Research article

# Artificial intelligence to estimate wine volume from single-view images

Miriam Cobo [a,*], Ignacio Heredia [a], Fernando Aguilar [a], Lara Lloret Iglesias [a], Daniel García [a], Begoña Bartolomé [b], M. Victoria Moreno-Arribas [b], Silvia Yuste [c], Patricia Pérez-Matute [d], Maria-Jose Motilva [c]

[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] *Center for Biomedical Research of La Rioja (CIBIR), 26006 Logroño (La Rioja), Spain*

## ARTICLE INFO

## ABSTRACT

In this paper, we present a method to determine the volume of wine in different types of glass liquid containers from a single-view image. The proposed model predicts red wine volume from a photograph of the glass containing the wine. Experimental results demonstrated satisfactory performance of our image-based wine measurement system, with a Mean Absolute Error lower than 10 mL. To train and evaluate our system, we introduced the WineGut_BrainUp dataset, a new dataset of glasses of wine that contains 24305 laboratory images, including a wide range of containers, volumes of wine, backgrounds, object distances, angles and lightning, with or without calibration object. The proposed methodology is a suitable analytical tool for automate measurement of red wine volume. Indeed, it has potential real life applications in diet monitoring and wine consumption studies.

## 1. Introduction

Accurate measurement of dietary intakes is crucial for researchers and public health institutions in studies aiming to improve general/specific health outcomes (De Rijk et al., 2021; González-Alzaga et al., 2022). Diet assessment is mainly based on Food Frequency Questionnaires (FFQs) that inquire individuals about the frequency at which they consume different food items from a predefined list (Sotos-Prieto et al., 2015). To calculate the grams of each food item consumed per day, frequency data are multiplied by the portion size of each food. Then, data (grams of food/day) are converted into daily nutrients intake by using food composition databases. Finally, daily nutrient intakes from the different food items are summed to obtain total daily intake of each nutrient. Main limitations of FFQs are that they rely on subject's recall and might not accurately estimate the portion size/volume, and, consequently, are associated with misreporting. In the case of wine, estimation of its consumption through FFQs is particular quite imprecise as a portion (glass of wine) consumed is standardized to 100 cc, when this is not always real. Besides, and depending above all on the type of glass, this volume may be lower or higher than that estimated through

surveys (Pechey et al., 2016). Thus, the use of images could improve the quality of the data obtained on dietary assessment improving the estimation of portion size/volume (Yang et al., 2021; Jia et al., 2019; Fang et al., 2015). In fact, the problem of measuring liquid volume in an image can be solved applying Machine Learning & Artificial Intelligence techniques, but is not deeply supported by research yet. One of the main challenges faced by this task is the fact that, in order to provide a measurement in milliliters, images shall include reference information, which is given by a calibration object inserted in the image, such as a checkerboard pattern (Fang et al., 2015, 2016; Siswantoro et al., 2014).

Food portion size/volume measurement from a single image of food items contained on a plate has been explored by (Chen et al., 2013). The proposed model includes a sophisticated implementation with adaptive thresholding and snake modelling. After food portion segmentation, a general 3D model that represented typical shapes of food had to be selected. The accuracy reached was quite high even though a single-view 2D photographic image does not contain the complete spatial information of the object (Chen et al., 2013). The authors presented an
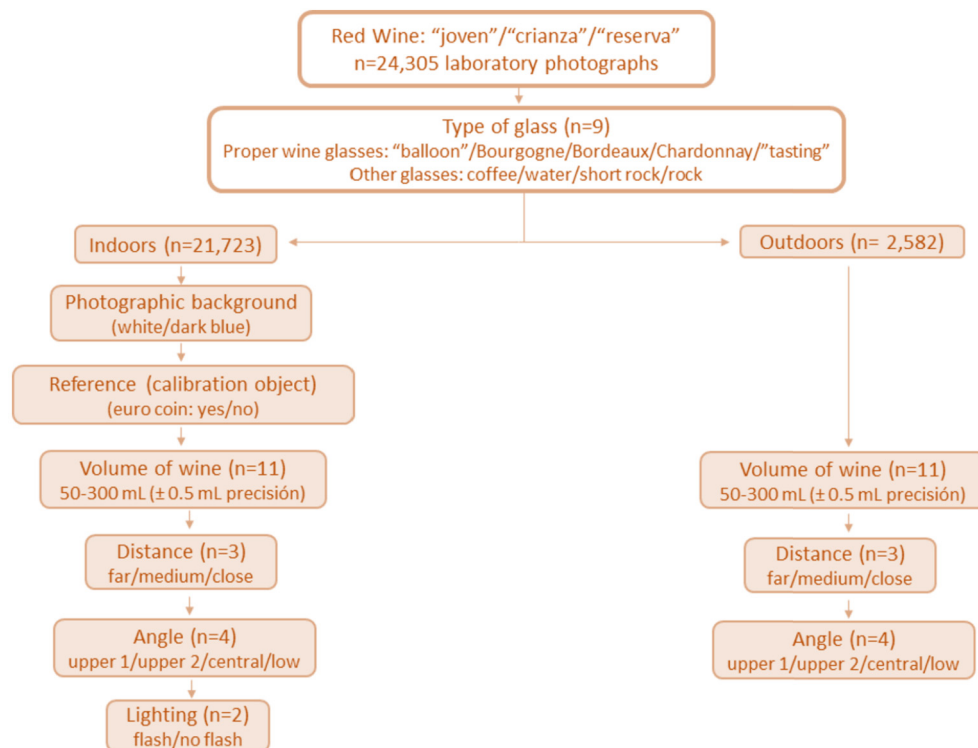
**Fig. 1.** Flowchart followed for the construction of WineGut_BrainUP dataset.

average error of 4% in food volume estimation. This method used the plate as scale reference and required measurements of its radius and depth parameters, that were provided by the user. Related research has been done towards volume measurement of irregular shape food using a computer vision system based on Monte Carlo method (Siswantoro et al., 2014). In spite of the refined hardware and software developed to carry out the measurements, the initial requirement of taking five images from different views of the object obscures real life diet monitoring applications of the system.

Container volume estimation and liquid content estimation are some of the tasks in which (Mottaghi et al., 2017) focused on. Their pictures included at least four extra objects in each image in order to avoid inconclusive results, as these objects were able to provide contextual information which was helpful in the task of estimating the volume of the container. The proposed convolutional neural network (CNN) followed a classification problem approach as it led to a better performance (Mottaghi et al., 2017). The suggested method achieved 32% per-class accuracy in content estimation, which is an insufficient overall performance for our use case, also due to the inherent challenges posed by the problem.

With this background, in this paper, we propose a method to estimate from a single-view RGB (Red, Green and Blue) photograph the volume of red wine in a glass without the need to incorporate a reference pattern in the image. For this purpose, we trained a deep learning model to estimate wine volume on the most common types of wine glasses. Our experimental evaluations on the specific dataset developed for our study, the WineGut_BrainUp dataset, show promising results, with a performance that reduces the Mean Absolute Error (MAE) to less than 10 mL. This method is suitable to measure in a fast, simple and efficient way the volume of wine in a glass, overcoming the limitations of FFQs, where volume estimations are subjective and error-prone. Instead, our proposed model is aimed to provide a generalized automatic tool for measuring liquid volume in nutritional studies and dietary assessment.

## 2. Material and methods

We propose a regression convolutional neural network (CNN) to carry out red wine volume estimation from photos of wine glasses. The code is based on the image classification module available in the DEEP Open Catalogue (García et al., 2020). The original classification model developed in the DEEP framework was adapted for regression[1] (Cobo, 2021), as such approach led to an improved performance in our task.

### 2.1. WineGut_BrainUP dataset

The WineGut_BrainUP dataset (Bartolomé et al., 2021) is the specific dataset that was created to develop our study. It includes 24305 laboratory photographs of glasses containing red wine that were taken in the laboratories of the Institute of Food Science Research (CIAL-CSIC), Institute of Grapevine and Wine Sciences (ICVV-CSIC) and the Center for Biomedical Research of La Rioja (CIBIR). Three commercial red wines representative of "joven", "crianza" and "reserva" wines were selected to take the photographs. For each wine, the same flowchart was followed, as represented in 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. The wine tasting glass shape used is the wine glass with specifications defined in ISO 3591:1977 (International Organization for Standardization, 1977). An example of each of the glasses is depicted in Fig. 2. The average measurements of these glasses are included in Table 1.
- Volume of wine (n = 11): 50, 75, 100, 125, 150, 175, 200, 225, 250, 275 and 300 mL. Measurements were done using a test tube with ±0.5 mL precision.

---

[1] https://github.com/MiriamCobo/BrainGut-WineUp.git.
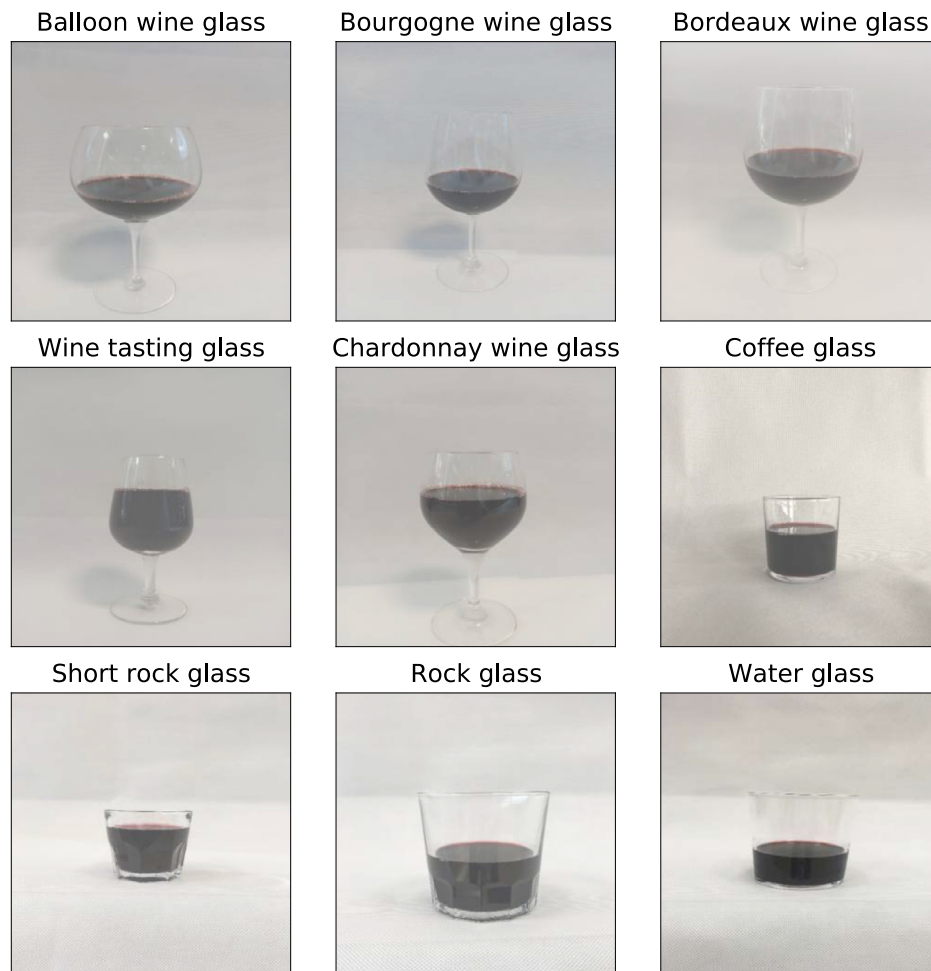
**Fig. 2.** Glasses used in the WineGut_BrainUp dataset. In these examples, glasses were filled with 150 mL (balloon, Bourgogne Bordeaux, tasting and Chardonnay wine glasses) or 100 mL (coffee, short rock, rock and water glasses) of red wine.

**Table 1**. Average size of the wine glasses used in this study.

| Type of glass | Volume (mL) | Height (cm) | Maximum diameter (cm) | Opening diameter (cm) |
|---|---|---|---|---|
| Balloon wine glass | 765 | 11.0 | 10.9 | 8.4 |
| Bourgogne wine glass | 815 | 13.5 | 10.8 | 7.0 |
| Bordeaux wine glass | 495 | 12.0 | 8.5 | 6.6 |
| Chardonnay wine glass | 315 | 9.8 | 8.0 | 6.0 |
| Wine tasting glass | 215 | 10.0 | 6.5 | 5.4 |
| Coffee glass | 185 | 5.8 | 8.1 | 8.1 |
| Water glass | 315 | 8.9 | 8.4 | 8.4 |
| Short rock glass | 135 | 7.4 | 7.2 | 7.2 |
| Rock glass | 235 | 9.5 | 7.6 | 7.6 |

- Object distance (n = 3): far (50-70 cm), medium (20-30 cm) and close (10-15 cm).
- Angle (n = 4): upper 1 (0, 30)°, upper 2 (30, 60)°, central (0°), low (-30, 0)°.

Additionally, the following fields were considered for indoors pictures:

- Photographic background (n = 2): white, dark blue.
- Reference (n = 2): yes, no. A 1€ coin was taken as possible reference (calibration object).
- Lighting (n =2): flash, no flash.

Fig. 3 shows four examples of laboratory images for different containers, backgrounds (outdoors, white and blue photographic backgrounds), angles, object distances and reference examples. This reference was not taken into consideration in the model as a calibration object. The purpose of the reference is to serve as an example of an extra object incorporated in the image and see if it has an influence in the predictions (see an example in Fig. 3).

### 2.2. Model training

The procedure to estimate wine volume in a glass container first consisted in training the convolutional neural network with laboratory images. Convolutional neural networks (also known as CNNs or ConvNets) are a type of deep learning neural networks particularly designed for analyzing image data, both with numerical or categorical labels. To predict continuous numerical data, convolutional neural networks include a regression layer at the end of the network. As a result, the output of our model is a real number (regression approach) instead of being an integer class (classification problem). The regression layer

**Table 2**. Number of images available in the WineGut_BrainUP dataset for every wine volume.

| Volume (mL) | 50 | 75 | 100 | 125 | 150 | 175 | 200 | 225 | 250 | 275 | 300 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| # train | 2187 | 2194 | 2194 | 2195 | 2025 | 2025 | 1508 | 1393 | 1261 | 1178 | 919 |
| # validation | 300 | 300 | 300 | 301 | 277 | 277 | 207 | 191 | 173 | 161 | 126 |
| # test | 299 | 301 | 301 | 300 | 277 | 278 | 206 | 191 | 173 | 161 | 126 |
| # all | 2786 | 2795 | 2795 | 2796 | 2579 | 2580 | 1921 | 1775 | 1607 | 1500 | 1171 |
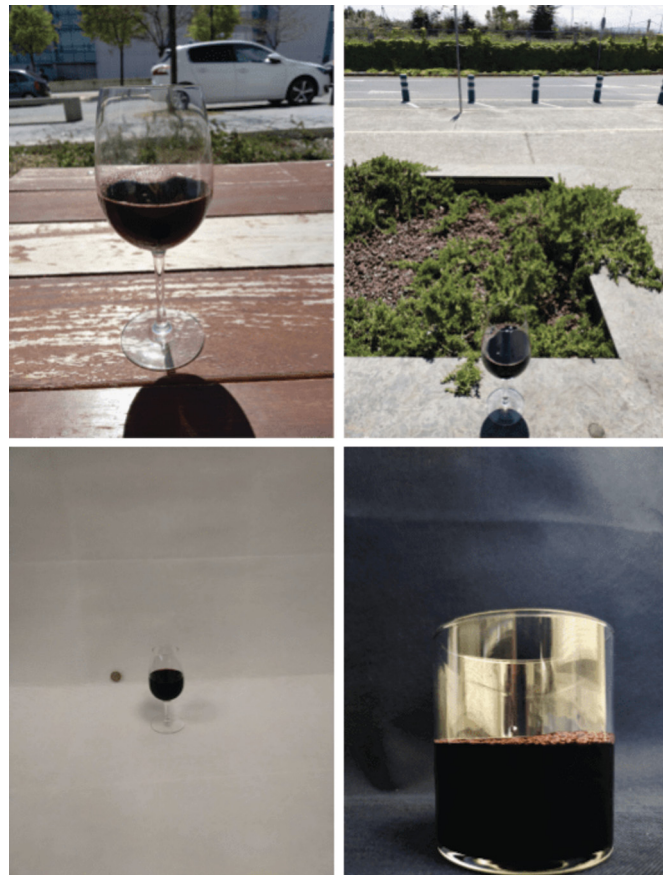


**Fig. 3.** Examples of laboratory images in the WineGut_BrainUP dataset.

is a fully-connected (dense) layer with a single node and a linear activation function, which replaces the fully-connected softmax classifier layer typically used for classification (Rosebrock, 2019). In addition, CNN regression models are trained with a continuous value prediction loss function, which in our case was the mean squared error function.

From the 24305 total number of photographs available, we separated around 80% for training (19079 photographs), 10% for validation (2613 photographs) and 10% for testing (2613 photographs). The number of images available for each volume and set is shown in Table 2. All sets were balanced with respect to the number of photographs belonging to each volume class. The training set is used to train the model. During the training phase, the hyperparameters are tuned in order to optimize the model's performance over the validation set. The test set contains photographs that the model has not previously seen and thus is used to assess the final unbiased accuracy of the model.

All types of liquid containers were included in the same network, as predictions showed similar validation metrics performance if separate models were trained for the proper wine glasses (balloon wine glass, Bourgogne wine glass, Bordeaux wine glass, Chardonnay wine glass and wine tasting glass) and the other glasses (coffee glass, water glass, short rock glass, and rock glass). In addition, training independent models for each type of image background gave similar results, and the network precision was comparable to the precision of the global CNN model presented in this paper.

**Table 3**. Regression metrics (MAE and RMSE) evaluation for wine volume predictions with our model.

| Set | MAE (mL) | RMSE (mL) |
|---|---|---|
| Training | 2 | 3 |
| Validation | 8 | 12 |
| Test | 8 | 11 |

We use a Xception (Chollet, 2017) neural network with images of size $528 \times 528$ pixels. In our method, dataset images were resized in order to meet this requirement. The batch size was set to 16, the number of training epochs was fixed up to 50 and we employed Adam optimizer (Kingma and Ba, 2017; Loshchilov and Hutter, 2019). A pretrained ImageNet base model was first loaded. Then, custom layers were included to adapt the model to the volume estimation task by using a linear activation function in the last layer. As the number of available images was quite high (24305 laboratory images), we disabled data augmentation in our model, also in order to reduce computation time.

The model was trained with a GPU Tesla V100-PCIE-32GB for 20 hrs. The model was coded using Keras (Chollet et al., 2015) and TensorFlow version 1.14.0 (Abadi et al., 2015) in Ubuntu 18.04.2 LTS.

## 3. Results and discussion

Different experiments were conducted to assess the quality of the developed models. Independent CNN were trained and evaluated with test laboratory images. The following results are referred to the final CNN model which was described in Section 2.2.

### 3.1. Performance evaluation of the proposed model in laboratory images

#### 3.1.1. Saliency maps examples for predicted images

In this section we present saliency maps for two sample test images, in order to detect those parts of the image in which the model focused to make the wine volume estimation. We depict gradient saliency (also known as vanilla gradient) (Simonyan et al., 2013) and integrated gradients maps (Sundararajan et al., 2017) both in its standard and smoothed (Smilkov et al., 2017) version (Anh, 2018), as it is illustrated in Fig. 4.

The explanation provided in Fig. 4(a) by the saliency maps highlights both the glass and its content, whereas the gradients' visualization avoids any misleadingness that could have occurred due to the outdoors background surrounding the glass.
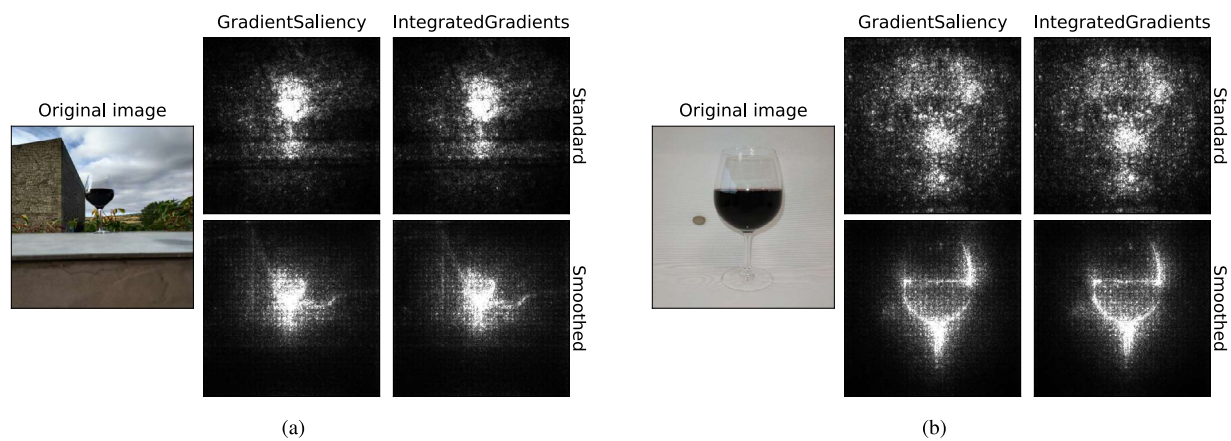
The explanation given by the saliency maps in Fig. 4(b) emphasizes the glass and the wine in it, while it does not pay attention to the coin. Therefore, our model is focusing on the container and its content in order to make the prediction, as it was expected.

#### 3.1.2. Regression metrics analysis on wine volume estimation results

Wine volume model's predictions were evaluated using the following regression metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Coefficient of Determination ($R^2$). MAE and RMSE results for training, validation and test sets are listed in Table 3. The coefficient of determination was calculated with test images, giving a value of $R^2 = 0.97$.

Fig. 5 shows violin plots for the predictions of each wine volume class available in the dataset. As it can be inferred from the shape of the violin plots distribution, predicted volumes are in general highly concentrated around the median.

GradientSaliency    IntegratedGradients

Original image                                    Standard

                                                  Smoothed

(a)

GradientSaliency    IntegratedGradients

Original image                                    Standard

                                                  Smoothed

(b)

**Fig. 4.** Standard and smoothed saliency maps examples. (a) Saliency maps of Chardonnay wine glass filled with 150 mL. The predicted volume is 146.5 mL (2.3% relative error). (b) Saliency maps of Bordeaux wine glass filled with 275 mL. The estimated volume is 259.6 mL (5.6% relative error).
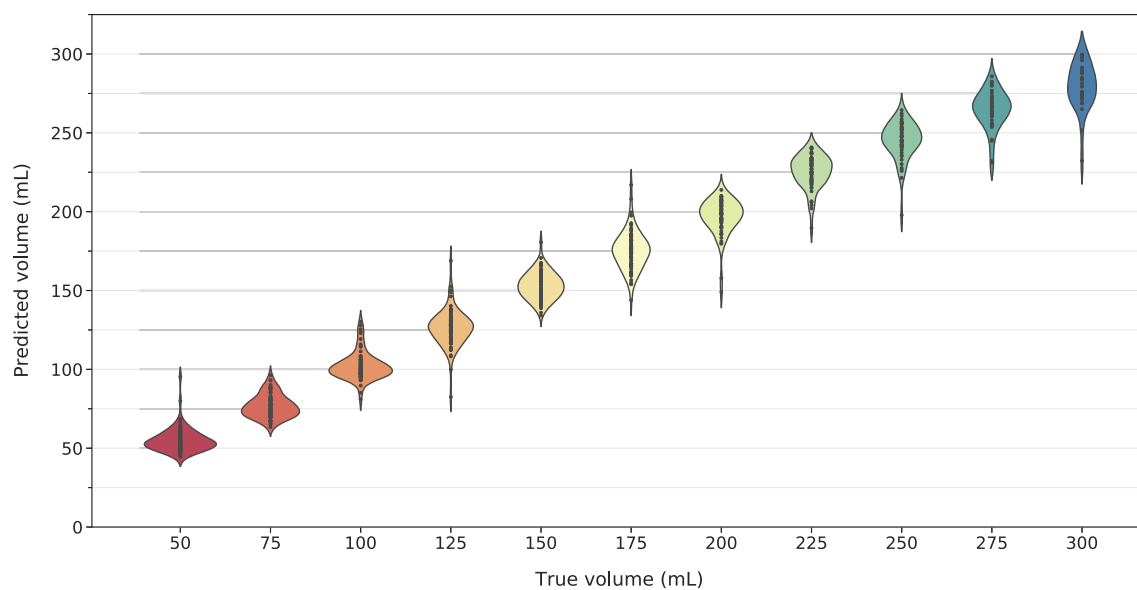


**Fig. 5.** Violin plots of the volumes distribution for model's predictions.

The amount of volume in the glass reduces the precision of the prediction when wine content is over 250 mL, usually predicting lower volumes than expected. One reason for this behaviour can be attributed to the lower number of images corresponding to higher volumes compared with the others, as listed in Table 2.

However, it should be noticed that, for real applications of the model, filling a glass of red wine with more than one third of its capacity is unusual, thus, volume content is hardly ever over 200 mL.

### 3.1.3. Effect of different image conditions on wine volume estimation

The above evaluations and saliency maps show an overview of our model's performance. In Appendix we present Fig. 6, which depicts various graphs representing the Mean Absolute Error under different conditions (type of glass, volume content, background, object distance, angle, light and reference).

The main results show that glass type, angle, object distance, lightning state and reference objects in the image have no significant influence in volume estimations, whereas background sometimes has a negative impact on the precision of the predictions. In fact, when the glass of wine is placed in an environment that includes other objects (i.e. outside background in the dataset), model's performance is more likely to drop because of the presence of these items, that sometimes confuse the CNN. Moreover, WineGut_BrainUP dataset is unbalanced

when taking into account the number of images for the different backgrounds that are available: white (10896), blue (10827) and outside (2582), where a significant reduced number of outdoors pictures were included. Thus, the performance of the model slightly decreases for images with outside backgrounds.

### 3.2. Comparison of the proposed model with existing methods

The precision of our model has been compared with the evaluation metrics of the state-of-the-art systems in liquid volume estimation found in the literature. As already mentioned in section 1, (Mottaghi et al., 2017) CNN model is the most comparable existing method with our model, in terms of operation and measurement conditions. The other referenced systems either incorporated a more sophisticated implementation that required more than one photograph in order to cover different views of the object to be measured, such as (Siswantoro et al., 2014), or were not developed specifically for liquids, like (Chen et al., 2013).

As a result, we compare the performance of our model with the suggested method by (Mottaghi et al., 2017) that followed a classification approach, reaching a 32% per class accuracy. We cannot directly compare this evaluation metric with our previous analysis, as we followed a regression approach, but it is straightforward to deduce that

our reported 8 mL MAE is lower than the cited precision, as it will be shown in the following lines. Briefly explained, (Mottaghi et al., 2017) divided the space of volumes into 10 classes, where the maximum volumes in each class were 50, 100, 200, 300, 500, 750, 1000, 2000, 3000, ∞ (unit of measurement: mL). Consequently, their method cannot exceed in performance the minimum space in mL between the volume classes with which it was trained, therefore, our approach outperforms the mentioned system.

### 3.3. Model limitations

The range of application of our model would be limited to the use of glass containers in the photos that are similar in shape and size to the ones with which our model was trained. However, the WineGut_BrainUP dataset covers all kinds of glass containers that are representative in the wine world, so this requirement is not likely to have an impact in real life applications. Indeed, our method predicts red wine volume from a photograph without requiring any information about the type of glass, nor its measurements. The reason for this behaviour is that our model was trained with all representative glasses in the wine world using photographs that were taken under different conditions, which cover a wide range of perspectives in terms of angles and object distances. Therefore, our model has learnt to recognize the outline of the glass container regardless of the possible small variations in size, height or shape that could occur between different manufacturers.

Additionally, our model provides a volume estimation for an image that contains a single glass of red wine. One challenging problem is automatic estimation of liquid volumes for all kinds of liquids and containers, or even in the case when there is more than one glass container in the image. To handle this issue, a larger dataset including enough samples of these elements would be required for estimating liquid volume.

Furthermore, from the proposed model (Cobo, 2021), the incorporation of new liquids is straightforward, the only requirement is to include in a new training phase of the model photographs of glasses containing the desired liquid. However, as the model has already been trained to recognize the glasses' outline, the number of images necessary for this new development would not be has high as in the WineGut_BrainUP dataset. In fact, in the testing phase of the model, in an independent test, we successfully estimated the liquid volume of a glass filled with white wine, although this was not the target wine in our method.

## 4. Conclusion

We have presented a new CNN based regression model to determine red wine volume in a glass container from single-view RGB images. This method does not require any reference object in the image, outperforming similar systems that were developed in the literature for related tasks. The proposed model efficiently estimates wine volume in almost any kind of wine glass container showing that in order to solve the liquid volume estimation challenge it is not needed to include a calibration object in the image. In contrast, this presumed deficiency can be overcome with a larger training dataset including enough photographs of all representative situations that could occur, so that the system is able to recognize the shape and size of the glass recipient containing the liquid. We introduced WineGut_BrainUP dataset to train and evaluate our system, which has potential real life applications in diet monitoring and wine consumption studies. In the future, we plan to incorporate volunteers' photographs in a subsequent study so that we can generalize our model with new real world backgrounds and setups, solving a long-standing problem in nutrition science, where FFQs-based dietary assessment often results subjective and time-consuming. This study aims to provide an automated tool for red wine volume estimation based on the proposed model, in a simple and efficient way that only requires the subject to take a photograph of the glass of wine with

a mobile phone, instead of having to carry a beaker or any other instrument to perform the measurement. Overall, this modelling will facilitate accurate measurement of liquid volume in diet and consumption studies.

## Declarations

### Author contribution statement

**Miriam Cobo:** Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper. **Ignacio Heredia, Fernando Aguilar:** Contributed reagents, materials, analysis tools or data; Wrote the paper. **Lara Lloret Iglesias, Begoña Bartolomé, M. Victoria Moreno-Arribas, Patricia Pérez-Matute, Maria-José Motilva:** Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper. **Daniel García, Silvia Yuste:** Contributed reagents, materials, analysis tools or data.

### Funding statement

### Data availability statement

Data associated with this study has been deposited at DIGITAL.CSIC under https://digital.csic.es/handle/10261/256232. URI: http://hdl.handle.net/10261/256232. DOI: https://doi.org/10.20350/digitalCSIC/14135.

### Declaration of interests statement

The authors declare no conflict of interest.

### Additional information

No additional information is available for this paper.

## Appendix A. Model results statistics for different image conditions

Fig. 6 shows mean absolute error bar plots for the estimated wine volume predictions in test set, broken down by the diverse image conditions available in the WineGut_BrainUP dataset.

## References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., 2015. TensorFlow: large-scale machine learning on heterogeneous systems. Software available from https://www.tensorflow.org/.

Anh, H.N., 2018. Deep-viz-keras GitHub software by experiencor. https://github.com/experiencor/deep-viz-keras.

Bartolomé, B., Moreno-Arribas, M.V., Lloret, L., Aguilar, F., Cobo, M., García, D., Heredia, I., Yuste, S., Pérez-Matute, P., Motilva, M.J., 2021. BrainGut_WineUp laboratory images [Dataset].

Chen, H.C., Jia, W., Yue, Y., Li, Z., Sun, Y.N., Fernstrom, J.D., Sun, M., 2013. Model-based measurement of food portion size for image-based dietary assessment using 3d/2d registration. Meas. Sci. Technol. 24, 105701.
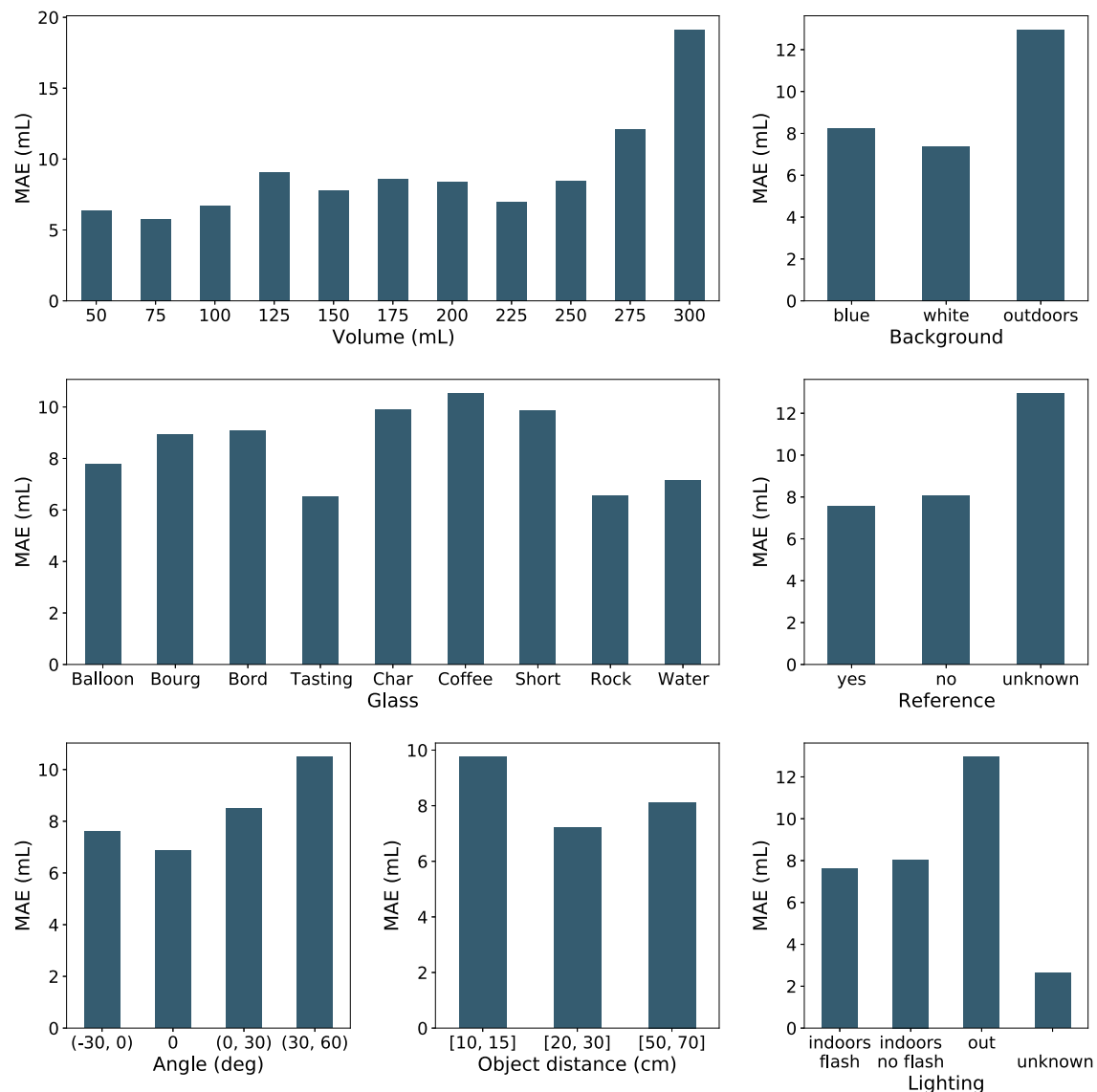
**Fig. 6.** Mean Absolute Error for estimated red wine volumes with our model.

Chollet, F., et al., 2015. Keras. https://keras.io.

Chollet, F., 2017. Xception: deep learning with depthwise separable convolutions. arXiv: 1610.02357.

Cobo, M., 2021. Liquid volume estimation from RGB images. WineGut_BrainUP Project Software. https://digital.csic.es/handle/10261/254875.

De Rijk, M.G., Slotegraaf, A.I., Brouwer-Brolsma, E.M., Perenboom, C.W., Feskens, E.J., De Vries, J.H., 2021. Development and evaluation of a diet quality screener to assess adherence to the Dutch food-based dietary guidelines. Br. J. Nutr., 1–11.

Fang, S., Liu, C., Zhu, F., Delp, E.J., Boushey, C.J., 2015. Single-view food portion estimation based on geometric models. In: 2015 IEEE International Symposium on Multimedia (ISM). IEEE, pp. 385–390.

Fang, S., Zhu, F., Jiang, C., Zhang, S., Boushey, C.J., Delp, E.J., 2016. A comparison of food portion size estimation using geometric models and depth images. In: 2016 IEEE International Conference on Image Processing (ICIP). IEEE, pp. 26–30.

García, Á.L., De Lucas, J.M., Antonacci, M., Zu Castell, W., David, M., Hardt, M., Iglesias, L.L., Moltó, G., Plociennik, M., Tran, V., et al., 2020. A cloud-based framework for machine learning workloads and applications. IEEE Access 8, 18681–18692.

González-Alzaga, B., Hernández, A.F., Pack, L.K., Iavicoli, I., Tolonen, H., Santonen, T., Vinceti, M., Filippini, T., Moshammer, H., Probst-Hensch, N., et al., 2022. The questionnaire design process in the European human biomonitoring initiative (hbm4eu). Environ. Int. 160, 107071.

International Organization for Standardization, 1977. Wine-tasting glass (ISO 3591:1977). https://www.iso.org/standard/9002.html.

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., 2019. Automatic food detection in egocentric images using artificial intelligence technology. Public Health Nutr. 22, 1168–1179.

Kingma, D.P., Ba, J., 2017. Adam: a method for stochastic optimization. arXiv:1412.6980.

Loshchilov, I., Hutter, F., 2019. Decoupled weight decay regularization. arXiv:1711.05101.

Mottaghi, R., Schenck, C., Fox, D., Farhadi, A., 2017. See the glass half full: reasoning about liquid containers, their volume and content. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1871–1880.

Pechey, R., Couturier, D.L., Hollands, G.J., Mantzari, E., Munafò, M.R., Marteau, T.M., 2016. Does wine glass size influence sales for on-site consumption? A multiple treatment reversal design. BMC Public Health 16, 1–6.

Rosebrock, A., 2019. Keras, regression, and CNNs. https://pyimagesearch.com/2019/01/28/keras-regression-and-cnns/.

Simonyan, K., Vedaldi, A., Zisserman, A., 2013. Deep inside convolutional networks: visualising image classification models and saliency maps. Preprint. arXiv:1312.6034.

Siswantoro, J., Prabuwono, A.S., Abdullah, A., 2014. Volume measurement algorithm for food product with irregular shape using computer vision based on Monte Carlo method. J. ICT Res. Appl. 8.

Smilkov, D., Thorat, N., Kim, B., Viégas, F., Wattenberg, M., 2017. Smoothgrad: removing noise by adding noise. Preprint. arXiv:1706.03825.

Sotos-Prieto, M., Santos-Beneit, G., Bodega, P., Pocock, S., Mattei, J., Peñalvo, J.L., 2015. Validation of a questionnaire to measure overall Mediterranean lifestyle habits for research application: the Mediterranean lifestyle index (medlife). Nutr. Hosp. 32, 1153–1163.

Sundararajan, M., Taly, A., Yan, Q., 2017. Axiomatic attribution for deep networks. In: International Conference on Machine Learning, PMLR, pp. 3319–3328.

Yang, Z., Yu, H., Cao, S., Xu, Q., Yuan, D., Zhang, H., Jia, W., Mao, Z.H., Sun, M., 2021. Human-mimetic estimation of food volume from a single-view rgb image using an ai system. Electronics 10, 1556.