

Inteligencia artificial y visión por ordenador para evaluar los componentes del rendimiento de la vid en viñedos comerciales

Artificial intelligence and computer vision to assess grape yield components in commercial vineyards

Rubén Íñiguez^{1,2}, Carlos Poblete-Echeverría^{1,2}, Inés Hernández^{1,2}, Salvador Gutiérrez³, Ignacio Barrio^{1,2} y Javier Tardáguila^{1,2}

¹Televitis Research Group, University of La Rioja, 26006 Logroño, Spain

²Institute of Grapevine and Wine Sciences (University of La Rioja, Consejo Superior de Investigaciones Científicas, Gobierno de La Rioja), 26007 Logroño, Spain

³Department of Computer Science and Artificial Intelligence (DECSAI), Andalusian Research Institute in Data Science and Computational Intelligence (DaSCI), University of Granada (UGR), 18071 Granada, Spain

Resumen. La estimación del rendimiento es muy importante para la industria vitivinícola, ya que proporciona información útil para la gestión de viñedos y bodegas. Los efectos del cambio climático, como el aumento de las temperaturas y la menor disponibilidad de agua, pueden afectar a los componentes del rendimiento del viñedo. En general, la previsión tradicional del rendimiento se basaba en el recuento manual y destructivo de los racimos y el peso de las bayas. Los métodos convencionales no proporcionan una estimación precisa y son lentos, caros y laboriosos. En este estudio, se utilizaron métodos novedosos que emplean tecnologías digitales como el uso de la detección próxima, la visión por ordenador y la inteligencia artificial para la estimación del rendimiento en viñedos comerciales. La visión por ordenador se utilizó para la detección automática de diferentes características del dosel y para la calibración de ecuaciones de regresión para la predicción del rendimiento por cepa. La inteligencia artificial se utilizó para el recuento automático de racimos. Los resultados mostraron que el algoritmo de aprendizaje profundo fue capaz de detectar racimos con una alta precisión. En conclusión, nuestros resultados demostraron la aplicabilidad de estos nuevos métodos para evaluar los componentes del rendimiento en viñedos comerciales.

Abstract. Yield estimation is very important for the wine industry since provides useful information for vineyard and winery management. Climate change effects such as higher temperatures and lower water availability can affect vineyard yield components. In general, traditional yield forecasts are based on destructive manual counting of bunches and berry weight. These conventional methods do not provide accurate estimations and are time-consuming, expensive, and labour-intensive. In this study, novel methods using digital technologies such as computer vision and artificial intelligence were used to estimate yield in commercial vineyards. Computer vision was used for the automatic detection of different canopy features and the calibration of regression equations for the prediction of yield per vine. Artificial intelligence was used for the automatic counting of bunches. The results showed that the deep learning algorithm was able to detect bunches with high accuracy. Our results demonstrate the applicability of these novel methods to assess yield components in commercial vineyards.

1 Introducción

El cultivo de la vid es importante económicamente y una evaluación precisa de sus componentes productivos es fundamental para una gestión eficiente y sostenible. Sin embargo, los métodos convencionales son laboriosos, destructivos, lentos y de baja precisión [1]. Por ello, se requieren nuevos métodos para la evaluación del rendimiento de la vid que reemplacen los procedimientos tradicionales.

Hoy en día, una de las técnicas más comunes y exitosas para la evaluación de la producción en viticultura ha sido la visión por ordenador. En la literatura científica se han descrito varias aplicaciones y métodos [2]. Los sistemas de visión por ordenador se han utilizado para estimar el rendimiento de la vid en diferentes etapas fenológicas, como la brotación [3], la floración [4], en baya tamaño guisante [5] y en vendimia [6]. Las técnicas de visión por ordenador utilizadas para la detección de racimos se clasifican principalmente en tres clases: i) características de color [7], ii) segmentación activa de contornos [8], y iii) segmentación de píxeles [9]. En general, la visión por ordenador ha mostrado buenos resultados para la detección de racimos; sin embargo, los resultados de estas técnicas están muy influenciados por las condiciones de adquisición de la imagen, como los efectos de fondo y las condiciones lumínicas, y las condiciones intrínsecas del dosel, como la oclusión de los racimos [9]. En este contexto, las nuevas técnicas de inteligencia artificial pueden ayudarnos a resolver estos problemas. Los métodos de aprendizaje profundo han demostrado ser muy eficaces en la detección de objetos [10]. Esta novedosa técnica ha mostrado resultados prometedores para la detección y el conteo de racimos [11].

Por lo tanto, el objetivo de este trabajo fue la evaluación de los componentes del rendimiento utilizando la visión artificial y la inteligencia artificial en viñedos comerciales.

2 Material y métodos

2.1 Diseño experimental

Los muestreos se realizaron en 2020 en varios viñedos comerciales de la variedad cv. Tempranillo (*Vitis vinifera* L.) situados en la denominación de origen Rioja, España. Todos los viñedos se formaron en un sistema de espaldera de posicionamiento vertical con dos pares de alambres móviles (VSP). Todos los viñedos fueron sometidos a prácticas culturales estándar similares durante la temporada de desarrollo. En cada viñedo se eligieron al azar 25 cepas antes de la vendimia. Todas las cepas se dividieron en dos segmentos y se etiquetaron en consecuencia. Se colocaron dos trozos de cinta de baliza

de 0,5 m a ambos lados de cada vid para delimitar la anchura de la región de interés (ROI).

2.2 Adquisición de imágenes

Las imágenes del dosel se tomaron con una cámara RGB convencional (Canon EOS 5D Mark IV RGB, Canon Inc. Tokio, Japón) sin iluminación artificial (condiciones ambientales no controladas). La cámara se montó en un trípode orientado hacia la espaldera, a 1,0 m del eje de la hilera y a 1,20 m por encima del suelo. Se colocó una pantalla blanca detrás de la cubierta para eliminar la influencia de la vegetación de fondo. Las imágenes se guardaron en formato JPG con el ajuste de calidad más alto disponible en la cámara. El tamaño de la imagen completa era de 6720 × 4480 píxeles.

2.3 Datos reales

Tras la adquisición de imágenes, se vendimiaron los racimos, se contaron manualmente por cada segmento de vid (número real de racimos) y se pesó la producción de uva por segmento.

2.4 Etiquetado de imágenes

2.4.1 Etiquetas segmentación por color

El algoritmo de segmentación de imágenes se entrenó mediante aprendizaje supervisado. Para ello, se definieron las siguientes cinco clases que cubren los objetos esperados en las imágenes: "racimo", "tronco", "pámpano", "hoja", "hueco" y "alambre". A continuación, se seleccionaron manualmente varias muestras de píxeles de cada clase.

2.4.2 Etiquetas detección automática de racimos

Se realizó un etiquetado manual seleccionando los racimos visibles dentro de la ROI de cada imagen con cuadros delimitadores mediante el software LabelImg [12]. La clase de racimo se utilizó para etiquetar el conjunto de datos de este estudio utilizando el formato de etiqueta YOLO.

2.5 Procesado de las imágenes

El algoritmo de análisis de imágenes por segmentación utiliza la distancia de Mahalanobis para clasificar cada píxel de una imagen en función de su color. El algoritmo utiliza una muestra conocida de valores de color para clasificar un lote desconocido de píxeles en grupos o clases basados en un vector característico, estos grupos fueron previamente definidos en las cinco clases mencionadas anteriormente (Fig. 1).



● Hojas ● Racimos ● Madera y pámpanos ○ Huecos



Figura 1. Ejemplo de un segmento de vid tras el procesamiento de la imagen mediante el algoritmo de visión por ordenador. (a) imagen adquirida en campo, (b) resultado de la imagen procesada con el algoritmo de visión por ordenador.

El algoritmo de detección de objetos se modeló a partir de la arquitectura YOLOv4 [13], implementada con darknet [14] y utilizando una configuración modificada del código original publicado por Bochkovskiy [15]. El proceso de entrenamiento completo tardó unas 24 horas en completarse.

3 Resultados y Discusión

3.1 Componentes de la producción

La Tabla 1 muestra los diferentes valores estadísticos de los componentes del rendimiento en los viñedos comerciales. El coeficiente de variación del rendimiento mostró diferencias considerables entre los distintos segmentos, y este comportamiento también estuvo representado en los valores de rendimiento mínimo y máximo, oscilando entre 0 y 6,63 kg/m. Además, se observó una alta variabilidad de rendimiento, con valores medios de producción de 2,53 kg/m y con un coeficiente de variación del 53,26% mostrando una alta variabilidad intraparcelaria e interparcelaria.

Otros componentes del rendimiento como el número de racimos por metro fueron más homogéneos, con aproximadamente 9 racimos por metro en todos los casos y un coeficiente de variación medio del 42,41%. Este coeficiente representa la alta variabilidad del número de racimos en cada segmento, que va de 0 a 20. En relación con el peso del racimo, 0,28 kg fue la media global, también se observó una considerable variabilidad con un coeficiente de variación medio del 38,87%.

Tabla 1. Estadísticas descriptivas de los componentes del rendimiento de la vid (producción, número de racimos y peso de los racimos) en los viñedos analizados en este estudio.

	Media	SD	CV (%)	Min	Max
Producción (kg/m)	2.53	1.35	53.26	0.00	6.63
Número de racimos (Racimos/m)	9.06	3.84	42.41	0.00	20.00
Peso del racimo (kg)	0.28	0.11	38.87	0.00	0.71

SD: desviación estándar; CV: coeficiente de variación; Min: mínimo; Max: máximo.

3.2 Evaluación del dosel mediante visión por ordenador

Las imágenes del dosel obtenidas se procesaron con algoritmos de visión por ordenador para obtener valores de porosidad, exposición de las hojas y exposición del racimo. Las estadísticas descriptivas de estas características se muestran en la Tabla 2.

La porosidad del dosel, la exposición de las hojas y la exposición de los racimos mostraron una alta variabilidad. La porosidad obtuvo valores medios de 19,80% con un coeficiente de variación cercano al 60% mostrando diferencias notables entre los distintos segmentos y esto se vio reflejado en el porcentaje de porosidad mínima y máxima oscilando entre 0,34% y 68,88%, donde se observó que hubo segmentos con baja porosidad y otros con grandes niveles de porosidad. La exposición foliar fue el parámetro más estable, con un valor medio de exposición foliar de 42,33% y un coeficiente de variación del 43,65%, aunque se observaron grandes diferencias entre el porcentaje mínimo y máximo de exposición foliar, yendo desde valores muy bajos cercanos al 3% hasta tener doseles muy vigorosos con exposición foliar superior al 80%. La exposición del racimo obtuvo valores más bajos. El valor medio de exposición de racimos fue cercano al 11% con un alto coeficiente de variación superior al 62% pudiendo tener segmentos con valores de 0,90% hasta de 32,08%.

Tabla 2. Estadística descriptiva de la porosidad del dosel (porcentaje de píxeles de hueco), exposición foliar (porcentaje de píxeles de hojas) y exposición del racimo (porcentaje de píxeles de racimos) calculados mediante visión por ordenador en los viñedos analizados en este estudio.

	Media	SD	CV (%)	Min	Max
Porosidad (%)	19.80	11.81	59.65	0.34	68.88
Exposición foliar (%)	42.33	18.48	43.65	3.17	82.32
Exposición del racimo (%)	10.68	6.63	62.10	0.90	32.08

SD: desviación estándar; CV: coeficiente de variación; Min: mínimo; Max: máximo.

3.3 Estimación de la producción mediante la visión por computador

La oclusión por hojas y por fruta supone un reto en la estimación de la producción. La estimación de la producción mediante la visión por ordenador no tuvo buenos resultados, ya que la presencia predominante de hojas impide que la exposición del racimo se correlacione con el rendimiento real de la planta con un R^2 de 0,42.

3.4 Detección de racimos mediante inteligencia artificial

Para evaluar el efecto de la oclusión en la detección de racimos, se realizó una comparación entre el número real de racimos por segmento (conteo manual en campo en la vendimia) frente al número de racimos visibles contados en las imágenes RGB. La regresión lineal indicó una fuerte correlación entre el número de racimos reales y el número de racimos visibles con un R^2 de 0,81 y un RMSE de 1,71 racimos.

El rendimiento del algoritmo de aprendizaje profundo (arquitectura YOLOv4) se analizó mediante la comparación con los racimos visibles contados manualmente en cada imagen. Los resultados de la validación mostraron unos valores de rendimiento con un mAP del 74,58%, un Recall de 0,74, un F1-score de 0,71 y una Precisión de 0,68.

En la Fig. 2 se muestran las imágenes etiquetadas manualmente por un experto y el resultado de la predicción realizada por el algoritmo de aprendizaje profundo. La regresión lineal entre el número de racimos visibles frente al número de racimos predicho confirmó el buen rendimiento de YOLOv4 con un R^2 de 0,72 y un RMSE de 1,14.

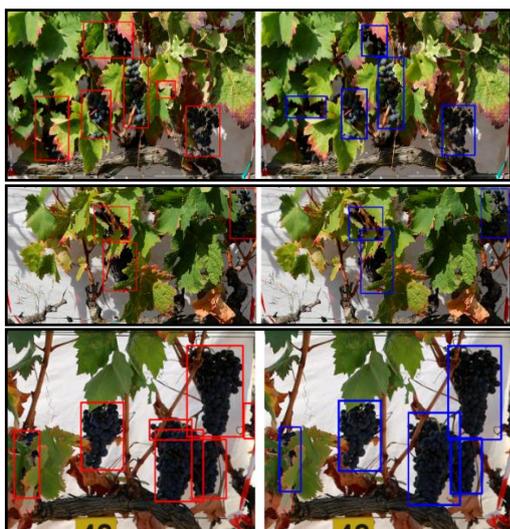


Figura 2. En la parte izquierda: imagen original etiquetada con el número de racimos visibles detectados y etiquetados por un experto (cajas rojas). En la parte derecha: imagen de los racimos predichos por el algoritmo (cajas azules).

4 Conclusiones

Los resultados presentados en este trabajo señalan que se pueden encontrar grandes diferencias en las condiciones y características del dosel de la vid, incluso de la misma variedad y región vinícola. El impacto de la oclusión sobre los racimos afectó de forma negativa a la capacidad de la visión por ordenador para estimar la producción del viñedo. Por ello, nuevas técnicas como la inteligencia artificial pueden ayudar a la detección y estimación del número de racimos en el viñedo. El algoritmo de aprendizaje profundo YOLOv4 fue capaz de detectar el número de racimos visibles con unos buenos resultados, a pesar de que la oclusión foliar y la oclusión de otros racimos afectó negativamente al rendimiento del modelo.

Referencias

1. S. Martin, R. Dunstone, G. Dunn. *GWRDC* **100** (2003)
2. L. Mohimont, F. Alin, M. Rondeau, N. Gaveau, L.A. Steffanel. *Agronomy* **12**, 2463 (2022)
3. S. Liu, S. Cossell, J. Tang, G. Dunn, M. Whitty. *Computers and Electronics in Agriculture* **137**, 88–101 (2017)
4. F. Palacios, G. Bueno, J. Salido, M. P. Diago, I. Hernández, J. Tardáguila. *Computers and Electronics in Agriculture* **178**, 105796 (2020)
5. F. Palacios, P. Melo-Pinto, M. P. Diago, J. Tardáguila. *Biosystems Engineering* **218**, 175–188 (2022)
6. B. Xin, S. Liu, M. Whitty. *Australian Journal of Grape and Wine Research* **26**(3), 207–219 (2020)
7. C. Hacking, N. Poona, C. Poblete-Echeverria. *OENO One* **54**, 793–812 (2020)
8. J. Xiong, Z. Liu, R. Lin, R. Bu, Z. He, Z. Yang, C. Liang. *Sensors* **18**, 969 (2018)
9. R. Íñiguez, F. Palacios, I. Barrio, I. Hernández, S. Gutiérrez, J. Tardáguila. *Agronomy* **11**(5), 1003 (2021)
10. A. Fuentes, S. Yoon, S.C. Kim, E.D.S. Park. *Sensors* **17** (2017)
11. M. Sozzi, S. Cantalamessa, A. Cogato, A. Kayad, F. Marinello. *Agronomy* **12**, 319 (2022)
12. D. Tzatalin. *LabelImg*. GitHub Repository 6 (2015)
13. A. Bochkovskiy, C. Y. Wang, H. Y. M. Liao. *arXiv preprint arXiv:2004.10934* (2020)
14. J. Redmon. *Darknet: Open source neural networks in C* (2013)
15. A. Bochkovskiy. *Yolo v4 repository [source code]*. <https://github.com/AlexeyAB/darknet> (2020)