Food detection and volume estimation for nutritional analysis

Komal Dilip Nale[1*], Dr.S.O. Rajankar[2], Prof. V.B. Baru[3], Dr. M.B. Mali[4]

Research Scholar[1] Sinhgad College of Engineering, Vadgaon Bk, Pune

Research Guide[2] Sinhgad College of Engineering, Vadgaon Bk, Pune

Research Co-Guide[3] Sinhgad College of Engineering, Vadgaon Bk, Pune

Head Of Department[4] Sinhgad College of Engineering, Vadgaon Bk, Pune

Mail id: komalnale1@gmail.com[1*]

Abstract

A healthy lifestyle and the prevention of chronic illnesses like diabetes and obesity depend on the precise calculation of dietary nutrition values, including calories, proteins, and carbs. Automated food recognition and volume estimate from photos is a viable solution to the problems with traditional nutritional assessment approaches, which rely on self-reporting and are prone to mistakes. Real-time nutrition estimate smartphone apps have been made possible by recent developments in machine learning (ML) and deep learning (DL). Techniques for food categorization, segmentation, and volume estimation that are used in automated dietary assessment systems are examined in this paper. Traditional machine learning (ML) techniques like support vector machines and k-nearest neighbors were used in early research, but convolutional neural networks (CNNs) and optimization techniques like fuzzy clustering and evolutionary algorithms are used in more recent studies. Additionally, the research emphasizes the difficulties in food segmentation, the use of picture datasets like Food-101, and the incorporation of depth estimate methods for volume measurement. Automated nutritional assessment systems provide increased accuracy and user engagement with the growing usage of AI-driven models, opening the door to better public health nutrition programs. For thorough nutrition estimate, future studies should concentrate on enhancing segmentation methods, diversifying datasets, and combining multi-modal data sources.

Keywords: *Food Detection, Nutritional Analysis, Food Volume Estimation, Machine Learning (ML), Deep Learning (DL), Convolutional Neural Networks (CNNs)*

1. Introduction

Identifying food values such as carbohydrate (CHO), protein, calorie, etc. are essential for a healthy living. In particular, it is crucial for a person (or a patient) to estimate the calorie intake from the food as overindulgence can lead to various lifelong diseases such as obesity, diabetes, heart-disease, etc. Automation of estimating food values from food images would be beneficial in maintaining physical and mental health. Recent development of smart phone-based applications. Has made it possible to deploy an efficient the associate editor coordinating the review of this manuscript and approving it for publication was Amin Zehtabian. real-time automated nutrition estimation framework. The general framework of the food value estimation from food images comprises of identification of the food items in the image, estimation of the volume of the identified food items, and retrieval of the nutritional information of food items, as shown in Figure 1. Moreover, for other smart health applications, food item identification, calorie approximation, etc. from meal images have attracted researchers' attention. The performance of our food value estimation framework depends on the results of the intermediate major steps along with several other factors such as quality and diversity of the food image dataset, relevant information to enhance the performance of the frameworks, etc. We observe considerable research activities in this area. In early research works, most of the studies like have used traditional Machine Learning (ML) methods to calculate the nutritional value from food images. However, from 2014, we have found a shift in utilizing Deep Learning (DL) based frameworks. Recently, the researchers are using optimization methods such as Genetic Algorithm (GA), Fuzzy Clustering for data filtering, Particle Swarm Optimization (PSO), etc. to improve the Deep Learning based frameworks for food classification. In case of food segmentation, which is a pre-processing step of food item identification, we find that researchers are mostly concerned with segmenting single food item from the serving plate

There are only a few review papers related to food value estimation methods from image datasets. have conducted a study on food computation in 2019. In their review, they have included quite a few things including food dataset acquisition, food perception, food recognition, food data retrieval, food recommendation, and prediction and monitoring of social issues. The food datasets include food images, food relevant texts, and multi-modal data of image and text. In their food recognition part, they have discussed only the food classification methods using mostly Machine Learning (ML) based techniques on hand-crafted features of meal images. have presented a literature review on existing food image datasets, food image segmentation, food item classification,

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 02 | Feb - 2025 SJIF Rating: 8.448 ISSN: 2582-3930

and volume estimation. In the food classification part, they describe feature selection, traditional ML techniques, and Deep Learning techniques. Estimating food value using Deep Learning techniques directly from food images has not been covered in their work. they have focused their review on different techniques only for image segmentation task. However, with the improved computational methods, researchers are now involved in segmenting food images from the images of multiple food items. In the volume estimation step, researchers have commonly used reference objects in the images. In recent years, researchers like have computed volume without reference in the images. In this particular work, researchers have used Generative Adversarial Networks (GAN) to map energy distribution in food images. Finally, to estimate the food value, researchers lookup the corresponding nutritional facts from some databases, e.g., US Department of Agriculture (USDA). Recently, in a few studies, the caloric values of food images are crowd sourced. However, this method is highly error-prone. A comprehensive literature review is greatly needed to assist the researchers due to their significant research activities in the area of food value estimation.

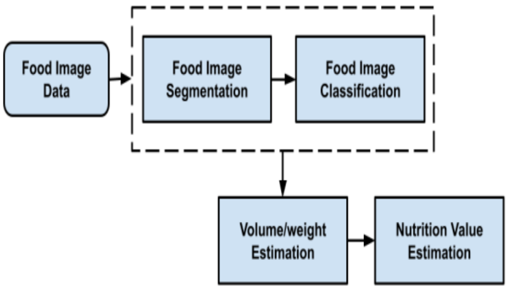

Fig 1. A generalized framework for food nutrition estimation from food images

1.1 Food Image Classification

To detect and identify the food in a meal, a simple automated dietary evaluation system is needed. A collection of unknown objects that correspond to a subset (class) that the classifier has learnt during the training phase are identified using the image classification machine learning approach. In this stage, the classifier is trained using food photographs as input data. Any food type that has been included into the learning process must be recognised by an ideal classifier. The correctness of digital photographs may be impacted by a variety of practical factors, such as rotation, distortion, colour distribution, lighting circumstances, and so on. The training procedure is a laborious undertaking that takes a long time to complete.

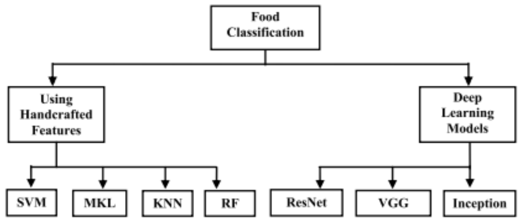

Fig 2. Common classification approaches for food images.

1. Traditional Machine Learning Approaches

In this category, feature extraction is done by hand by looking into the visual characteristics of the food photos, such colour, shape, and texture. A prediction model based on well-known methods, including support vector machines (SVM), K-Nearest Neighbours (KNN), Bag of Features (BoF), Multiple Kernel Learning (MKL), and Random Forests (RF), is then trained using these features. The three successive tasks of segmentation, feature extraction, and classification are essentially carried out using the conventional classification techniques. In order to identify various areas of a picture and then extract the positions of the objects, segmentation is a crucial step. When it comes to food identification, a suitable segmentation strategy should be used to locate food items in the picture and exclude other things such.

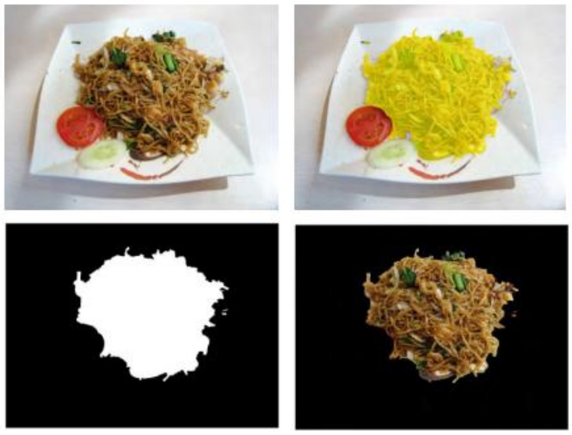

Fig 3. Food image segmentation using GrabCut algorithm.

2. Deep Learning Approaches

A novel method for learning and training a more efficient neural network is deep learning, a subset of machine learning. Deep learning algorithms' built-in process automatically extracts characteristics via a sequence of linked layers, culminating in a completely connected layer that determines the final classification. In comparison to other conventional approaches, it has lately gained popularity because to its slightly better performance with larger datasets, improved processing capabilities, and remarkable classification capacity. One of the most popular methods in deep learning is the Convolutional Neural Network (CNN). introduced it for the purpose of classifying handwritten numbers. CNNs are

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 02 | Feb - 2025 SJIF Rating: 8.448 ISSN: 2582-3930

often used for computer vision applications because of their remarkable capacity to learn operations on visual.

1.2 Food Image Datasets

Training a food image classifier relies on an inclusive collection of food images. An assembled image dataset can be used subsequently to benchmark the recognition performance of other approaches. Several food image datasets have been created for this purpose. It has been a common practice to verify new classifier performance in contrast with the previous methods by training it with a large food image datasets such as Food–101, PFID, UEC Food–100, and UEC Food–256. Existing food image datasets have diverse characteristics, such as food categories, cuisine type, and the total images in the dataset/per food class. For example, PFID has (61) classes of food with a total of 1098 images acquired from fast food restaurants and captured in laboratory conditions. While Food–101dataset contains 101 food classes and a total of 101000 images, 1000 images per food class, captured in three different restaurants. Table 5 summarizes different food datasets with their respective characteristics.

1.3 Food Volume Estimation

Once the food items in a given image have been identified, the volume/weight of the detected food is estimated, so that its corresponding nutrients information, such as sugar, carbohydrates or calories, could be determined. In practice, the process of estimating the total calories without an accurate instrument can be challenging, even to most nutritionists. An image-based calorie assessment must recognize all food regions, segment the food objects in the image, and classify these regions accurately, followed by the calculation of the volume of each segmented item. The nutrient information can be estimated by calculating the actual mass of the food according to the estimated volume (V) and the density of the classified food (d) as in (4), shown at the bottom of this page. The calorie and density information can be acquired as in (5), shown at the bottom of this page, from food nutritional database, such as the USDA Food Composition Database.

Estimating the volume of a food object can be challenging when a single 2-dimensional image is the only source of information, as the case of capturing an image with a smartphone or a handheld camera. These images normally do not contain any additional real-world information such as the scale or the depth of the objects in the scene. To estimate the depth, a synthesized image that contains information relating to the distance of the objects in a scene from the camera is usually generated using special hardware components such as depth sensors or by using stereo vision cameras with known focal length (f) and known baseline length (B) as the distance between the two cameras centers (4). The depth can also be estimated using multiple images from different views with known scene information, such as plates or containers with known size.

Fig 4. A checkerboard reference object is used to estimate the real dimensions of food items.

2. Literature review

Wenyan Jia et.al (2024) Image-assisted dietary assessment has become popular in dietary monitoring studies in recent years. However, food volume estimation is still a challenging problem due to the lack of 3D information in a 2D image and the occlusion of the food by itself or container (e.g., bowl, cup). This study aims to investigate the relationship between the observable surface of food in a bowl and a normalized index (i.e., bowl fullness) to represent its volume. A mathematical model is established for describing different shapes of bowls, and a convenient experimental method is proposed to determine the bowl shape. An image feature called Food Area Ratio (FAR) is used to estimate the volume of food in a bowl based on the relationship between bowl fullness and the FAR calculated from the image.

Fotios S. Konstantakopoulos et.al (2023) The daily healthy diet and balanced intake of essential nutrients play an important role in modern lifestyle. The estimation of a meal's nutrient content is an integral component of significant diseases, such as diabetes, obesity and cardiovascular disease. Lately, there has been an increasing interest towards the development and utilization of smartphone applications with the aim of promoting healthy behaviours. The semi – automatic or automatic, precise and in real-time estimation of the nutrients of daily consumed meals is approached in relevant literature as a computer vision problem using food images which are taken via a user's smartphone. Herein, we present the state-of-the-art on automatic food recognition and food volume estimation methods starting from their basis, i.e., the food image databases. First, by methodically organizing the extracted information from the reviewed studies, this review study enables the comprehensive fair assessment of the methods and techniques applied for segmenting food images, classifying their food content and computing the food volume, associating their results with the characteristics of the used datasets.

Wei Wang et.al (2022) This review presents Vision-Based Dietary Assessment (VBDA) architectures, including multi-stage architecture and end-to-end one. The multi-stage dietary assessment generally consists of three stages: food image analysis, volume estimation and nutrient derivation. The prosperity of deep learning makes VBDA gradually move to an end-to-end implementation, which applies food images to a single network to directly estimate the nutrition. The recently proposed end-to-end methods are also discussed. We further analyze existing dietary assessment datasets, indicating that one large-scale benchmark is urgently needed, and finally highlight critical challenges and future trends for VBDA.

Kalliopi V Dalakleidi et.al (2022) Dietary assessment can be crucial for the overall well-being of humans and, at least in some instances, for the prevention and management of chronic, life-threatening diseases. Recall and manual record-keeping methods for food-intake monitoring are available, but often inaccurate when applied for a long period of time. On the other hand, automatic record-keeping approaches that adopt mobile cameras and computer vision methods seem to simplify the process and can improve current human-centric diet-monitoring methods. Here we present an extended critical literature overview of image-based food-recognition systems (IBFRS) combining a camera of the user's mobile device with computer vision methods and publicly available food datasets (PAFDs). In brief, such systems consist of several phases, such as the segmentation of the food items on the plate, the classification of the food items in a specific food category, and the estimation phase of volume, calories, or nutrients of each food item.

Lameck Mbangula Amugongo et.al (2022) The growing awareness of the influence of "what we eat" on lifestyle and health has led to an increase in the use of embedded food analysis and recognition systems. These solutions aim to effectively monitor daily food consumption, and therefore provide dietary recommendations to enable and support lifestyle changes. Mobile applications, due to their high accessibility, are ideal for real-life food recognition, volume estimation and calorific estimation. In this study, we conducted a systematic review based on articles that proposed mobile computer vision-based solutions for food recognition, volume estimation and calorific estimation. In addition, we assessed the extent to which these applications provide explanations to aid the users to understand the related classification and/or predictions. Our results show that 90.9% of applications do not distinguish between food and non-food. Similarly, only one study that proposed a mobile computer vision-based application for dietary intake attempted to provide explanations of features that contribute towards classification. Mobile computer vision-based applications are attracting a lot of interest in healthcare.

Ghalib Ahmed Tahir et.al (2021) Dietary studies showed that dietary problems such as obesity are associated with other chronic diseases, including hypertension, irregular blood sugar levels, and increased risk of heart attacks. The primary cause of these problems is poor lifestyle choices and unhealthy dietary habits, which are manageable using interactive mHealth apps. However, traditional dietary monitoring systems using manual food logging suffer from imprecision, underreporting, time consumption, and low adherence. Recent dietary monitoring systems tackle these challenges by automatic assessment of dietary intake through machine learning methods. This survey discusses the best-performing methodologies that have been developed so far for automatic food recognition and volume estimation. Firstly, the paper presented the rationale of visual-based methods for food recognition. Then, the core of the study is the presentation, discussion, and evaluation of these methods based on popular food image databases. In this context, this study discusses the mobile applications that are implementing these methods for automatic food logging. Our findings indicate that around 66.7% of surveyed studies use visual features from deep neural networks for food recognition.

Wenyan Jia et.al (2021) As electronic and AI technologies advance rapidly, dietary assessment can now be performed using food images obtained from a smartphone or a wearable device. One of the challenges in this approach is to computationally measure the volume of food in a bowl from an image. This problem has not been studied systematically despite the bowl being the most utilized food container in many parts of the world, especially in Asia and Africa. In this paper, we present a new method to measure the size and shape of a bowl by adhering a paper ruler centrally across the bottom and sides of the bowl and then taking an image. When observed from the image, the distortions in the width of the paper ruler and the spacings between ruler markers completely encode the size and shape of the bowl. A computational algorithm is developed to reconstruct the three-dimensional bowl interior using the observed distortions. Our experiments using nine bowls, colored liquids, and amorphous foods demonstrate high accuracy of our method for food volume estimation involving round bowls as containers. A total of 228 images of amorphous foods were also used in a comparative experiment between our algorithm and an independent human estimator.

Maninder Meenu et.al (2021) This review paper is focused on the quality determination of grains, vegetables, fruits, beverages, meat, sea food and edible oils using Digital Image Processing (DIP). Several studies have reported the successful applications of DIP techniques for feature extraction, classification and quality prediction of foods. DIP algorithms are used to extract the significant features from images which are further used as input for machine learning (ML) algorithms to classify them based on different criteria. These feature extraction methods have been improved by Deep Learning

(DL) algorithms. Features can be automatically extracted by DL algorithms resulting in higher accuracy. DL algorithms require huge data management and computational resources which can be a major limitation.

Frank Po Wen Lo et.al (2020) A daily dietary assessment method named 24-hour dietary recall has commonly been used in nutritional epidemiology studies to capture detailed information of the food eaten by the participants to help understand their dietary behaviour. However, in this self-reporting technique, the food types and the portion size reported highly depends on users' subjective judgement which may lead to a biased and inaccurate dietary analysis result. As a result, a variety of visual-based dietary assessment approaches have been proposed recently. While these methods show promises in tackling issues in nutritional epidemiology studies, several challenges and forthcoming opportunities, as detailed in this study, still exist. This study provides an overview of computing algorithms, mathematical models and methodologies used in the field of image-based dietary assessment. It also provides a comprehensive comparison of the state-of-the-art approaches in food recognition and volume/weight estimation in terms of their processing speed, model accuracy, efficiency and constraints.

Wesley Tay et.al (2020) Obesity is a global health problem with wide-reaching economic and social implications. Nutrition surveillance systems are essential to understanding and addressing poor dietary practices. However, diets are incredibly diverse across populations and an accurate diagnosis of individualized nutritional issues is challenging. Current tools used in dietary assessment are cumbersome for users, and are only able to provide approximations of dietary information. Given the need for technological innovation, this paper reviews various novel digital methods for food volume estimation and explores the potential for adopting such technology in the Southeast Asian context. We discuss the current approaches to dietary assessment, as well as the potential opportunities that digital health can offer to the field. Recent advances in optics, computer vision and deep learning show promise in advancing the field of quantitative dietary assessment.

Prachi Kadam et.al (2020) This study represents an overview of the research in the food volume estimation methods using deep learning and 3D reconstruction methods using works from 1995 to 2020. The findings present the five different popular methods which have been used in the image based food volume estimation and also shows the research trends with the emerging 3D reconstruction and deep learning methodologies. Additionally, the work emphasizes the challenges in the use of these approaches and need of developing more diverse, benchmark image data sets for food volume estimation including raw food, cooked food in all states and served with different containers. image-based methods to estimate food volume in dietary management systems and projects. Deep learning and 3D reconstruction methods show better accuracy in the estimations over other approaches. The work also discusses importance of diverse and robust image datasets for training accurate learning models in food volume estimation.

Mohammed Ahmed Subhi et.al (2019) Consuming the proper amount and right type of food have been the concern of many dieticians and healthcare conventions. In addition to physical activity and exercises, maintaining a healthy diet is necessary to avoid obesity and other health-related issues, such as diabetes, stroke, and many cardiovascular diseases. Recent advancements in machine learning applications and technologies have made it possible to develop automatic or semi-automatic dietary assessment solutions, which is a more convenient approach to monitor daily food intake and control eating habits. These solutions aim to address the issues found in the traditional dietary monitoring systems that suffer from imprecision, underreporting, time consumption, and low adherence. In this paper, the recent vision-based approaches and techniques have been widely explored to outline the current approaches and methodologies used for automatic dietary assessment, their performances, feasibility, and unaddressed challenges and issues.

Hamid Hassannejad et.al (2017) A balanced diet is the key to a healthy lifestyle and is crucial for preventing or dealing with many chronic diseases such as diabetes and obesity. Therefore, monitoring diet can be an effective way of improving people's health. However, manual reporting of food intake has been shown to be inaccurate and often impractical. This paper presents a new approach to food intake quantity estimation using image-based modeling. The modeling method consists of three steps: firstly, a short video of the food is taken by the user's smartphone. From such a video, six frames are selected based on the pictures' viewpoints as determined by the smartphone's orientation sensors. Secondly, the user marks one of the frames to seed an interactive segmentation algorithm. Segmentation is based on a Gaussian Mixture Model alongside the graph-cut algorithm. Finally, a customized image-based modeling algorithm generates a point-cloud to model the food. At the same time, a stochastic object-detection method locates a checkerboard used as size/ground reference. The modeling algorithm is optimized such that the use of six input images still results in an acceptable computation cost. In our evaluation procedure, we achieved an average accuracy of 92%92% on a test set that includes images of different kinds of pasta and bread, with an average processing time of about 23 s.

C. J. Boushey et.al (2016) When using static pictures, each participant sees the same image and is required to relate it to a personal experience in order to estimate the amount of food ingested or to evoke a culinary memory. Images now play a dynamic role instead of a static one due to mobile technology. The distinctive photos that are produced when someone records their personal dining experiences provide a genuine account of the items they have really consumed. A dynamic compilation of pertinent information is produced by these pictures. In addition to serving as a recollection of the items eaten and the circumstances surrounding an eating occasion, such the time of intake, a distinctive, more dynamic picture might serve as a reference for portion size. When it comes to nutritional evaluation, a dynamic image's usefulness probably surpasses.

Expected Outcomes

Several noteworthy results are anticipated from the use of image-assisted nutritional assessment technology into public health nutrition efforts. Above all, these technologies should increase the precision of nutritional data gathering. The systems may reduce the mistakes often seen in conventional self-reporting techniques by using sophisticated picture recognition algorithms, which will provide more accurate information on food patterns at the individual and group levels. Through mobile apps with user-friendly interfaces, these technologies may increase user engagement in addition to improving accuracy. Because it is simpler for users to track their food consumption and get real-time feedback, this accessibility may promote greater adherence to dietary requirements. Including interactive features and gamification might encourage consumers to adopt healthy eating habits even more. The capacity to evaluate enormous datasets produced by these technologies, enabling tailored nutrition treatments, is another expected result. Dietary guidelines that are specific to each person's eating habits may improve the efficacy of public health initiatives that target obesity and chronic illnesses. In order to meet the distinct nutritional requirements of different populations, this individualised approach is essential.

Conclusion

In conclusion, a major step forward in encouraging good eating habits is the use of image-assisted nutritional evaluation tools. These techniques efficiently overcome the shortcomings of conventional nutritional assessment methods, which often depend on self-reporting and recollection biases, by using computer vision and machine learning to improve the accuracy of food identification and portion size estimates. Analysing these methods has shown how well they work to increase user engagement and dietary compliance. These technologies are used in mobile apps that provide an intuitive interface for tracking food consumption. Better dietary choices are also encouraged by the possibility of real-time feedback, which is crucial for addressing nutrition-related public health issues.

References

[1]. Jia, W., Zhang, X., Li, Y., Wu, H., & Wang, J. (2024). Image-assisted dietary assessment: A mathematical model for bowl fullness estimation. Journal of Food Engineering, 292, 110578. https://doi.org/10.1016/j.jfoodeng.2024.110578

[2]. Konstantakopoulos, F. S., Koutsou, A., & Sarri, A. (2023). A state-of-the-art review on automatic food recognition and volume estimation methods. Nutrition Reviews, 81(3), 243-257. https://doi.org/10.1093/nutrit/nuz108

[3]. Wang, W., Xu, Y., Liu, J., & Chen, Y. (2022). Vision-Based Dietary Assessment architectures and future trends. IEEE Transactions on Image Processing, 31, 1453-1466. https://doi.org/10.1109/TIP.2022.3158604

[4]. Dalakleidi, K. V., Tsapakis, I. S., & Pavlidis, G. A. (2022). Image-based food-recognition systems and their potential applications. Journal of Food Science, 87(10), 4333-4342. https://doi.org/10.1111/1750-3841.16620

[5]. Amugongo, L. M., Ngwenya, B. T., & Rukanga, F. (2022). Mobile computer vision-based solutions for food recognition. Journal of Artificial Intelligence in Agriculture, 6, 1-10. https://doi.org/10.1016/j.aiia.2022.04.003

[6]. Tahir, G. A., Khalil, K. A., & Iqbal, J. (2021). Survey of automatic food recognition methodologies for dietary assessment. Journal of Food Quality, 2021, Article ID 3659832. https://doi.org/10.1155/2021/3659832

[7]. Jia, W., Zhang, X., Li, Y., & Wang, J. (2021). A novel method for food volume estimation using bowl shape analysis. Sensors, 21(1), 150. https://doi.org/10.3390/s21010150

[8]. Meenu, M., Kaur, G., & Gupta, R. (2021). Quality determination of foods using Digital Image Processing techniques. Food Quality and Safety, 5(1), 1-11. https://doi.org/10.1093/fqsafe/fyaa024

[9]. Lo, F. P. W., Shiu, S. Y., & Lee, K. H. (2020). Visual-based dietary assessment and its challenges in nutritional epidemiology studies. Nutrition Journal, 19, Article 84. https://doi.org/10.1186/s12937-020-00638-4

[10]. Tay, W., Ang, W., & Chan, W. (2020). Novel digital methods for food volume estimation in dietary assessment. Food Research International, 132, 109066. https://doi.org/10.1016/j.foodres.2020.109066

[11]. Kadam, P., Dandge, N., & Prabhu, K. (2020). Overview of food volume estimation methods using deep learning. Machine Learning and Applications, 3, 100049. https://doi.org/10.1016/j.mlwa.2020.100049

[12]. Subhi, M. A., Ahmed, A. A., & Ibraheem, A. (2019). Automatic dietary assessment solutions using machine learning technologies. Journal of Food Science and Technology, 56(9), 3964-3972. https://doi.org/10.1007/s11483-019-01410-6

[13]. Hassannejad, H., Zohrevand, H., & Mohammadi, R. (2017). Image-based modeling for food intake quantity estimation. Computers and Electronics in Agriculture, 139, 141-147. https://doi.org/10.1016/j.compag.2017.05.020

[14]. Boushey, C. J., Delp, E. J., & Parikh, D. (2016). Dynamic role of images in dietary assessment: A comprehensive review. Public Health Nutrition, 19(10), 1801-1813. https://doi.org/10.1017/S1368980016000800