

Propuesta metodológica para la detección del virus de la mancha anular en frutos de papaya con procesamiento digital de imágenes

Jorge Ochoa Somuano^{1*}, Elizabeth Hernández Juárez ,
Francisco Ruíz Ruíz², Julieta Cruz Vázquez² & Verónica Ortega Baranda³

Resumen

La detección de objetos en procesamiento de imágenes digitales es un tema de interés que no ha disminuido con el paso del tiempo, por el contrario, cada vez son más las investigaciones orientadas a la detección e identificación de objetos. Aún con el creciente desarrollo de soluciones que permiten identificar objetos a partir de imágenes digitales, no se han abarcado todos los sectores. Un problema que se ha detectado en pláticas con productores de la región costa del Estado de Oaxaca, es que, en algunas ocasiones los cultivos de papaya presentan síntomas de virus de diferentes tipos, una vez que hacen la detección de las afectaciones, los productores invierten recursos económicos para curar los frutos, controlar el virus y eliminarlo de sus plantas y frutos. Sin embargo, en ocasiones la detección del virus se hace de forma tardía y pueden perder su producción o, por lo menos, parte de ella. Con base en lo anterior, surgió la idea de proponer una metodología de solución empleando el Procesamiento Digital de Imágenes, como primera etapa se pretende realizar la identificación del virus de la mancha anular en frutos de papaya con el objetivo de identificar de manera temprana el virus.

Palabras clave: Virus, detección, procesamiento digital, afectaciones, productores.

Recibido: 30 de mayo de 2019

Abstract

Object detection in digital image processing is a subject of interest has not diminished over time, indeed, more and more research oriented to the detection and identification of objects. Even with the growing development of solutions that allow to identify objects from digital images, not all sectors have been covered. A problem that has been detected in talks with producers in the coastal region of the State of Oaxaca, is that, sometimes papaya crops show symptoms of viruses of different types, once they make the detection of the affectations, farmers invest economic resources to heal the fruits, control the virus and eliminate it from their plants and fruits. However, sometimes the detection of the virus is done late and may lose its production or, at least, part of it. Based on the above, the idea of proposing a solution methodology using Digital Image Processing arose, that is to say, it is intended to perform the identification of the papaya ringspot virus in papaya fruits in order to identify the affectation early.

Key words: Virus, detection, digital processing, affectations, farmers.

Aceptado: 08 de agosto de 2019

¹ Instituto de Industrias, Universidad del Mar campus Puerto Escondido. Km. 2.5 Carretera Federal Puerto Escondido-Sola de Vega, Puerto Escondido 71980, Oaxaca, México.

² Instituto de Genética, Universidad del Mar campus Puerto Escondido. Km. 2.5 Carretera Federal Puerto Escondido-Sola de Vega, Puerto Escondido 71980, Oaxaca, México.

³ Instituto de Ecología, Universidad del Mar campus Puerto Escondido. Km. 2.5 Carretera Federal Puerto Escondido-Sola de Vega, Puerto Escondido 71980, Oaxaca, México.

* **Autor de correspondencia:** ochoa@zicatela.umar.mx (JOS)

Introducción

El inicio de la Inteligencia Artificial (IA) se atribuye a John McCarthy, Marvin Minsky, Nathaniel Rochester y Claude Shannon quienes acuñaron el término en una conferencia que ellos mismos organizaron en la Universidad Dartmouth College (Estados Unidos) en 1956 (Andresen 2002). Desde entonces y hasta la fecha, la IA ha tenido sus altibajos en cuanto al desarrollo de una máquina capaz de simular el comportamiento humano (Minsky 1961). Su enfoque ha sido utilizado en aplicaciones como: aprendizaje, razonamiento, cálculo, percepción, memorización e incluso descubrimiento científico o creatividad artística (Ganascia 2019). Cada una de las áreas mencionadas anteriormente ha sido importante para llegar al desarrollo del procesamiento de imágenes por medio de computadoras. Sin embargo, el origen de las imágenes se remonta a 1920 con la industria del periódico (Gonzalez & Woods 2001), al enviar una imagen digital por medio de un cable submarino entre Londres y Nueva York.

En torno a la identificación de afectaciones por virus, se han realizado algunas investigaciones que utilizan procesamiento digital de imágenes, entre ellas Singh *et al.* (2018) describen propuesta de solución para la identificación y clasificación de enfermedades fúngicas en las hojas de una planta. Para ello, utilizaron preprocesamiento de imagen, extracción de características, entrenamiento de una Red Neuronal de Base Radial empleando la Optimización de Búsqueda Bacteriana y segmentación de imagen por color, textura y forma. El algoritmo puede identificar enfermedades fúngicas como la roya común, la roya de la manzana de cedro, el tizón tardío, el rizo de la hoja, la mancha de la hoja y el tizón temprano.

En Kisan *et al.* (2017) plantean una solución para identificar enfermedades causadas por hongos o bacterias aplicada a hojas de tomate y uva en cultivos de la India. Los autores proponen una aplicación móvil para Android que utiliza técnicas de procesamiento digital de imágenes como: redimensión de la imagen, extracción interactiva de primer plano con el

algoritmo GrabCut, conversión RGB a HSV (Hue, Saturation, Value-Matiz, Saturación, Valor), umbralización y comparación de plantillas. Los resultados para la detección de enfermedades por hongos en hojas de tomate es de 94% de efectividad y por bacteria de 91.5%, en cuanto a las hojas de uva la detección de enfermedades por hongos representa un 94.7% y por bacterias un 93.2%.

Patel & Gamit (2016) identificaron el problema que existe con los métodos de acceso para recuperar imágenes de una gran base de datos de imágenes, utilizando anotaciones de palabras clave que describen el contenido de la imagen. Para solucionar dicha problemática, presentan un método que lleva a cabo la extracción de características con base al contenido de la imagen y posteriormente ubicarlas en un vector. Las técnicas utilizadas en función del color fueron: histograma, correlograma, descriptor de color dominante y matriz de co-ocurrencia del color; y en función de la textura: características de Tamura, pirámide dirigible, transformada de *wavelet* y, transformada *wavelet* de Gabor. Siendo la última técnica la más eficiente para extraer las características.

En Elrefaei *et al.* (2015) puntualizan sobre el método actual de lectura de un medidor de electricidad ubicado en Arabia Saudita, se realiza de forma manual, por lo que es propenso a errores. A este planteamiento, los autores ofrecen un sistema basado en procesamiento digital de imágenes, que consta de tres partes; el preprocesamiento de la imagen que determina el recorte del área de lectura numérica, la segmentación de dígitos individuales mediante el escaneo horizontal y vertical del área numérica recortada, y el reconocimiento de la lectura al comparar cada dígito segmentado con las plantillas de dígitos. Los resultados obtenidos demuestran un porcentaje de 85.71% de precisión sobre la lectura del medidor.

También hay en la literatura algunos trabajos relacionados con el procesamiento digital de imágenes. Una de las investigaciones plantea el desarrollo de un software para detectar objetos en movimiento, a partir

de una secuencia de imágenes y efectuar un seguimiento de trayectoria lineal, utilizando algoritmos de escala de grises, filtro de la mediana, binarización, operadores de Sobel, resta de imágenes y extracción de características de objetos (Toscano 2009). Posteriormente, Pérez (2012) aborda la extracción de características de objetos dentro de un ambiente controlado, el uso del clasificador *K-means* y momentos invariantes de *Hu*. Más adelante, Pereyra (2011) habla sobre un sistema de reconocimiento de rostros con imágenes de expresiones faciales no invariantes, aplicando técnicas de procesamiento digital de imágenes como: escala de grises, ecualización del histograma, umbralización, filtros: pasa bajo, pasa alto, sobel y prewitt, operadores: de Roberts, de Frei-Chen y Laplaciano, y para el entrenamiento utilizó la Red Neuronal Perceptrón con cuatro neuronas.

Sin embargo, hasta la fecha no ha existido un enfoque utilitario económico de este tipo de análisis. El cultivo de papaya en México se ha posicionado en el primer puesto en producción entre 2017 y 2018, siendo Oaxaca uno de los principales estados que contribuyeron con este logro (SADER 2017, SAGARPA 2018). No obstante, es importante mencionar que la pérdida de ganancias en la producción es a causa de las enfermedades que atacan tanto a las plantas como a los frutos de papaya.

Por tanto, como una forma de apoyo a los productores de papaya, se pretende implementar una herramienta que permita

disminuir costos de producción y contribuir a la región Costa con procesos tecnológicos, haciendo uso de los dispositivos móviles. El objetivo principal de la presente investigación es plantear una propuesta de solución que permita determinar si el fruto de papaya se encuentra enfermo con el virus de la mancha anular, a partir de una imagen digital con una aplicación móvil para Android.

Materiales y métodos

En la figura 1 se muestra la metodología propuesta para la identificación de virus de la mancha anular en frutos de papaya, en seguida se explica la forma de desarrollo.

Adquisición de la imagen. Para la adquisición de la imagen se utilizará la cámara digital que tiene el dispositivo móvil, el formato aceptado es JPG y se captura a una distancia aproximada de 20 cm entre el celular y el fruto de papaya, para capturar el fruto completo. Es importante señalar que se han tomado imágenes de frutos sin afectaciones e imágenes de frutos que tienen visiblemente el virus de la mancha anular. Las muestras se tomaron como parte del desarrollo de un proyecto titulado “Estandarización preliminar de la identificación de agentes virales en plantas de papaya de la región costa de Oaxaca con técnicas moleculares y procesamiento digital de imágenes” en el cual participan los autores del presente artículo.

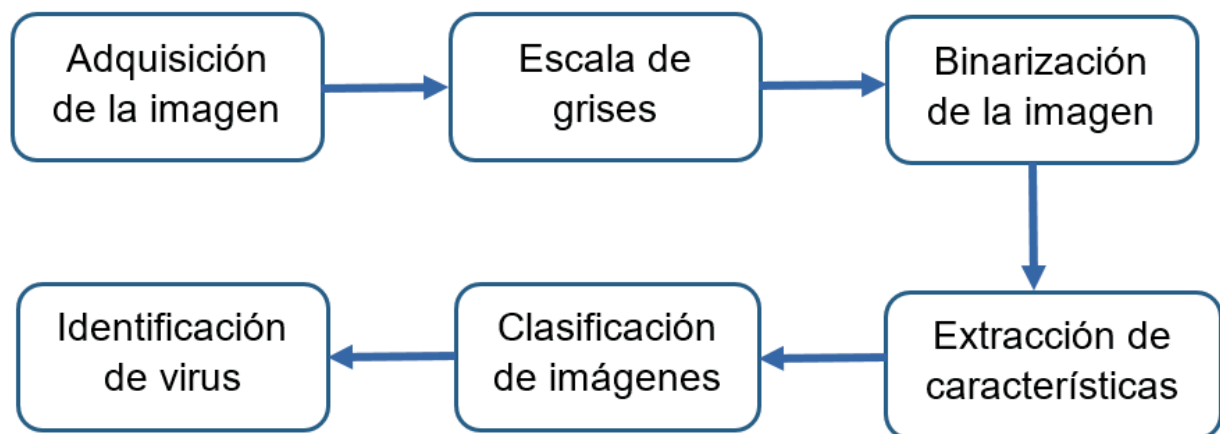


Figura 1. Propuesta de metodología de solución para la identificación del virus de la mancha anular en frutos de papaya.

Escala de grises. Una vez que se obtenga la imagen, se le aplica el algoritmo de conversión del sistema RGB a escala de grises con la ecuación 1:

$$f(x, y) = \frac{(r(x,y))+(g(x,y))+(b(x,y))}{3}$$

Donde:

$f(x, y)$ es la coordenada de cada pixel de la imagen que tomará el nuevo valor de la escala de grises.

$r(x, y)$ corresponde al valor del color rojo (red) en la coordenada indicada por (x, y) .

$g(x, y)$ es el valor del color verde (green) en la coordenada indicada por (x, y) .

$b(x, y)$ corresponde al valor del color azul (blue) en la coordenada indicada por (x, y) .

Binarización de la imagen: A la imagen resultante en el paso anterior se le aplica el algoritmo de Binarización (ecuación 2) con un valor de umbralización definido por el símbolo u :

Ec. 2: $f(x,y) \geq u$ entonces $p=255$

$f(x,y) < u$ entonces $p=0$

Donde:

$f(x, y)$ es la coordenada de cada pixel de la imagen

p es el nuevo valor que se le asignará a $f(x, y)$.

Extracción de características: Después de aplicar los algoritmos de preprocesamiento a la imagen, se procede a realizar la transformación de Hough, específicamente para hacer detección de manchas circulares en el fruto de papaya, característica común del virus de la mancha anular.

Clasificación de imágenes: Para hacer el entrenamiento del sistema que se encargará de hacer la identificación del virus, se tiene que hacer una clasificación con base en sus características, para agrupar los frutos que no están enfermos y los frutos que sí están enfermos,

para lograr dicho objetivo se utilizará la técnica Máquinas de Vectores de Soporte (SVM - Support Vector Machine).

Identificación de virus: Una vez que el sistema se ha entrenado con las imágenes de frutos infectados y frutos no infectados, se puede hacer la identificación del virus en nuevas imágenes que se adquieran por medio del dispositivo móvil. La identificación supone hacer todos los procesos anteriores hasta la extracción de características, y se hace la clasificación de la nueva imagen en una de las dos clases del proceso anterior, utilizando el modelo previamente entrenado.

Discusión

Se presenta una metodología como propuesta de solución para lograr el desarrollo de una aplicación móvil como herramienta para la detección de enfermedades en frutos de papaya. Primero, se obtendrán imágenes de los frutos para aplicarles los algoritmos y técnicas descritas anteriormente con la finalidad de extraer sus características, con lo cual se creará un archivo que contenga dicha información. Posteriormente, las características extraídas se ingresarán a una SVM para el entrenamiento y clasificación de objetos. Una vez construida la base de conocimiento (modelo entrenado) se exportará al dispositivo móvil, al ingresar una nueva imagen para detectar la enfermedad de un fruto de papaya en la aplicación, se le aplicarán las etapas de preprocesamiento, segmentación y extracción de características, para que se compare con la información que contiene el modelo entrenado y obtener el resultado. Para el desarrollo del sistema se utilizará el IDE Android Studio, con librerías OpenCV y lenguaje de programación Java para la aplicación móvil, para las pruebas en escritorio se utilizará el IDE NetBeans.

A nivel mundial, México es uno de los principales exportadores de papaya, y entre los estados que tienen mayor aporte al cultivo se encuentra Oaxaca. Puntualmente, la Costa se posiciona entre las cinco regiones primordiales que cultivan la papaya. Sin embargo, a pesar del trabajo y dedicación de los productores, se

generan pérdidas monetarias causadas por las enfermedades que se presentan en frutos de papaya. Además, existen ocasiones en las que el cultivo se pierde parcial o totalmente por la detección tardía de enfermedades en los frutos, razón por la cual ofrecer a los productores una herramienta que ayude y facilite la detección de dichas enfermedades es de gran importancia.

Con base en lo anterior, se plantea una metodología como propuesta de solución para obtener un resultado en la detección exitosa de enfermedades en frutos de papaya, que a su vez beneficie en la toma de decisiones a los productores y ello les permita disminuir la pérdida de sus cultivos.

Conclusión

Con base en la investigación realizada y teniendo como base los reportes encontrados en la literatura, se considera que la metodología de solución que se propone en la presente investigación, puede ser útil para la identificación de virus en papaya, específicamente, el virus de la mancha anular. La solución que se propone no se ha realizado aún, de acuerdo con la búsqueda que se realizó, en la que se buscó la posible existencia de un software que permitiera hacer la identificación de virus de la mancha anular en frutos de papaya, búsqueda que no dio resultados a favor. Como se puede observar en la introducción del presente documento, se han encontrado algunas propuestas de solución que prometen estar cerca de la solución que aquí se plantea, sin embargo, el ámbito de aplicación es diferente. Debido a lo expuesto anteriormente, se considera que la metodología que se propone en el presente artículo es viable.

Agradecimientos

Agradecemos las observaciones y recomendaciones realizadas por un revisor anónimo para mejorar el presente documento.

Referencias

- Andresen, S. 2002. John McCarthy: padre de la IA. *IEEE Intelligent Systems*, 17(5): 84-85. doi:10.1109/MIS.2002.1039837.
- Deitel, P. & H. Deitel. 2016. *Cómo programar Java*, Pearson Educación (9ª Edición), ISBN: 6073238029.
- Elrefaei, L., A. Bajaber, S. Natheir, N. AbuSanab, & M. Bazi. 2015. Automatic Electricity Meter Reading Based on Image Processing. Conferencia IEEE, Universidad King Abdulaziz, Universidad de Shoubra Benha, Departamento de Ciencias de la Computación Facultad de Informática y Tecnología de la Información y Departamento de Ingeniería Eléctrica Facultad de Ingeniería, Egipto.
- Ganascia, J.G. 2019. UNESCO. Consultado el 4 de abril de 2019: <https://es.unesco.org/courier/2018-3/inteligencia-artificial-mito-y-realidad>.
- Gonzalez, R., & R. Woods. 2001. *Digital Image Processing* (2da ed.). Nueva Jersey: Prentice Hall.
- Kisan, A., A. Priyanka, & S. Kisan. 2017. Detection, Categorization and suggestion to cure infected plants of tomato and grapes by using OpenCV framework for android environment. 2nd International Conference for Convergence in Technology (I2CT). Mumbai. DOI: 10.1109/I2CT.2017.8226270.
- M. Minsky, 1961. Steps toward Artificial Intelligence, in *Proceedings of the IRE*, vol. 49, no. 1, pp. 8-30. DOI: 10.1109/JRPROC.1961.287775.
- OpenCV. 2019. OpenCV library. Consultado el 8 de abril de 2019: <https://opencv.org/>
- Oracle. 2019. Consultado el 30 de mayo de 2019: <https://www.java.com>
- Patel, J., & N. Gamit. 2016. A Review on Feature Extraction Techniques in Content Based Image Retrieval. Conferencia, Universidad de Uka Tarsadia, Departamento de Ingeniería Informática, Bardoli, India.
- Pereyra, C. 2011. Reconocimiento de rostros en un ambiente de iluminación controlado no invariante a expresiones faciales. Tesis, Universidad del Mar, Puerto Escondido, México.
- Pérez, C. 2012. Clasificación de objetos rígidos a partir de imágenes digitales, empleando los momentos invariantes de HU. Tesis, Universidad del Mar, Puerto Escondido, México.
- SADER. 2017. Delegación SADER Oaxaca. Consultado el 29 de marzo de 2019: <https://www.gob.mx/sader/oaxaca/articulos/oaxaca-el-principal-productor-de-papaya-a-nivel-nacional?idiom=es>
- SAGARPA. 2018. Quadratin Oaxaca. Consultado el 29 de marzo de 2019: <https://oaxaca.quadratin.com.mx/oaxaca-principal-productor-de-papaya-a-nivel-nacional-sagarpa/>

Singh, S., A. Kaul, U. Pratap, & S. Jain. 2018. Bacterial Foraging Optimization Based Radial Basis Function Neural Network (BRBFNN) for Identification and Classification of Plant Leaf Diseases: An Automatic Approach Towards Plant Pathology, in *IEEE Access*, vol. 6, pp. 8852-8863. DOI: 10.1109/ACCESS.2018.280068

Toscano, J. 2009. Detección y seguimiento lineal de objetos no flexibles invariantes en color en una secuencia de imágenes. Tesis, Universidad del Mar, Puerto Escondido, México.