

QUALITY ASSESSMENT OF POTTED PETUNIA BASED ON A PROBABILISTIC NEURAL NETWORK CLASSIFIER

EVALUACIÓN DE LA CALIDAD DE PLANTAS DE PETUNIA EN MACETA CON BASE EN UN CLASIFICADOR DE RED NEURONAL PROBABILÍSTICA

José A. Carrillo-Salazar, Juan M. González-Camacho*, Erwin San Juan-Martínez, Víctor A. González-Hernández

Campus Montecillo, Colegio de Postgraduados. 56230. Montecillo, Estado de México.
(jmgc@colpos.mx)

ABSTRACT

Quality of ornamental plants relies on foliage and flower appearance and evaluation of such characteristics depends on people experience; therefore, more reliable scoring methods are needed. The objective of this research was to explore the usefulness of digital image segmentation with a probabilistic neural network (PNN) classifier from which quality indicators can be obtained for quantifying time-dependent ornamental characteristics. A petunia variety (*Petunia multiflora* cv. F1 'Glistening Pearls mixed') grown in pots in the Central High Valleys of Mexico was used as model. The variety includes a mixture of genotypes with flowers in many colors. The PNN classifier was trained and tested on six color classes (four flower classes, one foliage class and one background class). From image segmentation, six quality indicators were calculated. For the test sets, the best classification scenario whose inputs were a combination of some of the channels of RGB, CIE-Lab, CIE-L*uv, HSV, and CMY color models reached an overall accuracy (OA) of 98% while it was 96% when using the RGB color model alone. The quality indicators identified differences on trends of plant quality associated to flower color. White-flowered plants had larger ground cover per plant and greater ratio of flower area to total plant area, as well as a lower roundness index. Pink-flowered plants had the lowest ground cover per plant and plant perimeter, while they had the largest roundness index. This approach allowed to identify different behaviors of the genotypes based on quality indicators, and it might be used for estimating the period this species keeps its exhibition quality.

Key words: Radial basis function neural networks, image analysis, *Petunia multiflora*, quality indicator, pattern recognition.

* Author for correspondence ♦ Autor responsable.

Received: abril, 2018. Approved: agosto, 2018.

Published as ARTICLE in *Agrociencia* 53: 895-910. 2019.

RESUMEN

La calidad de las plantas ornamentales se basa en el follaje y el aspecto de las flores, y la evaluación de tales características depende de la experiencia de las personas; por lo tanto, se necesitan métodos de calificación más confiables. El objetivo de esta investigación fue explorar la utilidad de la segmentación de imágenes digitales con un clasificador de red neuronal probabilística (PNN), a partir del cual se pueden obtener indicadores de calidad para cuantificar las características ornamentales dependientes del tiempo. Como modelo se usó una variedad de petunia (*Petunia multiflora* cv. F1 'Glistening Pearls mixed') cultivada en macetas en el Valle de México. La variedad incluye una mezcla de genotipos con flores de varios colores. El clasificador PNN fue entrenado y probado con seis clases de color (cuatro clases de flor, una clase de follaje y una clase de fondo). A partir de la segmentación de imágenes, se calcularon seis indicadores de calidad. Para los conjuntos de prueba, el mejor escenario de clasificación, cuyas entradas fueron una combinación de algunos canales de los modelos de color RGB, CIE-Lab, CIE-L*uv, HSV y CMY, alcanzó una precisión global (OA) de 98%; mientras que ésta fue de 96% cuando se utilizó el modelo de color RGB por sí solo. Los indicadores de calidad identificaron diferencias en las tendencias de la calidad de las plantas asociadas al color de la flor. Las plantas de flores blancas tuvieron valores mayores de cobertura de suelo por planta y área de flores sobre el área total de planta, así como un índice de redondez menor. Las plantas de flores rosadas tuvieron los valores menores de cobertura de suelo por planta y perímetro de planta, pero tuvieron el índice de redondez mayor. Este enfoque permitió identificar comportamientos diferentes de los genotipos basados en indicadores de calidad, y podría utilizarse para estimar el período en el cual esta especie mantiene su calidad de exhibición.

Palabras clave: redes neuronales con funicones de base radial, análisis de imágenes, *Petunia multiflora*, indicador de calidad, reconocimiento de patrones.

INTRODUCTION

Quality of ornamental plants is commonly scored on foliage and flower appearance. Evaluation of quality is difficult when large number of plant structures is processed, since it requires intense plant manipulation. Additionally, the classification process inconveniently relies on personal expertise and varies from one person to another. Therefore, classification method can be imprecise and ambiguous. In this context, it is useful to develop quantitative methods for evaluating ornamental plant quality for breeding programs and for finding critical points in the production process (Parsons *et al.*, 2009).

Image analysis avoids the physical manipulation of plants and allows to develop classification strategies. A pixel is the basic unit composing any digital image. Each pixel can therefore be grouped with classification models that are able to segment areas of interest, flowers, foliage or any other structure.

Artificial neuronal networks (ANN) are used as classification algorithms to segment images (Littmann and Ritter, 1997; Kavzoglu and Mather, 2003; Parsons *et al.*, 2009). In turn, a segmented image can be used for evaluating agronomic variables such as plant size, number of flowers, roundness index, color of flowers and foliage, ground cover per plant, and plant perimeter. These variables can then be associated to physiological variables or used as plant quality indicators.

Hemming and Rath (2001) identified weeds in cabbage (*Brassica oleracea*) and carrot (*Daucus carota*) crops using leaf characteristics detectable in images such as area, area/contour length ratio, length/width ratio, circularity, convexity, maximum diameter, roundness and spikes. Granitto *et al.* (2002) identified 57 weed species based on image analysis of seeds size and shape (length of the main axis and several moments of the planar mass distribution with respect to that axe), texture (contrast along the main axis) and color (variation of intensity). Fruit number and diameter of apple (*Malus pumila*) in an orchard were also estimated by using different image processing algorithms (Stajniko *et al.*, 2004).

The ANN are also used to create models that associate visual traits obtained from digital images with morphological measured variables like area,

INTRODUCCIÓN

La calidad de las plantas ornamentales se califica comúnmente con base en el aspecto del follaje y de las flores. La evaluación de la calidad es difícil cuando se procesa una gran cantidad de estructuras de las plantas, ya que ello requiere una intensa manipulación del proceso. Además, un inconveniente es que el proceso de clasificación depende de la experiencia personal de un evaluador y varía de una persona a otra. Por lo tanto, el método de clasificación puede ser impreciso y ambiguo. En este contexto, es útil desarrollar métodos cuantitativos para evaluar la calidad de las plantas ornamentales para programas de mejoramiento y para identificar puntos críticos en el proceso de producción (Parsons *et al.*, 2009).

El análisis de imágenes evita la manipulación física de las plantas y permite desarrollar estrategias de clasificación. Un píxel es la unidad básica que compone cualquier imagen digital. Por lo tanto, cada píxel puede agruparse con modelos de clasificación para segmentar áreas de interés, flores, follaje o cualquier otra estructura.

Las redes neuronales artificiales (ANN) se utilizan como algoritmos de clasificación para segmentar imágenes (Littmann y Ritter, 1997; Kavzoglu y Mather, 2003; Parsons *et al.*, 2009). A su vez, se puede utilizar una imagen segmentada para evaluar variables agronómicas, tales como tamaño de la planta, número de flores, índice de redondez, color de las flores y follaje, cobertura del suelo por planta y perímetro de la planta. Estas variables pueden asociarse a variables fisiológicas o usarse como indicadores de calidad de la planta.

Hemming y Rath (2001) identificaron arvenses en los cultivos de col (*Brassica oleracea*) y de zanahoria (*Daucus carota*) con el uso de características foliares que son detectables en imágenes, tales como el área, la relación entre el área y la longitud del contorno, relación longitud/ anchura, circularidad, convexidad, diámetro máximo, redondez y espículas. Granitto *et al.* (2002) identificaron 57 especies arvenses con base en el análisis de imágenes del tamaño y forma de semillas (longitud del eje principal y varios momentos del plano de la distribución de masa con respecto a ese eje), textura (contraste a lo largo del eje principal) y color (variación de intensidad). También se han utilizado diferentes algoritmos de procesamiento de

height, width, root length, fresh biomass and dry biomass in seedlings of *Phalaenopsis* sp. var 'Sogo Vivien F819' (Hsieh and Weng, 2005). Parsons *et al.* (2009) used ANN to segment images of greenhouse-grown ornamental plants and for extracting quantitative information of leaf and flower coverage, uniformity of coverage of leaves and flowers, foliage height and shape, and leaf and flower color. Flowering progress and number of flowers in *Lesquerella fendleri* were monitored by Thorp and Dierig (2011).

Petunia (Petunia multiflora) is an herbaceous and attractive plant with large flowers of many colors, across a long flowering period. This plant species has commercial hybrids with different shapes and sizes adequate for pots, ground covering or hanging baskets (García-Albarado *et al.*, 2010). In 2015, this crop represented 44 million dollars for 15 states in the United States (USDA, 2016), whereas in 2015 the income in Mexico from growing pot petunias was around 15.5 million pesos (approx. 860 000 dollars) (SIAP, 2016).

The objective of this research was to evaluate time-dependent changes in quality indicators in petunia plants grown in pots, using differently colored flowers during the blooming period. These quality indicators were obtained from digital images segmented, by a probabilistic neural network (PNN) classifier, in three regions depending on flower color classes (four classes), leaf (one class), and background (one class).

MATERIALS AND METHODS

Crop conditions

Plants of *P. multiflora* F1 'Glistering Pearls mixed' were grown in a greenhouse (19° 29' N and 98° 54' W, 2240 m) in the Central High Valleys of Mexico, covered with light diffuser polyethylene film during spring. Seeds were sown on March 30 in 100-pit plastic trays. At 29 d after sowing, seedlings were transplanted into a 15×15 cm black polyethylene container filled with grounded volcanic stone as substrate. Seedlings were daily fertilized with 50% of the salt-strength Steiner solution; every week, plain water was used to wash salts. At flowering, plants were separated by flower color into four groups or treatments (blue, white, fuchsia and pink, as shown in Figure 1). A completely random experimental design was set up with 4 treatments with 5 replications.

imágenes para estimar en un huerto, el número de frutas y diámetro de manzana (*Malus pumila*) (Stajanko *et al.*, 2004).

Las ANN también se utilizan para crear modelos que asocian rasgos visuales obtenidos de imágenes digitales con medidas de variables morfológicas, como área, altura, ancho, longitud de raíz, biomasa fresca y biomasa seca en plántulas de *Phalaenopsis* sp. var 'Sogo Vivien F819' (Hsieh y Weng, 2005). Parsons *et al.* (2009) aplicaron ANN para segmentar imágenes de plantas ornamentales cultivadas en invernadero y extraer información cuantitativa de cobertura de hojas y flores, uniformidad de cobertura de hojas y flores, altura y forma del follaje y color de hojas y flores. Thorp y Dierig (2011) monitorearon el desarrollo de la floración y el número de flores en *Lesquerella fendleri*.

La petunia (*Petunia multiflora*) es una planta herbácea atractiva, con flores grandes de varios colores, y con un período de floración largo. Esta especie tiene híbridos comerciales, con formas diferentes y tamaños adecuados para macetas, jardines o cestas colgantes (García-Albarado *et al.*, 2010). En 2015, este cultivo representó una entrada de 44 millones de dólares para 15 estados en Estados Unidos (USDA, 2016); mientras que, en ese mismo año, el ingreso para México por el cultivo de petunias en maceta fue cercano a 15.5 millones de pesos (aproximadamente 860 000 dólares) (SIAP, 2016).

El objetivo de esta investigación fue evaluar cambios dependientes del tiempo, en los indicadores de calidad, en las plantas de petunia cultivadas en macetas, con el uso de genotipos con flores de diferentes colores, durante el período de floración. Estos indicadores de calidad se obtuvieron a partir de imágenes digitales segmentadas por medio de un clasificador de red neuronal probabilística (PNN) en tres regiones, de acuerdo con las clases de color de las flores (cuatro clases), hoja (una clase) y fondo (una clase).

MATERIALES Y MÉTODOS

Condiciones de cultivo

Las plantas de *P. multiflora* F1 'Glistering Pearls mixed' se cultivaron durante la primavera en un invernadero cubierto con polietileno difusor de luz en el Valle de México (19° 29' N y 98° 54' O, 2240 m). Las semillas se sembraron el 30 de marzo en bandejas de plástico con 100 cavidades. A los 29 d después de

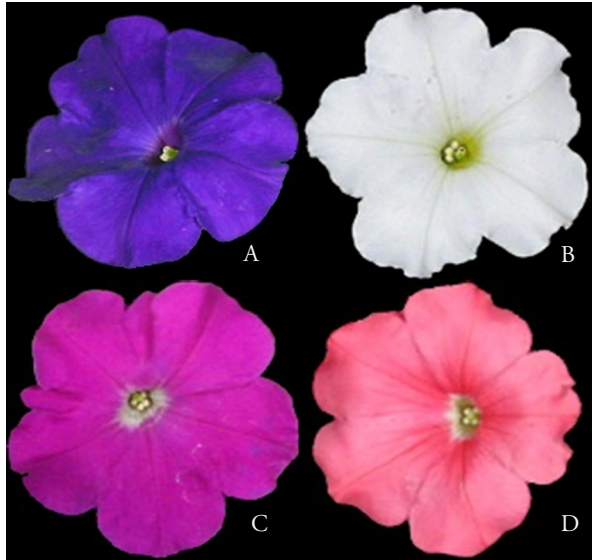


Figure 1. Color classes for petunia flowers: A) blue, B) white, C) fuchsia, and D) pink.

Figura 1. Clases de color en las flores de petunia: A) azul, B) blanca, C) fucsia y D) rosada.

Image analysis

For image acquisition, plants were placed inside a 1 m³ wooden box lined with black velvet, which homogenized brightness intensity in the images. In the top of the box, two 55 W, 60 Hz, General Electric fluorescent lamps were placed for providing 3690 lux lighting. Each image was taken on the horizontal plane using a Canon PowerShot G2 camera, in manual mode, fluorescent white-balance, 1/40 shutter speed and 8.0 aperture. Plants were photographed starting at the beginning of flowering at six sampling dates: 72, 77, 84, 103, 119 and 135 d after transplant (dat); at each sampling date 20 digital images were taken. Images were stored as JPEG format, at 640x480 resolution.

The channel values of the RGB (red, green, blue) color space were obtained from samples of 10x10 pixels of areas of digital images belonging to the four flower colors, background, and leaf (six classes). For this purpose, a program in Visual Basic 6.0 was developed. The RGB refers to a system of coordinates and sub-spaces where each color is represented by one single point (Gonzalez and Woods, 2002). Each channel can take values from 0 to 255 (none to full intensity), and their combination produces an array of more than 16 million of possible color values.

The RGB values for pixels in each image sample were transformed to RGB values as indicated by the Commission Internationale de L'Eclairage through International Electrotechnical Commission standards (IEC61966-2-1)

la siembra, las plántulas se trasplantaron a macetas de polietileno negro de 15x15 cm llenas con tezontle como sustrato. Las plántulas se fertilizaron diariamente con la solución de Steiner al 50%; cada semana se usaba agua corriente para lavar las sales. En la floración, las plantas se separaron por el color de la flor en cuatro grupos o tratamientos (azul, blanco, fucsia y rosado, como se muestra en la Figura 1). Un diseño completamente al azar fue el arreglo experimental con 4 tratamientos y 5 repeticiones.

Análisis de imágenes

Para obtener las imágenes, las plantas se colocaron dentro de una caja de madera de 1 m³, forrada con terciopelo negro, el cual homogeneizó la intensidad del brillo en las imágenes. En la parte superior de la caja, se colocaron dos lámparas fluorescentes General Electric de 55 W, 60 Hz para proporcionar 3690 lux de iluminación. Cada imagen se tomó en plano horizontal con una cámara Canon PowerShot G2, en modo manual, con balance de blancos fluorescentes, velocidad del obturador de 1/40 y apertura de 8.0. Las plantas se fotografiaron a partir del inicio de la floración en seis fechas de muestreo: 72, 77, 84, 103, 119 y 135 d después del trasplante (dat); en cada fecha de muestreo se tomaron 20 imágenes digitales. Las imágenes se almacenaron en formato JPEG, con una resolución de 640x480.

Los valores de los canales del espacio de color RGB (rojo, verde, azul) se obtuvieron a partir de muestras de área de 10x10 píxeles de las imágenes digitales pertenecientes a los cuatro colores de flores, al fondo y a la hoja (seis clases). Para ello, se desarrolló un programa en Visual Basic 6.0. El RGB se refiere a un sistema de coordenadas y subespacios donde cada color se representa por un solo punto (González y Woods, 2002). Cada canal puede tener valores de 0 a 255 (de intensidad nula a máxima) y su combinación produce una matriz de más de 16 millones de valores posibles de color.

Los valores RGB de los píxeles en cada muestra de imagen se transformaron en valores RGB según lo indicado por la Comisión Internacional de la Iluminación (o Commission Internationale de L'Eclairage, en francés) con los estándares de la Comisión Electrotécnica Internacional (International Electrotechnical Commission, IEC61966-2-1) (Mendoza *et al.*, 2006). Estos valores se convirtieron al modelo CIE XYZ (Poynton, 1996) como un paso previo para obtener los valores de los canales de color para los espacios de color CIE-Lab (Robertson, 1977), Hunter-Lab, CIE-L *uv (CIE Colorimetry, 1986), HSV (Smith, 1978), CMY y CMYK (González y Woods, 2002). Con los resultados de estas conversiones, se creó una base de datos de 1850 registros (píxeles) con 22 variables de entrada que representan los siete espacios de color, y se creó la clase de color objetivo asociada a cada registro (seis clases posibles). El conjunto de datos resultante se usó para entrenar y probar el clasificador PNN (Specht, 1991).

(Mendoza *et al.*, 2006). These values were again converted to the CIE XYZ model (Poynton, 1996) as a previous step to obtain the channel values for the CIE-Lab (Robertson, 1977), Hunter-Lab, CIE-L^auv (CIE Colorimetry, 1986), HSV (Smith, 1978), CMY and CMYK (Gonzalez and Woods, 2002) space colors. As a result of these conversions, a dataset containing 1850 records (pixels) with 22 input variables representing seven color spaces, and the target color class associated to each record (six possible classes) was created. The resulting dataset was used for training and testing the PNN classifier (Specht, 1991).

Probabilistic neural network

The PNN classifier was trained and tested using the computational system, Neuroshell Classifier, from AI Trilogy (Ward Systems Group, Inc.). The dataset generated was randomly divided in two subsets: a training set (90%) and a test set (10%). Using the training set, parameters of the PNN classifier were estimated through the genetic algorithms training strategy (Ward Systems Group, Inc., 1997 - 2007).

The objective function was to minimize the mean percent of incorrect classifications over all categories; optimization stopped after 10 iterations without improvement. This training strategy ends when the PNN classifier reaches its maximum performance for classifying inputs (channel values of the different color spaces) into their target color classes. The input scenarios varied from 3 to 22 variables, according to the color spaces (RGB, CIE-Lab, Hunter-Lab, CIE-Luv, HSV, CMY and CMYK), and the output variable had six possible classes: flower (four colors), foliage or image background. Each input scenario was replicated 10 times with random partitions (90% training, 10% test) from the full dataset. Training time of the PNN classifier for each classification scenario was measured (Zeinali and Story, 2017).

Performance criteria of the PNN classifier

Two criteria were used for assessing the predictive performance of the PNN classifier: the overall accuracy (OA) which is defined as the number of individuals correctly classified into each class divided by the total number of individuals in the test set, and the area under the ROC curve (area under the receiver operating characteristic curve, AUC) which is calculated by integrating the ROC curve defined by plotting true-positives rates (*tpr*) versus false-positives rates (*fpr*) at different threshold levels.

The *tpr* (or sensitivity) for each class is defined as the number of true positives divided by the sum of true positives plus false negatives, and the *fpr* is defined as the number of false positives divided by the sum of false positives plus true negatives. If AUC is close to 1, the PNN classifies better, and it refers to the

Red neuronal probabilística

El clasificador PNN fue entrenado y probado con el sistema computacional Neuroshell Classifier, de AI Trilogy (Ward Systems Group, Inc.). La base de datos generada se dividió aleatoriamente en dos subconjuntos: un conjunto de entrenamiento (90%) y uno de prueba (10%). Con el conjunto de entrenamiento, se estimaron los parámetros del clasificador PNN a través de una estrategia de entrenamiento con algoritmos genéticos (Ward Systems Group, Inc., 1997 - 2007).

La función objetivo fue minimizar el porcentaje promedio de clasificaciones incorrectas de todas las categorías; la optimización se detuvo después de 10 iteraciones sin mejora. Esta estrategia de entrenamiento terminó cuando el clasificador PNN alcanzó su desempeño máximo al clasificar las variables de entrada (los valores de los canales de los diferentes espacios de color) dentro de su clase de color objetivo asociada. Los escenarios de entrada variaron de 3 a 22 variables, según los espacios de color (RGB, CIE-Lab, Hunter-Lab, CIE-Luv, HSV, CMY y CMYK), y la variable respuesta tuvo seis clases posibles: flor (cuatro colores), follaje o fondo de la imagen. Cada escenario de entrada se repitió 10 veces con particiones aleatorias (90% de entrenamiento, 10% de prueba) de la base de datos completa. Para cada escenario de clasificación, se midió el tiempo de entrenamiento del clasificador PNN (Zeinali y Story, 2017).

Criterios de desempeño del clasificador PNN

Dos criterios se utilizaron para evaluar el desempeño del clasificador PNN en la predicción: la precisión global (OA), la cual se define como el número de individuos clasificados correctamente en cada clase, dividido por el número total de individuos en el conjunto de prueba y el área bajo la curva ROC (área bajo la curva operativa característica del receptor, AUC) la cual se calcula por medio de integrar, en diferentes umbrales, el área de la curva ROC definida por el trazo de las tasas de positivos verdaderos (*tpr*) contra las tasas de positivos falsos (*fpr*).

La *tpr* (o sensibilidad) para cada clase se define como el número de positivos verdaderos dividido por la suma de positivos verdaderos más negativos falsos, y *fpr* se define como el número de positivos falsos dividido por la suma de positivos falsos más los negativos verdaderos. Si el valor AUC es cercano a 1, el PNN clasificó mejor y esto refiere a la probabilidad de que el clasificador asigne correctamente un individuo a su clase objetivo (Fawcett, 2006).

Procesamiento de imágenes antes de obtener indicadores de calidad

Para cada proceso de segmentación de imágenes digitales, el mejor clasificador PNN se guardó en un archivo y se le asignó

probability that the classifier correctly assigns an individual to a target class (Fawcett, 2006).

Image processing before obtaining quality indicators

For each segmentation process of digital images, the best PNN classifier was saved in a file and accessed via a dynamic library link using a program written in Visual Basic 6.0 language and then, it was executed in a desktop computer equipped with an Intel Core Duo processor at 2.33 GHz and 4.0 GB RAM.

Images segmented by the best PNN classifier were processed with a median filter with a mask of 5×5 to eliminate random ("salt and pepper") and negative exponential noise. This spatial filter is a nonlinear filtering technique that replaces a pixel value by the median of the values of neighbor pixels with the aid of a convolute mask or array. This filter improves the image without blurring it (Gonzalez and Woods, 2002) and preserves edges of objects. This filtering was followed by dilation (for object expansion) and erosion (for object shrinking) using a structuring element of 5×5 pixels wide to eliminate small holes and join separate objects.

Quality indicators

Area covered by flowers (ACF, cm^2) and area covered by plants (ACP, cm^2) refer to projected areas on the soil surface. These indicators were calculated by counting the number of pixels belonging to each flower color or plant class. The percentage of pixels in each class was calculated, and the corresponding surface in cm^2 for each class was determined using the total analyzed surface of $3,173 \text{ cm}^2$ per image. Then, the area covered with flowers/plant ratio (RACF) was calculated as (ACF / ACP) .

Plant perimeter (PP, cm) was calculated by counting the number of pixels on the periphery of the area covered by plant. The following equation was used for such purpose (Gonzalez and Woods, 2002):

$$\beta(A) = A - (A \ominus B) \quad (1)$$

where $\beta(A)$ is the image of an object with a 1-pixel thick perimeter; A is the original binary image; and $(A \ominus B)$ is the erosion process of the binary image A with a structuring element B of 3×3 pixels. The PP was calculated counting the pixels on the perimeter in image $\beta(A)$ and scaling them to the object dimensions.

The number of open flowers per plant (NFP) was counted in each sampling date. The Euler number (EN) was used to calculate the number of objects in a binary image. The EN is defined as the

acceso a través de un enlace a biblioteca dinámica con un programa escrito en lenguaje Visual Basic 6.0; y luego, se ejecutó en una computadora de escritorio equipada con un procesador Intel Core Duo a 2.33 GHz y 4.0 GB de RAM.

Las imágenes segmentadas por el mejor clasificador PNN se procesaron con un filtro de mediana con máscara de 5×5 para eliminar los ruidos, aleatorio ("sal y pimienta") y exponencial negativo. Este filtro espacial es una técnica de filtrado no lineal que reemplaza el valor de un píxel, por la mediana de los valores de los píxeles contiguos, con la ayuda de una matriz o máscara convoluta. Este filtro mejora la imagen sin difuminarla (González y Woods, 2002) y conserva los bordes de los objetos. Al filtrado lo siguió la dilatación (por expansión del objeto) y la erosión (por reducción del objeto), con el uso de un elemento estructurante de 5×5 píxeles de ancho para eliminar agujeros pequeños y unir objetos separados.

Indicadores de calidad

El área cubierta por flores (ACF, cm^2) y el área cubierta por plantas (ACP, cm^2) se refieren a las áreas proyectadas sobre la superficie del suelo. Estos indicadores se calcularon contando el número de píxeles que pertenecen a cada clase de color de flor o clase de planta. El porcentaje de píxeles en cada clase se calculó y se determinó la superficie correspondiente en cm^2 para cada clase utilizando una superficie total analizada de 3173 cm^2 por imagen. Luego, la proporción entre el área cubierta con flores/área cubierta de plantas (RACF) se calculó como (ACF / ACP) . El perímetro de la planta (PP, cm) se calculó contando el número de píxeles en la periferia del área cubierta por la planta. La siguiente ecuación se utilizó para tal fin (González y Woods, 2002):

$$\beta(A) = A - (A \ominus B) \quad (1)$$

donde $\beta(A)$ es la imagen de un objeto con un perímetro de 1 píxel de espesor; A es la imagen binaria original; y $(A \ominus B)$ es el proceso de erosión de la imagen binaria A con un elemento estructurante B de 3×3 píxeles. El PP se calculó contando los píxeles en el perímetro en la imagen $\beta(A)$ proyectados a la escala de las dimensiones del objeto.

El número de flores abiertas por planta (NFP) se contó en cada fecha de muestreo. El número de Euler (EN) se usó para calcular el número de objetos en una imagen binaria. El EN se define como el número total de objetos en la imagen menos el número de agujeros en esos objetos. La imagen se convirtió a una imagen binaria al discriminar los píxeles asignados a la imagen de la flor del resto de las clases (Figura 2C), por lo que EN es

total number of objects in the image minus the number of holes in those objects. The image was converted to a binary image by discriminating the pixels assigned to the flower image from the rest of the classes (Figure 2C), thus EN is an estimator for NFP. First, the image was scanned using the patterns of two matrices of dimension (2×2) to find convexities and concavities. Then, EN was calculated by subtracting the number of holes in the objects (concavities) from the number of objects (convexities) (Umbaugh, 1998).

Roundness ratio (T) was calculated using the equation proposed by Umbaugh (1998):

$$T = 4\pi \left(\frac{ACP}{PP} \right) \quad (2)$$

where ACP is the area covered by the plant (cm^2); PP is the perimeter of the petunia plant (cm); and T is a quality index, that

un estimador para NFP. Primero, la imagen se digitalizó usando los patrones de dos matrices de dimensión (2×2) para encontrar convexidades y concavidades. Luego, el EN se calculó restando el número de agujeros en los objetos (concavidades) del número de objetos (convexidades) (Umbaugh, 1998).

El índice de redondez (T) se calculó utilizando la ecuación propuesta por Umbaugh (1998):

$$T = 4\pi \left(\frac{ACP}{PP} \right) \quad (2)$$

donde ACP es el área cubierta por la planta (cm^2); PP es el perímetro de la planta de petunia (cm); y T es un índice de calidad que calcula la calidad de la planta de petunia con base en la compacidad o la redondez de la planta, dado que una planta de petunia debe tener ramificaciones más cortas (PanAmerican Seed, 2007).

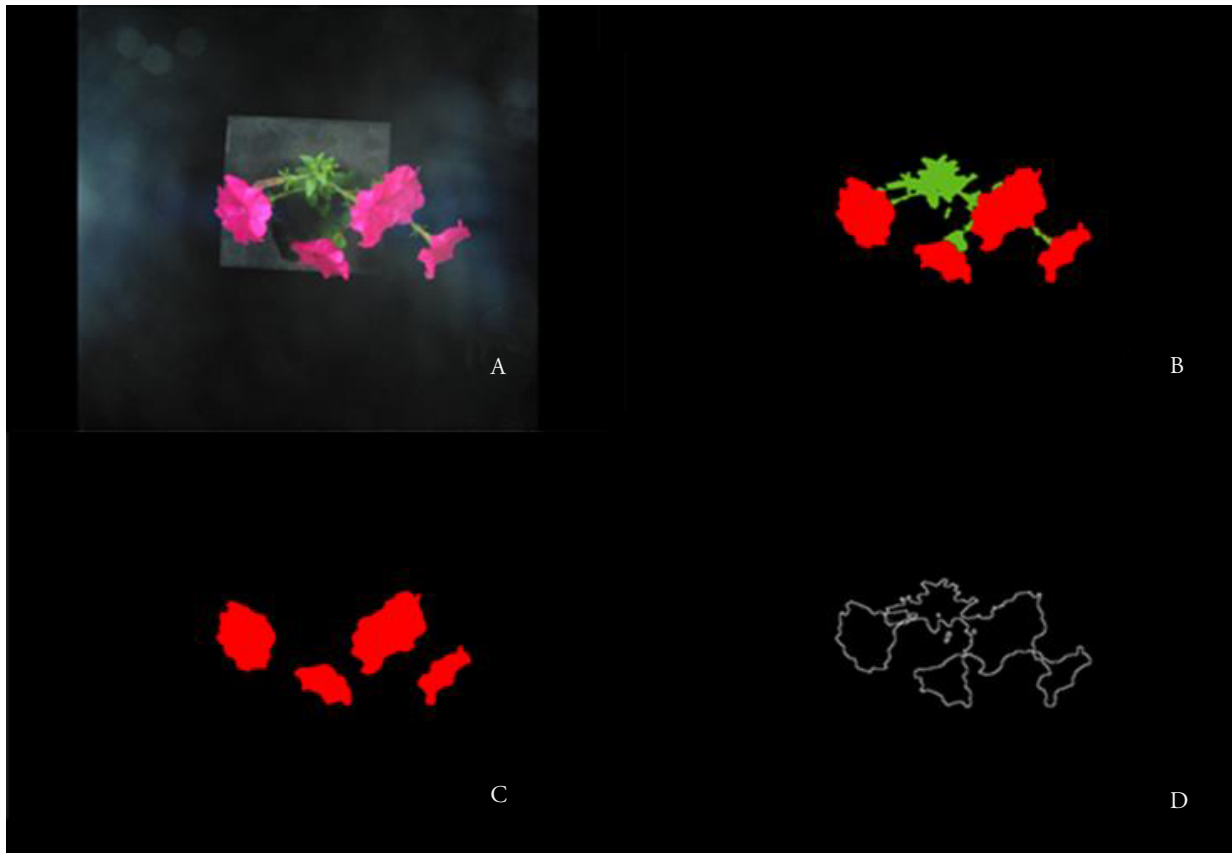


Figure 2. Processing of digital images of petunia plants to calculate the desired variables: A) original image on the horizontal plane, B) image classified by categories flower, leaf and background, C) selection of area covered by flowers, and D) plant perimeter.

Figura 2. Procesamiento de imágenes digitales de plantas de petunia para calcular las variables deseadas: A) imagen original en el plano horizontal, B) imagen clasificada por las categorías de flor, hoja y fondo, C) selección del área cubierta por flores y D) perímetro de planta.

calculates petunia plant quality based on compactness or plant roundness, since a petunia plant must have shorter ramifications (PanAmerican Seed, 2007).

For illustrating, the tendency of the calculated variables through time, linear, cubic spline and sigmoidal models were adjusted by using the CurveExpert software (Hyams, 2018); and the standard deviation (SD) for output variables was also calculated.

RESULTS AND DISCUSSION

Comparison of classification scenarios

Five of the best classification scenarios are presented in Table 1; the maximum difference between them, in their OA was around 2%, and all of them performed well in the classification of pixels belonging to each target class. Therefore, the color space chosen here as input, for segmentation, was the RGB color space, since this is the native color space for the images and its acquisition in the computer program is straightforward. Cubero *et al.* (2014) also used the RGB color space as input in a discriminant Bayesian model. In the test set, the mean OA was $96.2 \pm 1.6\%$, the PNN classifier shows a good performance and it is similar to the values

Con propósitos ilustrativos, la tendencia de las variables calculadas a través del tiempo se ajustó a modelos de interpolación lineal, cúbica y sigmoidal con el uso del software CurveExpert (Hyams, 2018); y también se calculó la desviación estándar (SD) de las variables de respuesta.

RESULTADOS Y DISCUSIÓN

Comparación de escenarios de clasificación

Cinco de los mejores escenarios de clasificación se presentan (Cuadro 1); la diferencia máxima entre ellos, en OA, fue de alrededor del 2%, y todos ellos obtuvieron buenos resultados en la clasificación de los píxeles que pertenecen a cada clase objetivo. Por lo tanto, el espacio de color seleccionado aquí, como entrada para la segmentación, fue el espacio de color RGB, ya que es el espacio de color nativo para las imágenes y su adquisición es directa en el programa de computadora. Cubero *et al.* (2014) también utilizaron el espacio de color RGB como entrada en un modelo Bayesiano discriminante. En el conjunto de prueba, la OA promedio fue de $96.2 \pm 1.6\%$, el clasificador PNN tuvo un buen desempeño y fue similar a los valores reportados por Parsons *et al.* (2009) de 93, 84 y 96% en flores de clavel (*Dianthus spp.*), viola (*Viola spp.*) y ciclamen (*Cyclamen spp.*), respectivamente. La precisión del clasificador PNN, obtenida con la estimación del AUC para todas las clases (blanco, rosado, fucsia, azul, hoja y fondo), fue mayor a 0.98 en este escenario (Cuadro 2), lo cual significa un buen desempeño. Algunos píxeles en flores de color oscuro como fucsia y azul, se clasificaron ocasionalmente como fondo de imagen, por lo que las sensibilidades de estas clases fueron las más bajas (Cuadro 2). Hernández-López *et al.* (2012) consideraron que la precisión en la identificación del color mejora al transformar el espacio de color RGB al CIE-Lab, así se logra una mejor segmentación del color. En nuestro estudio, esta transformación dio como resultado un aumento del 1% en OA. Según Ewing y Horton (1999), los canales en el espacio de color RGB se correlacionan con la asignación de color, de modo que el matiz no se puede separar de la intensidad del color; esto se considera un inconveniente en este espacio de color, ya que dificulta la identificación de los colores entre sí.

Table 1. Input classification scenarios for the PNN classifier with the highest overall accuracy (OA) of pixels in the test set from images of petunia plants in flowering.

Cuadro 1. Escenarios de clasificación de entrada para el clasificador PNN con la mayor precisión global (OA) de píxeles en el conjunto de prueba, a partir de imágenes de plantas de petunia en floración.

Input	OA (%)	Training time
R, G, B	96.2	2 min 45 s
Hunter-L, Hunter-a, Hunter-b	96.9	3 min 15 s
CIE-L, CIE-a, CIE-b	97.3	2 min 12 s
R, Hunter-a, CIE-a, CIE-b, C(u), M	97.6	4 min 40 s
R, G, B, CIE-L, CIE-a, CIE-b, C(u), H, C	98.1	7 min 14 s

R: Red channel; G: Green channel; B: Blue channel; CIE-L, CIE-a, CIE-b: Channels from CIE-Lab space; Hunter-L, Hunter_a, Hunter-b: Channels from Hunter-Lab space; C(u): Channel from CIE-Luv space; H: Hue; C: Chroma. ♦ R: canal rojo; G: canal verde; B: canal azul; CIE-L, CIE-a, CIE-b: Canales del espacio CIE-Lab; Hunter-L, Hunter_a, Hunter-b: Canales del espacio de Hunter-Lab; C (u): Canal del espacio CIE-Luv; H: matiz; C: croma.

reported by Parsons *et al.* (2009) of 93, 84 and 96% in flowers of carnation (*Dianthus* spp.), viola (*Viola* spp.) and cyclamen (*Cyclamen* spp.), respectively. The accuracy of the PNN classifier measured by AUC for all classes (white, pink, fuchsia, blue, leaf, and background) with this scenario was greater to 0.98 (Table 2), showing a good performance. Some pixels in dark colored flowers such as fuchsia and blue were occasionally classified as image background, thus sensitivities of these classes were the lowest (Table 2). Hernández-López *et al.* (2012) considered that the precision in color identification improves by transforming the RGB to CIE-Lab color space, thus a better color segmentation is achieved. In our study, this transformation resulted in 1% increase in the OA. According to Ewing and Horton (1999), channels in the RGB color space are correlated to color assignation so that hue cannot be separated from color intensity; this is considered a drawback in this color space, because it makes difficult to identify colors from each other.

Quality indicators and flower color

The segmentation process of the original plant image is illustrated in Figure 2A for calculating the ACP in the horizontal plane (Figure 2B). The ACP varied among flower colors. The ACP increased in white or pink flowered plants to 80 - 90 dat and thereafter stayed constant until 119 dat (Figure 3A). In blue or fuchsia flowered plants, ACP increased until 119 dat and then decreased due to leaf senescence at the base of the plant; this decrease is interpreted as a reduction in plant quality. However, at 135 dat ACP

Indicadores de calidad y color de flores

El proceso de segmentación de la imagen original de la planta se ilustra en la Figura 2A para calcular el ACP en el plano horizontal (Figura 2B). El ACP varió entre los colores de las flores; aumentó en plantas de flores blancas o rosadas a los 80-90 días después del trasplante (dat) y luego permaneció constante hasta 119 dat (Figura 3A). En plantas de flores azules o fucsias, el ACP aumentó hasta 119 dat y luego disminuyó debido a la senescencia de las hojas en la base de la planta; esta disminución se interpreta como una reducción en la calidad de la planta. Sin embargo, a 135 dat el ACP se duplicó en todos los tratamientos debido al crecimiento de brotes axilares. Las plantas de flores blancas tuvieron mayor ACP, y las de flores rosadas mostraron el ACP más bajo.

El PP se comportó de manera diferente entre los colores de flor, y aumentó continuamente en todos los colores, durante el período de evaluación (75 a 135 dat). El valor final de PP fue al menos el doble del inicial (Figura 3B). La planta aumentó en altura al formar nuevas hojas en el ápice de la rama, sin cambios aparentes en el diámetro de la planta; por lo tanto, este crecimiento de la planta es un incremento en el perímetro, pero no en el área proyectada. También se detectaron diferencias en el PP: las plantas con flor azul o fucsia registraron los valores mayores, mientras que las plantas con flor blanca o rosada tuvieron valores menores. Dos tendencias diferentes pueden distinguirse claramente en las plantas de petunia, utilizando las curvas *ACP* y *PP*: el comportamiento de las plantas de flores blancas o rosadas y el de las plantas de flores azules o fucsias.

Table 2. Mean (\bar{x}), standard deviation (SD) and coefficient of variation (CV) (n=10) for the classification scenario of RGB input “red” (R), “green” (G), y “blue” (B), and a multiclass output variable (with six classes).

Cuadro 2. Media (\bar{x}), desviación estándar (SD) y coeficiente de variación (CV) (n=10) para el escenario de clasificación de la entrada RGB “rojo” (R), “verde” (G), y “azul” (B), y una variable de salida multiclase (con seis clases).

Performance criteria		Classes					
		Blue	White	Fuchsia	Pink	Leaf	Background
AUC	\bar{x}	0.997	0.990	0.986	0.9940	0.998	0.994
	SD	0.004	0.000	0.009	0.0050	0.004	0.005
	CV (%)	0.480	0.000	0.979	0.5195	0.422	0.519
Sensitivity	\bar{x}	0.939	0.964	0.919	0.9530	0.979	1.000
	SD	0.053	0.042	0.067	0.0470	0.029	0.000
	CV (%)	5.733	4.406	7.316	5.0210	2.985	0.000

AUC: Area under the receiver operating characteristic curve. ♦ AUC: Área bajo la curva operativa característica del receptor.

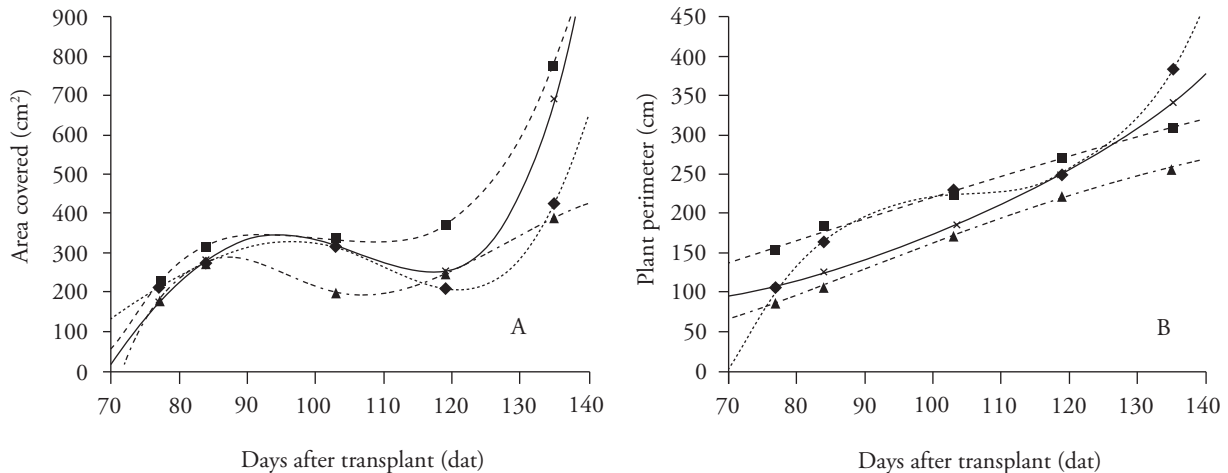


Figure 3. Mean of covered area (3A) and perimeter (3B) of petunia plants var. 'Glistening Pearl mixed' grown in pots with flowers colored blue (◆,), white (■, ----), fuchsia (×, —) and pink (▲, ---), transplanted in April at the Valley of Mexico; curves are added to show trends in data, not in order to interpret the resultant model parameters, symbols represent observed data.

Figura 3. Media del área cubierta (3A) y el perímetro (3B) de las plantas de petunia var. 'Glistening Pearl mixed', cultivadas en macetas con flores de color azul (◆,), blanco (■, ----), fucsia (×, —) y rosado (▲, ---), trasplantadas en abril en el Valle de México. Las curvas se agregan para mostrar las tendencias en los datos, no para interpretar los parámetros resultantes del modelo; los símbolos representan los datos observados.

had double size in all treatments due to axillary shoot sprouting. White-flowered plants had greater ACP, and pink-flowered plants showed the smallest ACP. The *PP* behaved differently among the flower colors, and it continuously increased in all colors during the evaluation period (75 to 135 dat). The final *PP* value at least doubled the initial reading (Figure 3B). The plant increased height by forming new leaves on the branch apex without apparently changing the plant diameter; therefore, this plant growth is an increment in perimeter but not as increment in projected area. Differences in *PP* were also detected: blue and fuchsia had the larger values, while white and pink had the lower values. Two different tendencies can be clearly distinguished in petunia plants using the *ACP* and *PP* curves: white or pink flowered plants behavior and blue or fuchsia flowered plants behavior.

In this ornamental plant, *NFP* increased rapidly until 85 dat when reached the value of 7 to 8 flowers per plant (Figures 4A, 4B and 4D), except in fuchsia flowered plants, which showed a slow rate of flower formation until 135 dat (Figure 4C). The *EN* y *NFP* followed a similar tendency, but in plants with blue flowers. The *EN* was consistently lower than *NFP*; therefore, it was not a good estimator for *NFP*.

The PNN classifier can group by color, but it cannot identify limits among same-colored objects;

En esta planta ornamental, el *NFP* se incrementó rápidamente hasta 85 dat, cuando alcanzó la cantidad de 7 a 8 flores por planta (Figuras 4A, 4B y 4D), excepto en plantas de flores fucsias, las cuales formaron flores lentamente, hasta 135 dat (Figura 4C). El *EN* y el *NFP* siguieron una tendencia similar, excepto en las plantas con flores azules. El *EN* fue consistentemente menor que el *NFP*; por lo tanto, no fue un buen estimador del *NFP*.

El clasificador PNN puede agrupar por color, pero no puede identificar límites entre objetos del mismo color; por lo tanto, las flores superpuestas no pueden detectarse. Esta limitación podría superarse con las técnicas de detección de bordes, como lo sugieren Hernández-López *et al.* (2012), ya que sus técnicas permiten segmentar objetos con base en el color y la profundidad de la imagen. Thorp y Dierig (2011) calcularon el número de flores ($R^2=0.75$) en *Lesquerella fendleri* a partir de la superficie segmentada de la cobertura de flores, con base en el espacio de color HSI y el análisis de Monte Carlo para reducir la incertidumbre en la estimación de parámetros para la segmentación de la imagen.

En todas las plantas de petunia, *T* disminuyó con el tiempo (Figura 5). Al inicio de la floración (119 dat), *T* cambió según el color de la flor, desde 0.14 hasta 0.29; mientras que a 135 dat todos

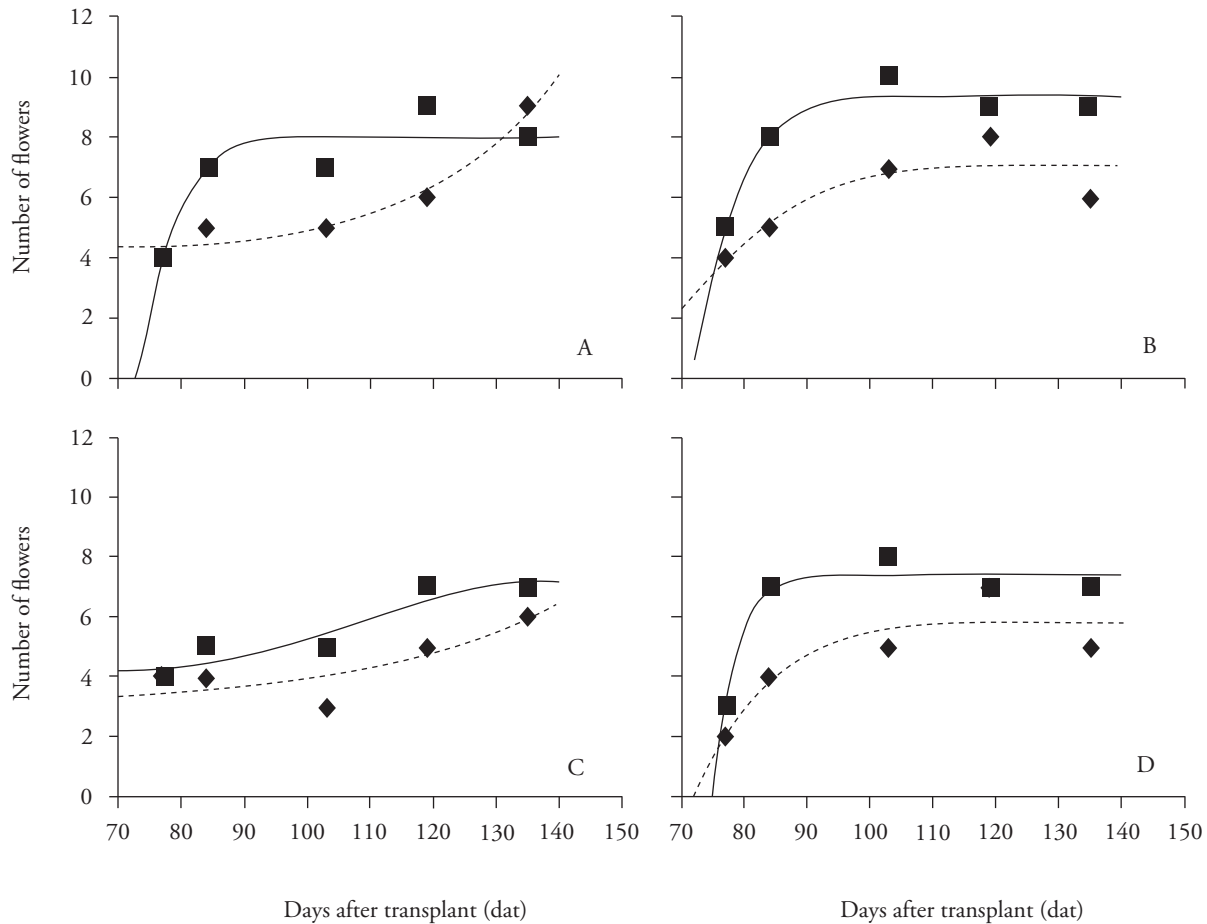


Figure 4. Mean number of completely opened flowers per plant of petunia var ‘Glistening Pearl mixed’ transplanted in April in the Valley of Mexico (■, —) and estimated mean number using the Euler number (◆, ----) of plants with A) blue, B) white, C) fuchsia, and D) pink flowers; curves are added to show trends in data, not in order to interpret the resultant model parameters, symbols represent observed or calculated data.

Figura 4. Media del número de flores completamente abiertas por planta de petunia var. ‘Glistening Pearl mixed’, trasplantada en abril en el Valle de México (■, —) y media estimada del número usando el número de Euler (◆, ----) en plantas con flores A) azules, B) blancas, C) fucsias y D) rosadas; se agregan curvas para mostrar las tendencias en los datos, no para interpretar los parámetros resultantes del modelo; los símbolos representan datos observados o calculados.

therefore, overlapped flowers cannot be detected. This limitation could be overtaken with border detection techniques as suggested by Hernández-López *et al.* (2012) because their techniques allow to segment objects based on image color and depth. Thorp and Dierig (2011) calculated flower number ($R^2=0.75$) in *Lesquerella fendleri* from the segmented surface of flower covering, based on the HSI color space and the Monte Carlo analysis to reduce uncertainty in the parameter estimation for image segmentation.

In all petunia plants, T decreased with time (Figure 5). At the start of flowering (119 dat), T changed depending on flower color from 0.14 to 0.29, whereas at 135 dat all treatments had smaller

los tratamientos tuvieron valores pequeños de 0.04 a 0.10, que fueron estadísticamente similares. Dos tendencias se identificaron en las curvas T : las plantas con flores rosadas y fucsias tuvieron una tendencia sigmoideal decreciente; mientras que las plantas con flores blancas y azules disminuyeron exponencialmente. Las plantas de flores blancas fueron más estables en T , ya que mostraron una tasa de disminución mínima. Para las plantas de petunia en maceta, el criterio de calidad se estableció como plantas compactas, con un buen número de ramas que no sean muy largas, aunque las formas irregulares son comunes en varias especies y cultivares (PanAmerican Seed, 2007; Meijón *et al.*, 2009). El índice T más alto (0.29) en

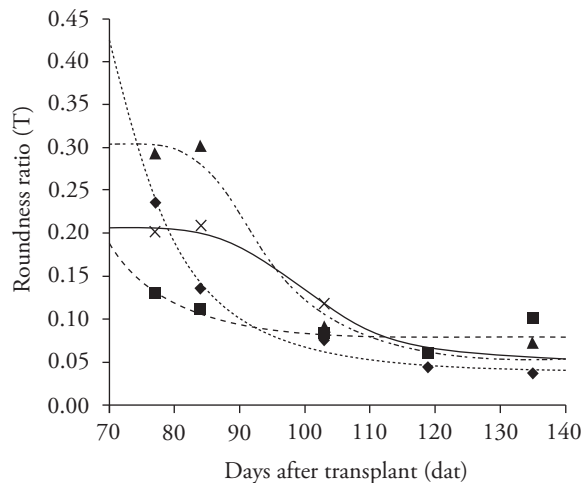


Figure 5. Roundness ratio (T) of plants of petunia var. ‘Glistening Pearl mixed’ plants transplanted in April at the Valley of Mexico, with blue (♦, ···), white (■, ----), fuchsia (×, —) and pink (▲, -.-) flowers; curves are added to show trends in data, not in order to interpret the resultant model parameters, symbols represent observed data.

Figura 5. Tasa de redondez (T) de plantas de petunia var. “Glistening Pearl mixed”, trasplantadas en abril en el Valle de México, con flores azules (♦, ···), blancas (■, ----), fucsias (×, —) y rosadas (▲, -.-); las curvas se agregan para mostrar las tendencias en los datos, no para interpretar los parámetros resultantes del modelo; los símbolos representan los datos observados.

values of 0.04 and 0.10 which were statistically similar. Two tendencies were identified in the T curves: pink and fuchsia flowered plants had a sigmoidal decreasing tendency, whereas white and blue flowered plants decreased exponentially. White flowered plants were more stable in T since they showed the minimal decrease rate. For potted petunia plants, the quality criterion was established as compact plants with a good number of branches that are not too long, although irregularly shaped forms are common in various species and cultivars (PanAmerican Seed, 2007; Meijón *et al.*, 2009). The highest T index (0.29) in petunia plants with pink flowers is very low compared to the T index in azalea (*Rhododendron simsii* hybrids) (0.68) (Meijón *et al.*, 2009) or to the T index of a circular object (1.0).

The T value provides with a quantitative measurement of petunia quality, and it can be useful for measuring the period in which plants keep their quality in terms of compactness. The period in which the maximum T value is reduced to half since

las plantas de petunia con flores rosadas es muy bajo en comparación con el índice T en azalea (híbridos de *Rhododendron simsii*) (0.68) (Meijón *et al.*, 2009) o con el índice T de un objeto circular (1.0).

El valor de T proporciona una medida cuantitativa de la calidad de la petunia y puede ser útil para medir el período en que las plantas mantienen su calidad en términos de compactidad. El período en el cual el valor máximo de T se reduce a la mitad desde el inicio de la floración (referido al período en que las plantas exhiben calidad) es de 11 d para las plantas de flores azules y 24 d para las plantas de flores fucsias o rosadas. En las plantas con flores blancas tomó 60 d que T disminuyera a la mitad. La curva T también podría usarse como una herramienta para aplicar prácticas oportunas de manejo de cultivos, como los tratamientos hormonales (Meijón *et al.*, 2009) o la poda.

Al inicio de la floración, el RACF fue de 60% en plantas con flores fucsias, y alcanzó 71% en las plantas con flores blancas (Figura 6). Esta proporción disminuyó a una tasa de 0.3% por día en plantas de flores blancas y de 0.7% en plantas de flores azules. Al final del período de evaluación, las plantas de flores azules registraron un RACF de 33% y las plantas de flores blancas 54%.

La técnica de análisis de imágenes utilizada en este estudio proporciona un procedimiento que permite calcular indicadores de calidad que pueden, de manera cuantitativa, discriminar genotipos, color de las flores y determinar el período en que las plantas de petunia en maceta pueden conservar su calidad. Parsons *et al.* (2009) propusieron otras variables obtenidas a partir del análisis de imágenes, como la uniformidad de la cobertura de flores y hojas, la altura y forma del follaje, el color de las hojas y las flores; su modelo de regresión pudo explicar 88.5, 81.7 y 70.4% de los puntajes del panel de evaluación para la calidad de la planta de clavel, viola y ciclamen, respectivamente.

Esta metodología proporciona un análisis no destructivo, no invasivo, económico, rápido y fácil para medir indirectamente la morfología, el crecimiento y otras variables fisiológicas, utilizando diferentes bandas capturadas con una cámara de visión superior; esto es útil para fenotipar (Humplík *et al.*, 2015) cuando el objetivo es encontrar relaciones entre genes y caracteres fenotípicos. De hecho, el fenotipado es una limitante para el análisis genético y la predicción

flowering initiation (referred to the period plants exhibit quality), it is 11 d for blue-flowered plants and 24 d for fuchsia or pink flowered plants. White flowered plants took 60 d to decrease T by half. The T curve could also be used as a tool for timely crop management practices, such as hormone treatments (Meijón *et al.*, 2009) or pruning.

At the start of flowering, $RACF$ was 60% in fuchsia-flowered plants, and it reached 71% with white-flowered plants (Figure 6). This ratio decreased at a rate of 0.3% per day in white-flowered plants, and by 0.7% in blue-flowered plants. At the end of the evaluation period, blue-flowered plants had an $RACF$ of 33 %, and white-flowered plants 54%.

The image analysis technique used in this study provides with a procedure to calculate quality indicators that can quantitatively discriminate genotypes, flower colors and determine the period in which the potted petunia plants can preserve their quality. Parsons *et al.* (2009) proposed other variables obtained from image analysis, like uniformity of flower and leaves cover, height and shape of foliage, leaf and flower color; their regression model was able to explain 88.5, 81.7 and 70.4% of the evaluation panel scores for the plant quality of carnation, viola and cyclamen, respectively.

This methodology provides a non-destructive, non-invasive, cheap, quick, and easy analysis for measuring indirectly morphology, growth, and other physiological variables, using different bands captured with a top-view camera; this is useful for phenotyping (Humplík *et al.*, 2015) when the objective is finding relationships between genes and traits. Actually, phenotyping limits genetic analysis and genomic prediction (Rahaman *et al.*, 2015) and image analysis proves to be a useful tool.

For instance, Crowell *et al.* (2014) identified a large number of quantitative trait loci (QTL) for panicle traits once they measured multiple architectural and branching phenotypes, usually correlated with yield and grain quality, from digital images. Also, Iwata *et al.* (2002) found that variation in citrus leaf shape characterized from descriptors obtained from digital images is defined by the genotype while the interaction between genotype and environment had a small effect on leaf shape and size. In a study in azalea, De Keyser *et al.* (2013) obtained data from image analysis for QTL mapping. Image analysis is also applicable when looking for drought tolerant

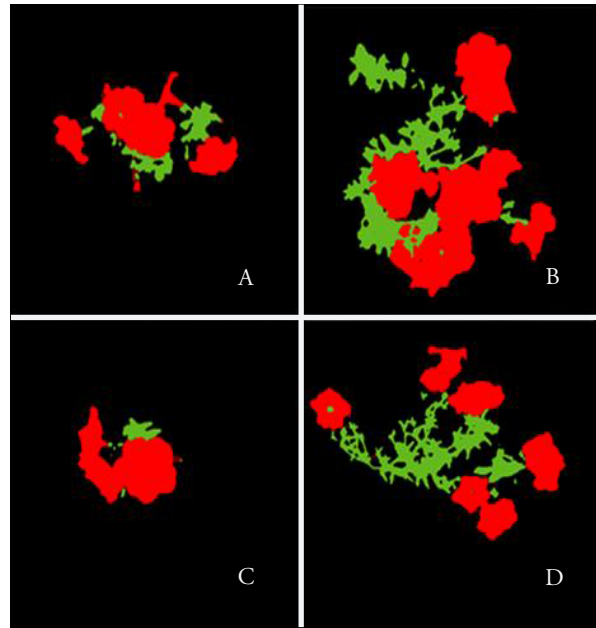


Figure 6. Area covered by flower (in red) and area covered by leaves (in green) of petunia plants with white flowers (A and B) and pink (C and D) at 77 dat (A and C) and 135 dat (B and D).

Figura 6. Área cubierta por flor (en rojo) y área cubierta por hojas (en verde) de plantas de petunia con flores blancas (A y B) y rosadas (C y D) a los 77 dat (A y C) y 135 dat (B y D).

genómica (Rahaman *et al.*, 2015) y el análisis de imágenes demuestra ser una herramienta útil.

Por ejemplo, Crowell *et al.* (2014) identificaron un gran número de loci de caracteres cuantitativos (QTL) de la panícula, los cuales se correlacionan en general con el rendimiento y la calidad del grano, después de medir la arquitectura y ramificación de un gran número de fenotipos, usando análisis de imágenes. También, Iwata *et al.* (2002) encontraron que la variación en la forma de la hoja de los cítricos, caracterizada mediante descriptores obtenidos a partir de imágenes digitales, está definida por el genotipo, mientras que la interacción entre el genotipo y el ambiente tuvo un efecto menor sobre la forma y el tamaño de la hoja. En un estudio en azalea, De Keyser *et al.* (2013) obtuvieron datos a partir del análisis de imágenes para el mapeo QTL . El análisis de imágenes también es aplicable cuando se buscan genotipos tolerantes a la sequía (Humplík *et al.*, 2015), marchitamiento de plantas (McCarthy *et al.*, 2010), crecimiento de plantas (Yeh *et al.*, 2014), cantidad de cobertura del suelo por planta (Velázquez -García *et*

genotypes (Humplík *et al.*, 2015), plant wilting (McCarthy *et al.*, 2010), plant growth (Yeh *et al.*, 2014), amount of soil covering by plant (Velázquez-García *et al.*, 2010), and grading of plant cuttings (McCarthy *et al.*, 2010).

The research in the use of image analysis in agriculture has increased remarkably (Humplík *et al.*, 2015), using new proposals on hardware from top-view camera to stereovision systems and three-views RGB imaging (Yeh *et al.*, 2014; Humplík *et al.*, 2015). Besides, more work was developed to extract information from multispectral to hyperspectral images at visible and near infrared wavelengths (Rahaman *et al.*, 2015).

Although phenotyping may be cheap and straightforward using commercial cameras in the visible band, it may be prone to errors due to the overlapping adjacent leaves and soil background during the segmentation process (Rahaman *et al.*, 2015). Besides, natural light makes it difficult to segment images since color is light intensity dependent.

CONCLUSIONS

A probabilistic neuronal network (PNN) classifier was used for segmenting the area covered by leaves and the area covered by flowers using digital images of potted plants of *Petunia multiflora* F1 'Glistening Pearls mixed'. The PNN classifier showed a good performance in terms of the overall accuracy (98.1%) for all color classes studied, the mean area under the ROC curve (0.99 ± 0.004) and the sensitivity (0.96 ± 0.04). The best overall accuracy (98.1%) of the PNN was obtained with a combination of channels of various color spaces as input, while using the RGB color space alone as input the overall accuracy of the PNN classifier was 96.2%. From this process, quality indicators were generated for detecting differences in behavior of the plant quality variables associated to flower color, as they were a mixture of genotypes.

White-flowered plants covered a greater area, greater ratio of area covered by flowers, and the highest stability in the roundness ratio. Pink-flowered plants behaved differently; they had lower area covered by plant, lower plant perimeter, larger roundness ratio at the start of flowering, and lost roundness quicker.

Euler number was not a good estimator of number of flowers per plant. Therefore, the use of quality

et al., 2010), y para clasificar esquejes de plantas (McCarthy *et al.*, 2010).

La investigación en el uso de análisis de imágenes en la agricultura se ha incrementado notablemente (Humplík *et al.*, 2015), utilizando nuevas propuestas de equipos digitales, desde cámaras de visión sobre el objeto hasta sistemas de estereovisión e imágenes RGB de tres vistas (Yeh *et al.*, 2014; Humplík *et al.*, 2015). Además, se ha realizado investigación en donde se obtienen resultados a partir de imágenes multiespectrales a hiperespectrales en el espectro visible e infrarrojo cercano (Rahaman *et al.*, 2015).

Aunque hacer fenotipado usando cámaras comerciales en la banda visible puede ser económico y directo, puede también conducir a cometer errores experimentales debido a la superposición de las hojas adyacentes y del suelo como trasfondo durante el proceso de segmentación (Rahaman *et al.*, 2015). Además, la luz natural dificulta la segmentación de las imágenes, ya que el color depende de la intensidad de la luz.

CONCLUSIONES

Un clasificador de red neuronal probabilística (PNN) se aplicó para segmentar el área cubierta por hojas y el área cubierta por flores, utilizando imágenes digitales de plantas en macetas de *Petunia multiflora* F1 "Glistening Pearls mixed". El clasificador PNN mostró un buen desempeño en términos de la precisión global (98.1%) para todas las clases de color estudiadas, la media del área bajo la curva ROC (0.99 ± 0.004) y la sensibilidad (0.96 ± 0.04). La mejor precisión global (98.1%) del PNN se obtuvo al usar como entrada una combinación de canales de varios espacios de color; mientras que al usar sólo el espacio de color RGB como entrada, la precisión global del clasificador PNN fue de 96.2%. A partir de este proceso, se generaron indicadores de calidad para detectar diferencias en el comportamiento de las variables de calidad de la planta asociadas al color de la flor, ya que fueron una mezcla de genotipos.

Las plantas de flores blancas cubrieron un área más amplia, una proporción mayor de área cubierta por flores y la estabilidad más alta en el índice de redondez. Las plantas de flores rosadas se comportaron de manera diferente: tuvieron menor área cubierta por planta, menor perímetro de planta, y un mayor índice de redondez al inicio de la floración, y perdieron la redondez más rápido.

indicators obtained from segmentation of color digital images by a PNN classifier enables to improve the accuracy for assessing the time-dependent changes in the quality of ornamental plants at a low cost as only a commercial camera and a dark space are needed to produce images from plants. The rest is acquired using image analysis.

LITERATURE CITED

- CIE. 2004. Colorimetry. Pub. No. 15.3. Third Edition. Central Bureau, CIE. Vienna, Austria. 72 p.
- Crowell S., A. X. Falcão, A. Shah, Z. Wilson, A. J. Greenberg, and S. R. McCouch. 2014. High-resolution inflorescence phenotyping using a novel image-analysis pipeline, PANorama. *Plant Physiol.* 165: 479-495.
- Cubero, S., M. P. Diago, J. Blasco, J. Tardáguila, B. Millán, and N. Aleixos. 2014. A new method for pedicel/peduncle detection and size assessment of grapevine berries and other fruits by image analysis. *Biosyst. Eng.* 117: 62-72.
- De Keyser, E., P. Lootens, E. Van Bockstaele, and J. de Riek. 2013. Image analysis for QTL mapping of flower color and leaf characteristics in pot azalea (*Rhododendron simsii* hybrids). *Euphytica* 189: 445-460.
- Ewing, R. P., and R. Horton. 1999. Quantitative color image analysis of agronomic images. *Agron. J.* 91: 148-153.
- Fawcett, T. 2006. An introduction to ROC analysis. *Pattern Recogn. Lett.* 27:861-874.
- García-Albarado, J. C., L. I. Trejo-Téllez, M.A. Velázquez-Hernández, A. Ruíz-Bello, y F.C. Gómez-Merino. 2010. Crecimiento de petunia en respuesta a diferentes proporciones de composta en sustrato. *Rev. Chapingo Serie Hort.* 16: 107-113.
- Gonzalez, R. C., and R. E. Woods. 2002. *Digital Imaging Processing*. Prentice Hall. Second Edition. New Jersey, USA. 793 p.
- Granitto, P. M., H. D. Navone, P. F. Verdes, and H. A. Ceccatto. 2002. Weed seeds identification by machine vision. *Comput. Electron. Agr.* 33: 91-103.
- Hemming, J., and T. Rath. 2001. Computer-vision based weed identification under field conditions using controlled lighting. *J. Agr. Eng. Res.* 78: 233-243.
- Hernández-López, J. J., A. L. Quintanilla-Olvera, J. L. López-Ramírez, F. J. Rangel-Butanda, M. A. Ibarra-Manzano, and D. L. Almanza-Ojeda. 2012. Detecting objects using color and depth segmentation with Kinect sensor. *Procedia Technol.* 3: 196-204.
- Hsieh, C. L., and S. F. Weng. 2005. Prediction of physical properties of orchid seedlings 'Phalaenopsis Sogo Vivien F819'. *Can. Biosystems Eng.* 47: 23-32.
- Humplík, J. F., D. Lazar, A. Husičová and L. Spíchal. 2015. Automated phenotyping of plants shoots using imaging methods for analysis of plant stress responses- a review. *Plant Methods* 11:29.
- Hyams, D. G. 2018. CurveExpert 2.6.5 software. <https://www.curveexpert.net> (Accessed. September 2018)
- Kavzoglu, T., and P.M. Mather. 2003. The use of backpropagating artificial neural networks in land cover classification. *Int. J. Rem. Sens.* 24: 4907-4938.
- El número de Euler no fue un buen estimador del número de flores por planta. Por lo tanto, el uso de indicadores de calidad obtenidos a partir de la segmentación de imágenes digitales en color por un clasificador PNN permite mejorar la precisión para evaluar los cambios dependientes del tiempo en la calidad de las plantas ornamentales con un costo bajo, ya que sólo se necesita una cámara comercial y un espacio oscuro para producir imágenes de plantas. El resto se obtiene mediante el análisis de imágenes.
- Fin de la versión en Español—
- *-----
- Littmann, E., and H. Ritter. 1997. Adaptive color segmentation a comparison of neural and statistical methods. *IEEE Trans. on Neural. Netw.* 8: 175-185.
- McCarthy, C. L., N. H. Hancock, and W. R. Raine. 2010. Applied machine vision of plants: a review with implications for field deployment in automated farming operations. *Intel. Serv. Robot.* 3: 209-217.
- Meijón, M., R. Rodríguez, M. J. Cañal, and I. Feito. 2009. Improvement of compactness and floral quality in azalea by means of application of plant growth regulators. *Sci. Hortic.* 119: 169-176.
- Mendoza, F., P. Dejmek, and J. M. Aguilera. 2006. Calibrated color measurements of agricultural foods using image analysis. *Postharvest Biol. Tec.* 41: 285-295.
- PanAmerican Seed. 2007. *Petunia Hedgiffora Serie Tidal Wave*. Grower Facts. Ball Horticultural Company, Chicago, USA. 4 p.
- Parsons, N. R., R. N. Edmondson, and Y. Song. 2009. Image analysis and statistical modelling for measurement and quality assessment of ornamental horticulture crops in glasshouses. *Biosyst. Eng.* 104: 161-168.
- Poynton, C. 1996. *A Technical Introduction to Digital Video*. John Wiley 1st Ed., New York. 356 p.
- Rahaman, M. M., D. Chen, Z. Gillani, C. Klukas, and M. Chen. 2015. Advanced phenotyping and phenotype data analysis for the study of plant growth and development. *Front. Plant Sci.* 6: 1-15.
- Robertson, A. 1977. The CIE color-difference formulae. *Color Res. Appl.* 2:7-11.
- SIAP (Servicio de Información Agroalimentaria y Pesquera). 2016. Anuario del Servicio de Información Agroalimentaria y Pesquera, ciclo 2015. Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación. <http://www.siap.gob.mx/cierre-de-la-produccion-agricola-por-cultivo/> (Accessed: June 2017).
- Smith, A. R. 1978. Color gamut transformations pairs. *Comput. Graph.* 12: 12-19.
- Specht, D. F. 1991. A general regression neural network. *IEEE Trans. Neural Netw.* 2: 568-576

- Stajanko, D., M. Lakota, and M. Hočvar. 2004. Estimation of number and diameter of apple fruits in an orchard during the growing season by thermal imaging. *Comput. Electron. Agr.* 42: 31-42.
- Thorp, K. R., and D. A. Dierig. 2011. Color image segmentation approach to monitor flowering in lesquerella. *Ind. Crop. Prod.* 34: 1150-1159.
- Umbaugh, S. E. 1998. *Computer Vision Image Processing*. Prentice Hall PTR, New Jersey, USA. 504 p.
- USDA (United States Department Agriculture). 2016. Floriculture crops 2015. Summary. U.S.D.A., U.S.A. <http://usda.mannlib.cornell.edu/usda/current/FlorCrop/FlorCrop-04-26-2016.pdf> (Accessed: May 2017).
- Velázquez-García, J., K. Oleschko, J. A. Muñoz-Villalobos, M. Velázquez-Valle, M. Martínez M., J.-F. Parrot, G. Korvin, and M. Cerca. 2010. Land cover monitoring by fractal analysis of digital images. *Geoderma* 160: 83-92.
- Yeh, Y. H. F., T. C. Lai, T. Y. Liu, C. C. Liu, W. C. Chung, and T. T. Lin. 2014. An automated growth measurement system for leafy vegetables. *Biosyst. Eng.* 117: 43-50.
- Zeinali, Y., and B. Story. 2017. Competitive probabilistic neural network. *Integ. Comput.-Aid. Eng.* 24: 1-14.