**e-ISSN:** *3048-7552*

**Volume 08 Issue 02**
**May-Aug, 2026**

*__Corresponding Author: Aishwarya Patil__,
*Student, Department of Information Science and Engineering, Poojya Doddappa Appa College of Engineering Kalaburagi, Karnataka, India*

# An AI-Based Food Calorie Estimation System Using Google Gemini API

***Aishwarya Patil [1], C. G. Balasubramanyam Singh [2], Madhushree[3], Pradeep S [4]***
*[1-4]Student, Department of Information Science and Engineering, Poojya Doddappa Appa College of Engineering Kalaburagi, Karnataka, India*

***ABSTRACT***
*Accurate monitoring of calorie intake is essential for maintaining a healthy lifestyle and preventing nutrition-related disorders. Conventional calorie tracking systems rely heavily on manual food logging, which is time-consuming and often inaccurate, particularly for mixed and homemade meals. This paper presents an AI-based food calorie estimation system using the Google Gemini multimodal API. The proposed system enables users to upload food images through a simple web interface and receive real-time calorie and nutritional analysis. By leveraging advanced image understanding and generative AI, the system reduces manual effort, improves usability, and supports informed dietary decisions.*

***Keywords:*** *Google Gemini multimodal API.*

## 1. INTRODUCTION

In recent years, lifestyle-related diseases such as obesity, diabetes, cardiovascular disorders, and nutritional deficiencies have increased at an alarming rate. One of the major causes of these health issues is poor dietary awareness and improper estimation of daily calorie intake. Maintaining a balanced diet requires not only consuming healthy food but also understanding the nutritional composition of meals.

Traditional calorie tracking methods depend on manual food entry, where users must specify food items, portion sizes, and preparation details. This process is inconvenient, error-prone, and unsuitable for mixed or homemade meals. As a result, many users discontinue diet tracking applications. Recent advancements in artificial intelligence, deep learning, and computer vision enable automated food recognition using images. These technologies offer an efficient alternative to manual calorie tracking. This work proposes an intelligent AI-based system for image-driven calorie estimation.

## 2. LITERATURE SURVEY

Food recognition and calorie estimation using artificial intelligence has gained significant attention due to its potential to automate dietary assessment. Early research relied on traditional image processing techniques which provided limited accuracy. With the evolution of deep learning, convolutional neural networks (CNNs) became widely used for food image classification.

Purandhar et al. (2025) – MobileNet-Based Mixed Food Recognition Purandhar et al. proposed a deep learning–based food recognition and calorie estimation system using the MobileNet architecture to handle mixed food items. The primary focus of the study was to reduce computational complexity while maintaining acceptable accuracy, enabling deployment on mobile devices. The model employed multi-label classification and segmentation techniques to identify multiple food components within a single dish. Although the system achieved faster inference and efficiency, accurate portion size estimation for complex meals remained a significant challenge, which affected calorie estimation accuracy.

**Mareddy et al. (2025) – CNN-Based Food Recognition and Calorie Prediction**

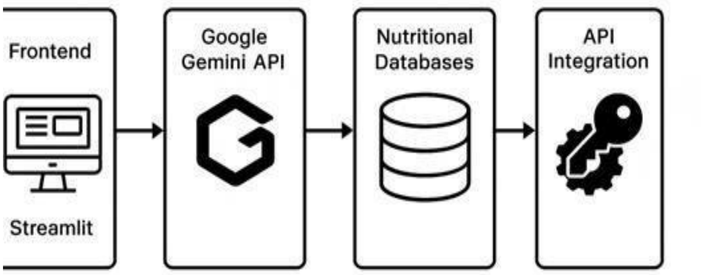
Mareddy et al. (2025) developed a convolutional neural network (CNN)–based food image recognition and calorie prediction system to support health and fitness monitoring. The system automatically identified food items from images and estimated calorie values using predefined nutritional databases, thereby reducing the need for manual food entry. The approach improved user convenience and achieved good classification accuracy for common food items. However, the system relied on static datasets, which limited its adaptability to regional cuisines and homemade dishes. Additionally, variations in cooking methods and portion sizes were not effectively addressed, affecting calorie estimation accuracy in real-world scenarios. These limitations indicate the need for more adaptive and context-aware calorie estimation techniques.

**Pothina et al. (2024) – Object Detection–Based Nutritional Assessment**

Pothina et al. (2024) introduced an AI-based nutritional assessment framework that utilized object detection techniques to identify food items within an image. The system improved food localization and enabled calorie estimation by quantifying detected items. Despite these advancements, the framework required controlled image capture conditions, including consistent lighting and backgrounds. This dependency reduced its robustness and effectiveness in real-world dietary monitoring scenarios.

**Parvathavarthini et al. (2025) – Multi-Model Deep Learning Approach**

Parvathavarthini et al. (2025) proposed a deep learning–based dietary tracking system that integrated CNN, AlexNet, and EfficientNet architectures. The system used transfer learning and fine-tuning to improve classification accuracy and supported real-time food recognition using webcam input. Although the approach demonstrated improved performance metrics, the use of multiple deep learning models increased training complexity and computational cost, limiting its scalability for lightweight applications.




**Sam Leo and Ruth Moly Benjamin (2024) – Web-camBased Nutrition Tracking**

Sam Leo and Ruth Moly Benjamin (2024) developed a webcam-based nutrition tracking system focused on real-time recognition of fruits and vegetables using convolutional neural networks. The system emphasized user convenience and real-time feedback. While it performed effectively for raw fruits and vegetables, it struggled with cooked foods and mixed meals. This limitation reduced its applicability to everyday dietary habits involving complex dishes.

**Sanitha et al. (2025) – Explainable AI-Based Diet Tracking System**

Sanitha et al. (2025) introduced an explainable AI- based diet tracking system designed to prevent nutrition- related disorders such as obesity and diabetes. The system combined deep learning for food recognition with explain ability techniques such as SHAP and LIME to enhance transparency. Although the explainable remain unresolved, motivating the proposed system.

Approach improved user trust and clinical relevance, the system required extensive dataset coverage and additional computational resources, posing challenges for real-time deployment.

**Indira Priyadharshini et al. (2024) – Ripeness Prediction and Nutrition Analysis**

Indira Priyadharshini et al. (2024) proposed an AI framework for predicting ripeness levels and nutritional content of regional fruits and vegetables. The system employed CNN, KNN, and Random Forest models to estimate nutrients such as calories, protein, iron, and calcium. The study emphasized region-specific food recognition but was limited to raw fruits and vegetables, excluding cooked and mixed food items.

**Zhang et al. (2022) – Survey on Image-Based Dietary Assessment**

Zhang et al. (2022) presented a comprehensive survey on image-based dietary assessment methods, covering traditional image processing, deep learning, and hybrid approaches. The study identified key challenges such as portion size estimation, dataset diversity, and occlusion in mixed food images. The authors highlighted that accurate calorie estimation requires both reliable food recognition and precise volume estimation, revealing gaps in existing systems.

**Yang et al. (2021) – Smartphone-Based Portion Size Estimation**

Yang et al. (2021) explored smartphone-based food portion size estimation without using fiducial markers. Their approach utilized geometric modeling and depth cues to estimate food volume. While the method addressed a critical challenge in calorie estimation, it faced limitations when handling irregular food shapes and complex meal compositions, affecting real-world accuracy.

**Multimodal Generative AI for Food Understanding (2024–2025)**

Recent studies (2024–2025) on multimodal generative artificial intelligence demonstrate improved food understanding by combining visual and textual information. These models reduce dependency on manual feature extraction and enhance contextual reasoning, enabling better handling of mixed meals. However, challenges such as real-time performance, dataset bias, and accurate portion size estimation remain open research problems. These limitations motivate the proposed system that leverages multimodal AI for robust calorie estimation.

Recent advancements in multimodal generative AI demonstrate improved food understanding by combining visual and textual information. However, challenges such as portion size estimation, real-time processing, and dataset diversity.

## 3. PROBLEM STATEMENT

Despite significant advancements in AI-based food recognition systems, accurate calorie estimation remains a challenging task. Most existing solutions rely on manual food entry, predefined nutritional databases, or controlled imaging environments. Such approaches reduce usability and accuracy in real-world scenarios, especially when dealing with mixed dishes, homemade meals, and region-specific foods.

Furthermore, many applications lack real-time feedback and require extensive user interaction, leading to low user adoption. Variations in portion size, lighting conditions, image quality, and food presentation further impact estimation accuracy. These limitations highlight the need for an intelligent, automated, and user-friendly system that can analyze food images directly and provide accurate calorie and nutritional information with minimal user effort.

## 4. PROPOSED SYSTEM AND METHODOLOGY

The proposed system aims to provide an intelligent, automated, and user-friendly solution for food recognition and calorie estimation using artificial intelligence. The primary objective of this system is to overcome the limitations of traditional calorie tracking methods that rely on manual food logging, predefined datasets, and controlled environments. By leveraging recent advancements in multimodal generative AI, the proposed approach enables accurate dietary assessment directly from food images.

### A. Overview of the Proposed System

The proposed system allows users to upload images of food items through a simple web-based interface. Once the image is uploaded, the system automatically analyzes the visual content to identify food items and estimate their calorie and nutritional values. The system eliminates the need for manual food selection, portion size entry, or reference objects, thereby reducing user effort and improving accuracy.

The methodology integrates image processing techniques with the **Google Gemini multimodal API**, which is capable of understanding both visual and contextual information. This integration enables the system to handle mixed meals, overlapping food items, and diverse food presentations commonly encountered in real-world scenarios.

### B. System Components

The proposed system consists of three main components:

- **User Interface Layer**
- **Backend Processing Layer**
- **AI Analysis Layer**

Each component plays a crucial role in ensuring operation and accurate nutritional estimation.

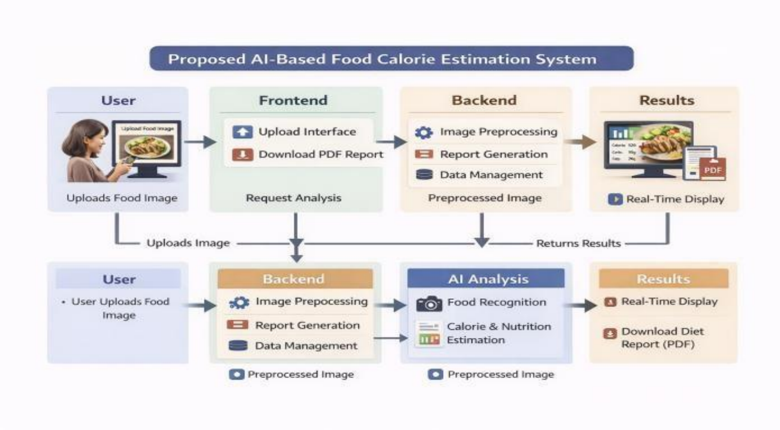



***Fig.1:-** Block Diagram of the Proposed AI-Based Food Calorie Estimation System*

### C. User Interface Layer

The user interface layer is developed using a lightweight web framework that provides a simple and intuitive platform for user interaction. Users can upload food images captured using smartphones or webcams. The interface is designed to minimize complexity, making the system accessible to users with minimal technical knowledge.

Once an image is uploaded, the interface displays the analyzed results, including identified food items, estimated calorie values, and nutritional breakdown. The interface also provides an option to download a detailed PDF report for recordkeeping and long-term dietary monitoring

### D. Backend Processing Layer

The backend processing layer acts as an intermediary between the user interface and the AI analysis layer. After receiving the uploaded image, this layer performs necessary preprocessing steps such as resizing, normalization, and format conversion to ensure compatibility with the AI model. Preprocessing improves image quality and enhances recognition accuracy.

The backend also manages communication with the Google Gemini API, handles response parsing, and organizes the nutritional data into a structured format. Additionally, it manages report generation and ensures real-time delivery of results to the user interface

### E. AI Analysis Layer Using Google Gemini API

The core intelligence of the proposed system lies in the AI analysis layer, which is powered by the **Google Gemini multimodal API**. Unlike traditional CNN-based approaches that rely solely on visual classification, Gemini combines image understanding with contextual reasoning to interpret complex food images.

When an image is submitted, the Gemini model analyzes visual features such as shape, texture, color, and composition to identify food items. It also applies contextual reasoning to estimate nutritional values, even for mixed or unfamiliar dishes. This multimodal capability allows the system to generalize better across diverse food types and regional cuisines.

The AI model outputs estimated calorie values along with nutritional components such as carbohydrates, proteins, and fats. This approach significantly improves flexibility and reduces dependency on static food databases.

### F. Workflow of the Proposed Methodology

The methodology follows a sequential workflow:

- User uploads a food image through the web interface.
- The image undergoes preprocessing in the backend.
- The processed image is sent to the Google Gemini API.
- The AI model identifies food items and estimates nutritional values.
- Results are displayed in real time to the user.
- A downloadable PDF report is generated for future reference.

This workflow ensures minimal user interaction while maintaining high accuracy and efficiency.

### G. Advantages of the Proposed System

The proposed system offers several advantages over existing methods:

- Eliminates manual food logging and portion size estimation
- Supports mixed meals and diverse food presentations
- Provides real-time calorie estimation
- Reduces dependency on predefined datasets
- Enhances user convenience and accessibility
- Scalable for future enhancements such as personalized recommendations

### H. Summary of Methodology

In summary, the proposed system integrates a user-friendly interface with a powerful AI backend to deliver an efficient and accurate food calorie estimation solution. By leveraging the multimodal capabilities of the Google Gemini API, the system addresses key challenges identified in existing literature and provides a practical solution for real-world dietary assessment. This methodology forms the core contribution of the research and demonstrates the potential of generative AI in healthcare and nutrition monitoring applications.

## 5. WORKFLOW

The workflow of the proposed AI-based food calorie estimation system begins when the user uploads an image of the food item through the frontend interface developed using Streamlit. The uploaded image is first validated and preprocessed to ensure appropriate size, format, and quality for further analysis. Preprocessing operations such as resizing and normalization help improve recognition accuracy.

After preprocessing, the image is forwarded to the Google Gemini multimodal API for analysis. The Gemini model examines the visual features of the image, including shape, texture, color, and composition, to identify the food items present. Using its multimodal capabilities, the model estimates the calorie content and nutritional values such as carbohydrates, proteins, and fats.

Once the analysis is complete, the generated nutritional information is sent back to the backend system. The results are then displayed to the user in real time through the web interface. Additionally, a detailed PDF report containing the food analysis and calorie estimation is generated, allowing users to download and store their dietary information for future reference.

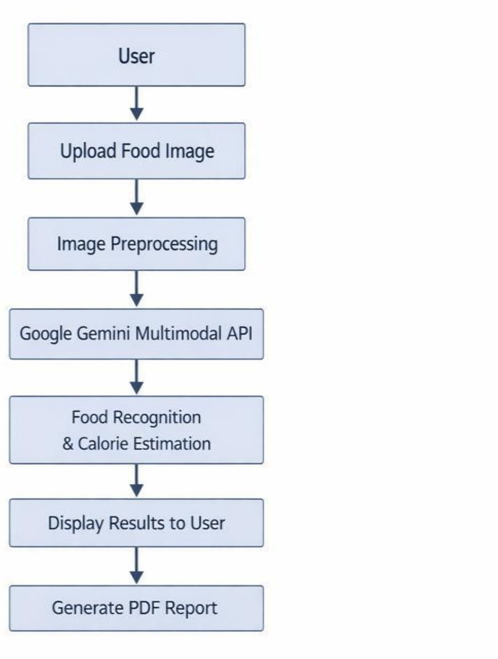



***Fig.2:-*** *Workflow of the Proposed AI-Based Food Calorie Estimation System*

## 6. SYSTEM ARCHITECTURE

The system architecture consists of a user interface layer for image upload, a backend processing layer for image handling, and an AI analysis layer powered by the Google Gemini API.

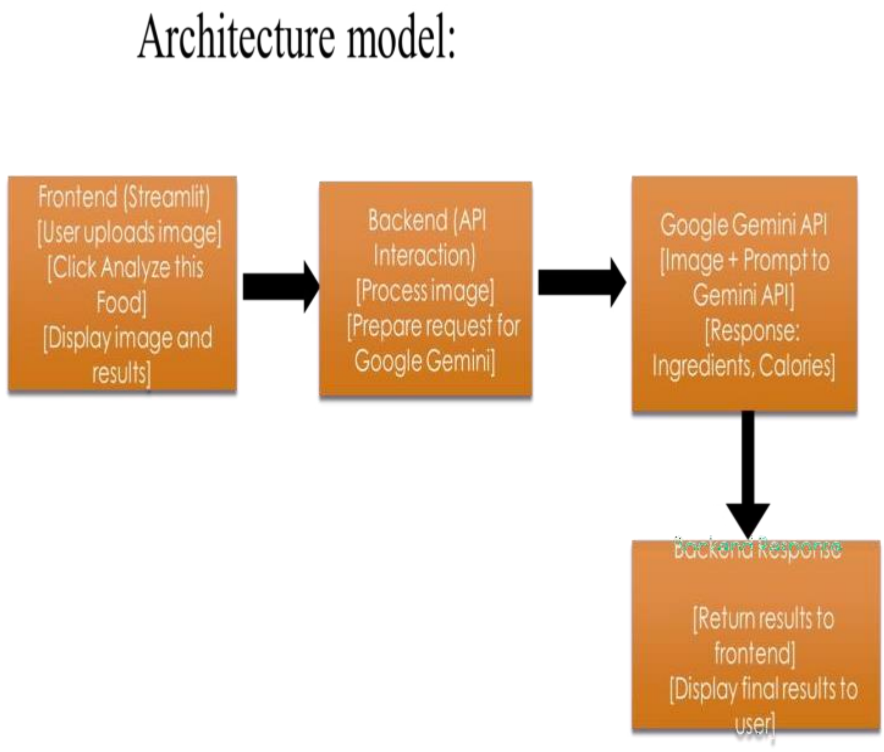



***Fig.3:-*** *System Architecture of Food Recognition and Calorie Estimation*

## 7. RESULTS AND DISCUSSION

Experimental evaluation shows that the proposed system accurately identifies food items and estimates calorie values in real time. The integration of the Google Gemini multimodal API improves recognition performance, particularly for mixed food items. Compared to manual calorie tracking methods, the system significantly reduces user effort and enhances usability. Minor variations in estimation accuracy were observed due to portion size and image quality; however, overall results demonstrate that the system provides an efficient and practical solution for automated dietary assessment.

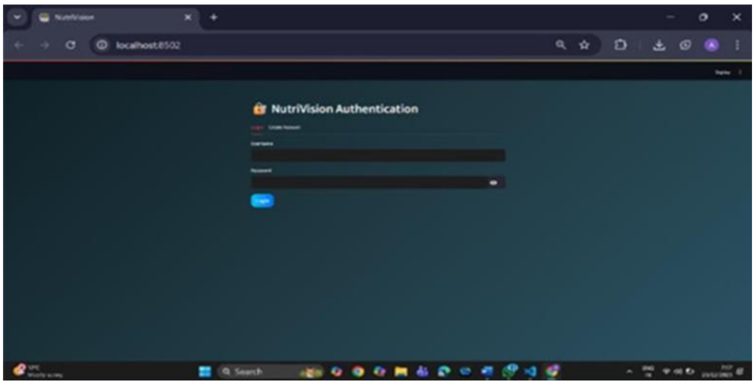



***Fig.4:-** Nutrivision User Login Interface*

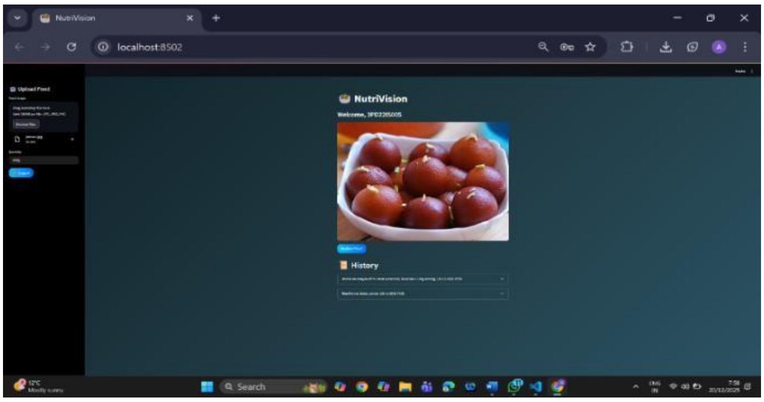



***Fig.5:-** Food Image Uploaded Interface*

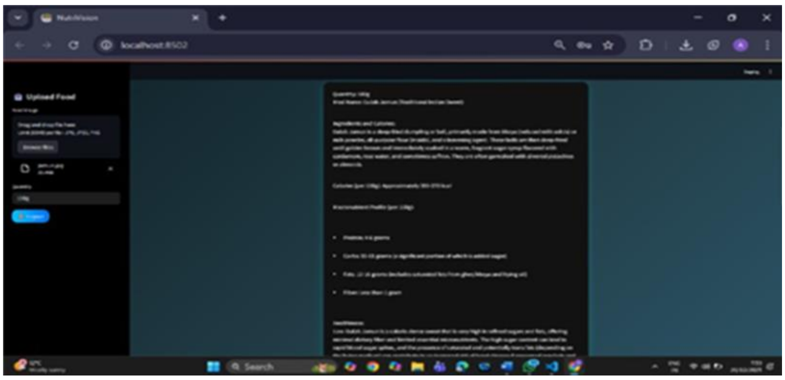


***Fig.6:-**Nutritional Analysis Output Screen*

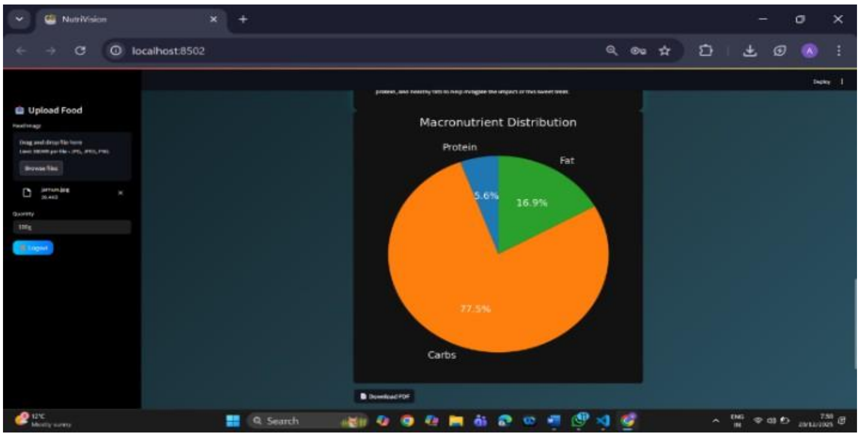



*Fig.7:-Macronutrient Distribution Chart*

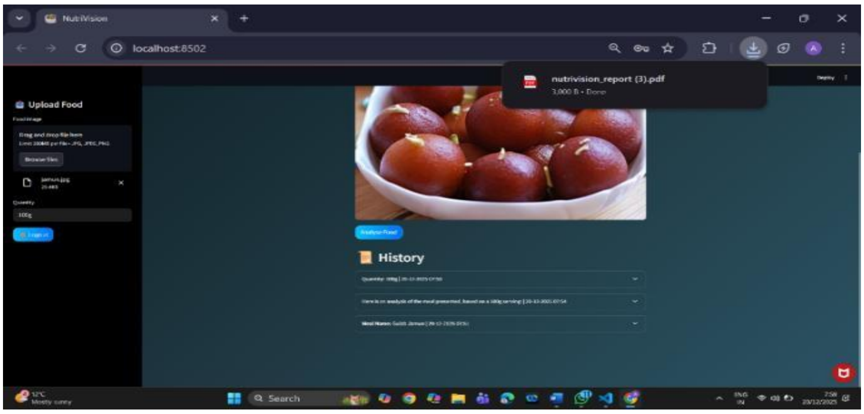



*Fig.8:-Generated PDF Report*

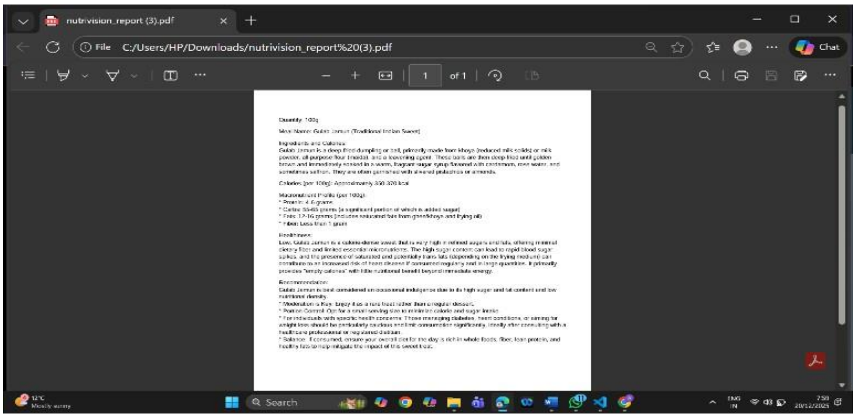



*Fig.9:- Sample PDF Output of the Proposed System*

## 8. CONCLUSION

This paper presented an AI-based food calorie estimation system using the Google Gemini multimodal API to automate dietary assessment. The proposed system enables users to analyze food images and obtain calorie and nutritional information in real time, reducing the need for manual food logging. By leveraging multimodal artificial intelligence, the system effectively handles mixed food items and improves usability compared to traditional approaches. Experimental results demonstrate that the system is efficient, user-friendly, and suitable for real-world dietary monitoring applications.

## REFERENCES


1. Purandhar, N., Poojitha, S., Hussain, S. M., Chowdary, M. P., & Rafi, M. (2025). Food recognition and calorie estimation in mixed food items using MobileNet. In *Proceedings of the 6th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*. Tirunelveli, India.
2. Anusha, M., Nandan, T. P. K., Shivani, M., Kaif, M., Mithra, P. S., & Kumar, K. C. (2025). Food image recognition and calorie prediction system using CNN. In *Proceedings of the 5th International Conference on Soft Computing for Security Applications (ICSCSA)*. Salem, India.
3. Sabarinath, P. M., Sudarshanreddy, G., Patnaik, M. S. P. K., & Rathnala, P. (2024). AI-based nutritional assessment for object detection and food quantification. In *Proceedings of the 5th International Conference on Electronics and Sustainable Communication Systems (ICESC)*.
4. Parvathavarthini, S., Madhuretha, R., Subashree, N. K., Vishnupriya, M., & Divya, K. (2025). A deep learning approach for food recognition and nutritional analysis. In *Proceedings of the 8th International Conference on Computing Methodologies and Communication (ICCMC)*. Erode, India.
5. Leo, S., & Benjamin, R. M. (2024). Enhancing nutrition tracking: Webcam-based fruit and vegetable recognition and calorie measurement using image processing and convolutional neural networks. In *Proceedings of the 5th International Conference on Emerging Technology (INCET)*. Belgaum, India.
6. Sanitha, P. C., Parveen, S. N., Thaherbasha, S., Shanmugapriya, M., Kalaivani, T., & Senthilkumar, R. (2025). Transparent nutrition: An explainable AI-based diet tracking system for preventing nutrition-related disorders. In *Proceedings of the 3rd International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)*. Coimbatore, India.
7. Priyadharshini, M. I., Anusha, G., Padmavathy, C., Lokshana, R. J., Pavithra, V., & Murugeshwari, M. (2024). An AI framework model for smart prediction of ripeness in regional fruits and vegetables. In *Proceedings of the 4th International Conference on Sustainable Expert Systems (ICSES)*. Kaski, Nepal.
8. Zhang, Y., Yang, M., & Chen, L. (2019). Image-based dietary assessment: A review of methods, challenges, and future directions. *IEEE Access*, *7*, 114285–114298. https://doi.org/10.1109/ACCESS.2019.2936117
9. Yang, Y., Jia, J., & Xu, L. (2021). Food portion size estimation using smartphone images without fiducial markers. *IEEE Journal of Biomedical and Health Informatics*, *25*(6).
10. Google. (2024). *Gemini API: Multimodal AI for image and text understanding*. Google AI Documentation. https://ai.google.dev