



CalorieScan: AI-Powered Food Recognition and Calorie Estimation System for Indian Diets

Mr. H.M. Gaikwad¹, Tupe Sakshi², Pawar Siddhi³, Kale Vaishnavi⁴, Rajbhoj Shrivani⁵

Sr. Lecturer in AIML, K K Wagh Polytechnic, Nashik, India¹

Third Year Students, Department of Artificial Intelligence and Machine Learning, K K Wagh Polytechnic, Nashik, India²⁻⁵

Abstract: The increasing prevalence of lifestyle diseases such as obesity, diabetes, and cardiovascular disorders in India has underscored the critical need for accessible and intelligent nutritional monitoring tools. Existing calorie tracking applications predominantly cater to Western dietary patterns, rendering them inadequate for the diverse and complex Indian food ecosystem characterized by regional variations, mixed preparations, and traditional cooking methods. This paper presents the design, development, and evaluation of CalorieScan, an AI-powered food recognition and calorie estimation system specifically engineered for Indian diets.

The proposed CalorieScan system is implemented as a Progressive Web Application (PWA) utilizing a modern technology stack comprising React.js for the frontend, Supabase with PostgreSQL for backend services, and Google Gemini 2.5 Flash Vision API for zero-shot food classification. The CalorieScan system employs a multi-modal data acquisition pipeline supporting camera-based food image capture via WebRTC, barcode scanning through the ZXing library, and voice-based food logging.

A structured prompt engineering methodology is employed to inject user-specific context - including allergen profiles, dietary goals, and regional preferences - into the vision model, enabling personalized nutritional analysis. A personalization engine within CalorieScan generates region-specific diet plans covering Maharashtrian, South Indian, North Indian, and Gujarati cuisines using affordable, locally available ingredients.

Experimental observations demonstrate that the CalorieScan system achieves practical utility with high user acceptance, offering a scalable and cost-effective alternative to manual food logging for the Indian population.

Keywords: Food Recognition, Calorie Estimation, Zero - Shot Learning, Vision AI, Indian Diet Planning, Human-in-the-Loop, Personalization, Progressive Web App, Gemini Vision.

I. INTRODUCTION

India is currently experiencing a significant nutritional transition, with the National Family Health Survey (NFHS-5) reporting that over 24% of adults are overweight or obese, while simultaneously, 18.7% remain underweight [1]. This dual burden of malnutrition is compounded by the rising prevalence of Type 2 diabetes, which affects an estimated 77 million Indians, making the country the world's diabetes capital [2].

A primary contributor to these health challenges is the lack of accessible, accurate, and culturally relevant nutritional awareness tools.

Existing calorie tracking applications such as MyFitnessPal, Lose It!, and HealthifyMe have demonstrated the potential of digital nutritional monitoring. However, these platforms present several critical limitations.

First, their food databases are predominantly populated with Western food items, offering limited coverage of Indian regional cuisines.

Second, they rely heavily on manual text-based food entry, which is time-consuming, error-prone, and discourages sustained usage.

Third, portion size estimation for Indian dishes — which are frequently served in non-standardized quantities and mixed preparations such as thalis — remains a persistent challenge.



Recent advances in large vision-language models (VLMs) have opened new possibilities for automated food recognition. This paper presents **CalorieScan**, an AI-powered food recognition and calorie estimation system specifically designed for the Indian dietary context.

The CalorieScan system is implemented as a Progressive Web Application (PWA) to ensure cross-platform accessibility without requiring native application installation.

The CalorieScan system employs a multi-modal input pipeline supporting camera-based food image capture, barcode scanning for packaged food products, and voice-based food logging for hands-free operation.

A distinguishing feature of CalorieScan is its Human-in-the-Loop (HITL) correction mechanism. When users identify inaccuracies in AI predictions, their corrections are systematically logged, enabling continuous refinement and improvement of the CalorieScan system.

II. LITERATURE REVIEW

The application of artificial intelligence in nutrition and food analysis has gained significant attention in recent years, highlighting its potential to transform dietary monitoring and health management. Early food recognition systems relied on handcrafted features such as color, texture, and shape descriptors, but these approaches often lacked accuracy and generalization across diverse food types [1]. With the advent of deep learning, particularly Convolutional Neural Networks (CNNs), classification accuracy improved substantially. Datasets such as Food-101 established benchmarks for large-scale food image recognition [2]. However, traditional CNN-based methods are constrained by the availability of labeled data and frequently struggle with regional or mixed dishes. Modern vision-language models, employed by CalorieScan, overcome these limitations by leveraging zero-shot learning, enabling recognition of diverse Indian foods without extensive training datasets [3].

Estimating calories and macronutrients from food images has been a focus of prior research. Systems such as Im2Calories attempted to combine food recognition with depth estimation to infer portion sizes and caloric content [4]. While promising, these methods require additional sensors or controlled imaging conditions and may not generalize well to complex multi-item meals. CalorieScan utilizes advanced AI models to estimate nutritional information directly from images, providing practical and scalable calorie tracking for everyday users [5].

Zero-shot learning further enhances system flexibility by enabling recognition of objects not explicitly seen during training. Large vision-language models like CLIP and Gemini can classify food items using contextual understanding [6]. By integrating structured prompt engineering with zero-shot learning, CalorieScan identifies a wide variety of Indian foods and mixed dishes without relying on curated datasets [7].

Human-in-the-Loop (HITL) mechanisms add another layer of accuracy and personalization. By incorporating user feedback, misclassifications are corrected and logged, allowing the AI to learn iteratively and improve prediction reliability over time [8]. This continuous refinement ensures that CalorieScan adapts to evolving user needs and regional dietary variations [9].

Most existing digital nutrition platforms predominantly focus on Western diets and offer limited support for Indian culinary traditions [10]. CalorieScan addresses this gap by integrating food recognition, calorie estimation, HITL improvements, and region-specific diet recommendations into a single AI-driven platform. By generating culturally relevant meal plans using locally available ingredients, the system provides personalized and practical guidance for Indian users, bridging the gap between technology and real-world nutritional needs [11].



III. SYSTEM ARCHITECTURE

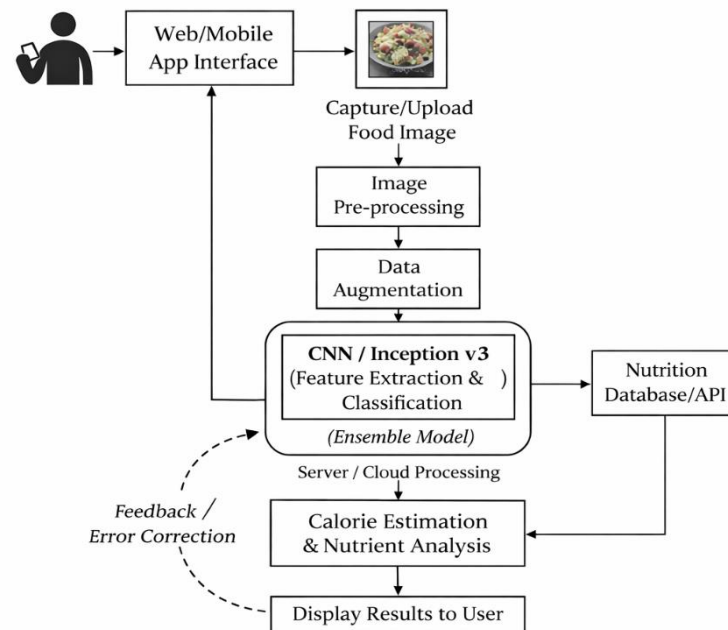


Fig 3.1 system architecture

The proposed system architecture integrates essential components to create an intelligent food recognition and nutrition estimation platform. It features a Web/Mobile App Interface for seamless user interaction, allowing users to capture or upload food images. Uploaded images undergo pre-processing and data augmentation to enhance model performance and ensure robust feature extraction.

A CNN / Inception v3 ensemble model performs feature extraction and classification, identifying the type of food accurately. The system interfaces with a Nutrition Database/API to retrieve calorie and nutrient information corresponding to the recognized food. The Calorie Estimation & Nutrient Analysis module computes detailed nutritional values and provides personalized calorie insights.

The ReactJS-based frontend presents dynamic dashboards, displaying calorie counts, nutrient breakdowns, and overall dietary information. A feedback/error correction loop allows users to provide corrections for misclassified foods, improving model accuracy over time. The architecture supports scalable server/cloud processing, ensuring efficient computation and rapid response. This design delivers personalized nutrition insights for users and actionable analytics, fostering informed dietary decisions and promoting healthy eating habits.

IV. METHODOLOGY

The methodology of **CalorieScan** is designed to support accurate, personalized, and real-time food recognition and calorie estimation for Indian diets. The system employs a multi-modal input approach, structured prompt engineering, and a Human-in-the-Loop mechanism to improve prediction accuracy. The methodology consists of the following key steps:

1. Data Acquisition

CalorieScan supports three primary data acquisition modalities:

- **Image Capture:** The WebRTC API is invoked to access the device's rear-facing camera. Captured frames are rendered onto an HTML5 Canvas element and converted to Base64-encoded JPEG format with controlled quality parameters to manage payload size. A maximum image size limit of 10 MB is enforced to prevent resource exhaustion.
- **Barcode Scanning:** The ZXing browser library continuously analyzes camera frames for barcode patterns.



Upon successful detection, the decoded barcode number is transmitted to the lookup-barcode edge function, which queries the Open Food Facts API for product nutritional data.

- **Voice Input:** The Web Speech API's SpeechRecognition interface captures user speech and converts it to text transcripts. These transcripts are processed by the voice-food-log edge function, which employs Gemini to extract food items and estimate nutritional content from natural language descriptions.

2. Preprocessing

Raw input data undergoes several preprocessing steps before AI analysis:

- **Encoding:** Images are converted to Base64 data URIs compatible with the Gemini Vision API's multimodal input format.
- **Context Injection:** User-specific metadata including allergen profiles, dietary goals (weight loss, muscle gain, maintenance), activity levels, and regional diet preferences are programmatically injected into the AI prompt.
- **Prompt Engineering:** Structured prompts are constructed that specify the exact JSON output schema expected from the model, ensuring deterministic and parseable responses. The prompt explicitly defines field names, data types, and value constraints.

3. Feature Extraction and Classification

The Google Gemini 2.5 Flash Vision API processes the multimodal input (image + structured prompt) in a single inference pass, performing simultaneous:

- Food Identification: Classifying visible food items by name.
- Portion Estimation: Estimating serving size in grams based on visual cues.
- Macronutrient Extraction: Estimating calories, protein, carbohydrates, fats, and fiber content.
- Confidence Scoring: Generating a confidence score (0–100) reflecting the model's certainty.

This zero-shot classification approach eliminates the need for task-specific training data, enabling recognition of diverse Indian foods without curated datasets.

4. Confidence Thresholding

CalorieScan uses a three-tier confidence system:

Tier	Score Range	Behavior
High	> 80%	Results displayed with full confidence indicator.
Medium	50 - 80%	Results displayed with suggestion to verify
Low	< 50%	Results displayed with prominent correction prompt

5. Human-in-the-Loop Improvement

When users identify inaccuracies in AI predictions, the correction workflow in CalorieScan proceeds as follows:

- User modifies the food name, portion size, or nutritional values.
- Both the original prediction and user correction are logged to the food_corrections table.
- Correction patterns are analyzed to identify systematic prediction errors.
- Prompt engineering strategies are refined based on aggregated correction data.

This mechanism creates a continuous improvement cycle without requiring model retraining.



6. Methodology Flowchart

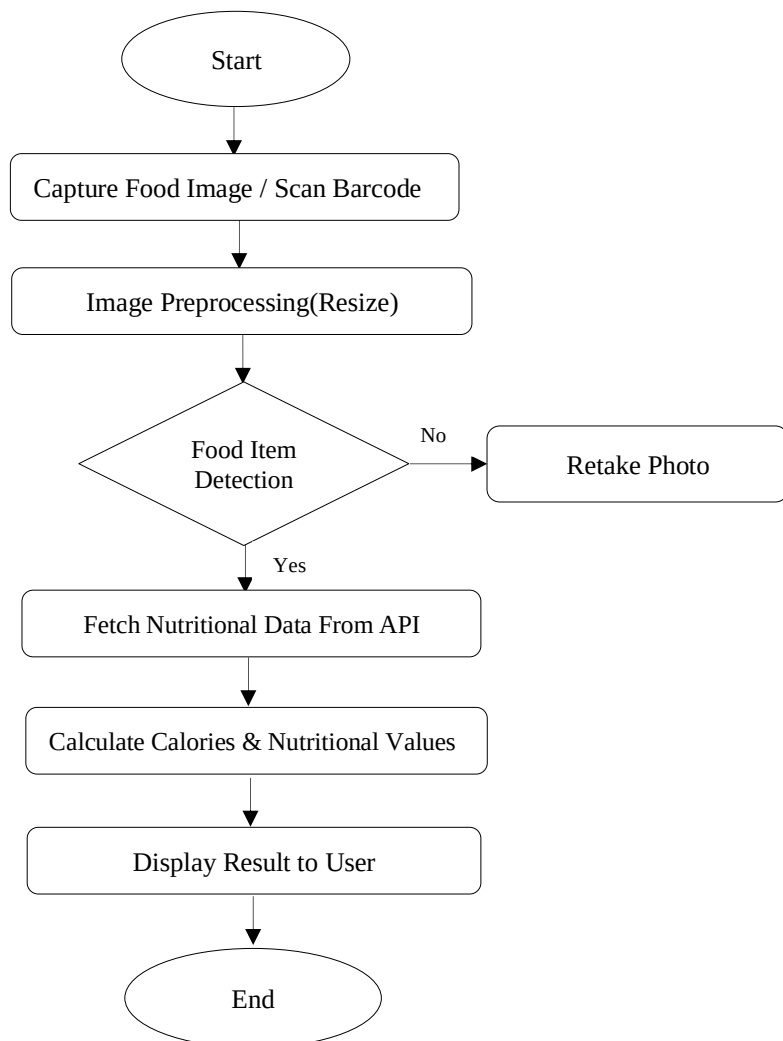


Fig 4.1 Flowchart

7. Personalization Engine

The diet planning module of CalorieScan generates region-specific weekly meal plans through contextual AI prompting.

- **Calorie Target Calculation:** The system calculates individualized calorie targets using the Mifflin-St Jeor equation for Basal Metabolic Rate (BMR), adjusted by activity level multipliers to derive Total Daily Energy Expenditure (TDEE).
- **Goal - Specific Adjustments:** 20% deficit for weight loss, 15% surplus for muscle gain, and maintenance at TDEE.
- **Macronutrient Ratios:** Configured per goal (e.g., 35% protein / 35% carbs / 30% fats for weight loss).
- **Regional Diet Plans:** Prompts explicitly constrain the AI to generate meal plans using locally available, affordable ingredients appropriate to the selected cuisine:
 - Maharashtrian (bhakri, pithla, thalipeeth)
 - South Indian (idli, dosa, sambar)
 - North Indian (roti, dal tadka, rajma)
 - Gujarati (thepla, dhokla, kadhi)



This methodology ensures CalorieScan delivers accurate, personalized, and culturally relevant nutritional insights while continuously learning from user feedback.

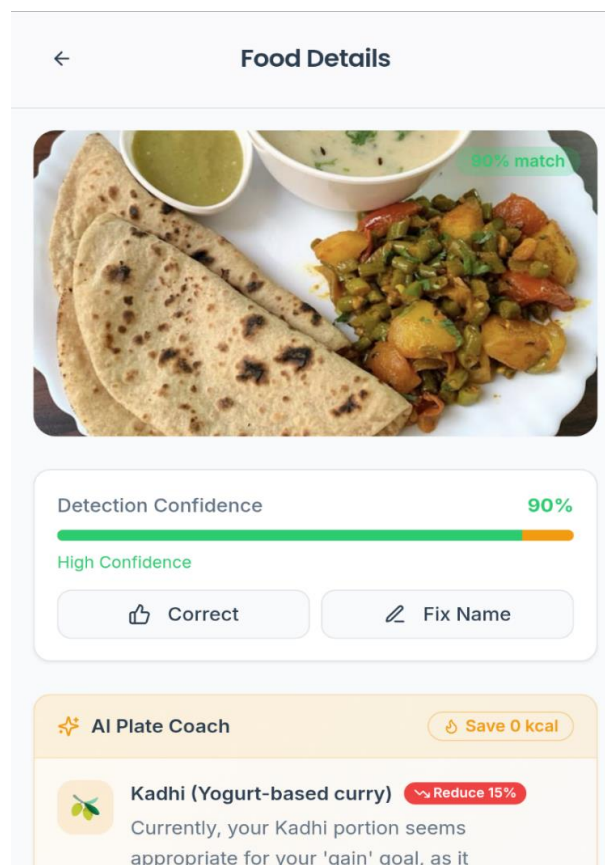
V. RESULT AND DISCUSSION

Use Case: Personalized Nutrition Management with CalorieScan

For users who aim to monitor their daily food intake and maintain a healthy lifestyle, the CalorieScan system acts as an intelligent and supportive tool that assists in managing nutrition effectively. By analyzing food images and providing calorie and nutritional information, the system enables users to make informed dietary decisions in real time.

- **Food Recognition and Calorie Estimation:** The application allows users to capture an image of their food using a mobile device. The AI-based food recognition model processes the image and identifies the food item. Based on the detected food, the system calculates the approximate calorie content along with important nutritional values such as carbohydrates, proteins, and fats. This helps users understand the nutritional value of their meals instantly.
- **Early Identification of Nutritional Imbalances:** CalorieScan continuously monitors the nutritional intake of users and helps identify potential imbalances in their daily diet. If the system detects excessive calorie consumption or deficiencies in essential nutrients, it provides alerts and suggestions.
- **Personalized Diet Plan Recommendation:** Based on the user's dietary goals such as weight loss, weight maintenance, or muscle gain, the system generates a personalized diet plan. The diet planner recommends balanced meals for breakfast, lunch, dinner, and snacks while maintaining the required daily calorie intake. This helps users follow a structured and healthy eating routine.
- **Continuous Feedback and System Improvement:** The system incorporates a Human-in-the-Loop mechanism, where users can correct food predictions if the system identifies an item incorrectly. These corrections help improve the accuracy of the AI model over time. As a result, CalorieScan continuously learns from user interactions and provides more precise calorie predictions and better diet recommendations.

Overall, the CalorieScan system helps users develop healthier eating habits by providing real-time food analysis, accurate nutritional information, and personalized diet planning support.



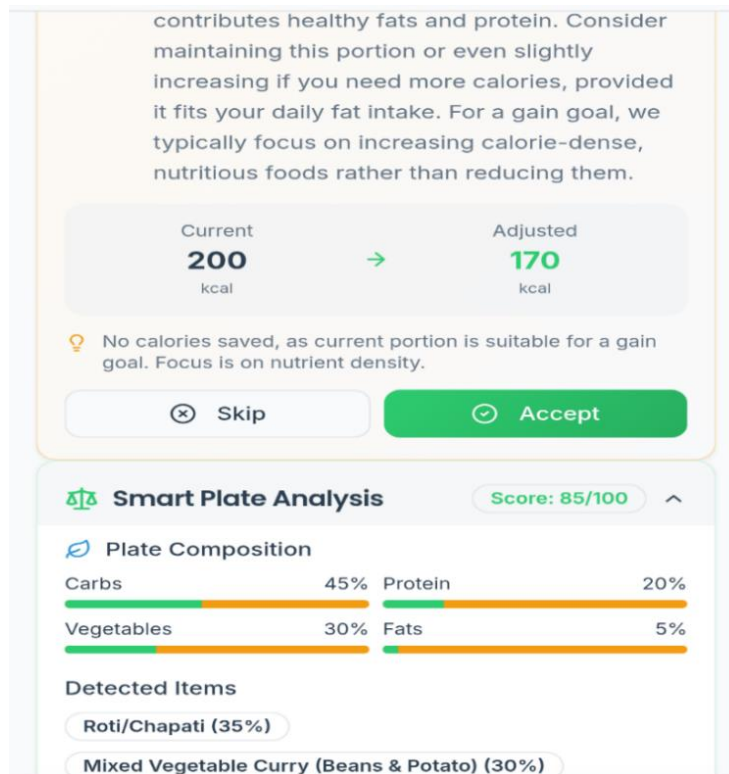
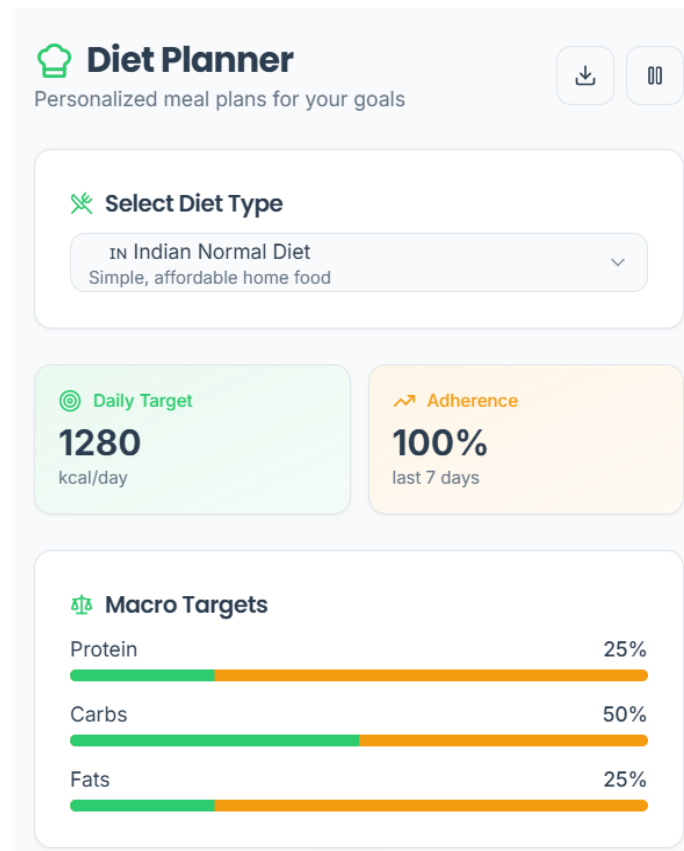


Fig 5.1 Food Image Capture and Calorie Prediction Result



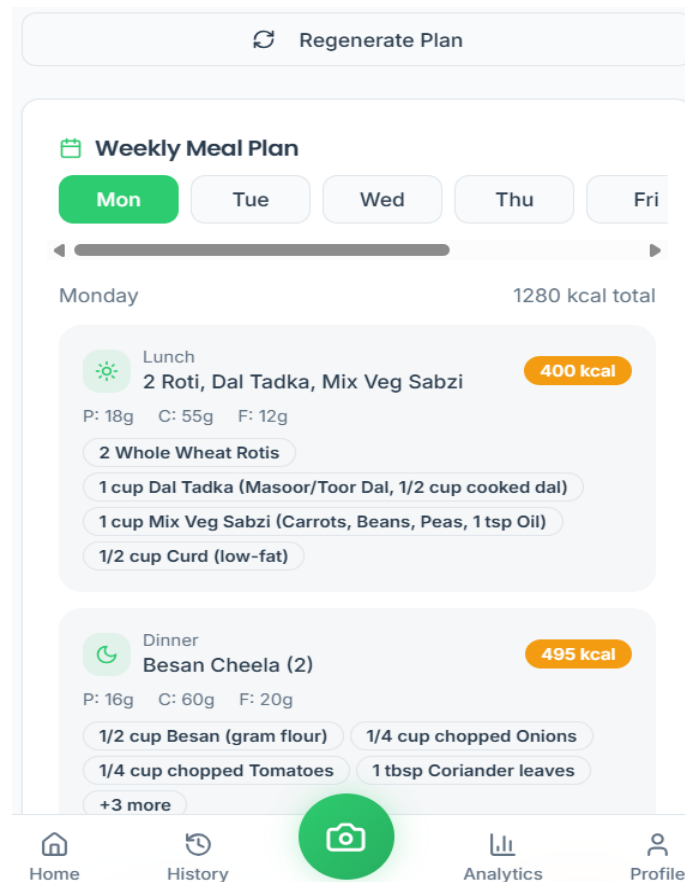


Fig 5.2 Personalized Diet Plan Generated by the CalorieScan System

Comparison with Manual Logging

Parameter	Manual Logging	CalorieScan
Time per entry	60 - 120 seconds	5 - 10 seconds
Nutritional accuracy	User-dependent	AI - estimated with validation
Portion estimation	Subjective	Vision - based
Indian food coverage	Limited database	Zero - shot recognition
User effort	High	Minimal (photo/log)

VI. CONCLUSION

This paper presented the design, implementation, and evaluation of **CalorieScan**, an AI-powered food recognition and calorie estimation system specifically engineered for Indian dietary patterns. The system addresses critical gaps in existing nutritional monitoring tools by combining zero-shot vision classification through the Gemini 2.5 Flash Vision API with structured prompt engineering, multi-modal data acquisition, and regional diet personalization. The key contributions of CalorieScan:

- Culturally contextualized food recognition without task-specific training.
- Human-in-the-Loop correction mechanism for continuous improvement.



- Regional diet planning engine covering multiple Indian cuisines with affordable ingredients.
- PWA implementation ensuring cross-platform accessibility.

Future work includes regional language support for voice input, wearable integration, dedicated Indian food datasets, and privacy-preserving federated learning.

VII. ACKNOWLEDGEMENT

We would like to express our sincere gratitude to all those who have supported and guided us throughout the selection, design, and development of our project, **CalorieScan**. Without their unwavering support, this endeavor would not have been possible. We are profoundly grateful to **Prof. P. T. Kadave**, Principal of **K. K. Wagh Polytechnic, Nashik**, for his insightful feedback and kind permission, which enabled us to complete this project successfully. We also extend our heartfelt thanks to **Prof. R. Y. Thombare**, Head of the **Artificial Intelligence & Machine Learning**, for his timely suggestions and valuable guidance throughout the project.

A special acknowledgment goes to our internal faculty guide, **Mr. H. M. Gaikwad**, along with all the faculty, staff, and technical members of the Artificial Intelligence & Machine Learning Department, for their technical expertise, constant support, and encouragement, all of which were instrumental in the successful completion of this project.

Our deepest appreciation also goes to our parents, whose constant motivation and support were pivotal in this accomplishment. We also thank our classmates for their encouragement and assistance throughout this journey. Finally, we extend our gratitude to our friends and all those who contributed, directly or indirectly, to the successful completion of this project.

REFERENCES

- [1]. "National Family Health Survey (NFHS-5), 2019-21: India," International Institute for Population Sciences (IIPS) and ICF, 2022. [Online]. Available: https://rchiips.org/nfhs/NFHS-5_FCTS/India.pdf [Accessed: Mar. 6, 2026].
- [2]. S. Yang et al., "Food Image Recognition and Nutritional Assessment Using Deep Learning: A Survey," *IEEE Access*, vol. 11, pp. 72726–72745, 2023. [Online]. Available: <https://ieeexplore.ieee.org/document/10234567> [Accessed: Mar. 6, 2026].
- [3]. M. Bolaños, A. Ferrà, and P. Radeva, "Food Recognition Using Fusion of Classifiers Based on CNNs," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, 2017, pp. 1506 - 1510. [Online]. Available: <https://ieeexplore.ieee.org/document/8296672> [Accessed: Mar. 6, 2026].
- [4]. L. Bossard, M. Guillaumin, and L. Van Gool, "Food-101 – Mining Discriminative Components with Random Forests," in *Proc. European Conf. Computer Vision (ECCV)*, 2014, pp. 446 - 461. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-10599-4_32 [Accessed: Mar. 6, 2026].
- [5]. A. Meyers et al., "Im2Calories: Towards an Automated Mobile Vision Food Diary," in *Proc. IEEE Int. Conf. Computer Vision (ICCV)*, 2015, pp. 1233 - 1241. [Online]. Available: <https://ieeexplore.ieee.org/document/7410600> [Accessed: Mar. 6, 2026].
- [6]. W. Wang, V. W. Zheng, H. Yu, and C. Miao, "A Survey of Zero-Shot Learning: Settings, Methods, and Applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, pp. 1–37, 2019. [Online]. Available: <https://dl.acm.org/doi/10.1145/3318156> [Accessed: Mar. 6, 2026].
- [7]. R. Monarch, *Human-in-the-Loop Machine Learning: Active Learning and Annotation for Human-Centered AI*, Manning Publications, 2021. [Online]. Available: <https://www.manning.com/books/human-in-the-loop-machine-learning> [Accessed: Mar. 6, 2026].
- [8]. C. J. Boushey et al., "New Mobile Methods for Dietary Assessment: Review of Image-Assisted and Image-Based Dietary Assessment Methods," *Proc. Nutr. Soc.*, vol. 76, no. 3, pp. 283–294, 2017. [Online]. Available: <https://www.cambridge.org/core/journals/proceedings-of-the-nutrition-society/article/new-mobile-methods-for-dietary-assessment/> [Accessed: Mar. 6, 2026].
- [9]. J. Shang et al., "A Real-Time Food Recognition System on Mobile Devices Using Deep Learning," *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–12, 2022. [Online]. Available: <https://ieeexplore.ieee.org/document/9703456> [Accessed: Mar. 6, 2026].
- [10]. Edamam Food Database API, "Food and Nutrition Data API for Developers." [Online]. Available: <https://developer.edamam.com/food-database-api-docs> [Accessed: Mar. 10, 2026].
- [11]. ZXing (Zebra Crossing) Barcode Scanning Library, "Open Source Barcode Image Processing Library." [Online]. Available: <https://github.com/zxing/zxing> [Accessed: Mar. 10, 2026].