Heliyon 10 (2024) e35689

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Novel digital-based approach for evaluating wine components' intake: A deep learning model to determine red wine volume in a glass from single-view images

Miriam Cobo [a], Edgard Relaño de la Guía [b], Ignacio Heredia [a], Fernando Aguilar [a], Lara Lloret-Iglesias [a], Daniel García [a], Silvia Yuste [c], Emma Recio-Fernández [d], Patricia Pérez-Matute [d], M. José Motilva [c], M. Victoria Moreno-Arribas [b], Begoña Bartolomé [b,*]

[a] *Institute of Physics of Cantabria (IFCA), CSIC - UC, 39005, Santander, Cantabria, Spain*
[b] *Institute of Food Science Research (CIAL), CSIC-UAM, 28049, Madrid, Spain*
[c] *Institute of Grapevine and Wine Sciences (ICVV), CSIC-University of La Rioja-Government of La Rioja, 26007, Logroño, La Rioja, Spain*
[d] *Infectious Diseases, Microbiota and Metabolism Unit, Center for Biomedical Research of La Rioja (CIBIR), CSIC Associated Unit, 26006, Logroño, La Rioja, Spain, USA*

ARTICLE INFO

Keywords:
Red wine
Polyphenols
Alcohol
Deep learning
Liquid volume estimation
Consumer study application

ABSTRACT

Estimation of wine components' intake (polyphenols, alcohol, etc.) through Food Frequency Questionnaires (FFQs) may be particularly inaccurate. This paper reports the development of a deep learning (DL) method to determine red wine volume from single-view images, along with its application in a consumer study developed via a web service. The DL model demonstrated satisfactory performance not only in a daily lifelike images dataset (mean absolute error = 10 mL), but also in a real images dataset that was generated through the consumer study (mean absolute error = 26 mL). Based on the data reported by the participants in the consumer study (n = 38), average red wine volume in a glass was 114 ± 33 mL, which represents an intake of 137–342 mg of total polyphenols, 11.2 g of alcohol, 0.342 g of sugars, among other components. Therefore, the proposed method constitutes a diet-monitoring tool of substantial utility in the accurate assessment of wine components' intake.

1. Introduction

The estimation of dietary intake in free-living environments is a major challenge in food science and nutrition research. Dietary intake can be assessed by direct methods, such as direct observation, duplicate diets and nutritional biomarkers, although the most common methods include indirect (self-report) methods, such as food diaries (weighed or estimated), 24-h dietary recall and Food Frequency Questionnaires (FFQs) [1,2]. In particular, FFQs are checklists of foods and beverages with a frequency response section to report how often each item is consumed over a specified period of time (usually the past year, but shorter periods can be used). FFQs, also classified as long-term methods, have several strengths they are less costly than other dietary assessment methods, allow for quick and automated data analysis, facilitate the establishment of individual dietary patterns, and can classify individuals in a population

* Corresponding author.
E-mail address: b.bartolome@csic.es (B. Bartolomé).

https://doi.org/10.1016/j.heliyon.2024.e35689
Received 6 July 2023; Received in revised form 30 July 2024; Accepted 1 August 2024
Available online 2 August 2024
2405-8440/

based on their habits [3]. However, FFQs depend primarily on long-term memory of the interviewed subject, do not account for intrapersonal variation in recording the daily consumption of food during the study period and, in particular, do not allow a precise estimation of serving sizes/volume of food consumed (see Ref. [4], for a detailed review). Due to these limitations, results obtained by these subjective methods over both the short- and long-term could lead to inexact conclusions and incorrect decisions [2,5,6]. Therefore, alternative technologies are needed to allow an assessment of food components' intake accurately and easily in real-world settings.

In this framework, new wearable imaging and computational technologies capable of performing objective and passive dietary assessments with a much more simplified procedure than traditional FFQs have emerged (see Ref. [7], for an exhaustive review). The architecture of the so-called Image-Based Food-Recognition Systems from the meal digital image to the final data of energy, nutrients and other nutritional features, implies several phases such as image collection, pre-processing, segmentation, feature extraction, dimensionality reduction, classification, volume estimation, and estimation [7]. Among these phases, the most critical task is the estimation of the food portion size or the volume contained in a cup/glass from the digital image [8]. In this context, the application of artificial intelligence and computer vision to foods or liquids and their containers opens new perspectives to improve and standardize size/volume information and food components' intake.

This problem in determining dietary intake may be particularly critical in the case of wine. A light-to-moderate consumption of red wine is one of the characteristics of the Mediterranean diet [9], probably one of the most widely evaluated and recommended dietary patterns [10]. Evaluation of the Mediterranean-diet adherence in epidemiological and observational studies through FFQs is particularly inaccurate, as a portion (glass of wine) consumed is standardized to 100 mL, which does not always correspond to real situations. Moreover, depending on the type of glass (container), wine consumption may be wrongly estimated because of inaccurate information on volume [11,12]. Accurately measurement of wine intake is a fundamental step to estimate bioactive wine components' intake (i.e., polyphenols) that could be associated with health outcomes [13].

As a first approach to solve the automatic liquid volume estimation challenge, a deep learning (DL) method to measure the red wine volume in a glass from a single-view RGB (Red, Green and Blue) photograph, without any size reference object, was proposed in a previous study [14]. In this former work, the algorithm was trained with laboratory photographs from the BrainGut_WineUp laboratory images dataset [15], as a first approximation albeit far from real conditions. The DL method achieved a mean absolute error (MAE) of 8 mL and was successfully validated under different image conditions taken at laboratory scale [14].

From this previous algorithm [14], in the present study we have developed a refined DL model to determine red wine volume in a glass from a single-view photograph, and present its real application based on an ad hoc consumer study, as a proof of concept. The main contributions of this work are: (1) creating and releasing an independent dataset of images, showing different glass containers with annotated red wine volumes, captured by the authors in everyday situations (the *BrainGut_WineUp daily lifelike images* dataset) [16]; (2) fine-tuning the previous DL laboratory model with this new dataset following a transfer learning approach; (3) collecting dataset of images taken by the participants of a red wine consumer study (the *BrainGut_WineUp real images* dataset) [17]; (4) evaluating the applicability of the proposed fine-tuned DL method on the real images taken in the consumer study; (5) providing accurate data about red wine volume in a glass (volume per serving) and wine consumption habits of the participants in the consumer study. Overall, this paper shows on a real validation dataset that the volume of red wine in a glass can be predicted from a single-view image through artificial intelligence approaches. The refined DL model together with the web application developed in this paper is suitable as a simple and effective automatic monitoring tool for red wine consumption, and therefore, for wine components intake.

2. Materials and methods

2.1. Consumer data

A web application by the name of ALIMENTA365 (https://alimenta365.csic.es/) was specifically designed and created to gather information about red wine consumption habits and to collect real photographs of wine intake. When a participant in the study first accessed the web application, they were asked to register and complete a questionnaire on their wine consumption habits (Supplementary Table 1). Once the questionnaire was completed, the participant received an email with a password to access the web platform and upload their photographs. Participants also received by post a 300 mL volume measuring glass with graduated markings at 25 mL intervals to accurately measure the red wine volume served in their glasses before their intake. The web application included a video tutorial for taking and uploading photographs, indicating the optimum distance and angle from the camera to the glass. The tutorial also asked participants to avoid including people as well as other containers (such as glasses, cups or bottles) different from their wine glass in their photographs.

Before uploading the photograph, participants ticked the right option corresponding to the type of wine (Young/Crianza/Reserva), the volume measured (50 mL/75 mL/100 mL/125 mL/150 mL/175 mL/200 mL/225 mL/250 mL/275 mL/300 mL), the type of glass used (balloon wine glass/Bourgogne wine glass/Bordeaux wine glass/Chardonnay wine glass/wine tasting glass/coffee glass/water glass/short rock glass/rock glass/others), and the time of consumption (Lunch/Dinner/Outside meals) (Supplementary Table 2). Once the photograph was uploaded (to the cloud storage system of the Institute of Physics of Cantabria, IFCA), the participant received a message confirming that the process had been successfully completed.

For this red wine consumer study, participants were recruited from the general population after advertising on social networks through the Institute of Food Science Research (CIAL, Madrid, Spain) Communication Unit. To participate in the study, volunteers had to be over 18 years of age and wine consumers, including both sporadic and regular wine consumers. Participants were asked to upload photographs of all wine servings during 3 weeks, starting from the day they uploaded their first photograph. Throughout this period, volunteers received a weekly follow-up email. At the end of the 3 weeks and for their own information, participants received a report with details of their wine consumption, including daily estimation of sugar, alcohol, calories, and polyphenol intake from their red wine consumption. The study was approved by the Ethics Committee of CSIC (Madrid, Spain) (Approval number 081/2021) and guaranteed the participant's data protection.

A total of 38 volunteers (20 men and 18 women) participated in the study (Supplementary Table 3). This population group ranged from 18 to 80 years, with the majority between 46 and 65 years old (60.5 %) (Supplementary Table 3). It was also characterized by a high level of education (71.1 %), a professional occupation in the Administration (47.4 %) or the private-sector business (23.7 %) (Supplementary Table 3).

2.2. Image datasets

Two independent datasets comprising images of different format of glasses containing red wine were considered to develop and evaluate the refined DL model that is presented in this paper: i) the *BrainGut_WineUp daily lifelike images* dataset [16] which contains photographs taken by members of our research groups under real conditions and situations, and ii) the *BrainGut_WineUp real images* dataset[17][17] which consists of photographs taken by wine consumers in their daily lives.

BrainGut_WineUp daily lifelike images dataset. This dataset was specifically designed and created to fine-tune the previous DL model [14]. The dataset included 1945 photographs that were taken by members of our research groups in lifelike situations to imitate consumers' photographs [16]. Three commercial red wines representative of young, "crianza" and "reserva" wines were selected to take the photographs. For each wine, the same flowchart was followed, as represented in Supplementary Fig. 1. Photographs were taken indoors and outdoors considering the following fields:

- Type of glass (n = 9): balloon wine glass, Bourgogne wine glass, Bordeaux wine glass, Chardonnay wine glass, wine tasting glass, coffee glass, water glass, short rock glass, and rock glass (Supplementary Fig. 2). The wine tasting glass shape used is the wine glass with specifications defined (ISO 3591:1977).
- Volume of wine (n = 7): 50, 75, 100, 125, 150, 175, 200 mL. Measurements were done using a test tube with ±0.5 mL precision.
- Object angle (n = 2): upper (0, 30o), central (0o).

BrainGut_WineUp real images dataset. This dataset contains 229 real photographs that were taken by 38 volunteers who took part in the consumer study [17]. This dataset was used to perform an independent test of the developed DL model, as well as to evaluate the quality of participants' photographs. For each photograph, participants had to fill in the following fields:

- Type of wine (n = 3): "joven" (young wine), "crianza" (after short oak wood-aging) and "reserva" (after long oak wood-aging) wines.
- Type of glass (n = 10): balloon wine glass, Bourgogne wine glass, Bordeaux wine glass, Chardonnay wine glass, wine tasting glass, coffee glass, water glass, short rock glass, rock glass (Supplementary Fig. 2), and others.
- Intervals for the allocation of the measured wine volume (n = 11): 50, 75, 100, 125, 150, 175, 200, 225, 250, 275 and 300 mL. Measurements were done using a beaker with ±25 mL precision.

2.3. Model training and validation

We followed a transfer learning approach [18] to fine-tune our previous convolutional neural network (CNN) laboratory model [14]. This pretrained laboratory model was firstly loaded to optimize the learning task. Then, we trained the last layers of our model with the daily lifelike photographs. We separated 96 % of them for training (1867 images), 1.4 % for validation (28 images) and 2.6 %

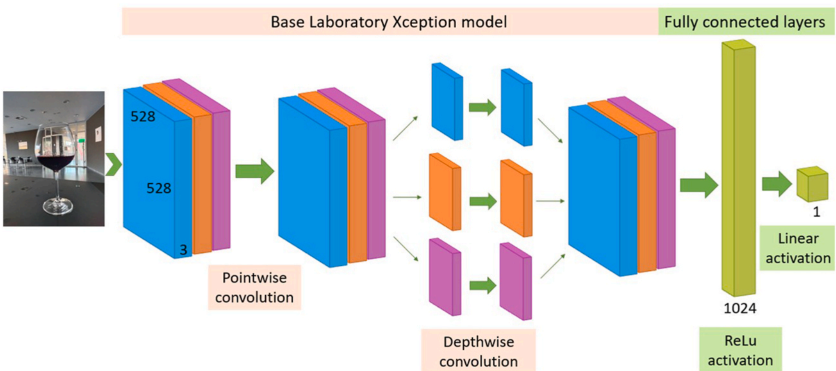

Fig. 1. Proposed CNN architecture for wine volume estimation. The numbers indicate the shape of the input images and the size of the final layers that result in the prediction.

for testing (50 images). We employed most of our images for training because further testing was going to be performed with consumers' photographs (real images dataset). In addition, we used early stopping to prevent overfitting. Thus, our model was trained until the validation loss failed to improve for 10 consecutive epochs.

The DL architecture utilized in this follow-up study to accurately measure wine volume is illustrated in Fig. 1. In brief, the proposed CNN architecture was pre-trained in the previous base laboratory Xception model [14], which consisted in several convolutional layers, subsequently connected with a global average pooling layer to the last fully connected layers, which were fine-tuned with the daily lifelike dataset training photographs. The model was trained with a GPU Tesla V100-PCIE-32GB using Keras and TensorFlow [19]. The code used in this study is publicly available [20].

Our model was evaluated on two independent tests: the subset of daily lifelike images (n = 50) and the real images dataset (n = 229). The later dataset served to perform an independent unbiased validation of our DL method, to test if the model was able to generalize to new non-optimal scenarios with different photographic cameras and lightning states.

2.4. Statistical analysis and interpretability saliency maps

To assess the performance of our new refined DL model, we evaluated wine volume predictions with the following regression metrics: Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE), as previously proposed [14]. We further evaluated the performance of the model for the different wine volumes available using violin plots and box plots. This statistical analysis was carried out using scikit-learn library [21] in Python 3.8.6.

The wine volume per serving recorded by the participants in the consumer study (real data) was compared with the values estimated by the novel DL method by a paired *t*-test. The differences in the wine volume per serving among individuals, time of consumption, type of wine, and type of glass were analyzed by the ANOVA test. The differences among gender and two types of glasses (balloon wine glass and short rock glass) were analyzed by the *t*-test. These statistical analyses were carried out using the program Excel of Microsoft Office plus 2019.

Interpretability saliency maps were computed to evaluate in which areas of the image the DL model was focusing to perform the wine volume predictions [22]. We assessed both gradient saliency and integrated gradients, following the same approach as in Cobo et al. [14].

3. Results

3.1. DL model evaluation

Table 1 depicts values of the regression metrics MAE (Mean Absolute Error) and RMSE (Root Mean Squared Error) for all the sets that were considered in this study: train, validation and test sets which come from the daily lifelike images dataset, as well as the real images dataset. Values for MAE and RMSE of our fine-tuned DL system in the testing subset (10 and 13 mL, respectively) were considered acceptable, indicating that the model was successfully optimized with the daily lifelike images. Supplementary Fig. 5 shows violin plots for the predictions of each wine volume values available in the daily lifelike images dataset.

To detect those parts of the image in which the model focused the most to make the wine volume estimation, saliency maps were evaluated for randomly selected sample test images. As an example of the daily lifelike images dataset, Fig. 2 depicts gradient saliency (also known as vanilla gradient) and integrated gradients both in its standard and smoothed version. Even though the glass container was partially covered by a finger, the DL model performed an adequate estimation, with 11 % relative error (Fig. 2).

Values for MAE and RMSE of our fine-tuned model in the real image dataset (participants' images) were slightly higher (26 and 33 mL, respectively) than in the daily lifelike image dataset (Table 1). Supplementary Fig. 6 shows violin plots for the predictions of each wine volume value available in the real images dataset. Additionally, we evaluated the percentage of photographs that were predicted with an error similar or higher than the precision of the beaker that was used to measure the volume (±25 mL). Thus, it was calculated that the wine volume in 91 out of the 229 images (39.7 %) was given with an error ≥25 mL.

As an example of the wide variability in wine volume estimation accuracy among consumers' images (real image dataset), Fig. 3 depicts saliency maps of photographs with various relative error percentages, which were attributed to different circumstances. Fig. 3A depicts saliency maps of a participant's image for which the model performed a very precise estimation (125 mL measured vs 125 mL estimated). In this case, the dark background behind the glass of wine was having a minor impact on the prediction. However, in other participants' images, background seemed to have a stronger impact on the prediction, especially when participants included other glass objects, even though they had specifically been advised to avoid them in their pictures, that induce a higher relative error in the

Table 1
Performance metrics (MAE and RMSE) evaluation for wine volume predictions with our DL model in daily lifelike images dataset (train, validation and test) and real images dataset.

Subset	*daily lifelike images* dataset			*real images* dataset
	Train	Validation	Test	Test
Images (n)	1867	28	50	229
MAE (mL)	5	14	10	26
RMSE (mL)	7	20	13	33

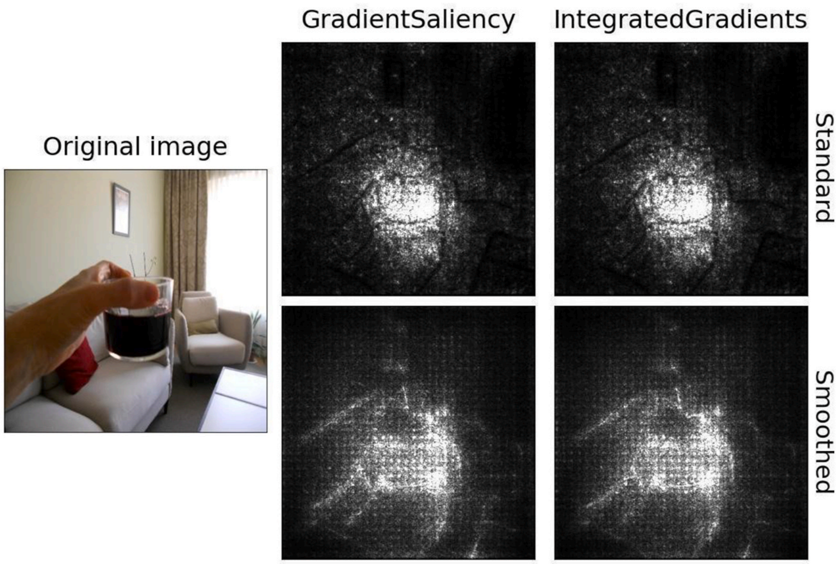

Fig. 2. Standard and smoothed saliency maps for daily lifelike image with a glass containing 125 mL of red wine. Estimated volume was 111.7 mL, (11 % relative error). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

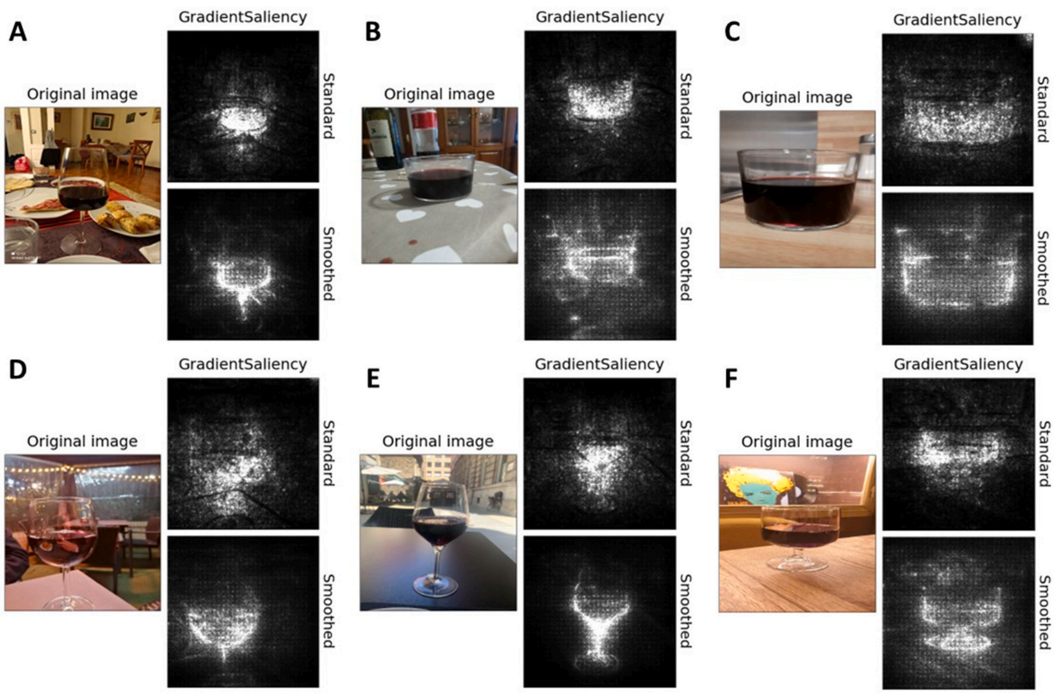

Fig. 3. Standard and smoothed saliency maps for real images: A) Glass containing 125 mL of red wine with an estimated volume of 125 mL; B) Glass containing 125 mL of red wine with an estimated volume of 100 mL (25 % relative error); C) Glass containing 100 mL of red wine with an estimated volume of 141 mL (41 % relative error); D) Glass containing 175 mL of rosé wine with an estimated volume of 132 mL (25 % relative error); E) Glass containing 75 mL of red wine with an estimated volume of 114 mL (52 % relative error); F) Glass containing 150 mL of red wine with an estimated volume of 102 mL (32 % relative error). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

estimation (Fig. 3B, relative error 25 %, 125 mL measured vs 100 mL estimated). In a few cases, participants did not seem to have followed the instructions indicated in the web application, because their photographs were taken at close distances/bad angles, such is the case in Fig. 3C (41 % relative error). In another example, the participant took a photograph of a glass filled with rosé wine instead of red wine, as depicted in Fig. 3D (25 % relative error, 125 mL measured vs 100 mL estimated), although the DL model was capable of predicting the volume for this new type of wine quite accurately, which could open up new applications of the model to other type of wines. In addition, in our visual inspection, we also detected possible error sources in the volume selected through the web application by the participant when the photograph was uploaded; this seemed to be the case in Fig. 3E (52 % relative error). Finally, there were also some photographs in which the participant used a different type of glass that the model had not seen before, this is the case in Fig. 3F (32 % relative error).

3.2. Assessment of individual wine consumption from the consumer study

In addition to the evaluation of the fine-tuned DL model, we analyzed the information collected about individual red wine consumption from the consumer study (real image dataset). The number of photographs (glasses of wine consumed) uploaded by each participant during the period of 3 weeks that lasted the study, varied between 17 and 1, with a mean of 6 photographs per participant (Supplementary Fig. 3). With regard to the time of red wine consumption, we observed that individual wine consumption mostly took place during the meals (lunch and dinner), with much less wine consumption outside mealtimes (Supplementary Fig. 4A). Of the three types of red wine consumed, the "crianza" wine was the most frequently consumed (Supplementary Fig. 4B). With regard to the type of glass used for the consumption, a considerable diversity was observed, although the Bordeaux type was the most frequent (Supplementary Fig. 4C).

Fig. 4 shows the distribution of the volume in a service (glass) estimated by our DL model from the photographs taken by the participants ($n = 229$) (real image dataset) in comparison with the volume directly measured by the participant using the graduated beaker. No significant differences ($p = 0.253$) between both variables were found by applying the paired *t*-test. The average volume in a service (glass) calculated from the estimations of our DL model was 116 ± 29 mL of red wine, close to the average of the volume measured by the participants (114 ± 33 mL of red wine).

The design of this consumer's study also allowed us to assess the influence of different factors in the volume of red wine served in a glass, such as individual, gender, time of consumption, type of wine, and type of glass (Fig. 5). As expected, significant ($p = 0.000$, ANOVA test) variability among individuals was observed in the volume of red wine in a service, from 50 to 275 mL (Fig. 5A). It should be noted that we suspected a possible upward error in the labelling of the only 275-mL-image, as also indicated by the volume estimated by our method (176 mL) (Supplementary Fig. 7). In relation to gender, in our study, no significant differences ($p = 0.881$, *t*-test) were found between male and female participants, although male participants showed wider variation intervals (Fig. 5B). No statistical differences were either observed in relation to the time of consumption ($p = 0.383$, ANOVA test) or the type of red wine ($p = 0.446$, ANOVA test) (Fig. 5C and D, respectively). No statistical differences for the type of glass were observed in general ($p = 0.235$, ANOVA test). However, when the larger capacity glasses were used (i.e., balloon wine glass), the tendency was to serve higher volumes of wine than when smaller capacity containers (i.e., short rock glass) were used ($p = 0.003$, *t*-test) (Fig. 5E).

The consumer's questionnaire included in the application web also allowed us to discover other qualitative information related to red wine consumption preferences. With regard to the frequency of red wine consumption, this was mainly between one or two days per week, following the same trend in men and in women (Supplementary Table 3). It should be noted that men reported drinking more frequently than women. The type of consumption, in all cases, was higher at weekends and the number of drinks was mostly around 1–2, with an increase in the percentage of 3–4 drinks in men (Supplementary Table 3). Both beer and other beverages were generally consumed occasionally, with beer consumption being higher (Supplementary Table 3). With regard to the circumstances of red wine consumption (time and place), the preferred consumption was at home; women were more likely to drink at home and at dinner, while men showed no clear preference (same percentage at home and away from home, as well as at lunch and dinner) (Supplementary Table 3). Around 78 % of the male participants in the study said that they never consumed low-alcoholic wine, compared to 41 % of the women, therefore, the results seem to point out a higher consumption of low-alcoholic wine by women (Supplementary Table 3).

When choosing red wine, the preferred factors, in descending order, were Denomination of Origin > Grape variety > Enjoying good

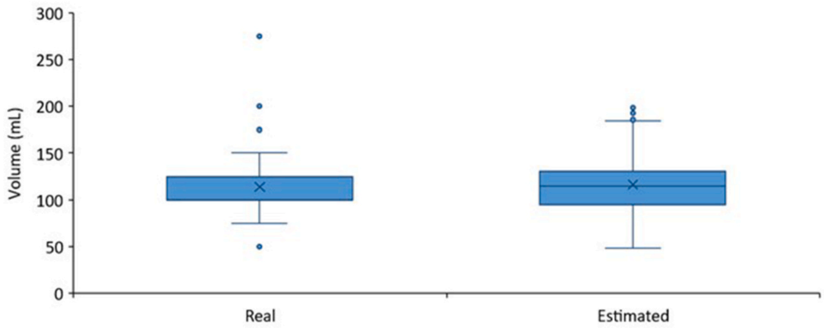

Fig. 4. Box plots of the volume of red wine in a service (glass) ($n = 229$) (real image dataset): A) real values, directly measured by the participant using the graduated glass, and B) values estimated by our DL model from the photographs taken by the participants. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

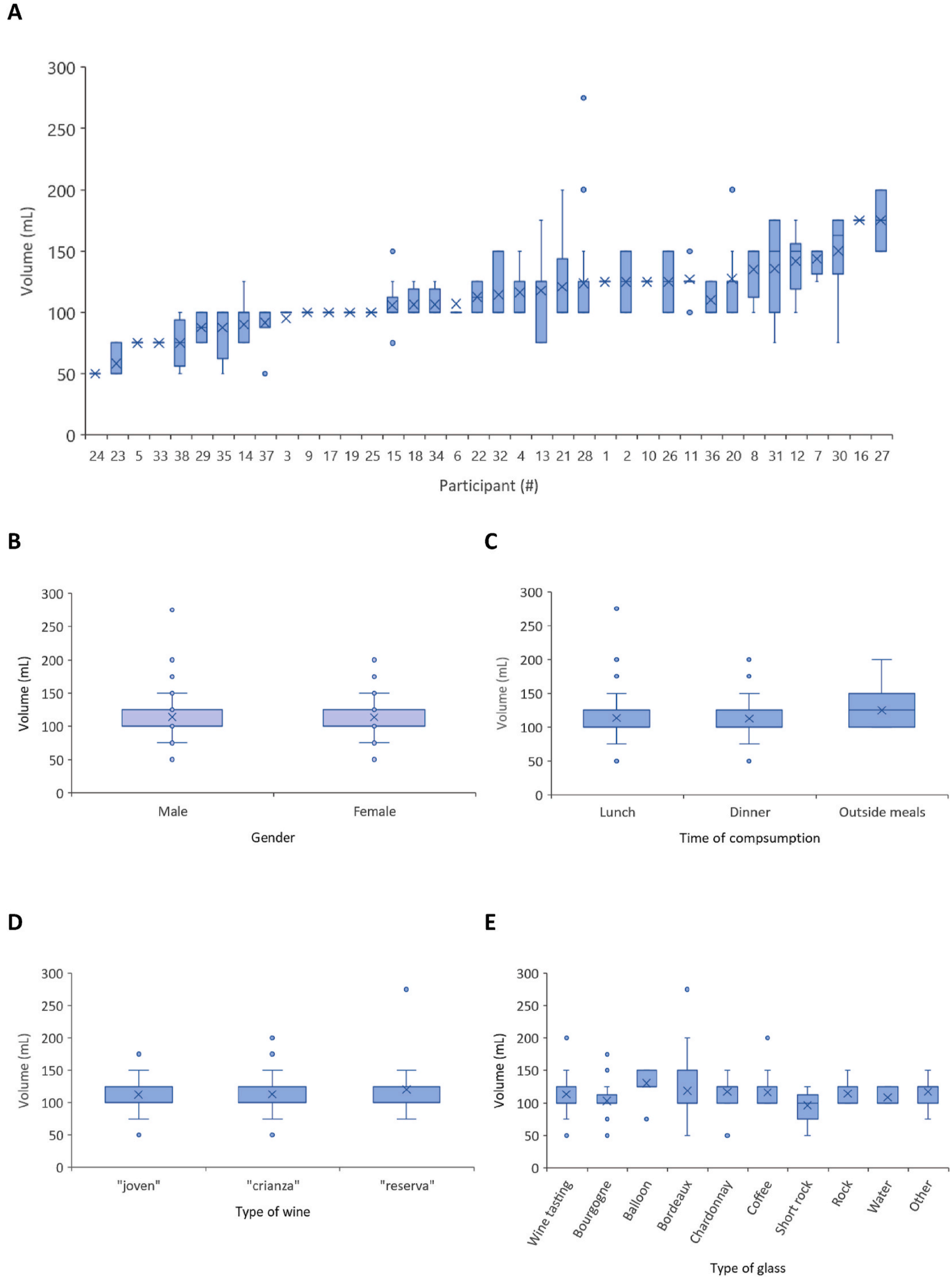

Fig. 5. Box plots of the volume of red wine in a service (glass) (n = 229) according to: A) participant, B) gender, C) time of consumption, D) type of wine, and E) type of glass. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

food > Wine production area (Supplementary Table 3). Women seem to take more into account the region than men, while around 17 % of the men surveyed were concerned about health when choosing wine, compared to 0 % of women (Supplementary Table 3). The general trend in wine-related activities was very similar for both men and women, with a predominance of purchases in supermarkets, followed by online purchases and tasting courses (Supplementary Table 3). Finally, with regard to knowledge of the world of wine, a higher percentage of men (58.8 %) looked at the information on the label compared to women (35.3 %) (Supplementary Table 3).

4. Discussion

Nowadays, the collection of data on beverage consumption patterns, particularly alcohol (ethanol) intake, is a sensitive topic in the field of public health. In fact, different epidemiological studies consistently found that light and moderate wine drinkers in the framework of the Mediterranean Diet seem to be at lower risk for cardiovascular disease (CVD) than non-drinkers, whereas heavy drinkers are at the highest risk [23]. In these circumstances, accurate measurements are needed to clearly establish what a light and/or moderate drinker is. The estimation of wine consumption through FFQs is particularly inaccurate, as a portion (glass of wine) consumed is standardized at 100 mL, which does not always correspond to reality.

Based on these limitations, in this paper we propose a strategy to estimate the red wine volume in a glass from a single-view image. For that, we have followed a previous strategy [14] based on regression CNNs. In fact, most of the known Image-Based Food-Recognition Systems have also adopted DL methods and especially CNNs in at least one phase of their development [24].

One of the first achievements of the present study was the generation of a public dataset of images of red wine consumption in lifelike situations (the *BrainGut_WineUp daily lifelike images* dataset) [15]. This is a free red wine glass labelled-image dataset composed of 1945 images that could be useful in future applications and modelling by the scientific community working in the field of wine components. In this line, one example of recently-published dataset related to whole food images and created for research studies in the field of nutrition was reported by [25]. In our study, the *BrainGut_WineUp daily lifelike images* dataset allowed us to fine-tune the previous DL model [14] following a CNN architecture, typical of computer vision tasks. This architecture has been widely used in deep learning, together with transfer learning [24,26], particularly in image classification and recognition tasks. The evaluation of the enhanced deep learning model using images from the daily lifelike dataset demonstrated reliable and accurate wine volume predictions (MAE = 10 mL and RMSE = 13 mL). These values were comparable to those reported for the previous DL model constructed with laboratory images (8 and 11 mL, respectively) [14]. Therefore, the initial laboratory model was successfully optimized, achieving performance in real life images comparable to the laboratory images, while requiring significantly fewer images due to the application of the transfer learning approach. To our knowledge, this is the first validated digital-based method for determining wine volume from single-view images. Among other advantages, this new model only requires one photograph to be taken (single-view images) and no reference object, which will significantly facilitate its future use [7].

The proposed deep learning (DL) method was specifically fine-tuned for red wine images, demonstrating its effectiveness in directly measuring wine volume from a single-view image. However, as evidenced in Fig. 3D, it also yields adequate results with rosé wine. Applying the method to other types of wine and beverages would involve fine-tuning the model to some extent, depending on the level of accuracy required for the task.

In parallel, we carried out an ad hoc consumer study via web service that allowed us to generate a new images dataset (the *BrainGut_WineUp real images* dataset) with which the new model was also tested. In this case, wine volume predictors showed higher values (MAE = 26 mL and RMSE = 33 mL), under or over-reporting in relation to "real" wine volume measured. In order to explain this finding and identify possible errors in real images, visual inspection was undertaken to evaluate the quality of participants' images. The relatively higher error observed was reasonably expected as photographs were taken by consumers in everyday situations without paying special attention to the indications that were given to them in the web application for taking the photographs. Moreover, the most common errors detected were related to the presence of additional objects, the distance/angle at which the photographs were taken or the use of rosé wine instead of red wine, although the error detected in this specific case was lower than in the other situations. Another plausible cause that could explain the high percentage of misestimated photos was the fact that the participants approximated the volume measured in the beaker to specific markings with 25 mL intervals (50, 75, 100, 125, 150, 175, 200, 225, 250, 275 and 300 mL) that were pre-set in the web application. In contrast, for the generation of the "daily lifelike" images, glasses were filled with volumes measured by test tube ($\pm$0.5 mL precision) at 25 mL intervals, which led to a higher precision in the estimation (Supplementary Figs. 5 and 6). Indeed, this is an aspect of the present work that should be improved in future studies in which participants would be asked to annotate the exact measured volume without using pre-set markings. Therefore, for future studies, we recommend including more detailed instructions for the photograph setting and allowing the participants freely set the volume; we believe that both changes would improve the accuracy of the wine volume estimation by the developed DL model.

The consumer study carried out via web service also provided valuable data about wine volume in a glass (volume per serving) as well as data about a pilot study on the intake of red wine components, which is another contribution of this paper. The average volume in a service (glass) from the 229 photographs taken by the 38 participants was 114 $\pm$ 33 mL of red wine, lower than the average volume reported in a study in the Netherlands carried out with 141 participants (131.7 mL) from 392 measurements of red wine servings [27]. Multiple factors would explain differences among studies, but, in any case, all findings suggest that the standardization of 100 mL per wine serving considered in many dietary recordings might be slightly underestimated and, in some dietary patterns, this may lead to non-accurate conclusions.

Among the factors considered in this study to affect the volume of red wine served in a glass (Fig. 5), only the type of glass seemed to influence this wine consumption's characteristic. Concretely, the balloon wine glass (largest capacity glass) tended to lead to greater servings than the short rock glass (smallest capacity glass) (Fig. 5E). Although there are not many studies in the literature to compare with, a study focused on wine consumption at home [12] found that households consumed on average 6.5 % less wine when drinking from smaller glasses (290 mL) than from larger (350 mL) wine glasses. In a laboratory study [28], the authors also found significantly lower volumes when wine was self-served in small and medium glasses in comparison to large glasses. In relation to gender, some authors [29], in an experiment of water and red wine-pouring, found that men tended to pour more liquid into the glasses than women, a factor that was weakly appreciated in our study.

Collection of daily wine intake through digital-based methods such as the one developed in the present study would allow accurate estimation of the dietary intake of alcohol and other beneficial components, as polyphenols, coming from wine consumption. This is of great interest in the context of the Mediterranean diet, which is probably one of the most widely evaluated and recommended dietary patterns to maintain health that includes a light-to-moderate consumption of red wine [9]. For example, a daily intake of 114 mL of red wine –the average value for a wine serving found in this study– would provide 80.9 Kcal (71 Kcal per 100 mL of red wine), 0.342 g of sugars (0.3 g per 100 mL or red wine), 11.2 g of alcohol (12.5° of ethanol per 100 mL of red wine) in the diet. Concerning wine active phytochemicals (i.e., polyphenols), a red wine serving of 114 mL would provide a total of 137–342 mg of total polyphenols if a concentration of 1.2–3.0 g per litre of wine is considered. Moreover, and considering an average value for daily polyphenol intake of 1011 mg/day in Mediterranean countries (such as Spain) [30], a red wine serving of 114 mL would represent the 14–34 % of dietary polyphenol intake, depending on the polyphenol richness of the wine. These data should be taken into account when, for example, making dietary recommendations to specific groups.

In practice, the use of an automated algorithm for diet monitoring such as the one developed in this study, either by researchers, dietetics and general public, requires its implementation in mobile phone applications [31]. In this sense, the present DL model will be the basis of a forthcoming app that will allow us to estimate the volume of wine from a photograph using the mobile phone.

In summary, this study reports the development of a new refined DL model for the volume estimation of red wine in a glass, with the findings demonstrating the performance and feasibility of this model in a small cohort of red wine consumer adults. The proposed DL model constitutes a simple and effective automatic tool to measure red wine volume and overcomes general drawbacks of conventional food frequency questionnaires, where liquid measurement is subjective and imprecise, as it relies on participant's recall. Moreover, when integrated with the web application utilized in this study, the developed deep learning model facilitates the estimation of wine component intake and its association with health outcomes. In the future, these imaging and computational technologies could be further extended to other foods, offering diverse applications in food science and nutrition.

Ethical approval statement

The study was approved by the Research Ethics Committee of CSIC (Madrid, Spain) (Approval number 081/2021) and guaranteed the participant's data protection. Written informed consent was obtained from the participants.

Data availability statement

Data associated with this study has been deposited at DIGITAL. CSIC under https://digital.csic.es/handle/10261/256232 (for BrainGut_WineUp laboratory images) and https://digital.csic.es/handle/10261/284780 (for BrainGut_WineUp real images).

CRediT authorship contribution statement

Miriam Cobo: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation. **Edgard Relaño de la Guía:** Writing – review & editing, Formal analysis, Data curation. **Ignacio Heredia:** Writing – review & editing, Formal analysis, Data curation. **Fernando Aguilar:** Writing – review & editing, Visualization, Validation, Supervision, Software, Methodology, Formal analysis, Data curation. **Lara Lloret-Iglesias:** Writing – review & editing, Visualization, Validation, Supervision, Software, Methodology, Formal analysis, Data curation. **Daniel García:** Writing – review & editing, Formal analysis, Data curation. **Silvia Yuste:** Writing – review & editing, Formal analysis, Data curation. **Emma Recio-Fernández:** Writing – review & editing, Formal analysis, Data curation. **Patricia Pérez-Matute:** Writing – review & editing, Writing – original draft, Project administration, Investigation, Funding acquisition, Conceptualization. **M. José Motilva:** Writing – review & editing, Writing – original draft, Project administration, Investigation, Funding acquisition, Conceptualization. **M. Victoria Moreno-Arribas:** Writing – review & editing, Writing – original draft, Project administration, Investigation, Funding acquisition, Conceptualization. **Begoña Bartolomé:** Writing – review & editing, Writing – original draft, Project administration, Investigation, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This study was supported by MCIN (Ministerio de Ciencia e Innovación)/AEI (Agencia Estatal de Investigación)/10.13039/501100011033 and the European Union NextGenerationEU/PRTR through the projects PID2019-108851RB-C21 and PID2019-108851RB-C22, and 'Prueba de Concepto' PDC2022-133861-C21 and PDC2022-133861-C22. M. Cobo acknowledges support from Consejo Superior de Investigaciones Científicas (CSIC) and Institute of Physics of Cantabria (IFCA) through the CSIC Interdisciplinary platform (PTI+) Global Health, and the Ministry of Education of Spain (FPU grant, reference FPU21-04458). I. Heredia was supported by the Gobierno de Cantabria with the project "Instrumentación y Ciencia de Datos para sondear la naturaleza del Universo". The authors would like to thank CSIC Interdisciplinary Thematic Platform (PTI+) Digital Science and Innovation).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e35689.

References

[1] M.C. Dao, A.F. Subar, M. Warthon-Medina, J.E. Cade, T. Burrows, R.K. Golley, N.G. Forouhi, M. Pearce, B.A. Holmes, Dietary assessment toolkits: an overview, Publ. Health Nutr. 22 (2019) 404–418, https://doi.org/10.1017/S1368980018002951.

[2] G. Ribeiro, A. Ferri, G. Clarke, J.F. Cryan, Diet and the microbiota - gut - brain-Axis: a primer for clinical nutrition, Curr. Opin. Clin. Nutr. Metab. Care 25 (2022) 443–450, https://doi.org/10.1097/MCO.0000000000000874.

[3] C. Whitton, C. Ramos-García, S.I. Kirkpatrick, J.D. Healy, S.S. Dhaliwal, C.J. Boushey, C.E. Collins, M.E. Rollo, D.A. Kerr, A systematic review examining contributors to misestimation of food and beverage intake based on short-term self-report dietary assessment instruments administered to adults, Adv. Nutr. 13 (2020) 2620–2665, https://doi.org/10.1093/advances/nmac085.

[4] G. Tabacchi, A.R. Filippi, E. Amodio, M. Jemni, A. Bianco, A. Firenze, C.A. Mammina, Meta-analysis of the validity of FFQ targeted to adolescents, Publ. Health Nutr. 19 (2016) 1168–1183, https://doi.org/10.1017/S1368980015002505.

[5] J.T. Kelly, P.F. Collins, J. McCamley, L. Ball, S. Roberts, K.L. Campbell, Digital disruption of dietetics: are we ready? J. Hum. Nutr. Diet. 34 (2021) 134–146, https://doi.org/10.1111/jhn.12827.

[6] K. v Dalakleidi, M. Papadelli, I. Kapolos, K. Papadimitriou, Applying image-based food-recognition systems on dietary assessment: a systematic review, Adv. Nutr. 13 (2022) 2590–2619, https://doi.org/10.1093/advances/nmac078.

[7] Z. Yang, H. Yu, S. Cao, Q. Xu, D. Yuan, H. Zhang, W. Jia, Z.-H. Mao, M. Sun, Human-mimetic estimation of food volume from a single-view RGB image using an AI system, Electronics (Basel) 10 (2021) 1556, https://doi.org/10.3390/electronics10131556.

[8] M.A. Martínez-González, A. Gea, M. Ruiz-Canela, The mediterranean diet and cardiovascular health, Circ. Res. 124 (2019) 779–798, https://doi.org/10.1161/CIRCRESAHA.118.313348.

[9] S. Minzer, R. Estruch, R. Casas, Wine intake in the framework of a mediterranean diet and chronic non-communicable diseases: a short literature review of the last 5 years, Molecules 25 (2020) 5045, https://doi.org/10.3390/molecules25215045.

[10] R. Pechey, D.-L. Couturier, G.J. Hollands, E. Mantzari, M.R. Munafò, T.M. Marteau, Does wine glass size influence sales for on-site consumption? A multiple treatment reversal design, BMC Publ. Health 16 (2016) 390, https://doi.org/10.1186/s12889-016-3068-z.

[11] E. Mantzari, M. Ventsel, J. Ferrar, M.A. Pilling, G.J. Hollands, T.M. Marteau, Impact of wine bottle and glass sizes on wine consumption at home: a within- and between- households randomized controlled trial, Addiction 117 (2022) 3037–3048, https://doi.org/10.1111/add.16005.

[12] Y. Probst, V. Guan, K. Kent, A systematic review of food composition tools used for determining dietary polyphenol intake in estimated intake studies, Food Chem. 238 (2018) 146–152, https://doi.org/10.1016/j.foodchem.2016.11.010.

[13] M. Cobo, I. Heredia, F. Aguilar, L. Lloret Iglesias, D. García, B. Bartolomé, M.V. Moreno-Arribas, S. Yuste, P. Pérez-Matute, M.-J. Motilva, Artificial intelligence to estimate wine volume from single-view images, Heliyon 8 (2022) e10557, https://doi.org/10.1016/j.heliyon.2022.e10557.

[14] B. Bartolomé, M.V. Moreno-Arribas, L. Lloret, F. Aguilar, M. Cobo, D. García, I. Heredia, S. Yuste, P. Pérez-Matute, M.J. Motilva, *BrainGut_WineUp laboratory images*. DIGITAL.CSIC. https://digital.csic.es/handle/10261/256232, 2021. (Accessed 13 June 2023).

[15] B. Bartolomé, M.V. Moreno-Arribas, L. Lloret, F. Aguilar, M. Cobo Cano, D. García Díaz, I. Heredia, S. Yuste, E. Recio-Fernández, P. Pérez-Matute, M.J. Motilva, BrainGut_WineUp daily lifelike images, DIGITAL.CSIC (2022). https://doi.org/10.20350/digitalCSIC/14816. (Accessed 13 June 2023).

[16] E. Relaño de la Guía, L. Lloret, F. Aguilar, M. Cobo Cano, D. García Díaz, I. Heredia, S. Yuste, E. Recio-Fernández, P. Pérez-Matute, M.J. Motilva, B. Bartolomé, M.V. Moreno-Arribas, BrainGut_WineUp real images. DIGITAL.CSIC (2022). https://digital.csic.es/handle/10261/284780. (Accessed 13 June 2023).

[17] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, Q. He, A comprehensive survey on transfer learning, Proc. IEEE 109 (2021) 43–76, https://doi.org/10.1109/JPROC.2020.3004555.

[18] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mane, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viegas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, X. Zheng, TensorFlow: large-scale machine learning on heterogeneous distributed systems, arXiv 1603 (2016) 04467, https://doi.org/10.48550/arXiv.1603.04467.

[19] M. Cobo, BrainGut_WineUp tool for wine volume estimation: consumer study validation, DIGITAL.CSIC (2022). https://digital.csic.es/handle/10261/288940. (Accessed 27 February 2023).

[20] O. Kramer, Scikit Learn-Machine Learning for Evolution Strategies first ed., 20, Springer, Cham, 2016.

[21] K. Simonyan, A. Vedaldi, A. Zisserman, Deep inside convolutional networks: visualising image classification models and saliency maps, in: Workshop at International Conference on Learning Representations, 2014. Citeseer.

[22] E.L. O'Keefe, J.J. DiNicolantonio, J.H. O'Keefe, C.J. Lavie, Alcohol and CV health: jekyll and hyde J-curves, Prog. Cardiovasc. Dis. 61 (2018) 68–75, https://doi.org/10.1016/j.pcad.2018.02.001.

[23] G. Tahir, C.K. Loo, A Comprehensive Survey of Image-Based Food Recognition and Volume Estimation Methods for Dietary Assessment, 9, 2021. Healthcare (Switzerland).

[24] K. Weiss, T.M. Khoshgoftaar, D.A. Wang, Survey of transfer learning, J. Big. Data 3 (2016) 9, https://doi.org/10.1186/s40537-016-0043-6.

[25] N. Clarke, E. Pechey, R. Pechey, M. Ventsel, E. Mantzari, K. De-Loyde, M.A. Pilling, R.W. Morris, T.M. Marteau, G.J. Hollands, Size and shape of plates and size of wine glasses and bottles: impact on self-serving of food and alcohol, BMC Psychol 9 (2021) 163, https://doi.org/10.1186/s40359-021-00645-z.

[26] R. Zamora-Ros, V. Knaze, J.A. Rothwell, B. Hémon, A. Moskal, K. Overvad, A. Tjønneland, C. Kyrø, G. Fagherazzi, M.-C. Boutron-Ruault, M. Touillaud, V. Katzke, T. Kühn, H. Boeing, J. Förster, A. Trichopoulou, E. Valanou, E. Peppa, D. Palli, C. Agnoli, F. Ricceri, R. Tumino, M.S. de Magistris, P.H.M. Peeters, H. B. Bueno-de-Mesquita, D. Engeset, G. Skeie, A. Hjartåker, V. Menéndez, A. Agudo, E. Molina-Montes, J.M. Huerta, A. Barricarte, P. Amiano, E. Sonestedt, L. M. Nilsson, R. Landberg, T.J. Key, K.-T. Khaw, N.J. Wareham, Y. Lu, N. Slimani, I. Romieu, E. Riboli, A. Scalbert, Dietary polyphenol intake in Europe: the European prospective investigation into cancer and nutrition (EPIC) study, Eur. J. Nutr. 55 (2016) 1359–1375, https://doi.org/10.1007/s00394-015-0950-x.

[27] B.N. Limketkai, K. Mauldin, N. Manitius, L. Jalilian, B.R. Salonen, The age of artificial intelligence: use of digital technology in clinical nutrition, Curr. Surg. Rep 9 (2021) 20, https://doi.org/10.1007/s40137-021-00297-3.

[28] Y.-L. Chen, Y.-C. Lee, Effects of lighting, liquid color, and drink container type on volume perception, I-perception 10 (2019), https://doi.org/10.1177/2041669519880916, 204166951988091.

[29] M.F. de Beukelaar, M.L. Janse, A. Sierksma, E.J. Feskens, J.H. de Vries, How full is your glass? Portion sizes of wine, fortified wine and straight spirits at home in The Netherlands, Publ. Health Nutr. 22 (2019) 1727–1734, https://doi.org/10.1017/S1368980019000442.

[30] S.P. Mohanty, G. Singhal, E.A. Scuccimarra, D. Kebaili, H. Héritier, V. Boulanger, M. Salathé, The food recognition benchmark: using DeepLearning to recognize food on images, Front. Nutr. 9 (2021) 875143, https://doi.org/10.3389/fnut.2022.875143.

[31] I. Trinidad Rodríguez, J. Fernández Ballart, G. Cucó, E. Biarnés Jordà, V. Arija Val, Validation of a short questionnaire on frequency of dietary intake: reproducibility and validity, Nutr. Hosp. 23 (2008) 242–252.