al., "Retrieval and classification of food images," Computers in Biology and Medicine, vol. 77, pp. 23 39, 2016.

[8] T. Ramachandran and V. Narayanan, "Full-stack web application development for AI-powered health monitoring using FastAPI and React," Int. J. Web Engineering and Technology, vol. 19, no. 2, pp. 112–134, 2024.

[9] C. Liu et al., "DeepFood: Deep learning-based food image recognition for computer-aided dietary assessment," in Proc. 14th Int. Conf. Computational Science, pp. 37–48, 2016.

[10] H. Hassannejad et al., "Food image recognition using very deep convolutional networks," in Proc. 2nd Int. Workshop Multimedia Assisted Dietary Management, pp. 41–49, 2016.

[11] G. Ciocca, P. Napoletano, and R. Schettini, "Food recognition: A new dataset, experiments, and results," IEEE J. Biomedical and Health Informatics, vol. 21, no. 3, pp. 588–598, 2017.

[12] S. Mezgec and B. Korousic Seljak, "NutriNet: A deep learning food and drink image recognition system for dietary assessment," Nutrients, vol. 9, no. 7, p. 657, 2017.

[13] N. Martinel, G. Foresti, and C. Micheloni, "Wide residual network for food recognition," in Proc. IEEE Int. Conf. Big Data, pp. 3933–3936, 2018.

[14] C. Panagiotakis, G. Tziritas, and P. Fragopoulou, "Multi scale feature fusion for fine-grained food recognition," Pattern Recognition Letters, vol. 168, pp. 53–61, 2023.

[15] L. Bossard, M. Guillaumin, and L. Van Gool, "Food- 101 – Mining discriminative components with random forests," in Proc. ECCV, pp. 446–461, 2014.

[16] X. Min et al., "ISIA Food-500: A dataset for large- scale food recognition via stacked global-local attention network," in Proc. ACM MM, pp. 393–401, 2023.

[17] A. Galoaa, K. Hamrouni, and S. Foufou, "Hybrid CNN Transformer model for real-time dietary assessment on mobile devices," IEEE Trans. Neural Systems and Rehabilitation Engineering, vol. 32, pp. 1102–1114, 2024.