International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 08 | Aug - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Gen AI based Nutrition analysis system.

Anusha kc 1,Dr.Geetha M 2

1 Student, 4th Semester MCA Department of MCA,BIET,Davanagere

2 Associate professor, Department of MCA, BIET,Davanagere

Abstract_Accurate and effortless nutritional tracking is a cornerstone of modern health and wellness, yet current methods are often tedious, inaccurate, and fail to handle complex, home-cooked meals. This paper presents a novel nutrition analysis system that leverages the power of Generative Artificial Intelligence (Gen AI) to overcome these limitations. The proposed system utilizes a state-of-the-art Large Multimodal Model (LMM) to perform a deep semantic analysis of food images. Unlike traditional systems that merely classify food items, our Gen AI-based approach generates a detailed, structured breakdown of a meal, including inferred ingredients, estimated portion sizes, and cooking methods. This generated data is then used to perform a precise nutritional calculation. Furthermore, the system employs a Large Language Model (LLM) to deliver personalized, conversational feedback and actionable recommendations to the user. This framework represents a paradigm shift from simple data logging to an interactive, intelligent nutritional partner, making personalized health management more accessible and effective.

Keywords—Generative AI, Nutrition Analysis, Large Multimodal Models (LMM), Computer Vision, Personalized Health, Natural Language Generation (NLG), Food Recognition.

I. INTRODUCTION

The growing global awareness of the link between diet and health has led to a surge in demand for personalized nutrition management tools. Accurate tracking of caloric intake, macronutrients (proteins, fats, carbohydrates), and micronutrients is essential for managing weight, controlling chronic diseases like diabetes, and optimizing athletic performance. However, the current landscape of nutrition applications, while useful, suffers from significant usability challenges.

Manual logging, the most common method, is laborious and prone to error. Users must meticulously search for each ingredient, estimate portion sizes, and log them individually. Barcode scanning simplifies the process for packaged goods but is ineffective for fresh produce or complex meals. While computer vision- based apps have emerged, they typically rely on discriminative AI models (e.g., CNN classifiers) that can identify "chicken" or "rice" but fail to understand context. They cannot easily differentiate between fried and grilled chicken, estimate the quantity accurately, or deconstruct a complex dish like a stew into its constituent ingredients.

This paper proposes a fundamentally different approach, built on the recent advancements in Generative AI. We introduce a system that moves beyond simple classification to sophisticated generation and reasoning. By using a Large Multimodal Model (LMM), our system analyzes an image of a meal and does not just classify it, but *generates a detailed description* of its components. This generative capability allows for a far more nuanced and accurate analysis of home-cooked and restaurant meals. Furthermore, the system uses a Large Language Model (LLM) to transform raw nutritional data into personalized, encouraging, and actionable advice, making the user's health journey more engaging.

II. RELATED WORK

The field of automated nutrition analysis has evolved in lockstep with advancements in AI.

Early computational approaches relied on structured databases like the USDA National Nutrient Database [1]. Users would manually search for and log food items. The primary limitation was the entirely manual and tedious nature of the process.

The first wave of AI integration involved **computer vision using discriminative models**. Researchers applied Convolutional Neural Networks (CNNs) to food recognition, training models on datasets like Food-101 or UEC-Food256 to classify images of different food types [2], [3]. These systems could successfully identify single-item foods but struggled with several key challenges:

- **Poor performance on mixed dishes:** A CNN might classify a picture of a casserole as "casserole" without being able to identify the ingredients within it.

- **Inaccurate portion size estimation:** Estimating volume from a 2D image is an ill-posed problem for standard classifiers.
- **Lack of cooking method inference:** The models could not reliably distinguish between different preparation methods (e.g., steamed vs. fried), which have a significant impact on nutritional content.

The second wave of innovation came with **Generative Artificial Intelligence**. The development of powerful foundation models like Transformers and, more recently, Large Multimodal Models (LMMs) such as GPT-4 with Vision, LLaVA, and Gemini has unlocked new capabilities [4], [5]. These models can perform complex reasoning tasks that bridge vision and language. They can "look" at an image and generate a detailed textual description, answer questions about its content, and perform "few-shot" learning on novel tasks. This ability to generate structured descriptions from an image, rather than just a single class label, is the key technological leap that our work harnesses for the specific domain of nutrition analysis.

III. METHODOLOGY

The proposed system is architected as an end-to-end pipeline that transforms a user's food photo into personalized nutritional insights. The core of the system is a specialized Gen AI module.

A. System Architecture

The system operates through the following stages:

1. **Multimodal Input:** The user provides an image of their meal and can optionally add textual context (e.g., "I cooked this with one tablespoon of olive oil").
2. **Generative Meal Analysis:** The input is processed by a fine-tuned LMM. The model generates a structured output (e.g., a JSON object) that details the inferred ingredients, estimated quantities, and cooking methods.
3. **Nutritional Calculation:** The structured data from the previous step is used to query a comprehensive nutritional database. The system then calculates the total calories, macronutrients, and key micronutrients for the meal.
4. **Personalized Feedback Generation:** The calculated nutritional data, along with the user's profile (goals, dietary restrictions, past meals), is fed into an LLM to generate conversational, actionable feedback.

B. Core Module: Generative Meal Analysis

This module is the heart of the system and performs the most innovative task.

- **Model:** We propose using a powerful LMM, fine-tuned specifically for the task of food analysis. Fine-tuning is performed on a custom dataset of food images paired with expert-annotated, structured descriptions of their contents.
- **Prompt Engineering:** For a given user image, a carefully crafted internal prompt is sent to the LMM. For example: Analyze the provided image of a meal. Identify every food component. For each component, estimate its weight in grams. Infer the cooking method used. Return the output as a structured JSON array of objects, with keys "item", "quantity_grams", and "cooking_method".
- **Generative Output:** The LMM does not classify; it *generates* the structured text. A sample JSON output for a meal might be:
- [
- {"item": "Salmon Fillet", "quantity_grams": 170, "cooking_method": "pan- seared"},
- {"item": "Asparagus", "quantity_grams": 100, "cooking_method": "roasted"},
- {"item": "Quinoa", "quantity_grams": 150, "cooking_method": "boiled"}
-]

This structured, generative approach allows the system to handle complex, multi-ingredient dishes with unprecedented accuracy.

C. Module 2: Personalized Feedback Generation

This module transforms data into a user- centric experience.

- **Model:** A general-purpose LLM (like GPT-3.5 or a similar model) is used for this task.
- **Contextual Prompting:** The LLM is provided with a rich context, including the user's health goals (e.g., "lose weight," "build muscle"), dietary preferences, and the nutritional summary of the meal just analyzed.
- **Generated Output:** The LLM generates a conversational response. For instance, instead of just showing "550 kcal, 40g Protein," it might generate: "Great choice! This meal packed 40g of protein, which is perfect for your muscle-building goal. It was a little low on fiber. For your next meal, consider adding some beans or a whole-wheat side to help you feel full longer."

IV. RESULTS AND DISCUSSION

This section describes the expected functional outputs and performance of the Gen AI-based system.

A. User Interface and System Output The system's effectiveness is demonstrated through its user interface.

- A snapshot would show a user having uploaded a photo of their lunch.
- The screen would display the **Generated Meal Components** as an itemized list (e.g., "1. Salmon Fillet

(~170g)", "2. Roasted Asparagus (~100g)").

- A section would show the **Nutritional Breakdown** with clear graphics, such as a pie chart for macronutrients.
- The most prominent section would be the **AI-Generated Insight**, displaying the conversational feedback as described in the methodology.

B. Discussion

The Gen AI-based approach provides a quantum leap in usability and accuracy over previous systems.

- **Handling Complexity:** The system's ability to deconstruct mixed dishes into their constituent ingredients is its primary advantage.
- **User Experience:** The move from manual data entry to a simple photo- and-receive-feedback loop dramatically lowers the barrier to consistent nutritional tracking. The conversational nature of the feedback fosters user engagement and adherence.

C. Limitations and Ethical Considerations

It is crucial to acknowledge the limitations of this technology:

- **Accuracy of Estimation:** While far better than previous methods, portion size estimation from a 2D image is still an approximation and not a replacement for a food scale. The model's accuracy is a key area for evaluation.
- **Hallucinations:** Generative models can sometimes "hallucinate" or invent ingredients that are not present. Fine- tuning and rigorous testing are required to minimize this.
- **Data Bias:** The model's performance will be best on the types of cuisine most represented in its training data. It may struggle with less common or ethnically diverse foods if not trained on them.
- **Medical Disclaimer:** The system must be positioned as a wellness and informational tool, **not a medical device**. Users must be clearly informed that the advice generated is not a substitute for consultation with a registered dietitian or doctor, especially for managing medical conditions.

V. CONCLUSION AND FUTURE WORK

This paper has introduced a Gen AI-based nutrition analysis system that represents a fundamental advancement in automated dietary tracking. By leveraging Large Multimodal Models to generate detailed, structured descriptions of meals from images, the system achieves a new level of accuracy and contextual understanding. Coupled with an LLM for personalized feedback, it creates an intuitive, engaging, and powerful tool for personal health management.

Future work will focus on several key areas to further enhance the system:

1. **Improving Portion Size Accuracy:** Incorporating techniques that use depth sensors (if available on the phone) or ask the user for a reference object (like a coin or their hand) in the photo to better calibrate a size estimate.
2. **Long-Term Trend Analysis and Generation:** Training the LLM to analyze a user's dietary patterns over weeks or months and generate comprehensive summary reports and long-term strategic advice.
3. **Meal Plan and Recipe Generation:** Extending the system to not only analyze past meals but also to generate complete, personalized weekly meal plans and recipes that align with the user's goals and preferences.
4. **Continuous Learning and Personalization:** Implementing a feedback loop where users can correct the model's generated ingredient list, allowing the system to continuously learn and personalize its performance for each individual user.

REFERENCES

[1] U.S. Department of Agriculture, Agricultural Research Service. FoodData Central, 2019. [Online]. Available: https://fdc.nal.usda.gov/

[2] L. Bossard, M. Guillaumin, and L. Van Gool, "Food-101 – Mining Discriminative Components with Random Forests," in *Proc. European Conference on Computer Vision (ECCV)*, 2014.

[3] K. Matsuda, Y. and K. Yanai, "Recognition of multiple-food images by detecting and recognizing individual food items," in *Proc. IEEE International Conference on Multimedia and Expo (ICME)*, 2012.

[4] A. Radford et al., "Learning Transferable Visual Models From Natural Language Supervision," in *Proc. International Conference on Machine Learning (ICML)*, 2021.

[5] A. Z. et al., "LLaVA: Large Language and Vision Assistant," *arXiv preprint arXiv:2304.08485*, 2023.