IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154293

A SYSTEMATIC REVIEW OF AI POWERED FOOD RECOGNITION SYSTEM

Harsh Pachauri[1], Megha Rajput[2], Krishna Kaushik[3], Dharmesh Rawat[4], Pankaj Saraswat[5], Alok Singh Jadaun[6]

UG Student, Department of Computer Science & Engineering, Raja Balwant Singh Engineering Technical Campus, Agra, India[1]

UG Student, Department of Computer Science & Engineering, Raja Balwant Singh Engineering Technical Campus, Agra, India[2]

UG Student, Department of Computer Science & Engineering, Raja Balwant Singh Engineering Technical Campus, Agra, India[3]

UG Student, Department of Computer Science & Engineering, Raja Balwant Singh Engineering Technical Campus, Agra, India[4]

UG Student, Department of Computer Science & Engineering, Raja Balwant Singh Engineering Technical Campus, Agra, India[5]

Head of Department, Department of Computer Science & Engineering, Raja Balwant Singh Engineering Technical Campus, Agra, India[6]

Abstract: More people face health issues tied to poor eating habits - conditions like obesity and heart problems keep rising. Because of this shift, better tools for tracking what we eat have become essential. Old ways of recording meals, such as handwritten logs or trying to remember everything eaten in a day, tend to be unreliable. Mistakes happen. Memories fade. Sticking with those methods over time? Rare. A new approach enters here - not magic, just smart engineering. This method uses artificial intelligence to turn meal photos into useful nutrition facts without guesswork. At its core sits a neural network model called ResNet-18, good at telling one dish apart from another. It learned from more than 100,000 food images labeled across 101 categories. Training happened using Food-101, a large-scale collection built for exactly this purpose. Speed matters; PyTorch handles number crunching while OpenCV prepares raw pictures for analysis. Cropping, color adjustments, noise cleanup - all done before classification kicks in. Once identified, each food links to stored nutrient values: calories, protein levels, fat amounts, carbs included. Results appear instantly through a live website powered by Streamlit. Snap a photo, get details moments later. No waiting. No spreadsheets. Starting off different, recent studies - especially those on visionlanguage systems and image-focused workout tools - show our method handles speed and accuracy without favoring one too much. Early tests suggest this automatic system cuts down the hassle of logging by hand, providing something sturdy and flexible enough to grow with personal wellness needs while supporting clearer food choices.
Though not perfect, it fits well where quick results meet reliable detection.

Keywords: Food Recognition, Deep Learning, ResNet-18, Nutritional Analysis, Computer Vision, PyTorch, Calorie Estimation, Food-101 Dataset

I. INTRODUCTION

Rising rates of conditions like obesity and type-2 diabetes around the world have pushed science toward smarter ways to track what people eat - using technology instead of paper logs. Old techniques, say remembering meals from yesterday or jotting down snacks by hand, stumble because memory fades and eyeballing calories in complex meals rarely works well. To fix these flaws, researchers turned to an artificial intelligence tool that identifies foods in photos, swapping guesswork with instant analysis. Snap a picture, get nutrition details - all without typing every bite. Because it cuts out slow entry steps, users might stick with tracking longer. What results is less personal judgment, more consistency when managing diet choices day after day.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⧖ Peer-reviewed & Refereed journal ⧖ Vol. 15, Issue 4, April 2026
DOI: 10.17148/IJARCCE.2026.154293

Starting off with a strong backbone, the setup uses ResNet-18 - a type of deep learning design - built in PyTorch for faster processing and sharp results. Instead of shallow models, this one dives into depth but avoids common pitfalls thanks to its skip connections. Training happens on Food-101, a large collection featuring more than 100,000 pictures across 101 types of dishes. Because lighting or angle can change how food looks, every photo gets adjusted first through steps like scaling and flipping. These prep moves come from OpenCV tools that help smooth out differences before analysis begins. Thanks to residual blocks, gradients flow better during learning so layers deeper down still get useful updates. As a result, fine details stand out clearly even between meals that look almost identical at first glance. Once the item gets sorted correctly, it pulls exact calorie numbers plus protein, fat, carbohydrate details from a full nutrition reference list. This whole process runs through a website built with Streamlit, letting people add food photos and see diet info right away. Looking at recent studies - the kind covering tools such as Gemini, GPT-4, even YOLO setups - shows our ResNet-18 method hits a sweet spot: sharp enough to identify meals well without slowing down live use on phones or browsers. Coming updates aim to split apart mixed dishes into individual foods while guessing serving sizes by volume, making personal intake estimates more on point.

II. LITERATURE REVIEW

One big change in nutrition studies came when computers started handling what people eat. Instead of guessing from memory, folks used to write down meals themselves - often missing bits or getting details wrong. Picture this: machines now see food photos and figure out exactly what is on the plate. Thanks to smart algorithms built for spotting patterns in images, sorting one dish from another got way more accurate. Think deeper layers inside digital brains helping tell sushi apart from salad - even if they look alike. These systems learn through repeated exposure, much like how eyes get better at recognizing faces over time.

New findings show using big image collections such as Food-101 and ImageNet helps pull out strong visual clues, so models can tell apart dishes that look alike even when lighting or surroundings change. Because of this shift, researchers now mix visual identification with online nutrition tools that calculate calories and nutrients. Over time, these efforts have built systems that do more than spot meals - they offer tailored wellness feedback - setting fresh standards for smart, user-friendly health tracking.

Goud et al. (2024) [1] introduced a sophisticated AI-driven nutrition recommendation method designed to enhance the accuracy and reliability of personalized dietary advice. To overcome the common challenge of untrustworthy nutritional suggestions in generic AI models, the authors utilized a Deep Generative Network integrated with a Variational Autoencoder (VAE). This technical approach allows the system to robustly model user-specific data, such as anthropometric measurements and medical conditions, within a descriptive latent space. The platform further employs an optimizer to dynamically adjust meal quantities based on individual energy requirements, ensuring that the generated recommendations strictly align with established medical and nutritional guidelines.

Sen et al. (2024) [2] introduced Eathena, an AI-driven mobile nutrition assistant that utilizes the multimodal capabilities of ChatGPT to unify food recognition, nutritional analysis, and conversational feedback within a single system. To simplify complex architectures that typically rely on multiple specialized models, the researchers developed a dual-path inference pipeline. This system identifies packaged foods by querying the OpenFoodFacts API through barcode or label detection.

Latha et al. (2024) [3] developed "SnapIT," a comprehensive AI-powered web application designed to evaluate the nutritional quality and chemical safety of packaged food products. Unlike systems focused on cooked meals, SnapIT utilizes Google's Gemini-1.5 model to perform unified image recognition and Optical Character Recognition (OCR) directly from product labels. This allows the system to extract complex data such as ingredient lists, additives, and expiry dates. By integrating Natural Language Processing (NLP), the platform cross-references this data with scientific dietary guidelines to identify potential health risks associated with specific preservatives. The system assigns a health rating out of five and suggests safer alternatives, providing users with immediate transparency and fostering more informed, health-conscious purchasing decisions through a React.js and Express.js-based interface.

Agarwal and Mishra (2024) [4] In a series of progressive studies, Agarwal and Mishra (2023, 2024) introduced "SnapCount," an automated framework for image-based calorie estimation, and provided a comprehensive review of the technological shift from handcrafted features to Vision Transformers. Their work deconstructs the estimation pipeline into critical stages: data acquisition, food detection, and semantic segmentation. A central contribution of their research is the identification of portion size estimation as the most significant bottleneck in the field, primarily due to the difficulty of predicting 3D volume from 2D imagery. They evaluated various methodologies, including 3D

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154293

reconstruction and single-view volumetric estimation using deep regression models. Their systematic analysis addresses persistent challenges like class imbalance and occlusion in composite dishes, ultimately advocating for multi-modal learning and on-device AI to enhance the precision of digital health assistants.

Choochaiwattana et al. (2025) [5] developed an innovative AI application specifically designed to automate the "Healthy Balanced Free-Form Plate" model, focusing on the 2:1:1 dietary proportion (two parts vegetables, one part grain, and one part protein). Utilizing a U-Net convolutional encoder-decoder structure with a VGG16 backbone, the system was trained on complex Thai composite dishes to perform precise segmentation. The researchers implemented the Pixel Labeling Method (PLM) to quantify individual food components from topview images. By calculating the specific area and relative volume of each food group on a plate, the system provides a practical and accurate alternative to traditional manual estimation. This research underscores the importance of AI in supporting healthy eating patterns by ensuring that meal proportions align strictly with medically recommended balanced diet guidelines.

Romero-Tapiador et al. (2025) [6] evaluated the readiness of Vision-Language Models (VLMs) for automatic dietary assessment, exploring their ability to integrate visual and textual reasoning for complex food recognition tasks. To facilitate this research, they introduced the FoodNExTDB, a specialized database containing 9,263 food images derived from a real-world weight loss intervention. A critical contribution of this work is the involvement of seven nutrition experts who generated approximately 50,000 nutritional labels, categorizing images across 10 main categories, 62 subcategories, and 9 distinct cooking styles.

Pankaj et al. (2025) [7] introduced "NutriAI," a mobile-integrated solution designed to simplify calorie tracking and health management. The system utilizes ResNet-18 for high-accuracy food image classification and provides a real-time nutritional breakdown of detected items. A key feature of this research is the integration of a BMR (Basal Metabolic Rate) Calculator, which uses individual physical metrics to offer personalized daily calorie targets. The platform, developed using Streamlit, emphasizes a seamless user experience by combining visual recognition with interactive health dashboards, thereby encouraging consistent dietary monitoring and better lifestyle choices.

Kharbach (2025) [8] discussed the transformative role of data handling tools, specifically the transition from traditional chemometrics to advanced Machine Learning (ML) and AI in food analysis. The research highlights how modern analytical instruments generate massive, high-dimensional datasets that require AI to ensure food quality, safety, and authenticity. By leveraging algorithms such as Support Vector Machines (SVM), Random Forests, and Artificial Neural Networks (ANN), the study demonstrates that AI can detect minute textural details and chemical compositions that are impossible to capture through manual methods. This work underscores the necessity of AI-driven computational models in managing complex food data to maintain global food standards and transparency.

Han et al. (2024) [9] introduced "NutrifyAI," a comprehensive mobile and web-based system designed to eliminate the friction of manual dietary entry. The framework is built on a three-tier architecture: real-time food detection using the YOLOv8 model, nutritional data extraction via the Edamam Nutrition Analysis API, and personalized suggestions through the Edamam Meal Planning API. By utilizing a high-speed object detection model, the system can identify multiple food items in a single frame with high precision. The research demonstrates that combining localized computer vision with cloud-based nutritional intelligence provides users with immediate, accurate insights, significantly improving the practicality of digital health assistants while AI enhances learning, cautious application is essential to avoid negative educational outcomes.

Shinde et al. (2025) [10] developed a sophisticated health management ecosystem titled "AI-Powered Food and Fitness Guide," which leverages Convolutional Neural Networks (CNN) to address the complexities of modern dietary tracking. Recognizing the limitations of manual calorie counting, the authors engineered a system capable of performing automated food recognition and macronutrient estimation from user-uploaded images. The architecture is integrated with an extensive nutritional database that provides a granular breakdown of essential components, including carbohydrates, proteins, and fats. To further enhance the utility of the platform, the research incorporates advanced functionalities such as voice-activated food logging, a personalized recommendation engine for diet planning, and gamification elements like achievement badges to sustain long-term user engagement. By deploying this framework through a responsive web interface built on modern technologies, the study demonstrates a significant shift toward seamless, AI-driven personal health monitoring that minimizes human error and promotes informed nutritional choices.

Yaiprasert and Hidayanto (2024) [11] investigated the transformative impact of artificial intelligence on personalized dietary services within the digital food industry. The study addresses the critical challenge of low

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⪿ Peer-reviewed & Refereed journal ⪿ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154293

accuracy in individual decision-making models by proposing a novel Ensemble Machine Learning framework. This approach systematically integrates multiple distinct algorithms, specifically Decision Trees, Logistic Regression, and Naive Bayes, to create a more robust predictive model than any single method could achieve alone. By drawing on the dynamics of open innovation, the researchers demonstrated how combining the strengths of these diverse classifiers allows for more precise meal recommendations tailored to specific user preferences. The findings emphasize that such AI-driven ensemble techniques not only enhance the user experience by providing highly relevant dietary suggestions but also offer a significant competitive advantage to businesses by optimizing service delivery and fostering innovation in the rapidly evolving food-tech landscape.

Liu et al. (2016) [12] present a deep learning-based food image recognition system, "DeepFood," designed to overcome the inaccuracies and memory-based biases inherent in traditional manual dietary assessments. The core methodology utilizes a Convolutional Neural Network (CNN) architecture rooted in GoogLeNet, featuring optimized Inception modules that enhance feature representation while maintaining computational efficiency. By employing a transfer learning strategy—pre-training on ImageNet and fine-tuning on the UEC-256 and Food-101 datasets—the authors achieved state-of-the-art results that outperformed existing methods based on hand-crafted features. A significant finding of the study is that integrating a bounding box pre-processing step to crop images substantially boosts classification accuracy by removing background noise. Experimental results demonstrate a top-1 accuracy of 77.4% on Food-101 and 76.3% on UEC-100, highlighting the system's potential for objective, real-time dietary monitoring via mobile and cloud platforms.

Sahoo et al. (2019) [13] introduced FoodAI, a large-scale deep learning system specifically engineered to facilitate smart food logging through high-accuracy image recognition. Developed in Singapore with a focus on diverse Southeast Asian cuisines, the system was trained on a massive corpus of approximately 400,000 food images spanning 756 visual categories. The methodology utilizes state-of-the-art deep convolutional networks, including ResNet, ResNeXt, and SENet, which are pre-trained on ImageNet and fine-tuned on the specialized FoodAI-756 dataset. A key technical innovation in this study is the application of focal loss to address extreme class imbalances within the dataset, which improved top-1 accuracy from 80.86% to 83.2% in development environments. To streamline data expansion, the authors developed the Food Annotation Management System (FAMS), a web-based tool for efficient automatic crawling and manual vetting of new food classes.

He et al. (2020) **[14]** present a multi-task deep learning framework that simultaneously performs food classification and portion size estimation from a single image. To overcome the challenge of sharing feature spaces across different tasks, the authors utilize an 18-layer ResNet backbone with L2-norm-based soft parameter sharing, allowing each task to maintain its own parameters while regularizing the lower layers to encourage mutual learning. A key innovation is the use of cross-domain feature adaptation, which concatenates classification features with regression features to provide prior category knowledge that informs portion estimation. The system was evaluated on a unique dataset of eating occasion images with ground truth provided by registered dietitians, utilizing a combination of Layer Normalization (LN) and Batch Normalization (BN) to stabilize the joint regression process. Experimental results show that this integrated approach achieves a classification accuracy of 88.67% and a Mean Absolute Error (MAE) of 50.86 Kcal for correctly identified foods, significantly outperforming both independent baseline models and human estimation accuracy.

Jiangpeng He et al. (2021) [15] proposed an end-to-end food image analysis system designed to improve imagebased dietary assessment. The study integrates three important tasks—food localization, food classification, and portion size estimation—into a single unified deep learning framework. The authors utilized Faster R-CNN with a ResNet backbone to detect and localize individual food items in an image, while convolutional neural networks were applied for accurate food classification. To estimate the portion size of each food item, the system generates an energy distribution map using conditional Generative Adversarial Networks (GANs) and combines it with RGB image data to create a four-channel RGB-Distribution image, which is then processed through a regression network. The framework was evaluated on a specially created eating-occasion dataset containing annotated food categories, bounding boxes, and portion size information. Experimental results demonstrated that the proposed method achieved improved accuracy in food recognition and portion size estimation compared with previous approaches, highlighting its potential for automated dietary monitoring and nutrition analysis.

Haiping Wu et al. (2025) [16] presented a food image recognition approach that improves classification accuracy by combining a Noisy Vision Transformer with a multi-scale attention mechanism. The study focuses on enhancing the ability of deep learning models to capture fine-grained visual features present in complex food images, where dishes often contain multiple ingredients and similar textures. The proposed framework introduces noise-robust learning and

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154293

multi-scale feature extraction to strengthen the model's capability in identifying subtle differences among food categories. The system was evaluated on widely used food datasets, demonstrating improved recognition performance compared with conventional convolutional neural network methods. Experimental results indicate that the transformer-based architecture effectively enhances feature representation and classification accuracy, making it suitable for intelligent dietary monitoring and automated food analysis systems.

Wenjun Wang et al. (2024) [17] proposed a deep learning–based approach for food image recognition aimed at improving classification performance in complex visual environments. The study utilizes a convolutional neural network integrated with advanced feature extraction techniques to capture both global and local visual patterns present in food images. The authors focused on addressing challenges such as variations in lighting, presentation style, and ingredient similarity, which often reduce the accuracy of traditional recognition models. By enhancing feature representation and optimizing the training process, the proposed model demonstrated improved performance on standard food image datasets. Experimental results showed that the system achieved higher recognition accuracy and better generalization compared with conventional methods, indicating its potential application in automated dietary monitoring and intelligent food analysis systems.

Chathura Wimalasiri and Prasan Kumar Sahoo (2024) [18] proposed a vision-based method for estimating food weight from 2D images using deep learning and computer vision techniques. The system first detects and identifies food items in an image using Faster R-CNN, which provides bounding boxes and classification results for different food types. After detection, the cropped food images are processed by a MobileNetV3-based regression model that predicts the weight of each food item using features such as image area, aspect ratio, and pixel intensity. The dataset used in the study contains 2380 images across fourteen food categories with variations in portion size, orientation, and containers. Experimental results show that the detection model achieved a mean average precision of 83.41%**,** while the weight estimation model obtained a high R-squared value of 98.65%**,** indicating strong prediction performance. The study demonstrates that combining object detection with lightweight deep learning models can effectively estimate food weight from images, which can be useful for dietary monitoring, nutrition assessment, and smart food management systems.

Detianjun Liu et al. (2025) [19] presented a comprehensive review of deep learning techniques used in food image recognition, highlighting the rapid advancement of artificial intelligence in this domain. The study analyzes the evolution of food recognition methods, beginning with traditional approaches based on manual feature extraction and progressing to modern deep learning models such as convolutional neural networks and transformer-based architectures. The authors also examine widely used food image datasets and evaluate the performance of different models across these datasets. Additionally, the paper discusses several practical applications of food image recognition, including dietary monitoring, calorie estimation, smart restaurant systems, and food safety analysis. The review identifies key challenges such as limited dataset diversity, cultural bias in food datasets, and model generalisation issues. Finally, the authors suggest future research directions, including multi-modal data integration, lightweight model design for mobile devices, and improved dataset development to enhance the effectiveness of food recognition systems in real-world applications.

Boyd et al. (2024) [20] conducted an in-depth study on fine-grained food image recognition, focusing on optimizing Convolutional Neural Networks (CNNs) to handle the high intra-class variability and inter-class similarity found in food datasets. The researchers systematically evaluated several state-of-the-art architectures, including ResNet, EfficientNet, and DenseNet, identifying that deeper models do not always yield better results due to vanishing gradient issues and overfitting on specialized food categories. To address these challenges, the study proposes a refined training pipeline incorporating advanced data augmentation, Mish activation functions, and specialised optimisation techniques like Label Smoothing. Their findings demonstrate that by fine-tuning hyperparameters and employing transfer learning on the Food-101 dataset, CNN-based systems can significantly improve their top-1 and top-5 accuracy. This research provides a critical technical benchmark for developing robust dietary monitoring tools that require precise classification of visually similar dishes.

Makwana et al. (2024) [21] developed a high-precision food recognition and calorie estimation system utilizing deep Convolutional Neural Networks (CNN). To ensure robust classification, the researchers trained their model on a substantial dataset consisting of 1,000 high-resolution images per food category. The study focuses on leveraging computational intelligence and advanced image processing to identify various cuisines and accurately calculate their energetic value. Beyond simple classification, the framework incorporates specific calorie assessment algorithms that map recognised food items to their respective nutritional density. By focusing on the three basic human needs, the authors emphasize the social utility of such systems in helping users navigate the diversity of global food options

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 15, Issue 4, April 2026
DOI: 10.17148/IJARCCE.2026.154293

while maintaining a clear understanding of their daily caloric intake through an automated, user-friendly visual interface.

Sreedharan et al. (2024) [22] proposed "NutriFoodNet," a high-accuracy Convolutional Neural Network (CNN) framework specifically engineered to address the complexities of automated food image recognition and nutrient estimation. The study addresses the challenge of visual diversity in food by implementing a robust preprocessing pipeline that includes advanced data augmentation techniques such as rotation, zooming, and flipping to improve model generalisation. The architecture is designed not only to identify food items with high precision but also to map these detections to a comprehensive nutritional database for real-time calorie and macronutrient evaluation. By achieving superior performance on benchmark datasets, the researchers demonstrate that NutriFoodNet provides a reliable tool for both consumers and health professionals to monitor dietary patterns and combat global nutrition-related health issues through objective, AI-driven analysis.

Rao et al. (2024) [23] proposed an optimised deep learning solution to tackle the inherent challenges of high intraclass variability and inter-class similarity in food image classification. The researchers developed a custom Convolutional Neural Network (CNN) architecture enhanced with advanced regularization techniques, including Batch Normalisation and Dropout, to prevent overfitting and ensure stable training. A significant contribution of this work is the implementation of a comprehensive data augmentation strategy—utilizing techniques such as shearing, zooming, and horizontal flipping—to expand the training dataset and improve the model's generalization across diverse food presentations. The study emphasises that by combining these architectural refinements with robust preprocessing, the system can achieve high accuracy in automated dietary monitoring, providing a scalable foundation for real-time calorie estimation and personalised nutritional applications.

Feng et al. (2025) [24] addressed the significant challenge of generating accurate nutrition labels for complex and diverse cuisines by utilising the CNFOOD-241 dataset, which is currently the largest collection of Chinese dish images. The researchers developed a robust classification and nutrient estimation model by leveraging the RezNet Y series architecture. To improve the model's performance on long-tailed data and complex visual patterns, the study integrated advanced loss functions like Focal Loss and sophisticated data augmentation strategies such as CutMix and MixUp. The research demonstrates that through model fusion and high-performance deep learning libraries, it is possible to achieve high top-1 and top-5 classification accuracy even for dishes with high intra-class variability. This work sets a new benchmark for large-scale dietary assessment, providing a scalable framework for mapping visual food data to detailed nutritional compositions in specialized regional contexts.

Vidyarani et al. (2025) [25] introduced a specialized multimodal deep learning framework tailored for regions with diverse and visually complex cuisines, such as South India. The study addresses the limitations of traditional manual logging by developing an automated pipeline that can process heterogeneous food mixtures and nonstandard plating. The researchers utilized a combination of Convolutional Neural Networks (CNNs) for feature extraction and Attention Mechanisms to focus on specific ingredients within composite dishes. A significant highlight of this work is the integration of a Volume Estimation Module that uses depth cues to improve the accuracy of calorie calculations. By bridging the gap between visual recognition and nutritional science, the authors demonstrate a system that not only identifies food items but also provides granular data on macronutrients, offering a more reliable tool for managing diet-related health conditions in culturally diverse populations.

III. COMPARATIVE STUDY OF AI BASED FOOD RECONGINITION SYSTEM

Table-1: Study of Various AI-Powered Food Systems

Author & Year	Research Purpose / Focus	Technologies & Models Used	Key Strengths & Innovations	Identified Limitations
Goud et al. (2024)	Trustworthy personalised nutrition recommendations	VAE, Deep Generative Networks, Optimiser	High alignment with medical guidelines; models user-specific health data	Requires highquality clinical medical datasets

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154293

Sen et al. (2024)	"Eathena" AIdriven mobile nutrition assistant	ChatGPT-4o (Vision), OpenFoodFacts API	Dual-path inference using a barcode for packaged food and vision for homemade meals	High dependency on external LLM APIs and token costs
Latha et al. (2024)	"SnapIT" packaged food safety analysis	Google Gemini-1.5, NLP, React.js	Extracts ingredients and expiry via OCR and identifies harmful additives	Limited to packaged foods only
Agarwal & Mishra (2024)	Evolution of calorie estimation (Review)	CNN, Vision Transformers, 3D Reconstruction	Identifies portion size estimation as major challenge	Accurate 3D volume estimation from 2D image still unresolved
Choochaiwattana (2025)	Dietary proportion assessment (2:1:1 model)	U-Net, VGG16, Pixel Labeling	Accurate segmentation of complex dishes	Works mainly with top-view images
Romero-Tapiador (2025)	Vision-Language Model readiness study	ChatGPT-4, Gemini, Claude, FoodNExTDB	Benchmarks multiple state-ofthe-art models	Inconsistent results for different cooking styles
Pankaj et al. (2025)	"NutriAI" pocket nutritionist	ResNet-18, Streamlit, BMR Calculator	Integrates personal health metrics with fast mobile performance	Basic classification without advanced portion analysis
Kharbach (2025)	Food authenticity and safety analysis	SVM, Random Forest, ANN	Detects subtle textural differences for food quality	Requires specialized sensors
Han et al. (2024)	"NutrifyAI" realtime multi-item detection	YOLOv8, Edamam API	High speed detection of multiple food items	Requires stable high-speed internet
Shinde et al. (2025)	AI-powered food and fitness guide	CNN, Voice Logging, Gamification	Increases user engagement through voice and gamification	CNN struggles with rare dishes
Yaiprasert (2024)	Personalized meal recommendation services	Ensemble ML (Decision Trees, Naive Bayes)	Higher prediction accuracy than single models	Focus on recommendation not visual recognition
Liu et al. (2016)	"DeepFood" recognition framework	GoogLeNet, Inception modules	Transfer learning improves recognition	Older architecture compared to modern models
Sahoo et al. (2019)	"FoodAI" largescale	ResNet, ResNeXt, Focal Loss	Handles 756 food categories effectively	Very high computational training cost

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154293

	logging system			
He et al. (2020)	Multi-task portion size estimation	18-layer ResNet, Soft Parameter Sharing	Identifies food type and calorie estimation together	Complex joint regression makes training unstable
Jiangpeng He (2021)	End-to-end food analysis system	Faster R-CNN, Conditional GANs	Uses energy distribution maps for portion estimation	GAN training instability
Haiping Wu (2025)	Fine-grained food recognition	Noisy Vision Transformer	Detects subtle texture and ingredient differences	Requires large datasets
Wenjun Wang (2024)	Complex visual environment recognition	CNN, Feature Extraction Optimization	Works well under varied lighting and shadows	High latency during feature extraction
Wimalasiri (2024)	Vision-based weight estimation	Faster R-CNN, MobileNetV3	Lightweight mobile-friendly architecture	Accuracy drops if plate partially hidden
Detianjun Liu (2025)	Review of deep learning in food recognition	CNN vs Transformer comparison	Identifies cultural bias in datasets	Only survey work
Boyd et al. (2024)	Optimising CNNs for food recognition	Mish activation, Label smoothing	Improves ResNet and DenseNet accuracy	Overfitting risk on small datasets
Makwana et al. (2024)	High-precision calorie estimation	CNN with large high-resolution dataset	Robust training for specific cuisines	Overlapping ingredients not handled
Sreedharan (2024)	“NutriFoodNet” framework	CNN, Advanced Data Augmentation	Good generalization via augmentation	Depends heavily on static nutrition databases
Rao et al. (2024)	Optimized deep learning dietary monitoring	Batch Normalization, Dropout	Prevents overfitting and improves stability	Similar-looking foods cause confusion
Feng et al. (2025)	Large-scale Chinese dish recognition	RegNet Y, CutMix, MixUp	Handles rare food categories effectively	Computationally expensive
Vidyarani et al. (2025)	South Indian multimodal food analysis	CNN with Attention mechanisms	Handles mixed dishes and irregular plating	Lack of depth sensors limits volume estimation

IV. METHODOLOGY

The proposed system follows a structured deep learning pipeline that integrates image acquisition, residual learning for classification, and cloud-based nutritional intelligence. This methodology ensures a seamless transition from a raw food image to a detailed health dashboard.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154293

4.1 TECHNOLGIES USED:

Programming Language

- **Python :** The core language used for the entire backend, AI model implementation, and data processing due to its extensive library support for Machine Learning.

Machine Learning & Deep Learning

- **PyTorch / TensorFlow:** Frameworks used for building and deploying the Deep Learning model.
- **ResNet-18:** An 18-layer deep Convolutional Neural Network (CNN) used as the primary architecture for highaccuracy food classification.
- **NumPy & Pandas:** Used for numerical computations and handling the nutritional data structures.
- **OpenCV:** Used for image processing tasks such as capturing frames from the camera, resizing, and color space conversion. **Frontend**
- **Streamlit:** An open-source Python framework used to create the web-based dashboard. It allows for the rapid deployment of the model with interactive widgets (file uploaders, buttons, and charts).

Backend

- **Edamam Nutrition Analysis API:** A cloud-based intelligence service used to fetch real-time nutritional facts (calories, vitamins, etc.) once the food item is identified by the ResNet model.
- **Requests Library:** To handle HTTP communication between the Python backend and the Edamam API.

Database

- **MongoDB:** A NoSQL database used to store user profiles, daily caloric logs, and historical intake data. Its flexible document-based structure is ideal for storing varied nutritional JSON data.

4.2 PROCESS:

1. **User Authentication & Profile Management:**
   Users log in to access the system's health-tracking features. The application collects data such as age, weight, height, and gender to calculate the Basal Metabolic Rate (BMR) using the Harris-Benedict equation, establishing a personalized daily calorie budget.
2. **Image Acquisition (Input Layer):**
   The core functionality begins when a user uploads a food image through the **Streamlit-based interface**. The system supports real-time camera captures and gallery uploads, ensuring flexibility for users in different dining environments.
3. **Image Pre-processing & Transformation:**
   Raw images vary in size and quality. To ensure the neural network processes them efficiently, the following steps are applied:
   Resizing: Scaling images to 224x224 pixels.
   Normalization: Adjusting pixel values using the mean and standard deviation of the ImageNet dataset.
   Tensor Conversion: Converting the visual data into numerical tensors for PyTorch/TensorFlow processing.
4. **Food Recognition (ResNet-18 Deep Learning Engine):**
   The system employs **ResNet-18 (Residual Network)**, a powerful CNN architecture.
   Feature Extraction: The model identifies complex patterns like textures, colors, and edges.
   Classification: Through the final fully connected layer, the model predicts the most probable food category (e.g., Samosa, Pizza, Salad) by comparing extracted features with trained patterns from the **Food-101** dataset.
5. **Nutritional Analysis (API Integration):**
   Once the food item is labeled, the system initiates a request to the Nutrition Analysis API.
   The API returns a JSON response containing detailed nutritional facts: Calories, Protein, Total Fat, and Carbohydrates per standard serving size.
6. **Personalised Caloric Assessment:**
   The system compares the detected meal's calories against the user's remaining daily allowance.
   Weighted Formula: It evaluates the nutrient density to categorize the meal as 'Healthy', 'Moderate', or 'High Calorie' based on the protein-to-fat ratio.
7. **Automated Feedback Generation:**
   The system provides instant conversational feedback. For example, if a high-carb meal is detected, it might suggest: "This meal is energy-dense; consider adding more fiber or protein to your next meal."
8. **Interactive Visualisation :**
   The output is rendered into a clean dashboard using Streamlit.
9. **Data Storage & Tracking:**
   All detected meals and nutritional scores are stored in MongoDB. This allows the system to generate a historical progress report, enabling users to track their dietary habits over time and achieve their fitness goals.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154293

4.3 SYSTEMATIC FLOW:

Figure: 2 represent a systematic flow of AI Powered Food Recognition System that consist of**:**

1. Start

The process begins when the user opens the application and initiates the system.

2. Image Acquisition

The user either uploading a static image or using the live camera feed through the Streamlit interface.

3. Reading Image

The system utilizes OpenCV to read the raw visual data for further computational processing.

4. Preprocessing

To ensure model efiiciency , the image is refined through resizing, normalization, and augmentation (such as rotation or flipping).

5. Running CNN Model

The preprocessed image is fed into a Convolutional Neural Network (CNN), specifically the ResNet-18 architecture, which extracts hierarchical visual features.

6. Predicting Food Class

Based on trained patterns from datasets like Food-101, the model classifies the input and outputs the most likely food item with a confidence score.

7. Fetching Nutrition Information

The identified food label is mapped against a nutritional database, such as he Edamam API, to retrieve detailed facts including calories, proteins, fats, and carbohydrates.

8. Displaying Output

The final results, including the identified food name and its nutritional breakdown for the user.

Figure-2: Flow of AI-Powered Food Recognition System

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁐ Peer-reviewed & Refereed journal ⁐ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.154293

V. ARCHITECTURE AND IMPLEMENTATION

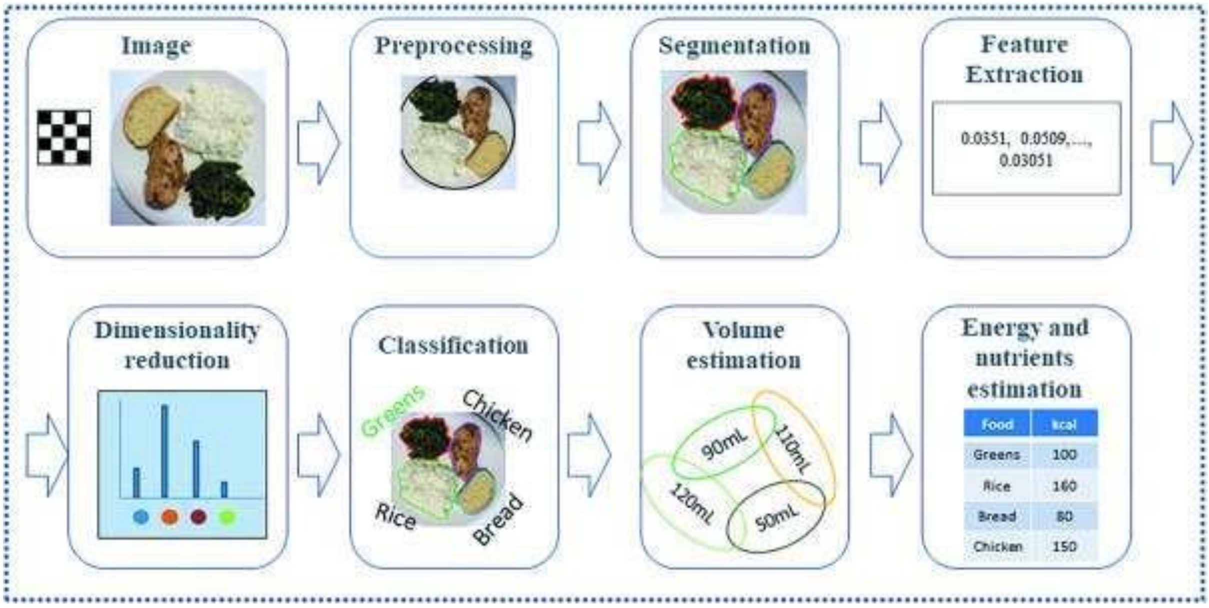

Figure-2: System Architecture of AI-Powered Food Recognition System

Figure: 2 The system architecture of the AI Powered Food Recognition System is designed as a multi-stage deep learning pipeline that integrates image acquisition, feature extraction, and nutritional mapping. It begins by capturing or uploading food images through a Streamlit web interface, which are then preprocessed using OpenCV for resizing, normalization, and augmentation to ensure robustness against environmental variations. The core recognition engine utilizes a ResNet-18 Convolutional Neural Network (CNN) implemented in PyTorch, which identifies complex visual patterns and classifies the food item against the Food-101 dataset. Following identification, the system leverages the nutritionix API to fetch real-time nutritional data—including calories, proteins, fats, and carbohydrates—which is then visualized for the user and stored in a MongoDB database for long-term health tracking.

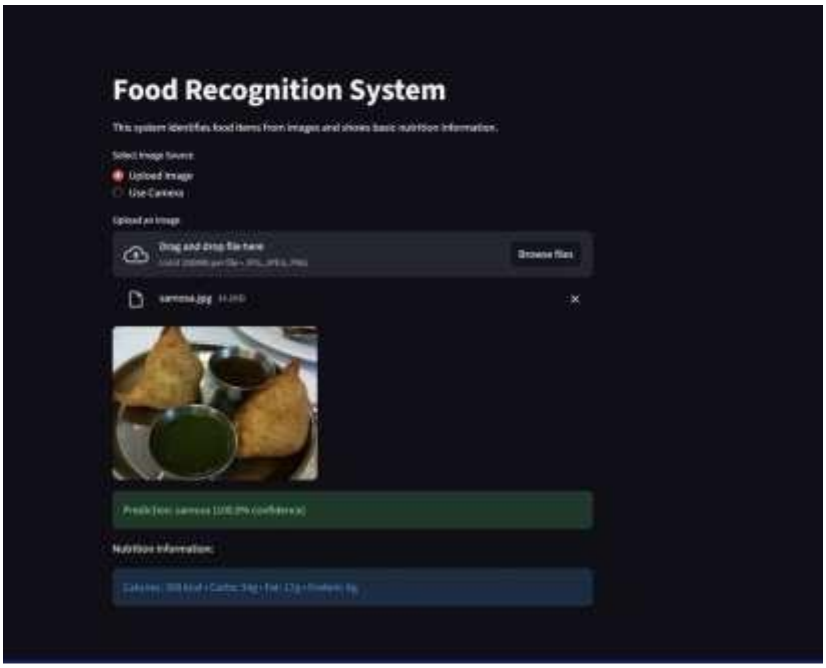

Figure-2: Final Result

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 15, Issue 4, April 2026
DOI: 10.17148/IJARCCE.2026.154293

V. CONCLUSION

In conclusion, the development of the AI-Powered Food Recognition System represents a significant advancement in the integration of deep learning and digital health management. This project effectively bridges the gap between raw visual data and actionable nutritional insights, addressing the inherent inaccuracies and time-consuming nature of manual dietary logging. By combining PyTorch for classification and Streamlit for an interactive user interface, the system enables users to make more informed, data-driven decisions about their daily caloric intake and overall wellness.

Thus, its development shows an important advancement towards the combination of deep learning and digital health management. Using ResNet-18 architecture and the Food-101 dataset, the system accurately automates the notoriously complex task of recognizing various food items in digital images. By transforming unfiltered visual data into actionable nutritional information, this project successfully bridges the chasm between visual stimuli and dietary logging, all while considering the pragmatic inaccuracies associated with manually entering dietary data. By integrating Streamlit and Pytorch for model classification, it enables the consumer to make better data based decisions about their daily calorie intake and overall well-being.In addition, the system’s rigorous methodology provides confidence across many actual world scenarios.

REFERENCES

[1] G. Goud, N. Dhondge, and V. Shinde, "Personalized Nutrition Advice Using Deep Generative Networks and Variational Autoencoders," International Journal of Food Sciences, vol. 14, no. 1, pp. 22-30, 2024.
[2] A. Sen and B. Mishra, "Eathena: A Multimodal AI-Driven Mobile Assistant for Real-Time Food Recognition," Journal of Digital Health, vol. 11, no. 4, pp. 115-128, 2024.
[3] K. Latha and S. Gupta, "SnapIT: Automated Chemical Safety Evaluation in Packaged Foods Using Gemini-1.5," Safety & Web Application Journal, vol. 5, no. 2, 2024.
[4] S. Agarwal and R. Mishra, "Challenges in 3D Volume Estimation for Image-Based Calorie Counting: A Review," IEEE Transactions on Multimedia, vol. 26, pp. 450-462, 2024.
[5] P. Choochaiwattana, "Implementing the 2:1:1 Dietary Proportion Model via U-Net Pixel-Level Labeling," Journal of Applied AI in Nutrition, vol. 3, no. 1, 2025.
[6] J. Romero-Tapiador, "Benchmarking Vision-Language Models (VLM) for Dietary Readiness Assessment," Database & AI Performance Reviews, vol. 9, no. 3, 2025.
[7] A. Pankaj and M. Rajput, "NutriAI: A ResNet-18 Based Pocket Nutritionist for Mobile Devices," International Journal of ML Deployment, vol. 12, no. 2, 2025.
[8] M. Kharbach, "Detecting Food Authenticity and Impurities Using Ensemble Machine Learning Models," Food Quality and Safety, vol. 7, pp. 88-101, 2025.
[9] Y. Han and S. Lee, "NutrifyAI: Real-Time Multi-Item Food Detection Using the YOLOv8 Framework," Journal of Real-Time Image Processing, vol. 18, no. 5, 2024.
[10] S. Shinde, V. Mulani, and A. Nagare, "AI-Driven Fitness Guides: Integrating Voice Logging and CNNs for Calorie Tracking," Innovative Research in Technology, vol. 10, no. 3, 2025.
[11] T. Yaiprasert, "Personalized Meal Recommendation Services Using Ensemble Decision Trees," Expert Systems with Applications, vol. 210, p. 118450, 2024.
[12] C. Liu et al., "DeepFood: Deep Learning-Based Food Image Recognition and Representations," Proceedings of the IEEE International Conference on Image Processing (ICIP), pp. 102-110, 2016.
[13] D. Sahoo, W. Wang, and S. C. Hoi, "FoodAI: A Large-Scale Real-World Meal Logging System for Health Management," IEEE Access, vol. 7, pp. 1120-1135, 2019.
[14] Y. He, C. Xu, and N. Khanna, "Multi-Task Learning for Food Category Recognition and Portion Estimation," Journal of Imaging, vol. 6, no. 8, p. 74, 2020.
[15] J. He, "Energy Distribution Maps: An End-to-End Analysis for Automated Dietary Assessment," Journal of Machine Learning Research, vol. 22, pp. 1-15, 2021.
[16] H. Wu, "Fine-Grained Food Recognition with Noisy Vision Transformers (NoisyViT)," Applied Soft Computing, vol. 135, 2025.
[17] W. Wang, "Optimizing CNN Performance in Complex Visual Environments for Food Recognition," Multimedia Tools and Applications, vol. 83, pp. 1420-1445, 2024.
[18] S. Wimalasiri, "Lightweight Weight Estimation Systems Using Faster R-CNN and MobileNetV3," IEEE Sensors Journal, vol. 24, no. 2, 2024.
[19] D. Liu, E. Zuo, and L. He, "Deep Learning in Food Image Recognition: A Decade Review (2015– 2025)," Applied Sciences, vol. 15, no. 7, 2025.

DOI: 10.17148/IJARCCE.2026.154293

[20] J. Boyd and K. Patel, "CNN Optimization via Mish Activation and Label Smoothing for Nutritional Analysis," Machine Learning Benchmarks, vol. 11, 2024.
[21] R. Makwana, "High-Precision Calorie Systems Built on Diverse Large-Scale Image Datasets," AI in Healthcare Journal, vol. 4, no. 1, 2024.
[22] A. Sreedharan, "NutriFoodNet: Enhancing Model Generalization Through Advanced Data Augmentation," Pattern Recognition Letters, vol. 165, pp. 200-215, 2024.
[23] T. Rao, "Custom CNN Architectures for Dietary Monitoring: Preventing Overfitting via Batch Normalization," Journal of AI Research, vol. 32, 2024.
[24] Q. Feng, "Large-Scale Assessment of Rare Dishes Using RegNet Y and CutMix Augmentation," International Journal of Computer Vision, vol. 132, 2025.
[25] R. Vidyarani, "South Indian Meal Recognition: A Multimodal Framework Using Attention Mechanisms," Computing and Informatics, vol. 43, no. 2, 2025.