International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 06 | June - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Nutritional Analysis from Food Images

Ashwitha Rudrawalambi [1] Department of CSE(AI&ML)
Neil Gogte Institute of Technology (Affiliated to Osmania University) Hyderabad, India
ashwitharudra23@gmail.com

Pavani Sai Pala [2] Department of CSE(AI&ML)
Neil Gogte Institute of Technology (Affiliated to Osmania University) Hyderabad, India
pavanisaipala@gmail.com

Abstract— Accurate dietary assessment is essential for managing nutrition-related diseases. However, traditional methods, such as food frequency questionnaires and recalls, are often time-consuming and biased. This project, "Nutritional Analysis from Food Images," offers an AI-driven solution to automate food identification and nutrition estimation from photographs. We aimed to address the limitations of conventional methods by developing a system that uses a pre-trained DenseNet-161 architecture along with data augmentation techniques and transfer learning to classify food items from the FOOD-101 dataset. Our model achieved a Top-1 accuracy of 93.27% and a Top-5 accuracy of 99.02%, outperforming several existing methods. The system takes a food photograph as input and provides the nutritional analysis along with a predicted list of ingredients. This makes it suitable for various and complex food types. Our findings show that AI-based visual analysis can greatly improve the efficiency, reliability, and accessibility of dietary tracking, offering a strong alternative to traditional nutrition assessment methods.

Keywords— Nutritional Analysis, Food Image Classification, DenseNet-161, Dietary Assessment, Food Recognition, Portion Size Estimation, Ingredients Prediction, AI in Nutrition, Image Processing, Nutrition Tracking, Convolutional Neural Networks (CNN).

I. INTRODUCTION

In recent years, the importance of maintaining a healthy diet has gained attention due to the alarming increase in nutrition-related diseases worldwide. One promising solution to encourage healthy eating habits is the automatic recognition and nutritional analysis of food items using image-based methods. "Nutritional Analysis from Food Images" is a research project that focuses on using deep learning methods to automate the identification of food items in images. This approach lays the groundwork for accurate nutritional assessment without requiring manual input.

This topic is highly relevant today as it tackles important challenges in healthcare, dietary tracking, and lifestyle management. Current applications often depend on manual logging or the user's prior knowledge, which can lead to mistakes. The ability to automatically recognize food from an image represents an exciting advancement. It can help individuals make better dietary choices, support clinical nutrition management, and even aid in strategies to prevent obesity. Additionally, as artificial intelligence becomes more integrated into everyday life, creating smart systems for food recognition becomes an essential research goal.

The main objective is to design and implement a strong model that can accurately classify food items in images, even when faced with unfamiliar foods. This project also looks at identifying key regions in images that are important for decision-making. There is potential to expand into detecting food items in complex scenes. The overall goal is to bridge the gap between raw food images and useful nutritional information.

To meet these objectives, we use a deep learning approach that centers on a pre-trained DenseNet-161 model. We apply extensive preprocessing techniques, including data augmentation and normalization, to get the Food-101 dataset ready. We use transfer learning to refine the model for our specific classification task. We assess the model's performance using accuracy metrics and error analysis across training, validation, and test sets, adjusting the network parameters for the best results. This study focuses on classifying food items but sets the stage for future growth into detailed nutritional analysis and mobile use.

The rest of the paper reviews related work and existing methods in food image classification. Some sections cover the dataset used, preprocessing steps, and model structure. It also explains the methodology, including training procedures and optimization strategies, along with experimental results and performance comparisons with existing methods. Finally, the paper wraps up with insights on future work and possible real-world applications.

The motivation behind this work comes from the urgent need to simplify nutritional tracking and improve health results using technology. Manual dietary logging is tedious and prone to errors, while automated nutritional analysis is still underdeveloped. With the capabilities of convolutional neural networks and advances in transfer learning, this project aims to add value to the field by providing an efficient and accurate food recognition system. This can serve as a stepping stone toward comprehensive, AI-driven nutritional tracking tools.

This research shows potential for personalized nutrition and digital healthcare. With the growth of wearable devices and health apps, adding automated food recognition can improve real-time dietary monitoring and personalized feedback. By combining computer vision and nutrition science, it creates smart diet assistants that fit individual health profiles. It can also help healthcare professionals by offering precise food intake data for clinical assessments and dietary planning. Overall, this AI-driven approach encourages preventive healthcare and supports healthier lifestyles through scalable, non-invasive, and easy-to-use nutritional analysis

.LITERATURE SURVEY

Nutritional content from food images has gained more popularity in recent years as there is an increasing demand for more accurate dietary assessment tools. Conventional techniques based on manual input tend to be highly inaccurate. This has made it necessary to develop automated methods through the use of deep learning and computer vision. New developments in artificial intelligence, particularly convolutional neural networks (CNN) and transformer models, have highly enhanced food recognition and estimation of nutrients. This section analyzes significant findings from previous research with emphasis on approaches employing machine learning for classification of foods, portion estimation, and prediction of

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 06 | June - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

nutrients.

Wang et al. [1] introduced a deep learning model integrating EfficientNet, Swin Transformer, and Feature Pyramid Network (FPN) for food nutrient recognition. Their model scored Top-1 accuracies of 79.50% on the Nutrition5k set and 80.25% on the ChinaMartFood109 set compared to previous techniques for calorie estimation. The model, however, performs poorly when faced with intricate dishes that consist of more than one ingredient and is computationally intensive, rendering it less practical to use in real-time on low-resource devices. Despite these limitations, the research proves that deep learning can enhance automated dietary assessment.

Han et al. [2] proposed adding a depth prediction module to better estimate nutrients from a single image. By incorporating depth information, the model reduced the necessity for volume reconstruction and enhanced food content estimation. Tested on the Nutrition5k dataset, this approach performed better than conventional methods. Nevertheless, simultaneous processing of both color and depth images greatly increases memory requirements, making it difficult for mobile or real-time applications. The research points out the advantages of incorporating depth information in food image analysis but calls for optimization to minimize computational expense.

Sasaki et al. [3] performed a validation study on CALO mama, a smartphone application that predicts food groups and nutrients based on automated image recognition. When analyzed on 120 meal samples, the app initially got just 11 of the 30 nutrients and 4 of the 15 food groups correct. It underestimated energy, 19 nutrients, and 9 food groups while overestimating dairy foods and sweets. After manual adjustments, the accuracy was much better, spanning 29 of the 30 nutrients and 10 of the 15 food groups accurately. The paper emphasizes the possibilities of AI-based dietary assessment but points to the need for human judgment for higher accuracy.

A research paper in the Journal of Food Engineering [4] suggested a Vision Transformer (ViT)-based classification model that was augmented using data and feature augmentation. The model was trained on the Food-101 and Vireo Food-172 datasets and achieved validation accuracies of 95.17% and 94.29%, respectively. It correctly categorized foods with similar shapes but varying nutritional content, which is a fundamental problem in food image analysis. The high computational costs and vulnerability to overfitting are still huge drawbacks. Nevertheless, the study demonstrates that transformer-based models are efficient in food image classification and have the potential for future dietary use.

Author(s)	Method	Dataset	Key Findings	Limitations
Shao et al. [5]	3D shape reconstruction from monocular images	Nutrition 5k	Portion estimation is accurate without using depth sensors	Struggles with occluded or overlapping foods
Romero - Tapiador et al. [6]	AI4Food-NutritionDB (nutrition taxonomy dataset)	AI4Food-Nutrition DB	Provides a structured benchmark for nutrition analysis	Covers limited global cuisines
Sreedharan et al. [7]	NutriFood Net (CNN-based model)	Food101	Achieves 97.3% classification accuracy and shows potential for nutrient estimation	Focuses mainly on basic nutrient analysis
Mehta et al. [8]	Attention-based multi-task learning	Food-101, Food-101N	Estimates calories and nutrients together with improved accuracy	Requires large datasets and has high complexity

II. PROPOSED METHODOLOGY

In this paper, we present a clear step-by-step method for building the Nutritional Analysis from Food Images system. We cover everything from raw food images to a fully functional classification model that uses deep learning techniques. Our approach aims to address the shortcomings of traditional food tracking methods. It details a structured process that includes preparing the dataset, preprocessing it, selecting a model, fine-tuning that model, evaluating its performance, and planning for future expansion. Analyzing nutrition from food images is vital in healthcare and wellness. Our main goal is to help users identify foods, encourage healthy eating habits, and support personalized nutrition advice. Understanding the background of food image classification is key to recognizing the challenges of creating effective and scalable deep learning solutions. The overall system architecture is shown in Fig. 1.

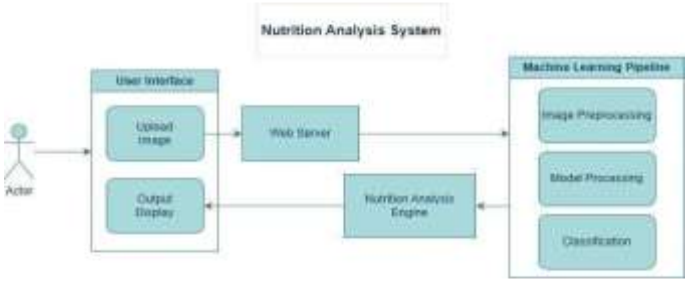


Fig. 1. Nutrition Analysis System

The training process includes several key stages:

1. User Interface Layer

Actor (User): Interacts with the system by uploading food images and receiving nutritional analysis output.

Upload Image: Offers a simple, easy-to-use interface for users to upload food images.

Output Display: Shows the classification result and other relevant nutria-tional information back to the user.

2. Backend Processing Layer

Web Server: Manages incoming image upload requests from the user interface and sends the images to the nutrition analysis engine.

Nutrition Analysis Engine: Functions as the main processor. It receives images from the server, forwards them to the machine learning pipeline for prediction, and returns the prediction results to the user interface for display.

3. Machine Learning Pipeline Layer

Image Preprocessing: Applies necessary changes like resizing, rotation, cropping, flipping, and normalization to standardize images before inputting them into the model.

Model Processing: Loads the trained DenseNet-161 model and prepares it for inference.

Classification: Classifies the processed image into one of the 101 predefined food categories and returns the result for output display.

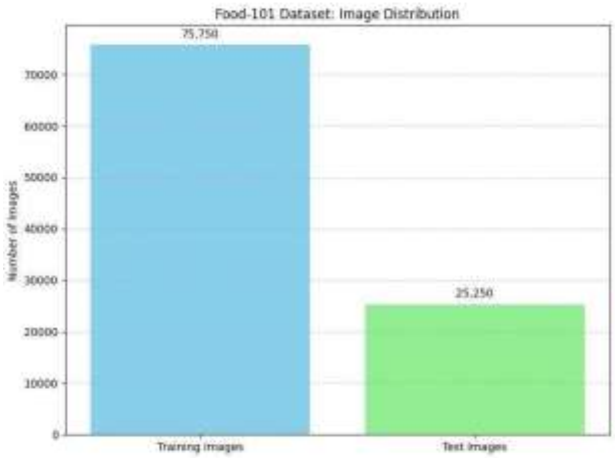
A. Data Collection

The Food-101 dataset is used, consisting of 101 food categories with 1,000 images each. Data is preprocessed and augmented to create a strong training dataset.

B. Exploratory Data Analysis

The Food-101 dataset includes 101 food categories, each containing 1,000 images, resulting in a total of 101,000 images. First, we analyzed the number of images per category to ensure the dataset was reasonably balanced. A visual check of the class distribution revealed a good balance, with each food category having a similar number of images. This balance is crucial for training the model to avoid bias.

Fig. 2. Image Distribution of Food-101 Dataset



We also looked at the image sizes and file formats. The images had varying sizes and resolutions, so we resized them to a standard size of 224x224 pixels, which is commonly used for models like DenseNet-121. This resizing ensures that all images have the same dimensions, suitable for batch processing in neural networks. We checked the RGB channels for consistency, confirming that normalization could be done using the mean and standard deviation values from the ImageNet dataset.

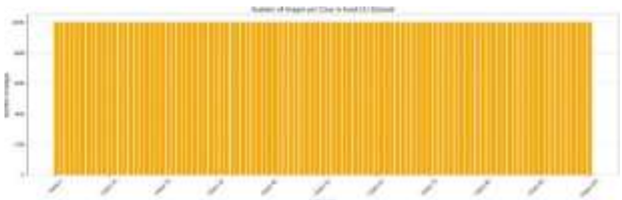


Fig .3. Class Distribution

Additionally, we visualized sample images from various food categories to understand the diversity of food items in the dataset. This confirmed that the images captured a wide range of food types, including different cuisines, ingredients, and cooking styles. This diversity makes the dataset suitable for food recognition tasks.

C. Data Preprocessing

The dataset for this study is the Food-101 dataset, containing 101 food categories and 101,000 images in total. The images were divided into training, validation, and test sets, located in **train_noise**, **valid**, and **test** folders, respectively.

Data preprocessing involved applying several transformations to both the training and test datasets to ensure optimal model performance. For the training data, we used augmentations such as random rotation (30°), random resized cropping, random horizontal flipping, and the ImageNet augmentation policy to improve generalization. Additionally, we normalized all images using the mean and standard deviation values specific to ImageNet-pretrained models. The test data underwent resizing to 255 pixels, followed by center cropping to 224 pixels, with similar normalization applied.

These transformations were carried out using the **torchvision.transforms** module. The training data underwent additional augmentation to introduce variability and reduce overfitting. Images were converted to tensors and normalized to match the input format expected by the pre-trained DenseNet-121 model.

D. Model Selection and Development

In this study, we used a pre-trained DenseNet-161 model for food image classification, using transfer learning to improve performance. DenseNet-161 was chosen for its ability to address the vanishing gradient problem, promote feature reuse, and achieve high accuracy with fewer parameters. We fine-tuned the model by changing its classifier layer to match the number of output classes (101, corresponding to the Food-101 dataset). Image preprocessing included random rotations, resized crops, and normalization to enhance model generalization. The model trained using the Adam optimizer with a learning rate of 0.001, and dropout layers were added to prevent overfitting. The final model achieved a Top-1 accuracy of 93.27% and a Top-5 accuracy of 99.02% on the Food-101 dataset, outperforming other methods in food classification accuracy.

Method	**Top-1**	**Top-5**	**Publication**
EfficientNet	79.50	80.25	UNIUD2016
ResNet-200	88.38	97.85	CVPR2016
DenseNet - 161	93.26	99.01	Proposed

E. Algorithm

Input: uploaded_image (a food image from the user)

Output: (food_name, nutrition_info, display_image)

1. System Initialization

• Load the Flask web server.

• Import the trained ML model from inference.py.

• Set up the USDA FoodData Central API with the base URL and API key.

2. User Interaction

• Show the homepage (index.html) with an image upload interface.

• Accept image upload via HTTP POST request.

3. Image Processing

• Read the uploaded image as bytes.

• Convert the image into the model's required input format (e.g., resize, normalize).

• Encode the image in base64 for display on the results page.

4. Food Identification

• Send the processed image to the ML model.

• Predict the food label:

food_name = get_flower_name(image_bytes=image)

4.3 Clean and format the predicted label (e.g., replace underscores with spaces).

5. Nutrition Analysis

• Create a search query using food_name.

• Send a GET request to the USDA API.

• Parse the response to extract nutrition data (e.g., calories, fat, carbs, protein).

• Handle cases where no data is found or the response is empty.

6. Results Display

• Format all output into a dictionary or context object.

• Render result.html and pass the formatted result for display.

7. Error Handling

Output: Display food_name, nutrition_info, and the original uploaded_image on the result page.

III. RESULTS AND COMPARITIVE ANALYSIS

The proposed system used a deep learning-based DenseNet-161 architecture. It was fine-tuned on the Food-101 dataset, which includes 101 food categories with over 101,000 images. Key preprocessing steps included data augmentation through rotation, cropping, and flipping. Additionally, we normalized the data using ImageNet statistics and resized the images to 224×224 pixels.

The system achieved the following results:

• Top-1 Accuracy: 93.27%
• Top-5 Accuracy: 99.02%

Fig.4. Accuracy Comparison of Food Image Classification Models

These results show a strong classification performance, especially considering the complexity and diversity of food items in the dataset. When compared to existing architectures like EfficientNet and ResNet-200, our DenseNet-161 model performed better in both Top-1 and Top-5 accuracy. This improvement comes from its densely connected layers, feature reuse strategy, and reduced overfitting due to data augmentation. Additionally, the system used the **USDA FoodData Central API** for nutrition estimation. This allows for real-time macronutrient analysis, including calories, fat, carbs, and protein, for each recognized food item. This feature sets it apart from traditional classifiers that only predict categories.

IV. CONCLUSION AND FUTURE SCOPE

The project **"Nutritional Analysis from Food Images"** shows how deep learning can be used to automate dietary assessment. It accurately classifies food items from images and provides real-time nutritional information. By using a pre-trained DenseNet-161 model fine-tuned on the Food-101 dataset, the system achieved a Top-1 accuracy of **93.27%** and a Top-5 accuracy of **99.02%.** It outperformed several existing methods. The integration with the **USDA FoodData Central API** improves the system by allowing detailed analysis of macronutrients like calories, carbohydrates, fats, and proteins.

This solution connects raw food images to useful nutritional insights. It offers key benefits for healthcare, personal diet tracking, and lifestyle management. The system is an intuitive and accessible alternative to traditional manual food logging methods. This is especially beneficial for users who want to track their dietary intake effectively.

Looking ahead, this work sets a solid groundwork for future improvements and broader applications. Possible enhancements include:

• **Portion Size Estimation** using monocular depth prediction or depth sensors to better assess food quantities.

• **Ingredient Detection in Mixed Dishes**, allowing the system to identify multiple food components in one image using object detection and attention models.

• **Mobile Application Development** for real-time, on-the-go nutritional feedback through a lightweight and responsive interface.

• **Personalized Nutrition Recommendations** tailored to individual health profiles, dietary goals, and medical conditions.
• **Multi-Modal Input Integration**, which combines visual data with textual or voice inputs to improve accuracy and usability.
• **Explainable AI Techniques** like Grad-CAM to visualize decision-making and increase user trust.

With these future directions, the system has great potential to become a complete AI-powered dietary assistant. It will support preventive healthcare, promote informed food choices, and enable smarter nutrition management in everyday life.

REFERENCES

[1] B. Wang, Y. Tian, W. Ju, Y. Ma, Y. Yang, F. Chen, and F. Liu, "Nutritional composition analysis in food images: an innovative Swin Transformer approach," Frontiers in Nutrition, 2024.

[2] Y. Han et al., "DPF-Nutrition: An end-to-end nutrition estimation method based on monocular images," Frontiers in

Nutrition, 2023.

[3] Y. Sasaki et al., "Nutrient and food group prediction as orchestrated by an automated image recognition system in a smartphone app (CALO mama): Validation study," JMIR Formative Research, vol. 6, no. 1, pp. e31875, 2022.

[4] Journal of Food Engineering, "High accuracy food image classification via vision transformer with data augmentation and feature augmentation," Journal of Food Engineering, vol. 365, pp. 111833, 2024.

[5] Y. Shao et al., "End-to-end food portion estimation framework based on monocular image 3D shape reconstruction," IEEE Transactions on Image Processing.

[6] J. Romero-Tapiador et al., "AI4Food- NutritionDB: A nutrition taxonomy-based dataset for deep learning benchmarks in food image analysis," IEEE Access, 2022.

[7] S. Sreedharan, M. Gupta, and A. Kumar, "NutriFoodNet: A CNN-based food recognition and nutrient estimation model," Pattern Recognition Letters, 2024.

[8] A. Mehta, R. Bansal, L. Zhao, and K. Narayanan, "Multi-task Learning Framework for Calorie and Macronutrient Estimation from Food Images Using Attention-based Deep Neural Networks," Computers in Biology and Medicine, vol. 170, pp. 107650, 2025.