

# IDENTIFICACIÓN DE VARIEDADES DE ACEITUNAS A PARTIR DEL ENDOCARPO UTILIZANDO VISIÓN POR COMPUTADOR

E. M. Ortega Vázquez\*, D. Martínez Gila, S. Satorres Martínez, J. Gómez Ortega, J. Gámez García  
Grupo de Robótica, Automática y Visión por Computador (GRAV), Universidad de Jaén, 23071 España

\*Autor en correspondencia. Email: emortega@ujaen.es

## Resumen

*La identificación varietal es fundamental en los distintos estadios relacionados con la producción de aceite de oliva virgen, comenzando por su cultivo y posterior procesado hasta su comercialización. Realizar la determinación varietal de manera precisa es un proceso que requiere de mucho tiempo, además de profesionales entrenados o la utilización de un equipo específico. Uno de los métodos más empleados es el tradicional, también conocido como morfológico, consistente en la evaluación, por parte de un experto, de ciertas características del hueso de la aceituna o endocarpo. En este trabajo se presenta una propuesta de automatización del método morfológico. En primer lugar, se obtendrán y procesarán, mediante Wilk's Lambda, características del endocarpo para, posteriormente, identificar la variedad mediante el uso de dos clasificadores: el análisis discriminante por mínimos cuadrados parciales y diferentes máquinas de soporte vectorial. En la validación de esta propuesta se han utilizado 250 ejemplares de 5 variedades procedentes del sur de España.*

**Palabras clave:** Características del endocarpo, Visión por computador, Identificación varietal, Wilk's Lambda, Análisis discriminante por mínimos cuadrados parciales, Máquinas de soporte vectorial

## 1 INTRODUCCIÓN

La identificación varietal en el sector del olivar es un aspecto clave para su planificación y gestión. El conocimiento de la variedad o variedades a cultivar permitirá una mejor adecuación de los requerimientos agronómicos que permitan maximizar la producción. En la almazara también es importante conocer las características de la materia prima de entrada. Esto permitirá un mejor ajuste de la maquinaria involucrada en el proceso de extracción del aceite de oliva virgen, lo que repercutirá directamente en la calidad final del producto y en su agotamiento (cantidad de aceite obtenido) [1]. Finalmente, en la comercialización del producto será necesaria la trazabilidad de los frutos, confirmación de su origen geográfico y autenticidad [2]. Esto posibilitará la obtención de aceite con Denominación de Origen Protegida, constituyendo este etiquetado una gran motivación para una amplia

gama de consumidores, ya que se considera estar asociado con productos de alta calidad [3].

Hoy en día se emplean dos métodos para determinar la variedad de origen: genéticos y morfológicos [4]. Los métodos genéticos son más precisos, pero requieren de un equipamiento muy específico. Los métodos morfológicos se centran en el análisis del árbol, hojas, flores, frutos y endocarpos [5]. De entre todas las fuentes anteriores la que ofrece una información más valiosa es el endocarpo. Esto es debido a que es independiente de las condiciones agroclimáticas, siendo, las características extraídas de él, muy similares para aceitunas de una misma variedad y ligeramente distintas para aquellas pertenecientes a diferentes variedades.

Por su coste, eficiencia y repetitividad, el método morfológico es el más empleado [6] y una tecnología factible de utilizar para su automatización sería la visión por computador. Varios autores han hecho uso de esta tecnología para la identificación varietal. Díaz *et al.* [7] clasificó aceitunas en cuatro categorías usando un análisis discriminante Bayesiano. Bari *et al.* [8] caracterizó nueve genotipos de oliva usando fractales en combinación con momentos y diversos clasificadores estadísticos. Los estudios de Beyaz *et al.* [9], [10] se concentraron en identificar cultivos procedentes de Turquía y España extrayendo características de frutos, hojas y endocarpos. Piras *et al.* [11] estudió el uso de características biométricas del endocarpo para determinar relaciones entre olivos silvestres y cultivados. Por otro lado, Vanloot *et al.* [12] clasificaron cinco cultivos franceses, muy diferentes entre sí, con una tasa de acierto del 100 %.

El propósito de este trabajo es presentar una propuesta de automatización del método morfológico de identificación varietal. En la evaluación de la propuesta se han empleado cinco de las variedades más extendidas en Andalucía siendo, en algunos casos, muy similares entre sí.

Este documento se organiza de la siguiente forma. La sección 2 presenta el método de preparación de las muestras, la adquisición de imágenes y los algoritmos de procesamiento de las mismas. En la sección 3 se muestran los resultados de los distintos clasificadores y, por último, en la sección 4 se detallan las conclusiones.

## 2 MATERIALES Y MÉTODOS

La propuesta de automatización del método morfológico se ha dividido en las cinco fases detalladas a continuación. Es importante destacar que se ha dedicado una especial atención a la selección de variedades, preparación de las muestras y adquisición de la imagen. Las variedades deben ser representativas de la zona geográfica y tratadas para que el endocarpo no presente ningún tipo de resto orgánico. El set-up de adquisición de imagen ha permitido disponer de imágenes libres de reflejos y sombras indeseadas que han facilitado el posterior procesado de las mismas.

### 2.1 FASE 1. SELECCIÓN DE ACEITUNAS EN CAMPO Y PREPARACIÓN DE LAS MUESTRAS

Se recolectaron cinco variedades de frutos de olivos durante enero-febrero de 2018 procedentes de diferentes cultivos localizados en Andalucía. Las aceitunas se recogieron manualmente cuando los frutos alcanzaron un índice de madurez de 5. En dicho estado de madurez, la piel del fruto es completamente negra y menos de la mitad de la pulpa es púrpura. En este estado, el endocarpo está totalmente desarrollado [13]. Para asegurar que las muestras son representativas, se escogieron varios árboles con una carga de cultivo y madurez homogéneas. Un total de 250 frutos, 50 por variedad, se recogieron por la zona media de la copa de los olivos.

Respecto a la preparación de endocarpos, se probaron tres diferentes métodos de limpieza sobre huesos de prueba: agua a presión, limpiador ultrasónico y una solución basada en hipoclorito sódico comúnmente conocida como lejía. El primer método de limpieza fue desechado por producir daños en la textura del endocarpo. El segundo método tampoco fue factible, al no conseguirse la eliminación completa de materia orgánica. Los mejores resultados se obtuvieron con el tercer método consistente en la aplicación de una solución del 5% de lejía durante una hora. Tal y como se muestra en la Figura 1, transcurrida una hora es posible disponer de endocarpos libres de pulpa y sin daños en su textura superficial.



Figura 1. Efecto de la solución de lejía en la textura del endocarpo.

### 2.2 FASE 2. ADQUISICIÓN DE IMÁGENES

La adquisición de imágenes se ha realizado mediante una cámara CMOS Genie Nano G3-GC11-C2420 con sensor de color, resolución 2464x2056 y tamaño de píxel de 3.45 $\mu$ m. La cámara, con una lente de 50 mm y anillos de extensión de 15 mm, se ha colocado a 180 mm del endocarpo.

Cada endocarpo se ha colocado en tres posiciones diferentes: alzado y perfil frente a la cámara y planta o posición vertical paralela al eje de la cámara. El campo de visión se obtuvo teniendo en cuenta el caso más restrictivo: la dimensión más grande en las imágenes en planta o alzado. Según la Ec.1, el campo de visión, FOV, se calculó como enuncia [14]:

$$FOV = \text{máximo tamaño de endocarpo} + \text{tolerancia en el posicionamiento} + \text{margen} + \text{adaptación a la relación de aspecto del sensor} \quad (1)$$

donde el máximo tamaño de endocarpo fue 2.5 mm, la tolerancia y el margen es 0.5 mm y la relación de aspecto, 4:3. Por tanto, el FOV se calculó de acuerdo a la Ec. 2:

$$\begin{aligned} FOV_{vertical} &= 2.5mm + 0.5mm + 0.5mm \\ &= 3.5mm \\ FOV_{horizontal} &= 3.5mm * \frac{4}{3} = 4.66mm \end{aligned} \quad (2)$$

### 2.3 FASE 3. SEGMENTACIÓN DE LA IMAGEN

El propósito principal de esta fase es el de extraer el endocarpo del fondo de la imagen. La Figura 2 muestra el procedimiento de segmentación de la imagen. En primer lugar, cuatro imágenes (correspondientes a los 3 canales RGB y el monocromo) se fusionan en una sola. El resultado de esta fusión será una imagen en la que se maximizará el contraste entre fondo y endocarpo, obtenida a partir de la Ec. 3. Donde  $H(i, j)$  será la imagen resultante y  $f(i, j, k)$ , la hipermatriz compuesta por los canales de color y monocromo.

$$H(i, j) = \max_{k=1, \dots, 4} f(i, j, k) \quad (3)$$

Posteriormente se realiza la segmentación de la imagen mediante una técnica de umbralización global, en concreto el método Ridler [19]. Se trata de un procedimiento iterativo en el que, partiendo de un umbral inicial  $U$ , se realiza una primera subdivisión de la imagen en dos grupos correspondientes al endocarpo y fondo. De cada grupo se obtiene el nivel medio de gris,  $m_e(T_n)$  y  $m_f(T_n)$  respectivamente, y el

siguiente umbral de segmentación será la media del nivel de gris de ambos grupos. El algoritmo finalizará en la iteración  $n$  cuando se cumpla que:  $|U_{n-1} - U_n| < \text{diferencia}$ , siendo *diferencia* un parámetro determinado empíricamente. Por tanto, el umbral óptimo se calcula aplicando la Ec. 4.

$$U_{opt} = \lim_{n \rightarrow \infty} \frac{m_e(T_n) + m_f(T_n)}{2} \quad (4)$$

Finalmente, el procesado morfológico de la imagen se utiliza para restablecer las regiones correspondientes al endocarpo (rellenar los huecos producidos tras realizar la umbralización global) y más específicamente la operación de cierre. Denotado por  $A \cdot B$ , la operación de cierre es una dilatación seguida de una erosión (Ec. 7):

$$A \cdot B = (A \oplus B) \ominus B \quad (7)$$

donde  $B$  es el elemento estructurante considerando la geometría de la región que se debe restaurar. Se ha utilizado un elemento estructurante con forma de disco para este paso del procesado de la imagen, con tamaño obtenido empíricamente para cerrar todos los huecos posibles sin modificar excesivamente el contorno del endocarpo.

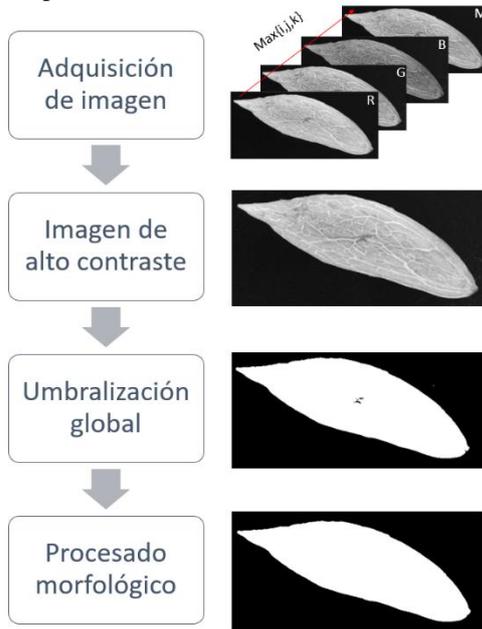


Figura 2. Procedimiento de segmentación de la imagen.

Si existe ruido en el fondo de la imagen, que puede haber sido magnificado por las operaciones morfológicas anteriores, se ha eliminado filtrando todos los objetos conectados de la imagen binaria con el criterio de selección de área y escogiendo el de mayor tamaño, correspondiente en este caso al endocarpo.

## 2.4 FASE 4. EXTRACCIÓN DE CARACTERÍSTICAS

Una metodología de identificación de variedades de olivo mediante características del endocarpo se publicó en [15]. La caracterización utilizada en dicho catálogo se basa en: forma, simetría en dos posiciones diferentes, ancho transversal máximo, ápice, base, superficie, mucrón (zona superior del ápice) y distribución de surcos fibrovasculares. Esta información es la utilizada por los expertos en la aplicación del método morfológico y, así mismo, ha sido la empleada en el presente trabajo.

A partir de los endocarpos, segmentados en el paso anterior, se extrajeron ocho conjuntos de características. La Tabla 1 presenta las características morfológicas del endocarpo, las posiciones en las que se miden y cómo se derivaron de los algoritmos de procesado de imágenes.

Tabla 1. Lista de las características más representativas del endocarpo.

Característica	Descripción	Implementación
[1] Forma	Determinado por longitud y ancho máximos del endocarpo	Relación entre longitud máxima y ancho (L/A)
[2] Simetría(A)	Determinado por la correspondencia entre dos mitades posicionado en perfil	Superposición de las mitades ( $C_{sim}$ descrita en Ec. 8)
[3] Simetría (B)		
[4] Ancho máximo transversal	Distancia del ápice hasta el máx. diámetro transversal	Distancia desde el endocarpo hasta el eje menor (ancho)
[5] Ápice	Estimado por la forma del ápice	Ángulo del ápice desde el ancho transversal máx.
[6] Base	Estimado por la forma de la base	Ángulo de la base desde el ancho transversal máximo
[7-26] Superficie (A)	Determinado por la textura del endocarpo	Textura de la superficie medida por Haralick (14 características) y medidas estadísticas (6 características)
[27-46] Distribución de surcos fibrovasculares (C)	Estimado por la textura del endocarpo en la posición C	Textura del mucrón medida por Haralick (14 características) y medidas estadísticas (6 características)

Un total de 46 características se extrajeron de tres imágenes diferentes. Las posiciones que se utilizaron en la adquisición y descritas en [16] son las siguientes:

- A) Posición en la que el fruto presenta mayor asimetría dispuesto en posición de perfil.
- B) Giro de 90° respecto la posición A, donde se observa la parte más desarrollada.
- C) Posición vertical desde la base, paralela al eje óptico, para observar la distribución de las fibras vasculares.

Como se resume en la Tabla 1, se han calculado seis características morfológicas. Todas ellas se apoyan en el cálculo de la longitud máxima (Figura 3b). Para el cálculo de la simetría ( $C_{sim}$ ) se realiza un abatimiento en el eje de la longitud máxima, con objeto de obtener la diferencia de áreas antes y después del abatimiento. La simetría se obtiene según la Ec. 8:

$$C_{sim} = 1 - \sum_{i,j=1}^{n,m} \frac{A_{M(i,j)} \& \overline{A_{m(i,j)}}}{A_{T(i,j)}} \quad (8)$$

donde  $A_M$  es el área mayor,  $A_m$  el área menor y  $A_T$  el área total sin divisiones. El área obtenida en la operación del numerador se corresponde al área magenta en la Figura 3c.

La textura se describe mediante 20 características basadas en propiedades estadísticas del histograma de intensidad y descriptores de Haralick [17] [18]. Estas características no son igualmente informativas entre variedades y algunas de ellas pueden ser ruidosas, correlacionadas o irrelevantes.

La selección de un conjunto de características con poder de discriminación, fiable e independiente es fundamental para el éxito del proceso de reconocimiento. Además, reducir la dimensionalidad de este conjunto simplificará el posterior diseño del clasificador.

La selección de características se ha realizado utilizando el criterio de Wilk's Lambda [19]. La figura 4 muestra el Wilk's Lambda ( $\lambda_j$ ) para cada una de las 46 características que se obtuvieron en etapas anteriores. Es importante destacar que la característica es más discriminante si  $\lambda_j$  es pequeña. Un estudio de las características que llevan a una mejor discriminación se incluye en la fase del diseño del clasificador.

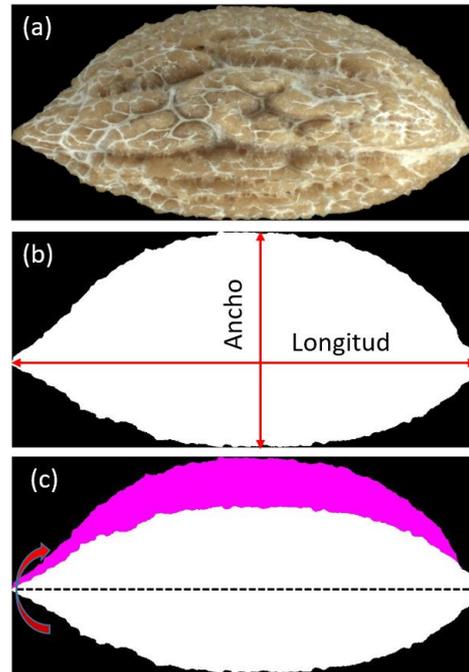


Figura 3. Extracción de características; (a) Imagen original; (b) Imagen segmentada con anchura y longitud máximas; (c) simetría en posición A (color magenta indica área no superpuesta después de abatir la zona inferior).

## 2.5 FASE 5. DISEÑO DE CLASIFICADORES

Se han implementado dos tipos de algoritmos de clasificación supervisada, evaluando los resultados de precisión y comprobando si éstos mejoraban introduciendo las características extraídas en el orden de Wilk's Lambda, incluyendo, en primer lugar, las características más discriminantes.

En ambos clasificadores se ha utilizado una validación *holdout*, con lo que se emplea un 20% de las muestras para validación y el resto para entrenar el modelo, realizando 1000 iteraciones y calculando la media aritmética para obtener el resultado de precisión.

### 2.5.1 Clasificador PLS-DA

El clasificador supervisado, análisis discriminante por mínimos cuadrados parciales (*Partial Least Squares Discriminant Analysis* o PLS-DA), basado en las características extraídas a los pasos anteriores, ha sido

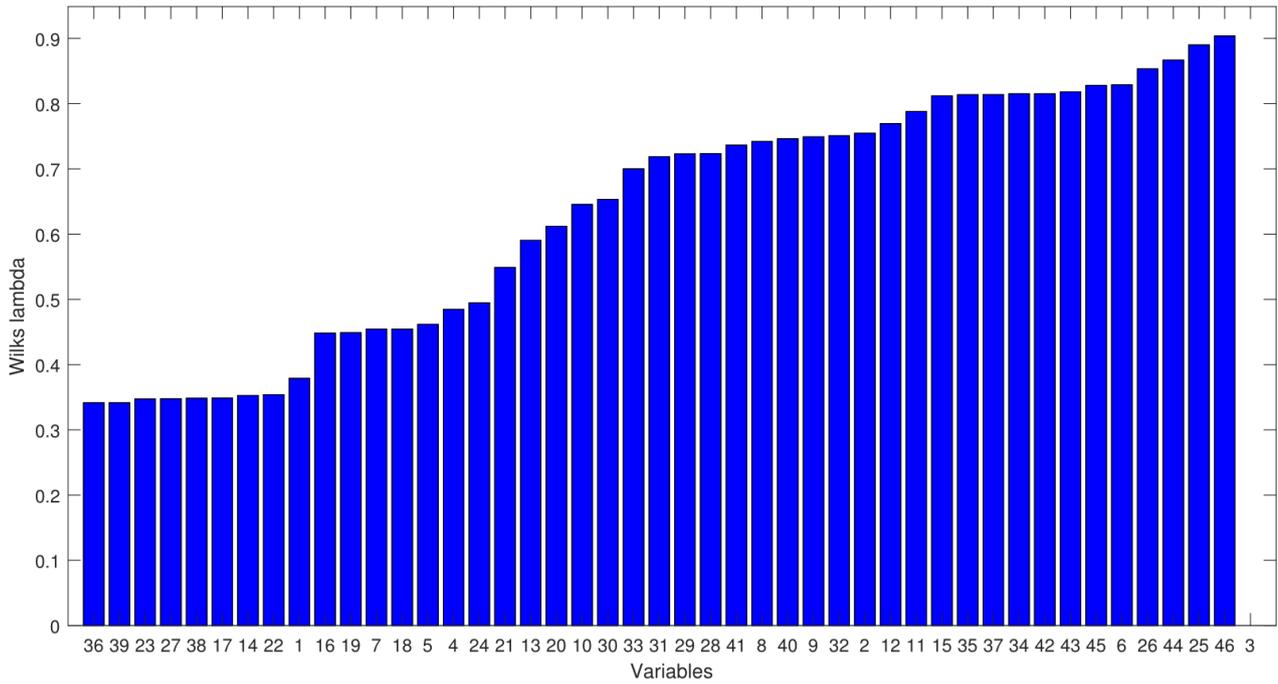


Figura 4. Evolución de Wilk's Lambda para características del endocarpo.

implementado para clasificar cada muestra de endocarpo según su variedad. PLS-DA es un método de clasificación lineal que combina las propiedades de una regresión de mínimos cuadrados parciales con la posibilidad de discriminación de un clasificador [20]. El algoritmo PLS-DA busca combinaciones lineales de las características, también llamadas variables latentes (*Latent Variables* o LVs), para maximizar la covarianza con las clases, en este caso, variedades de aceituna. El modelo multivariante general, PLS-DA se describe en la Ec. 9.

$$\begin{aligned} X &= TP^T \\ Y &= UQ^T \\ U &= T\beta \end{aligned} \quad (9)$$

donde  $T$  representa la matriz de puntuaciones (*scores*) de  $X$ ,  $P$  es la matriz de pesos (*loadings*) de  $X$ ,  $U$  contiene los *scores* de  $Y$ .  $Q$  es la matriz de *loadings* de  $Y$ . Los *loadings* pueden ser interpretados como los pesos de cada característica para determinar las LVs, mientras que las puntuaciones representan las coordenadas de las muestras en el hiperespacio de proyección LV. Asimismo,  $\beta$  contiene los coeficientes de regresión que proporcionan la relación entre  $U$  y  $T$ . Sustituyendo el modelo resultante se puede expresar como muestra la Ec. 10.

$$\begin{aligned} Y &= UQ^T = T\beta Q^T = X\beta Q^T \\ \tilde{A} &= \beta Q^T \end{aligned} \quad (10)$$

De acuerdo con los coeficientes  $\tilde{A}$ , el algoritmo PLS-DA devuelve los valores estimados ( $\hat{y}_{ig}$ ) para cada  $i_{th}$  (Ec. 11). Por tanto, la probabilidad de pertenecer a una clase específica  $g_{th}$  se obtiene mediante la función de densidad de probabilidad de cada clase (Ec. 12).

$$\hat{y}_i = x_i^T \tilde{A} \quad (11)$$

$$p(\hat{y}_i | g) = \frac{1}{N} \sum_{n=1}^N g_n(\hat{y}_i) \quad (12)$$

donde  $N$  es el número de muestras de calibración para la clase  $g$  y  $g_i(\hat{y})$  es la función potencial sobre cada predicción PLS [23]. De esta forma, valores bajos de  $p(\hat{y}_i | g)$  significan que la muestra  $i_{th}$  probablemente no pertenezca a la clase  $g_{th}$ , por el contrario, valores altos cercanos a uno indicarían lo opuesto. Asimismo, la clasificación se lleva a cabo eligiendo, entre todos los casos, el de mayor probabilidad. Bajo este enfoque, las muestras siempre se clasifican en una de las clases.

## 2.5.2 MÁQUINAS DE SOPORTE VECTORIAL

El segundo clasificador implementado está basado en máquinas de soporte vectorial (*Support Vector Machines*, SVM) [21]. Las SVM se basan en que un límite de decisión puede describirse como un hiperplano, que se expresa en términos de una combinación lineal de funciones parametrizadas por

vectores de soporte y que proporcionan el mejor hiperplano usando una función kernel (Ecs. 13 y 14).

$$\min_{\beta, \beta_0} \frac{1}{2} \|\beta\|^2 + C \sum_{i=1}^N \xi_i \quad (13)$$

$$\text{sujeto a } y_i(x_i^T \beta + \beta_0) \geq 1 - \xi_i \quad \forall i \quad (14)$$

En este contexto, se han comparado tres kernels: lineal, polinómico de grado 2 o cuadrática y sigmoidea.

Además,  $K(x_i, x_j) \equiv \phi(x_i)^T \phi(x_j)$  es la llamada función kernel.

Los kernel comparados fueron los siguientes:

- Lineal:  $K(x_i, x_j) = x_i^T x_j$ .
- Polinómico:  $K(x_i, x_j) = (\gamma x_i^T x_j + r)^d$ ,  
 $\gamma > 0$
- Sigmoidea:  $K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r)$

Donde  $\gamma$ ,  $r$  y  $d$  son parámetros del kernel.

El valor de coste óptimo de SVM y los parámetros del kernel se seleccionaron minimizando el error de validación cruzada.

### 3 RESULTADOS

Un total de 250 muestras para cinco variedades andaluzas—Picudo, Lucio, Cortijuelo, Manzanillo de Montefrío y Negrillo de Estepa— se utilizaron para evaluar la propuesta. Un simple examen visual de los endocarpos fue suficiente para comprender la dificultad del problema. Algunas de estas variedades eran muy similares en forma y tampoco era posible ver, a simple vista, diferencias significativas en la textura. Esto ocasionaba que en ciertos casos las características de geometría y/o textura no fuesen lo suficientemente discriminantes entre clases.

Debido a ello, se valoró la precisión del clasificador introduciendo iterativamente características de endocarpo en el orden dado por Wilk's Lambda. La Figura 5 muestra la evolución del error en la clasificación, en relación con las características utilizadas, para los distintos clasificadores implementados en esta propuesta.

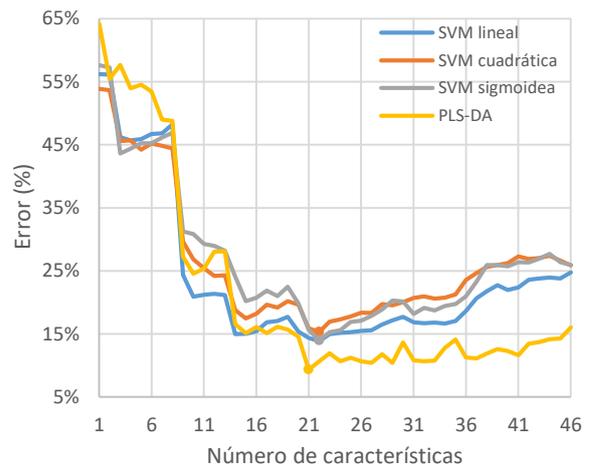


Figura 5. Comparación del error obtenido al incluir diferentes características.

Como se puede observar, los resultados mejoraron al elegir entre 21 y 22 características para los diferentes clasificadores, en lugar de utilizar el total de las 46 características inicialmente extraídas. Los resultados quedan resumidos en la Tabla 2, donde se compara la diferencia de precisión obtenida con y sin selección de características.

Tabla 2. Comparación de clasificadores.

Clasif.	Sin selección			Con selección		
	N	Pr	Er	N	Pr	Er
PLSDA	46	84,0%	16,0%	21	<b>90,7%</b>	9,3%
SVM-L	46	75,3%	24,7%	22	86,2%	13,8%
SVM-C	46	74,1%	25,9%	22	84,7%	15,3%
SVM-S	46	74,1%	25,9%	22	86,0%	14,0%

La columna  $N$  representa el número de características, las columnas  $Pr$  y  $Er$  representan el porcentaje de precisión y error, respectivamente.

Los mejores resultados de precisión media se obtuvieron con un 90.7% para el análisis parcial discriminante de mínimos cuadrados con selección de características. Las máquinas de soporte vectorial ofrecieron peores resultados, pero la selección de características tuvo una influencia positiva ya que mejoraron su desempeño en un 10%.

## 4 CONCLUSIONES

La identificación varietal es un aspecto clave y de vital importancia en el sector del aceite de oliva. Los métodos empleados hoy en día para tal fin requieren de un equipamiento muy específico y costoso o de expertos. Por su coste y eficiencia la identificación de variedades suele realizarse mediante expertos que, básicamente, analizan la geometría y textura del hueso de la aceituna. El objetivo de este trabajo ha sido el de automatizar este proceso. Extrayendo mediante visión por computador las características del hueso analizadas por el experto en el proceso manual, se han implementado distintos clasificadores para inferir la variedad: análisis parcial discriminante de mínimos cuadrados y máquinas de soporte vectorial. Los resultados son mejores para el caso del análisis parcial discriminante de mínimos cuadrados cuando, previamente, se seleccionan las características más discriminantes, llegándose a alcanzar un 90.7% de precisión en la clasificación. Es importante destacar que las 5 variedades analizadas son muy similares en forma y textura. Para obtener resultados similares a los presentados en esta propuesta, el proceso llevado a cabo por expertos incorpora información adicional del fruto, floración y árbol. En trabajos futuros se tendrá en cuenta esta información para conseguir tasas de precisión cercanas al 100%.

### Agradecimientos

Este trabajo ha sido parcialmente financiado por el proyecto DPI2016-78290-R. Los autores agradecen a la denominación de origen Poniente de Granada por las muestras de aceitunas proporcionadas sin las cuales no habría sido posible esta investigación.

### English summary

## OLIVE IDENTIFICATION BASED IN ENDOCARP FEATURES USING COMPUTER VISION

### Abstract

*The identification of olives is of great importance for a multitude of factors, including harvesting, olive oil production process and trade exchanges. Precisely identifying varieties is a time-consuming task in addition to trained experts or specific and expensive equipment. When applying the traditional method, also known as morphological, a specialist assesses*

*morphological features using the olive endocarp. In this paper a proposal to automate this identification methodology is presented. Endocarp images are acquired and analyzed to extract the endocarp features, processed by Wilk's Lambda. Then, the varieties are identified by two classifiers: partial least squares discriminant analysis and different support vector machines. The proposal has been tested on a set of 250 samples from 5 varieties from the south of Spain.*

**Keywords:** Endocarp features, Computer vision, Olive varietal identification, Wilk's Lambda, Partial least squares discriminant analysis, Support Vector Machines.

### Referencias

- [1] T. Ram, Z. Wiesman, I. Parmet, and Y. Edan, "Olive oil content prediction models based on image processing," *Biosyst. Eng.*, vol. 105, no. 2, pp. 221–232, Feb. 2010.
- [2] Z. Likudis, "Olive Oils with Protected Designation of Origin (PDO) and Protected Geographical Indication (PGI)," in *Products from Olive Tree*, InTech, 2016.
- [3] C. Parra-López, A. Hinojosa-Rodríguez, S. Sayadi, and C. Carmona-Torres, "Protected Designation of Origin as a Certified Quality System in the Andalusian olive oil industry: Adoption factors and management practices," *Food Control*, vol. 51, pp. 321–332, May 2015.
- [4] I. Trujillo *et al.*, "Identification of the Worldwide Olive Germplasm Bank of Córdoba (Spain) using SSR and morphological markers," *Tree Genet. Genomes*, vol. 10, no. 1, pp. 141–155, Feb. 2014.
- [5] K. N. Blazakis, M. Kosma, G. Kostelenos, L. Baldoni, M. Bufacchi, and P. Kalaitzis, "Description of olive morphological parameters by using open access software," *Plant Methods*, vol. 13, no. 1, p. 111, Dec. 2017.
- [6] M. Laaribi, I. Gouta, H. Ayachi, M. M., Labidi, F., Mars, "Combination of morphological and molecular markers for the characterization of ancient native olive accessions in Central-Eastern Tunisia.," *Biologies*, vol. 340, no. 2017, pp. 287–297.
- [7] R. Diaz, "Classification and Quality Evaluation of Table Olives," in *Computer Vision Technology for Food Quality*

- Evaluation*, Elsevier, 2016, pp. 351–367.
- [8] J. L. Bari, A. Martin, A., Boulouha, A., Gonzalez Andujar, “Characterization and identification of olive genotypes using an image feature extraction approach,.” in *IX Conferencia Espanola de Biometria La Coruna*, 2003, pp. 5–8.
- [9] A. BEYAZ and R. ÖZTÜRK, “Identification of olive cultivars using image processing techniques,” *TURKISH J. Agric. For.*, vol. 40, pp. 671–683, 2016.
- [10] A. Beyaz, M. T. Özkaya, and D. İçen, “Identification of some spanish olive cultivars using image processing techniques,” *Sci. Hortic. (Amsterdam)*, vol. 225, pp. 286–292, Nov. 2017.
- [11] F. Piras, O. Grillo, G. Venora, G. Lovicu, M. Campus, and G. Bacchetta, “Effectiveness of a computer vision technique in the characterization of wild and farmed olives,” *Comput. Electron. Agric.*, vol. 122, pp. 86–93, Mar. 2016.
- [12] P. Vanloot, D. Bertrand, C. Pinatel, J. Artaud, and N. Dupuy, “Artificial vision and chemometrics analyses of olive stones for varietal identification of five French cultivars,” *Comput. Electron. Agric.*, vol. 102, pp. 98–105, Mar. 2014.
- [13] A. Dag, Z. Kerem, N. Yogev, I. Zipori, S. Lavee, and E. Ben-David, “Influence of time of harvest and maturity index on olive oil yield and quality,” *Sci. Hortic. (Amsterdam)*, vol. 127, no. 3, pp. 358–366, Jan. 2011.
- [14] A. Hornberg, *Handbook of machine vision*. Wiley-VCH, 2006.
- [15] International Olive Council, “World Catalogue of Olive Varieties,” International Olive Council, Madrid, 2000.
- [16] D. Barranco, I. Trujillo, L. Rallo, J. M. Caballero, C. Del Río, and A. Martín, *Varietades de olivo en España*, Mundi-Pren. Madrid: Elaiografia Hispanica, 2005.
- [17] R. C. Gonzalez and R. E. (Richard E. Woods, *Digital image processing*, 4 edition. Pearson, 2017.
- [18] R. M. Haralick, K. Shanmugam, and I. Dinstein, “Textural Features for Image Classification,” *IEEE Trans. Syst. Man. Cybern.*, vol. SMC-3, no. 6, pp. 610–621, 1973.
- [19] A. el Ouardighi, A. el Akadi, and D. Aboutajdine, “Feature Selection on Supervised Classification Using Wilks Lambda Statistic,” in *2007 International Symposium on Computational Intelligence and Intelligent Informatics*, 2007, pp. 51–55.
- [20] N. F. Pérez, J. Ferré, and R. Boqué, “Calculation of the reliability of classification in discriminant partial least-squares binary classification,” *Chemom. Intell. Lab. Syst.*, vol. 95, no. 2, pp. 122–128, Feb. 2009.
- [21] C. Cortes and V. Vapnik, “Support-Vector Networks,” *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.



© 2018 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution CC-BY-NC 3.0 license (<https://creativecommons.org/licenses/by-nc/3.0>).