IJERST International Journal of Engineering Research and Science & Technology

MEASURING NUTRITION AND CALORIES FROM FOOD IMAGE CALCULATION IN DAY TO DAY LIFE

K. ISHWARYA DEVI, K. YAMUNA , M.SAIKRISHNA , M.SANTOSH

Assistant professor ,UG Student, Department of Computer Science and Engineering(Aritificial Intelligence And Machine Learning)

TKR College Of Engineering and Technology ,Medbowli,Meerpet ,Balapur,Hyderabad,

Telangana 500097,India

Abstarct:

Nutrition estimation from food images helps in understanding calorie and nutrient intake without using laboratory methods. Traditional chemical techniques are slow, costly, and destroy the food sample, whereas image-based methods are fast and non-destructive. This project uses RGB and Depth images along with a deep-learning model to accurately estimate calories, protein, fat, and carbohydrates. The system adopts the Swin Transformer for feature extraction and an Adaptive Feature Fusion and Enhancement (ADFE) module to combine RGB–D information effectively. Detailed-information enhancement and semantic information enhancement further improve feature quality. Experiments on the Nutrition5k dataset show that the model provides reliable nutrition prediction, making it useful for daily food analysis and health monitoring.

Keywords: Food Nutrition Estimation, RGB-D Images, Swin Transformer, Adaptive Feature Fusion (ADFE), Deep Learning Feature Enhancement, Multimodal Data, Nutrition5k Data.

1.Introduction

Healthy eating habits play an important role in preventing physical and mental health problems. Many studies show that poor dietary patterns are strongly linked to diseases such as obesity, diabetes, depression, and heart-related issues. To improve eating behaviour, it is essential to accurately estimate the nutritional content of foods, including calories, fat, protein, and carbohydrates. Traditional nutrition analysis methods such as chromatography, spectrophotometry, and chemical testing are accurate but require special equipment, professional expertise, and often destroy the food sample. Therefore, they are not suitable for daily food analysis. With advancements in artificial intelligence and computer vision, image-based nutrition estimation has become a fast, non-destructive, and more practical solution. Vision-based systems use food images captured from mobile devices to estimate nutritional values, making them convenient for personal health monitoring. However, many early methods relied only on RGB images and failed to utilize important 3D spatial information. To overcome these limitations, recent research focuses on multimodal approaches that combine RGB images with depth information to improve accuracy. Models such as Swin Nutrition, Google-Nutrition, and fusion networks have shown better performance by integrating spatial and visual features. Motivated by these developments, our project uses RGB–Depth images along with the Swin Transformer and an Adaptive Feature Fusion and Enhancement (ADFE) model to extract richer features and produce more accurate nutrition predictions. This system aims to support healthy eating habits and contribute to the development of modern food analysis and precision nutrition.

2. Literature Survey:

[1] Food Item Calorie Estimation using YOLOv4 and Image Processing . This work focuses on estimating food calories using deep learning models like YOLO and image processing techniques. It detects food items, estimates portion size, and calculates calories using nutrition databases. The system improves accuracy compared to manual methods. It supports automated and efficient dietary monitoring.

[2] Food Volume Estimation using 3D Projection and Wire Mesh . This study estimates food volume using 3D image projection and geometric modeling techniques. YOLO-based detection identifies food items, while wire-mesh and calibration methods estimate portion size. The system improves volume accuracy from 2D images. It enables better calorie estimation with minimal user effort.

[3] Food Recognition and Volume Estimation using Deep Learning.This approach uses CNNs and models like YOLO and Swin Transformers for food recognition. It estimates portion size using geometric and 3D techniques. Nutritional values are calculated using databases like USDA. The system enhances automated dietary tracking accuracy.

[4] Ingredient-Based Food Calorie Estimation System . This system focuses on ingredient-level analysis for accurate calorie estimation. It combines image processing, nutrition knowledge, and databases to calculate food values. Multi-modal features like brightness and heat improve accuracy. It provides detailed nutrition insights beyond simple dish recognition.

[5] Food Image-Based Calorie Estimation using Multi-Feature Fusion . This method integrates deep learning with ingredient analysis and multi-feature fusion. It uses visual cues like texture, brightness, and structure for better estimation. The system calculates calories, proteins, and fats accurately. It improves performance over traditional single-feature models.

[6] Advanced Multi-Feature Fusion for Nutrition Estimation . This study combines deep learning, ingredient recognition, and multi-feature fusion techniques. It uses models like CNN, YOLO, and Swin Transformers for detection. Volume estimation and nutrition mapping are done using databases. The system provides detailed and accurate dietary analysis.

[7] Ingredient-Based Food Recognition using Deep Learning . This research focuses on recognizing ingredients using region-based and multi-task learning. It improves accuracy in complex food images with mixed ingredients. Volume estimation techniques are also applied for calorie calculation. It enhances real-world dietary assessment systems.

[8] Measuring Calorie and Nutrition from Food Image . This work uses image processing and

machine learning to estimate food calories. It analyzes features like color, texture, and size for food identification. Advanced segmentation improves accuracy. The system reduces manual effort in diet tracking.

[9] Estimating Food Amount in a Circular Bowl This study focuses on estimating food quantity in bowls using single images. It models bowl geometry and applies image processing techniques. The approach reduces dependency on reference objects. It improves portion estimation in real-world scenarios.

[10] Multi-Task Deep Learning for Ingredient Recognition . This work uses multi-task and region-wise deep learning to identify ingredients in food images. It handles complex dishes with mixed components. Large datasets improve model performance. It enhances calorie estimation accuracy through better ingredient detection.

[11] Prediction of Mental State from Food Images . This study explores the relationship between food habits and mental health using meal images. It analyzes patterns like portion size and meal frequency. Deep learning helps extract behavioral insights. It extends food analysis beyond nutrition to lifestyle monitoring.

[12] Review of Food Recognition and Volume Estimation Systems. This paper reviews AI-based food recognition and volume estimation methods. It discusses techniques like segmentation, classification, and 3D reconstruction. Strengths and limitations of existing approaches are analyzed. It highlights challenges in real-world applications.

[13] Food Recognition and Calorie Prediction using Deep Learning . This system uses CNNs and classifiers like SVM for food recognition. It detects food items and estimates calorie values automatically. The approach simplifies dietary tracking. It supports healthier lifestyle management.

[14] Food Nutrition Estimation using Machine Learning. This study uses machine learning to estimate food calories from images. It applies CNNs for segmentation and volume estimation. The system automates dietary tracking. It helps reduce obesity and improve health awareness.

[15] Food Recognition and Calorie Estimation using Deep Learning.This approach uses CNN-based models for accurate food identification. It calculates calorie values based on recognized items. The system is user-friendly and efficient. It promotes healthy eating habits through automation.

[16] Food Calorie Estimation with Spoilage Detection . This system combines deep learning with IoT sensors to detect food spoilage and estimate calories. It uses image recognition for classification. Sensors monitor freshness levels. It ensures both food safety and nutritional analysis.

[17] Img2Calories – Image-Based Nutrition Estimation . This application estimates calories from food images using deep learning. It works with smartphone images and social media photos. The system provides instant nutritional information. It simplifies calorie tracking for users.

[18]Smart Nutrition Monitoring using Edge Computing. This system integrates AI with edge computing for real-time food analysis. It performs food recognition, segmentation, and calorie estimation. Edge devices improve speed and efficiency. It reduces dependency on manual tracking.

[19]Adaptive Feature Fusion for Food Nutrition Estimation .This study proposes a Swin Transformer-based model for improved feature extraction. It combines RGB and depth data using adaptive fusion techniques. The system enhances accuracy in nutrition estimation. It outperforms traditional CNN-based methods.

3. Architecture

The architecture of the Food Nutrition & Calorie Estimation System is designed to process food images and accurately predict nutritional values using deep learning-based feature and e extraction and fusion techniques. The system follows a modular pipeline structure, ensuring scalability, high accuracy, and efficient processing. The architecture begins with the User Interface, where the user uploads a food image. This image then passes to the Image Preprocessing Module, where noise removal, resizing,normalization, and enhancement operations are performed to improve model readiness.

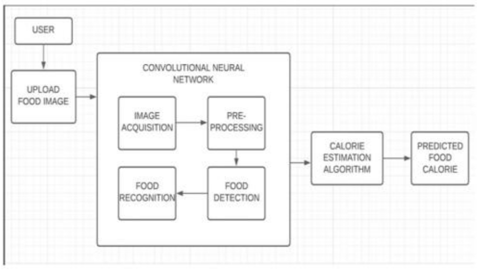


Fig 1 Architecture Diagram

4. Algorithm and Techinques

1) Swin Transformer

Swin Transformer is used for feature extraction using a hierarchical vision transformer architecture. It processes images in shifted windows to capture both local and global features efficiently. It performs better than CNNs in handling complex patterns like texture and shape variations. It extracts meaningful features from both RGB and depth images.

2) Detail Information Enhancement (DIE) DIE enhances low-level features such as edges, textures, and fine details in food images. It helps differentiate between visually similar food items. This improves feature clarity and precision. As a result, overall model accuracy is increased.

3) Semantic Information Enhancement (SIE) SIE focuses on improving high-level semantic features like shape and category of food items. It helps the model understand the overall structure and context of the image. This improves food recognition capability. It works together with detailed features for better predictions.

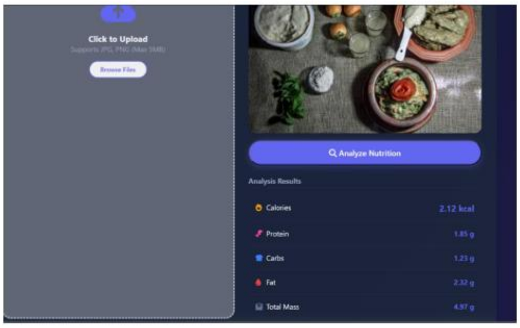
4) Adaptive Feature Fusion and Enhancement (ADFE)
ADFE combines features from RGB and depth images into a unified representation. It assigns importance to different features adaptively. This improves the integration of visual and spatial information. As a result, prediction accuracy and system performance are enhanced.

5. Proposed system:

The system provides a well-structured, simple, and user-friendly graphical interface that enables smooth interaction between the user and the application. Each form is designed to perform a specific function in the workflow, ensuring ease of use and efficient processing of food images.

Home Page:

The Home Page serves as the entry point of the application. It displays the project title, brief description, and navigation options to access different functionalities of the system. It may also include instructions or guidelines for users on how to use the application. This page ensures that even first-time users can understand the workflow easily.

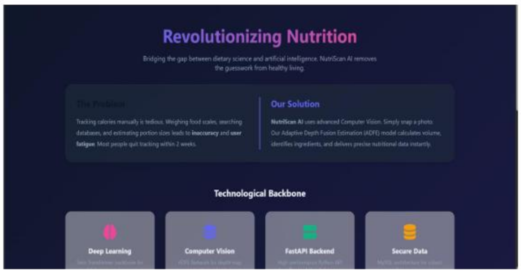


Upload Form:

The Upload Form allows users to input food images into the system. Users can upload RGB images and, if available, corresponding depth images for better accuracy. The form supports file selection, preview options, and validation to ensure correct image formats. This step is crucial as the quality of input images directly affects the accuracy of the results.

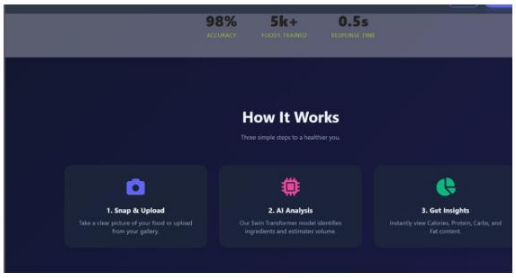
Processing Page:

The Processing Page displays the current status of the system while the uploaded images are being analyzed. It shows messages such as “Processing,” “Feature Extraction,” or “Model Running” to keep users informed. This page improves user experience by providing feedback and preventing confusion during computation.



Result page

The Result Page presents the final output of the system in a clear and structured format. It displays predicted nutritional values such as calories, proteins, fats, and carbohydrates. Additional details like identified food items and estimated portion size may also be shown. The results are organized in a readable format (tables or labels) to help users easily understand their dietary information.



6.Result Analysis:

The result analysis of the proposed system is carried out to evaluate its performance in estimating nutritional values from food images. The system processes both RGB and Depth images and predicts values such as calories, fat, carbohydrates, and protein. The performance of the model is analysed by comparing the predicted nutritional values with the actual values available in the dataset. The results show that the system produces accurate and consistent predictions for most of the food items. The use of Depth information along with RGB images helps in better understanding of food portion size, which improves the estimation accuracy. The integration of the Swin Transformer plays a significant role in extracting meaningful features from the images. It captures both local details and global patterns effectively. Additionally, the enhancement modules such as Detail Information Enhancement (DIE) and Semantic Information Enhancement (SIE) improve the quality of features by focusing on fine details and high-level information. The Adaptive Feature Fusion and Enhancement (ADFE) module further improves the performance by combining features from different sources. This fusion helps the system to generate more precise and reliable outputs. The system also performs well under different testing conditions. It handles normal inputs efficiently and provides appropriate outputs for low-quality or unclear images. In cases of invalid input, the system responds with suitable error messages, ensuring robustness. Overall, the experimental results indicate that the proposed system achieves good accuracy with low error rates. The model is efficient, stable, and suitable for real-time food nutrition estimation applications.

7. Conclusion:

The testing and validation results indicate that the proposed system performs effectively under different conditions. The system is capable of processing various types of food images and generating consistent and reliable nutritional predictions. During testing, the system successfully handled different scenarios such as valid inputs, invalid files, low-quality images, and missing inputs. In all cases, the system responded appropriately by either generating accurate results or displaying suitable error messages. The validation results show that the model achieves good accuracy with low error rates. The use of both RGB and Depth images improves the prediction performance by providing better understanding of food structure and portion size. The integration of Swin Transformer for feature extraction, along with DIE, SIE, and ADFE modules, enhances the overall efficiency and accuracy of the system. These components help in extracting meaningful features and combining them effectively for 20 better predictions. Overall, the system is stable, reliable, and efficient. It can be effectively used for real time food nutrition estimation and has potential applications in diet monitoring and health management.

8. References:

1. B. Štiti´c, L. Urbinati, G. Di Guglielmo, L. P. Carloni, and M. R. Casu, "Enhanced machine learning flow for microwave-sensing systems for contaminant detection in food," IEEE Trans. AgriFood Electron.,vol.2, no. 2, pp. 181–189, Sep./Oct. 2024

2. S. Riaz et al., "Electrical impedance spectroscopy-based detection of internal browning disorder in apples," IEEE Trans. AgriFood Electron., vol. 3, no. 1, pp. 26–33, Mar./Apr. 2025.

3. M. A. Rahman, A. A. Khan, M. M. Hasan, M. S. Rahman, and M. T. Habib, "Deep learning modeling for potato breed recognition," IEEE Trans. AgriFood Electron., vol. 2, no. 2, pp. 419–427, Sep./Oct. 2024

4.D.Kirk,C.Catal, and B.Tekinerdogan, "Precision nutrition: A systematic literature review," Comput. Biol. Med., vol. 133, 2021, Art. no. 104365

5. K. Kalinowska, W. Wojnowski, and M. Tobiszewski, "Simple analytical method for total biogenic amines content determination in wine using a smartphone," Anal. Methods, vol. 15, no. 11, pp. 1395–1401, 2023.

6. Q.Thamesetal.,"Nutrition5k: Towards automatic nutritional understanding of generic

food," in Proc. IEEE/CVF Conf. Compute. Vis. Pattern Recognit., 2021, pp. 8903–8911.

7.D.Zhang,X.-J.Wu,andJ.Yu,"Discrete bidirectional matrix factorization hashing for zero-shot cross media retrieval," in Proc. Chin. Conf. Pattern Recognit. Comput. Vis., 2021, pp. 524–536.

8.W.Min,S.Jiang,L.Liu,Y.Rui,andR.Jain,"Asurve yonfoodcomputing," ACMComput. Surv., vol. 52, no. 5, pp. 1–36, 2019.

9. Y. Lu, T. Stathopoulou, M. F. Vasiloglou, S. Christodoulidis, Z. Stanga, and S. Mougiakakou, "An artificial intelligence-based system to assess nutrient intake for hospitalised patients," IEEE Trans. Multimedia, vol. 23, pp. 1136–1147, May 2021.

10. Y. Li et al., "A scheme for pest-dense area localization with solar in secticidal lamps internet of things under asymmetric links," IEEE Trans. AgriFood Electron., vol. 1, no. 2, pp. 71–85, Dec. 2023

11. S. J. Lehotay and J. Hajšlová, "Application of gas chromatography in food analysis," TrAC Trends Anal. Chem., vol. 21, no. 9/10, pp. 686–697, 2002

12. . H. M. Merken and G. R. Beecher, "Measurement of food flavonoids by high-performanceliquid chromatography:Areview,"J.Agricultural Food Chem., vol. 48, no. 3, pp. 577–599, 2000.

13. G. Mamone, G. Picariello, S. Caira, F. Addeo, and P. Ferranti, "Analysis of food proteins and 22 peptides by mass spectrometry-based techniques," J. Chromatogr. A, vol. 1216, no. 43, pp. 7130 7142, 2009.

14. N. Taprab and Y. Sameenoi, "Rapid screening of formaldehyde in food usingpaper basedtitration,"AnalyticalChimicaActa,vol.1069, pp. 66–72, 2019.

15. D. Zhang, X.-J. Wu, T. Xu, and H.-F. Yin, "DAH: Discrete asymmetric hashingforefficientcross-mediaretrieval,"IEEETrans.Knowl.DataEng., vol. 35, no. 2, pp. 1365–1378, Feb. 2023.

16. D. Zhang and X.-J. Wu, "Robust and discrete matrix factorization hashing for cross-modal retrieval," Pattern Recognit., vol. 122, 2022, Art. no. 108343 23

17. Z. Wang et al., "Ingredient-guided region discovery and relationship modeling for food category ingredient prediction," IEEE Trans. Image Process., vol. 31, pp. 5214–5226, Aug. 2022.

18. W. Min et al., "Large scale visual food recognition," IEEE Trans. Pattern Anal. Mach. Intel., vol. 45, no. 8, pp. 9932–9949, Aug. 2

19. 1.Afshin et al., "Health effects of dietary risks in 195 countries, 1990 2017: A systematic analysis for the global burden of disease study 2017," Lancet, vol. 393, no. 10184, pp. 1958–1972, 2019.

20. C.-J. Wang, T.-F. Yang, G.-S. Wang, Y.-Y. Zhao, L.-J. Yang, and B.-N. Bi, "Association between dietary patterns and depressive symptoms among middle-aged adults in China in 2016 2017," Psychiatry Res., vol. 260, pp. 123–129, 2018.