

Multimedia Data-Based Mobile Applications for Dietary Assessment

Journal of Diabetes Science and Technology
2023, Vol. 17(4) 1056–1065
© 2022 Diabetes Technology Society
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/19322968221085026
journals.sagepub.com/home/dst


Maria F. Vasiloglou, PhD^{1*} , Isabel Marcano, MD^{2*},
Sergio Lizama, MD², Ioannis Papathanail, MSc¹,
Elias K. Spanakis, MD^{2,3} , and Stavroula Mouggiakakou, PhD^{1,4} 

Abstract

Diabetes mellitus (DM) and obesity are chronic medical conditions associated with significant morbidity and mortality. Accurate macronutrient and energy estimation could be beneficial in attempts to manage DM and obesity, leading to improved glycemic control and weight reduction, respectively. Existing dietary assessment methods are subject to major errors in measurement, are time consuming, are costly, and do not provide real-time feedback. The increasing adoption of smartphones and artificial intelligence, along with the advances in algorithms and hardware, allowed the development of technologies executed in smartphones that use food/beverage multimedia data as an input, and output information about the nutrient content in almost real time. Scope of this review was to explore the various image-based and video-based systems designed for dietary assessment. We identified 22 different systems and divided these into three categories on the basis of their setting for evaluation: laboratory (12), preclinical (7), and clinical (3). The major findings of the review are that there is still a number of open research questions and technical challenges to be addressed and end users—including health care professionals and patients—need to be involved in the design and development of such innovative solutions. Last, there is a clear need that these systems should be validated under unconstrained real-life conditions and that they should be compared with conventional methods for dietary assessment.

Keywords

AI, apps, dietary assessment, mHealth, nutrition, smartphones

Introduction

More than 34.2 million patients live with diabetes mellitus (DM).¹ Individuals with DM use carbohydrate (CHO) counting techniques to estimate the appropriate prandial dose of insulin. Prior studies have shown that insulin-treated patients with DM experience difficulties in estimating CHO^{2,3} and their accuracy may be as low as 59%.⁴ Therefore, there is a need to provide patients with resources to improve their accuracy in counting CHO. Moreover, the prevalence of obesity has exponentially risen to 39.6% over five years⁵ which has been the driving force behind the rapid emerge of smartphone applications (apps).⁶ Accurate estimation of CHO and other nutrients could be beneficial in attempts to manage DM and obesity, leading to improved glycemic control and weight reduction, respectively, as well as preventing patient malnutrition in hospitals and geriatric clinics.⁷ There are various standardized methods of measuring dietary intake, such as 24-hour food recall and food frequency questionnaires.⁸ However, existing methods are time consuming, costly, lack precision,⁹ and have difficulties with estimating portion sizes.¹⁰

For these reasons, new methods have been introduced to document nutritional intake based on food photos or videos acquired using smartphone cameras and can be divided according to the technological features that they incorporate. The first group are the apps where the user needs to manually insert the type and portion size of the food item or beverage after capturing its image. In the second group, the user takes a photo/video of a meal, and a dietitian analyses the data and

¹ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland

²Division of Endocrinology, Diabetes, and Nutrition, University of Maryland School of Medicine, Baltimore, MD, USA

³Division of Endocrinology, Baltimore Veterans Affairs Medical Center, Baltimore, MD, USA

⁴Department of Emergency Medicine, Bern University Hospital, Bern, Switzerland

*These authors contributed equally to this work.

Corresponding Author:

Stavroula Mouggiakakou, PhD, ARTORG Center for Biomedical Engineering Research, University of Bern, Murtenstrasse 50, Bern 3008, Switzerland.

Email: stavroula.mouggiakakou@unibe.ch

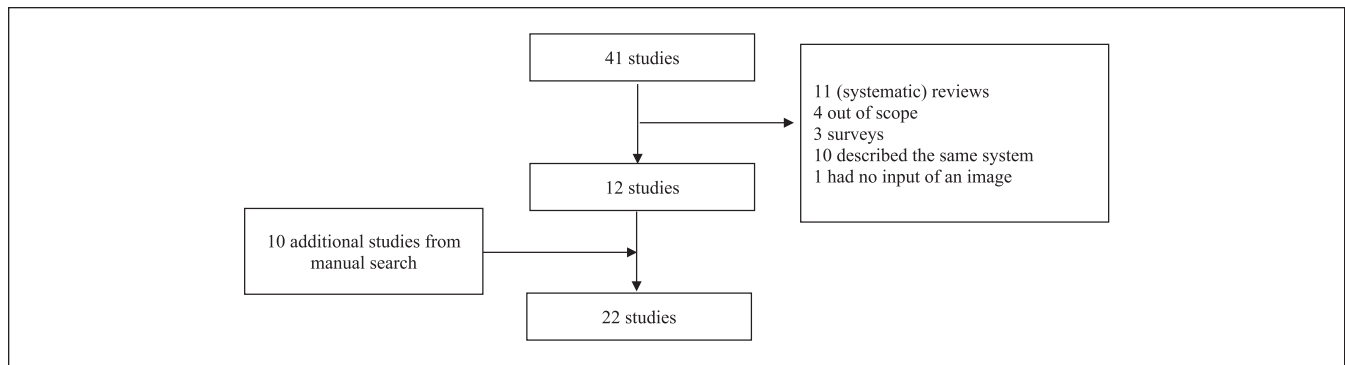


Figure 1. Flowchart of the study selection process.

returns the output to the user. The third group are apps that incorporate some degree of automation, such as automatic food identification of a captured photo. The recent advances in the field of artificial intelligence (AI) and smartphone technologies allowed the development of apps that are completely automatic. Specifically, the user takes a photo/video of a meal, and the app automatically detects and identifies the various food items, estimates the volume, and, by using information from food composition databases, outputs the nutrient values.¹¹

In this review article, we present the available systems that use image/video as input and employ AI-based features to estimate nutritional information for food/beverages.

Methodology

PubMed was used to identify suitable studies conducted during the period January 2010 to January 2022 and published in English. The following queries were used: (((Image) OR (video)) AND (((dietary assessment) OR (diet record)) OR (dietary monitoring)) AND (((smartphone) OR (mobile phone)) OR (mobile telephone)) OR (mobile device)) OR (mHealth))) AND (((((AI) OR (artificial intelligence)) OR (machine learning)) OR (computer vision)) OR (augmented reality)) OR (food recognition)) OR (volume estimation)) and 41 articles were found (Figure 1).

Results

The systems were divided into three categories based on the environment for which they were mainly evaluated: laboratory, preclinical, and clinical evaluation. In Table 1, we present some of the published software apps that have been developed based on image/video analysis and focus on their performance and a few of their most common characteristics. Clinical setting studies include complete systems, that is, from image/video capture to estimation of nutrient content and have been tested in the framework of a trial. Preclinical evaluation studies include prototypes evaluated within the framework of feasibility studies and involve a small number

of participants ($n < 10$). Studies in the laboratory settings mostly include approaches that are evaluated in terms of technical performance and focus on the automatic or semiautomatic features for food recognition and/or volume estimation. In the following sections, we present in chronological order the systems that exist, starting with those tested in a clinical setting, then those tested in the preclinical setting, and finally those tested in a laboratory setting.

Studies Conducted in a Clinical Setting

The systems described below are considered as first generation, in the sense that they mainly use classical AI approaches along with image processing methods aiming at semi-automation or full automatization of the dietary assessment pipeline. They provide significant algorithmic approaches that have facilitated research on more accurate and computationally efficient approaches that are closer to the end users' needs.

Three studies that are presented below use either semiautomatic or fully automatic systems. Two of the systems use semiautomatic apps^{16-19,40} and one uses an automatic app.^{12,13,15,41} One of the apps incorporated a semiautomatic feature for the food recognition part and asked the user to manually insert information on portion size.¹⁶ One app asks the user to confirm food type and portion size,^{17,19} while one app gives the user the possibility of correcting food classification and food segmentation once the respective results have been provided.¹²

GoCARB is an Android app for people living with DMI to help them estimate the CHO content of plated meals. The user takes two food images of the dish before eating, while a credit-card-sized fiducial marker is placed next to the dish. The food is automatically detected and recognized, the respective volume is estimated,⁴¹ and the CHO content is calculated using the United States Department of Agriculture (USDA) food composition database. GoCARB has been studied in a variety of settings. In the preclinical setting, the repeatability of the system's results and the consistency of the modules were tested by evaluating 24 dishes, achieving

Table 1. Published Research Works on Applications for Dietary Assessment Based on Image/Video Analysis.

Study	Input	Intelligent features	Output	Main results
Clinical setting Anthimopoulos et al. ¹² Rhyner et al. ¹³ Bally et al. ¹⁴ Vasiloglou et al. ¹⁵	Photos	Automatic food recognition, segmentation, volume estimation, CHO calculation	Carbohydrates	Preclinical testing of 24 dishes: mean relative error of CHO estimation 10 ± 12% (6 ± 8 g). Clinical setting GoCARB vs six experienced dietitians, CHO intake of 54 plates: no difference in mean absolute error (14.9 vs 14.8 g, P = .93). Clinical testing in DMI subjects (n = 19): GoCARB had better mean absolute error than self-performance (12.28 vs 27.89 g, P = .001). Clinical testing in DMI on insulin pumps + Continuous Glucose Monitoring (n = 20): GoCARB compared with conventional method had less time spent in postprandial hyperglycemia (P = .039) and less glycemic variability (P = .007). Significant differences between Keenoa-participant and Keenoa-dietitian data for 12 of the 22 nutrients analyzed. Participant reports of dietary fat and protein (% difference: +31.9% and +6.9%, respectively) were lower but for carbohydrates was higher (% difference: -24.7%), compared with the edited version by the dietitian. Clinical testing in adolescents aged 11 to 18 years (n = 78). Feasibility study in which the participants agreed that the app was easy to use. With 10% training data, app reported within 10% margin of correct nutrient information; with 50% training data, margin improved to 1%. Clinical testing of healthy individuals (n = 45), mFR compared with doubly labeled water for estimation of energy expenditure correlated significantly (Spearman correlation coefficient 0.58, P < .0001).
Ji et al. ¹⁶	Photos	Semiautomatic food recognition, manual food volume estimation, automatic nutrient calculation	Macronutrients and micronutrients	
Zhu et al. ¹⁷ Six et al. ¹⁸ Boushey et al. ¹⁹	Photos	Automatic food recognition, segmentation, volume estimation, CHO calculation	Calories	
Preclinical setting Domhardt et al. ²⁰	Photos	Manual food recognition, automatic food volume estimation, automatic nutrient calculation	Carbohydrates	DMI (n = 8) patients eating in a laboratory: compared with visual estimation of CHO, the app decreased absolute error in 11/18 estimations. In 8/18 estimations, the absolute error decreased by at least 6 g of CHO. Four estimations showed a worse absolute error of grams of CHO.
Kong and Tan ²¹	Photos/video	Automatic food recognition, segmentation, volume estimation, CHO calculation	Calories	Accuracy in food recognition reported to be 92%, based on 20 different food types.
Lu et al. ²²	Photos/video	Automatic food recognition, segmentation, volume estimation, nutrient calculation	Energy and macronutrients	When compared with ground truth, goFOOD outperforms the estimations of experienced dietitians in the European meals (goFOOD; CHO (g): 7.2 (3.2-15.3), protein (g): 4.5 (2.0-10.9), fat (g) 5.2 (2.0-10.06), calories (kcal): 74.9 (40.4-139.3), dietitians; CHO (g): 27 (10.6-37.7), protein (g): 8.7 (4.7-13.5), fat (g): 5.2 (2.3-9.7), calories (kcal): 180 (119-271)). Comparable performance was achieved on standardized fast-food meals (goFOOD; CHO (g): 7.9 (4.2-15.7), protein (g): 2.8 (1.3-4.2), fat (g): 5.8 (1.4-14.6), calories (kcal): 75.9 (27.9-124.7), dietitians; CHO (g): 5.3 (2.9-7), protein (g): 1.5 (0.5-3.3), fat (g): 3.8 (1.5-6.3), calories (kcal): 55.5 (17-83)).
Makhous et al. ²³	Video	Manual food recognition, automatic food volume estimation, automatic nutrient calculation	Not specified	The results showed improved volume estimation by almost 40% when compared with manual calculations and also a higher error of almost 15% for calculated calories when compared with ground truth.
Sasaki et al. ²⁴	Photos	Automatic food recognition, portion size estimation, automatic nutrient calculation	Calories, macronutrients and micronutrients	The automatic results were similar to the ground truth for the weight, 11/30 nutrients, and 4/15 food groups. The manually adjusted results were similar to the ground truth for the energy, 29/30 nutrients, and 10/15 food groups.

(continued)

Table 1. (continued)

Study	Input	Intelligent features	Output	Main results
Herzig et al ²⁵	Photos	Automatic food recognition, segmentation, volume estimation, nutrient calculation	Calories, macronutrients	For all meals, the absolute error (SD) for the system: for CHO (g) 5.5 (5.1), for protein (g) 2.4 (5.6), for fat (g) 1.3 (1.7), for calories (kcal) 41.2 (42.5). Breakfast: for CHO (g) 7.1 (5.5), for protein (g) 1.0 (1.1), for fat (g) 1.2 (1.3), for calories (kcal) 40.4 (30.5). Snacks: for CHO (g) 2.9 (3.7), for protein (g) 0.5 (0.8), for fat (g) 1.1 (1.6), for calories (kcal) 24.1 (36.9). Cooked meals: for CHO (g) 6.5 (5.2), for protein (g) 5.6 (8.9), for fat (g) 1.6 (2.3), for calories (kcal) 59.1 (58.0). The relative weight error for nine items was <10%. The relative CHO error was <10% for all items except nigiri.
Laboratory setting Huang et al ²⁷	Photos	Automatic food recognition, segmentation, volume estimation, CHO calculation	Carbohydrates	A small set of fruit images taken under real living conditions were used ($n = 6$, three different fruits: three apples, one peach, and two tomatoes) and compared with ground truth. The overall accuracy of the SVM-based classification was 90% and the average error in volume estimation was 6.86%. For CHO estimation, the average error was 8.18%. The classification performance for the combination of four different features was between 66.9% and 93.2%. The average volume estimation error was 12.01% for the plate method and 29.01% for the elliptical spotlight method.
Farinella et al ²⁸	Photos	Automatic food recognition, volume estimation, CHO estimation	Carbohydrates	The experimental results showed a classification rate close to 80% (79.2) for the first five food categories from the 100-category food data set.
Jia et al ²⁹	Photos	Automatic food recognition	Out of scope	The recognition accuracy on the data set was 86.72% and the detection model achieved an accuracy of 94.47%. On the other hand, the real-life tests with a data set of self-recorded images from people living with Parkinson's, all of which were taken with a smartphone camera, achieved a maximum accuracy of 55%. Their segmentation networks achieved 92.18% pixel accuracy, 59.2% average precision, and 82.1% average recall on a data set with images that contain multiple food/beverage items.
Kawano and Yanai ³⁰	Photos	Automatic food recognition	Calories	The volume efficiency was tested with only eight food objects and the percentage of error in the calculated volume ranged from 3.3% for a banana to 9.4% for a pudding box.
Mezgec et al, ³¹ Mezgec et al ³²	Photos	Automatic food recognition	No information	Recognition accuracy is 91.3% in Korean foods in comparison to other models.
Lo et al ³³	Photos	Automatic food volume estimation	Out of scope	Recognition accuracy in experimental condition is 0.47.
Park et al ³⁴	Photos	Automatic food recognition	Out of scope	The food recognition performance using deep learning was 87.2%. The results showed that there was an average absolute error of <20% in large-volume items, but a >20% error in small-volume items.
Probst et al ³⁵	Photos	Automatic food recognition	Macronutrients and energy	85% classification accuracy was achieved detecting 15 different kinds of foods.
E Silva et al ³⁶	Photos	Semiautomatic food image recognition	Out of scope	With vs without depth refinement: segmentation accuracy: 81.9% vs 93.4%, intersection over union: 87.9% vs 86.7%, volume estimation errors (mL): 11.0 vs 10.8, % intake error: -4.2% vs -5.3%.
Yang et al ³⁷	Photos	Automatic food volume estimation	Out of scope	
Zhang et al ³⁸	Photos	Automatic food recognition, segmentation, portion size estimation	Energy and nutrient	
Pfisterer et al ³⁹	Photos	Automatic food segmentation, volume estimation	Out of scope	

Abbreviations: CHO, carbohydrate; mFR, Mobile Food Record; DM1, diabetes mellitus 1; SVM, support vector machine; SD, standard deviation.

low errors.¹² Another study found that the GoCARB system gave lower mean absolute error for CHO estimation than self-estimations by participants living with DM1.¹³ GoCARB has also been compared with dietitians with respect to the accuracy of CHO estimation¹⁵ and achieved similar results. Last, the use of GoCARB utility was evaluated in the clinical setting in a randomized controlled crossover study versus standard methods of counting CHO. Postprandial glucose control was studied in 20 adults with DM1, using sensor-augmented insulin pump therapy.¹⁴

Mobile Food Record (mFR) is an iOS app by which the user takes a picture before and after meal consumption to estimate food intake.¹⁷ From the meal images, the food is segmented, and food type is identified based on textural and color features along with a support vector machine (SVM). To estimate food volume, a fiducial marker of known dimensions must be placed next to the food. In this app, the results are sent back to the user who confirms and/or adjusts the information with regard to food identification and portion size and sends it back to the server. The data are indexed with a nutrient database, the Food and Nutrient Database for Dietary Surveys, to retrospectively calculate the content of the foods consumed. The mFR's accuracy in collecting reported energy intake has been demonstrated by comparing it with the doubly labeled water (DLW) technique.¹⁹ The participants were asked to capture photos of their food before and after eating for 7.5 days. The energy estimates using mFR correlated significantly with energy expenditure using DLW. The system was also tested in 78 adolescents taking part in a summer camp.¹⁸ The energy intake measured from the known food items for each meal was used to validate the performance of the system.¹⁷

Keenoa¹⁶ is a smartphone image-based dietary assessment app where the user takes a photo of a single food item, and the app automatically presents a number of food options to the user to select the correct food type. The user estimates and manually enters the serving size using visual aids (eg, cups) and selecting the unit (eg, weight, volume). Nutrient estimations are produced automatically, as the app is linked to the Canadian Nutrient File (2015) database. Keenoa app is only accessible to registered dietitians in Canada so that they can accurately generate individuals' dietary intake. A study in 72 healthy adults has assessed the relative validity of Keenoa app compared with a three-day food diary and the estimations of dietitians.¹⁶ The results reflected significant differences between the output of the Keenoa app and the analysis of the dietitians. The majority of nutrients consumed by the Keenoa-participants were under-recorded compared with the results of the Keenoa dietitian. In a qualitative analysis of the app's evaluation, participants ($n = 50$) pointed out positive characteristics of the app including food recognition and easier data collection than with paper and pencil methods. However, the barcode scanner, the limited food database, and the time needed to enter food items were challenging for them.⁴⁰

Studies Conducted in a Preclinical Setting

These systems belong to new-generation apps which are still under development. The researchers developed the results of the first-generation systems and employed various enhancements to improve them. Along with the availability of powerful hardware and the recent advances in sensor technologies, the smartphones extended their capabilities and, for example, made it possible to bypass the fiducial marker requirement, together with the synergistic use of information from smartphone's embedded sensors beyond cameras.

Seven studies have been identified that were conducted in a preclinical setting. Two of these were tested in a small group of participants who were asked to estimate portion size or nutrient values.^{20,23} Five systems only used real foods that were weighed during the study.^{21,22,24-26}

BE_{ar} is a semiautomatic Android app designed to estimate CHO. The user identifies food on their plate from a list and points the camera to the food plate with a fiducial marker in front of it.²⁰ A red circle marks the food drawing area in three dimensions and the user can accept or redraw the suggested shape. The app then measures the food volume, which is converted to CHO estimation using the USDA database. BE_{ar} app was integrated in a diabetes diary app that allowed users to keep track of food intake, blood glucose levels, insulin injections, and physical activity which are then accessible to health care providers. BE_{ar} app was tested in the preclinical setting in a study where eight patients with DM1 attended training sessions and visually estimated the CHO content of meals in laboratory setting at the beginning and the end of the study (one-month difference). In 8/18 of the estimations, the absolute error decreased by at least 6 g of CHO when comparing beginning against ending study estimations.

DietCam²¹ recognizes food and calculates calorie content from three images or a short video. It is automatic but also has the option of being semiautomatic if food is not correctly recognized. The app consists of three main modules: image manager, food classifier, and volume estimator. The image manager either receives three images from the user or extracts them from a video input. The food classifier segments and classifies each food part against local images and remote databases, which is assisted by optical character recognition techniques or user inputs. Finally, the volume is estimated by means of a reference object or by user input. To obtain calorie content, the USDA database is used. The recognition accuracy of DietCam was tested with real food items ($n = 20$).

goFOOD is an app that supports input from images and/or short videos.²² It automatically outputs energy and macronutrient content of meals by detecting, segmenting, and identifying more than 300 food items, and reconstructing the 3D model—after being linked to nutrient databases. The system also supports food item identification for packaged products using the smartphone's camera as a barcode scanner. The

system has been validated on two multimedia databases: one containing Central European meals and the other containing fast-food meals. goFOOD outperformed the estimations of experienced dietitians in the European meals and achieved comparable performance on standardized fast-food meals.²² A modified version of goFOOD has been used for the automatic assessment of adherence to Mediterranean diets and was able to recognize multiple food items and their corresponding portion sizes simultaneously from a single image. The results showed that the algorithm performed better than a widely used algorithm and gave similar results to those of experienced dietitians ($n = 4$).⁴² One version of the system, named goFOOD Lite, has been also used as a tracking tool without providing output to the user—in a study aiming at determining the common human mistakes made by users.⁴³

A system was initially aimed at measuring dietary intake for people living with diabetes.²³ The user needs to take a 360° video of the food and the system asks the user to insert the ingredients for one portion size of the meal or alternatively its recipe. The “Digital Dietary Recording System” algorithm then uses the structured light system and the recorded video to calculate the volume of the food. 3D mapping is achieved by attaching a laser module that projects a matrix onto the surface of the food item and, once the volume has been calculated, the nutritional content of the meal is produced using the Nutritional Data Systems for Research. In a pilot study, ten participants were recruited to use a prototype of the device to record their food intake for three days. The results showed improved volume estimation by almost 40% when compared with manual calculations.

Another system estimates the weight, energy, and macronutrient content in a plated meal on the basis of an RGB-depth image.²⁵ The user captures a single photo with a smartphone with a built-in depth sensor. The system first segments the image into the different food items and estimates their volumes, utilizing the depth map. Finally, the volume of each food item is converted to its weight and, thus, to its energy and macronutrient content, using a food database. A study was conducted to evaluate the system’s performance on 48 test meals. The accuracy was generally lower for the cooked meals because food items that appear in them are not well separated. One limitation of this approach is that the smartphone needs to be equipped with a depth sensor to perform volume estimation, a feature provided by few smartphones, and which, if at all, is mainly used for facial recognition (front camera).

CALO mama is a smartphone health app that provides automatic food recognition, while nutrient and food estimation are based on meal images, as well as manual correction of the output by the user. A validation study has been conducted to evaluate its performance in a controlled environment.²⁴ The study involved 120 sample meals from 15 food groups, with known weight, energy, and 30 nutrients. The results showed that manual adjustments by the user to the

automatic output of the app can significantly improve the performance of the app.

In another study, the authors created a system that automatically estimates the weight and CHO content of a meal using a color image and a depth map, acquired from an iPhone’s camera.²⁶ As in Herzig et al’s study,²⁵ the user needs to capture the meal photo with the smartphone’s front camera to utilize its depth sensor. The system first divides the image into food and nonfood segments using a deep neural network, and then, the food segments are further partitioned into the different food items. The depth map and the camera calibration information are used to estimate the volume of each food item, while a commercially available app module is used to recognize each segmented food item and calculate its CHO content. The authors created their own density library to convert food volume to weight.

Studies Conducted in a Laboratory Setting

Twelve systems have been described in only a laboratory setting. Three of the apps offer a complete system from image capture to its conversion into nutrients,^{27,36,38} five systems are restricted to food image recognition,^{28,30,31,34,35} and four systems only focus on estimating food volume.^{29,33,37,39}

Complete Systems

An Android app uses image processing techniques and incorporates image classification of food photos, and estimation of volume and CHO content.²⁷ A small set of fruit images taken under real-life conditions was used to test the classification performance and compared with ground truth. A standard credit card was placed next to the fruits as reference. The significance of this study is unclear, as the used food-image database was limited, and its poor variety does not reflect real-life conditions. No information was provided on the nutrient database which was used to convert fruits’ volume into CHO value.

An additional app was intended to provide a recommendation for real-time energy balance with semiautomatic image recognition.³⁶ The user takes four pictures of meals and is asked to use his/her fingerprint as a reference object. The system selects the top-ranked food predictions, and the user verifies these. The user is then asked to place each food item inside a partitioned plate with three fixed areas, to improve localization and segmentation. The system contains a calculation of food database (correlated volume/weight), based on the USDA database, as this can be used to calculate nutrient content (macronutrients and energy).

The snap-n-eat is an Android and web food recognition system that automatically estimates energy and nutrient content of foods.³⁸ The user takes a photo of a food plate that can contain different food items, and once most of the food is inside a circle it clicks on the screen. Cropped images are

sent to the server and automatically recognized by the system, which hierarchically segments the image and removes the background correspondingly. The system then estimates the size of the food portion by counting the pixels in each corresponding food segment, which can then be used to estimate the caloric and nutrient content of the food on the plate. However, no information is provided on the food databases used.

Image Recognition

One publication focuses on the image recognition function of a baseline prototype of an automated food record.³⁵ This uses image processing and pattern recognition algorithms. The prototype supports automatic recording and recognition and facilitates (future) determination of portion sizes. After matching with food composition databases, this gives the food content. The advantages of the system are that there is no need for a fiducial marker or manual user input of text or voice. The system also aims to recognize multiple foods from a single image, as this is more realistic in practice.

The aim of this study was to develop a model for image detection and recognition of Korean foods, to be used in mobile devices for the accurate estimation of dietary intake.³⁴ The researchers collected 4000 food images by taking pictures and searching the web to build a training data set for recognition of Korean foods. The food images were categorized into 23 food groups, based on dishes commonly consumed in Korea. After using augmentation techniques, a data set of 92 000 images has been created. The study results showed that K-foodNet achieved higher recognition accuracy in Korean foods than other models.

Kawano and Yanai proposed a semiautomatic Android app which would be able to automatically recognize food type and identify food calories and nutrition ingredients.³⁰ The user points the smartphone camera at the food and draws a bounding box around the food item. The app performs food item recognition within the indicated bounding box, while the user receives a list of top five candidate food items and selects the appropriate one. Subsequent classification is performed by a linear SVM which assigns the food image features to 100 food categories. The users liked the app's usability but disliked the quality of the direction from the app during the recognition process. The authors also claim that the system could output calorie and protein estimation, though they did not present any information on the food database used or results on kcal/g of protein estimations.

In another study, the system evaluation included about 2000 images of 15 predefined food classes, containing 100 to 400 sample images of each food in the training set.²⁸ The authors emphasize the importance of personalization of such systems and mention that information on the user's location can improve the accuracy of the classification. The authors present a study of food image processing from the computer vision perspective and propose the UNICT-FD1200 data set,

which consists of 4754 food images of 1200 different plates obtained in real meals which are categorized into eight groups.

The NutriNet detection and recognition system has been proposed for both food and beverage images that contain a single item and aims at high classification accuracy.³¹ The authors created a 520-class data set containing images from Google search. NutriNet is part of an app designed for the dietary assessment of people living with Parkinson's disease and incorporates training material which can automatically update the model with new images, as well as new categories of food and drinks. More recently, the authors also proposed two segmentation networks that are able not only to classify the food and beverage items that appear in an image, but also to identify their location in a meal image.³² Therefore, the segmentation networks can also recognize multiple foods/beverages in the image.

Food Volume Estimation

This study aims to investigate the ability to estimate food volume from a mobile's single photo.²⁹ It uses two distinctive approaches, both of which are based on an elliptical reference pattern. The proposed automated method works by sending all the images of the user's food to the dietitian before and after consumption. Then there are three steps: first, the dietitian identifies the food and ingredients; second, the app recognizes the consumed volume; and third, the app evaluates nutrients and calories by using the USDA database. The study evaluates only the volume recognition algorithms. The location and orientation of food objects and their volumes are subsequently calculated using this reference pattern and image processing techniques. An important problem is that many foods are not regularly shaped and this ambiguity may bias the methods.

A novel smartphone-based imaging approach for estimating the food volume overcomes a major current limitation, the need for a fiducial marker.³⁷ This model shows that the smartphone-based imaging system can be adequately calibrated with a special recording strategy if the smartphone length and the motion sensor output in the device are known. However, in the current system, the volume estimation is established manually. In their pilot study, 69 participants with no prior experience in visual estimation manually estimated the portion size of 15 food models. Some of the participants ($n = 29$) followed a training process to improve their visual estimations. The results showed that average absolute error was smaller for large-volume items and that volume estimation was improved by training.

A deep learning-based approach was proposed to synthesize views, and, thus, to reconstruct 3D point clouds of food and estimate the volume from a single depth image.³³ In a typical scenario, a mobile phone with depth sensors or a depth camera is used, and RGB-depth images are captured from any convenient angle. The images are segmented and

the corresponding regions in the depth image are labeled accordingly. Then 3D reconstruction is implemented, and food volume is estimated. Finally, the information is linked with the USDA nutrient database to output the nutrient content.

The final study of this category focused on the segmentation of the food items as well as their volume estimation, using an RGB-depth camera and utilizing the known diameter of the plate.³⁹ The testing data set consisted of 689 images and 36 unique foods, and it was divided into two subsets: the “regular texture” data set and the “modified texture” data set. They are experimented with two approaches, with and without utilizing the depth maps created by the depth sensor of the camera.

The basic characteristics of the aforementioned studies are summarized in Table 1.

Discussion

Calculating energy intake and specific nutrients is a demanding task that requires specific education and training. Although several apps have already been developed for this purpose, this technology has not yet been widely adopted. In an international survey, health care professionals ($n = 1001$) stated that they would recommend a nutrition app to their clients/patients, if it was validated, easy to use, free of charge, and if it supported automatic food recording and automatic nutrient and energy estimation.⁴⁴ In another survey, potential nutrition app users ($n = 2382$) mentioned that they would not select a nutrition app if it produced incorrect calorie and nutrient results, if local foods were not supported, and if it produced unconvincing estimations of portion size.⁴⁵

There are still some challenges related to image-based systems, as foods are highly variable and have different shapes, and mixed foods may be highly complex.¹⁵ Some nutrients, such as types of fat, cannot be detected via images without user input and the plate's shape or color may affect the system's performance. Apps are less accurate with more complex foods, such as an entire meal with multiple ingredients on a plate.³³ In the ideal case, large numbers of well-annotated pictures are needed that contain information on the ground truth to train the system and enhance its accuracy.³⁶

Apart from the challenges related to the apps' nature, ways to improve users' motivation should be investigated. In a study in which users captured photos of their meals using a free self-reporting app, it was mentioned that only 2.6% (4895/189 770) of them could be identified as active users.⁴⁶ In a randomized controlled trial where a weight loss app was compared with Web site use or paper diary, it was found that using the app increases user's recording/input rate.⁴⁷ However, it was not determined for how long the app has to be used to achieve changes in eating habits, body weight, or nutritional knowledge.⁴⁸

More studies are required to address long-term effects of those apps, since the existing interventions had a duration of three to six months.⁴⁹ This systematic review and meta-analysis

of commercially available dietary self-monitoring apps showed encouraging results of those apps with regard to weight loss when compared with other self-monitoring tools (eg, paper-based diaries), other technologies (eg, bite counter), or no self-monitoring approach. The authors indicated that although the findings were positive, the quality and effectiveness of commercially available apps were questionable; therefore, they should be interpreted with caution.⁴⁹

If this technology is to be widely adopted, the AI systems should be trustworthy and accompanied by clinical trials and studies performed in a realistic clinical environment. Localization in terms of nutrient databases and local food incorporated to them is of high importance⁴⁵ to support the eating patterns of a specific population. Last, there is a need for open access, clean, and labeled databases that will boost the research and development not only in the field of health, but also in nutrition and food industry, toward the provision of tailored advice on dietary scheme personalized to the eating culture, habits, preferences, health status, and needs of people with diabetes.

Conclusion

Smartphone apps that provide nutritional information based on food pictures and videos are a promising field of research. However, most of the presented systems have been designed and used for research purposes. Apps that can be tested under real-life conditions are required and can then be integrated in the toolbox of patients living with nutrition-related chronic diseases or who just wish to change their lifestyle.

Abbreviations

AI, artificial intelligence; app, application; BMI, body mass index; CHO, carbohydrate; DLW, doubly labeled water; DM, diabetes mellitus; FNDDS, Food and Nutrient Database for Dietary Surveys; mFR, Mobile Food Record; NDSR, Nutritional Data Systems for Research; OCR, optical character recognition; SD, standard deviation; SVM, support vector machine; USDA, United States Department of Agriculture.

Declaration of Conflicting Interests


The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iDs

Maria F. Vasiloglou  <https://orcid.org/0000-0002-2013-2858>

Elias K. Spanakis  <https://orcid.org/0000-0002-9352-7172>

Stavroula Mougiakakou  <https://orcid.org/0000-0002-6355-9982>

References

1. Centers for Disease Control and Prevention. National Diabetes Statistics Report, 2020. Atlanta, GA: U.S. Department of

- Health and Human Services, Centers for Disease Control and Prevention; 2020.
2. Brazeau AS, Mircescu H, Desjardins K, et al. Carbohydrate counting accuracy and blood glucose variability in adults with type 1 diabetes. *Diabetes Res Clin Pract.* 2013;99(1):19-23.
 3. Deeb A, Al Hajeri A, Alhmoudi I, Nagelkerke N. Accurate carbohydrate counting is an important determinant of postprandial glycemia in children and adolescents with type 1 diabetes on insulin pump therapy. *J Diabetes Sci Technol.* 2017;11(4):753-758.
 4. Meade LT, Rushton WE. Accuracy of carbohydrate counting in adults. *Clin Diabetes.* 2016;34(3):142-147.
 5. Hales CM, Fryar CD, Carroll MD, Freedman DS, Ogden CL. Trends in obesity and severe obesity prevalence in US youth and adults by sex and age, 2007-2008 to 2015-2016. *JAMA.* 2018;319(16):1723-1725.
 6. Semper HM, Povey R, Clark-Carter D. A systematic review of the effectiveness of smartphone applications that encourage dietary self-regulatory strategies for weight loss in overweight and obese adults. *Obes Rev.* 2016;17(9):895-906.
 7. Papatthanail I, Brühlmann J, Vasiloglou MF, et al. Evaluation of a novel artificial intelligence system to monitor and assess energy and macronutrient intake in hospitalised older patients. *Nutrients.* 2021;13(12):4539.
 8. Food and Agriculture Organization of the United Nations. *Dietary Assessment: A Resource Guide to Method Selection and Application in Low Resource Settings.* Rome, Italy: Food and Agriculture Organization; 2018.
 9. Stumbo PJ. New technology in dietary assessment: a review of digital methods in improving food record accuracy. *Proc Nutr Soc.* 2013;72(1):70-76.
 10. Eldridge AL, Piernas C, Illner AK, et al. Evaluation of new technology-based tools for dietary intake assessment-an ILSI Europe dietary intake and exposure task force evaluation. *Nutrients.* 2018;11(1):55.
 11. Reber E, Gomes F, Vasiloglou MF, Schuetz P, Stanga Z. Nutritional risk screening and assessment. *J Clin Med.* 2019;8:1065.
 12. Anthimopoulos M, Dehais J, Shevchik S, et al. Computer vision-based carbohydrate estimation for type 1 patients with diabetes using smartphones. *J Diabetes Sci Technol.* 2015;9(3):507-515.
 13. Rhyner D, Loher H, Dehais J, et al. Carbohydrate estimation by a mobile phone-based system versus self-estimations of individuals with type 1 diabetes mellitus: a comparative study. *J Med Internet Res.* 2016;18(5):e101.
 14. Bally L, Dehais J, Nakas CT, et al. Carbohydrate estimation supported by the GoCARB system in individuals with type 1 diabetes: a randomized prospective pilot study. *Diabetes Care.* 2017;40(2):e6-e7.
 15. Vasiloglou MF, Mouggiakakou S, Aubry E, et al. A comparative study on carbohydrate estimation: GoCARB vs. dietitians. *Nutrients.* 2018;10(6):741.
 16. Ji Y, Plourde H, Bouzo V, Kilgour RD, Cohen TR. Validity and usability of a smartphone image-based dietary assessment app compared to 3-day food diaries in assessing dietary intake among Canadian adults: randomized controlled trial. *JMIR Mhealth Uhealth.* 2020;8(9):e16953. doi:10.2196/16953.
 17. Zhu F, Bosch M, Woo I, et al. The use of mobile devices in aiding dietary assessment and evaluation. *IEEE J Sel Top Signal Process.* 2010;4(4):756-766.
 18. Six BL, Schap TE, Zhu FM, et al. Evidence-based development of a mobile telephone food record. *J Am Diet Assoc.* 2010;110(1):74-79.
 19. Boushey CJ, Spoden M, Delp EJ, et al. Reported energy intake accuracy compared to doubly labeled water and usability of the mobile food record among community dwelling adults. *Nutrients.* 2017;9(3):312.
 20. Domhardt M, Tiefengrabner M, Dinic R, et al. Training of carbohydrate estimation for people with diabetes using mobile augmented reality. *J Diabetes Sci Technol.* 2015;9(3):516-524.
 21. Kong F, Tan J. 2012. DietCam: automatic dietary assessment with mobile camera phones. *Pervasive Mob Comput.* 2012;8(1):147-163. doi:10.1016/j.pmcj.2011.07.003.
 22. Lu Y, Stathopoulou T, Vasiloglou MF, et al. goFOOD™: an artificial intelligence system for dietary assessment. *Sensors.* 2020;20:4283. doi:10.3390/s20154283.
 23. Makhssous S, Mohammad HM, Schenk JM, Mamishev AV, Kristal AR. A novel mobile structured light system in food 3D reconstruction and volume estimation. *Sensors (Basel).* 2019;19(3):564. doi:10.3390/s19030564.
 24. Sasaki Y, Sato K, Kobayashi S, Asakura K. Nutrient and food group prediction as orchestrated by an automated image recognition system in a smartphone app (CALO mama): validation study. *JMIR Form Res.* 2022;6(1):e31875. doi:10.2196/31875.
 25. Herzig D, Nakas CT, Stalder J, et al. Volumetric food quantification using computer vision on a depth-sensing smartphone: preclinical study. *JMIR Mhealth Uhealth.* 2020;8(3):e15294.
 26. Huang HW, You SS, Di Tizio L, et al. An automated all-in-one system for carbohydrate tracking, glucose monitoring, and insulin delivery. *J Control Release.* 2022;343:31-42.
 27. Huang J, Ding H, McBride S, Ireland D, Karunanithi M. Use of smartphones to estimate carbohydrates in foods for diabetes management. *Stud Health Technol Inform.* 2015;214:121-127.
 28. Farinella GM, Allegra D, Moltisanti M, Stanco F, Battiato S. Retrieval and classification of food images. *Comput Biol Med.* 2016;77:23-39. doi:10.1016/j.combiomed.2016.07.006.
 29. Jia W, Yue Y, Fernstrom JD, et al. Imaged based estimation of food volume using circular referents in dietary assessment. *J Food Eng.* 2012;109(1):76-86. doi:10.1016/j.jfoodeng.2011.09.031.
 30. Kawano Y, Yanai K. FoodCam: a real-time food recognition system on a smartphone. *Multimed Tools Appl.* 2015;74:5263-5287. doi:10.1007/s11042-014-2000-8.
 31. Mezgec S, Koroušić Seljak B. NutriNet: a deep learning food and drink image recognition system for dietary assessment. *Nutrients.* 2017;9(7):657. doi:10.3390/nu9070657.
 32. Mezgec S, Koroušić Seljak B. Deep neural networks for image-based dietary assessment. *J Vis Exp.* 2021(169). doi:10.3791/61906.
 33. Lo FP, Sun Y, Qiu J, Lo B. Food volume estimation based on deep learning view synthesis from a single depth map. *Nutrients.* 2018;10(12):2005. doi:10.3390/nu10122005.
 34. Park SJ, Palvanov A, Lee CH, Jeong N, Cho YI, Lee HJ. The development of food image detection and recognition model of Korean food for mobile dietary management. *Nutr Res Pract.* 2019;13(6):521-528. doi:10.4162/nrp.2019.13.6.521.
 35. Probst Y, Nguyen DT, Tran MK, Li W. Dietary assessment on a mobile phone using image processing and pattern recognition techniques: algorithm design and system prototyping. *Nutrients.* 2015;7(8):6128-6138. doi:10.3390/nu7085274

36. E Silva BVR, Rad MG, Cui J, McCabe M, Pan K. A mobile-based diet monitoring system for obesity management. *J Health Med Inform.* 2018;9(2):307. doi:10.4172/2157-7420.1000307.
37. Yang Y, Jia W, Bucher T, Zhang H, Sun M. Image-based food portion size estimation using a smartphone without a fiducial marker. *Public Health Nutr.* 2019;22(7):1180-1192. doi:10.1017/S136898001800054X.
38. Zhang W, Yu Q, Siddiquie B, Divakaran A, Sawhney H. "Snap-n-eat": food recognition and nutrition estimation on a smartphone. *J Diabetes Sci Technol.* 2015;9(3):525-533. doi:10.1177/1932296815582222.
39. Pfisterer KJ, Amelard R, Chung AG, et al. Automated food intake tracking requires depth-refined semantic segmentation to rectify visual-volume discordance in long-term care homes. *Sci Rep.* 2022;12:83. doi:10.1038/s41598-021-03972-8
40. Bouzo V, Plourde H, Beckenstein H, Cohen TR. Evaluation of the diet tracking smartphone application Keenoa™: a qualitative analysis. *Can J Diet Pract Res.* 2021;1-5. doi:10.3148/cjdr-2021-022.
41. Dehais J, Anthimopoulos M, Shevchik S, Mougiakakou S. Two-view 3D reconstruction for food volume estimation. *IEEE Transactions on Multimedia.* 2017;19(5):1090-1099.
42. Vasiloglou MF, Lu Y, Stathopoulou T, et al. Assessing Mediterranean Diet Adherence with the smartphone: the mediaplatto project. *Nutrients.* 2020;12:3763. doi:10.3390/nu12123763
43. Vasiloglou MF, van der Horst K, Stathopoulou T, et al. The human factor in automated image-based nutrition apps: analysis of common mistakes using the goFOOD Lite App. *JMIR Mhealth Uhealth.* 2021;9(1):e24467. doi:10.2196/24467
44. Vasiloglou MF, Christodoulidis S, Reber E, et al. What health-care professionals think of "nutrition & diet" apps: an international survey. *Nutrients.* 2020;12(8):2214. doi:10.3390/nu12082214.
45. Vasiloglou MF, Christodoulidis S, Reber E, et al. Perspectives and preferences of adult smartphone users regarding nutrition and diet apps: web-based survey study. *JMIR Mhealth Uhealth.* 2021;9(7):e27885.
46. Helander E, Kaipainen K, Korhonen I, Wansink B. Factors related to sustained use of a free mobile app for dietary self-monitoring with photography and peer feedback: retrospective cohort study. *J Med Internet.* 2020;16:1-13.
47. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. *J Med Internet Res.* 2013;15:e32.
48. Holzmann SL, Pröll K, Hauner H, Holzapfel C. Nutrition apps: quality and limitations. An explorative investigation on the basis of selected apps. *Ernahrungs Umschau.* 2017;64(5): 80-89.
49. El Khoury CF, Karavetian M, Halfens RJ, Crutzen R, Khoja L, Schols JM. the effects of dietary mobile apps on nutritional outcomes in adults with chronic diseases: a systematic review and meta-analysis. *J Acad Nutr Diet.* 2019;119(4):626-651. doi:10.1016/j.jand.2018.11.010