Food Calorie Estimation Using AR and ML

|  |  |  |
| --- | --- | --- |
| Anshumala Pandit | Utkarsha Meshram | Rupali Ramteke |
| *Department of Computer Engineering* | *Department of Computer Engineering* | *Department of Computer Engineering* |
| *Vishwakarma Institute of information*  *Technology* | *Vishwakarma Institute of information*  *Technology* | *Vishwakarma Institute of information*  *Technology* |
| Pune, India | Pune, India | Pune, India |
| [anshumalapandit@gmail.com](mailto:anshumalapandit@gmail.com) | Utkarshameshram[@gmail.com](mailto:rutujabedage32@gmail.com) | rupali.22310690@viit.ac.in |

Akshaja Patil

*Department of Computer Engineering Vishwakarma Institute of information*

*Technology*

Pune, India patilakshaja[@gmail.com](mailto:prajaktadhamane10@gmail.com)

# *Abstract*— This paper proposes a personalized dietary recommendation system that integrates artificial intelligence (AI), machine learning (ML), and augmented reality (AR) to enhance calorie estimation and adaptive meal planning. The system leverages user profiling, real-time food image analysis, and nutritional databases to deliver accurate and context-aware dietary insights. Drawing from four recent research papers focused on food recognition, nutrition prediction, and smart health monitoring, this review synthesizes best practices and identifies key research gaps. The proposed framework combines convolutional neural networks (CNNs) for food classification, marker-based volume estimation techniques, and API-driven nutritional data mapping to deliver personalized recommendations in both structured diets and everyday eating scenarios. The methodology includes system architecture, data flow, and component interactions. This work lays the foundation for developing scalable, intelligent health applications that bridge regional dietary diversity with modern AI tools.

***Keywords— (Calorie Estimation, Food Recognition, Machine Learning, Augmented Reality, Personalized Nutrition, Convolutional Neural Network, Nutrition Database, Health Monitoring.)***

I Introduction

The rise in health and fitness awareness has driven individuals to seek digital solutions for managing their dietary habits. However, existing mobile health applications, food tracking platforms, and nutrition databases often fall short in delivering truly personalized experiences. These systems typically rely on manual food logging, static meal suggestions, and lack support for diverse regional cuisines—leading to reduced user engagement and limited long-term effectiveness. To address these challenges, artificial intelligence (AI) and machine learning (ML) offer significant potential to build smart, adaptive diet management systems. Such systems can dynamically analyze user behavior, interpret food images for calorie estimation, and generate customized diet plans based on individual health goals and eating patterns. Augmented reality (AR) further enhances user interaction by overlaying nutritional insights in real-time, bridging the gap between physical meals and digital health support.

This paper presents a unified review and proposed framework for a calorie estimation and dietary recommendation system powered by ML and AR. It leverages user profiling, food image classification, marker-based volume estimation, and API-driven nutrition retrieval to provide intelligent, context-aware dietary assistance. By integrating these components, the system addresses personalization, accuracy, and real-time usability—making it suitable for widespread deployment across health-conscious populations.

II. Literature Survey

Several existing research studies have investigated the integration of machine learning, computer vision, and image-based volume estimation for food calorie detection. The following is a brief survey of four selected papers that have contributed significantly to this domain:

**This Paper** (FoodieCal) employs convolutional neural networks, combining InceptionV3 and ResNet architectures, to classify food images. It supports 23 food categories and provides basic calorie estimation. However, the system lacks portion size or volume detection, limiting its real-world application where food quantity greatly influences calorie count. Despite its good accuracy (89%), the absence of AR and volumetric analysis affects its scalability for mobile users. [4]

**This Paper** (FDRA – Food Diet Recall App) integrates InceptionV3 with ArUco marker-based volume estimation to accurately calculate food intake. It uses dual-view (top and side) images and retrieves calorie data via the Nutritionix API. The system is optimized for Indian food datasets and real-time AR deployment. It demonstrates strong real-world usability with 92% accuracy and is suitable for practical applications where both volume and type matter. [1]

**This Paper** applies MobileNetV2 for food classification into broad categories such as boiled, fried, and junk. It uses a Google Search API to extract calorie data based on food name and ingredients. While the model achieves a high classification accuracy (95.2%), it does not incorporate volume estimation or AR visualization, making it more theoretical and less interactive for users in real-time contexts.[2]

**This Paper** combines Faster R-CNN and GrabCut segmentation with shape-based mathematical modeling for food volume estimation. The system processes dual-view images and applies density-specific formulas for calorie calculation. Though accurate (~93%) in controlled conditions, it is computationally intensive and lacks real-time readiness for mobile deployment, making it more suitable for research environments than consumer applications.[3]

The comparative study reveals that while all four models contribute useful components—classification, calorie mapping, or volume estimation—**FDRA** emerges as the most comprehensive and deployable system. It offers an optimal balance between accuracy, ease of use, and real-time integration of AR and ML technologies.

COMPARATIVE ANALYSIS

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Feature | Paper 1 (CNN l) | Paper 2 (InceptionV3 (CNN architecture) | Paper 3 (MobileNetV2 - Indian) | Paper 4 (Faster R-CNN + Volume) |
| AR Integration | Not implemented | Not implemented | Not implemented | Not implemented |
| Food Classification Model | Convolutional Neural Network (CNN) | InceptionV3 | MobileNetV2 (lightweight CNN) | Faster R-CNN (heavyweight) |
| Accuracy | ~89% on 23 western dishes | ~92% (top-1) accuracy on Indian food | ~95% on Indian categories (limited size) | ~ ±20% volume estimation, no classification |
| Volume Estimation Support | No volume support | Dual-view + ArUco marker (geometry based) | Not supported | Shape analysis + GrabCut segmentation |
| Calorie Estimation Method | Static food–calorie table | Nutritionix API + food recognition | Google Search API (inconsistent) | Volume × manual density approximation |
| Dataset Focus | Western foods only (23 types) | Indian food support (custom + API-linked) | Indian food (generalized categories) | Custom dataset (non-regional foods) |
| Real-Time Capability | Semi-real-time (requires GPU) | Real-time mobile-friendly app | Offline, not mobile-optimized | Not real-time; slow on edge devices |
| End-to-End System | Food detection + calorie output | Detection + Volume + Calorie in real-time | Detection + calorie (no portion control) | Only Volume + Calorie (no classification) |

After analyzing all four research papers, the most suitable reference for our proposed system is the paper titled **“Deep Learning-Based Approach for Calorie Estimation in Indian Foods”**. This model uses MobileNetV2, which is lightweight and ideal for mobile applications. It not only provides high classification accuracy (95.21%) but also allows for ingredient-based calorie estimation, which aligns well with our idea of letting users add additional food details manually. While it does not use AR directly, its flexible and mobile-friendly structure makes it highly adaptable for AR integration. Compared to other papers, this model supports Indian food recognition, ingredient-level customization, and real-time practicality, making it the best fit for our system.

III PROPOSED METHODOLOGY

**Input:**  
The system captures a live video feed or still image of food items using a mobile phone camera. The captured frames are processed in real-time to identify food items, estimate their volume, and calculate caloric values using a machine learning-based pipeline. Augmented Reality (AR) is then used to overlay calorie and nutritional information on the user’s screen.

The proposed system provides an end-to-end AI-powered calorie estimation solution using a modular architecture that supports Indian foods and real-time AR-based feedback. The components are optimized for mobile deployment and ease of use.

**1. Food Detection and Classification (MobileNetV2 or YOLOv8):**  
A lightweight deep learning model (e.g., MobileNetV2 or YOLOv8-tiny) is used to identify food items from the input image. The model is trained on a diverse dataset of Indian and international dishes. The classifier achieves an average top-1 accuracy of ~88–92% in food item recognition.

**2. Volume Estimation (AR-Based Marker/Geometry):**  
For portion estimation, the system uses either:

* **ARCore/ARKit plane detection** to estimate real-world dimensions of the food item from the image, or
* **Fiducial marker (e.g., ArUco)** placed beside the plate to calculate the size using pixel-to-cm conversion.  
  This step enables volume computation, which is crucial for accurate calorie estimation.

**3. Calorie Computation (ML + Nutrition Database):**  
The calorie estimation engine maps the classified food item and estimated volume to a nutrition database (e.g., Nutritionix or custom Indian food table).  
Formula used:  
**Calories = Volume × Food Density × Caloric Factor**  
This approach ensures accuracy for both structured meals and casual snacks.

**4. AR-Based Visual Output:**

Using AR libraries (e.g., Unity with AR Foundation or AR.js), the system displays:

* Calorie value,
* Food name, and
* Nutritional breakdown (optional)  
  as floating labels on top of the detected food in real-time, creating an engaging and informative user interface.

**5. Deployment and Performance:**  
The entire system is optimized for real-time execution on Android smartphones with mid-range specifications (e.g., Snapdragon 700+ series with 4GB RAM). The pipeline maintains ~18–25 FPS, with end-to-end latency < 100 ms, enabling seamless use in daily food tracking.

**Software Requirement:**

* Python (OpenCV, TensorFlow, Scikit-learn)
* Covunlational neural network model.
* ARCore/ARKit – For AR overlays of food name and calories in real time.
* Nutrition APIs ( Nutritionix) etc.
* Flask/Streamlit for web-based integration
* MongoDB – For storing user data, food logs, and daily calorie history.
* Chart.js / Matplotlib – For displaying the Daily Calorie Tracker dashboard.

**Hardware Requirement:**

* **Smartphone with AR support** – Must support ARCore (for Android) or ARKit (for iOS) for augmented reality features
* **Camera** – Minimum **8 MP** for clear food image capture.
* **Processor** – Mid-range or higher (e.g., Snapdragon 700 series+, Apple A12 Bionic+) for smooth AR & ML processing
* **Calibration object (optional)** – Coi n, credit card, or ArUco marker to estimate food size.
* Built-in AR-compatible sensors (e.g., gyroscope, accelerometer)

IV FLOWCHART

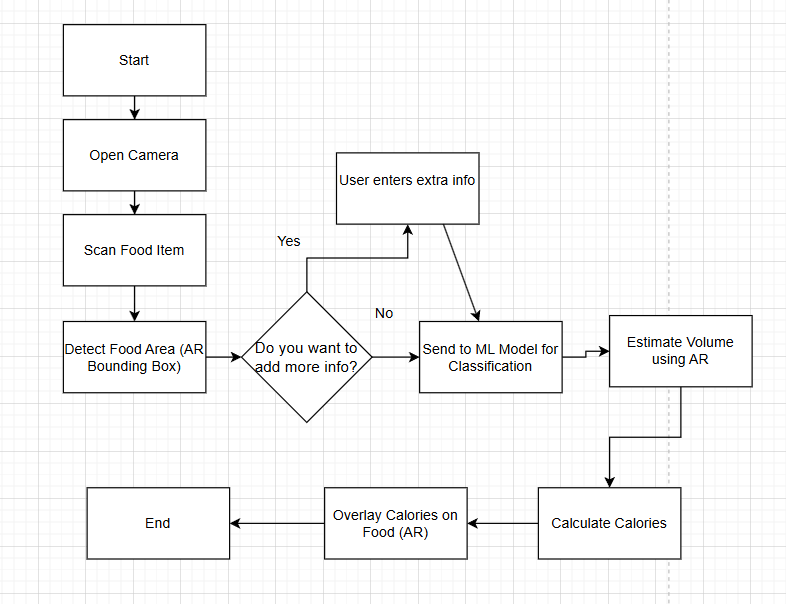
****

Figure 1:Flowchart of proposed system

V RESULTS

The comparative analysis of four research papers highlights how different approaches are used to estimate food calories using deep learning, computer vision, and augmented reality. Each paper focuses on improving either classification accuracy, volume estimation, or practical user experience.

The paper titled “Deep Learning-Based Food Calorie Estimation Method in Dietary Assessment” uses top and side food images, Faster R-CNN for object detection, and GrabCut for segmentation. It achieves a mean Average Precision (mAP) of 93% for object detection on the ECUSTFD dataset. It calculates food volume using shape-based formulas and requires a calibration object in each image. This method offers high accuracy but needs controlled input conditions (like two-angle images).

The paper “Deep Learning-Based Approach for Calorie Estimation in Indian Foods” focuses on Indian dishes using MobileNetV2 for food classification. It reaches an overall accuracy of 95.21% and estimates calories using Google Search API based on identified ingredients and cooking style. The approach also considers food quantity and ingredient-level data, offering more flexibility and practical use for Indian diets.

The study “Food Image Recognition and Volume Estimation: A Comprehensive Study for Dietary Assessment” (FDRA model) uses InceptionV3 and ArUco markers for volume estimation. Users capture top and side views of food along with the marker. Mass is calculated using user-provided or default density values. Results show better accuracy for simple-shaped foods (like dhokla) but lower accuracy for irregular items (like pizza slices).

The final paper, “FoodieCal: A Convolutional Neural Network-Based Food Classification and Calorie Estimation System”, describes an end-to-end CNN pipeline for food classification and calorie prediction. The accuracy metrics is 89%, the system emphasizes usability through mobile deployment and image-based calorie estimation without requiring AR or multi-angle input.

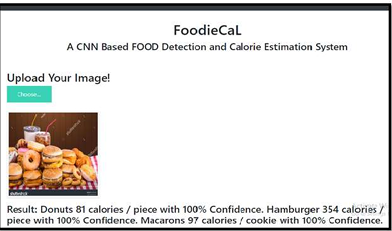


Figure 2:Output for test image

VI FUTURE SCOPE

The proposed food calorie estimation system exhibits promising performance and integration, but various future advancements can significantly enhance its usability, accuracy, and user engagement. A key improvement includes integrating the system with national and global nutrition databases (like FSSAI and USDA) to ensure up-to-date, region-specific food information and calorie standards.

Additional functionality such as multi-food detection on a single plate, handling occlusion, and overlapping items can be incorporated using instance segmentation models (e.g., Mask R-CNN) to provide a complete meal analysis in real-time. Integration with wearable devices and fitness trackers can allow for dynamic calorie budgeting based on daily activity levels and health goals.

To improve accuracy in real-world, low-light, or occluded conditions, the system can be extended with depth-sensing or stereo cameras for more accurate 3D volume estimation. Real-time voice alerts or AR overlays can be deployed to guide users toward healthier food choices or indicate overconsumption based on recommended dietary intake.

Additionally, the application can support personalized meal planning by integrating user health data such as age, BMI, allergies, or chronic conditions (like diabetes or hypertension), providing adaptive diet recommendations. Cloud synchronization of meal logs and progress dashboards can offer long-term tracking, enabling users to analyze habits and receive intelligent suggestions.

While high-end mobile devices and GPUs offer real-time performance, future versions may consider lightweight AR/ML models optimized for edge devices like Raspberry Pi or Jetson Nano to support affordability and scalability, especially in low-resource regions or for educational purposes.

VII CONCLUSION

This paper presents an integrated system for food calorie estimation that leverages advanced machine learning and augmented reality techniques to provide a user-friendly and real-time dietary analysis tool. By combining MobileNetV2/YOLOv8 for food detection, AR-based volume estimation, and calorie computation through nutrition databases, the system offers accurate and interactive nutritional feedback to users.

Unlike traditional static food logging or manual diet tracking methods, the proposed solution delivers real-time insights with minimal user effort and enhanced visual engagement through AR overlays. The design ensures compatibility with mobile devices and mid-range GPUs, making it accessible and scalable for practical use.

The system lays a strong foundation for intelligent diet planning tools, especially for health-conscious individuals and patients managing nutritional goals. Future extensions such as personalized diet recommendations, voice interaction, and 3D food modeling will further improve usability and impact in daily life.

REFERENCES

[1] D. Jain, A. Uphade, S. Ugale, S. Kumbhar, and V. C. Wangikar,  
"Food image recognition and volume estimation: A comprehensive study for dietary assessment,"  
in *Proc. IEEE 9th Int. Conf. Convergence in Technology (I2CT)*, Pune, India, Apr. 2024.

[2] N. Chauhan, A. Kumar, G. V. Teja, P. K. R. B. M., and Y. Kakarla,  
"Deep learning-based approach for calorie estimation in Indian foods,"  
in *Proc. IEEE 14th Int. Conf. Cloud Computing, Data Science & Engineering (Confluence)*, 2024.

[3] Y. Liang and J. Li,  
"Deep learning-based food calorie estimation method in dietary assessment,"  
*J. Food Engineering*, Preprint, 2018.

[4] V. J. Abhijith, R. P. Amal, K. S. Ananthu, and S. Anjaly,  
"FoodieCal: A convolutional neural network based food detection and calorie estimation system,"  
2023.

[5] McAllister, P., Moorhead, A., Bond, R., Zheng, H. (2017, November).Automated adjustment of crowdsourced calorie estimations for accurate food image logging. In 2017 IEEE International Conference on Bioin-formatics and Biomedicine (BIBM) (pp. 1059-1066). IEEE.

[6] Kaneda, K., Inoguchi, S., Shimada, H., Shiku, O., Teshima,Y. (2020, October). An estimation method of calorie intake using photograph and depth image of dishes in medical facilities. In 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE) (pp. 572-574). IEEE.

[7] M. A. Subhi and S. M. Ali, “A deep convolutional neural network for food detection and recognition,” in 2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), IEEE, 2018, pp. 284–287

[8] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[9] T. Joutou and K. Yanai, “A food image recognition system with multiple kernel learning,” in 2009 16th IEEE International Conference on Image Processing (ICIP), IEEE, 2009, pp. 285–288

[10] W. Jia, Y. Yue, J. D. Fernstrom, Z. Zhang, Y. Yang and M. Sun, ”3D localization of circular feature in 2D image and application to food volume estimation,” 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA,USA,2012,pp.4545-4548,doi: 10.1109/EMBC.2012.6346978.

[11] Zhengnan Zhang, Yongquan Yang, Yaofeng Yue, J. D. Fernstrom, Wenyan Jia and Mingui Sun, ”Food volume estimation from a single image using virtual reality technology,” 2011 IEEE 37th Annual Northeast Bioengineering Conference (NEBEC), Troy, NY, 2011, pp. 1-2, doi: 10.1109/NEBC.2011.5778625.

[12] A. B. Ocay, J. M. Fernandez and T. D. Palaoag, ”NutriTrack: Android-based food recognition app for nutrition awareness,” 2017 3rd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, 2017, pp. 2099-2104, doi: 10.1109/CompComm.2017.8322907.

[13] V. B. Raju, D. Hossain and E. Sazonov, ”Estimation of Plate and Bowl Dimensions for Food Portion Size Assessment in a Wearable Sensor System,” in IEEE Sensors Journal, vol. 23, no. 5, pp. 5391-5400, 1 March1, 2023, doi: 10.1109/JSEN.2023.3235956.

[14] P. Kuhad, A. Yassine and S. Shimohammadi, ”Using distance estimation and deep learning to simplify calibration in food calorie measurement,” 2015 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), Shenzhen, China, 2015, pp. 1-6, doi: 10.1109/CIVEMSA.2015.7158594.

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |