Convolutional Neural Network Based Food Calorie Estimation System for Dietary Tracking

Conference Paper · November 2024

DOI: 10.1109/NIGERCON62786.2024.10927328

Convolutional Neural Network Based Food Calorie Estimation System for Dietary Tracking

Abiodun OGUNTIMILEHIN
Department of Computing
Afe Babalola University
Ado-Ekiti, Nigeria
abiodunoguns@abuad.edu.ng

Ometie AYIRIORITSE
Department of Computing
Afe Babalola University
Ado-Ekiti, Nigeria
ayirioritseometie@gmail.com

Aondoaseer James ATACHIN
Department of Computing
Afe Babalola University
Ado-Ekiti, Nigeria
atachinjames@abuad.edu.ng

Okebule TOYIN
Department of Computing
Afe Babalola University
Ado-Ekiti, Nigeria
okebulet@abuad.edu.ng

Adebusola M. TOPE-OKE
Department of Computing
Afe Babalola University
Ado-Ekiti, Nigeria
topeokea @abuad.edu.ng

Oluwatoyin Bunmi ABIOLA
Department of Computing
Afe Babalola University
Ado-Ekiti, Nigeria
abiolaob@abuad.edu.ng

Boluwaji Gbenga JAIYESIMI
Department of Computing
Afe Babalola University
Ado-Ekiti, Nigeria
jayesimibg@abuad.edu.ng

Oluwafemi A. SANYA
Department of Computing
Afe Babalola University
Ado-Ekiti, Nigeria
sanyaluwafemi@abuad.edu.ng

Tolani Grace OLUWATOKI
Department of Computing
Afe Babalola University
Ado-Ekiti, Nigeria
oluwatokitg@abuad.edu.ng

Abstract - This study provides a novel approach that enhances the efficiency and accuracy of dietary tracking through automated image recognition technology. This is aimed at promoting the consumption of a healthy diet. This is important due to the fact that if the correct calories are not taken, it could lead to overweight or obesity which causes certain cardiovascular diseases like hypertension, stroke, and coronary artery disease. A dataset named Food 101 containing 101,000 food images was sourced from Kaggle.com and was split into 70% for training, 15% for testing, and 15% for validation. The data went through normalization, resizing, and augmentation for preprocessing. Feature extraction was then carried out using edge detection method. Two trained convolutional neural network models- ResNet50 and google gemini pro vision were used in this work. ResNet50 was used for food type recognition while google gemini pro vision was used to calculate the calories in the food(images). The overall model was deployed as a web application using python programming language. When put to test, the system was able to validate 85% accuracy, ultimately providing a valuable tool for individuals to manage their dietary intake effectively.

Keywords — tracking, diet, artificial intelligence, smartphone, deep learning.

I. Introduction

The emergence of digital health technologies has remarkably revolutionized how individuals track their dietary intake and manage their nutrition, with image-based food calorie estimation standing out as a notable innovation according to Limketkai et al. [1]. Currently, regarding health and lifestyle management, the role of food consumption, individual well-being, and the rising concerns regarding general wellbeing has continued to gain significant attention globally according to Limketkai et al. [1], Puli et al. [2].

According to Tahir and Loo. [3], data from The World Health Organization shows that over 1.9 billion obese adults exist globally with a proportionate and a staggering increase in related health risks. As societies grapple with the repercussions of unhealthy dietary habits and their impact on overall health, there's an intensified focus on understanding, monitoring, and optimizing the right calorie intake to promote better health outcomes.

Dietary tracking has long been a cornerstone of nutrition management and weight control strategies is presented in Ming et al. [4]. Traditional methods have relied heavily on self-reporting, using food diaries or logs to track daily intake [5]. However, these methods are often marred by inaccuracies due to human error or intentional underreporting, leading to less effective dietary monitoring.

The advent of mobile health (mHealth) technologies introduced digital tools for dietary tracking, which have shown promise in improving the accuracy and ease of logging food intake. Despite these advancements, the challenge of simplifying the logging process while maintaining accuracy remains significant. The integration of image recognition technology into dietary tracking represents a pivotal shift towards addressing these challenges.

By allowing users to capture photographs of their meals, these systems can automatically identify food items and estimate their nutritional content, including calorie counts. The potential for image-based calorie estimation tools to transform dietary tracking lies in their ability to reduce the burden on users, making it easier to maintain accurate and consistent logs of food intake.

Moreover, accurate calorie estimation plays a crucial role in public health initiatives aimed at combating obesity and diet-related diseases. By providing individuals with a more accessible and reliable means of tracking their dietary intake, these technologies can support healthier eating habits and contribute to broader efforts to improve nutritional health.

Hence, there is need to develop an image-based food calorie estimation system which will serve as an advisory for individuals about healthy eating so as to avoid health challenges that come with consuming wrong diets.

2024 IEEE 5th International Conference on Electro-Computing Technologies for Humanity (NIGERCON) | 979-8-3315-4255-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/NIGERCON62786.2024.10927328

II. Related Work

According to Puli et al. [2], food plays a key role in the human anatomy, as it affects the general health of an individual. Hence, people are becoming more and more interested in what they eat as unhealthy foods can lead to numerous diseases. In order to avoid eating in an unhealthy manner, it is important to eat what contains the right amount of calories as excessive calories can lead to obesity. In order to avoid being overweight, individuals need to monitor the amount of calories they consume. Hence a machine learning approach is adopted in this study for determining the amount of calories in a meal. In this study, data will be loaded data into system , then processing is done, data splitting is done having 70% for training and 30% for testing. Using CNN, SVM, KNN, Random Forest, MobileNetv2, Inceptionv3 and DenseNet, the accuracy of the models is compared with Random Forest model performing better with mean error of 13.12%.

In Ramkumar et al. [6], the study submits that modern lifestyle practices reveals the significance of maintaining a daily healthy diet and also ensuring a balanced intake of essential nutrients. By estimating nutrient content within meals a lot of critical health issues like diabetes, obesity, and cardiovascular diseases can be addressed. The proposed system in this study functions as a comprehensive food recognition solution, given a satisfactory amount of relevant data; it enables users to meticulously monitor their daily caloric intake. Hence users are able to submit their own images or from their phone cameras for processing. The main goal is not just to identify the food category but also predict its calorie content. Interestingly, the model produces exceptional performance with an accuracy of 0.98%, signifying its capability to correctly classify 98% of instances within the dataset. This study adopts only the use of shape in estimating which is a limitation.

From MV. [7], it is obvious that dietary habits have significant impact on human health and wellbeing. The surge in cases of ailments that manifest as a result of lifestyle practices is on the increase. These ailments include obesity, diabetes, and cardiovascular disease, are linked to dietary patterns. This study presents a novel methodology aimed at addressing this persistent problem by utilizing artificial intelligence, specifically deep learning and image analysis methodologies. In this approach, the dimensions of the food items are computed first while using the known objects present in the image, such as a plate or a fork, as size references. Hence, the size of the known objects is taken as the actual size and is compared with the area of objects from the camera therefore providing a scale for measuring the food in the image. This model uses pre trained models like VGG16, InceptionV3, and ResNet50. The convolutional neural networks (CNNs) produced a probability of 0.9 on a scale of 0 and 1 in recognizing apple and banana. The limitation of this work is its dependence on fruits solely.

According to Lee [8], monitoring the types and amount of foods consumed daily is a very necessary in maintaining a good health. Understanding the nutritional content of foods to be consumed in very helpful in properly treating patients with weight related diseases or cases of obesity. With the current limitations in monitoring food which depends on manual computations, there is the need to develop a food classification system for estimating calorie content in food using convolutional neural networks (CNN). In this study, 10,909 images were used for CNN training. Validation tests were carried out on 3636 images of the food types that were used in training the CNN. This study produced 99.81% accuracy in food classification. The inability of the study is the failure to deploy third system online.

From Al-Saffar and Baiee [9], the study aimed at addressing the lack of accurate nutrition calculations and the increasing prevalence of obesity and related health issues by developing a Nutrition information estimation technique using food photos and machine learning. The method involved a machine learning model for predicting food ingredients and compute essential health metrics for each ingredient, combining them to obtain nutrition data for the food. A pretrained convolutional neural network model for ingredient prediction and nutrition information estimation is used in this work. the dataset used here is the Yelp and Nutrition5k. The study achieves a high accuracy of 85% in predicting food ingredients and extracting nutrition information, surpassing previous research in the field.

According to Haque et al. [10], in Lightweight and Parameter-Optimized Real-Time Food Calorie Estimation from Images Using CNN-Based Approach, the research addresses the challenge of accurately estimating food calories in real time for combating obesity and unhealthy dietary habits. The study involves data set selection, preprocessing, augmentation, and model construction, using the Fruit 360 and Food 101 data sets. The best-performing model, Model 44, achieves an accuracy of 85% and can process images at a rate suitable for real-time applications.

From Deshmukh et al. [11], the study introduces Calorimeter, a machine learning (ML)-based system designed to estimate calories and nutrients in food automatically, addressing the limitations of manual calorie measurement methods. The system uses Machine Learning algorithms like Faster R-CNN for object detection, canny edge detection, and GrabCut segmentation to identify and measure food volume from images, and then estimate caloric content. The system's efficiency is validated with a dataset of 19 types of food, showing satisfactory results when the measurement error is less than 10%.

Also, in Konstantakopoulos et al., [12] the study addressed the increasing interest in smartphone applications for promoting healthy behaviors and the use of computer vision and AI for automatic, precise, and real-time estimation of nutrient content in daily consumed meals. The study highlights the potential of deep learning techniques for food image classification and volume estimation and addresses the challenges and opportunities for future research and development in the field of vision-based dietary assessment systems. As for the results, the review highlights the high accuracy achieved by users familiar with the application in segmenting food images, reaching up to 93% accuracy.

According to Kasyap and Jayapandian [13] the study addressed the challenge of accurately estimating food calorie intake and volume, emphasizing the importance of monitoring calorie intake for maintaining good health. The proposed methodology involved the use of various algorithms such as Convolutional Neural Network (CNN), Random Forest, and Support Vector Machines (SVM) for food recognition and detection. The results show that the proposed CNN model achieved an accuracy of 97% in estimating food calorie intake.

In Jiang et al. [14], the problem addressed is the need for an automatic system to analyze food images and assess dietary habits due to the increasing availability of food-related data from social and mobile networks. The proposed method involved a deep learning-based system for food item detection and nutrition analysis, consisting of three main steps: region proposal, object classification, and dietary assessment.

III. Materials and Methods

The methodology encompasses data collection, image processing techniques, and the deployment of machine learning models for calorie estimation. The architecture of the system is presented in Fig 1.

A. Research and Review

A thorough research and review of related literature was carried out to understand existing methodologies and limitations in this work.

B. Data Collection and description of dataset

A data named the Food 101 dataset was sourced from kaggle.com which was used as the primary source of image data, it consists of 101,000 images categorized into 101 food categories, with each category containing 1,000 images. Nutritional standard metrics was sourced from the USDA National Nutrient Database, which offers comprehensive information on the nutritional content of thousands of food items. This standard metrics includes details on calorie counts, macronutrients (proteins, fats, carbohydrates), and micronutrients (vitamins and minerals), providing a well-rounded nutritional profile for each food item. The process involved mapping each food item identified from the Food 101 dataset images to its corresponding entry in the nutritional standard metrics.

In each of these classes, there are 250 test images that are manually cleaned while 750 images are unclean and are left for training purposes. Hence, there is need for data cleaning to reduce the noise which is basically in the form of intense colors and wrong labels.

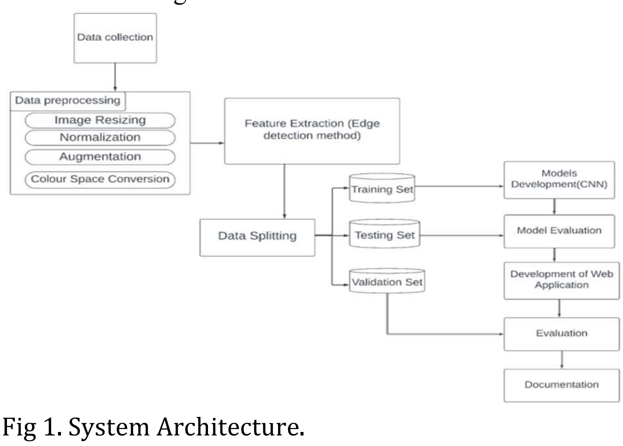

Fig 1. System Architecture.

C. Image Processing Techniques

Image processing plays a crucial role in enhancing the performance and accuracy of the system. This section outlines the methodologies and techniques employed in the preprocessing and feature extraction stages, which are essential for preparing the images for subsequent analysis by machine learning models.

1) Pre-processing

Following data collection, the image dataset underwent several preprocessing steps to improve their quality and ensure consistency which includes resizing using standard dimension, then normalization using a common scale between 0 and 1 followed by augmentation using techniques such as rotation, flipping, scaling, and cropping to expand the training dataset after which colour space conversion was carried out using greyscale.

2) Feature Extraction

Feature extraction was carried out using edge detection method, sobel and canny edge detection. While the sobel edge detection produces thicker edges around the area of interest, canny edge detection produces thinner and precise edges. A combination of these produces a better area of interest with less noise left in the dataset.

3) Data Splitting

This process involved dividing the generated image dataset into separate subsets for training, validation, and testing purposes. The dataset was divided into three subsets: training set, validation set, and test set, split in the ratio 70:15:15, 70% for training, 15% for validation, and 15% for testing.

D. Development of Models for Calorie Estimation

Two trained convolutional neural network models were used, specifically ResNet50 which is used in [7] and google gemini pro vision. Google gemini pro vision, developed using python programming library was used to estimate the calorie content of the food image using standard nutritional metrics. Amount of calories is dependent of food type and food size. Gemini was used to identify foods which addressed the issue of labeling of images in the dataset by concatenating the food classes identified with the images. Then, the ResNet50, was used to train the 70% of the food images for recognition of food types while 15% of the dataset was used for model validation. In Benchaji et al., [12], an artificial intelligent package is used for food identification with accuracy of 97%, hence the use of the gemini pro vision in this study.

The learning rate was adjusted 0.1, 0.01, 0.001, 0.0001, 0.00001 with the best performance being 85% accuracy using learning rate of 0.0001. The result is an advancement of the work in Ramkumar et al. [6] which used a small dataset hence prone to over fitting of the model. Also, in Awoyemi et al. [7], the study adopts the use of a dataset which is predominantly composed of fruits whereas this current study focuses on larger variety of foods.

1) Model Evaluation Criteria

Mean Absolute Error (MAE) in Calorie Estimation: For evaluating the calorie estimation aspect, a MAE of 15.56 is measured which is the average magnitude of errors in calorie predictions compared to true values, offering a clear indication of prediction accuracy. Mean Absolute Error is presented in (1).

$$\mathbf{MAE}=\frac{\sum_{i=1}^{n}|y_i-x_i|}{n} \quad (1)$$

E. Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNNs) are powerful deep learning methods used in image processing, machine learning and computer vision problems. Classification of images, object recognition and digital character recognition are some examples of fields where Convolutional Neural Networks have already been adopted successfully.

In Awoyemi et al. [7], Convolutional Neural Networks (CNNs) achieved an accuracy of 97% hence it is used in this current study. A Convolutional Neural Network simple architecture mainly consists of an input layer, a convolutional layer, a pooling layer and a fully connected layer. The number of layers changes according to the complexity of the studied task. Convolutional Neural Network consist of five layers which are:

1) Input Layer

This layer represents the input data, usually in the form of images but can be other types of data as well. For images, each pixel value might be a node in this layer.

2) Convolutional Layer:

This layer applies uses filters to breakdown the input image. These filters detect features like edges, textures, or patterns within the input image. Convolution involves sliding a filter/kernel over the input data and computing dot products at each position, producing feature maps. It is represented in (2):

$$O_{i,j,k} = \sigma \sum_{l=1}^{C_{in}} \sum_{m=1}^{H_k} \sum_{n=1}^{W_k} I_{i+m-1,j+n-1,l} \cdot K_{m,n,l,k} + B_k \quad (2)$$

3) Pooling Layer

Pooling layers down sample the feature maps, reducing their dimensionality while retaining important information. Max pooling is a common technique, which selects the maximum value from each patch of the feature map. It is represented in (3).

$$Y_{i,j,k} = max_{m,n} X_{S.i+m,S.j+n,k} \quad (3)$$

4) Fully Connected Layer (Dense Layer)

After several convolutional and pooling layers, the network may have one or more fully connected layers. These layers connect every neuron in one layer to every neuron in the next layer, essentially performing classification based on the high-level features learned by the convolutional layers. It is represented in (4).

$$Y = \sigma(W \, . X + B) \quad (4)$$

5) Output Layer

The final layer of the network is responsible for producing the desired output. In classification tasks, this layer often consists of neurons corresponding to each class label, with the softmax activation function applied to produce probabilities for each class. It is represented in (5):

$$Output_i = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}} \quad (5)$$

IV. RESULT

The developed system is shown in Fig. 2 and Fig. 3. This enables the user to capture images of foods for further processing with the results shown in Table I.

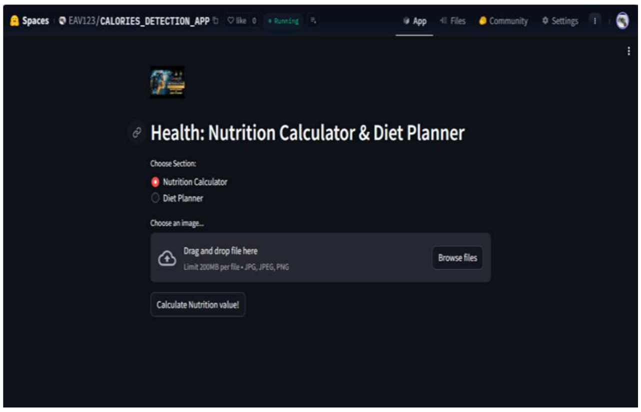

Fig 2. User interface of the web application

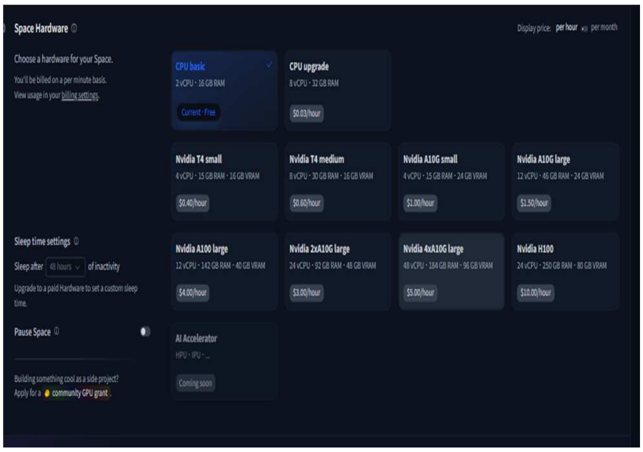

Fig 3. Necessary libraries used in the requirements.txt for smooth running of the system.

TABLE I. RESULT OF NUTRITION CALCULATOR

No.	Food items	Using Colour Features	Using Texture Features	Using Size Features	Using Shape Features	Using All Features
1	Apple	60.33	31.78	81.50	97.47	98.01
2	Orange	65.38	31.22	71.33	89.56	89.45
3	Corn	52.00	48.37	94.85	95.73	93.70
4	Tomato	71.29	55.00	52.91	89.99	76.75
5	Carrot	74.61	28.02	58.55	93.53	89.67
6	Bread	78.65	97.64	69.81	97.47	90.58
7	Pasta	75.43	76.71	56.11	22.59	93.13
8	Sauce	83.45	65.11	44.61	85.00	79.82
9	Chicken	65.43	22.59	94.85	89.99	92.34
10	Egg	70.45	99.79	91.41	98.90	90.19
11	Cheese	75.6	32.35	84.48	89.15	33.46
12	Meat	74.56	22.55	77.53	98.76	91.41
13	Onion	65.45	45.81	97.79	99.57	93.50
14	Bean	71.56	71.33	31.22	34.27	84.48
15	Fish	78.27	48.09	35.20	77.70	93.43
Total Average	63.76	76.28	44.888	45.90	92.21	90.41

V. DISCUSSION

A. **System Requirements**

For this system to operate efficiently, a 2vCPU 4GB RAM, NVIDIA H100 24VCPU 250GB RAM is recommended.

B. User Interface Setup

The interface as shown in Fig. 2 and Fig. 3 enables interaction with the functions of the system. A visit to the site will present the following functions.

1) Title and Logo Display

The application's title, "Health: Nutrition Calculator & Diet Planner," is prominently displayed at the top of the interface using the `st.header()` function. The logo of the application, represented by the "LOGO.jpg" image file, is displayed using the `st.image()` function.

2) Section Selection

Users are presented with a radio button labeled "Choose Section," allowing them to select between "Nutrition Calculator" and "Diet Planner." This interactive element provides a clear indication of the two main functionalities of the application and allows users to choose the section they want to interact with.

3) Nutrition Calculator Section

Upon selecting the "Nutrition Calculator" section, users are prompted to upload an image of food items using the `st.file_uploader()` function. Once an image is uploaded, it is displayed within the interface using the `st.image()` function, allowing users to visually confirm the uploaded content. A prompt template for analyzing nutrition values is provided to guide users on what input is expected. This text is displayed using the `st.subheader()` and `st.write()` functions. A button labeled "Calculate Nutrition value!" is provided for users to trigger the analysis process. The interactive nature of the button encourages user engagement.

4) Diet Planner Section

Upon selecting the "Diet Planner" section, users are presented with a text area where they can input either a list of food items or the desired calorie intake for the day. The input area is clearly labeled with instructions using the `st.text_area()` function, guiding users on how to interact with the application. A button labeled "Plan my Diet!" can be clicked to activate the process of planning a healthy diet. The button serves as a call-to-action, prompting users to take the next step in the interaction flow.

5) Output Display

The results of the analysis or diet planning are displayed below the corresponding sections. The responses generated by the application, such as nutrition analysis or diet plans, are presented using the `st.subheader()` and `st.write()` functions. The output is formatted and displayed in a clear and readable manner, ensuring that users can easily understand and interpret the information provided by the application.

The user interface is designed to be intuitive, interactive, and user-friendly. It guides users through the interaction flow, provides clear instructions and prompts, and presents results in a visually appealing format.

VI. Conclusion and Future work

In conclusion, this study has successfully developed and optimized convolutional neural networks (CNNs) to automate calorie estimation from food images. This approach addresses the limitations of manual input methods by accurately identifying and analyzing food items, considering factors like portion size and composition. The user-friendly application interface enhances usability and adherence to dietary tracking protocols. These findings demonstrate the potential of CNN-based calorie estimation technology to revolutionize dietary management, promoting healthier habits and facilitating personal health goals. Further research is needed to refine make this result adaptable across diverse contexts, which would contribute to an improved public health awareness. This can be done by increasing the dataset and not just the food 101 dataset.

Declaration

The authors declare that there is no conflict of interest.

Acknowledgement

The authors are grateful to Afe Babalola University-Nigeria for supporting this research.

References

[1] B. N. Limketkai, K. Mauldin, N. Manitius, L. Jalilian and B. R. Salonen. "The age of artificial intelligence: use of digital technology in clinical nutrition." Current surgery reports 9, no. 7 (2021): 20.

[2] M. S. Puli, M. S. S. Surarapu, , K. Prajitha, A. Shreshta, G. N. Reddy, and K. V. simha Reddy. “Food Calorie Estimation Using Convolutional Neural Network”. *Journal of Survey in Fisheries Sciences*, p. 2665-2671, 2023.

[3] G. A. Tahir and C. K. Loo. “A comprehensive survey of image-based food recognition and volume estimation methods for dietary assessment”. In *Healthcare* (Vol. 9, No. 12, p. 1676). MDPI. December, 2021.

[4] Z. Y. Ming, J. Chen, Y. Cao, C. Forde, C. W. Ngo and T. S. Chua. “Food photo recognition for dietary tracking: System and experiment”. In *MultiMedia Modeling: 24th International Conference, MMM 2018, Bangkok, Thailand, February 5-7, 2018, Proceedings, Part II 24* (pp. 129-141). Springer International Publishing.

[5] C. K. Suryadevara, “Revolutionizing dietary monitoring: A comprehensive analysis of the innovative mobile app for tracking dietary composition”. *International Journal of Innovations in Engineering Research and Technology*, *10*(88). (2023)

[6] G. Ramkumar, C. B. Venkatramanan, V. Manonmani, S. Palanivel and A.K. Kumar. “A Real-time Food Image Recognition System to Predict the Calories by Using Intelligent Deep Learning Strategy”. In *2023 International Conference on Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering (RMKMATE)* (pp. 1-7). IEEE.

[7] MV D, Prasad. "Deep Learning-based Food Calorie Estimation Method in Dietary Assessment: An Advanced Approach using Convolutional Neural Networks." International Journal of Advanced Computer Science & Applications 15, no. 3 (2024).

[8] K. S. Lee. "Multispectral food classification and caloric estimation using convolutional neural networks." Foods 12, no. 17 (2023): 3212.

[9] M. Al-Saffar and W.R Baiee. "Nutrition information estimation from food photos using machine learning based on multiple datasets." Bulletin of Electrical Engineering and Informatics 11, no. 5 (2022): 2922-2929.

[10] R. U. Haque, R. H. Khan, , A. S. M. Shihavuddin, M. M. Syeed, and M. F. Uddin. "Lightweight and Parameter-Optimized Real-Time food calorie estimation from images using CNN-Based approach." Applied Sciences 12, no. 19 (2022): 9733.

[11] P. B. Deshmukh, V. A. Metre and R. Y. Pawar. "Caloriemeter: food calorie estimation using machine learning." In 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), pp. 418-422. IEEE, 2021.

[12] F. S. Konstantakopoulos, E. I. Georga, and D. I. Fotiadis. "A review of image-based food recognition and volume estimation artificial intelligence systems." IEEE Reviews in Biomedical Engineering 17 (2023): 136-152.

[13] V. B. Kasyap and N. Jayapandian. "Food calorie estimation using convolutional neural network." In 2021 3rd International Conference on Signal Processing and Communication (ICPSC), pp. 666-670. IEEE, 2021.

[14] L. Jiang, B. Qiu, X. Liu, C. Huang and K. Lin. "DeepFood: food image analysis and dietary assessment via deep model." IEEE Access 8 (2020): 47477-47489.