

Un nuevo método basado en inteligencia artificial para evaluar la ingesta individual de vino

Novel approach based on artificial intelligence to evaluate individual wine intake

Miriam Cobo¹, Edgard Relaño de la Guía², Ignacio Heredia¹, Fernando Aguilar¹, Lara Lloret-Iglesias¹, Daniel García¹, Silvia Yuste^{3,4}, Emma Recio-Fernández⁵, Patricia Pérez-Matute⁵, M. José Motilva³, M. Victoria Moreno-Arribas² y Begoña Bartolomé²

¹Institute of Physics of Cantabria (IFCA), CSIC - UC, 39005 Santander (Cantabria), Spain

²Institute of Food Science Research (CIAL), CSIC-UAM, 28049 Madrid, Spain

³Institute of Grapevine and Wine Sciences (ICVV), CSIC-University of La Rioja-Government of La Rioja, 26007 Logroño (La Rioja), Spain

⁴Grupo de investigación Antioxidantes, Departamento de Tecnología de Alimentos, Centro de Investigación Agrotecnio, Universidad de Lleida, Av. Alcalde Rovira Roure, 191, 25198 Lleida, Spain

⁵Infectious Diseases, Microbiota and Metabolism Unit, Center for Biomedical Research of La Rioja (CIBIR), CSIC Associated Unit. 26006 Logroño (La Rioja), Spain

Resumen. Este estudio surge de la necesidad de nuevas metodologías que permitan cuantificar el consumo de vino con mayor precisión, para posteriormente utilizar esta información en estudios observacionales de alimentación-salud y estudios de intervención de dieta. Se ha desarrollado un algoritmo basado en un método de “aprendizaje profundo”, que permite determinar el volumen de vino en una copa/vaso a partir de una fotografía, y se ha validado en un estudio de consumidores realizado a través de una aplicación web. La aplicación del modelo a imágenes “cuasi-reales” y a imágenes “reales” (obtenidas a partir del estudio de consumidores), ha mostrado una precisión satisfactoria con un error absoluto medio (MAE) de 10 mL y 26 mL, respectivamente. En relación a las pautas de consumo de vino observadas en el estudio de consumidores (n=38), el volumen medio de vino tinto servido en una copa fue de 114±33 mL, sin estar condicionado por factores como el sexo del consumidor, el momento de consumo, el tipo de vino, o el formato de copa/vaso. En síntesis, el sistema de aprendizaje profundo desarrollado junto con la aplicación web, constituyen una herramienta de gran valor para la estimación precisa del volumen de vino consumido diariamente, así como las pautas de su consumo, de gran utilidad para estudios poblacionales.

Abstract. This study arises from the need to propose new methodologies to quantify wine consumption more precisely in order to use subsequently this information in observational food-health studies and dietary intervention studies. It has been developed an algorithm based on a “deep learning” method to determine wine volume from a single-view image, and it has been validated through a consumer study developed via a web application. The new model demonstrated satisfactory performance not only in a “daily lifelike” images dataset but also in “real” images (obtained from the consumer study), with a mean absolute error (MAE) of 10 and 26 mL, respectively. In relation to the data reported by the participants in the consumer study (n=38), average red wine volume in a glass was 114±33 mL, without being affected by factors such as gender, time of consumption, type of wine or type of glass. Therefore, the deep learning system together with the web application developed in this study constitute a diet monitoring tool of substantial value in the accurate assessment of daily wine intake, as well as in the habits of its consumption, with relevant applications in observational studies.

1 Introducción

La valoración precisa de los patrones de dieta, además de la estimación de ingesta diaria de macro- y micronutrientes es crucial para investigadores e instituciones de salud pública en los estudios destinados a mejorar la salud de la población [1, 2]. Estas estimaciones se basan principalmente en cuestionarios de frecuencia alimentaria (*Food Frequency Questionnaires*, FFQ) que recogen la frecuencia con la que consumen diferentes alimentos de una lista predefinida [3]. Para calcular los gramos consumidos diariamente de cada alimento, los datos de frecuencia se multiplican por el tamaño de la porción. A partir de esta información, los gramos diarios de alimento se transforman en ingesta diaria de nutrientes (macro- y micro-) utilizando las bases de datos de composición de alimentos. Sin embargo, los FFQ tienen dos limitaciones importantes, por un lado se basan en la memoria del sujeto, y además consideran un peso/volumen de ración estándar para cada alimento/bebida, dando lugar a errores de estimación. Especialmente sensible es la estimación del volumen de la ingesta diaria de vino a través de los FFQ, ya que una ración (copa de vino) consumida se estandariza a 100 cc (100 mL), lo que no siempre es real. En este sentido, el vino se consume en diferentes formatos de copa/vaso lo que condiciona que este volumen (100 cc) pueda ser inferior o superior al estimado a través de los cuestionarios [4].

En este contexto, la aplicación de la inteligencia artificial (IA) para generar algoritmos que permita estimar el peso/volumen de un alimento/bebida a partir de una imagen abre nuevas perspectivas para mejorar y estandarizar la información de dieta en estudios nutricionales [5].

Como una nueva aproximación de la IA para resolver el reto de la estimación precisa de la ingesta de vino, en este estudio se ha desarrollado un modelo de aprendizaje profundo (*deep learning*, DL) que permite estimar el volumen de vino servido en una copa a partir de una única fotografía. Para el desarrollo del modelo se partió de una red neuronal convolucional de regresión, que en una primera fase se entrenó con fotografías de copas/vasos de vino obtenidas por los investigadores en el laboratorio (*BrainGut_WineUp laboratory images* dataset [6]). Este primer modelo [7] se optimizó en una segunda fase con fotografías de copas de vino tomadas por los investigadores en situaciones reales de consumo (*BrainGut_WineUp daily lifelike images* dataset [8]) para desarrollar un modelo optimizado [9]. Finalmente, el modelo DL optimizado se ha aplicado a un conjunto de imágenes (*BrainGut_WineUp real images* data set [10]) tomadas por los participantes en un estudio de consumidores *ad hoc*. En conjunto, este estudio muestra, por primera vez, que es posible estimar el volumen de vino tinto contenido en una copa a partir de una fotografía mediante la aplicación de métodos de inteligencia artificial.

2 Materiales y métodos

2.1 Bancos de imágenes

Para desarrollar y evaluar el modelo DL, se utilizaron tres conjuntos de datos (*image dataset*) independientes que comprendían fotografías de diferentes formatos de copas que contenían vino tinto: *i) BrainGut_WineUp laboratory images* dataset [6], con imágenes obtenidas en el laboratorio, *ii) BrainGut_WineUp daily lifelike images* dataset [8], con imágenes obtenidas por los investigadores en situaciones reales, e *iii) BrainGut_WineUp real images* dataset [10], con imágenes tomadas por consumidores de vino en su vida cotidiana.

2.1.1 Banco de imágenes *BrainGut_WineUp laboratory*

Este banco de imágenes *BrainGut_WineUp laboratory* [6] incluía 24.305 fotografías de laboratorio de copas/vasos conteniendo diferentes volúmenes medidos (50-300 mL) de vino tinto que fueron tomadas en los laboratorios del Instituto de Investigación en Ciencias de la Alimentación (CIAL-CSIC), el Instituto de Ciencias de la Vid y el Vino (ICVV-CSIC) y el Centro de Investigación Biomédica de La Rioja (CIBIR). Para ello, se seleccionaron tres vinos tintos comerciales representativos de los vinos "joven", "crianza" y "reserva", y para cada vino se siguió el mismo diagrama de flujo, representado en la Fig. 1.

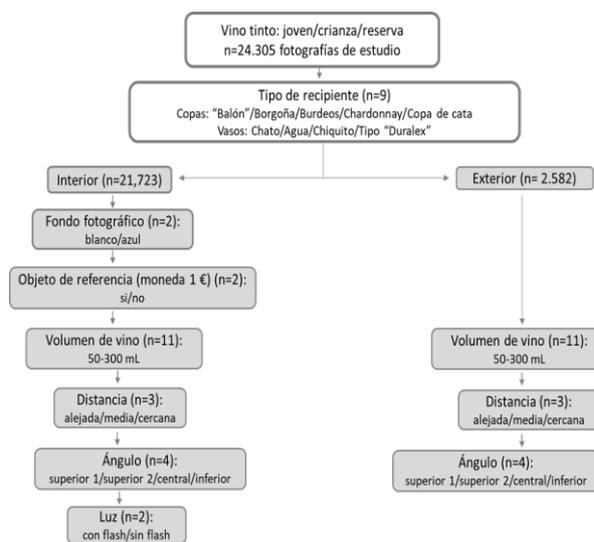


Figura 1. Diagrama de flujo seguido para la creación del banco de imágenes *WineGut_BrainUp laboratory*.

Las fotografías se tomaron en interiores y exteriores, teniendo en cuenta las siguientes variables:

-Tipo de copa (n = 5): “balón”, Borgoña, Burdeos, de cata, Chardonnay; Tipo de vaso (n = 4): chato, de agua, chiquito y “Duralex”. En la Figura 2 se muestra un ejemplo de cada una de las copas/vasos utilizados.

-Volumen de vino (n = 11): 50, 75, 100, 125, 150, 175, 200, 225, 250, 275 y 300 mL. Las mediciones se realizaron con una probeta con una precisión de $\pm 0,5$ mL.

-Distancia del objeto (n = 3): alejada (50-70 cm), media (20-30 cm), cercana (10-15 cm).

-Ángulo (n = 4): superior 1 (0-30°), superior 2 (30- 60°), central (0°), bajo (-30-0°).

Además, se consideraron las siguientes variables para las fotografías en interiores:

-Fondo fotográfico (n = 2): blanco, azul.

-Referencia (n = 2): sí, no. Se tomó una moneda de 1 euro como posible referencia (objeto de calibración).

-Iluminación (n = 2): con flash, sin flash.



Figura 2. Tipos de copas/vasos empleados en este estudio.

2.1.2 Banco de imágenes *BrainGut_WineUp daily lifelike*

Este banco *BrainGut_WineUp daily lifelike images* [8] se diseñó y creó específicamente para optimizar el primer modelo DL [7]. El conjunto de datos incluía 1.945 fotografías tomadas por miembros de los grupos de investigación en situaciones reales de consumo. Se emplearon los mismos vinos tintos utilizados en la creación del banco de imágenes *BrainGut_WineUp laboratory*, y se siguió un diagrama de flujo similar. Las fotografías se tomaron en interior y exterior, teniendo en cuenta los siguientes campos:

-Tipo de copa (n = 5): “balón”, Borgoña, Burdeos, de cata, Chardonnay; Tipo de vaso (n = 4): chato, de agua, chiquito y “Duralex” (Fig. 2).

-Volumen de vino (n=7): 50, 75, 100, 125, 150, 175, 200 mL. Las mediciones se realizaron utilizando una probeta con una precisión de $\pm 0,5$ mL.

-Ángulo (n=2): superior (0-30°), central (0°).

2.1.2 Banco de imágenes *BrainGut_WineUp real*

Este banco *BrainGut_WineUp real images* [10] contiene 229 fotografías reales que fueron tomadas por 38 voluntarios que participaron en el estudio de consumidores (ver 2.3). Este conjunto de datos se utilizó para testar, con una aproximación realista, el modelo DL desarrollado. Antes de enviar la fotografía, los

participantes tenían que etiquetarla con los siguientes campos:

-Tipo de vino (n=3): "joven", "crianza" y "reserva".

-Tipo de copa/vaso (n=10): “balón”, Borgoña, Burdeos, de cata, Chardonnay, chato, de agua, chiquito, “Duralex” (Fig. 2) y otros.

-Intervalos para la asignación del volumen de vino medido (n=11): 50, 75, 100, 125, 150, 175, 200, 225, 250, 275 y 300 mL. Las mediciones se realizaron utilizando un vaso de precipitados con una precisión de ± 25 mL.

2.2 Entrenamiento y validación del modelo

El procedimiento para estimar el volumen de vino en una copa/vaso consistió primero en entrenar una red neuronal convolucional (también conocidas como CNN o Conv-Nets) con imágenes de laboratorio. Las redes neuronales convolucionales son un tipo de redes neuronales de aprendizaje profundo especialmente diseñadas para analizar datos de imágenes, tanto con etiquetas numéricas como categóricas. Del total de 24.305 fotografías disponibles, el 80% se utilizaron para el entrenamiento (19.079 fotografías), el 10% para la validación (2.613 fotografías) y el 10% para el *test* de precisión (2.613 fotografías).

Se utilizó una red neuronal Xception [11] con imágenes de tamaño 528×528 píxeles. Las imágenes del conjunto de datos se redimensionaron para cumplir este requisito. El tamaño del batch se fijó en 16, el número de épocas de entrenamiento se fijó en 50 y se empleó el optimizador Adam [7]. Primero, se utilizó un modelo base preentrenado en ImageNet. A continuación, se incluyeron capas personalizadas para adaptar el modelo a la tarea de estimación de volumen mediante el uso de una función de activación lineal en la última capa. Como el número de imágenes disponibles era bastante elevado, se desactivó el aumento de datos del modelo (data augmentation) para reducir el tiempo de cálculo.

Este modelo se entrenó con una GPU Tesla V100-PCIE-32GB durante 20 horas. El modelo se codificó utilizando Keras [11] y Ten-sorFlow versión 1.14.0 [12] en Ubuntu 18.04.2 LTS. mL.

Este primer modelo [7] se optimizó posteriormente con imágenes de copas de vino tomadas en situaciones reales, siguiendo un enfoque de transferencia del aprendizaje [13]. Para ello, se utilizó el 96% de las imágenes para el entrenamiento (1.867 fotografías), el 1,4% para la validación (28 fotografías) y el 2,6% para el *test* de precisión (50 fotografías). Se empleó la mayoría de las imágenes para el entrenamiento porque las pruebas posteriores se realizaron con fotografías reales de consumidores (*BrainGut_WineUp real images* dataset [10]). Además, se utilizó “early stopping” para evitar el sobreajuste. Así, nuestro modelo se entrenó hasta que la pérdida de validación (“validation loss”) no mejoraba durante 10 épocas consecutivas.

La arquitectura DL utilizada en este estudio de seguimiento para medir con precisión el volumen del vino se ilustra en la Figura 3. En resumen, se utilizó el modelo anterior preentrenado (Xception) [7], que

consistía en varias capas convolucionales, conectadas posteriormente con las últimas capas totalmente conectadas tras realizar una agrupación de promedios globales (Global Average Pooling). Los pesos de las últimas capas se ajustaron con las fotografías de entrenamiento del conjunto de datos realistas. Este modelo optimizado se entrenó también con una GPU Tesla V100-PCIE-32GB utilizando Keras [11] y TensorFlow versión 1.14.0 [12] en Ubuntu 18.04.2 LTS.

El modelo optimizado [9] se evaluó en dos pruebas independientes: el subconjunto de test de imágenes reales diarias (*BrainGut_WineUp daily lifelike*, n=50) y el conjunto de datos de imágenes reales (*BrainGut_WineUp real* n=229). Este último conjunto de datos sirvió para realizar una validación insesgada para comprobar si el modelo era capaz de generalizar a nuevos escenarios no óptimos con diferentes cámaras fotográficas y estados de iluminación.

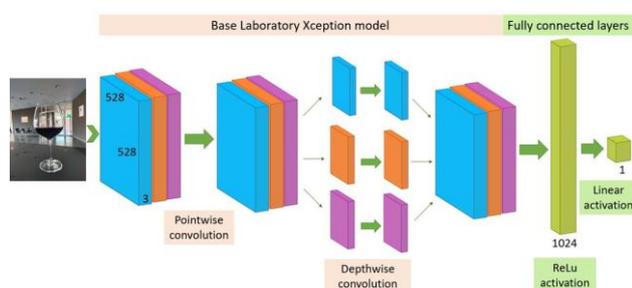


Figura 3. Arquitectura CNN propuesta para la estimación del volumen del vino. Los números indican el tamaño de las imágenes de entrada y el tamaño de las capas finales que dan lugar a la predicción.

2.3 Estudio de consumidores

Se diseñó y creó específicamente una aplicación web con el nombre de ALIMENTA365 (<https://alimenta365.csic.es/>) para recabar información de los participantes en el estudio sobre sus hábitos de consumo de vino tinto, así como para recoger fotografías reales de la ingesta de vino. Cuando un participante en el estudio accedía por primera vez a la aplicación web, se le pedía que se registrara y rellenara un cuestionario sobre sus hábitos de consumo de vino. Una vez completado el cuestionario, el participante recibía un correo electrónico con una contraseña para acceder a la plataforma web y subir sus fotografías. Los participantes también recibieron por correo un vaso medidor de 300 mL de volumen con marcas graduadas a intervalos de 25 mL para medir el volumen de vino tinto servido en la copa/vaso antes de su ingesta. La aplicación web incluía un tutorial en vídeo para tomar y subir las fotografías, en el que se indicaba la distancia y el ángulo óptimos desde la cámara a la copa. En el tutorial también se pedía a los participantes que evitaran incluir en sus fotografías a personas, así como otros recipientes (como vasos, copas o botellas) distintos de su copa de vino.

Antes de subir la fotografía, los participantes marcaron la opción correcta correspondiente al tipo de vino, el volumen medido, el tipo de copa utilizada y el momento de consumo. Para este estudio de consumo de vino tinto, los participantes se reclutaron entre la población general tras anunciarse en redes sociales a través de la Unidad de Comunicación del Instituto de Investigación en Ciencias de la Alimentación (CIAL, CSIC, Madrid, España). Para participar en el estudio, los voluntarios debían ser mayores de 18 años y consumidores de vino, tanto esporádicos como habituales. Se pidió a los participantes que subieran fotografías de todas las raciones de vino durante 3 semanas, a partir del día en que subieron su primera fotografía. El estudio fue aprobado por el Comité Ético del CSIC (Madrid, España) (Número de aprobación 081/2021) y garantizó la protección de los datos de los participantes.

Un total de 38 voluntarios (20 hombres y 18 mujeres) participaron en el estudio.

2.4 Análisis estadístico

Para evaluar el rendimiento del nuevo modelo DL optimizado, se evaluaron las predicciones del volumen de vino con las siguientes métricas de regresión: Error Medio Absoluto (MAE) y Error Medio Cuadrático (RMSE). Este análisis estadístico se llevó a cabo utilizando la biblioteca scikit-learn [14] en Python 3.8.6.

El volumen de vino por servicio registrado por los participantes en el estudio de consumidores (datos reales) se comparó con los valores estimados por el nuevo modelo DL mediante una prueba t pareada. Las diferencias en el volumen de vino por servicio entre individuos, momento de consumo, tipo de vino y tipo de copa se analizaron mediante la prueba ANOVA. Las diferencias entre sexos y entre dos tipos de contenedores (copa de vino “balón” y vaso chiquito) se analizaron mediante la prueba t. Estos análisis estadísticos se realizaron con el programa Excel de Microsoft Office plus 2019.

3 Resultados y Discusión

2.1 Desarrollo del modelo DL

La Tabla 1 muestra los valores de las métricas de regresión MAE (error absoluto medio) y RMSE (error cuadrático medio) para todos los conjuntos de imágenes considerados en este estudio: entrenamiento, validación y prueba procedentes de las imágenes de laboratorio (*BrainGut_WineUp laboratory images* dataset [6]), “cuasi-reales” (*BrainGut_WineUp daily lifelike images* dataset [8]) y reales (*BrainGut_WineUp real images* dataset [10]).

Tabla 1. Evaluación de las métricas de rendimiento (MAE y RMSE) para las predicciones del volumen de vino con el primer modelo y el modelo optimizado.

	Images (n)	MAE (mL)	RMSE (mL)
Imágenes de laboratorio (primer modelo)			
Entrenamiento	19079	2	3
Validación	2613	8	12
Prueba	2613	8	11
Imágenes cuasi-reales (modelo optimizado)			
Entrenamiento	1867	5	7
Validación	28	14	20
Prueba	50	10	13
Imágenes reales (aplicación del modelo optimizado)			
Prueba	229	26	33

El primer modelo DL alcanzó un valor de MAE de 8 mL y se validó con éxito en diferentes condiciones de imagen tomadas a escala de laboratorio (Tabla 1). Los valores de MAE y RMSE del sistema DL optimizado en el subconjunto de prueba (10 y 13 mL, respectivamente) también se consideraron aceptables, lo que indicaba que el modelo se había optimizado con éxito con las imágenes reales diarias. La Fig. 3 muestra gráficos de violín para las predicciones de cada uno de los valores de volumen de vino para las imágenes de laboratorio y “cuasi-reales”, respectivamente.

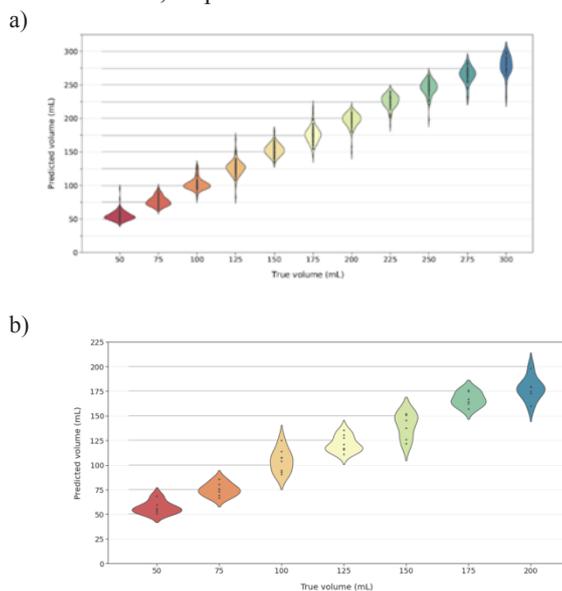


Figura 3. Gráficos de violín de la distribución de volúmenes de vino predichos por el primer modelo DL (a) y por el modelo optimizado (b).

A modo de ejemplo, la Fig. 4 recoge los mapas de saliencia y los gradientes integrados, tanto en su versión estándar como suavizada, para dos imágenes de prueba correspondientes al primer modelo DL (Fig. 4a) y al modelo optimizado (Fig. 4b). Especialmente para este último caso, se observa que el modelo omite el fondo y los objetos que están en la fotografía, integrando exclusivamente los puntos que afectan al volumen de vino contenido en la copa. Incluso, aunque el recipiente de cristal estaba parcialmente cubierto por un dedo, el

modelo DL realizó una estimación adecuada, con un error relativo del 11% (Fig. 4b).

Los valores de MAE y RMSE del modelo DL optimizado en el conjunto de datos de imágenes reales (imágenes de los consumidores de vino participantes en el estudio) fueron ligeramente superiores (26 y 33 mL, respectivamente) que en el conjunto de datos de imágenes de laboratorio e imágenes “cuasi-reales” (Tabla 1).

Como ejemplo de la gran variabilidad en la precisión de la estimación del volumen de vino entre las imágenes de los consumidores (conjunto de datos de imágenes reales), la Fig. 5 muestra varios mapas de saliencia de fotografías con diversos porcentajes de error relativo, que se atribuyeron a diferentes circunstancias. La Fig. 5a muestra los mapas de saliencia de la imagen de un participante en la que el modelo realizó una estimación muy precisa (125 mL medidos frente a 125 mL estimados). En este caso, el fondo oscuro detrás de la copa de vino tenía un impacto menor en la predicción. Sin embargo, en las imágenes de otros participantes, el fondo parecía tener un mayor impacto en la predicción, especialmente cuando los participantes incluían otros objetos de cristal, a pesar de que se les había aconsejado específicamente que los evitaran en sus imágenes, ya que inducían un mayor error relativo en la estimación (Fig. 5b, error relativo 25%, 125 mL medidos frente a 100 mL estimados).

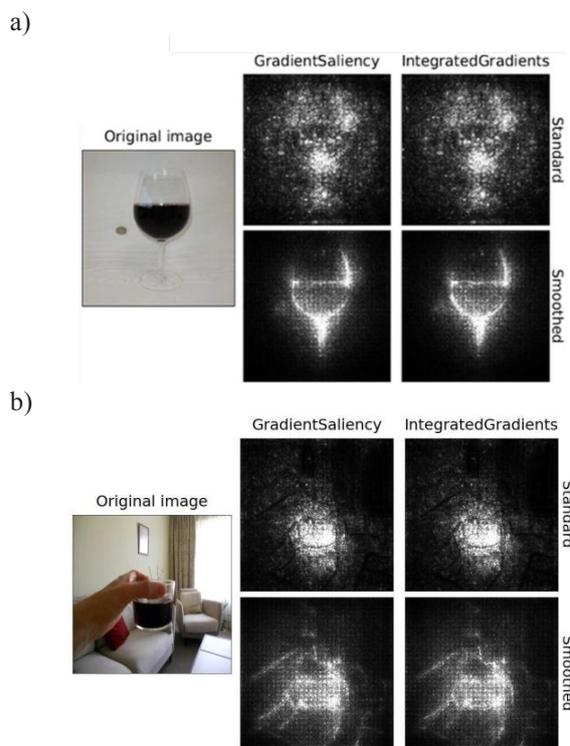


Figura 4. Mapas de características de dos imágenes de prueba: a) Copa Burdeos con 275 mL de vino (Volumen estimado según el primer modelo DL: 259,6 mL, 5,6% error relativo), b) Vaso chiquito con 125 mL de vino (Volumen estimado según el modelo DL optimizado: 111.7 mL, (11% error relativo).

El error relativo más elevado observado en las imágenes reales era razonablemente previsible, ya que las

fotografías fueron tomadas por los consumidores en situaciones cotidianas sin prestar especial atención a las indicaciones que se les daban en la aplicación web para tomar las fotografías. Además, los errores más comunes detectados estaban relacionados con la presencia de objetos adicionales, la distancia/ángulo a la que se tomaron las fotografías o la utilización de vino rosado en lugar de tinto, aunque el error detectado en este caso concreto fue menor que en el resto de situaciones. Otra causa plausible que podría explicar el alto porcentaje de fotos mal estimadas fue el hecho de que los participantes aproximaron el volumen medido en el vaso de precipitados a marcas específicas con intervalos de 25 mL (50, 75, 100, 125, 150, 175, 200, 225, 250, 275 y 300 mL) que estaban preestablecidas en la aplicación web. En cambio, para la generación de las imágenes “cuasi-reales”, los vasos se llenaron con volúmenes medidos con probeta (precisión de $\pm 0,5$ mL) a intervalos de 25 mL, lo que dio lugar a una mayor precisión en la estimación (Tabla 1). De hecho, este es un aspecto del presente trabajo que debería mejorarse en futuros estudios en los que se pidiera a los participantes que anotaran el volumen medido exacto sin utilizar marcas preestablecidas. Por tanto, para futuros estudios y como propuesta de mejora, se plantea incluir instrucciones más detalladas para el ajuste de la fotografía y dejar que los participantes fijen libremente el volumen; creemos que ambos cambios mejorarían la precisión de la estimación del volumen del vino mediante el modelo DL desarrollado.

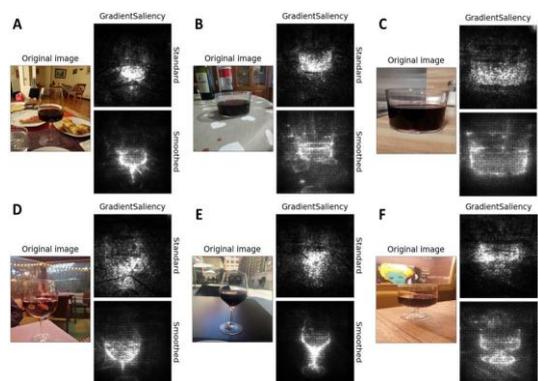


Figura 5. Mapas de características de dos imágenes reales tomadas por consumidores de vino participantes en el estudio: a) Copa que contiene 125 mL de vino tinto con un volumen estimado de 125 mL, b) Vaso que contiene 125 mL de vino tinto con un volumen estimado de 100 mL (error relativo del 25%).

3.2 Aplicación del modelo DL optimizado

Además de la evaluación del modelo DL optimizado, analizamos la información recogida sobre el consumo individual de vino tinto a partir del estudio de consumo (conjunto de datos de imágenes reales). El número de fotografías (copas de vino consumidas) subidas por cada participante durante el periodo de 3 semanas que duró el estudio, varió entre 17 y 1, con una media de 6 fotografías por participante (datos no mostrados). En

cuanto al momento del consumo de vino, se observó que el consumo individual de vino tuvo lugar mayoritariamente durante las comidas (almuerzo y cena), con un consumo mucho menor de vino fuera de las comidas. De los tres tipos de vino tinto consumidos, el vino "crianza" fue el más frecuentemente consumido. En cuanto al tipo de copa utilizado para el consumo, se observó una considerable diversidad, aunque el tipo Burdeos fue el más frecuente.

La Fig. 6 muestra la distribución del volumen en un servicio (copa/vaso) estimada por nuestro modelo DL a partir de las fotografías tomadas por los participantes en el estudio ($n=229$) en comparación con el volumen medido directamente por el participante utilizando el vaso graduado. No se encontraron diferencias significativas ($p = 0,253$) entre ambas variables aplicando la prueba t pareada.

El volumen medio en un servicio (copa/vaso) calculado a partir de las estimaciones de nuestro modelo DL fue de 116 ± 29 mL de vino tinto, cercano a la media del volumen medido por los participantes (114 ± 33 mL de vino tinto). Estos primeros resultados sugieren que la estandarización de 100 mL por ración de vino considerada en muchos registros dietéticos podría estar ligeramente infravalorada y, en algunos patrones dietéticos, podría llevar a conclusiones imprecisas.

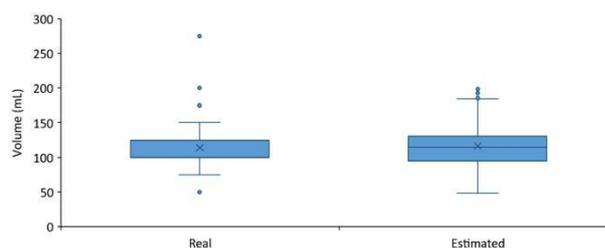


Figura 6. Box-plot del volumen de vino tinto en un servicio (copa/vaso) ($n=229$) comparando los valores reales, medidos directamente por el participante utilizando el vaso graduado, y los valores estimados por nuestro modelo DL a partir de las fotografías tomadas por los participantes.

El diseño de este estudio de consumidores también nos permitió evaluar la influencia de distintos factores en el volumen de vino tinto servido en una copa, como el individuo, el sexo, la hora de consumo, el tipo de vino y el tipo de copa. Como era de esperar, se observó una variabilidad significativa ($p = 0,000$, prueba ANOVA) entre individuos en el volumen de vino tinto en un servicio, de 50 a 275 mL. En relación con el sexo, no se encontraron diferencias significativas ($p = 0,881$, prueba t), aunque los participantes varones mostraron intervalos de variación más amplios. Tampoco se observaron diferencias estadísticamente significativas en relación con la hora de consumo ($p = 0,383$, prueba ANOVA) ni con el tipo de vino tinto ($p = 0,446$, prueba ANOVA). En general, no se observaron diferencias estadísticamente significativas para el tipo de copa ($p = 0,235$, prueba ANOVA). Sin embargo, cuando se utilizaron copas de mayor capacidad (es decir, copa tipo “balón”), se tendía a

servir mayores volúmenes de vino que cuando se utilizaron recipientes de menor capacidad (por ejemplo, vaso chiquito) ($p = 0,003$, prueba t) (Fig. 1).

En la práctica, el uso de un algoritmo de imagen para la mejora de la estimación del tamaño de ración en estudios de dieta, ya sea por parte de investigadores, dietistas y público en general, requiere su implementación en aplicaciones para teléfonos móviles. En este sentido, el presente modelo de DL será la base de una próxima APP que permitirá estimar el volumen de vino a partir de una fotografía utilizando el teléfono móvil.

3 Conclusiones

Uno de los primeros logros del presente estudio ha sido la creación de diferentes bancos de imágenes de copas/vasos de vino tinto, servido en diferentes volúmenes (*BrainGut_WineUp images*), disponibles en acceso abierto que podrían ser útiles en futuras aplicaciones y modelizaciones por parte de la comunidad científica que trabaja en el campo del vino.

No obstante, el reto más importante alcanzado en este trabajo ha sido el desarrollo de un nuevo modelo DL para la estimación del volumen de vino tinto en una copa, que se ha aplicado satisfactoriamente a una pequeña cohorte de adultos consumidores de vino tinto. Consideramos que el modelo DL desarrollado constituye una herramienta automática apropiada para medir el volumen de vino tinto, y resuelve las limitaciones de los cuestionarios convencionales de frecuencia de alimentos (i.e., FFQ) en los que la medición del líquido es subjetiva e imprecisa, ya que depende de la memoria del participante. Además, en combinación con una aplicación web, como la realizada en este estudio, el modelo DL desarrollado permite relacionar los datos de estimación del volumen de vino con otros hábitos de consumo de vino. En el futuro, estas tecnologías de imagen y computación podrían ampliarse a otros alimentos con diferentes aplicaciones en estudios epidemiológicos y nutricionales, así como estudios de consumidores.

Este trabajo ha sido financiado por el Ministerio de Ciencia e Innovación (MCIN) a través de la Agencia Estatal de Investigación (AEI)/10.13039/501100011033 y por la "Unión Europea Generation EU/PRTR": proyectos PID2019-108851RB-C21 y PID2019-108851RB-C22, y 'Prueba de Concepto' PDC2022-133861-C21 y PDC2022-133861-C22.

M. Cobo agradece el apoyo del Consejo Superior de Investigaciones Científicas (CSIC) y del Instituto de Física de Cantabria (IFCA) a través de la plataforma Interdisciplinar del CSIC (PTI+) Salud Global. I. Heredia ha sido financiado por el Gobierno de Cantabria con el proyecto "Instrumentación y Ciencia de Datos para sondear la naturaleza del Universo". Los autores agradecen también a la Plataforma Temática

Interdisciplinar (PTI+) Ciencia e Innovación Digital del CSIC. Silvia Yuste posee una ayuda postdoctoral Margarita Salas financiada por la Unión Europea – NextGenerationEU a través del Ministerio de Universidades y la Universitat de Lleida.

Referencias

1. M.G. De Rijk, A.I. Slotegraaf, E.M. Brouwer-Brolsma, C.W. Perenboom, E.J. Feskens, J.H. De Vries, Br. J. Nutr. **128**, 1615-1625 (2022)
2. B. González-Alzaga, A.F. Hernández, L.K. Pack, I. Iavicoli, H. Tolonen, T. Santonen, M. Vinceti, T. Filippini, H. Moshhammer, N. Probst-Hensch, M. Kolossa-Gehring, Environ. Inter. **160**, 107071 (2022)
3. M. Sotos-Prieto, G. Santos-Beneit, P. Bodega, S. Pocock, J. Mattei, J.L. Peñalvo, Nutricion Hospitalaria. **32**, 1153-1163 (2015)
4. R. Pechey, D.L. Couturier, G.J. Hollands, E. Mantzari, M.R. Munafò, T.M. Marteau, BMC Public Health **16**, 1-6 (2016)
5. K. Dalakleidi, M. Papadelli, I. Kapolos, K. Papadimitriou, Adv. Nutr. **13**, 2590–2619 (2022)
6. B. Bartolomé, M.V. Moreno-Arribas, L. Lloret, F. Aguilar, M. Cobo, M., D. García, I. Heredia, S. Yuste, P. Pérez-Matute, M.J. Motilva, *WineGut_BrainUp laboratory images* [Dataset], <https://digital.csic.es/handle/10261/256232> (2021)
7. 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, Heliyon **8**, e10557 (2022)
8. 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* [Dataset], <https://digital.csic.es/handle/10261/256232> (2022)
9. M. Cobo, E. Relaño de la Guía, I. Heredia, F. Aguilar, L. Lloret-Iglesias, D. García, S. Yuste, E. Recio-Fernández, P. Pérez-Matute, M. J. Motilva, M. V. Moreno-Arribas, B. Bartolomé (aún por publicar)
10. 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* [Dataset], <https://digital.csic.es/handle/10261/284780?mode=simple> (2022)
11. F. Chollet, arXiv **1610**, 02357 (2017)
12. M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro *et al.* arXiv **1603**, 04467 (2016)
13. F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, Q. He, Proceedings of the IEEE **109**, 43–76 (2021)
14. O. Kramer, Learn-Machine Learning for Evolution Strategies, 1st ed.; Springer Cham **20** (2016)