

## Article

# A Cross-Sectional Reproducibility Study of a Standard Camera Sensor Using Artificial Intelligence to Assess Food Items: The FoodIntech Project

Virginie Van Wymelbeke-Delannoy<sup>1,2,\*</sup> , Charles Juhel<sup>3</sup>, Hugo Bole<sup>3</sup>, Amadou-Khalilou Sow<sup>4</sup> , Charline Guyot<sup>1</sup>, Farah Belbaghdadi<sup>1</sup>, Olivier Brousse<sup>5</sup> and Michel Paindavoine<sup>5</sup>

<sup>1</sup> Elderly Unit, University Hospital Center Dijon Bourgogne F Mitterrand, F-21000 Dijon, France; charline.guyot@chu-dijon.fr (C.G.); farah.belbaghdadi@chu-dijon.fr (F.B.)

<sup>2</sup> Centre des Sciences du Goût et de l'Alimentation, INRAE, Université de Bourgogne Franche-Comté, CNRS, Agrosup, F-21000 Dijon, France

<sup>3</sup> ATOL Conseils & Développements (AtolCD), ZAE les Terres d'Or, Route de Saint Philibert, F-21220 Gevrey-Chambertin, France; cju@atolcd.com (C.J.); hbo@atolcd.com (H.B.)

<sup>4</sup> CHU Dijon Bourgogne, Inserm, Université de Bourgogne, CIC 1432, Module Épidémiologie Clinique, F-21000 Dijon, France; amadou-khalilou.sow@chu-dijon.fr

<sup>5</sup> Yumain, 14 Rue Pierre de Coubertin, F-21000 Dijon, France; olivier.brousse@yumain.fr (O.B.); michel.paindavoine@yumain.fr (M.P.)

\* Correspondence: virginie.vanwymelbeke@chu-dijon.fr; Tel.: +33-00-3-80-29-31-55

**Abstract:** Having a system to measure food consumption is important to establish whether individual nutritional needs are being met in order to act quickly and to minimize the risk of undernutrition. Here, we tested a smartphone-based food consumption assessment system named FoodIntech. FoodIntech, which is based on AI using deep neural networks (DNN), automatically recognizes food items and dishes and calculates food leftovers using an image-based approach, i.e., it does not require human intervention to assess food consumption. This method uses one-input and one-output images by means of the detection and synchronization of a QRcode located on the meal tray. The DNN are then used to process the images and implement food detection, segmentation and recognition. Overall, 22,544 situations analyzed from 149 dishes were used to test the reliability of this method. The reliability of the AI results, based on the central intra-class correlation coefficient values, appeared to be excellent for 39% of the dishes ( $n = 58$  dishes) and good for 19% ( $n = 28$ ). The implementation of this method is an effective way to improve the recognition of dishes and it is possible, with a sufficient number of photos, to extend the capabilities of the tool to new dishes and foods.

**Keywords:** artificial intelligence; portion evaluation; mobile phone images; machine learning; reliability



**Citation:** Van Wymelbeke-Delannoy, V.; Juhel, C.; Bole, H.; Sow, A.-K.; Guyot, C.; Belbaghdadi, F.; Brousse, O.; Paindavoine, M. A Cross-Sectional Reproducibility Study of a Standard Camera Sensor Using Artificial Intelligence to Assess Food Items: The FoodIntech Project. *Nutrients* **2022**, *14*, 221. <https://doi.org/10.3390/nu14010221>

Academic Editor: Megan A. McCrory

Received: 23 November 2021

Accepted: 1 January 2022

Published: 5 January 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Measuring patient food consumption and food waste are important for the goals of healthy eating and sustainability. In hospitals, having a system to measure food consumption is a key element to know if patients' nutritional needs are being properly covered and to decrease food waste. However, recording food intake is challenging to implement in the hospital context, and often suboptimal [1,2], and nutritional monitoring is thus rarely part of the clinical routine [3]. Potential barriers include the lack of knowledge and poor awareness of caregivers, the medical prescription of the three-day food record, and the absence of a quick and easy-to-use monitoring tool that is accurate and precise. The most used method is a direct visual estimation [4]. This method consists of quantifying the remaining part of the food intake of subjects during mealtimes by trained staff (care staff, dieticians or experimenters). Direct visual estimation is easy to organize, and has proven to be reliable in different institutions, including geriatric units [5–7]. This method can be used to differentiate each item of food served and leftovers by portion (1, 3/4, 1/2, 1/4 or 0 of

food served) because employees can mentally separate different items on a plate. However, direct visual estimation is often performed quickly and with a low degree of precision because staff are not always trained. In addition, the estimation of the food before and after consumption is not always done by the same employee. This approach is also subjective, which may lead to errors. Additionally, while a large number of subjects can be followed, this method also requires the availability of multiple employees at the same time.

In a clinical research context, the weighing method remains the “gold standard” to measure actual food consumption. This method is based on the difference between the weights of foods offered before consumption and those not consumed after consumption by the participant [8,9]. However, the weighing method is complex in real-life in facilities, and the delivery modes are difficult to reconcile with data requirements. This method is time-consuming, requires significant staff resources, and remains unsuitable for the hospital environment [8,9].

Other methods using photography have been developed. Martin et al. developed the Remote Food Photography Method (RFPM), which involves participants capturing images of their food selection and leftovers [10,11]. These images were then sent to the research center via a wireless network, where they are analyzed by dietitians to estimate food intake. These same authors were able to show that there was a good correspondence between the RFPM method and the weighing method, with a difference ranging from  $-9.7\%$  for the starter to  $6.2\%$  for the vegetables [12]. This technique does not require a large number of personnel on site during mealtimes. It was compared to the visual estimation process in collective dining facilities [13], to the weighing method in an experimental restoration situation [14], in a cafeteria situation [15] and in a geriatric institution [16]. It avoids the bias associated with the presence of experimenters at mealtimes and, by capturing images, the analyses can be repeated by different people. Although these methods are reliable, they all require a professional to perform the data analysis, and they remain time-consuming.

Several novel techniques have recently been proposed thanks to advances in Artificial Intelligence (AI) applications. AI has expanded in different domains using images with new opportunities in nutrient science research [17]. In a review, mobile applications based on systems using AI were of significant importance in the different fields of studies on biomedical and clinical nutrients research and nutritional epidemiology. Among the available AI applications, two algorithms can be used: machine learning (ML) algorithms, widely used in studies on the influence of nutrients on the functioning of the human body in health and disease; and deep learning (DL) algorithms, used in clinical studies on nutrient intake [17]. ML is an AI domain related to algorithms that improve automatically through gathered experience, making it possible to create mathematical models for decision-making. DL is a subtype of ML, with the advantage of program autonomy that can build functions used in recognition. Recently, Lu et al. proposed a dietary assessment system, named goFOOD<sup>TM</sup>, based on AI using DL [18]. This method uses the deep neural networks to process two images taken by a single press of the camera shutter button. Even if the results demonstrated that goFOOD<sup>TM</sup> performed better with detection using DL algorithms than experienced dietitians did, this method still requires human intervention.

In this paper, we propose a smartphone-based food consumption assessment system, called FoodIntech. FoodIntech, which is based on AI using DL, automatically recognizes food items and dishes and calculates portion leftovers using an image-based approach, i.e., it does not require human intervention to assess food consumption. This method uses one-input and one-output images by means of the detection and synchronization of a QRcode label stuck to the meal tray. Then, the deep neural networks are used to process the images and implement food detection, segmentation and recognition. The aim of this paper is to test the reliability of this method in laboratory conditions, but as similar as possible to routine clinical practice in a hospital environment.

## 2. Materials and Methods

### 2.1. General Procedure

The tool developed is the subject of an industrial program whose technical details cannot be fully disclosed.

The implemented method is based on the analysis of food tray pictures taken by a standard out-of-the box mobile phone (i.e., Android Samsung S8), before and after the simulated consumption of a standard meal tray taken from the central kitchen of the Dijon University Hospital, and processed by laboratory research staff. However, the method implemented here attempted to emulate routine clinical practice in a hospital environment: the smartphone was attached by a system adapted to the heating trolley used to distribute the meal trays in the hospital, the food quantities measured were identical to those served to patients, and the same dishes, the same trays, the same crockery and the same arrangement of dishes were also used. The protocol was designed to simulate patient consumption in a large range of situations to build a consolidated vision of the liability and repeatability of the process (detailed in Weighed food method paragraph).

Each pair of pictures of the tray (before/after) was recorded in an experimental database and linked to the weight values of each food item on each plate corresponding to the captured images. The food items were weighed before and after the simulation by the research team, for all the plates and in each of the experimental conditions established in the protocol. The collected data constitute the reference dataset for comparison of the study on which the technical process of the solution has been evaluated.

AI based on deep learning received the input of 13,152 pictures showing 26,584 food item consumption situations produced throughout the study to ensure learning based on image segmentation. The AI program therefore needed a very large number of pictures per food item or dish to learn to recognize it, around 200 images for each.

The raw data returned by AI cannot immediately be processed by the system; it has to be interpreted. As such, a transcoding overlay was added. The sole purpose of the overlay is to process the AI results and map the using the list of known dishes from the menu.

Sometimes, the transcoding overlay is unable to map the dishes properly, especially when the AI returns a relatively vague result (e.g., the same element several times). In this case, a refinement increment is required to be able to discern which is which, and to return an accurate response, enabling a flawless transcoding process.

The results of the system were also challenged to obtain a precise percentage of each food component portion remaining from one picture to another compared to the weighted method.

Each new increment of the AI was thoroughly tested to track any possible regression. To do so, each dataset line result was reprocessed and then compared with the experimental value to produce a comprehensive vision of the AI's capabilities for this version.

Both the AI core and the transcoding overlay were modified in the process so they could deliver the best results. Many iterations were made, focusing on different priorities each time, to be able to come up with a satisfying AI result. Monitoring the different results release after release made it possible to follow the evolution of the AI and to establish the optimal process of teaching the neural network to recognize new dishes/components. This also made it possible to manage inconsistencies with the transcoding system to achieve a technically reliable system capable of digital detection and weighing (not in this paper).

### 2.2. Algorithm Used for Deep Learning

The measurement of the quantity of food ingested was based on an analysis of images taken before and after meals. This analysis consisted of a measurement of the pixel surface of the different food items present in the menu. Thanks to the progress that has been made over the last few years with deep learning techniques, image analysis has also become increasingly powerful. Three main approaches are used to analyze images with these techniques: the first one, called "Classification", indicated the presence or not of a particular object (here food) in the analyzed image. The second more precise approach, called "Detection", indicated in which area of the image the recognized object was located.

Finally, the third approach, which is the most precise, consisted in a fine clipping of the detected object. This approach, called “Instance Segmentation”, was well adapted to our application.

There are mainly two types of Deep Learning neural networks dedicated to instance segmentation: U-Net [19] and Mask R-CNN [20]. Based on study of Vuola and al, we decided to use Mask R-CNN rather than U-Net for its better ability to identify and separate small and medium-sized objects that compose a large part of the food items present in the analyzed plates [21].

For both training and inference, the spatial resolution of the images was between 800 and 1365 pixels per side. The Mask-RCNN neural networks were trained and used on a computer equipped with a NVidia Titan RTX graphics card with 24 GB VRAM.

For the implementation of Mask-RCNN, we used the `mask_rcnn_inception_resnet_v2_atrous_coco` one from Tensor Flow Object Detection Model Zoo [22]. It is based on the inception-resnet-v2 backbone [23], and provides great precision masks at the cost of a high computing power requirement. Classical data augmentation was used for the training with a randomized image horizontal flip and a randomized brightness adjustment at each period of the training. In order to achieve satisfactory performance, it was necessary to train the Mask-RCNN with an average of 200 examples per food item; items were accurately annotated and included in the training database.

The details of the learning parameters are provided in Appendix B.

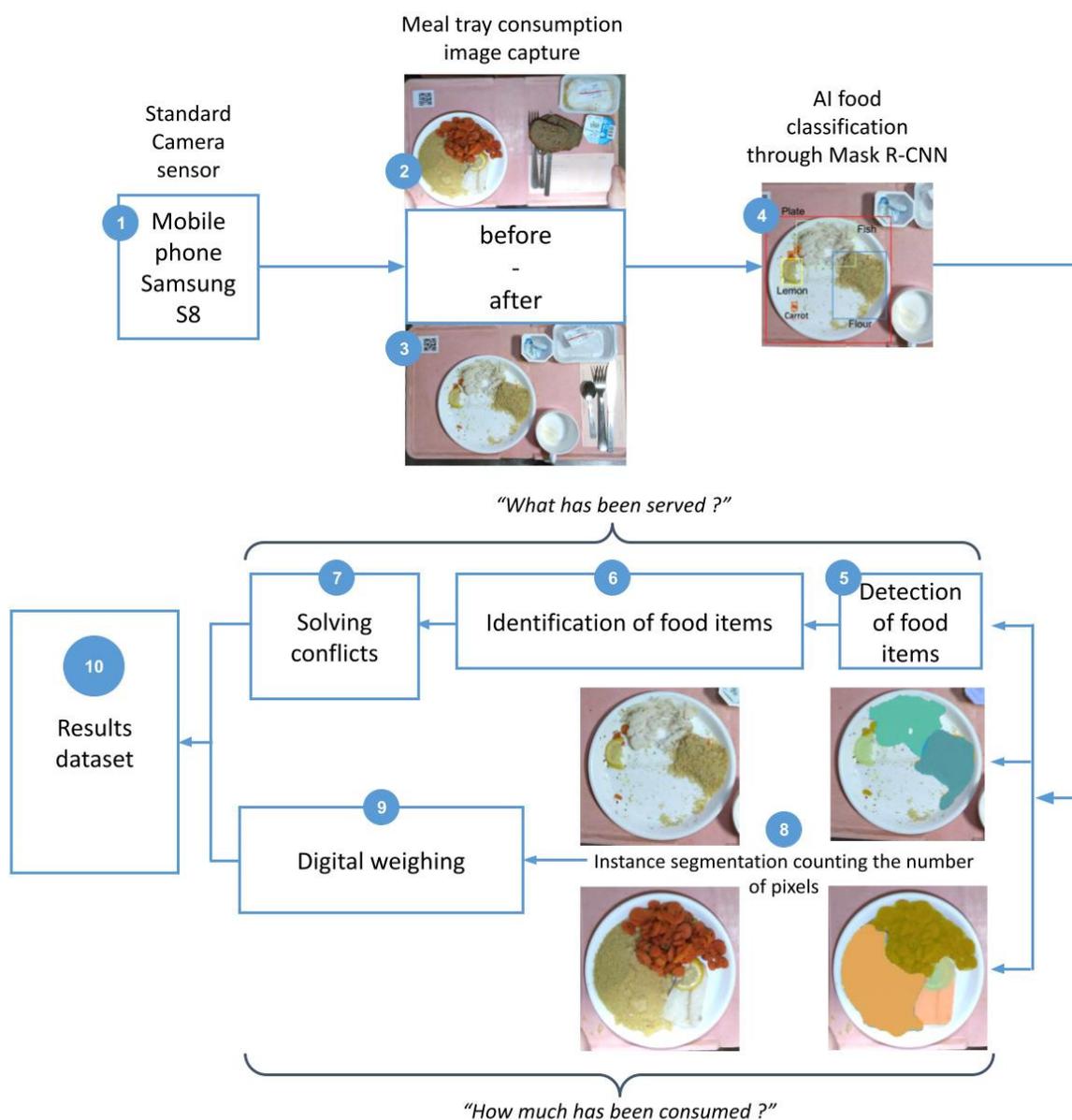
Figure 1a,b gives an illustration of this segmentation technique. The left part of the figure corresponds to the image of the plate acquired with the mobile phone and the right part to the image of the plate analyzed with the Mask-RCNN segmentation algorithm. We can see that the different foods contained in the plate are very well separated. From this segmentation, the surface of each food item is measured by the number of pixels needed to cover the item.



**Figure 1.** Food segmentation example applied to (a) plate with food served, (b) segmentation of each food of the plate.

### 2.3. Data Collection and Process or Procedure of Deep Learning View Synthesis Approach

The Foodintech procedure of using deep learning synthesis to determine food portions is shown in Figure 2:



**Figure 2.** Implementing computer vision technology and a deep learning process to recognize food items and calculate the amount of missing food between two meal tray pictures.

(1) A standard mobile phone was used, and the camera zoom of the device allowed to automatically focus the frame around the border of the tray; (2) and (3) two captured images before and after simulated consumption were taken and synchronized thanks to the QRcode label on the tray; (4) the bounding boxes show the results of Mask R-CNN to identify food items; (5) detection of food items: precision enhanced by deep learning annotations; (6) identification of food items: each food item is identified among 77 food categories, the exact recipe is defined by the menu which is provided in advance; (7) solving conflicts: transcoding layer applies decision rules to AI results conflicts or errors between items; (8) segmentation counting the numbers of pixels corresponding to different regions on the plate; (9) digital weighing: food intake calculated by % of missing portion for each item; (10) results dataset: displays consumption % of each plate on each tray applied to known served portion weight to show results in grams of food intake compared to real portion weight measures.

**Dish preparation**—The dishes were prepared in the central kitchen of the Dijon University Hospital according to the standardized recipes and the same procedures for each day of measurement. The plates were prepared in a meal tray as if they were being served

to patients, with the same quantities, presentation and daily menus. The quantities served adhered to the nutritional guidelines for adults in France. The captured images of 169 different labels of dishes were used to test the reliability of the FoodIntech method. They were produced between 22 December 2020, and 15 May 2021, at the Dijon University Hospital.

**Weighed food method**—Each meal component was weighed before and after each simulated consumption using the same electronic scales (TEFAL, precision  $\pm 1$  g). The experimenter, always the same person, initially weighed the dish with the quantities of the nutritional guidelines provided by the central kitchen and then removed 1 spoonful at a time until there was nothing left on the plate. The obtained weight allowed us to define the different experimental conditions for each dish. Then, the experimental conditions were reproduced between 4 to 14 different simulated consumptions. Each food component was weighed between 5 to 20 times for each experimental condition, and 30 to 200 pictures were taken per food item depending on the initial quantity served.

**Food Image acquisition**—Each meal tray was pictured before and after experimental conditions were applied using a standard mobile smartphone (Samsung Galaxy S8) with a zoom function on the camera. To estimate the reliability of the AI algorithm results, i.e., the estimated consumption percentage of a given dish, each of the 169 dishes studied had to be photographed 4 to a maximum of 14 times. A total of 13,152 pairs of images of 26,584 food items in various consumption conditions were produced. Two captured images, one input and one output, were synchronized thanks to a QRcode label (Figure 3a,b).



**Figure 3.** Meal trays with QRCode label (a) before consumption and (b) after consumption.

#### 2.4. Statistical Analysis

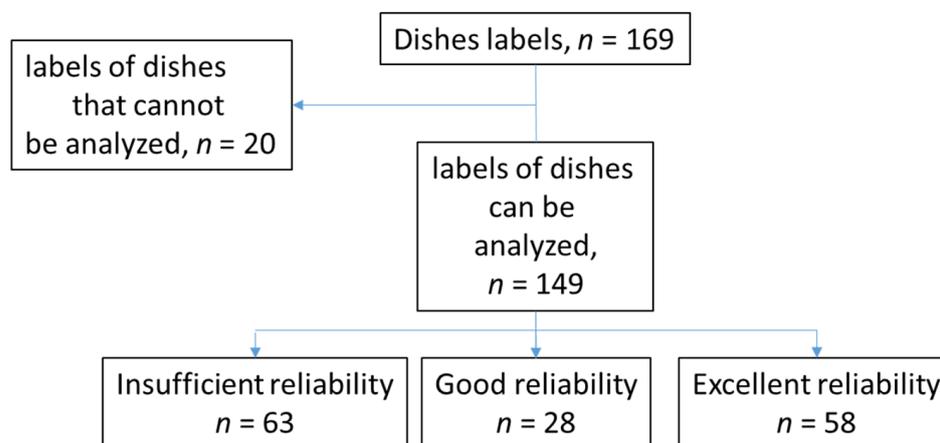
The sample size of this study was determined using the PASS software. A sample size of 20 subjects (camera angles) with 14 repetitions (food conditions) per subject achieves 80% power. This can detect an intra-class correlation of 0.90 under the alternative hypothesis when the intraclass correlation under the null hypothesis is 0.80, using an F-test with a significance level of 0.05 [24].

The estimation of the reliability of the AI results for each dish was assessed by the Type 3 Intra-class Correlation Coefficient (ICC) [25]. We used type 3 ICC issued from a two-way mixed model as recommended by Koo and Li 2016 [26]. Indeed, the reliability of a measurement refers to its reproducibility when it is repeated on the same subject. We wanted to have a complete overview with all possible camera angles, so positions were randomly selected among all the possible camera angles (infinite population). This corresponded to the random subject effect. Concerning repetitions, we were interested in assessing the 14 possible conditions, which reflect the finite, well-defined, possible measurement circumstances, and they were not randomly selected in an infinite population. This corresponded to the repetition factor, which was considered as a fixed effect in our model. All of these considerations were applied for each given food.

These ICCs were determined using a random effect model from the irrNA (R package version 0.2.2; Berlin Germany) R software package (version 4.1.0, R Core Team; R Foundation for Statistical Computing, Vienna, Austria). It is important to note that the calculation of the ICC does not require the normality of the measurements and that the package allows the estimates to be made despite a varying number of conditions and images. The values of the ICC are between 0 and 1. The closer the ICC is to 1, the more similar the results are using AI under the same condition; however, the lower the ICC is (close to zero), the more the results of the AI diverge for the same condition. The central estimates for each ICC are accompanied by their 95% confidence interval, providing an estimate of the range in which the true value of the ICC could lie at the risk of a type I error of 5%. According to Shrout and Fleiss, an ICC greater than or equal to 0.8 supports excellent reliability; an ICC between 0.7 and 0.8 indicates good reliability. Values below 0.7 indicate moderate reliability at best [25,27].

### 3. Results

Dishes for which data was only available for one condition, or dishes for which measurements were not available for all conditions were excluded (exclusion of 20 dishes and 15% of all the pictures/situations, i.e., 4040 pictures). Thus, 22,544 analyzed situations for 149 dishes were considered for this reproducibility analysis. In total, based on the central ICC values obtained and the classification proposed by Shrout and Fleiss, the reliability of the AI results appeared to be excellent for 39% of the dishes ( $n = 58$  dishes) and good for 19% ( $n = 28$ ). The reliability appeared insufficient for 42% of the dishes (Figure 4). The ICCs for each dish are available in Appendix A (Table A1).



**Figure 4.** Flow chart of the design.

An additional analysis was performed for dishes for which the number of pictures was greater than or equal to 200 ( $n = 60$  dishes). The ICCs obtained were in favor of excellent reliability for 45% of the dishes selected ( $27/60$ ), of good reliability for 17% of the dishes ( $n = 10$ ) and for insufficient reliability for 38% of the dishes.

### 4. Discussion

In this paper, we report our testing of FoodIntech, which is a dietary portion assessment system that estimates the consumed portion size of a meal using images automatically captured by a standard smartphone. This study is a pioneer study including 149 different labels of dishes, with single or composite food items. We covered high volume images of 22,544 food items using automatic detection and AI. The images were taken in experimental conditions approaching routine clinical practice in a hospital environment: a smartphone was attached to a heating trolley used to distribute meal trays, the quantities of food were identical to those served to patients, and we used the same dishes, trays, crockery, and arrangement of dishes on the trays.

A review of the literature shows that there is has been real interest in using image-based dietary assessments in different situations: in free-living individuals [10,28], and in specific environments such as hospitals, [29–31], laboratories, [10,32] and cafeterias [33]. This method is often combined with food records or voice recording describing served and consumed meals [28,34], or associated with video recordings to be compared to weighed food records in free-living young adults [35]. However, even if the performance was found to have good reliability compared to traditional methods by some researchers, these image-based dietary assessments require the capture of images by participants and/or trained staff members to observe or calculate the individual's consumption [10,29,31,36,37].

The main advantage of AI is that human intervention is not required, and the results are obtained instantly. This is why the automatic detection of food in real-life contexts with data acquired by a wearable camera or smartphone in association with an AI analysis is challenging: image quality depends on variable factors such as appropriate lighting, image resolution, and limited blur caused by the movement of person taking the pictures. Many studies have demonstrated the validity of AI to automatically detect food items. To our knowledge, none have tested reproducibility on a number of complex foods or dishes as large as in our study. In addition, few studies have specified the number of analyzed images or the conditions under which the images were analyzed. Ji et al. [38] assessed the relative validity of an image-based dietary assessment app—Keenoa—and a three-day food diary in a sample of 102 healthy Canadian adults, but no information was given about the sample of images used and the authors showed that the validity of Keenoa was better at the group level than the individual level. In Fang et al. [39], the authors estimated food energy based on images and the generative adversarial network architecture. They validated the proposed method using data collected by 45 men and women between 21–65 years old, and obtained accurate food energy estimation with an average error of 209 kilocalories for eating occasion images (1875 paired images) collected from the Food in Focus study using the mobile food record. However, the authors specify the need to combine automatically detected food labels, segmentation masks, and contextual dietary information to further improve the accuracy of their food portion estimation system. Jia et al. [40] developed an artificial intelligence-based algorithm which can automatically detect food items from 1543 food-images acquired by an egocentric wearable camera, called the eButton. Even in the absence of reproducibility data, they reported that accuracy, sensitivity, specificity and precision were 98.7, 98.2, 99.2 and 99.2%, respectively for food images. In a recent publication, Lu et al. [18], proposed goFOODTM as a dietary assessment system based on AI to estimate the calorie and macronutrient content of a meal with food images captured by a smartphone. goFOODTM uses deep neural networks to process the two images and implements food detection, segmentation and recognition, while a 3D reconstruction algorithm estimates the food's volume. Thus, the calorie and macronutrient content was calculated from 319 fine-grain food categories, and the authors specified that the validation was performed using two multimedia databases containing non-standardized and fast food meals (one contained 80 central-European style meals and another one contained 20 fast-food type meals). However, in this study, there was no indication of reproducibility or the number of images analyzed. Other authors showed the error estimation of their system: Lu et al. [41] showed an estimation error of 15% with their system allowing a sequential semantic food segmentation and estimation of the volume of the consumed food with 322 meals. Sudo et al. [42] showed an estimation error close to 16.4% with their system presenting a novel algorithm that can estimate healthiness from meal images without requiring manual inputs.

It is still impossible for the system to recognize all the food categories in the world and in real-life, and this remains a limit for all existing applications. The limitation of our system is mainly determined by the training data.

We showed that 86 dishes were correctly detected by the system in repeatable conditions over an ICC value of 0.7/1. The 63 non-repeatable dishes (ICC < 0.7) show the limitations of the system: 10 were served in containers (cups or tubes), 18 had more than

200 pictures but were specially shaped food items or dishes (e.g., oranges or other fruits with skin leftover) and 35 were reproduced in less than 200 pictures. In order to address the limitations associated with food recognition, we will try to analyze their origins and the possible solutions:

- First of all, is the AI able to improve its results on the identified less detected food items and dishes? If we look precisely at the identified items that were the least reliable in the field of the study (for example yogurts or purées in containers) it appears that it was difficult to build a volumetric vision from a single sight view at  $90^\circ$ . It produces, with containers that are higher than wide, a drop shadow that leads to the difficult identification of height and volume. This issue could probably be improved by the use of new virtual 3D sensors embedded in recent market smartphones based on, for example, the Time of Flight (ToF) technology proposed by Samsung from the Galaxy S10 version. This technology calculates the speed of photons to access the surface of the target objects in order to construct a virtual depth vision of the objects in three dimensions. Additionally, some of the products might be transferred to glass or plastic containers in order to avoid the identified issues. Lo et al. [43] created an objective dietary assessment system based on a distinct neural network. They used a depth image, the whole 3D point cloud map and iterative closest point algorithms to improve dietary behavior management. They demonstrated that the proposed network architecture and point cloud completion algorithms can implicitly learn the 3D structures of various shapes and restore the occluded part of food items to allow better volume estimation.
- Secondly, is the AI able to improve its results on the fruit with a peel? Inedible leftovers are currently interpreted as the fruit itself in most cases. However, we are confident that the AI can learn to recognize these leftovers from the flesh of the fruit with new learning. In addition, we can code some specific rules to the transcoding overlay to help with the reproducibility of correct identifications.
- Finally, is the size of the sample enough, especially when we consider that some food items with more than 200 pictures in records had insufficient reliability? Some dishes, as well as some food containers, are more difficult to recognize and segment than others due to their shape, colors or ingredients, with the mixture on the plate making them more difficult to identify. Like human vision, computer vision has limitations that will never achieve 100% performance. However, just like humans, AI can improve the recognition of certain dishes or situations through a wider learning process and thus increase the number of reference pictures and segmentations. The method validated in this study, in particular, can obtain high-performance results for complex food items, allowing us to extrapolate that significant improvements could be obtained for dishes that are still poorly recognized or poorly reliable. Examples of these complex items include the “Colombo of veal with mangoes”, which had an ICC of 0.897 with 201 photos, the “Nicoise salad”, which had an ICC of 0.899 with 199 photos, and the “gourmet mixed salad”, which had an ICC of 0.949 with 198 photos.

## 5. Conclusions

These results were obtained using different learning steps, demonstrating that the method used to improve the recognition of dishes is effective, and that with a sufficient number of photos it can be extended to new dishes and foods. This study is, to our knowledge, one of the first to have tested such a large sample (22,544 images), and we obtained a high level of precision for more than half (57.8%) of a wide range of foods (149). To our knowledge, this method represents a paradigm shift in dietary assessment. AI technology can automatically detect foods with camera-acquired images, reducing both the burden of data processing and paper transcription errors. An additional benefit of AI is the ability to immediately analyze the data and obtain results. This study demonstrates that

this image recognition technique could be an exploitable clinical tool for monitoring the food intake of hospital patients.

## 6. Future Work

Future work will aim to show the validity and the usability of the FoodIntech artificial intelligence system for the evaluation of food consumption in hospitalized patients. The system will provide a measure of patient food consumption, which could then be associated with the patient's age, gender and body weight. Dietitians or physicians could use this information to adapt in-hospital menus to patient needs.

**Author Contributions:** Conceptualization, writing, visualization, obtaining funding, V.V.W.-D.; research, resource provision, data collection, editing and proofreading, F.B. and C.G.; validation, editing and proofreading, obtaining funding, C.J.; methodology, editing and proofreading, H.B.; methodology and statistical analysis, A.-K.S.; deep learning training and validation, methodology, O.B.; methodology, editing and proofreading, obtaining funding, M.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This study is part of FOODINTECH project funded by the European Funding for Regional Economic Development (FEDER), the "Grand Dijon", BPI France, the Regional Council of Burgundy France and the National Deposit and Consignment Fund. This study was also supported by a grant from the French Ministry of Health (PHRIP 2017).

**Institutional Review Board Statement:** Not applicable for studies not involving humans or animals.

**Informed Consent Statement:** Not applicable for studies not involving humans. Ethics approval was not needed because only food was measured and none of the data from patients were analyzed.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank Suzanne Rankin for editing the manuscript.

**Conflicts of Interest:** At the time of the study, the authors affiliated with ATOLcd and YUMAIN were working for these companies, who were the private partners of the project.

## Appendix A

**Table A1.** The Intra-class Correlation Coefficients (ICCs) for each dish studied.

Labelled Dishes in English	Labelled Dishes in French	ICC (95%CI)	Number of Images	p Values
Apple turnover	Chausson aux pommes	0.993 (0.985–0.998)	200	<0.001
breaded fish	poisson pané	0.991 (0.98–0.997)	200	<0.001
parsley potatoes	potatoes persillées	0.985 (0.952–0.998)	29	<0.001
Pear	poire	0.981 (0.94–0.997)	30	<0.001
chopped steak	steak haché	0.977 (0.951–0.993)	201	<0.001
Clamart rice	riz Clamart	0.972 (0.94–0.991)	199	<0.001
light cream baba	baba crème légère	0.968 (0.927–0.992)	201	<0.001
saithe fillet with saffron sauce	filet de lieu sauce safran	0.968 (0.9–0.995)	29	<0.001
potatoes	potatoes de terre	0.962 (0.92–0.988)	197	<0.001
Ratatouille	ratatouille	0.959 (0.914–0.987)	200	<0.001
green beans salad	haricots verts salade	0.958 (0.91–0.988)	200	<0.001
Flemish apples	potatoes flamande	0.957 (0.908–0.988)	200	<0.001
Gruyère Cream	Crème de gruyère	0.952 (0.882–0.992)	198	<0.001
Poultry Nuggets	nuggets de volaille	0.95 (0.897–0.985)	200	<0.001
Eastern pearl salad	salade mélangée gourmande	0.949 (0.895–0.984)	198	<0.001

Table A1. Cont.

Labelled Dishes in English	Labelled Dishes in French	ICC (95%CI)	Number of Images	<i>p</i> Values
Brussels sprouts	chou de Bruxelles	0.945 (0.836–0.991)	30	<0.001
Béchamel spinach	épinards béchamel	0.943 (0.883–0.982)	200	<0.001
Kiwi	kiwi	0.942 (0.826–0.99)	30	<0.001
quenelle with aurore sauce	quenelle sauce aurore	0.94 (0.877–0.982)	200	<0.001
veal fricassee	fricassé de veau	0.939 (0.872–0.983)	200	<0.001
couscous semolina	semoule couscous	0.938 (0.876–0.979)	225	<0.001
mortadelle	mortadelle	0.938 (0.873–0.981)	197	<0.001
Apple	pomme	0.936 (0.803–0.99)	26	<0.001
grated carrots	carottes râpées	0.931 (0.855–0.98)	201	<0.001
rustic lentils	lentilles paysanne	0.926 (0.85–0.977)	195	<0.001
rhubarb pie	tarte à la rhubarbe	0.924 (0.83–0.984)	201	<0.001
mashed broccoli	purée brocoli	0.923 (0.844–0.976)	201	<0.001
Hoki fillet sorrel sauce	filet de hoki sauce oseille	0.921 (0.841–0.975)	198	<0.001
troppezian pie	tropezienne	0.921 (0.675–0.985)	16	<0.001
parsley endive	endives persillées	0.91 (0.821–0.972)	199	<0.001
juice spinach	épinards au jus	0.906 (0.814–0.97)	201	<0.001
Nicoise salad	salade niàSoise	0.899 (0.801–0.968)	199	<0.001
Colombo of veal with mangoes	colombo de veau aux mangues	0.897 (0.785–0.973)	201	<0.001
Dijon Lentils	lentilles dijonnaise	0.892 (0.704–0.982)	30	<0.001
meal bread	pain repas	0.892 (0.69–0.982)	26	<0.001
Macedonia Mayonnaise	macédoine mayonnaise	0.89 (0.773–0.971)	201	<0.001
sautéed lamb	sauté d’agneau	0.875 (0.747–0.967)	201	<0.001
sautéed Ardéchoise	poêlée ardéchoise	0.87 (0.409–0.98)	13	0.0027
parsley youth carrots	carottes jeunes persillées	0.865 (0.743–0.956)	200	<0.001
lemon fish	poisson citron	0.865 (0.736–0.96)	200	<0.001
paella trim	garniture paëlla	0.864 (0.641–0.976)	30	<0.001
Italian dumplings	boulettes à l’italienne	0.862 (0.739–0.955)	192	<0.001
peas with juice	petits pois au jus	0.861 (0.75–0.948)	231	<0.001
salted plain yogurt pie	tarte au fromage blanc salée	0.86 (0.62–0.982)	30	<0.001
pasta salad	salade de pâtes	0.859 (0.738–0.954)	365	<0.001
Parisian pudding pie	tarte au flan parisien	0.854 (0.699–0.966)	176	<0.001
Natural yogurt	Yaourt nature	0.851 (0.623–0.964)	29	<0.001
turkey fricassee	fricassée de dinde	0.847 (0.603–0.973)	29	<0.001
colored pasta	pâtes de couleur	0.845 (0.711–0.949)	199	<0.001
blood sausage with apples	boudin aux pommes	0.835 (0.565–0.978)	29	<0.001
devilled chicken	poulet à la diable	0.829 (0.686–0.943)	200	<0.001
Braised celery heart	coeur de céleri braisé	0.829 (0.68–0.948)	201	<0.001
braised fennel	fenouil braisé	0.829 (0.68–0.948)	201	<0.001
donuts	beignets	0.825 (0.679–0.941)	197	<0.001
parsley carrots duet	duo de carottes persillées	0.822 (0.666–0.941)	109	<0.001

Table A1. Cont.

Labelled Dishes in English	Labelled Dishes in French	ICC (95%CI)	Number of Images	<i>p</i> Values
Bavarian apricot	bavarois abricot	0.82 (0.672–0.939)	200	<0.001
beef carbonnade	carbonnade de boeuf	0.81 (0.522–0.966)	27	<0.001
fish with mushroom sauce	poisson sauce champignons	0.805 (0.527–0.964)	30	<0.001
Mayonnaise tuna	thon mayonnaise	0.798 (0.639–0.931)	182	<0.001
fennel with basil	fenouil au basilic	0.798 (0.638–0.931)	183	<0.001
Tagliatelle with vegetables	tagliatelles aux légumes	0.791 (0.631–0.928)	201	<0.001
Sautéed Veal Marengo	sauté de veau marengo	0.789 (0.625–0.927)	182	<0.001
Natural fesh cheese	Fromage frais nature	0.788 (0.482–0.961)	27	<0.001
apricot pie	tarte aux abricots	0.784 (0.628–0.919)	229	<0.001
black forest cake	forêt noire	0.782 (0.155–0.965)	13	0.012
beef with 2 olives	boeuf aux 2 olives	0.78 (0.596–0.937)	201	<0.001
bread	pain	0.775 (0.587–0.919)	74	<0.001
pasta in shell	coquillettes	0.761 (0.449–0.955)	29	<0.001
flaky pastry with vanilla cream	feuilleté vanille	0.76 (0.435–0.966)	29	<0.001
Small grilled sausages	petites saucisses grillées	0.758 (0.584–0.915)	201	<0.001
chicken breast	escalope de poulet	0.755 (0.444–0.954)	30	<0.001
parsley ham	jambon persillé	0.755 (0.444–0.954)	30	<0.001
beef goulasch	goulasch de boeuf	0.749 (0.0784–0.96)	13	0.017
parsley chard	bettes persillées	0.748 (0.562–0.918)	200	<0.001
Assorted green vegetables	légumes verts assortis	0.746 (0.568–0.909)	199	<0.001
southern vegetables flan	flan de légumes du soleil	0.745 (0.534–0.935)	175	<0.001
roast chicken	poulet rôti	0.737 (0.566–0.899)	225	<0.001
juice lentils	lentilles au jus	0.736 (0.554–0.905)	197	<0.001
Vegetable bouquette	bouquetière de légumes	0.733 (0.551–0.904)	201	<0.001
leeks Vinaigrette	poireaux vinaigrette	0.716 (0.539–0.883)	162	<0.001
bow-tie pasta	papillons	0.715 (0.528–0.896)	198	<0.001
Dijon chicken cutlet	escalope de poulet dijonnaise	0.712 (0.492–0.906)	83	<0.001
melting apples	pommes fondantes	0.709 (0.521–0.893)	201	<0.001
beef sirloin	faux filet de boeuf	0.704 (0.515–0.891)	200	<0.001
rustic mix	mélange champêtre	0.701 (0.528–0.874)	240	<0.001
Green salad mixed	salade verte mélangée	0.701 (0.452–0.913)	60	<0.001
stuffed tomatoes	tomates farcies	0.693 (0.492–0.895)	190	<0.001
Pudding English Cream	pudding crème anglaise	0.682 (0.488–0.881)	201	<0.001
parsley celery root	céleri rave persillé	0.676 (0.481–0.878)	200	<0.001
sautéed vegetables	poelée de légumes	0.675 (0.33–0.934)	30	<0.001
chocolate meringue	meringue chocolatée	0.668 (0.471–0.874)	199	<0.001
lemon cake	cake citron	0.664 (0.458–0.882)	200	<0.001
penne	penne	0.663 (0.458–0.881)	200	<0.001
Parsley salsifis	salsifis persillés	0.66 (0.463–0.87)	198	<0.001
Coleslaw	salade coleslaw	0.652 (0.453–0.866)	199	<0.001

Table A1. Cont.

Labelled Dishes in English	Labelled Dishes in French	ICC (95%CI)	Number of Images	<i>p</i> Values
apricot flaky pie	tarte feuilletée aux abricots	0.652 (0.444–0.876)	201	<0.001
Cider ham	jambon au cidre	0.632 (0.431–0.856)	196	<0.001
fish with cream sauce	poisson sauce crème	0.601 (0.398–0.839)	198	<0.001
parsley wax beans	haricots beurre persillés	0.594 (0.383–0.848)	200	<0.001
Green cabbage Vinaigrette	chou vert vinaigrette	0.575 (0.38–0.812)	229	<0.001
gingerbread chicken	poulet au pain d'épices	0.571 (0.209–0.905)	30	<0.001
chicken with supreme sauce	poulet sauce supreme	0.562 (0.195–0.925)	30	<0.001
natural omelette	omelette nature	0.535 (0.327–0.802)	160	<0.001
parsley cauliflower	chou fleur persillé	0.534 (0.342–0.796)	398	<0.001
orange	orange	0.48 (0.3–0.734)	283	<0.001
Cooked and dry sausage	saucisson cuit et sec	0.479 (0.282–0.763)	199	<0.001
sautéed deer	sauté de biche	0.479 (0.282–0.763)	199	<0.001
Fresh celery remould	céleri frais rémoulade	0.478 (0.252–0.773)	94	<0.001
bolognese trim	garniture bolognaise	0.457 (0.244–0.756)	113	<0.001
couscous vegetables	légumes couscous	0.438 (0.23–0.774)	163	<0.001
Comté slice	Comté portion	0.433 (0.25–0.714)	214	<0.001
papet from Jura	papet jurassien	0.429 (0.24–0.726)	201	<0.001
spaghetti	spaghetti	0.426 (0.0793–0.852)	30	0.0053
hard boiled egg with Mornay sauce	oeuf dur sce mornay	0.406 (0.211–0.714)	139	<0.001
Cheese Pie	tarte au fromage	0.386 (0.206–0.691)	200	<0.001
Coffee mousse	Mousse café	0.382 (0.203–0.688)	200	<0.001
cauliflower flan	flan de chou fleur	0.371 (0.19–0.699)	196	<0.001
farmer pâté	pâté de campagne	0.338 (0.176–0.63)	213	<0.001
parsley green beans	haricots verts persillés	0.337 (0.186–0.62)	398	<0.001
raspberry pie	tarte aux framboises	0.323 (0.157–0.655)	201	<0.001
Fig pastry with vanilla cream	Figue	0.318 (0.157–0.628)	200	<0.001
Camembert slice	Camembert portion	0.3 (0.0733–0.761)	57	0.001
cauliflower salad	chou fleur en salade	0.29 (0.144–0.579)	229	<0.001
Clementine	clémentine	0.289 (0.0352–0.767)	42	0.009
Hedgehog	hérisson	0.235 (0.102–0.536)	201	<0.001
raspberry pastry	framboisier	0.218 (0.0872–0.52)	175	<0.001
cheese omelette	omelette au fromage	0.213 (0.0785–0.519)	149	<0.001
Strasbourg salad	salade strasbourgeoise	0.198 (0.0515–0.519)	104	<0.001
choux pastry with whipped cream	chou chantilly	0.186 (0.0294–0.708)	110	0.0038
choux pastry with vanilla cream	chou vanille	0.177 (0.0648–0.459)	191	<0.001
Applesauce	Compote de pommes	0.169 (0.0573–0.497)	199	<0.001
salted cake	cake salé	0.168 (0.061–0.444)	200	<0.001
Liège coffee	Café liégeois	0.131 (0.0401–0.386)	200	<0.001
liver pâté	pâté de foie	0.129 (–0.112–0.671)	29	0.18
Chocolate flan	Flan chocolat	0.099 (0.0224–0.329)	199	0.0014

Table A1. Cont.

Labelled Dishes in English	Labelled Dishes in French	ICC (95%CI)	Number of Images	p Values
Flavored yogurt	Yaourt aromatisé	0.0929 (0.0221–0.286)	230	0.0013
chocolate eclair	éclair chocolat	0.0756 (−0.109–0.665)	30	0.24
milk chocolate mousse	Mousse chocolat lait	0.0637 (0.00361–0.259)	198	0.016
clafoutis with cherries	clafoutis aux cerises +	0.0527 (−0.17–0.779)	21	0.33
20% fat plain yogurt	Fromage blanc 20%	0.0521 (−0.00104–0.267)	199	0.029
Garlic and herbs cheese	fromage Ail et fines herbes	0.0167 (−0.0173–0.192)	200	0.21
pear pie	tarte aux poires	0.00198 (−0.0444–0.163)	129	0.42
endive with ham	endives au jambon	0	11	1
Hoki fillet Crustacean sauce	filet de hoki sauce crustacés	0	4	1
mackerel fillet	filets de maquereaux	0	4	1
Semolina Cake	Gâteau de semoule	0	27	0.56
Lemon mousse	Mousse citron	0	29	0.55

## Appendix B. Details of the Learning Parameters Using Mask-RCNN Deep Learning Algorithm

We used default `mask_rcnn_inception_resnet_v2_atrous_coco` learning parameters, except for the learning rate scheduler, which we set to cosine decay scheduling instead of the default manual step scheduling. The cosine decay parameters are set as follows: `total_steps` to 200,000 iterations, `warmup_steps` to 50,000 iterations, `warmup_learning_rate` to 0.0001 and `learning_rate_base`: 0.002.

## References

- Allison, S.; Stanga, Z. Basics in clinical nutrition: Organization and legal aspects of nutritional care. *e-SPEN Eur. e-J. Clin. Nutr. Metab.* **2009**, *4*, e14–e16. [[CrossRef](#)]
- Williams, P.; Walton, K. Plate waste in hospitals and strategies for change. *e-SPEN Eur. e-J. Clin. Nutr. Metab.* **2011**, *6*, e235–e241. [[CrossRef](#)]
- Schindler, K.; Pernicka, E.; Laviano, A.; Howard, P.; Schütz, T.; Bauer, P.; Grecu, I.; Jonkers, C.; Kondrup, J.; Ljungqvist, O.; et al. How nutritional risk is assessed and managed in European hospitals: A survey of 21,007 patients findings from the 2007–2008 cross-sectional nutritionDay survey. *Clin. Nutr.* **2010**, *29*, 552–559. [[CrossRef](#)] [[PubMed](#)]
- Björnsdóttir, R.; Óskarsdóttir, E.S.; Thordardóttir, F.R.; Ramel, A.; Thorsdóttir, I.; Gunnarsdóttir, I. Validation of a plate diagram sheet for estimation of energy and protein intake in hospitalized patients. *Clin. Nutr.* **2013**, *32*, 746–751. [[CrossRef](#)] [[PubMed](#)]
- Kawasaki, Y.; Sakai, M.; Nishimura, K.; Fujiwara, K.; Fujisaki, K.; Shimpo, M.; Akamatsu, R. Criterion validity of the visual estimation method for determining patients' meal intake in a community hospital. *Clin. Nutr.* **2016**, *35*, 1543–1549. [[CrossRef](#)]
- Amano, N.; Nakamura, T. Accuracy of the visual estimation method as a predictor of food intake in Alzheimer's patients provided with different types of food. *Clin. Nutr. ESPEN* **2018**, *23*, 122–128. [[CrossRef](#)] [[PubMed](#)]
- Palmer, M.; Miller, K.; Noble, S. The accuracy of food intake charts completed by nursing staff as part of usual care when no additional training in completing intake tools is provided. *Clin. Nutr.* **2015**, *34*, 761–766. [[CrossRef](#)]
- Gibson, R.S. *Principles of Nutritional Assessment*; Oxford University Press: Oxford, UK, 2005; ISBN 978-0-19-517169-3.
- Lee, R.D.; Nieman, D.C. *Nutritional Assessment*; Mosby: St. Louis, MO, USA, 1996; ISBN 9780815153191.9.
- Martin, C.K.; Correa, J.; Han, H.; Allen, H.R.; Rood, J.C.; Champagne, C.M.; Gunturk, B.; Bray, G.A. Validity of the Remote Food Photography Method (RFPM) for Estimating Energy and Nutrient Intake in Near Real-Time. *Obesity* **2012**, *20*, 891–899. [[CrossRef](#)] [[PubMed](#)]
- Martin, C.K.; Han, H.; Coulon, S.M.; Allen, H.R.; Champagne, C.M.; Anton, S.D. A novel method to remotely measure food intake of free-living individuals in real time: The remote food photography method. *Br. J. Nutr.* **2009**, *101*, 446–456. [[CrossRef](#)]

12. Martin, C.K.; Nicklas, T.; Gunturk, B.; Correa, J.; Allen, H.R.; Champagne, C. Measuring food intake with digital photography. *J. Hum. Nutr. Diet.* **2014**, *27*, 72–81. [[CrossRef](#)] [[PubMed](#)]
13. McClung, H.L.; Champagne, C.M.; Allen, H.R.; McGraw, S.M.; Young, A.J.; Montain, S.J.; Crombie, A.P. Digital food photography technology improves efficiency and feasibility of dietary intake assessments in large populations eating ad libitum in collective dining facilities. *Appetite* **2017**, *116*, 389–394. [[CrossRef](#)] [[PubMed](#)]
14. Hinton, E.C.; Brunstrom, J.M.; Fay, S.H.; Wilkinson, L.L.; Ferriday, D.; Rogers, P.J.; de Wijk, R. Using photography in ‘The Restaurant of the Future’. A useful way to assess portion selection and plate cleaning? *Appetite* **2013**, *63*, 31–35. [[CrossRef](#)]
15. Swanson, M. Digital Photography as a Tool to Measure School Cafeteria Consumption. *J. Sch. Health* **2008**, *78*, 432–437. [[CrossRef](#)] [[PubMed](#)]
16. Pouyet, V.; Cuvelier, G.; Benattar, L.; Giboreau, A. A photographic method to measure food item intake. Validation in geriatric institutions. *Appetite* **2015**, *84*, 11–19. [[CrossRef](#)]
17. Sak, J.; Suchodolska, M. Artificial Intelligence in Nutrients Science Research: A Review. *Nutrients* **2021**, *13*, 322. [[CrossRef](#)] [[PubMed](#)]
18. Lu, Y.; Stathopoulou, T.; Vasiloglou, M.F.; Pinault, L.F.; Kiley, C.; Spanakis, E.K.; Mougiakakou, S. goFOOD™: An Artificial Intelligence System for Dietary Assessment. *Sensors* **2020**, *20*, 4283. [[CrossRef](#)]
19. Ronneberger, O.; Fischer, P.; Brox, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*; Springer: Cham, Switzerland, 2015; pp. 234–241.
20. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. *arXiv* **2018**, arXiv:1703.06870. Available online: <https://arxiv.org/pdf/1703.06870.pdf> (accessed on 20 December 2021).
21. Vuola, A.O.; Akram, S.U.; Kannala, J. Mask-RCNN and U-Net Ensembled for Nuclei Segmentation. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging, ISBI 2019, Venice, Italy, 8–11 April 2019; Computer Science: Venice, Italy, 2019; pp. 208–212. Available online: <https://arxiv.org/pdf/1901.10170.pdf> (accessed on 20 December 2021). [[CrossRef](#)]
22. Welcome to the Model Garden for TensorFlow; Tensorflow. 2021. Available online: <https://github.com/tensorflow/models/> (accessed on 17 December 2021).
23. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A.A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. In Proceedings of the 31st AAAI Conference on Artificial Intelligence, AAAI 2017, San Francisco, CA, USA, 4–9 February 2017; pp. 4278–4284. Available online: <http://www.cs.cmu.edu/~jjeanoh/16-785/papers/szegedy-aaai2017-inception-v4.pdf> (accessed on 20 December 2021).
24. Walter, S.D.; Eliasziw, M.; Donner, A. Sample size and optimal designs for reliability studies. *Stat. Med.* **1998**, *17*, 101–110. [[CrossRef](#)]
25. Shrout, P.E.; Fleiss, J.L. Intraclass correlations: Uses in assessing rater reliability. *Psychol. Bull.* **1979**, *86*, 420–428. [[CrossRef](#)]
26. Koo, T.K.; Li, M.Y. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J. Chiropr. Med.* **2016**, *15*, 155–163. [[CrossRef](#)]
27. Lebreton, J.M.; Senter, J.L. Answers to 20 questions about interrater reliability and interrater agreement. *Organ. Res. Methods* **2008**, *11*, 815–852. [[CrossRef](#)]
28. Rollo, M.E.; Ash, S.; Lyons-Wall, P.; Russell, A.W. Evaluation of a Mobile Phone Image-Based Dietary Assessment Method in Adults with Type 2 Diabetes. *Nutrients* **2015**, *7*, 4897–4910. [[CrossRef](#)]
29. Sullivan, S.C.; Bopp, M.M.; Roberson, P.K.; Lensing, S.; Sullivan, D.H. Evaluation of an Innovative Method for Calculating Energy Intake of Hospitalized Patients. *Nutrients* **2016**, *8*, 557. [[CrossRef](#)]
30. Monacelli, F.; Sartini, M.; Bassoli, V.; Becchetti, D.; Biagini, A.L.; Nencioni, A.; Cea, M.; Borghi, R.; Torre, F.; Odetti, P. Validation of the photography method for nutritional intake assessment in hospitalized elderly subjects. *J. Nutr. Health Aging* **2017**, *21*, 614–621. [[CrossRef](#)]
31. Winzer, E.; Luger, M.; Schindler, K. Using digital photography in a clinical setting: A valid, accurate, and applicable method to assess food intake. *Eur. J. Clin. Nutr.* **2018**, *72*, 879–887. [[CrossRef](#)]
32. Jia, W.; Chen, H.-C.; Yue, Y.; Li, Z.; Fernstrom, J.; Bai, Y.; Li, C.; Sun, M. Accuracy of food portion size estimation from digital pictures acquired by a chest-worn camera. *Public Health Nutr.* **2014**, *17*, 1671–1681. [[CrossRef](#)]
33. Williamson, D.A.; Allen, H.R.; Martin, P.D.; Alfonso, A.J.; Gerald, B.; Hunt, A. Comparison of digital photography to weighed and visual estimation of portion sizes. *J. Am. Diet. Assoc.* **2003**, *103*, 1139–1145. [[CrossRef](#)]
34. Casperson, S.L.; Sieling, J.; Moon, J.; Johnson, L.; Roemmich, J.N.; Whigham, L.; Yoder, A.B.; Hingle, M.; Bruening, M. A Mobile Phone Food Record App to Digitally Capture Dietary Intake for Adolescents in a Free-Living Environment: Usability Study. *JMIR mHealth uHealth* **2015**, *3*, e30. [[CrossRef](#)] [[PubMed](#)]
35. Naaman, R.; Parrett, A.; Bashawri, D.; Campo, I.; Fleming, K.; Nichols, B.; Burleigh, E.; Murtagh, J.; Reid, J.; Gerasimidis, K. Assessment of Dietary Intake Using Food Photography and Video Recording in Free-Living Young Adults: A Comparative Study. *J. Acad. Nutr. Diet.* **2021**, *121*, 749–761.e1. [[CrossRef](#)]
36. Boushey, C.J.; Spoden, M.; Zhu, F.; Delp, E.; Kerr, D.A. New mobile methods for dietary assessment: Review of image-assisted and image-based dietary assessment methods. In Proceedings of the Nutrition Society; Cambridge University Press: Cambridge, UK, 2017; Volume 76, pp. 283–294. [[CrossRef](#)]
37. Saeki, K.; Otaki, N.; Kitagawa, M.; Tone, N.; Takachi, R.; Ishizuka, R.; Kurumatani, N.; Obayashi, K. Development and validation of nutrient estimates based on a food-photographic record in Japan. *Nutr. J.* **2020**, *19*, 1–13. [[CrossRef](#)] [[PubMed](#)]

38. Ji, Y.; Plourde, H.; Bouzo, V.; Kilgour, R.D.; Cohen, T.R. 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*, e16953. [[CrossRef](#)] [[PubMed](#)]
39. Fang, S.; Shao, Z.; Kerr, D.A.; Boushey, C.J.; Zhu, F. An End-to-End Image-Based Automatic Food Energy Estimation Technique Based on Learned Energy Distribution Images: Protocol and Methodology. *Nutrients* **2019**, *11*, 877. [[CrossRef](#)] [[PubMed](#)]
40. Jia, W.; Li, Y.; Qu, R.; Baranowski, T.; E Burke, L.; Zhang, H.; Bai, Y.; Mancino, J.M.; Xu, G.; Mao, Z.-H.; et al. Automatic food detection in egocentric images using artificial intelligence technology. *Public Health Nutr.* **2018**, *22*, 1–12. [[CrossRef](#)] [[PubMed](#)]
41. Lu, Y.; Stathopoulou, T.; Vasiloglou, M.F.; Christodoulidis, S.; Blum, B.; Walser, T.; Meier, V.; Stanga, Z.; Mougiakakou, S. An artificial intelligence-based system for nutrient intake assessment of hospitalised patients. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* **2019**, 5696–5699. [[CrossRef](#)]
42. Sudo, K.; Murasaki, K.; Kinebuchi, T.; Kimura, S.; Waki, K. Machine Learning–Based Screening of Healthy Meals from Image Analysis: System Development and Pilot Study. *JMIR Form. Res.* **2020**, *4*, e18507. [[CrossRef](#)] [[PubMed](#)]
43. Lo, F.P.-W.; Sun, Y.; Qiu, J.; Lo, B. Food Volume Estimation Based on Deep Learning View Synthesis from a Single Depth Map. *Nutrients* **2018**, *10*, 2005. [[CrossRef](#)]