



## ALGORITMOS DE APRENDIZAJE AUTOMÁTICO PARA IDENTIFICAR VARIEDADES DE DURAZNO CON BASE EN DESCRIPTORES CROMÁTICOS Y MORFOLÓGICOS

### MACHINE LEARNING ALGORITHMS TO IDENTIFY PEACH VARIETIES BASED ON CHROMATIC AND MORPHOLOGICAL DESCRIPTORS

Daniel Ayala-Niño y Juan Manuel González-Camacho\*

Colegio de Postgraduados, Campus Montecillo, Posgrado SEI-Cómputo Aplicado, Montecillo, Estado de México, México.

\*Autor de correspondencia (jmgc@colpos.mx)

#### RESUMEN

La inteligencia artificial ha permitido desarrollar herramientas para el reconocimiento automático de frutos y hortalizas con mayor precisión y rapidez. El desarrollo de nuevos genotipos de árboles frutales requiere del uso de herramientas tecnológicas para la identificación de variedades con mayor robustez que los métodos convencionales. En esta investigación se aplicaron algoritmos de aprendizaje automático para identificar seis variedades de durazno (*Prunus persica* L.) CP-03-06, Oro Azteca, Oro San Juan, Cardenal, Colegio y Robin a partir de imágenes digitales de hojas. Los modelos máquina de soporte vectorial (SVM), bosque aleatorio (RF) y perceptrón multicapa (MLP) se entrenaron y evaluaron con base en tres descriptores cromáticos y 14 morfológicos extraídos de imágenes digitales. La evaluación del desempeño en predicción de los modelos se realizó con base en métricas globales y específicas para cada clase objetivo (variedad de durazno). Los cinco descriptores más importantes para identificar las variedades de durazno fueron los tres canales de color HSV (*hue, saturation, value*), la redondez y la excentricidad de las hojas. SVM obtuvo la mayor precisión global de clasificación con Acc de 98.7 % y  $F1_{macro}$  de 98 %. SVM obtuvo el mayor puntaje  $F1$  (99.2 %) para identificar la variedad de durazno CP-03-06 y el menor puntaje  $F1$  (96.1 %) para identificar la variedad Cardenal. La utilización conjunta de descriptores cromáticos y morfológicos mejoró el desempeño de los algoritmos de aprendizaje para identificar las seis variedades de durazno. Los modelos SVM, RF y MLP obtuvieron un Acc de 98.7, 98.6 y 97 %, respectivamente. Este estudio muestra el potencial de las técnicas de aprendizaje automático para su aplicación en el reconocimiento de descriptores de interés en cultivos agrícolas y su aplicación a procesos automatizados en la agricultura.

**Palabras clave:** *Prunus persica* L., aprendizaje automático, inteligencia artificial, reconocimiento de patrones, visión computacional.

#### SUMMARY

Artificial intelligence has allowed the development of tools for automatic recognition of fruits and vegetables with greater precision and speed. The development of new genotypes of fruit trees requires the use of technological tools to identify varieties with greater robustness than conventional methods. In this research, machine learning algorithms were applied to identify six peach varieties (*Prunus persica* L.) CP-03-06, Oro Azteca, Oro San Juan, Cardenal, Colegio and Robin from digital leaf images. The support vector machine (SVM), random forest (RF) and multilayer perceptron (MLP) models were trained and evaluated based on three chromatic and 14 morphological descriptors

extracted from digital images. The evaluation of the prediction performance of the models was based on global metrics and specific for each target class (peach variety). The five most important descriptors to identify peach varieties were three HSV color channels (*hue, saturation, value*), roundness and eccentricity of the leaves. SVM achieved the highest overall classification accuracy with Acc of 98.7 % and  $F1_{macro}$  of 98 %. SVM obtained the highest  $F1$  score (99.2 %) to identify the peach variety CP-03-06 and the lowest  $F1$  score (96.1 %) to identify the Cardenal variety. The joint use of chromatic and morphological descriptors improved the performance of learning algorithms to identify the six peach varieties. The SVM, RF and MLP models obtained an Acc of 98.7, 98.6 and 97 %, respectively. This study shows the potential of machine learning methods for their application in recognizing descriptors of interest in agricultural crops and their application to automated processes in agriculture.

**Index words:** *Prunus persica* L., artificial intelligence, computer vision, machine learning, pattern recognition.

#### INTRODUCCIÓN

El reconocimiento de frutos y vegetales con base en técnicas de inteligencia artificial e imágenes digitales ha permitido desarrollar tecnologías de visión artificial y automatización para mejorar la producción agrícola (Xiao et al., 2023). En México, el durazno es un fruto de alto consumo, en 2022 la producción fue de 239,133 Mg (SIAP, 2023). Para incrementar la producción agrícola se requiere desarrollar nuevas variedades de alta calidad y con bajos requerimientos de frío (Calderón-Zavala et al., 2019). Las investigaciones sobre detección y reconocimiento de frutos y vegetales se han desarrollado en dos ejes, el desarrollo de sensores [cámaras blanco y negro, cámaras RGB (*red, green, blue*), espectrales y térmicas] para la adquisición de imágenes digitales y el desarrollo de algoritmos de aprendizaje automático y aprendizaje profundo (Xiao et al., 2023).

Aakif y Khan (2015) utilizaron descriptores morfológicos de hojas de árboles frutales, como la excentricidad y la redondez, para entrenar una red neuronal artificial y

reportaron una precisión global de clasificación de 96 %. Apriyanti *et al.* (2013) identificaron 10 géneros de orquídeas con un modelo máquina de soporte vectorial (SVM) a partir de la distancia al centroide de la hoja, redondez y modelo de color HSV, SVM obtuvo una precisión global de 79.3 %. Wäldchen y Mäder (2018) realizaron una revisión de literatura sobre la aplicación de técnicas de visión por computadora para identificar plantas por medio de hojas y otros órganos; estos autores reportaron que el color de la hoja asociado con descriptores morfológicos mejoró el desempeño de los modelos, con precisiones globales que variaron de 50 a 99 %. Ayala-Niño y González-Camacho (2022) aplicaron algoritmos de aprendizaje automático para identificar variedades de durazno a partir de descriptores cromáticos del haz y envés de hojas, con los que se obtuvo una precisión global máxima de 84.1 %. García *et al.* (2016) aplicaron un modelo Bayesiano para identificar 20 frutos diferentes con base en descriptores cromáticos, texturales y morfológicos; con este modelo se alcanzó una precisión global de 95 %. Similarmente, Septiarini *et al.* (2021) utilizaron modelos Naïve Bayes, SVM y perceptrón multicapa para clasificar la madurez de racimos de fruta fresca de palma con base en descriptores cromáticos y de textura; con el perceptrón multicapa se obtuvo una precisión global máxima de 98.4 %.

Barré *et al.* (2017) utilizaron una red neuronal convolucional (CNN) para identificar nuevas variedades en diversas especies. Li *et al.* (2020) utilizaron una CNN pre-entrenada VGG-16 (Simonyan y Zisserman, 2015) como capa inicial acoplada a SVM y bosque aleatorio (RF) para identificar enfermedades en maíz, manzana y uva con una precisión global de 94 %. La extracción automática de descriptores por medio de CNN facilita la identificación de variedades; sin embargo, estos enfoques demandan mayores recursos de cómputo que los algoritmos de aprendizaje automático (Wäldchen y Mäder, 2018).

En esta investigación se evaluó el desempeño de tres modelos de aprendizaje automático para identificar seis variedades de durazno (*Prunus persica* L. Batsch) con base en descriptores cromáticos y morfológicos extraídos de imágenes digitales del haz de las hojas; se asume que los modelos de aprendizaje automático mejoran su desempeño para identificar las variedades de durazno con el uso conjunto de descriptores cromáticos y morfológicos extraídos de imágenes digitales de sus hojas.

## MATERIALES Y MÉTODOS

### Conjunto de datos

El conjunto de datos que se utilizó en este estudio proviene del *peach dataset*, que cuenta con 2530 imágenes

de hojas de seis variedades de durazno: CP-03-06 (V1), Oro Azteca (V2), Oro San Juan (V3), Cardenal (V4), Colegio (V5) y Robin (V6), que se obtuvieron en 2019 en el campo experimental del Colegio de Postgraduados, Campus Montecillo, México (<https://bit.ly/3cvA9MF>). Las imágenes de hojas se capturaron con una cámara EOS REBEL T5i (Canon, Tokio, Japón), en un entorno controlado con fondo blanco e iluminación artificial con una resolución de 72 pixeles por pulgada (ppp) y tamaño de 720 × 480 pixeles en formato JPG (Ayala *et al.*, 2020). Un conjunto de 480 imágenes de hojas de durazno se seleccionaron al azar (80 para cada variedad).

### Recursos de cómputo

Se utilizó una computadora portátil con procesador Intel Core i7 8va generación, con un procesador gráfico NVIDIA GTX 1070 MaxQ, lenguaje Python 3.9.5, y el programa OpenCV 4.5.3 para el análisis de imágenes. Los modelos de aprendizaje automático se implementaron en Scikit-learn 0.24.2 y se ejecutaron en la plataforma de modelación Google Colaboratory Pro.

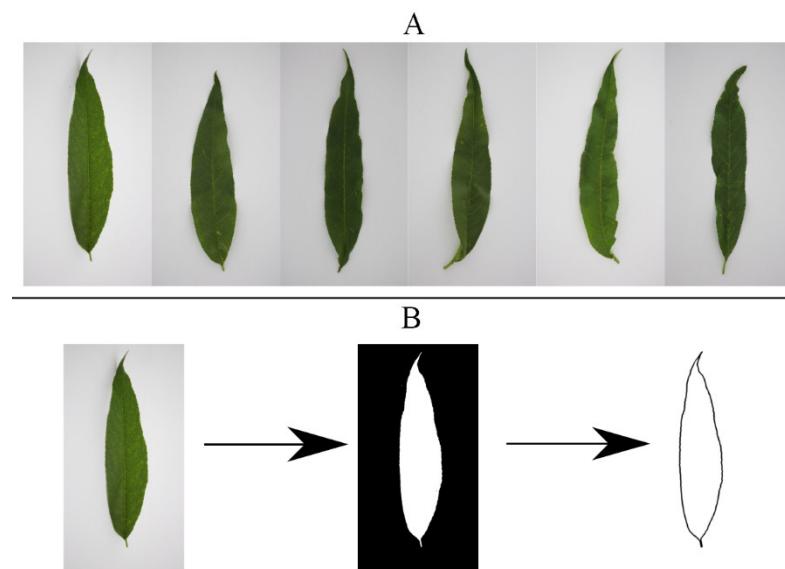
### Preprocesamiento de imágenes

Las imágenes se segmentaron para separar la hoja del fondo por medio del algoritmo de Otsu (Kumar y Sodhi, 2020). La segmentación consistió en aplicar un umbral a una imagen en escala de grises para obtener una imagen binaria, donde el fondo de la imagen se representa con pixeles iguales a 1 y la hoja con pixeles iguales a 0 (Otsu, 1979). Las imperfecciones debidas a variaciones de intensidad en las imágenes se eliminaron con un filtro Gaussiano (Marr y Hildreth, 1980) (Figura 1).

### Extracción de descriptores cromáticos y morfológicos

Los descriptores cromáticos se obtuvieron a partir de muestras de imagen de tamaño 30 × 30 pixeles del área de la venación primaria y secundaria del haz de la hoja (Lee *et al.*, 2017). Las muestras de píxeles obtenidas con el modelo de color RGB (*red, green, blue*) se transformaron al modelo de color HSV (*hue, saturation, value*) (García-Mateos *et al.*, 2015). La extracción de descriptores morfológicos se realizó con el programa OpenCV a partir del contorno de la hoja.

Los descriptores morfológicos fueron: ancho (*w*, distancia máxima perpendicular al eje mayor de la hoja), largo (*l*, distancia entre la base y la punta de la hoja, perímetro de la hoja [*p*, longitud del contorno de la hoja (Suzuki y Be, 1985)], área [*a*, superficie de la lámina foliar (Wäldchen y Mäder, 2018)], redondez (*r*, es la relación entre el área de la lámina foliar y el área de un círculo con igual



**Figura 1.** Descripción de variedades de durazno, A) de izquierda a derecha, variedad CP-03-06 (V1), variedad Oro Azteca (V2), variedad Oro San Juan (V3), variedad Cardenal (V4), variedad Colegio (V5) y variedad Robin (V6); B) segmentación de la imagen: clase V1 imagen original, obtención del contorno de la hoja y extracción de la hoja.

perímetro, que varía de 0 a 1 y se calcula como  $r = \frac{4\pi a}{P^2}$ , donde  $r = 1$  representa un círculo perfecto y  $r$  tiende a 0 en formas altamente no circulares), excentricidad [ $\varepsilon$ , es la distancia entre los focos de una elipse y su eje mayor, toma valores entre 0 y 1,  $\varepsilon = 0$  si el objeto es redondo, y  $\varepsilon = 1$  si se tiene un solo eje (Wäldchen y Mäder, 2018)]; los siete momentos (o promedios ponderados) de Hu son valores que describen la forma del objeto de interés en una imagen, y son invariantes a la traslación, escala, rotación y reflexión de una imagen (Hu, 1962).

El vector de descriptores de entrada se representa como:

$$\mathbf{x} = [x_h, x_s, x_v, x_w, x_l, x_{p'}, x_{d'}, x_r, x_{e'}, x_{w/l}, x_{hu1}, x_{hu2}, \dots, x_{hu7}]$$

Donde  $x_h$ ,  $x_s$  y  $x_v$  representan los valores cromáticos y las variables restantes son los descriptores morfológicos. En el conjunto de datos de entrada se eliminaron las muestras repetidas dentro de cada clase objetivo y entre clases objetivo.

El conjunto de datos de entrada consistió de 5976 muestras y 17 descriptores con diferente número de muestras por clase objetivo; esto es, V1 (2340), V2 (243), V3 (825), V4 (240), V5 (1726) y V6 (602). Los descriptores cromáticos y morfológicos más relevantes se seleccionaron a prueba y error con diferentes subconjuntos de descriptores y su impacto en el desempeño de los modelos; en cada subconjunto se utilizaron como base los descriptores cromáticos H, S, y V y se agregaron descriptores morfológicos.

## Modelos de aprendizaje automático

### Máquina de soporte vectorial

La optimización de un modelo máquina de soporte vectorial (SVM) consiste en encontrar un hiperplano óptimo que define los límites de decisión de los vectores de descriptores de entrada (Chauhan *et al.*, 2019). La función objetivo a minimizar se define como:

$$J(\mathbf{w}, b, \xi) = \min_{w, b, \xi} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i$$

sujeto a  $y_i(\mathbf{w}_0 \Phi(\mathbf{x}_i) + b) \geq 1 - \xi_i, \xi_i \geq 0$

Donde  $\mathbf{w}$  es el vector de pesos del hiperplano,  $b$  es el sesgo,  $\xi_i, i = 1, \dots, n$  son variables de holgura que ayudan a evitar problemas de convergencia en casos de no linealidad,  $\mathbf{x}_i \in R^p$  es el vector de  $p$  descriptores de entrada,  $y_i$  es la clase objetivo de la muestra  $\mathbf{x}_i$ , y  $n$  es el número de restricciones.

Las predicciones se obtienen por:

$$\hat{y}(\mathbf{x}) = \text{sgn} \left( \sum_{i=1}^N a_i K(\mathbf{x}_i, \mathbf{x}) + w_0 \right)$$

Donde  $a_i$  es un multiplicador de Lagrange,  $w_0$  es un parámetro de sesgo, y  $K(\mathbf{x}_i, \mathbf{x})$  es el kernel que transforma los datos a una dimensión mayor para resolver problemas no lineales; el kernel puede ser Gaussiano  $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2)$ , donde  $\|\mathbf{x}_i - \mathbf{x}_j\|^2$  es la norma cuadrática,  $\gamma$  es un hiperparámetro a optimizar; lineal  $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{x}_i - \mathbf{x}_j \rangle$ ; o

sigmoide  $K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(K(\mathbf{x}_i, \mathbf{x}_j) + \theta)$ .

### Bosque aleatorio

El modelo bosque aleatorio (RF) consiste en una colección de árboles de decisión independientes. La independencia entre árboles se obtiene por medio de una selección aleatoria de muestras (*bootstrap*) y la selección de un conjunto de descriptores de tamaño  $m = \sqrt{p}$  para cada árbol, donde  $p$  es el número de descriptores de entrada (Biau y Scornet, 2016).

RF busca incrementar la ganancia de información (IG) en los nodos de cada árbol; los datos se dividen de manera jerárquica de acuerdo con IG, la cual se define como:

$$IG(D_p, f) = I(D_p) - \frac{N_i}{N_p} I(D_i) - \frac{N_d}{N_p} I(D_d)$$

Donde  $D_p$ ,  $D_i$  y  $D_d$  son conjuntos de datos en el nodo padre, nodo izquierdo y nodo derecho, respectivamente;  $f$  es la función que divide los datos;  $N_p$ ,  $N_i$  y  $N_d$  son el número de datos en el nodo padre, nodo izquierdo y nodo derecho respectivamente; e  $I(\cdot)$  es el criterio de impureza. Los criterios de impureza se definen por medio del índice de Gini o de entropía. Gini se define como:

$$I_G(D_t) = 1 - \sum_{i=1}^c p(i|t)^2$$

Y la entropía como:

$$I_E(D_t) = -\sum_{i=1}^c p(i|t) \log_2 p(i|t)$$

donde  $p(i|t)$  es la proporción de muestras que pertenecen a la clase objetivo  $i$  en el nodo  $t$  y  $c$  es el número total de clases objetivo.

La predicción de una clase se obtiene por medio de mayoría de votos, para una clase binaria se tiene:

$$\hat{Y} = \begin{cases} 1, & \text{si } \frac{1}{nar} \sum_{j=1}^{nar} f_j(\mathbf{x}) > 1/2 \\ 0, & \text{de otra forma} \end{cases}$$

donde  $f_j$  es la predicción del  $j$ -ésimo árbol,  $\mathbf{x}$  es el vector de entrada,  $nar$  es el número máximo de árboles, el cual es un hiperparámetro y  $pmax$  es la profundidad máxima de un árbol.

### Perceptrón multicapa

El modelo perceptrón multicapa (MLP) es un algoritmo que se estructura de la entrada a la salida en capas de neuronas conectadas hacia adelante. MLP aprende a partir de un conjunto de datos de entrenamiento para obtener una respuesta de salida; esto es, se define una función

$f : \mathbb{R}^p \rightarrow \mathbb{R}^c$ , donde  $p$  es la dimensión de la entrada  $x$  y  $c$  es el número de clases objetivo. El proceso de propagación hacia adelante en MLP con una capa de entrada, una capa oculta con  $nn$  neuronas y una capa de salida; consiste en dados una entrada  $x$  y un vector de pesos  $w_j$  asociado con la neurona  $j$  ( $j = 1, 2, \dots, nn$ ), la capa oculta transforma la entrada  $x$  como:

$$\mathbf{z} = \mathbf{Wx}$$

donde  $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{nn}]$ , y  $\mathbf{z} = [z_1, z_2, \dots, z_{nn}]$ . El vector  $\mathbf{z}$  se transforma por medio de una función de activación:

$$\begin{bmatrix} 1 \\ f(\mathbf{z}) \end{bmatrix}$$

donde  $f(\mathbf{z})$  es una función que se evalúa para cada componente del vector  $z$  más un término de sesgo, y el vector de clases objetivo predichas  $\hat{\mathbf{y}} \in \mathbb{R}^c$ . Las funciones de activación más comunes son *relu*, que se define como  $f(\mathbf{z})_{\text{relu}} = \max(0, \mathbf{z})$  (Nair y Hinton, 2010) y la función tangente hiperbólica (*tanh*) que se define como  $f(\mathbf{z})_{\text{tanh}} = a \tanh(bz/2)$ , donde  $a$  y  $b$  son parámetros de control.

El algoritmo de propagación del error hacia atrás se utilizó para obtener los parámetros óptimos de MLP (Rumelhart et al., 1986). Este algoritmo utiliza el gradiente descendente, o variantes de éste para minimizar la función de costo  $J(\mathbf{W})$ :

$$J(\mathbf{W}) = \frac{1}{n} \left( \sum_{i=1}^n \mathcal{L}_w(y_i, \hat{y}_i) + \lambda \sum_j w_j^2 \right)$$

donde  $n$  es el número de muestras,  $y_i$  es la clase objetivo que corresponde a la muestra  $i$ ,  $\hat{y}_i$  es la clase objetivo predicha para la muestra  $i$ ,  $\mathcal{L}_w$  es la función de costo que se define como entropía cruzada (*cross-entropy*). En el caso de una clasificación multiclase  $\mathcal{L}_w$  se define como  $\mathcal{L}_w = \sum_{i=1}^c -\ln(\hat{y}_i)$  donde  $c$  es el número de clases,  $\mathbf{W}$  es la matriz de pesos de la capa oculta;  $\lambda$  es un hiperparámetro de ajuste con un valor inicial de 0.001,  $\lambda \sum_j w_j^2$  es el término de regularización  $\mathcal{L}_2$  y  $w_j$  es el vector de pesos  $j$  de  $\mathbf{W}$ .

Los métodos de optimización que se utilizaron fueron gradiente descendente estocástico (*sgd*), *adam* (*adaptive moment estimation*) (Kingma y Ba, 2014), y *l-bfgs* (*limited-memory Broyden-Fletcher-Goldfarb-Shanno*).

### Métricas de desempeño de los modelos de aprendizaje

Las métricas de desempeño fueron la precisión global de clasificación (*Acc, accuracy*), puntaje promedio global *F1 macro*, precisión (*P*), sensibilidad (*S*), puntaje *F1* y área bajo la curva *ROC* (*AUC*).

$Acc$  es la proporción de muestras que se clasifican correctamente,  $Acc = (TP + TN)/N$ , donde  $TP$  es el total de muestras predichas como positivas que son positivas  $TN$  es el total de muestras predichas como negativas que son negativas,  $N$  es el número total de muestras que se predicen.

$F1$  macro es el promedio de  $F1$  para cada clase,  $F1_{macro} = 1/c \sum_{i=1}^c F1_i$ , donde  $c$  es el número total de clases,  $F1_i$  es el valor de  $F1$  para la clase  $i$ .

$P$  se define como  $P = TP / (TP + FP)$ , donde  $FP$  (falsos positivos) es el total de muestras predichas como positivas que son negativas.

$S$  se define como  $S = TP / (TP + FN)$ , donde  $FN$  son los falsos negativos (muestras positivas predichas como negativas).

$F1$  se define como  $F1 = 2PS / (P + S)$ ,  $F1$  toma valores entre 0 y 1, valores cercanos a 1 indican mejor desempeño en  $P$  y  $S$ , y cercanos a 0 indican pobre desempeño en  $P$  y  $S$ .

AUC es el área bajo la curva ROC, AUC indica el desempeño del modelo para identificar cada una de las clases objetivo y toma valores entre 0 y 1 (Fawcett, 2006).

### Entrenamiento de los modelos de aprendizaje

#### Selección de hiperparámetros óptimos

Previo a la selección de hiperparámetros, los conjuntos de entrenamiento se estandarizaron con media  $\mu = 0$  y desviación estándar  $\sigma = 1$  (Obaid et al., 2019). La selección de hiperparámetros óptimos de cada modelo se realizó por medio de validación cruzada (VC) y búsqueda por retícula en el conjunto de entrenamiento. Para cada

combinación de hiperparámetros (Cuadro 1), VC consiste en crear  $k$  particiones aleatorias disjuntas del conjunto de entrenamiento; cada conjunto de entrenamiento consiste en  $k - 1$  particiones y cada conjunto de prueba se crea con la partición restante. Para cada conjunto de entrenamiento y prueba se obtiene  $Acc$ , luego se obtiene el promedio de los  $k$  valores de  $Acc$ , al término de la búsqueda se selecciona la combinación de hiperparámetros que genera el máximo desempeño del modelo (máximo  $Acc$  promedio). El entrenamiento de los modelos se realizó con una partición aleatoria estratificada por clase objetivo (80 % del conjunto de datos original) y la búsqueda de hiperparámetros se realizó con VC ( $k = 5$ ).

### Evaluación del desempeño de los modelos de aprendizaje

La evaluación del desempeño en predicción de cada modelo se realizó con los hiperparámetros óptimos VC ( $k = 10$ ) con 100 % de los datos. Para cada ejecución  $k$ , se calcularon las métricas de desempeño  $Acc$ ,  $P$ ,  $S$  y  $F1$ . Para la ejecución  $k = 10$  se obtuvieron AUC de la curva ROC y la curva de validación del modelo con el mejor desempeño.

## RESULTADOS Y DISCUSIÓN

### Hiperparámetros óptimos seleccionados

Los descriptores más relevantes  $x_h$ ,  $x_s$ ,  $x_v$ ,  $x_r$  y  $x_e$  se utilizaron en el entrenamiento y prueba de los modelos para seleccionar los hiperparámetros óptimos de SVM, RF, y MLP (Cuadro 2).

### Desempeño en predicción de los modelos

SVM, RF y MLP alcanzaron una  $Acc$  de 98.7, 98.6 y 97 %, respectivamente (Cuadro 3). La prueba de Wilcoxon

**Cuadro 1. Hiperparámetros e intervalos de búsqueda por retícula de los modelos máquina de soporte vectorial (SVM), bosque aleatorio (RF) y perceptrón multicapa (MLP).**

Modelo	Hiperparámetros <sup>†</sup>		
SVM	$C$	$\gamma$	$K$
	[0.0001, 0.001, 0.01, 0.1, 1.0, 10.0, 100.0]	[0.0001, 0.001, 0.01, 0.1, 1.0, 10.0, 100.0, 1000.0]	Sigmoide y Gaussiano
RF	$nar$	$pmax$	$crit$
	[50, 100, 150, 200]	[10, 15, 20]	Gini y Entropía
MLP	$fa$	$opt$	$nn$
	$relu$ y $tanh$	$sgd$ , $adam$ , $l-bfgs$	[100, 150, 200]
			[0.001, 0.5, 1]

<sup>†</sup>y: hiperparámetro gamma,  $K$ : kernel,  $nar$ : número de árboles,  $pmax$ : profundidad máxima,  $crit$ : criterio de impureza,  $fa$ : función de activación,  $opt$ : optimizador,  $nn$ : número de neuronas,  $a$ : hiperparámetro alfa.

**Cuadro 2. Hiperparámetros óptimos de los modelos máquina de soporte vectorial (SVM), bosque aleatorio (RF) y perceptrón multicapa (MLP).**

Modelo	Hiperparámetros <sup>†</sup>		
SVM	C	$\gamma$	K
	100.0	1.0	rbf
RF	<i>nar</i>	<i>pmax</i>	<i>crit</i>
	100	15	entropy
MLP	<i>fa</i>	<i>a</i>	<i>opt</i>
	<i>relu</i>	0.001	<i>adam</i>
			<i>nn</i>
			150

<sup>†</sup>y: hiperparámetro gamma, K: kernel, *nar*: número de árboles, *pmax*: profundidad máxima, *crit*: criterio de impureza, *fa*: función de activación, *a*: hiperparámetro alfa, *opt*: optimizador, *nn*: número de neuronas.

se aplicó para comparar Acc por pares de modelos. Los resultados fueron: SVM vs. RF, estadístico W = 45, p-value = 0.733; SVM vs. MLP, W = 100, p-value = 0.0001736; RF vs. MLP, W = 94, p-value = 0.0009773. SVM y RF no presentaron diferencias significativas y fueron superiores a MLP.

El desempeño a nivel de clases objetivo, SVM obtuvo un *F1* de 99.2 % para clasificar V1 (variedad CP-03-06) y *F1* de 96.1 % para clasificar V4 (variedad Cardenal). F1 proporciona una estimación más conservadora que AUC debido al desbalance entre clases objetivo (Cuadro 4).

De acuerdo con la matriz de confusión, SVM obtuvo las tasas más bajas de FN para clasificar las variedades V1 y V5, y las tasas más altas de FN para clasificar las variedades V2, V3 y V4. SVM alcanzó mayor Acc para clasificar V1 y menor Acc para clasificar V3; sin embargo, se observa un alto desempeño de SVM para clasificar las seis variedades (Figura 2).

Los modelos SVM, RF y MLP obtuvieron un Acc superior a 97 %. MLP es un buen aproximador no lineal; sin embargo, en este estudio el espacio de búsqueda de los hiperparámetros (número de capas ocultas, número de neuronas) se limitó debido a los tiempos de cómputo que se requieren para su optimización (Basha *et al.*, 2020).

Los descriptores morfológicos  $x_r$  y  $x_e$  son invariantes a

las transformaciones aplicadas a una imagen y pueden tener un uso práctico (Gómez-Flores y Hernández-López, 2020), a diferencia de descriptores como área, perímetro o ancho, que no son invariantes a transformaciones (Bojamma y Shastry, 2021).

Yanikoglu *et al.* (2014) identificaron plantas del ImageCLEF'2012; (Goëau *et al.*, 2013) y exploraron el uso de otros órganos vegetales como corteza, flores y tallos para identificar plantas con base en descriptores cromáticos y de textura. (Ayala-Niño y González-Camacho, 2022) aplicaron modelos de aprendizaje automático para identificar cuatro variedades de durazno con base únicamente en descriptores cromáticos HSV, estos autores reportaron un desempeño máximo de Acc de 84.1 % con SVM. La similitud de los tonos de verde en el haz de las hojas inducen mayor homogeneidad entre clases objetivo, lo cual limita alcanzar altos valores de clasificación; asimismo, Shrivastava y Pradhan (2021) aplicaron diferentes modelos de aprendizaje para clasificar cuatro enfermedades en arroz con base en 172 descriptores cromáticos (14 espacios de color), donde SVM fue el mejor y obtuvo un Acc de 94.65 %.

Los valores cromáticos del espacio de color HSV, la excentricidad y redondez del haz de las hojas fueron suficientes para identificar con alta precisión global las seis variedades de durazno. Los scripts de los algoritmos de aprendizaje y conjuntos de datos utilizados en esta

**Cuadro 3. Desempeño en predicción de los modelos máquina de soporte vectorial (SVM), bosque aleatorio (RF) y perceptrón multicapa (MLP).**

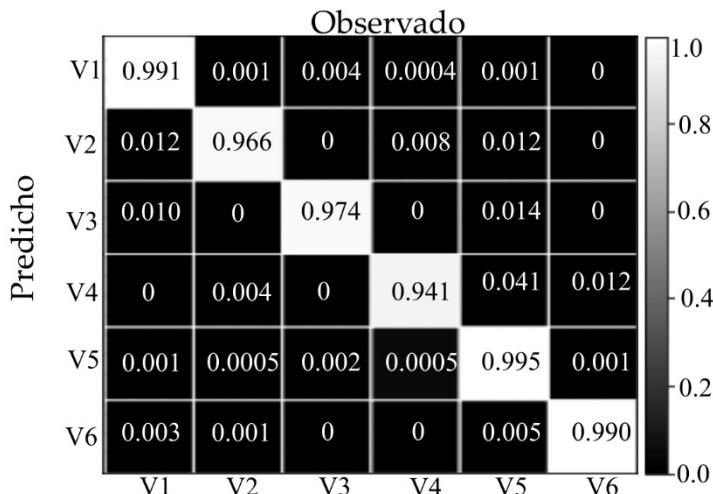
Modelos	Acc <sup>†</sup>	<i>F1 macro</i> <sup>††</sup>
SVM	0.987 ± 0.004	0.980 ± 0.006
RF <sup>‡</sup>	0.986 ± 0.008	0.979 ± 0.009
MLP	0.970 ± 0.004	0.958 ± 0.008

<sup>†</sup>Acc: precisión global de clasificación correcta, <sup>††</sup>*F1 macro*: promedio de *F1*.

**Cuadro 4.** Métricas de desempeño promedio en predicción del modelo máquina de soporte vectorial (SVM).

Clases <sup>†</sup>	P <sup>††</sup>	S <sup>‡</sup>	F1 <sup>¶</sup>	AUC <sup>×</sup>
V1	0.992 ± 0.004	0.991 ± 0.007	0.992 ± 0.005	0.999 ± 0.001
V2	0.975 ± 0.027	0.966 ± 0.031	0.970 ± 0.023	0.992 ± 0.012
V3	0.980 ± 0.015	0.974 ± 0.012	0.977 ± 0.011	0.996 ± 0.006
V4	0.982 ± 0.020	0.941 ± 0.042	0.961 ± 0.025	0.992 ± 0.016
V5	0.981 ± 0.008	0.993 ± 0.004	0.987 ± 0.004	0.998 ± 0.001
V6	0.991 ± 0.010	0.990 ± 0.016	0.990 ± 0.010	0.999 ± 0.001
macro			0.980 ± 0.006	0.996 ± 0.008

<sup>†</sup>V1: CP-03-06, V2: Oro Azteca, V3: Oro San Juan, V4: Cardenal, V5: Colegio, V6: Robin, <sup>††</sup>P: precisión, <sup>‡</sup>S: sensibilidad, <sup>¶</sup>F1: F1, <sup>×</sup>AUC: área bajo la curva ROC.



**Figura 2.** Matriz de confusión normalizada acumulada del modelo máquina de soporte vectorial (SVM) evaluado en predicción con validación cruzada ( $k = 10$ ); para identificar las variedades de durazno: V1: CP-03-06, V2: Oro Azteca, V3: Oro San Juan, V4: Cardenal, V5: Colegio, V6: Robin.

investigación se encuentran disponibles en el enlace <https://github.com/DanielResearch/Identificacion-de-Durazno>

## CONCLUSIONES

Los modelos máquina de soporte vectorial (SVM), bosque aleatorio (RF) y red neuronal perceptrón multicapa (MLP) incrementaron su capacidad predictiva para identificar variedades de durazno. El modelo SVM obtuvo una precisión global de clasificación correcta (Acc) de 98.7%; la variedad CP-03-06 se identificó con la mayor precisión de 99.2 %; y la variedad Cardenal con la menor de 96.1 %.

## AGRADECIMIENTOS

Al Consejo Mexiquense de Ciencia y Tecnología por la beca otorgada al primer autor para realizar esta investigación con código EESP2021-007.

## BIBLIOGRAFÍA

- Aakif A. and M. F. Khan (2015) Automatic classification of plants based on their leaves. *Biosystems Engineering* 139:66–75, <https://doi.org/10.1016/j.biosystemseng.2015.08.003>
- Apriyanti D. H., A. M. Arymurthy and L. T. Handoko (2013) Identification of orchid species using content-based flower image retrieval: In: 2013 International Conference on Computer, Control, Informatics and Its Applications. Institute of Electrical and Electronics Engineers. Jakarta, Indonesia. pp:53–57, <https://doi.org/10.1109/IC3INA.2013.6819148>
- Ayala N. D., J. Cervantes C., F. García L., J. Ayala V. and G. Calderón Z. (2020) A hybrid convolutional neural network for complex leaves identification: In: Intelligent Computing Theories and Application. Lecture Notes in Computer Science. Vol. 12463. D. S. Huang, V. Bevilacqua and A. Hussain (eds.). Springer. Cham, Switzerland. pp:291–301, [https://doi.org/10.1007/978-3-030-60799-9\\_25](https://doi.org/10.1007/978-3-030-60799-9_25)
- Ayala-Niño D. and J. M. González-Camacho (2022) Evaluation of machine learning models to identify peach varieties based on leaf color. *Agrociencia* 56:1-17, <https://doi.org/10.47163/agrociencia.v56i1.2810>
- Barré P., B. C. Stöver, K. F. Müller and V. Steinhage (2017) LeafNet: a computer vision system for automatic plant species identification. *Ecological Informatics* 40:50–56, <https://doi.org/10.1016/j.ecoinf.2017.05.005>

- Basha S. H. S., S. R. Dubey, V. Pulabaigari and S. Mukherjee (2020) Impact of fully connected layers on performance of convolutional neural networks for image classification. *Neurocomputing* 378:112-119, <https://doi.org/10.1016/j.neucom.2019.10.008>
- Biau G. and E. Scornet (2016) A random forest guided tour. *TEST* 25:197-227, <https://doi.org/10.1007/s11749-016-0481-7>
- Bojamma A. M. and C. Shastry (2021) A study on the machine learning techniques for automated plant species identification: current trends and challenges. *International Journal of Information Technology* 13:989-995, <https://doi.org/10.1007/s41870-019-00379-7>
- Calderón-Zavala G., J. Rodríguez-Alcazar, M. C. Espíndola-Barquera y A. García-Ávila (2019) Variedades mejoradas de durazno (*Prunus persica* L.). *Agroproductividad* 12:81-83, <https://doi.org/10.32854/agrop.v12i9.1492>
- Chauhan V. K., K. Dahiya and A. Sharma (2019) Problem formulations and solvers in linear SVM: a review. *Artificial Intelligence Review* 52:803-805, <https://doi.org/10.1007/s10462-018-9614-6>
- Fawcett T. (2006) An introduction to ROC analysis. *Pattern Recognition Letters* 27:861-874, <https://doi.org/10.1016/j.patrec.2005.10.010>
- García F., J. Cervantes, A. López and M. Alvarado (2016) Fruit classification by extracting color chromaticity, shape and texture features: towards an application for supermarkets. *IEEE Latin America Transactions* 14:3434-3443. <https://doi.org/10.1109/TLA.2016.7587652>
- García-Mateos G., J. L. Hernández-Hernández, D. Escarabajal-Henarejos, S. Jaén-Terrones and J. M. Molina-Martínez (2015) Study and comparison of color models for automatic image analysis in irrigation management applications. *Agricultural Water Management* 151:158-166, <https://doi.org/10.1016/j.agwat.2014.08.010>
- Goëau H., A. Joly, P. Bonnet, V. Bakic, D. Barthélémy, N. Boujemaa and J. F. Molino (2013) The imageCLEF plant identification task 2013. Proceedings of the 2nd ACM International Workshop on Multimedia Analysis for Ecological Data, October 22, 2013. MAED@ACM Multimedia. Barcelona, Spain. pp:23-28, <https://doi.org/10.1145/2509896.2509902>
- Gómez-Flores W. and J. Hernández-López (2020) Assessment of the invariance and discriminant power of morphological features under geometric transformations for breast tumor classification. *Computer Methods and Programs in Biomedicine* 185:105173, <https://doi.org/10.1016/j.cmpb.2019.105173>
- Hu M. K. (1962) Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory* 8:179-187, <https://doi.org/10.1109/TIT.1962.1057692>
- Kingma D. P. and L. J. Ba (2014) Adam: a method for stochastic optimization. *arXiv* 1412.6980v9, <https://doi.org/10.48550/arXiv.1412.6980>
- Kumar A. and S. S. Sodhi (2020) Comparative analysis of Gaussian filter, median filter and denoise autoencoder. 2020 7th International Conference on Computing for Sustainable Global Development. New Dheli, India. pp:45-51, <https://doi.org/10.23919/INDIACom49435.2020.9083712>
- Lee S. H., C. S. Chan, S. J. Mayo and P. Remagnino (2017) How deep learning extracts and learns leaf features for plant classification. *Pattern Recognition* 71:1-13, <https://doi.org/10.1016/j.patcog.2017.05.015>
- Li Y., J. Nie and X. Chao (2020) Do we really need deep CNN for plant diseases identification? *Computers and Electronics in Agriculture* 178:105803, <https://doi.org/10.1016/j.compag.2020.105803>
- Marr D. and E. Hildreth (1980) Theory of edge detection. *Proceedings of the Royal Society of London. Series B. Biological Sciences* 207:187-217, <https://doi.org/10.1098/rspb.1980.0020>
- Nair V. and G. E. Hinton (2010) Rectified linear units improve restricted Boltzmann machines. *Proceedings of the 27th International Conference on Machine Learning*, 21 June 2010. Haifa, Israel. pp:807-814.
- Obaid H. S., S. A. Dheyab and S. S. Sabry (2019) The impact of data pre-processing techniques and dimensionality reduction on the accuracy of machine learning. 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference. Jaipur, India. pp:279-283, <https://doi.org/10.1109/IMECONX.2019.8877011>
- Otsu N. (1979) A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics* 9:62-66, <https://doi.org/10.1109/TSMC.1979.4310076>
- Rumelhart D., G. Hinton and R. Williams (1986) Learning representations by back-propagating errors. *Nature* 323:533-536, <https://doi.org/10.1038/323533a0>
- Septiarini A., A. Sunyoto, H. Hamdani, A. A. Kasim, F. Utaminingrum and H. R. Hatta (2021) Machine vision for the maturity classification of oil palm fresh fruit bunches based on color and texture features. *Scientia Horticulturae* 286:110245, <https://doi.org/10.1016/j.scienta.2021.110245>
- Shrivastava V. K. and M. K. Pradhan (2021) Rice plant disease classification using color features: a machine learning paradigm. *Journal of Plant Pathology* 103:17-26, <https://doi.org/10.1007/s42161-020-00683-3>
- SIAP, Servicio de Información Agroalimentaria y Pesquera (2023) Anuario estadístico de la producción agrícola. Servicio de Información Agroalimentaria y Pesquera. Ciudad de México. <https://nube.siap.gob.mx/cierreagrico/> (Enero 2023).
- Simonyan K. and A. Zisserman (2015) Very deep convolutional networks for large-scale image recognition. 3rd International Conference on Learning Representations. May 7-9 2015. San Diego, California, USA. pp:1-14, <https://doi.org/10.48550/arXiv.1409.1556>
- Suzuki S. and K. be (1985) Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing* 30:32-46, [https://doi.org/10.1016/0734-189X\(85\)90016-7](https://doi.org/10.1016/0734-189X(85)90016-7)
- Wäldchen J. and P. Mäder (2018) Plant species identification using computer vision techniques: a systematic literature review. *Archives of Computational Methods in Engineering* 25:507-543, <https://doi.org/10.1007/s11831-016-9206-z>
- Xiao F., H. Wang, Y. Li, Y. Cao, X. Lv and G. Xu (2023) Object detection and recognition techniques based on digital image processing and traditional machine learning for fruit and vegetable harvesting robots: an overview and review. *Agronomy* 13:639, <https://doi.org/10.3390/agronomy13030639>
- Yanikoglu B., E. Aptoula and C. Tirkaz (2014) Automatic plant identification from photographs. *Machine Vision and Applications* 25:1369-1383, <https://doi.org/10.1007/s00138-014-0612-7>