

Vision-Based Methods for Food and Fluid Intake Monitoring

Subjects: Engineering, Biomedical | Nursing

Contributor: Xin Chen, Ernest N. Kamavuako

Food and fluid intake monitoring are essential for reducing the risk of dehydration, malnutrition, and obesity. The existing research has been preponderantly focused on dietary monitoring, while fluid intake monitoring, on the other hand, is often neglected. Food and fluid intake monitoring can be based on wearable sensors, environmental sensors, smart containers, and the collaborative use of multiple sensors. Vision-based intake monitoring methods have been widely exploited with the development of visual devices and computer vision algorithms. Vision-based methods provide non-intrusive solutions for monitoring. They have shown promising performance in food/beverage recognition and segmentation, human intake action detection and classification, and food volume/fluid amount estimation. However, occlusion, privacy, computational efficiency, and practicality pose significant challenges.

Keywords: intake monitoring ; drinking action detection ; dietary monitoring

1. Introduction

Maintaining healthy food intake and adequate hydration is significant for humans' physiological and physical health ^{[1][2][3]}.

The quality of food intake was proven to be associated with the metabolic function of the human body ^[4]. Unbalanced nutrition intake increases the risk of many diseases, including diabetes, obesity, cardiovascular disease, and certain cancers ^{[1][5]}. When understanding human body dynamics associated with underweight, overweight, and obesity, it is important to objectively assess energy intake (EI); energy intake assessment is related to food type recognition, amount consumed estimation, and portion size estimation ^[6]. Being underweight can result from energy expenditure exceeding energy intake over an extended period, which leads to health risks such as malnutrition and premature death ^[7]. Being overweight and obesity are associated with energy intake exceeding energy expenditure, leading to chronic diseases such as type 2 diabetes, cardiovascular diseases, cancers, and musculoskeletal disorders ^{[6][7][8]}. A dietary assessment system could be used to monitor daily food intake and control eating habits by triggering a just-in-time intervention during energy intake to prevent health issues ^[9].

Low-intake dehydration, caused by inadequate fluid intake, has endangered public health and is often underemphasised ^{[9][10]}. Mild dehydration happens commonly among people and increases the risk of chronic diseases ^{[11][12]}. A notable example is a significant association between urolithiasis (kidney stone) and low daily water intake ^{[4][5]}. Furthermore, dehydration is closely associated with disability, hospitalisation and mortality ^[13] in hospitals ^{[14][15][16][17]} and long-term care systems ^{[10][11][13][18]}. In the hydration and outcome in older patients (HOOP) prospective cohort study of 200 older adults in a large UK teaching hospital, 37% of the participants admitted as emergencies were dehydrated, 7% of the participants died in the hospital, and 79% of those who died were dehydrated at admission ^[15]. Dehydration and drinking status could also be related to children's and adults' attention and memory performance ^[19].

2. Overview of Vision-Based Intake Monitoring

2.1. Active and Passive Methods

In vision-based methods, there are two approaches to capturing images: active and passive ^{[20][21]}. Active methods require the user to take pictures and record their intake manually, while passive methods automatically access the food or fluid intake information. Active methods are widely used in practice. Traditionally, active food intake monitoring was in the form of food records, recalls, or questionnaires ^[22]. For active fluid intake monitoring, a fluid balance chart is used as a self-reporting tool to identify a positive (fluid input higher than output) or negative (fluid output higher than input) balance in hospitals or nursing homes ^{[18][23]}. A fluid balance chart includes information on the time, approach, and amount of body fluid input (oral, intravenous, etc.) and output (urine, tube, etc.), which can be completed by trained nurses, doctors, or patients themselves ^[23].

With the development of cameras, images of meals and drinks are more commonly used for dietary monitoring. In visual-based monitoring methods, active methods are not as widely seen as passive methods and mostly rely on mobile phone cameras. For example, ^[24] proposed a food and nutrition measurement approach by analysing the images taken by the

users before and after a meal, which provided up to 97.2% of correct classification for food type and 1% of misreported nutrient information.

Nevertheless, manually recording the intake information by writing, weighing, or triggering a camera can be time-consuming and burdensome [24][25], hence not ideal for daily application. Moreover, self-reporting is not an option for patients with actional difficulties or the elderly with cognition degeneration. Therefore, in recent years, passive sensing methods with different devices and automatic strategies predominated over traditional active methods.

2.2. Environmental Settings

The environmental settings were categorised into the free-living environment, pseudo-free-living environment, laboratory environment, and others. In free-living-based studies, the systems were assessed with sufficient data collected by sensors configured into the user's natural living environment. The pseudo-free-living environment tried to replicate the user's natural living environment in a laboratory. The controlled laboratory environment only covers the specific actions needed as input data for the system (e.g., biting or drinking). In contrast, the others include methods based on existing datasets without considering the experimental environment. A camera's viewing angle can be either first-person or third-person. It is noticed that most of the third-person methods were only considered in a controlled testing environment, not testified in a free-living scenario. As for the first-person camera, in free-living, the highest accuracy achieved on food and non-food classification is 95% [26][27], and on eating action detection [28] was 89.68%. One recent free-living research reached the F1-score of 89% on drinking and eating episode detection but only 56.7% precision at the fluid intake level estimation [29]. Therefore, there is still a gap in harnessing cameras in the free-living scenario. Facts identified which can affect the performance include unstable light conditions, occlusion, low framerate, and motion blur.

2.3. Privacy Issue

Most papers investigating vision-based monitoring failed to discuss privacy issues even though some of the concerns were evident with the participant's face and body shown in the figures in the paper [30][31][32][33][34][35]. In active methods, cameras can be controlled manually to avoid taking images with privacy concerns, which is inconvenient and labour-demanding [36][37]. Another approach seen in both active and passive methods was reviewing the photos after they were taken and deleting the ones with privacy concerns, which could also be time-consuming and burdensome [28][38]. Hence, passive methods with approaches to eliminate privacy issues were considered the most. Passive methods are more likely to face privacy concerns because most images captured passively are not related to food or drink consumption only [20]. Therefore, in some designs, the intake action was detected by a smart watch or glasses, and the camera was only turned on when the eating or drinking episode was highly probable [20][29][39]. In the survey of privacy concerns for users of AIM-2, the average level of concerns was reduced from 5.0 to 1.9 when images were captured only during intake action rather than continuously [20]. However, this method requires the users to wear multiple sensors, which can be cosmetically unpleasant, uncomfortable, and intrusive, especially for the elderly or groups with disabilities [37].

3. Viewing Angles and Devices in Monitoring Systems

In passive vision-based intake monitoring, a camera's viewing angle can be either first-person or third-person. A first-person camera (egocentric camera/wearable camera) is typically attached to the human body, pointing out at the food or container. In contrast, a third-person camera is mounted in the living environment, pointing at the subject. In the included papers that proposed a monitoring system or nutrition log application, 49 studies were based on first-person cameras, 39 on external third-person cameras, and 28 took advantage of the users' smartphones. What is worth mentioning is that most of the phone-based methods were active, meaning the users needed to take the food/drink picture manually, e.g., the food/nutrition/dietary log proposed in [40][41][42].

There were four tasks identified in intake monitoring methods: binary classification, to distinguish food/drink intake from other activities; food/drink type classification, to detect the type of items consumed; food/fluid amount estimation, crucially related to energy intake; and intake action recognition, to recognize the human body movement. In the binary classification task, elements such as fingers, hands, containers, cutleries, and food can be detected, and different criteria can be set and followed as an indication of an intake activity.

Regarding the placement of devices, first-person cameras could be in the form of glasses, watch and pendant, while third-person cameras could be mounted on the ceiling for a top-down view or placed around the subjects. The selected devices vary from studies, and cameras seen in the existing papers were mainly RGB and RGB-D cameras. RGB cameras have been spotted and used in combination with other non-vision sensors. The pattern of viewing angles and devices found in the papers is shown in **Figure 1**.

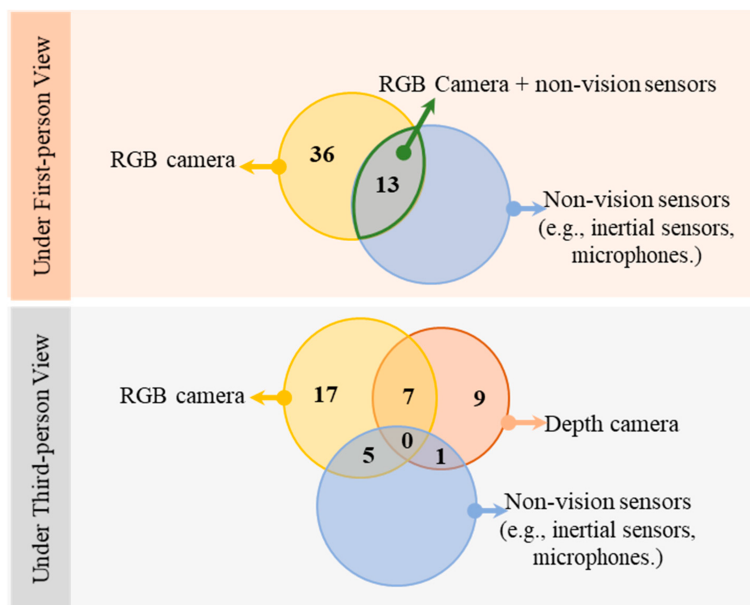


Figure 1. Pattern of the viewing angles and devices selected.

From the pattern of device selection shown in **Figure 1**, it is evident that the RGB cameras were the most used, primarily as first-person. In contrast, Depth cameras were not used as first-person cameras and were also barely used collaboratively with non-vision sensors. Moreover, there has been no system that covered all three sensors: RGB camera, Depth camera, and any of the non-vision sensors.

3.1. First-Person Approaches

As shown in **Figure 1**, of the 49 first-person methods, 36 relied on RGB information alone. The remaining 13 used RGB cameras collaboratively with other non-vision sensors, including accelerometers [20][43][44][45], gyroscopes, flex sensors [20][43], load cells [46], proximity sensors and IMUs [29]. The most common technology setting is an inertial-smart watch and a wearable RGB camera [26][27][39][47][48]. For example, Annapurna [26][27][48] is a smartwatch with a built-in camera proposed for autonomous food recording. The inertial sensor of the watch was used for gesture recognition to identify the eating activity, and then the camera took only pictures which were likely to be useful. Thus, compared to methods with a camera constantly in operation, redundant images were reduced, and so was storage requirement, privacy issue, computation, and camera power consumption. However, one fundamental problem with an inertial smartwatch is that the intake action could be missed when the user is drinking with a straw or using a hand that is not wearing the watch. Unlike the approaches mentioned above, which mainly focus on food intake detection was an intake monitoring system for fluid combining glasses, smartwatches, and phones [39].

Smart glass is another form of wearable device. Automatic Ingestion Monitor Version 2 (AIM-2) [20] was proposed with an accelerometer, a flex sensor and an RGB camera. However, in this design, images captured by the camera were only for validating the performance of other wearable sensors on intake detection; no visual methods were considered. FitByte [29] was a glasses-based diet monitoring system that applied six IMUs for chewing and swallowing detection, a proximity sensor for hand-to-mouth gesture detection, and an RGB camera pointing downward to capture potential food images. Both eating and drinking episodes were detected in this design. However, only 56.7% precision was achieved in fluid intake detection, while it was 92.8% for food intake. This means that Fitbyte performed significantly lower in fluid intake detection because it was only sensitive in simple and continuous drinking scenarios, other than short sips or when drinking happens with other irrelevant activities.

Assessments have been made on the efficiency of first-person cameras for dietary monitoring. Thomaz et al. (2013) proposed and evaluated a dietary monitoring system based on a neck-worn and human computation. Images were taken by the camera every thirty seconds and sent on Amazon's Mechanical Turk (AMT) (a platform providing human intelligence labour) for identifying food by human labour. This design resulted in 89.68% accuracy in identifying eating activities [28].

The feasibility evaluations mentioned above revealed the limitations of utilising first-person cameras for passive dietary monitoring. What came first was the occlusion of view. For example, if the image did not provide a complete observation of food, the estimation accuracy of portion size could be low [49]. The uncertainty of wear time, battery sustainability and noncompliance in wearing the camera were other problems, especially when faced with the elderly or patients with cognition decline. As for image acquisition, dark and blurry images obtained in poor light conditions could make classification difficult.

In summary, the common system architecture of first-person methods was combining one first-person RGB camera with other sensors. Cameras can be placed in the form of smartwatches [26][27], glasses [20][29], or even caps [47]. Combining

cameras with other sensors can reduce the energy consumption of cameras, extend the use time of batteries, save storage space and rule out privacy concerns by turning the camera on only when a candidate movement is detected [26][27][48]. One fundamental limitation of inertial smartwatches is that the intake action could be missed when the user drinks with a straw or uses the contralateral hand with no watch. The incontinence of wearable devices was another limitation.

3.2. Third-Person Approaches

Compared to first-person cameras, third-person cameras have the advantage of being non-intrusive to the user's [50]. The placement of cameras is one of the primary issues to consider. Most research had only one position of a single camera, placed on the ceiling for a top-down view [51][52][53], or placed pointing at the subject with a fixed distance from 0.6 m to 2 m [21][30][31][54][55]. Multiple cameras could be placed around for different viewing angles and complement the possible occlusion to achieve a more robust system [32][56][57].

In third-person methods, RGB and depth information can be used individually or collaboratively for action detection. Specifically, 17 papers used RGB information only, nine were with depth information from an RGB-D camera, and seven were based on the fusion of RGB and depth information, as seen in **Figure 2**. Unlike first-person cameras, non-vision sensors are used less frequently with third-person cameras. The main reason was that the kinematic information or distance information provided by IMUs and proximity sensors could also be obtained from the visual information of the third-person camera [30][55].

Microsoft Kinect was dominantly adopted in existing research, which can work day and night with the infrared sensor generating the depth images, and the skeleton tracking tool kit providing the joint coordinates [55][58]. The effectiveness of MS Kinect was tested for detecting the eating behaviour of older adults by placing the camera in front of the subject, resulting in an average of 89% success rate [30].

Regarding reducing privacy and image data concerns, some studies only used depth information from RGB-D cameras. For example, Kinect skeletal tracking was used for counting bites by tracking the jaw face point [59], and wrist roll joint of users based on depth information, achieving an overall accuracy of around 94% [55]. A system with an average accuracy of 96.2% was proposed, relying on the depth information of wrist joint and elbow joint motion obtained by a Kinect camera. However, although the study was presented for free-living calorie intake monitoring, only one camera position was tested, and no occlusion problem was considered [60]. The fusion of depth information and RGB information was another option, with the depth information for skeleton definition and body movement tracking while the RGB data for specific intake-related object detection [52].

RGB cameras were also popular devices in intake monitoring used as third-person cameras. It can be embedded in the ceiling, pointing down [53], or put on the dining table, pointing at the subject [31]. The fusion of RGB and depth information has the potential to reach higher accuracy compared to using a single modality of information. An example can be seen in [50], where an adapted version of the self-organized map algorithm was applied to the skeleton model obtained from depth information for movement tracking. The RGB stream was for recognising eating-related items such as glass. This method achieved a 98.3% overall accuracy. All RGB-D cameras were used as third-person cameras (as seen in **Figure 1**).

4. Algorithms by Task

4.1. Binary Classification

Eliminating unrelated images was a preliminary step for identifying candidate intake activities. This was commonly proposed as a binary classification approach to distinguish food/drink from other objects or to classify low-quality images and delete them. For example, to distinguish sharp images from blurry images for adequate image quality, Fast-Fourier Transform (FFT) for images was calculated to analyse the sharpness, resulting in a 10–15% misclassification [43].

Im2 Calories was a food intake monitoring system proposed in 2015, where a GoogLeNet CNN was trained with a modified Food101 dataset. One of the tasks for Im2 Calories was to determine whether the image was related to a meal, at which an accuracy of 99.02% was achieved [61].

The GoogLeNet in Im2 Calories [61] was tuned on a Titan X GPU with 12 GB memory; then implemented into an Android APP less than 40 MB, which could classify an image within one second. iLog could also operate on edge-level, low-performance computing paradigms, such as mobile phones, sensors and single-board computers [62]. Apart from the networks mentioned above, for real-time and portable monitoring, a derived MobileNet was proposed and implemented into a Cortex-M7 microcontroller for dietary image capturing, which achieved an average precision of 82% in identifying food-related images [63]. The training was conducted on Google Colab using 400 food images and 400 non-food images, taking up to 5.5 h, while only 761.99 KB of flash and 501.76 KB of RAM were needed to implement this algorithm.

Annapurna was a multimodel system with a camera mounted on an inertial smartwatch for dietary recording [26][27]. In this design, the camera was only switched on when the watch detected intake action. A mobile phone was first used as a lightweight computing platform to eliminate images with human faces and blurred edges. Then, 37% of the remaining images with food items in them were transferred to a server for further processing, where the Clarifai API was used to

identify the presence of food items in pictures based on CNN, and a depth map was created to detect food too far from the camera (considered as unrelated to the meal). As a result, 95% of the meals could be recalled by the proposed system in a free-living environment. The computation was firstly on mobile phones for Annapurna to remove blank, blurry, and misleading images to reduce runtime for further computing. However, the latency was around 0.9 s for the smartwatch to capture an image, which limited the response speed of the whole system [27].

The server used in Annapurna [26][27], Clarifai API, was also used in [37], where it generated tag outputs (e.g., 'food', 'car', 'dish'.) of an input image for determining whether the image was food-related. This method was tested on both Food-5K and e-Button and reached the specificity of 87% on Food-5K (created in [64]), higher than the results on e-Button. This was because e-Button was an egocentric free-living dataset with 17.7% blurred images, complex backgrounds, and more diverse objects. According to the authors, although the burden of manually observing and recording dietary activities in previous work [65] was reduced, the effectiveness of automatic monitoring was still limited due to the quality of the captured images.

Only limited papers addressed the binary classification of fluid/drink/beverage. An example covering food and fluid was [66], which trained a YOLOv5 network to detect and localize food and beverage items from other objects. The study aimed to distinguish food and beverages from other objects and added 'screen' and 'person' as extra classes. As a result, an overall mean average precision of 80.6% was achieved for classifying these four objects, which was still far from being used in practice. NutriNet was another deep Neural Network proposed for both food and beverage; the detection model's output was either 'food/drink' or 'other' [67].

4.2. Food/Drink Type Classification

ML methods included support vector machines (SVM), principal component analysis (PCA) [62][68][69], K-means classifiers, random forest (RF), fully connected neural networks (NN), artificial neural networks (ANN) [70], and some image matching and retrieving methods such as content-based image retrieval (CBIR) [71], dynamic time wrapping (DTW) [72], bag of features (BoF) [33][73][74], which clusters the features into visual words. Features of the image could be extracted by methods including speeded-up robust features (SURF) and scale-invariant feature transform (SIFT). In these methods, SVM was most seen and often used collaboratively with other DL methods. Different networks were seen in DL methods, such as GoogLeNet [61][64][67][70][72][75][76][77][78], MobileNetV2 [68][69], AlexNet [62][75][76], Inception-V3 [72][78][79][80][81], NutriNet [67][76][82], K-foodNet, and Very deep convolutional neural network [75], DenseNet161 [72], fully convolutional networks (FCN) [76][82][83][84], YOLO [83][85][86], extreme learning machine (ELM) extreme [87][88][89], neural trees [87], graph convolutional networks (GCN) [90], deep learning PDE model (DPM) [73], SibNet [91], VGG16 or VGG365 [72][75][84][92], ResNet, ResNet50 and ResNet152 [67][70][72][75][76][78][79][80][93][94], EfficientNet [95], EfficientDet [96], faster-RCNN [39][80][94]. GoogLeNet and ResNet, with their variant, were the most popular. Except for learning-based methods, other methods, including the region growing algorithm [68], mean-shift [33] algorithm, template matching [97][98], and other image processing algorithms were also used for image segmentation, recognition and even amount estimation.

An early study harnessed SVM with a Gaussian radial basis kernel for training a classifier on food type achieved an accuracy of 97.3% when the training data took up to 50% of the dataset [24]. Surprisingly, only 1% of misreported nutrient information was found in the study. However, there was only one food item in each image, so the robustness of the proposed algorithm could be limited when tested on images with multiple food items or complex backgrounds [24].

For automatic and larger-scale image analysis, computer vision algorithms were then used in later research. CNNs trained by labelled image data provided another method for food classification. Im2Calories mentioned in the last section were examples of a GoogLeNet CNN being trained with different datasets created from existing datasets online. Im2Calories [61] trained the GoogLeNet with a self-made multi-label dataset and achieved an average precision of 80% [61]. In Ref. [64], Food-11 was created for training, validation, and evaluation, resulting in recognition by 83.6% of food categories.

Deep neural networks were most likely to achieve extremely high performance (over 99% accuracy) in the classification and recognition tasks. The networks could be used for both food and fluid classification in which Inception ResNet V2, ResNet50 V2, ResNet152, MobileNet V2 and V3 and GoogleNet were seen with over 95% accuracy. Apart from deep neural networks, machine learning methods such as RF, SVM, KNN etc, could also reach over 90% accuracy. However, DL methods could require high-performance devices and be time-consuming when training and the performance of models rely on a sufficient amount of training data with variety. The value of a deep neural network lies in the trade-off between its performance and simplicity.

4.3. Intake Action Recognition

The process of an intake activity can be segmented into preparing, delivering, and swallowing, where the preparing phase includes the action of grasping a container and delivering refers to lifting hands to one's mouth. Most of the methods took the observation of food or fluid in human hands as a representation of intake, which turned the action recognition problem into a simple object detection problem. However, taking the presence of food/drink objects as the representation of intake activities has a high false positive rate.

Most of the action detection tasks depended on third-person cameras rather than first-person cameras; in those third-person cameras, depth cameras were more popular than RGB cameras. Microsoft Kinect was the most used device, of which the SDK could provide skeleton tracking for 25 joints on each body for up to six people, from 0.8 to 4 m, as well as six types of streams, including depth, infrared, colour, skeleton and audio [51][55]. As for hardware settings, [51] was tested on a computer running Windows 8 with an Intel Core i5 processor and 8 GB of RAM.

Starting with third-person methods, RGB information was first used for intake detection before the development of the depth camera. One example was the method based on fuzzy vector quantization proposed in 2012, in which activities were considered as 3D volumes formed by a sequence of human poses [54]. Fuzzy vector quantization was for associating the 3D volume representation of an activity video with 3D volume prototypes; the linear discriminant analysis was then used to map activity representations in a low-dimensional discriminant feature space.

The Naive Bayes classifier was first used with Kinect in [3] in 2013 to classify the input images for patient fluid intake monitoring. The performance with different positions of the subject and partial occlusions of the camera was tested. However, the limitation found in this method was that Naive Bayes classifier was only applicable to a relatively small dataset and test case, so the effectiveness of this approach in large-scale free-living scenarios remained to be validated. Moreover, the experimental test set for the method was insufficient, with only three replications and 10 s data for each [3].

In later years, more information from the camera, rather than a single indicator, was used for more accurate detection. In [30], the skeleton coordinates from the depth image provided by Kinect were used for analysing the movements of drinking soup, drinking water, and eating the main course. The distance from both hands to the head and the plate to the head were used as characterisation for classifying the gestures, resulting in an 89% average success rate for three subjects [30]. However, no algorithm was presented in the study, no occlusion was considered during the test, and only three subjects were observed and evaluated, which could lead to bias because of personal dietary habits. Disregarding the limitations, the study validated the feasibility of using the distance between hands, head, and plate for intake monitoring.

4.4. Intake Amount Estimation

The studies mentioned above were mostly related to intake detection and classification rather than intake amount estimation. However, food volume estimation is another problem to be considered, which is the 'how much' problem [61][99][100][101][102][103]. Meal estimation could be realised based on the respective number of intaking gestures for consuming liquid, soup, and meal [34], but the accuracy was not evaluated. Volume estimation based on 3D reconstruction algorithms from images taken by phone was seen in [103], resulting in less than 0.02 inch absolute error for radius estimation (for radius ranging from 0.8 inches to 1.45 inches). Im2Calories was another example, which firstly predicted the distance of each pixel from the camera using a CNN trained on the NYUv2 RGBD dataset, resulting in an average relative error of 0.18 m, which was too high [61]. Then, the depth map was converted to voxel representation for food size estimation, resulting in less than 400 mL absolute volume error [61].

As for fluid amount estimation, a design called 'Playful Bottle' was proposed, which combined the camera and accelerometer on the phone to realise fluid intake tracking and reminding [104]. The accelerometer was used for drinking action detection, in which case 21.1% of false-positive detections could be caused by shaking the bottle without actually drinking from it.

References

1. Vu, T.; Lin, F.; Alshurafa, N.; Xu, W. Wearable Food Intake Monitoring Technologies: A Comprehensive Review. *Computers* 2017, 6, 4.
2. Jayatilaka, A.; Ranasinghe, D.C. Towards unobtrusive real-time fluid intake monitoring using passive UHF RFID. In *Proceedings of the 2016 IEEE International Conference on RFID (RFID)*, Orlando, FL, USA, 3-5 May 2016; IEEE: Orlando, FL, USA, 2016; pp. 1–4.
3. Burgess, R.A.; Hartley, T.; Mehdi, Q.; Mehdi, R. Monitoring of patient fluid intake using the Xbox Kinect. In *Proceedings of the CGAMES'2013 USA*, Louisville, KY, USA, 30 July–1 August 2013; pp. 60–64.
4. Cao, E.; Watt, M.J.; Nowell, C.J.; Quach, T.; Simpson, J.S.; De Melo Ferreira, V.; Agarwal, S.; Chu, H.; Srivastava, A.; Anderson, D.; et al. Mesenteric Lymphatic Dysfunction Promotes Insulin Resistance and Represents a Potential Treatment Target in Obesity. *Nat. Metab.* 2021, 3, 1175–1188.
5. Lauby-Secretan, B.; Scoccianti, C.; Loomis, D.; Grosse, Y.; Bianchini, F.; Straif, K. Body Fatness and Cancer—Viewpoint of the IARC Working Group. *N. Engl. J. Med.* 2016, 375, 794–798.
6. Doulah, A.; Mccrory, M.A.; Higgins, J.A.; Sazonov, E. A Systematic Review of Technology-Driven Methodologies for Estimation of Energy Intake. *IEEE Access* 2019, 7, 49653–49668.
7. Schoeller, D.A.; Thomas, D. Energy balance and body composition. In *Nutrition for the Primary Care Provider*; Karger Publishers: Basel, Switzerland, 2014.

8. Wang, W.; Min, W.; Li, T.; Dong, X.; Li, H.; Jiang, S. A Review on Vision-Based Analysis for Automatic Dietary Assessment. *Trends Food Sci. Technol.* 2022, 122, 223–237.
9. Lacey, J.; Corbett, J.; Forni, L.; Hooper, L.; Hughes, F.; Minto, G.; Moss, C.; Price, S.; Whyte, G.; Woodcock, T.; et al. A Multidisciplinary Consensus on Dehydration: Definitions, Diagnostic Methods and Clinical Implications. *Ann. Med.* 2019, 51, 232–251.
10. Volkert, D.; Beck, A.M.; Cederholm, T.; Cruz-Jentoft, A.; Goisser, S.; Hooper, L.; Kiesswetter, E.; Maggio, M.; Raynaud-Simon, A.; Sieber, C.C.; et al. ESPEN Guideline on Clinical Nutrition and Hydration in Geriatrics. *Clin. Nutr.* 2019, 38, 10–47.
11. Armstrong, L. Challenges of Linking Chronic Dehydration and Fluid Consumption to Health Outcomes. *Nutr. Rev.* 2012, 70, S121–S127.
12. Manz, F.; Wentz, A. 24-h Hydration Status: Parameters, Epidemiology and Recommendations. *Eur. J. Clin. Nutr.* 2003, 57, S10–S18.
13. Hooper, L.; Bunn, D.; Jimoh, F.O.; Fairweather-Tait, S.J. Water-Loss Dehydration and Aging. *Mech. Ageing Dev.* 2014, 136–137, 50–58.
14. El-Sharkawy, A.M.; Sahota, O.; Maughan, R.J.; Lobo, D.N. The Pathophysiology of Fluid and Electrolyte Balance in the Older Adult Surgical Patient. *Clin. Nutr.* 2014, 33, 6–13.
15. El-Sharkawy, A.M.; Watson, P.; Neal, K.R.; Ljungqvist, O.; Maughan, R.J.; Sahota, O.; Lobo, D.N. Hydration and Outcome in Older Patients Admitted to Hospital (The HOOP Prospective Cohort Study). *Age Ageing* 2015, 44, 943–947.
16. Kim, S. Preventable Hospitalizations of Dehydration: Implications of Inadequate Primary Health Care in the United States. *Ann. Epidemiol.* 2007, 17, 736.
17. Warren, J.L.; Bacon, W.E.; Harris, T.; McBean, A.M.; Foley, D.J.; Phillips, C. The Burden and Outcomes Associated with Dehydration among US Elderly, 1991. *Am. J. Public Health* 1994, 84, 1265–1269.
18. Jimoh, F.O.; Bunn, D.; Hooper, L. Assessment of a Self-Reported Drinks Diary for the Estimation of Drinks Intake by Care Home Residents: Fluid Intake Study in the Elderly (FISE). *J. Nutr. Health Aging* 2015, 19, 491–496.
19. Edmonds, C.J.; Harte, N.; Gardner, M. How Does Drinking Water Affect Attention and Memory? The Effect of Mouth Rinsing and Mouth Drying on Children's Performance. *Physiol. Behav.* 2018, 194, 233–238.
20. Doulah, A.; Ghosh, T.; Hossain, D.; Imtiaz, M.H.; Sazonov, E. "Automatic Ingestion Monitor Version 2"—A Novel Wearable Device for Automatic Food Intake Detection and Passive Capture of Food Images. *IEEE J. Biomed. Health Inf.* 2021, 25, 568–576.
21. Rouast, P.; Adam, M.; Burrows, T.; Chiong, R.; Rollo, M. Using deep learning and 360 video to detect eating behavior for user assistance systems. In *Proceedings of the Twenty-Sixth European Conference on Information Systems (ECIS2018)*, Portsmouth, UK, 23–28 June 2018.
22. Block, G. A Review of Validations of Dietary Assessment Methods. *Am. J. Epidemiol.* 1982, 115, 492–505.
23. Fluid Balance Charts: Do They Measure Up? Available online: <https://www.magonlinelibrary.com/doi/epdf/10.12968/bjon.1994.3.16.816> (accessed on 5 April 2022).
24. Zhu, F.; Bosch, M.; Boushey, C.J.; Delp, E.J. An image analysis system for dietary assessment and evaluation. In *Proceedings of the 2010 IEEE International Conference on Image Processing*, Hong Kong, China, 26–29 September 2010; pp. 1853–1856.
25. Krebs, P.; Duncan, D.T. Health App Use Among US Mobile Phone Owners: A National Survey. *JMIR mHealth uHealth* 2015, 3, e101.
26. Sen, S.; Subbaraju, V.; Misra, A.; Balan, R.; Lee, Y. Annapurna: Building a real-world smartwatch-based automated food journal. In *Proceedings of the 2018 IEEE 19th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*, Chania, Greece, 12–15 June 2018; pp. 1–6.
27. Sen, S.; Subbaraju, V.; Misra, A.; Balan, R.; Lee, Y. Annapurna: An Automated Smartwatch-Based Eating Detection and Food Journaling System. *Pervasive Mob. Comput.* 2020, 68, 101259.
28. Thomaz, E.; Parnami, A.; Essa, I.; Abowd, G.D. Feasibility of identifying eating moments from first-person images leveraging human computation. In *Proceedings of the SenseCam '13: 4th International SenseCam & Pervasive Imaging Conference*, San Diego, CA, USA, 18–19 November 2013; ACM Press: New York, NY, USA, 2013; pp. 26–33.
29. Bedri, A.; Li, D.; Khurana, R.; Bhuwarka, K.; Goel, M. FitByte: Automatic diet monitoring in unconstrained situations using multimodal sensing on eyeglasses. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, Honolulu, HI, USA, 25–30 April 2021; ACM: New York, NY, USA, 2020; pp. 1–12.
30. Cunha, A.; Pádua, L.; Costa, L.; Trigueiros, P. Evaluation of MS Kinect for Elderly Meal Intake Monitoring. *Procedia Technol.* 2014, 16, 1383–1390.
31. Obaid, A.K.; Abdel-Qader, I.; Mickus, M. Automatic food-intake monitoring system for persons living with Alzheimer's-vision-based embedded system. In *Proceedings of the 2018 9th IEEE Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*, New York, NY, USA, 8–10 November 2018; pp. 580–584.

32. Hossain, D.; Ghosh, T.; Sazonov, E. Automatic Count of Bites and Chews from Videos of Eating Episodes. *IEEE Access* 2020, 8, 101934–101945.
33. Hafiz, R.; Islam, S.; Khanom, R.; Uddin, M.S. Image based drinks identification for dietary assessment. In *Proceedings of the 2016 International Workshop on Computational Intelligence (IWCI)*, Dhaka, Bangladesh, 12–13 December 2016; pp. 192–197.
34. Automatic Meal Intake Monitoring Using Hidden Markov Models | Elsevier Enhanced Reader. Available online: <https://reader.elsevier.com/reader/sd/pii/S1877050916322980?token=11F68A53BC12E070DB5891E8BEA3ACDCC32BEB1B24D82F1AAEFBBA65AA6016F2E9009A4B70146365DDA53B667BB161D7west-1&originCreation=20220216180556> (accessed on 16 February 2022).
35. Al-Anssari, H.; Abdel-Qader, I. Vision based monitoring system for Alzheimer's patients using controlled bounding boxes tracking. In *Proceedings of the 2016 IEEE International Conference on Electro Information Technology (EIT)*, Grand Forks, ND, USA, 19–21 May 2016; pp. 821–825.
36. Davies, A.; Chan, V.; Bauman, A.; Signal, L.; Hosking, C.; Gemming, L.; Allman-Farinelli, M. Using Wearable Cameras to Monitor Eating and Drinking Behaviours during Transport Journeys. *Eur. J. Nutr.* 2021, 60, 1875–1885.
37. Jia, W.; Li, Y.; Qu, R.; Baranowski, T.; Burke, L.E.; 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.* 2019, 22, 1168–1179.
38. Gemming, L.; Doherty, A.; Utter, J.; Shields, E.; Ni Mhurchu, C. The Use of a Wearable Camera to Capture and Categorise the Environmental and Social Context of Self-Identified Eating Episodes. *Appetite* 2015, 92, 118–125.
39. Du, B.; Lu, C.X.; Kan, X.; Wu, K.; Luo, M.; Hou, J.; Li, K.; Kanhere, S.; Shen, Y.; Wen, H. HydraDoctor: Real-Time liquids intake monitoring by collaborative sensing. In *Proceedings of the 20th International Conference on Distributed Computing and Networking*, Bangalore, India, 4–7 January 2019; Association for Computing Machinery: New York, NY, USA, 2019; pp. 213–217.
40. Kong, F.; Tan, J. DietCam: Automatic Dietary Assessment with Mobile Camera Phones. *Pervasive Mob. Comput.* 2012, 8, 147–163.
41. Hongu, N.; Pope, B.T.; Bilgiç, P.; Orr, B.J.; Suzuki, A.; Kim, A.S.; Merchant, N.C.; Roe, D.J. Usability of a Smartphone Food Picture App for Assisting 24-Hour Dietary Recall: A Pilot Study. *Nutr. Res. Pract.* 2015, 9, 207–212.
42. Ahmad, Z.; Khanna, N.; Kerr, D.A.; Boushey, C.J.; Delp, E.J. A mobile phone user interface for image-based dietary assessment. In *Mobile Devices and Multimedia: Enabling Technologies, Algorithms, and Applications 2014*; SPIE—The International Society for Optical Engineering: Bellingham, WA, USA, 2014; Volume 9030, p. 903007.
43. Raju, V.; Sazonov, E. Processing of egocentric camera images from a wearable food intake sensor. In *Proceedings of the 2019 SoutheastCon*, Huntsville, AL, USA, 11–14 April 2019; pp. 1–6.
44. Vinay Chandrasekhar, K.; Imtiaz, M.H.; Sazonov, E. Motion-Adaptive image capture in a body-worn wearable sensor. In *Proceedings of the 2018 IEEE SENSORS*, New Delhi, India, 28–31 October 2018; pp. 1–4.
45. Gemming, L.; Rush, E.; Maddison, R.; Doherty, A.; Gant, N.; Utter, J.; Mhurchu, C.N. Wearable Cameras Can Reduce Dietary Under-Reporting: Doubly Labelled Water Validation of a Camera-Assisted 24 h Recall. *Br. J. Nutr.* 2015, 113, 284–291.
46. Seiderer, A.; Flutura, S.; André, E. Development of a mobile multi-device nutrition logger. In *Proceedings of the 2nd ACM SIGCHI International Workshop on Multisensory Approaches to Human-Food Interaction*, Glasgow, UK, 13 November 2017; ACM: New York, NY, USA, 2017; pp. 5–12.
47. Schiboni, G.; Wasner, F.; Amft, O. A privacy-preserving wearable camera setup for dietary event spotting in free-living. In *Proceedings of the 2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, Athens, Greece, 19–23 March 2018; pp. 872–877.
48. Sen, S.; Subbaraju, V.; Misra, A.; Balan, R.K.; Lee, Y. Experiences in building a real-world eating recogniser. In *Proceedings of the 4th International Workshop on Physical Analytics—WPA '17*, Niagara Falls, NY, USA, 19 June 2017; ACM Press: New York, NY, USA, 2017; pp. 7–12.
49. 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.
50. Cippitelli, E.; Gasparrini, S.; Gambi, E.; Spinsante, S. Unobtrusive intake actions monitoring through RGB and depth information fusion. In *Proceedings of the 2016 IEEE 12th International Conference on Intelligent Computer Communication and Processing (ICCP)*, Cluj-Napoca, Romania, 8–10 September 2016; pp. 19–26.
51. Cippitelli, E.; Gambi, E.; Spinsante, S.; Gasparrini, S.; Florez-Revuelta, F. Performance analysis of self-organising neural networks tracking algorithms for intake monitoring using Kinect. In *Proceedings of the IET International Conference on Technologies for Active and Assisted Living (TechAAL)*, London, UK, 5 November 2015; Institution of Engineering and Technology: London, UK, 2015; p. 6.
52. Blechert, J.; Liedlgruber, M.; Lender, A.; Reichenberger, J.; Wilhelm, F.H. Unobtrusive Electromyography-Based Eating Detection in Daily Life: A New Tool to Address Underreporting? *Appetite* 2017, 118, 168–173.
53. Maekawa, T. A sensor device for automatic food lifelogging that is embedded in home ceiling light: A preliminary investigation. In *Proceedings of the 2013 7th International Conference on Pervasive Computing Technologies for*

54. Iosifidis, A.; Marami, E.; Tefas, A.; Pitas, I. Eating and drinking activity recognition based on discriminant analysis of fuzzy distances and activity volumes. In *Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, 25–30 March 2012; pp. 2201–2204.
55. bin Kassim, M.F.; Mohd, M.N.H. Food Intake Gesture Monitoring System Based-on Depth Sensor. *Bull. Electr. Eng. Inform.* 2019, 8, 470–476.
56. Das, S.; Dai, R.; Koperski, M.; Minciullo, L.; Garattoni, L.; Bremond, F.; Francesca, G. Toyota smarhome: Real-World activities of daily living. In *Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Republic of Korea, 27 October–2 November 2019.
57. Lo, F.P.-W.; Sun, Y.; Qiu, J.; Lo, B. A Novel Vision-Based Approach for Dietary Assessment Using Deep Learning View Synthesis. In *Proceedings of the 2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, Chicago, IL, USA, 19–22 May 2019; pp. 1–4.
58. Farooq, M.; Doulah, A.; Parton, J.; McCrory, M.A.; Higgins, J.A.; Sazonov, E. Validation of Sensor-Based Food Intake Detection by Multicamera Video Observation in an Unconstrained Environment. *Nutrients* 2019, 11, 609.
59. qm13 Azure Kinect Body Tracking Joints. Available online: <https://docs.microsoft.com/en-us/azure/kinect-dk/body-joints> (accessed on 23 June 2022).
60. Kassim, M.F.; Mohd, M.N.H.; Tomari, M.R.M.; Suriani, N.S.; Zakaria, W.N.W.; Sari, S. A Non-Invasive and Non-Wearable Food Intake Monitoring System Based on Depth Sensor. *Bull. Electr. Eng. Inform.* 2020, 9, 2342–2349.
61. Myers, A.; Johnston, N.; Rathod, V.; Korattikara, A.; Gorban, A.; Silberman, N.; Guadarrama, S.; Papandreou, G.; Huang, J.; Murphy, K. Im2Calories: Towards an automated mobile vision food diary. In *Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, 7–13 December 2015; IEEE: New York, NY, USA, 2015; pp. 1233–1241.
62. Rachakonda, L.; Mohanty, S.P.; Kougianos, E. iLog: An Intelligent Device for Automatic Food Intake Monitoring and Stress Detection in the IoMT. *IEEE Trans. Consum. Electron.* 2020, 66, 115–124.
63. Hossain, D.; Imtiaz, M.H.; Ghosh, T.; Bhaskar, V.; Sazonov, E. Real-Time food intake monitoring using wearable egocentric camera. In *Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, Montreal, QC, Canada, 20–24 July 2020; pp. 4191–4195.
64. Singla, A.; Yuan, L.; Ebrahimi, T. Food/Non-Food image classification and food categorization using pre-trained GoogLeNet model. In *Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management*, Amsterdam, The Netherlands, 16 October 2016; ACM: New York, NY, USA, 2016; pp. 3–11.
65. Beltran, A.; Dadabhoy, H.; Chen, T.A.; Lin, C.; Jia, W.; Baranowski, J.; Yan, G.; Sun, M.; Baranowski, T. Adapting the eButton to the Abilities of Children for Diet Assessment. *Proc. Meas. Behav.* 2016, 2016, 72–81.
66. Ramesh, A.; Raju, V.B.; Rao, M.; Sazonov, E. Food detection and segmentation from egocentric camera images. In *Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Guadalajara, Mexico, 1–5 November 2021; pp. 2736–2740.
67. Mezgec, S.; Koroušić Seljak, B. NutriNet: A Deep Learning Food and Drink Image Recognition System for Dietary Assessment. *Nutrients* 2017, 9, 657.
68. Zhang, Z.; Mao, S.; Chen, K.; Xiao, L.; Liao, B.; Li, C.; Zhang, P. CNN and PCA based visual system of a wheelchair manipulator robot for automatic drinking. In *Proceedings of the 2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Kuala Lumpur, Malaysia, 12–15 December 2018; pp. 1280–1286.
69. Lee, J.; Banerjee, A.; Gupta, S.K.S. MT-Diet: Automated smartphone based diet assessment with infrared images. In *Proceedings of the 2016 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, Sydney, NSW, Australia, 14–19 March 2016; pp. 1–6.
70. McAllister, P.; Zheng, H.; Bond, R.; Moorhead, A. Combining Deep Residual Neural Network Features with Supervised Machine Learning Algorithms to Classify Diverse Food Image Datasets. *Comput. Biol. Med.* 2018, 95, 217–233.
71. Jiang, H.; Starkman, J.; Liu, M.; Huang, M.-C. Food Nutrition Visualization on Google Glass: Design Tradeoff and Field Evaluation. *IEEE Consum. Electron. Mag.* 2018, 7, 21–31.
72. Matei, A.; Glavan, A.; Radeva, P.; Talavera, E. Towards Eating Habits Discovery in Egocentric Photo-Streams. *IEEE Access* 2021, 9, 17495–17506.
73. Matsuda, Y.; Hoashi, H.; Yanai, K. Recognition of multiple-food images by detecting candidate regions. In *Proceedings of the 2012 IEEE International Conference on Multimedia and Expo*, Melbourne, VIC, Australia, 9–13 July 2012; pp. 25–30.
74. Anthimopoulos, M.M.; Gianola, L.; Scarnato, L.; Diem, P.; Mougiakakou, S.G. A Food Recognition System for Diabetic Patients Based on an Optimized Bag-of-Features Model. *IEEE J. Biomed. Health Inform.* 2014, 18, 1261–1271.
75. Park, S.-J.; Palvanov, A.; Lee, C.-H.; Jeong, N.; Cho, Y.-I.; Lee, H.-J. The Development of Food Image Detection and Recognition Model of Korean Food for Mobile Dietary Management. *Nutr. Res. Pract.* 2019, 13, 521–528.
76. Mezgec, S.; Koroušić Seljak, B. Deep Neural Networks for Image-Based Dietary Assessment. *J. Vis. Exp.* 2021, 169, e61906.

77. Bolaños, M.; Radeva, P. Simultaneous food localization and recognition. In Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR), Cancun, Mexico, 4–8 December 2016; pp. 3140–3145.
78. Ciocca, G.; Micali, G.; Napoletano, P. State Recognition of Food Images Using Deep Features. *IEEE Access* 2020, 8, 32003–32017.
79. Fuchs, K.; Haldimann, M.; Grundmann, T.; Fleisch, E. Supporting Food Choices in the Internet of People: Automatic Detection of Diet-Related Activities and Display of Real-Time Interventions via Mixed Reality Headsets. *Futur. Gener. Comp. Syst.* 2020, 113, 343–362.
80. Ruenin, P.; Bootkrajang, J.; Chawachat, J. A system to estimate the amount and calories of food that elderly people in the hospital consume. In Proceedings of the 11th International Conference on Advances in Information Technology, Bangkok, Thailand, 1–3 July 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–7.
81. Lu, Y.; Stathopoulou, T.; Vasiloglou, M.F.; Pinault, L.F.; Kiley, C.; Spanakis, E.K.; Mougiakakou, S. goFOODTM: An Artificial Intelligence System for Dietary Assessment. *Sensors* 2020, 20, 4283.
82. Mezgec, S.; Seljak, B.K. Using deep learning for food and beverage image recognition. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019; pp. 5149–5151.
83. Aguilar, E.; Remeseiro, B.; Bolaños, M.; Radeva, P. Grab, Pay, and Eat: Semantic Food Detection for Smart Restaurants. *IEEE Trans. Multimed.* 2018, 20, 3266–3275.
84. Gao, J.; Tan, W.; Ma, L.; Wang, Y.; Tang, W. MUSEFood: Multi-Sensor-Based food volume estimation on smartphones. In Proceedings of the 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), Leicester, UK, 19–23 August 2019; pp. 899–906.
85. Yumang, A.N.; Banguilan, D.E.S.; Veneracion, C.K.S. Raspberry PI based food recognition for visually impaired using YOLO algorithm. In Proceedings of the 2021 5th International Conference on Communication and Information Systems (ICCIS), Chongqing, China, 15–17 October 2021; pp. 165–169.
86. Li, H.; Yang, G. Dietary Nutritional Information Autonomous Perception Method Based on Machine Vision in Smart Homes. *Entropy* 2022, 24, 868.
87. Martinel, N.; Piciarelli, C.; Foresti, G.L.; Micheloni, C. Mobile food recognition with an extreme deep tree. In Proceedings of the ICDSC 2016: 10th International Conference on Distributed Smart Camera, Paris, France, 12–15 September 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 56–61.
88. Martinel, N.; Piciarelli, C.; Micheloni, C.; Foresti, G.L. A structured committee for food recognition. In Proceedings of the 2015 IEEE International Conference on Computer Vision Workshop (ICCVW), Santiago, Chile, 7–13 December 2015; pp. 484–492.
89. Martinel, N.; Piciarelli, C.; Micheloni, C. A Supervised Extreme Learning Committee for Food Recognition. *Comput. Vis. Image Underst.* 2016, 148, 67–86.
90. Zhao, H.; Yap, K.-H.; Chichung Kot, A. Fusion learning using semantics and graph convolutional network for visual food recognition. In Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 3–8 January 2021; pp. 1710–1719.
91. Nguyen, H.-T.; Ngo, C.-W.; Chan, W.-K. SibNet: Food Instance Counting and Segmentation. *Pattern Recognit.* 2022, 124, 108470.
92. Martinez, E.T.; Leyva-Vallina, M.; Sarker, M.M.K.; Puig, D.; Petkov, N.; Radeva, P. Hierarchical Approach to Classify Food Scenes in Egocentric Photo-Streams. *IEEE J. Biomed. Health Inform.* 2020, 24, 866–877.
93. Sarapisto, T.; Koivunen, L.; Mäkilä, T.; Klami, A.; Ojansivu, P. Camera-Based meal type and weight estimation in self-service lunch line restaurants. In Proceedings of the 2022 12th International Conference on Pattern Recognition Systems (ICPRS), Saint-Etienne, France, 7–10 June 2022; pp. 1–7.
94. Qiu, J.; Lo, F.P.-W.; Gu, X.; Jobarteh, M.L.; Jia, W.; Baranowski, T.; Steiner-Asiedu, M.; Anderson, A.K.; McCrory, M.A.; Sazonov, E.; et al. Egocentric image captioning for privacy-preserved passive dietary intake monitoring. In *IEEE Transactions on Cybernetics*; IEEE: New York, NY, USA, 2021.
95. Tomescu, V.-I. FoRConvD: An approach for food recognition on mobile devices using convolutional neural networks and depth maps. In Proceedings of the 2020 IEEE 14th International Symposium on Applied Computational Intelligence and Informatics (SACI), Timisoara, Romania, 21–23 May 2020; pp. 000129–000134.
96. Liu, Y.-C.; Onthoni, D.D.; Mohapatra, S.; Irianti, D.; Sahoo, P.K. Deep-Learning-Assisted Multi-Dish Food Recognition Application for Dietary Intake Reporting. *Electronics* 2022, 11, 1626.
97. Chae, J.; Woo, I.; Kim, S.; Maciejewski, R.; Zhu, F.; Delp, E.J.; Boushey, C.J.; Ebert, D.S. Volume estimation using food specific shape templates in mobile image-based dietary assessment. In *Computational Imaging IX*; Bouman, C.A., Pollak, I., Wolfe, P.J., Eds.; SPIE—The International Society for Optical Engineering: Bellingham, WA, USA, 2011; Volume 7873, p. 78730K.
98. Bellandi, P.; Sansoni, G.; Vertuan, A. Development and Characterization of a Multi-Camera 2D-Vision System for Enhanced Performance of a Drink Serving Robotic Cell. *Robot. Comput.-Integr. Manuf.* 2012, 28, 35–49.

99. Rachakonda, L.; Kothari, A.; Mohanty, S.P.; Kougianos, E.; Ganapathiraju, M. Stress-Log: An IoT-Based smart system to monitor stress-eating. In Proceedings of the 2019 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 11–13 January 2019; pp. 1–6.
100. Lu, Y.; Allegra, D.; Anthimopoulos, M.; Stanco, F.; Farinella, G.M.; Mougiakakou, S. A multi-task learning approach for meal assessment. In Proceedings of the CEA/MADiMa2018: Joint Workshop on Multimedia for Cooking and Eating Activities and Multimedia Assisted Dietary Management in conjunction with the 27th International Joint Conference on Artificial Intelligence IJCAI, Stockholm, Sweden, 15 July 2018; pp. 46–52.
101. Fang, S.; Liu, C.; Zhu, F.; Delp, E.J.; Boushey, C.J. Single-View food portion estimation based on geometric models. In Proceedings of the 2015 IEEE International Symposium on Multimedia (ISM), Miami, FL, USA, 14–16 December 2015; pp. 385–390.
102. Puri, M.; Zhu, Z.; Yu, Q.; Divakaran, A.; Sawhney, H. Recognition and volume estimation of food intake using a mobile device. In Proceedings of the 2009 Workshop on Applications of Computer Vision (WACV), Snowbird, UT, USA, 7–8 December 2009; pp. 1–8.
103. Woo, I.; Otsmo, K.; Kim, S.; Ebert, D.S.; Delp, E.J.; Boushey, C.J. Automatic portion estimation and visual refinement in mobile dietary assessment. In Proceedings of the Computational Imaging VIII, San Jose, CA, USA, 4 February 2010; Bouman, C.A., Pollak, I., Wolfe, P.J., Eds.; SPIE—The International Society for Optical Engineering: Bellingham, WA, USA, 2010; pp. 188–197.
104. Chiu, M.-C.; Chang, S.-P.; Chang, Y.-C.; Chu, H.-H.; Chen, C.C.-H.; Hsiao, F.-H.; Ko, J.-C. Playful bottle: A mobile social persuasion system to motivate healthy water intake. In Proceedings of the 11th International Conference on Ubiquitous Computing, Orlando, FL, USA, 30 September–3 October 2019; Association for Computing Machinery: New York, NY, USA, 2009; pp. 185–194.

Retrieved from <https://encyclopedia.pub/entry/history/show/106027>