NutriVision: A System for Automatic Diet Management in Smart Healthcare

Madhumita Veeramreddy[1], Ashok Kumar Pradhan[2], Swetha Ghanta[3], Laavanya Rachakonda[4], Saraju P Mohanty[5]

[1,2,3]Department of Computer Science & Engineering, SRM University AP, Amaravati, 522502, Andhra Pradesh, India.
[4]Department of Computer Science & Engineering, University of North Carolina , Wilmington, USA.
[5]Department of Computer Science & Engineering, University of North Texas, Denton, USA.

Contributing authors: madhumita_v@srmap.edu.in; ashokkumar.p@srmap.edu.in; swetha_ghanta@srmap.edu.in; rachakondal@uncw.edu ; saraju.mohanty@unt.edu ;

Abstract

Effective diet management is crucial for preventing non-communicable diseases, including heart disease, diabetes, and obesity. This paper introduces NutriVision, an enhanced system featuring an interactive chatbot that delivers real-time dietary guidance and personalized recommendations. NutriVision enables users to inquire about nutritional content, receive customized recipe suggestions, and track nutritional goals through natural language interaction. By integrating user data such as health conditions, dietary preferences, and nutritional requirements, it generates tailored responses that enhance user engagement and decision-making. With the help of computer vision and machine learning, NutriVision accurately identifies food items and estimates quantities from smartphone-captured images, providing instant nutritional analysis that includes both macronutrient and micronutrient details with a 94% accuracy rate. It also offers step-by-step cooking instructions, suggests dietary alternatives, and answers queries related to nutrition and health. NutriVision refines its recommendations based on user feedback, creating a dynamic and responsive dietary management experience. Unlike conventional tools that rely on manual data entry or static food databases, NutriVision delivers a seamless, AI-driven interaction by combining food recognition with personalized dialogue. NutriVision exhibits high engagement, with a user engagement rate of 87% and response accuracy of 91%, demonstrating its effectiveness in fostering sustained interaction and delivering contextually relevant and accurate information. Designed with a user-friendly interface, NutriVision supports a variety of dietary frameworks, including vegetarian, vegan, and low-carb plans, ensuring its recommendations align with diverse user needs. By merging intelligent healthcare with interactive digital support, NutriVision empowers users to make informed dietary choices, promoting healthier eating habits. This paper also presents the system's design, performance evaluation, and potential real-world applications, highlighting its contributions to personalized nutrition.

Keywords: Smart Healthcare, Personalized Dietary Guidance, Computer Vision, Real-Time Nutritional Analysis, Generative AI

1 Introduction

Nutrition is crucial for overall well-being and disease prevention. Essential nutrients support bodily functions, while dietary imbalances can lead to chronic conditions like diabetes, heart disease, kidney disease, hypertension, obesity, cancer, and many more [3]. Fig. 1 illustrates the impact of an unbalanced diet, particularly excessive consumption of unhealthy foods, on the development of various chronic diseases. Processed foods, trans fats, refined sugars, excess salt, and additives drive obesity and inflammation, accelerating disease progression [4]. As a result, there is a growing focus on dietary awareness and regulation to mitigate health risks. Balancing dietary intake with nutritional monitoring is vital for long-term health and disease prevention.

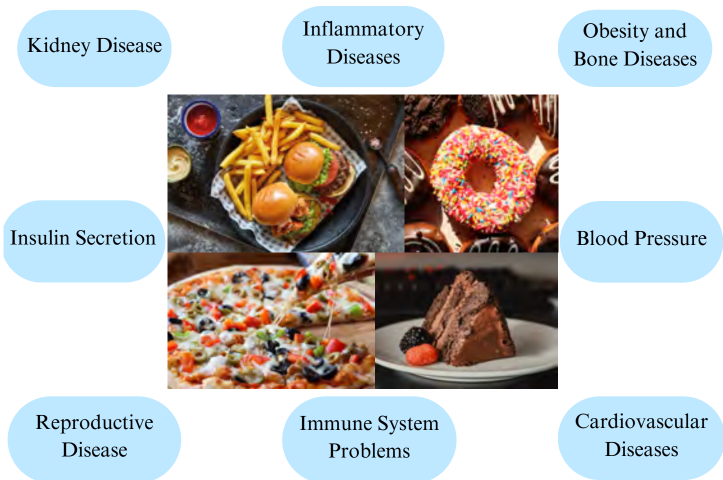

Fig. 1: Impact of Unbalanced Diet

Although nutritional tracking devices and apps are increasingly available, many still rely on manual entry of food intake ([52], [53]), making them time-consuming and discouraging regular use while increasing the risk of inaccurate data. Most systems emphasize calorie counting and diet adherence, but true health requires a holistic view, addressing macronutrient balance, micronutrient intake, and individual lifestyle factors. Individuals managing conditions such as hypertension or diabetes require personalized recommendations that limit sugar, sodium, and saturated fat, underscoring the need for more intuitive and tailored dietary management [54].

In order to overcome the above issues, our work introduces NutriVision, an AI-powered system for enhanced dietary management using computer vision (CV), machine learning (ML), and natural language processing (NLP) techniques. It leverages Faster R-CNN[16] to detect food items from smartphone images and estimate nutritional content, including macronutrients, micronutrients, and calories. Similar to AI-driven IoT-based health monitoring frameworks, such as the AE-GRU-HBA system proposed in [57], NutriVision integrates deep learning as a core analytical component of the cyber-physical infrastructure. A key advancement is the integration of a generative AI (GenAI) chatbot that supports real-time, multi-turn interactions, delivering personalized advice based on health conditions and dietary preferences, as well as offering cooking instructions and ingredient substitutions.

Fig. 2 illustrates NutriVision's role within a Healthcare Cyber-Physical System (H-CPS). Operating in the Processing Hub, NutriVision employs AI, cloud computing, and data processing techniques to analyze dietary information. Through the Connectivity Hub, it interfaces with health monitoring devices, including fitness trackers, glucometers, and blood pressure monitors, to aggregate multi-source physiological data and generate a comprehensive health profile. Based on this integrated data, NutriVision provides real-time feedback and personalized nutritional recommendations, supporting informed dietary decisions and healthier lifestyle choices. As a key component of H-CPS, NutriVision underscores the role of nutrition in preventive healthcare by delivering timely insights into food choices and their health implications. Leveraging communication technologies such as 5G and Wi-Fi, it seamlessly integrates with broader healthcare systems, enabling instant feedback, promoting healthy habits, and helping mitigate diet-related health risks.

The innovative aspects of the proposed solution are detailed in Section 2. A review of related works and background information is provided in Section 3. A comprehensive explanation of the proposed framework, along with the test results, is presented in Section 4. The validation results and comparative analysis are discussed in Section 5. Finally, Section 6 concludes the paper with insights into future work.

2 Contributions

This work presents NutriVision, an innovative AI-driven dietary management system that supports preventive healthcare and personalized nutrition. NutriVision combines food recognition with conversational artificial intelligence to enable users to track dietary intake, receive real-time nutritional guidance, and make informed food choices. The design, implementation, and evaluation of the system demonstrate its potential

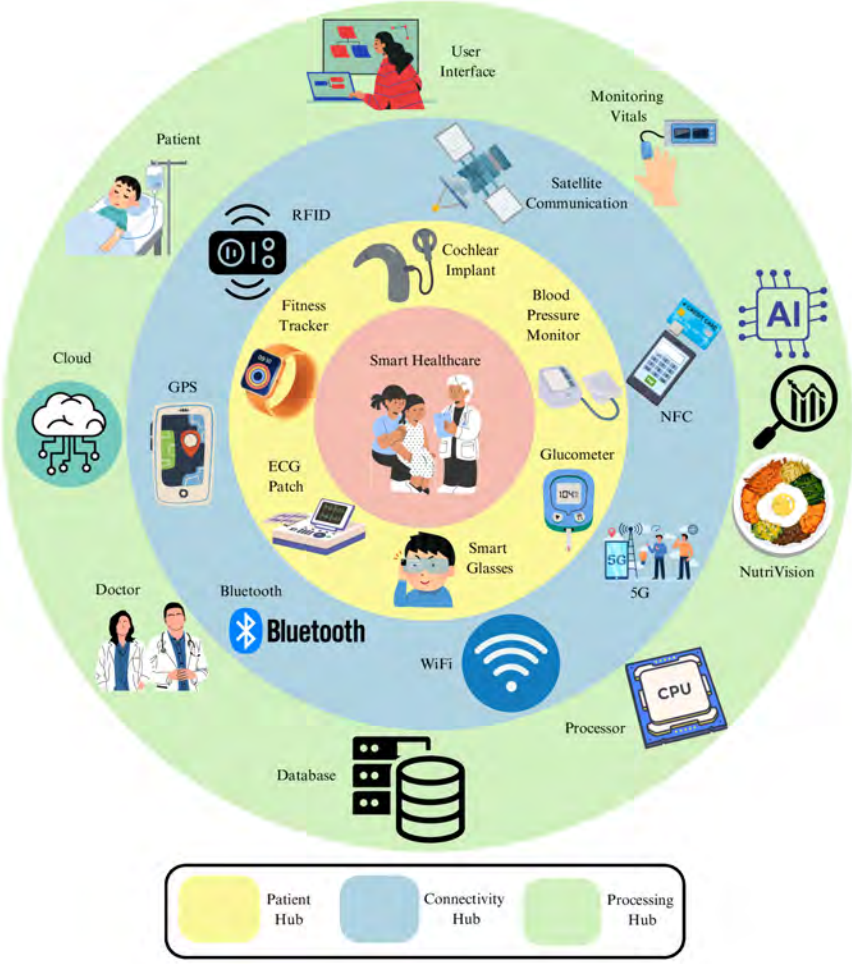

Fig. 2: Smart Healthcare as a Healthcare Cyber-Physical System (H-CPS)

to enhance user engagement and improve health outcomes. The key contributions of this work are summarized as follows:

- Development of a smart conversational chatbot that provides real-time assistance for food, nutrition, and health-related queries.
- Provides diet recommendations based on BMI, allergies, and food preferences.
- Integrates AI to detect food items and estimate their nutrition without manual input.
- Provides instant feedback on daily food intake and meal planning.
- Improves recommendations based on user interactions over time.
- Makes food tracking simpler and more interactive through conversations.
- Helps users make better food choices with AI-driven insights.

3 Related Work

Recent advances in high-performance computing, edge platforms, and AI have significantly improved smart healthcare and dietary analysis, especially through mobile app-based automated dietary intake estimation. Dietary monitoring methods are broadly classified into manual and automated approaches, as shown in Fig. 3. Manual methods, such as food journals or calorie counting via nutritional labels and databases, rely on user input and are prone to errors like incorrect portion estimates and forgotten meals, despite mobile tracking systems like [1] and [2] simplifying the process. Automated methods using CNN-based food recognition, as in [8] and [9], improved detection but often lacked nutritional analysis. Mask-RCNN was applied [10] for food weight and calorie estimation, but did not provide detailed nutrient profiles or personalized feedback.

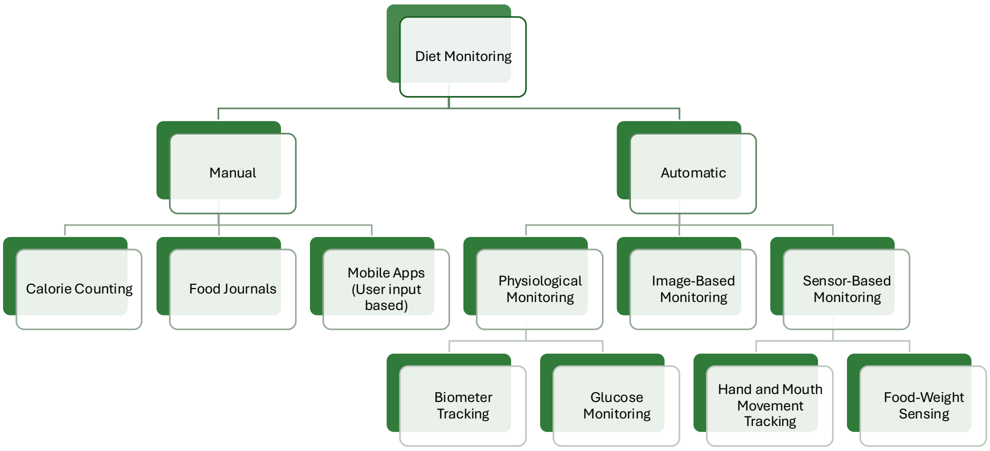

Fig. 3: Diet Monitoring Approaches

Ingredient recognition and mixed food analysis have been explored in several works. For example, [18] used the Bag-of-Features model for diabetic patient diet tracking through HSV color analysis, but lacked nutritional computation. Similarly, [19] proposed a CNN-based ingredient recognition system without nutrient analysis. A hybrid model in [38] combined manual input and image-based analysis, but relied heavily on user data. Multimodal approaches, like the visual and acoustic model in [42], focused on calorie monitoring but lacked a detailed nutrient breakdown. [8] integrated edge-cloud computing for dietary assessment and used image texture features for food recognition. Similarly, [9] applied deep learning for food attribute recognition from text. However, both focused only on calorie estimation without real-time feedback or personalized advice. Although [11] combined CNN-based food recognition with text attribute estimation, it was limited to calorie assessment. Similarly, [40] used Mask-RCNN for food segmentation and portion estimation but lacked detailed nutritional analysis and personalization.

Sensor-based dietary monitoring has been explored through wearable systems. For instance, [23] utilized piezoelectric sensors to detect food intake and estimate calories, but did not provide nutrient profiling or adaptive feedback. Real-time metabolic tracking using continuous glucose monitoring (CGM) and biomarker analysis [20, 21] delivers detailed insights but depends on specialized equipment, limiting widespread use. Edge-based CNN models [24, 25] estimate food volume primarily for calorie counting, with less focus on comprehensive nutritional analysis. CNN-based methods are also widely used for dietary analysis. [26] estimates food volume and calories, while [27] extends this to ingredient identification and cooking instructions using multitask CNN models. [28] classifies food types but focuses mainly on calories. [39] handles ingredient recognition in mixed dishes but lacks detailed nutritional analysis. [43] explored automated nutritional estimation, but limited training data restricted its real-world use. An edge-cloud system in [51] focused on real-time food identification but provided limited nutritional data, mainly estimating calories.

Table 1: Comparative Analysis of Self-Tracking Systems

Sl. No.	**System**	**Input Type**	**Capability**	**Fully Automated**	**Nutritional Estimation**	**Personalized Guidance**	**Interactive Functionality**
1	Pouladzadeh et al. (2015) [34]	Image	Limited Effectiveness	Semi-Automated	No	No	No
2	L. Rachakonda et al. (2019) [35]	Image	Effective Analysis	Semi-Automated	No	No	No
3	Taichi, et al. (2009) [32]	Image	No Nutritional Information	No	No	No	No
4	M.-L.Chian, et al. (2019) [10]	Image	Provides Nutritional Estimation	Semi-Automated	Yes	No	No
5	Jiang, et al. (2018) [31]	Image	Not Very Beneficial	Semi-Automated	No	No	No
6	Harrison, et al. (2010) [29]	Image	Not Feasible	No	No	No	No
7	**NutriVision (Proposed)**	Image	Nutrition and Personalized Analysis	Yes	Yes	Yes, based on health profile and BMI	Yes, answers dietary and health-related queries

Several approaches have been explored for dietary intake estimation, for example, an inertial sensor-based system was developed [45] to track eating gestures and estimate calories, but lacked real-time feedback and personalized advice. Smartphone-based deep learning models like [49] automate food detection but struggle with complex dishes and limit analysis to calories without customization. Cloud-based apps such as [48] estimate macronutrients via CNN-based recognition but may face connectivity issues. Similarly, [46] combines food diaries with AI for detailed nutrient analysis but requires manual input, limiting automation and ease of use. Advanced AI enables detailed analysis of macronutrients, micronutrients, and portion sizes. While systems like [30, 31] offered only basic nutritional info, and others excelled at food recognition [22] but lacked personalized dietary plans, automated tracking with personalized advice and real-time interaction addresses these gaps. Adaptive learning improves recommendations over time, unlike [41, 47], which lacked feedback-based refinement. Although [44, 50] included portion estimation and ingredient recognition, they focused mainly on calorie tracking, whereas our approach emphasizes real-time nutrient profiling and interactive feedback to enhance user engagement and long-term healthy diet adherence.

Table 1 summarizes key findings from various studies alongside the features offered by the proposed NutriVision system. It presents a comparative analysis that highlights how different approaches address nutritional analysis and user engagement, and how NutriVision builds upon the limitations of existing systems.

4 The Innovative Framework: NutriVision

4.1 Framework

The user captures an image of the food with a smartphone, from which features are extracted for detection. The food is matched with entries in a nutrition database to calculate its nutritional value. Discrepancies in portion size are addressed through food quantification. NutriVision enhances interaction with a GenAI-powered dietary assistant, providing real-time, tailored recommendations based on preferences and health conditions. It refines suggestions based on feedback, improving analysis and supporting healthier eating habits. Fig. 4 illustrates the complete workflow of NutriVision.

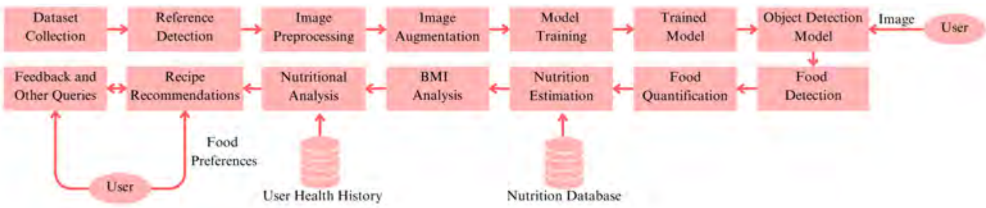

Fig. 4: Development Workflow of NutriVision

4.1.1 Acquisition of Dataset

A customized food dataset is constructed by combining food images from the COCO dataset [15] with additional sandwich images from GitHub's Food Recognition dataset [14]. Out of COCO's 91 classes, our customized dataset includes 10 food-specific classes: Apples, Oranges, Pizza, Broccoli, Cake, Donut, Sandwich, Carrot, Banana, and Hotdog. This customized dataset contains 1000 images, with 800 used for training and 200 for testing, all annotated using Roboflow [33].

4.1.2 Reference Detection

For proof-of-concept evaluation, a reference object is essential for food quantification, as plate size varies with camera distance, affecting accurate measurement. NutriVision uses a one-rupee coin (21.93 mm diameter) as the reference, placed next to the plate during photo capture. The known dimensions of the coin enable accurate estimation of portion size.

4.1.3 Image Preprocessing and Augmentation

The images have been subjected to preprocessing, involving size and normalization adjustments, before undergoing data augmentation. This involved cropping the images using a minimum scaling factor of 0.1 and a maximum scaling factor of 2.0, along with horizontal flipping. The setting for maximum dimension padding has been retained as True.

4.1.4 Model Training

The object detection model utilizes a pre-trained Faster R-CNN [16] trained on the COCO dataset [7], leveraging transfer learning to enhance accuracy and reduce training time. It combines a Region Proposal Network (RPN) with a ResNet-50 backbone [6] to refine bounding boxes for precise object localization, thereby improving detection performance [17]. This approach enables accurate identification of food items even in complex, real-world scenarios, contributing significantly to NutriVision's dietary analysis.

4.1.5 Food Quantification and Nutrition Estimation

After identifying the food and reference coin in the image, the food label is matched with entries in the nutritional database, using predefined portion sizes. The food's actual dimensions are determined using the coin as a reference, with the volume adjusted by a factor of 0.8 to account for empty space in the bounding box, following the approach described in [1, 2]. The estimated volume is then converted to weight using data from the Food-A-Pedia [12] database, and nutritional values are mathematically calculated based on this custom nutrition database for 100 grams of the standard serving. While this approach provides practical volume-to-weight estimation, accuracy may be affected by irregular shapes, occlusions, or camera distortions. The complete reference-based food portion and nutritional estimation procedure, including coin detection, scale calibration, volume computation, and nutrient calculation, is summarized in Algorithm 1.

Algorithm 1 Reference-Based Food Quantity and Nutrition Estimation

Input: Food image I containing a reference coin
Output: Estimated nutritional values N

```
Convert I from RGB to HSV color space
Segment the coin region using predefined color thresholds
Extract contours from the segmented image
Select the contour corresponding to the coin
Compute the minimum bounding box of the coin
Measure the coin pixel diameter d_c
Compute pixel-to-mm scale factor: S = D_c / d_c    ▷ D_c = 21.93 mm is the actual coin diameter
Apply the food detection model on I to obtain food bounding boxes B_i
for each detected food object i do
    Obtain bounding box width w_p and height h_p in pixels
    Compute real-world food dimensions: W_i = w_p × S,  L_i = h_p × S
    Retrieve predefined height H_i for food class i from class-specific lookup table
    Estimate food volume: V_i = k × W_i × L_i × H_i    ▷ k is the bounding-box correction factor
    Convert volume to mass using the cm^3 to gram conversion factor ρ_i: M_i = V_i × ρ_i
    Retrieve nutritional values per 100 g from the Food-A-Pedia database
    Compute nutritional intake: N_i = (M_i / 100) × N_{100,i}
end for
Compute total nutritional values: N = Σ_i N_i
return N
```

4.1.6 Nutritional Analysis

NutriVision provides a personalized nutritional analysis system that integrates collaborative filtering [36] and content-based filtering[13] with an AI-powered chatbot. NutriVision employs the Gemini 1.5 Flash large language model (LLM) as the base generative AI engine for its conversational and recommendation components. The chatbot is integrated into the NutriVision system through a Flask-based backend, with all system components implemented in Python and executed on Google Colab. Personalization is achieved through a hybrid recommendation framework that combines content-based filtering and collaborative filtering. The content-based filtering leverages NLP [37] to extract pertinent information from user health reports, while collaborative filtering enhances the relevance of recommendations through data-driven matrix factorization. NutriVision dynamically adapts to user preferences and health conditions, offering personalized suggestions, such as adjusting for gluten intolerance, diabetes, and hypertension, thereby ensuring real-time, contextually relevant recommendations.

NutriVision utilizes a prompt to guide the Gemini model to interpret user queries and facilitate multi-turn conversations for iterative preference refinement. It incorporates health data to generate tailored dietary recommendations, such as low-sugar options for diabetics and low-sodium meals for individuals with hypertension. The system also provides real-time feedback during meal preparation, suggesting ingredient substitutes based on availability and dietary restrictions. To enhance user engagement, NutriVision delivers comprehensive nutritional breakdowns, including macronutrient and micronutrient analysis, as well as caloric content. It generates recommendations that align with user health goals such as weight management, muscle hypertrophy, or balanced nutrition. The system's fusion of sophisticated recommendation algorithms and an intuitive user interface enables seamless user interaction, adaptive real-time support, and personalized dietary management.

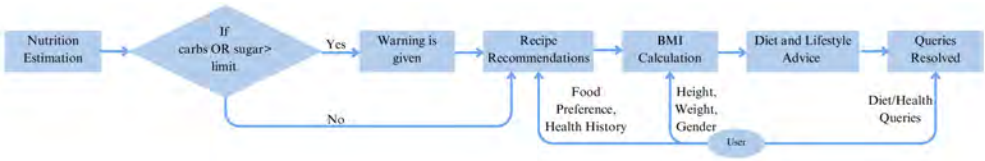

Fig. 5: Workflow of Nutritional Analysis

The nutritional analysis module workflow, as shown in Fig. 5, illustrates how personalized recipe recommendations are generated based on individual health profiles. These suggestions are tailored to users' medical histories, with automatic alerts triggered when nutrient intake exceeds safe limits. Continuous user feedback refines these recommendations, and BMI calculations enable additional meal options based on nutritional status. Users can also query NutriVision for real-time, context-aware dietary guidance aligned with their health goals. This highlights NutriVision as a cutting-edge, personalized nutritional analysis system delivering timely and tailored recommendations informed by individual health data.

5 Validation of NutriVision

5.1 Validation

To validate our model, 200 images are used to evaluate food classification accuracy. The model showed strong accuracy even after localization, but challenges remain when food items are closely positioned, lowering confidence scores. Expanding the dataset with diverse meal presentations, plate designs, and background settings can improve accuracy and robustness in real-world scenarios. Fig. 6 illustrates the model's capability to estimate nutritional values from detected food items, demonstrating its potential for accurate dietary analysis. NutriVision includes two primary nutritional visualizations: the macronutrient distribution chart, as shown in Fig. 7, and the macronutrient composition chart in grams, as shown in Fig. 8. The distribution graph shows the percentage breakdown of carbohydrates, proteins, sugars, and fats, helping users quickly assess meal balance and spot dietary imbalances. The composition graph provides exact gram values for each macronutrient, supporting precise tracking and meal planning. Both graphs offer an easy-to-understand interface for users with limited nutritional knowledge. Combining clarity with accuracy, NutriVision empowers informed dietary choices, with real-time updates ensuring users access the latest nutritional data for effective planning and goal tracking.

Fig. 9 shows the NutriVision interface, which analyzes the user's health history to set dietary limits like carbohydrate and sugar intake per meal. Foods exceeding these limits trigger real-time warnings in red with exact nutritional values, reducing manual effort and improving diet tracking. The system supports preventive healthcare by helping manage or prevent conditions like diabetes and obesity. It also offers alternative meal suggestions to meet dietary goals, empowering users to make healthier choices. Overall, the figure highlights how NutriVision uses AI-driven chatbots to provide personalized, accessible health guidance.

Food item	Quantity (grams)	Calories (kcal)	Carbs (grams)	Protein (grams)	Saturated Fat (grams)	Sugar (grams)	Potassium (mg)	Iron (mg)	Fiber (grams)	Calcium (mg)
Pizza	1221.65	3368.15	370.34	150.57	61.08	43.49	1965.85	22.72	23.21	2565.47
Sand wich	453.54	1378.77	149.13	44.21	19.72	18.63	757.42	6.80	6.80	1247.23

Fig. 6: Nutritional Value Estimation by NutriVision

Fig. 10 highlights the NutriVision Chatbot's ability to assist with various cooking-related questions, offering practical guidance on ingredient preparation, cooking techniques, food storage, and dietary substitutions. It provides step-by-step instructions for cutting vegetables, cooking at optimal temperatures, preserving freshness, and enhancing flavors. The chatbot also shares food safety tips, suggests ingredient alternatives, and advises on recipe modifications for dietary needs. Additionally, users can ask about nutritional values, nutrient-retaining cooking methods, and ways to balance meals for better health.

NutriVision offers clear, structured responses that simplify complex culinary concepts into easy steps. Users can provide feedback to help the system refine queries and deliver personalized recommendations. It provides instant, expert-backed guidance for both beginners and experienced cooks to confidently handle kitchen challenges. Leveraging AI, NutriVision learns from interactions to improve accuracy, making meal preparation more efficient and adaptable to diverse dietary needs. Fig. 11 illustrates

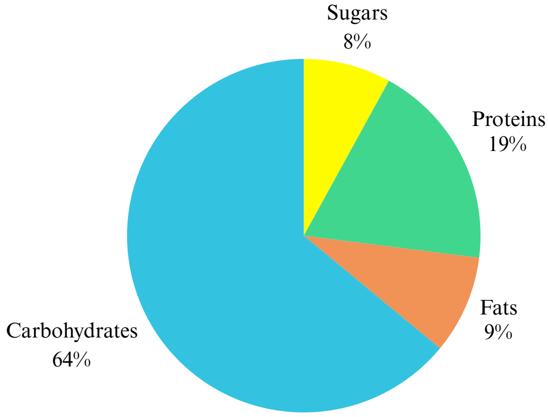

Fig. 7: Macronutrient Distribution

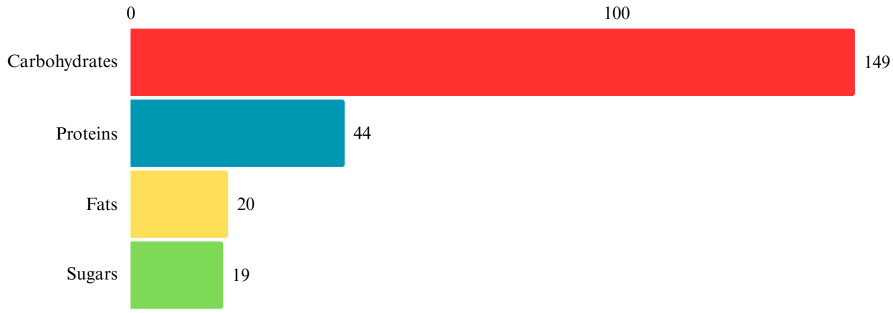

Fig. 8: Macronutrient Composition in grams

the BMI calculation feature, which uses a checklist allowing users to select their height (120–214 cm) and weight (30–154 kg) from predefined ranges. This approach simplifies data input, accommodates diverse body types, and protects user privacy by avoiding the need for exact values. The system calculates BMI using the standard formula, classifies users into weight categories, and offers personalized dietary, activity, and lifestyle advice accordingly.

NutriVision incorporates a robust system for generating recipe suggestions, designed to provide users with personalized and tailored dietary options. This system employs a multi-select checklist interface, enabling users to specify a variety of criteria to refine their recipe search. As depicted in Fig. 12, NutriVision offers an extensive selection of cuisines, including those of all Indian states and numerous countries (a total of 65), allowing users to explore a diverse range of culinary options and align with specific regional preferences.

NutriVision allows users to specify their dietary preferences by selecting from options such as vegetarian, non-vegetarian, or vegan, ensuring that recipe suggestions adhere to their restrictions or choices. It further allows users to indicate the type of meal they are seeking, such as breakfast, lunch, snacks, or dinner, facilitating the

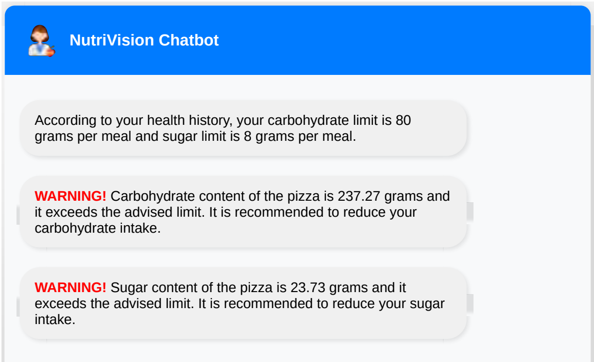

Fig. 9: Real-time sugar and carbohydrate alerts in NutriVision

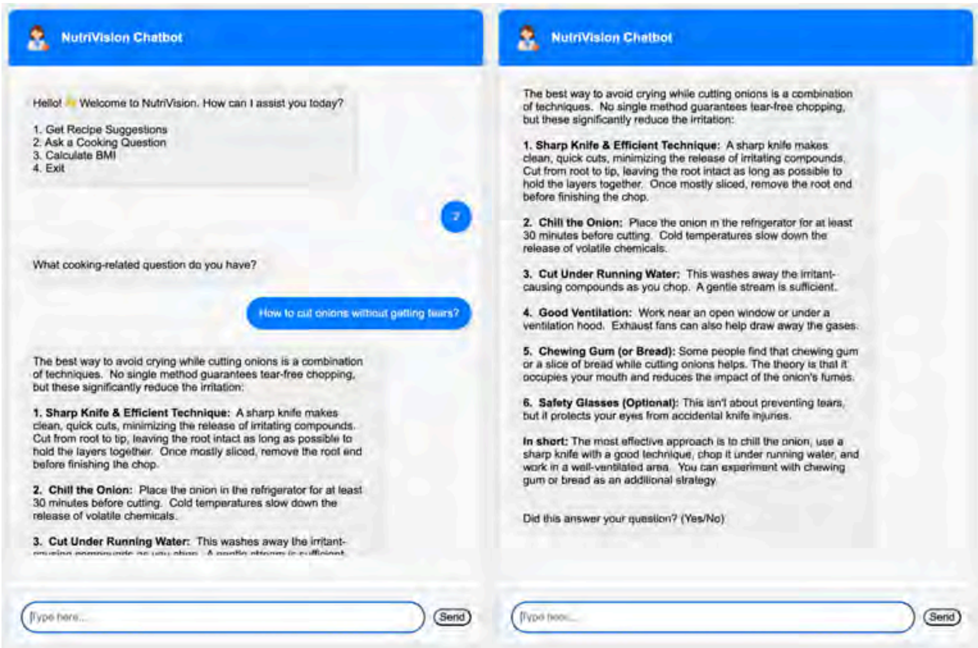

Fig. 10: NutriVision offering real-time cooking guidance and recipe assistance.

discovery of recipes suitable for specific times of day or occasions. This multi-level filtering enhances the relevance of recommendations by aligning them with nutritional needs and contextual factors.

A critical aspect of the system is its detailed allergy specification feature, which includes 35 types of allergies, enabling users to exclude recipes containing ingredients that could trigger allergic reactions and promote healthy eating. NutriVision addresses both common and uncommon allergies, such as tree nuts, wheat, dairy, and gluten. Furthermore, users can input their health conditions from a list of 68 conditions,

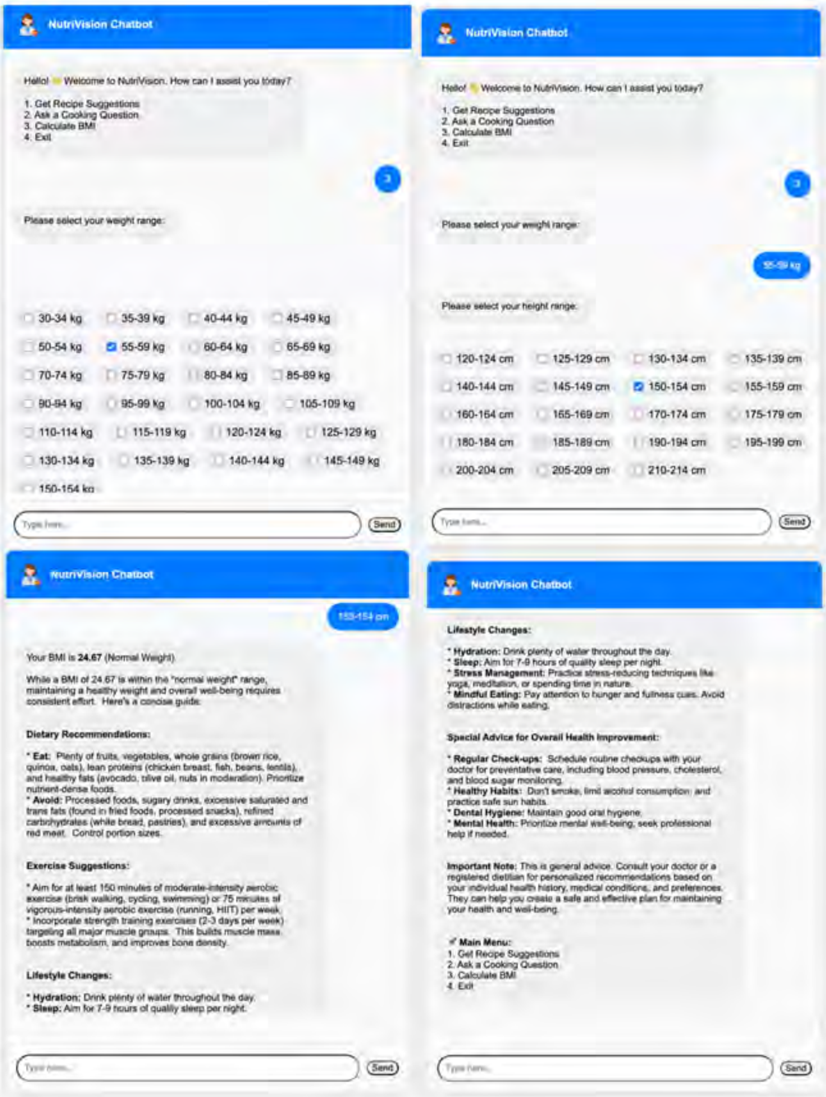

Fig. 11: BMI calculation interface with height and weight range selection for health insights.

spanning general, cardiovascular, gastrointestinal, allergies, autoimmune, neurological, respiratory, bone and joint, kidney and urinary, liver and pancreas, cancer-related, and miscellaneous conditions. This functionality allows NutriVision to provide recipe suggestions suited to users' health needs.

The system also allows users to select ingredients they have available from a list of 176 ingredients, helping users find recipes they can prepare with existing resources, thus reducing food waste and enhancing convenience. In addition to these criteria,

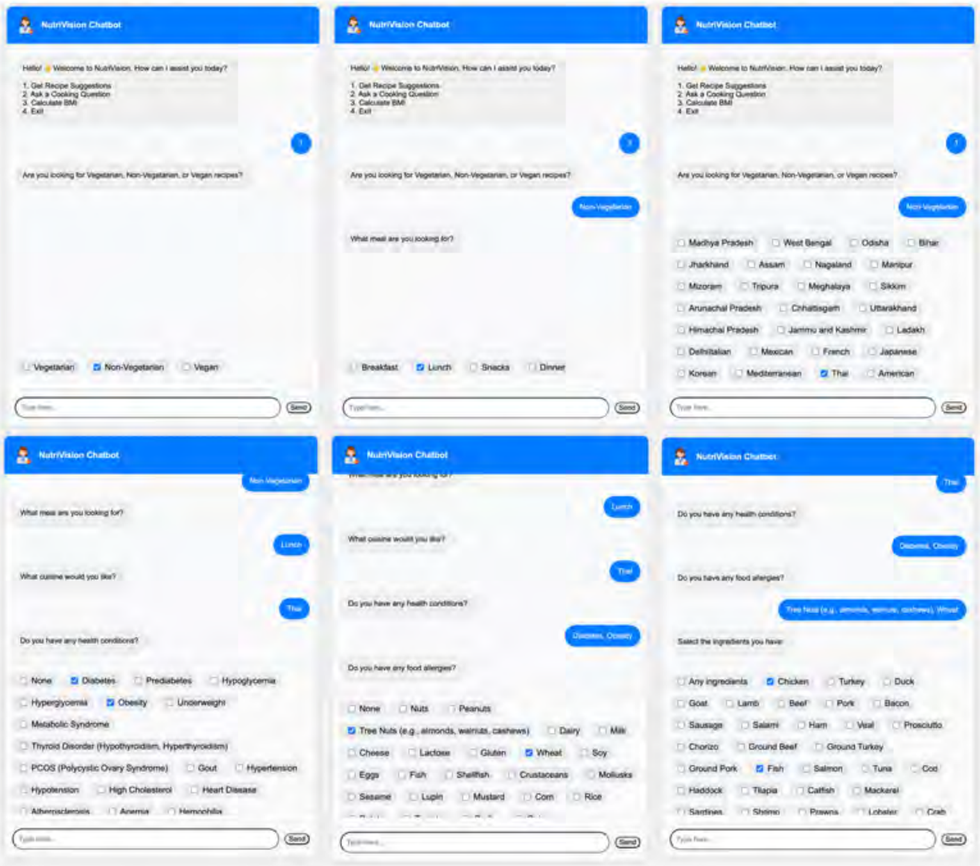

Fig. 12: Recipe customization options for dietary preferences and personalized meal suggestions.

NutriVision also allows users to specify the number of servings they require for a recipe. Recognizing the potential for user needs and preferences to extend beyond the provided lists, NutriVision's design includes the flexibility for users to type in options, accommodating a wider range of specific requests. These features create a highly personalized recipe suggestion experience. The multi-select checklist allows users to specify multiple criteria at once, ensuring tailored recipe recommendations. With options for cuisines, dietary needs, meal types, allergies, health conditions, and ingredients, users gain greater control over their diet and can find recipes that support their health goals.

Fig. 13 illustrates how NutriVision excels not only in generating recipe suggestions but also in how it presents these suggestions and facilitates refinement based on user input. Upon providing recipe suggestions, NutriVision includes estimated nutritional information for each recipe, offering estimates for key nutritional components such as calories, protein, carbohydrates, and fat content per serving. Each recipe is presented

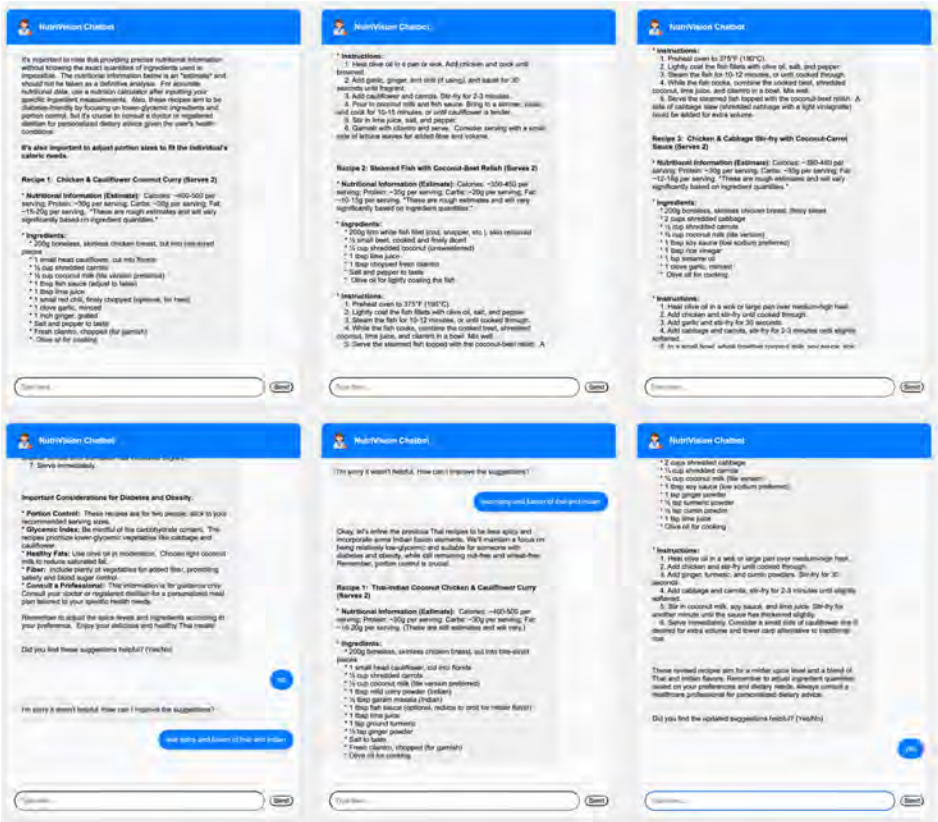

Fig. 13: Recipe recommendations with step-by-step instructions and user feedback integration.

with a detailed list of ingredients and step-by-step cooking instructions, structured in a clear, numbered format to facilitate ease of use during the cooking process.

Furthermore, NutriVision enhances the user experience by providing additional tips related to the recipes. These tips may include serving suggestions, modifications for specific dietary needs, advice on portion control, information on the glycemic index, guidance on using healthy fats, and the importance of including vegetables for fiber. This added layer of information helps users make more informed choices and align the meals with their health goals.

A notable feature of NutriVision is its use of a feedback mechanism to refine recipe suggestions based on user input. After presenting initial recipe suggestions, NutriVision solicits user feedback on their helpfulness. In cases of negative feedback, NutriVision prompts users for specific ways to improve the suggestions. This iterative process enables NutriVision to adapt its recommendations, aligning them more closely with the user's preferences and dietary needs. For instance, NutriVision can refine recipes to be less spicy or incorporate Indian fusion elements, showcasing its capacity for personalized adaptation. To further validate its effectiveness, a performance evaluation of NutriVision is conducted through a simulation-based evaluation using 40 synthetic user profiles representing diverse demographic and health conditions.

We constructed each synthetic user profile to represent a set of dietary preferences, health conditions, and nutritional goals. These profiles are converted into structured sequences of queries such as meal choices, dietary constraints, and ingredient substitution requests, which we submitted to the chatbot in a multi-turn conversational setting. The chatbot responses are evaluated automatically using a second Gemini model acting as an independent evaluator [59], which scored user engagement, response accuracy, and system usability (SUS) across the 40 synthetic user profiles based on dialogue consistency, contextual relevance, and nutritional validity. The results are presented in Table 2.

Table 2: NutiVision Performance Evaluation

Metric	Value	Description
User Engagement Rate	87%	Percentage of simulated user profiles that resulted in multi-turn conversational interactions.
Response Accuracy	91%	Correctness of nutrition-related responses evaluated by an independent Gemini model using reference nutrition data.
System Usability Scale (SUS)	83.5	Simulated usability score computed from automated dialogue evaluation on a 0-100 scale.
Number of User Profiles	40	Total number of synthetic user profiles used for evaluation.

The user engagement rate of 87% indicates that most synthetic user profiles resulted in multi-turn interactions with NutriVision during the automated evaluation. A high engagement rate is important for health-focused conversational systems, as sustained interaction reflects the system's ability to provide coherent, relevant, and context-aware responses over multiple dialogue turns. The user engagement rate [58] is computed as:

$$\text{User Engagement Rate} = \frac{\text{Total Interactions}}{\text{Total Active Profiles}} \times 100 \tag{1}$$

where:

- **Total Interactions** represents the total number of dialogue exchanges between the chatbot and the synthetic user profiles during the evaluation.
- **Total Active Profiles** denotes the number of distinct synthetic user profiles used in the evaluation.

NutriVision's response accuracy of 91% indicates that the system consistently delivered correct and contextually relevant answers to nutrition and health-related queries.

This is especially important in dietary recommendation systems, where inaccurate information could negatively affect user health decisions. Response Accuracy [60, 61] is calculated as:

$$\text{Response Accuracy} = \frac{\text{Correct Responses}}{\text{Total Responses}} \times 100 \tag{2}$$

Where:

- Correct Responses refers to the number of times NutriVision provided accurate and relevant answers to user queries.
- Total Responses refers to the total number of responses generated by NutriVision during the evaluation study.

The System Usability Scale (SUS) [56] score of 83.5 indicates a high level of simulated usability and interaction quality. In this study, SUS is estimated automatically using the independent Gemini evaluator based on the coherence, clarity, and effectiveness of the chatbot's responses across the 40 synthetic user profiles. Scores above 80 are generally interpreted as excellent usability, indicating that the system provides intuitive and efficient interactions within the evaluated scenarios.

The SUS score is computed as:

$$\text{SUS Score} = \frac{(\text{Sum of all SUS response scores}) - 25}{0.5} \tag{3}$$

where:

- Sum of all SUS response scores represents the aggregated ratings assigned by the Gemini evaluator for each SUS questionnaire item based on the generated dialogues.
- The constant 25 is subtracted as part of the standard SUS normalization.
- Division by 0.5 scales the result to the standard 0-100 SUS range.

Furthermore, qualitative analysis of the generated dialogue responses showed that NutriVision is able to refine its recommendations based on iterative user inputs, demonstrating adaptability to dietary preferences such as low-carbohydrate or vegan diets. Together, the quantitative metrics and automated qualitative assessments validate the system's ability to provide accurate, coherent, and context-aware dietary guidance within the evaluated scenarios. These results indicate the practical potential of NutriVision for real-world deployment in digital health and wellness platforms.

In addition, a simulated use-case demonstration illustrated how NutriVision can generate a balanced meal plan tailored to specific health objectives, such as weight loss or muscle gain. During this multi-turn interaction, the system dynamically adjusted its recommendations in response to profile-driven constraints, including reduced spice levels or ingredient substitutions, with an end-to-end response latency of approximately 3.5 seconds per request, highlighting its capability for real-time adaptive guidance.

Overall, these findings demonstrate NutriVision's effectiveness in delivering intuitive, engaging, and personalized dietary support across a range of synthetic user scenarios. The combination of quantitative performance metrics and automated qualitative evaluation highlights the system's potential to support adaptive, data-driven nutrition management in practical applications.

5.2 Comparison with existing research

Table 3 presents a comparison between NutriVision and several existing research works in the domain of meal identification, nutritional analysis, and personalized dietary guidance. The table evaluates various features, such as meal identification, calorie estimation, portion size estimation, and the capability to provide personalized recommendations. Existing systems are assessed based on their functionality in terms of dietary goal adaptation, user interaction through natural language, and real-time feedback-based learning. NutriVision is highlighted for its advanced features, including personalized recommendations, real-time alerts, multi-turn conversational capability, and real-time feedback learning. These aspects distinguish NutriVision from traditional systems, which often lack real-time adaptation and personalized guidance, making it a more dynamic and user-centric solution.

Table 3: Comparison of NutriVision with Existing Research Works

Feature	[1]	[2]	[5]	[10]	[30]	**NutriVision**
Meal Identification	Yes	Yes	Yes	Yes	Yes	**Yes**
Calorie Estimation	Yes	Yes	Yes	Yes	No	**Yes**
Portion Size Estimation	No	Yes	No	Yes	No	**Yes**
Personalized Recommendations	No	No	Limited	No	No	**Yes**
Dietary Goal Adaptation	Limited	Limited	No	No	Limited	**Yes**
User Interaction (Natural Language)	No	No	No	No	No	**Yes**
Feedback-based Learning	No	No	No	No	No	**Yes**
Real-Time Alerts (e.g., sugar/carb)	No	No	No	No	No	**Yes**
Recipe Guidance	No	No	No	No	No	**Yes**
Multi-turn Conversations	No	No	No	No	No	**Yes**

Table 4: Comparison of Object Detection Algorithms

Metric	Faster R-CNN	YOLO	CNN	Mask R-CNN
Accuracy	**94%**	85%	82%	88%
Recall	**85%**	75%	72%	80%
Precision	**84%**	77%	75%	81%
F1-Score	**84%**	75%	73%	80%
Loc. Precision	**92%**	85%	80%	88%

In contrast to existing systems that primarily focus on meal identification and calorie estimation, NutriVision introduces advanced features for dynamic and personalized dietary analysis within a controlled experimental setting. Unlike many systems that lack advanced object detection, NutriVision employs Faster R-CNN to accurately detect multiple food items on a single plate and distinguish visually similar foods among the selected food categories used in this study. This model enhances both speed and accuracy for the evaluated dataset, which is critical for reliable dietary analysis in the proposed framework. The performance metrics presented in Table 4 demonstrate how Faster R-CNN's improved localization capability translates into higher precision and accuracy on the customized dataset considered. With Faster R-CNN, NutriVision is able to distinguish subtle visual differences between similar food items within this dataset, ensuring that nutritional estimates are based on precise and consistent detections.

Additionally, NutriVision with reference-based portion estimation, automatically estimates portion sizes, mathematically calculates macronutrient and micronutrient values using the nutrition database, and adjusts recommendations based on factors like BMI and medical history using GenAI. Its integration of fine-grained food quantification with intelligent guidance provides a highly personalized dietary experience that evolves with the user's health needs.

6 Conclusion and Future Work

NutriVision represents a significant advancement in intelligent dietary monitoring systems by integrating CV, ML, and GenAI technologies. The system enhances user interaction through a conversational assistant that understands natural language queries related to food, nutrition, and health. It provides real-time responses, personalized dietary guidance, and dynamic meal suggestions based on individual health profiles, preferences, and nutritional needs. The assistant supports multi-turn dialogues and adapts its responses based on user feedback, ensuring evolving personalization. Beyond nutritional analysis of macronutrients and micronutrients, NutriVision offers features such as recipe suggestions, cooking tips, and alerts for excessive intake of sugar or carbohydrates, empowering users to make healthier food choices.

However, the current study also has limitations. The food dataset is restricted to 10 classes, and the nutritional estimation relies on a reference object for portion size, which may limit generalizability in uncontrolled real-world settings. Future work will focus on enhancing food quantity estimation without the need for reference objects by integrating reference-free depth estimation and camera calibration techniques, thereby improving accuracy in unconstrained real-world settings. The food database can also be expanded to include a broader range of dishes from regional and global cuisines, promoting cultural inclusivity. Additionally, the system can be extended by incorporating advanced large-scale LLMs for complex food detection and nutritional analysis, as well as refining portion estimation methods to further reduce dependence on reference objects. These improvements will enhance the adaptability, accuracy, and real-world applicability of NutriVision in smart healthcare scenarios.

Declarations

- **Funding**
  The authors did not receive any funding for the submitted work.

- **Conflict of Interest**
  On behalf of all authors, the corresponding author states that there is no conflict of interest.
- **Ethics approval and consent to participate**
  Not Applicable
- **Author contribution**
  Conceptualization: Madhumita Veeramreddy, Ashok Kumar Pradhan, Swetha Ghanta, Laavanya Rachakonda, **Methodology**: Madhumita Veeramreddy, Ashok Kumar Pradhan, Swetha Ghanta, Laavanya Rachakonda, **Formal analysis and investigation**: Madhumita Veeramreddy, Ashok Kumar Pradhan, Swetha Ghanta, Laavanya Rachakonda, **Writing - original draft preparation**: Madhumita Veeramreddy, Ashok Kumar Pradhan, Swetha Ghanta, Laavanya Rachakonda, Saraju P Mohanty **Writing - review and editing**: Madhumita Veeramreddy, Ashok Kumar Pradhan, Swetha Ghanta, Laavanya Rachakonda, Saraju P Mohanty, **Supervision**: Ashok Kumar Pradhan, Laavanya Rachakonda, Saraju P Mohanty.
- **Research Involving Human and/or Animals**
  Not Applicable
- **Informed Consent**
  Not Applicable

References

[1] L. Rachakonda, S. P. Mohanty, and E. Kougianos, "iLog: An Intelligent Device for Automatic Food Intake Monitoring and Stress Detection in the IoMT", IEEE Transactions on Consumer Electronics (TCE), Vol. 66, No. 2, May 2020, pp. 115–124.

[2] A. Mitra, S. Goel, S. P. Mohanty, E. Kougianos, and L. Rachakonda, "iLog 2.0: A Novel Method for Food Nutritional Value Automatic Quantification in Smart Healthcare", in Proceedings of the IEEE International Symposium on Smart Electronic Systems (iSES), 2022, pp. 683–688

[3] H. Cena and P. C. Calder, "Defining a healthy diet: evidence for the role of contemporary dietary patterns in health and disease," Nutrients, vol. 12, no. 2, p. 334, 2020.

[4] D. R. Wahl, K. Villinger, L. M. Konig, K. Ziesemer, H. T. Schupp, and B. Renner, "Healthy food choices are happy food choices: Evidence from a real life sample using smartphone based assessments," Scientific Reports, vol. 7, no. 1, pp. 1–8, 2017.

[5] P. Pouladzadeh, P. Kuhad, S. V. B. Peddi, A. Yassine, and S. Shir- mohammadi, "Mobile cloud based food calorie measurement," in Proc. of IEEE International Conference on Multimedia and Expo Workshops (ICMEW), 2014, pp. 1–6.

[6] Tahir, Hassam & Khan, Muhammad Shahbaz & Tariq, Muhammad Owais. (2021). Performance Analysis and Comparison of Faster R-CNN, Mask R-CNN and ResNet50 for the Detection and Counting of Vehicles. 587-59.

[7] Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., & Dollár, P. (2015). Microsoft COCO: Common objects in context. arXiv. https://arxiv.org/abs/1405.0312 last accessed on September 20,2023

[8] C. Liu, Y. Cao, Y. Luo, G. Chen, V. Vokkarane, M. Yunsheng, S. Chen, and P. Hou, "A New Deep Learning-Based Food Recognition System for Dietary Assessment on An Edge Computing Service Infrastructure," IEEE Transactions on Services Computing, vol. 11, no. 2, pp. 249–261, 2018.

[9] R. Yunus, O. Arif, H. Afzal, M. F. Amjad, H. Abbas, H. N. Bokhari, S. T. Haider, N. Zafar, and R. Nawaz, "A Framework to Estimate the Nutritional Value of Food in Real Time Using Deep Learning Techniques," IEEE Access, vol. 7, pp. 2643–2652, 2019.

[10] M.-L.Chiang,C.-A.Wu,J.-K.Feng,C.-Y.Fang,andS.-W.Chen,"Food calorie and nutrition analysis system based on mask r-cnn," in Proc. of IEEE 5th International Conference on Computer and Communications (ICCC), 2019, pp. 1721–1728.

[11] Z. Shen, A. Shehzad, S. Chen, H. Sun, and J. Liu, "Machine learning based approach on food recognition and nutrition estimation," Procedia Computer Science, vol. 174, pp. 448–453, 2020.

[12] "Data.Gov:Food-a-pedia," https://catalog.data.gov/dataset/food-a-pedia.

[13] Meteren, R.V. (2000). Using Content-Based Filtering for Recommendation.

[14] "Food Boundary Box Detection Dataset," https://github.com/stressGC/Food-Boundary-Box-Detection-Dataset.

[15] Tsung-Yi Lin, Maire, M., Belongie, S. J., Bourdev, L. D., Girshick, R. B., Hays, J., Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. CoRR, abs/1405.0312. Retrieved from http://arxiv.org/abs/1405.0312

[16] Girschick, Ross. (2015). Fast r-cnn. 10.1109/ICCV.2015.169

[17] Lokanath, M & Kumar, K & Keerthi, E. (2017). Accurate object classification and detection by faster-RCNN. IOP Conference Series: Materials Science and Engineering. 263. 052028. 10.1088/1757-899X/263/5/052028.

[18] M. M. Anthimopoulos, L. Gianola, L. Scarnato, P. Diem, and S. G. Mougiakakou, "A Food Recognition System for Diabetic Patients Based on an Optimized Bag-of-Features Model," IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 4, pp. 1261–1271, 2014.

[19] J. Chen, B. Zhu, C.-W. Ngo, T.-S. Chua, and Y.-G. Jiang, "A Study of Multi-Task and Region-Wise Deep Learning for Food Ingredient Recognition," IEEE Transactions on Image Processing, vol. 30, pp. 1514–1526, 2021.

[20] H. Jiang, J. Starkman, M. Liu, and M.-C. Huang, "Food Nutrition Visualization on Google Glass: Design Tradeoff and Field Evaluation," IEEE Consumer Electronics Magazine, vol. 7, no. 3, pp. 21–31, 2018.

[21] K. J. Pfisterer, R. Amelard, B. Syrnyk, and A. Wong, "Towards Computer Vision Powered Color-Nutrient Assessment of Pure'ed Food," in Proc. of IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2019, pp. 490–492.

[22] C.Liu,Y.Liang,Y.Xue,X.Qian,andJ.Fu,"FoodandIngredientJoint Learning for Fine-Grained Recognition," IEEE Transactions on Circuits and Systems for Video Technology, vol. 31, no. 6, pp. 2480–2493, 2021.

[23] G. Hussain, B. Ali Saleh Al-rimy, S. Hussain, A. M. Albarrak, S. N. Qasem, and Z. Ali, "Smart piezoelectric-based wearable system for calorie intake estimation using machine learning," Applied Sciences, vol. 12, no. 12, p. 6135, 2022.

[24] R. D. Kumar, E. G. Julie, Y. H. Robinson, S. Vimal, and S. Seo, "Recognition of food type and calorie estimation using neural network," The Journal of Supercomputing, vol. 77, no. 8, pp. 8172–8193, 2021.

[25] T. Miyazaki, G. C. de Silva, and K. Aizawa, "Image-based calorie con- tent estimation for dietary assessment," in Proc. of IEEE International Symposium on Multimedia, 2011, pp. 363–368.

[26] V. B. Kasyap and N. Jayapandian, "Food calorie estimation using convolutional neural network," in Proc. of 3rd International Conference on Signal Processing and Communication (ICPSC), 2021, pp. 666–670.

[27] T. Ege and K. Yanai, “Image-based food calorie estimation using knowledge on food categories, ingredients and cooking directions,” in Proceedings of the on Thematic Workshops of ACM Multimedia, 2017, p. 367–375.

[28] T. Ege and K. Yanai, “Estimating food calories for multiple-dish food photos,” in Proc. of 4th IAPR Asian Conference on Pattern Recognition (ACPR), 2017, pp. 646–651.

[29] A. Harrison, S. Sullivan, K. Tchanturia, and J. Treasure, “Emotional Functioning in Eating Disorders: Attentional Bias, Emotion Recognition and Emotion Regulation,” Psych. Med., vol. 40, no. 11, pp. 1887–1897, 2010.

[30] O. Beijbom, N. Joshi, D. Morris, S. Saponas, and S. Khullar, “Menu- Match: Restaurant-Specific Food Logging from Images,” in Proc. IEEE Winter Conf. on App. of Comp. Visn., 2015, pp. 844–851.

[31] H. Jiang, J. Starkman, M. Liu, and M. Huang, “Food Nutrition Visual- ization on Google Glass: Design Tradeoff and Field Evaluation,” IEEE Consum. Electron. Mag., vol. 7, no. 3, pp. 21–31, May 2018.

[32] J. Taichi and K. Yanai, “A food image recognition system with Multiple Kernel Learning,” in Proc. 16th IEEE ICIP, 2009, pp. 285–288.

[33] “Roboflow: Give your software the power to see objects in images and video.” https://roboflow.com accessed on September 22,2023

[34] P. Pouladzadeh, S. Shirmohammadi, A. Bakirov, A. Bulut, and A. Yas- sine, “Cloud-based SVM for food categorization,” Multimedia Tools and App., vol. 74, no. 14, pp. 5243–5260, Jul 2015.

[35] L. Rachakonda, A. Kothari, S. P. Mohanty, E. Kougianos, and M. Gana- pathiraju, “Stress-Log: An IoT-based Smart System to Monitor Stress- Eating,” in Proc. IEEE ICCE, 2019, pp. 1–6.

[36] Schafer, Ben & J, Ben & Frankowski, Dan & Dan, & Herlocker, & Jon, & Shilad, & Sen, Shilad. (2007). Collaborative Filtering Recommender Systems.

[37] Khurana, D., Koli, A., Khatter, K. et al. Natural language processing: state of the art, current trends and challenges. Multimed Tools Appl 82, 3713–3744 (2023).

[38] Bi, Y., Lv, M., Song, C., Xu, W., Guan, N., & Yi, W. (2019). AutoDietary: A wearable acoustic sensor system for food intake recognition in daily life. IEEE Transactions on Mobile Computing, 15(9), 2230–2243.

[39] Boyd, L., Nnamoko, N., & Lopes, R. (2024). Fine-Grained Food Image Recognition: A Study on Optimising Convolutional Neural Networks for Improved Performance. Journal of Imaging, 10(6), 126.

[40] Jeyaraj, A., Ajay, M., George, R., & Parthiban, P. (2021). Mask-RCNN based portion size estimation and caloric intake tracking. IEEE Access, 9, 38244–38253.

[41] Kong, Fanyu & Tan, Jindong. (2012). DietCam: Automatic Dietary Assessment with Mobile Camera Phones. Pervasive and Mobile Computing, 8, 147-163. 10.1016/j.pmcj.2011.07.003.

[42] Lee, Changmin, & Jin, Sung. (2019). Multimodal dietary intake estimation using visual and acoustic data. Proceedings of the IEEE Conference on Multimedia and Expo (ICME), 788–793.

[43] Liu, Yang, Cheng, Wei, & Fu, Chen. (2020). Deep learning-based nutrition estimation from food images. Journal of Food Engineering, 45(3), 109-121.

[44] Perez, Alexander, Thompson, John, & Walker, Mark. (2018). Graph-based segmentation and caloric estimation for mixed food items. Computer Vision and Image Understanding, 169, 25–39.

[45] Thomaz, Edison, Pering, Tony, & Oliver, Nuria. (2020). A wearable solution for real-time dietary intake recognition. ACM Transactions on Sensor Networks, 11(4), 78–87.

[46] Vasiloglou, Marios, Christoph, Samuel, & Staub, Andreas. (2019). The role of artificial intelligence in food diary and dietary monitoring. Journal of Nutrition and Health, 14(3), 189–199.

[47] Wu, Ping, Qiao, Zhen, & Zhang, Lei. (2020). Food ingredient recognition using deep learning models. Computers in Biology and Medicine, 128, 104096.

[48] Yang, Yu, Zhang, Ling, & Cheng, Xiaoping. (2021). Real-time food recognition and macronutrient estimation using cloud-based CNN models. IEEE Transactions on Neural Networks, 25(6), 2195-2207.

[49] Zhang, Yong, Wu, Kai, & Zhou, Fan. (2020). Smartphone-based deep learning models for automated food detection. Journal of Health Informatics, 12(2), 209-215.

[50] Zhou, Qin, Wang, Xuan, & Chen, Jie. (2021). A CNN approach for portion size and ingredient estimation in food recognition. Pattern Recognition Letters, 143, 112–119.

[51] Liu, Chang & Cao, Yu & Luo, Yan & Chen, Guanling & Vokkarane, Vinod & Ma, Yunsheng & Chen, Songqing & Hou, Peng. (2017). A New Deep Learning-Based Food Recognition System for Dietary Assessment on An Edge Computing Service Infrastructure. IEEE Transactions on Services Computing. PP. 1-1. 10.1109/TSC.2017.2662008.

[52] S. Banerjee and A. C. Mondal, "Nutrient Food Prediction Through Deep Learning," 2021 Asian Conference on Innovation in Technology (ASIANCON), PUNE, India, 2021, pp. 1-5, doi: 10.1109/ASIANCON51346.2021.9545014.

[53] Pouladzadeh, Parisa & Kuhad, Pallavi & Peddi, Sri Vijay Bharat & Yassine, Abdulsalam & Shirmohammadi, Shervin. (2016). Food calorie measurement using deep learning neural network. 1-6. 10.1109/I2MTC.2016.7520547.

[54] Tsolakidis, Dimitris & Gymnopoulos, Lazaros & Dimitropoulos, Kosmas. (2024). Artificial Intelligence and Machine Learning Technologies for Personalized Nutrition: A Review. Informatics. 11. 62. 10.3390/informatics11030062.

[55] Veeramreddy, M., Pradhan, A. K., Ghanta, S., Rachakonda, L., & Mohanty, S. P. (2024). NUTRIVISION: A system for automatic diet management in smart healthcare. arXiv preprint, arXiv:2409.20508. https://arxiv.org/abs/2409.20508

[56] Brooke, J. (1996). SUS-A quick and dirty usability scale. Usability evaluation in industry, 189(194), 4-7.

[57] Addula, S. R., Meesala, M. K., Ravipati, P., & Sajja, G. S. (2025). A hybrid autoencoder and gated recurrent unit model optimized by honey badger algorithm for enhanced cyber threat detection in iot networks. Security and Privacy, 8(6), e70086.

[58] Miller, J. K., & Tang, W. (2025). Evaluating llm metrics through real-world capabilities. arXiv preprint arXiv:2505.08253.

[59] Ghanta, S., Thiriveedhi, A., Boyapati, P. and Pradhan, A.K., 2025. Federated transfer learning for chest X-ray classification: an explainable and generative AI framework with reliability assessment. SN Computer Science, 6(7), p.795.

[60] Xu, Z., Liu, J., Wang, Y., & Gu, Y. (2025). Latency-Response Theory Model: Evaluating Large Language Models via Response Accuracy and Chain-of-Thought Length. arXiv preprint arXiv:2512.07019.

[61] Hu, T., & Zhou, X. H. (2024). Unveiling llm evaluation focused on metrics: Challenges and solutions. arXiv preprint arXiv:2404.09135.