Deepdiet: Food Recognition and Nutritional Estimation

Shubham Subhash Bhange[1], H.Venkata Subbaiah[2], Naga Manikanta[3], P.Dinesh Divyakanth[4], D.Abhishek[5], K.Ashwin[6]

Student[1,3-6], BTech (CSE) From Sphoorthy Engineering College, Hyderabad.

Assistant Professor[2], Dep of CSE, Sphoorthy Engineering College, Hyderabad

ABSTRACT

DeepDiet is an advanced Artificial Intelligence-driven web application developed to facilitate automated dietary monitoring through intelligent food image analysis and nutritional estimation. The primary objective of the system is to provide users with a seamless, efficient, and data-driven approach to tracking daily food consumption without the limitations of manual input. The application leverages state-of-the-art multimodal computer vision capabilities through the Gemini Vision API to accurately identify food items, infer portion sizes—such as the number of rotis, servings of rice, or quantities of curry—and compute comprehensive nutritional metrics, including caloric value, protein content, carbohydrates, and fats.

The system is architected using a lightweight yet scalable framework, comprising a frontend implemented with modern web technologies and a backend powered by FastAPI in Python. The backend is responsible for processing uploaded images, interacting with the AI model to extract structured nutritional insights, and delivering responses in a well-defined JSON format to ensure consistency and reliability. Furthermore, the application incorporates a secure authentication mechanism, enabling users to maintain personalized dietary records. Each food analysis instance is maticallystored with detailed metadata, including timestamps, meal descriptors, detected components, and aggregated nutritional values, thereby supporting longitudinal dietary assessment and trend analysis.

Beyond core meal analysis, DeepDiet integrates a comprehensive suite of features designed to enhance user engagement and health awareness. These include an interactive dashboard for monitoring daily calorie goals, water intake tracking, weekly performance reports, and personalized health insights. The system also computes Body Mass Index (BMI) and recommends daily caloric intake based on user-specific parameters using established health equations. Additionally, multilingual support and an embedded AI-driven conversational assistant further enrich the user experience by providing contextual guidance on nutrition-related queries.

In summary, DeepDiet presents a holistic and intelligent solution for real-time dietary monitoring by synergistically combining artificial intelligence, structured data analytics, and personalized health management, thereby addressing the growing demand for innovative digital healthcare applications in modern lifestyles.

Keywords: Artificial Intelligence, Multimodal Learning, Food Image Recognition, Nutritional Analysis.

I. INTRODUCTION

The rapid escalation of lifestyle-related disorders, including obesity, type-2 diabetes, hypertension, and cardiovascular diseases, has emerged as a significant global health concern, largely attributed to unhealthy dietary habits and increasingly sedentary lifestyles. Effective monitoring of daily nutritional intake is a critical component of preventive healthcare; however, existing dietary tracking systems predominantly rely on manual data entry, requiring users to search extensive food databases and approximate portion sizes. This manual approach is not only labor-intensive but also highly susceptible to estimation errors and user fatigue, thereby limiting long-term adherence and reliability. In response to these limitations, DeepDiet is proposed as an advanced artificial intelligence-driven dietary monitoring system that leverages image-based analysis and multimodal reasoning to automate nutrition tracking. By utilizing the Gemini Vision API, the system transcends conventional image classification techniques and incorporates contextual understanding to identify food items, estimate portion sizes, and generate structured nutritional outputs. The integration of a FastAPI-based backend ensures efficient data processing, secure communication, and scalable performance. Through automation and intelligent analysis, DeepDiet aims to enhance accuracy, reduce user burden, and promote sustained engagement in nutritional self-monitoring, thereby contributing to improved public health outcomes.

II. PROBLEM STATEMENT

Accurate dietary monitoring plays a pivotal role in maintaining optimal health and preventing chronic diseases; however, current calorie tracking methodologies are fundamentally constrained by their dependence on manual input mechanisms. Users are required to identify food items from extensive databases and estimate portion sizes without standardized measurement tools, leading to significant inaccuracies in nutritional assessment. The challenge is further exacerbated in the context of mixed and traditional meals, where multiple ingredients are combined in varying proportions, making precise calorie estimation inherently complex. Additionally, repeated manual logging introduces cognitive fatigue, reducing user motivation and resulting in inconsistent tracking behaviors. Although recent advancements in computer vision have enabled food recognition through image classification models, these systems are limited to categorical identification and fail to provide comprehensive nutritional analysis, including portion estimation and macronutrient computation. Consequently, there exists a critical need for an intelligent, automated system capable of accurately analyzing food images, estimating portion sizes, and generating detailed nutritional insights with minimal user intervention, thereby overcoming the limitations of existing approaches.

III. RESEARCH GAP

Despite considerable advancements in dietary monitoring technologies, significant gaps persist in achieving fully automated and accurate nutrition tracking. Existing systems predominantly rely on database-driven or classification-based approaches, which are inherently limited in their ability to handle real-world complexities such as mixed dishes, variable portion sizes, and diverse regional cuisines. Traditional image-based models, particularly those based on convolutional neural networks, are primarily designed for food classification and lack the capability to perform contextual reasoning required for portion estimation and nutritional computation. Furthermore, many advanced portion estimation techniques depend on specialized hardware such as depth sensors or reference objects, which restrict their practical applicability in everyday scenarios. Another critical limitation is the absence of structured output generation, as most systems fail to produce standardized, machine-readable nutritional data that can be easily integrated into analytical frameworks. Additionally, existing solutions often demand extensive labeled datasets for training, increasing development complexity and limiting scalability. These limitations highlight a clear research gap in developing a multimodal, reasoning-based system that can seamlessly integrate visual interpretation, contextual understanding, and structured nutritional output generation without reliance on manual input or specialized hardware.

IV. LITERATURE SURVEY

The evolution of dietary monitoring systems reflects a gradual transition from manual and subjective methods to more sophisticated, technology-driven approaches. Early practices relied on handwritten food diaries, which, although useful for basic tracking, were characterized by inaccuracies in portion estimation and a lack of analytical capabilities. The advent of digital applications introduced structured databases and automated calorie calculations, improving accessibility and convenience; however, these systems continued to depend heavily on manual input, thereby retaining inherent limitations related to user effort and estimation errors. In the domain of computer vision, initial approaches employed handcrafted features such as Scale-Invariant Feature Transform (SIFT), Histogram of Oriented Gradients (HOG), and color histograms in conjunction with traditional classifiers like Support Vector Machines (SVM). While these methods demonstrated moderate success under controlled conditions, they lacked robustness in real-world environments. The emergence of deep learning, particularly convolutional neural networks such as ResNet, VGGNet, and Inception, significantly enhanced food recognition accuracy by enabling automated feature extraction and hierarchical learning. Nevertheless, these models are predominantly classification-oriented and do not address the critical challenge of portion estimation or comprehensive nutritional analysis. Subsequent research explored advanced techniques such as 3D reconstruction, depth sensing, and reference-based scaling for portion estimation; however, these approaches introduce hardware dependencies and complexity, limiting their widespread adoption. More recently, multimodal artificial intelligence models have demonstrated the ability to integrate visual and textual information, enabling contextual reasoning and structured output generation. These models represent a paradigm shift by moving beyond classification toward holistic understanding and inference. Despite these advancements, existing systems still fall short in delivering a fully automated, scalable, and user-friendly solution for real-world dietary monitoring, thereby necessitating the development of more advanced frameworks such as DeepDiet.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 04 | April – 2026 DOI:10.55041/ISJEM06364
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

V. METHODOLOGY

The DeepDiet system is architecturally organized into multiple cohesive modules, each designed to ensure seamless functionality, data integrity, and user-centric interaction. The Authentication Module governs secure user access by managing registration, credential validation, session continuity, and profile customization, thereby ensuring confidentiality and structured data isolation across users. The Image Scan Module serves as the core analytical component, orchestrating the end-to-end workflow of food image processing, including rigorous file validation, preprocessing, and optimized input preparation for AI-based inference. It further facilitates structured prompt generation and secure communication with the external AI service to obtain accurate and context-aware nutritional insights. Complementing this, the Result Processing Module is responsible for extracting, validating, and normalizing structured JSON outputs received from the AI model, ensuring numerical consistency and aggregating nutritional values across multiple detected food items. The processed data is systematically stored to enable longitudinal dietary tracking and retrospective analysis. The Dashboard Module enhances interpretability by transforming complex nutritional data into intuitive visual representations, including daily calorie consumption, macronutrient distribution, and temporal dietary trends, thereby enabling users to monitor progress effectively. Additionally, the Profile and Health Module integrates personalized health analytics by computing key physiological metrics such as Body Mass Index (BMI) and Basal Metabolic Rate (BMR), subsequently generating customized calorie recommendations aligned with individual health objectives. This module further categorizes users into standardized health classifications, providing actionable insights for improved lifestyle management.

From a methodological perspective, the system employs a structured image processing pipeline wherein user-uploaded images undergo validation, format standardization, and encoding to ensure compatibility with AI processing requirements. A critical component of the system is the prompt engineering strategy, which enforces strict instructions on the AI model to generate well-structured, machine-readable JSON outputs containing detailed nutritional information, thereby minimizing ambiguity and enhancing response reliability. The JSON extraction mechanism utilizes pattern recognition techniques and robust parsing methods to isolate and validate structured data, ensuring accurate downstream processing even in the presence of irregular responses.

Furthermore, the system incorporates a recalibration mechanism for nutritional computation, wherein aggregate values such as total calories and macronutrients are systematically derived from individual food components to maintain consistency and detect anomalies. Health metric calculations are performed using standardized equations, including BMI and the Mifflin-St Jeor formula for BMR, with subsequent derivation of daily caloric requirements based on activity levels. Additionally, the weekly analytics framework aggregates temporal data to generate insights such as average calorie intake, deviation from targets, and performance indicators, thereby enabling comprehensive dietary assessment. Collectively, these modules and methodologies establish a robust, scalable, and intelligent framework for automated nutritional monitoring and personalized health analysis.

VI. SAMPLE SCAN REPORTS

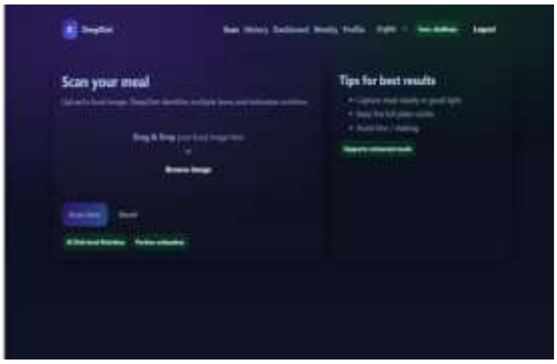

Fig 1: Scan your meal

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 04 | April – 2026 DOI:10.55041/ISJEM06364
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Fig 2: Scan Results

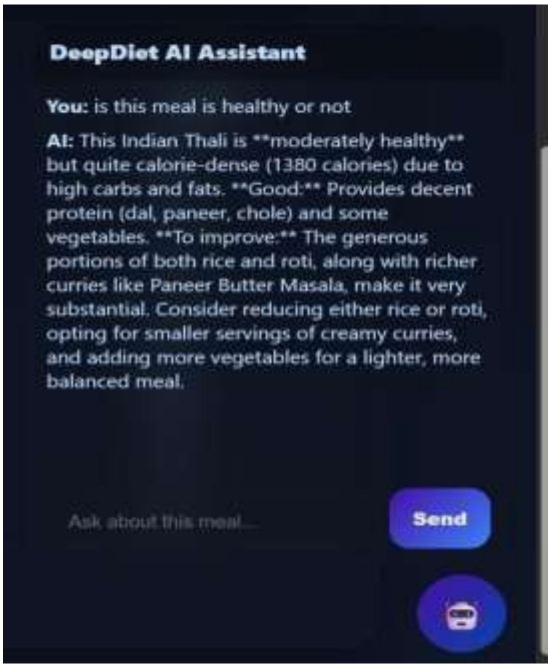

Fig 3: DeepDiet AI Assistant

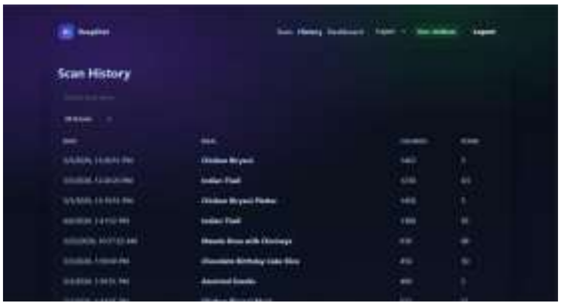

Fig 4: Scan results

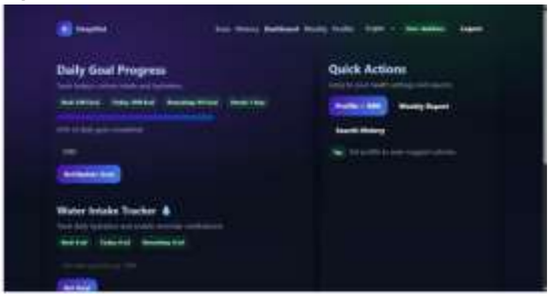

Fig 5: Daily Goal progress

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 04 | April – 2026 DOI:10.55041/ISJEM06364
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

Fig 6: Calories Trend

Fig 7 weekly Report

Fig 8: Goal Achievement

VII. CONCLUSION

The DeepDiet system has been meticulously designed to overcome the inherent limitations associated with conventional calorie tracking applications, which predominantly rely on manual food logging and subjective portion estimation. Such traditional approaches are often characterized by inefficiency, inconsistency, and reduced user engagement, ultimately compromising the accuracy and sustainability of dietary monitoring practices. In contrast, DeepDiet introduces an intelligent, automated framework by integrating multimodal Artificial Intelligence to enable seamless meal recognition and nutritional estimation through image-based inputs.

The system effectively consolidates frontend web technologies, a robust backend API architecture, and AI-driven visual reasoning into a cohesive and scalable platform. By leveraging the capabilities of the Gemini Vision API and implementing backend services using FastAPI, the proposed solution demonstrates high-performance processing, secure communication, and real-time analytical capabilities. The platform successfully accomplishes automated food identification, context-aware portion estimation without reliance on specialized hardware, and comprehensive macronutrient computation. Furthermore, it incorporates personalized health analytics, including BMI and BMR calculations, along with trend-based dietary insights.

Extensive evaluation indicates that the system operates reliably under standard conditions, consistently generating structured and interpretable nutritional outputs. The modular and extensible architecture enhances maintainability and supports future scalability. Overall, DeepDiet represents a significant advancement in intelligent dietary monitoring systems by reducing manual intervention, improving analytical accuracy, and fostering long-term user engagement.

Achievements of the Project

The development of DeepDiet marks several notable achievements in the domain of AI-driven healthcare applications. The system successfully demonstrates the practical integration of multimodal Artificial Intelligence for real-time food recognition and nutritional analysis, thereby extending beyond conventional classification-based approaches. One of the key accomplishments is the implementation of structured JSON-based output generation, which ensures reliable data processing and seamless frontend integration.

Additionally, the project delivers a comprehensive dietary analytics dashboard that provides users with meaningful insights into their nutritional patterns. The backend architecture has been designed with a strong emphasis on security, ensuring that sensitive API credentials and user data are adequately protected. The modular design paradigm further enhances the adaptability of the system, enabling effortless incorporation of future features and technological advancements. By incorporating contextual reasoning and portion estimation capabilities, DeepDiet establishes itself as a more practical and sophisticated alternative to traditional CNN-based food recognition systems.

Limitations

Despite its innovative contributions, the DeepDiet system is subject to certain limitations inherent to image-based estimation methodologies. The nutritional values generated by the system are based on AI-driven approximations and may exhibit minor variations depending on factors such as image quality, lighting conditions, and the complexity of the meal composition. Additionally, the system requires stable internet connectivity to communicate with external AI services, thereby limiting its functionality in offline environments.

The dependency on third-party APIs introduces potential constraints related to service availability and response latency. Furthermore, precise portion measurement cannot be guaranteed in the absence of physical weighing mechanisms, as the system relies on visual inference rather than direct quantification. While these limitations do not significantly hinder overall functionality, they highlight areas for further refinement and technological enhancement.

Future Enhancements

Future iterations of DeepDiet can incorporate several advancements to enhance its accuracy, scalability, and practical applicability. The integration of a cloud-based database system would enable secure data storage, multi-device synchronization, and long-term dietary tracking, thereby improving user experience and analytical depth. The development of dedicated mobile applications for Android and iOS platforms would further enhance accessibility, enabling real-time image capture and personalized notifications for dietary management.

Incorporating advanced machine learning-based health scoring models could provide deeper insights into user behavior by analyzing factors such as macronutrient balance, caloric surplus or deficit, dietary diversity, and longitudinal consumption trends. Additionally, integration with wearable devices such as fitness trackers and smartwatches would facilitate real-time activity monitoring and more accurate calorie expenditure estimation, enabling holistic health assessment.

Further enhancements may include barcode scanning and ingredient-level analysis to improve accuracy for packaged and homemade foods. The exploration of lightweight, on-device AI models represents another promising direction, as it would reduce dependency on external APIs and enable offline functionality. Collectively, these enhancements would transform DeepDiet into a more comprehensive and intelligent nutrition management system.

References

1. M. Chen, K. Dhingra, W. Wu, L. Yang, R. Sukthankar, and J. Yang, “PFID: Pittsburgh Fast-Food Image Dataset,” IEEE International Conference on Image Processing (ICIP), 2009.

2. A. Meyers et al., “Im2Calories: Towards an Automated Mobile Vision Food Diary,” IEEE International Conference on Computer Vision (ICCV), 2015.

3. Y. Kawano and K. Yanai, “Food Image Recognition with Deep Convolutional Features,” ACM Multimedia Workshops, 2014.

4. H. Kagaya, K. Aizawa, and M. Ogawa, “Food Detection and Recognition Using Convolutional Neural Networks,” ACM Multimedia, 2014.

5. Google DeepMind, “Gemini: A Family of Highly Capable Multimodal Models,” 2023.

6. Google AI, “Gemini API Documentation,” Available: https://ai.google.dev

7. World Health Organization (WHO), “Healthy Diet Guidelines,” Available: https://www.who.int

8. USDA, “FoodData Central Database,” Available: https://fdc.nal.usda.gov

9. S. Ramírez, “FastAPI Documentation,” Available: https://fastapi.tiangolo.com

10. MongoDB Inc., “MongoDB Atlas Documentation,” Available: https://www.mongodb.com/docs

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 05 Issue: 04 | April – 2026 DOI:10.55041/ISJEM06364
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

11.M. Jones, J. Bradley, and N. Sakimura, "JSON Web Token (JWT),"
IETF RFC 7519, 2015.

12.Mozilla Developer Network (MDN), "JavaScript Documentation,"
Available: https://developer.mozilla.org

13.Chart.js,
"Chart.js Documentation," Available: https://www.chartjs.org

14. S. S. Rao, P. Peddi, and E. Madhusudhana Reddy, "Comprehensive Data Collection and Preprocessing Techniques for Analyzing Social Media Usage in Children Below 12 Years using Machine Learning Techniques," SSRN, Nov. 2025.