

# Método segmentando basado en los valores de los canales de color del modelo RGB.

Juan Pablo Guerra Ibarra  
Centro de Investigaciones en Óptica  
A.C., León Guanajuato 37150, México  
[juangi@cio.mx](mailto:juangi@cio.mx)

Francisco Javier Cuevas de la Rosa  
Centro de Investigaciones en Óptica,  
A.C., León Guanajuato 37150, México  
[fjcuevas@cio.mx](mailto:fjcuevas@cio.mx)

**Abstract—** La agricultura fue, es y será una actividad básica para la supervivencia y desarrollo del ser humano, por tal motivo, siempre ha sido potenciada por medio de desarrollos tecnológicos, lo que ha mejorado su productividad y eficiencia. En la actualidad se desarrollan sistemas computacionales que usan algoritmos para apoyar a diversos procesos agrícolas. Este tipo de sistemas tienen en común, la necesidad de ser alimentados con información filtrada o procesada, preservando las características específicas y distintivas de los objetos. En las siguientes cuartillas se presenta un método de segmentación de imágenes con cultivos de jitomate en ellas, cuyo objetivo es el separar los elementos presentes en la imagen en tres clases: hojas, frutos y fondo. El método se basa en la superioridad numérica de uno de los canales de color sobre los otros, en el modelo de representación de color RGB, este tipo de técnicas tiene entre sus ventajas la facilidad de implementación y el bajo costo computacional de implementarla, en comparación como otros métodos o técnicas, como lo son las de Deep learning, ya que estas requieren de tiempo para el etiquetado de datos y su posterior entrenamiento del modelo. Es importante el desarrollo y perfeccionamiento de los métodos de segmentado de imágenes de bajo costo computacional y tiempo de respuesta cortos, los cuales puedan ser ejecutados en hardware de características limitadas, que son accesibles en los campos agrícolas. Los resultados de métodos de segmentado serán base para sistemas de apoyo para la optimización a los procesos agrícolas, en lo que se denomina agricultura de precisión.

**Keywords—** Agricultura, Segmentación de imágenes, Umbrales, Redes neuronales convolucionales, RGB

## I. INTRODUCCIÓN

La agricultura es una actividad básica para el ser humano, esta se ha desarrollado a la par de las diferentes civilizaciones a lo largo de la historia [1][2], teniendo un rol fundamental para su sustento alimenticio. Se estima que en el año 2050, el número de seres humanos a nivel mundial será de cerca de diez billones de personas, según los datos de la Organización de las Naciones Unidas [3], por lo tanto la producción de alimentos en cantidad y calidad debe incrementarse. La agricultura tendrá gran importancia para la continuidad y desarrollo de la civilización moderna, ya que la producción de los alimentos requeridos para alimentar al creciente número de la individuos a nivel mundial, depende en gran parte de la agricultura.

Debido a la importancia que tiene la producción de alimentos para las civilizaciones a lo largo de la historia, está siempre se ha visto fortalecida con avances tecnológicos aplicados en la agricultura, por ejemplo: el establecimiento de campos de cultivos fijos, permitió abandonar el estilo de vida nómada, el uso del arado, redujo los tiempos de preparación de la tierra

para la siembra, los sistemas de riego, permite el cultivo de productos fuera de temporada de lluvia, los cultivos bajo techo, aíslan las plantaciones de las condiciones climatológicas, etc. En años recientes se han desarrollado sistemas computacionales que hacen uso de algoritmos para mejorar diferentes procesos agrícolas, por ejemplo: el riego, fertilización, detección de enfermedades y carencias nutricionales [4][2][5][6][7]. La primera etapa de este tipo de sistemas consiste en proveerles de información lo más filtrada de ruido, tanto como sea posible, sin eliminar las características distintivas de la misma, lo anterior independientemente de cuál sea su origen de los datos, ya sean de imágenes, datos estadísticos o de otro tipo.

En el presente trabajo se describe la primera etapa de un método de segmentado aplicado a imágenes de cultivos de jitomate en agricultura protegida, el objetivo es agrupar los píxeles de la imagen en tres clases: hojas, frutos y fondo. Las imágenes son manejadas en el modelo de representación de color de tres canales: rojo, verde y azul o RGB (*Red, Green, Blue*) por sus siglas en inglés. El método de segmentado presentado se basa en la superioridad numérica de alguno de los canales de color del modelo RGB para una determinada clase, el canal verde para las hojas y el canal rojo para los frutos.

Las ventajas del método propuesto en las siguientes cuartillas son: la facilidad de implementación, rapidez a la hora de realizar el proceso de segmentado, no requiere de un etiquetado previo de las imágenes y no requiere de un hardware especializado para realizar el proceso de segmentado.

## II. ANTECEDENTES

Existen en la literatura gran diversidad de aplicaciones para los algoritmos o métodos de segmentación de imágenes, estas aplicaciones se dan en diferentes ámbitos, pasando por aplicaciones industriales, entre ellas la agrícola.

El objetivo de cualquier método de segmentado es la separación de los elementos de interés de los que no lo son. Los métodos de segmentación por color han sido aplicados hace ya varios años. Los distintos formatos o modelos de representación de color que existen, son usados para manejar numéricamente las características que apreciamos con la vista en las imágenes, dichos formatos poseen características propias que son usadas para un fin determinado. El color es uno de los atributos más usado para discriminar el follaje y frutos del resto de la imagen.

En años recientes el desarrollo de nuevo hardware ha permitido realizar algoritmos computacionales más robustos, como lo son los de Deep learning, que pueden ser usados

#### A. Estado del arte.

En el trabajo [8] se propone un método para segmentar plantas de arroz del fondo de la imagen, se basa en la resta del valor del canal verde menos el valor del canal rojo para los píxeles de la imagen, el resultado de la resta es comparado contra un umbral, establecido basándose en conocimiento a priori. Los resultados del proceso de segmentación son usados para estimar la cantidad de nitrógeno presente en las hojas. Un trabajo en el que se utilizan umbrales de color para determinar el crecimiento de cultivos es planteado en [5], se utiliza el sistema de representación de color en formato RGB y HSL (*Hue*, *Saturation*, *Lightness*) por sus siglas en inglés, se logra extraer texturas, morfologías de los cultivos y déficit de calcio. El trabajo [9], se procesan imágenes tomadas con cámaras hiperespectrales para la estimación de nitrógeno presente en cultivos de maíz. Los investigadores del trabajo [10] estiman la cantidad de vegetación presente en tomas de cultivos de cereales en etapas tempranas de desarrollo fenológico, un trabajo similar es presentado en [11], este tiene el objetivo de evaluar la cantidad de desarrollo foliar en campos limpios de maleza, se reporta la importancia de las condiciones con las que se adquieren las imágenes. Adicionalmente existen trabajos que detectan malas hierbas en condiciones de luz no controladas [12], mediante el uso de un robot autónomo para la adquisición de imágenes. En [13] miden la cantidad de clorofila que se encuentra en cultivos de papa. El trabajo [14] se realizan comparaciones de segmentación usando diferentes modelos de representación de color. Los trabajos mencionados aplican alguna técnica de segmentación basados en el color presente en las imágenes.

Las redes neuronales convolucionales pertenecen al área de conocimiento llamada “Deep learning”; las cuales son una rama de algoritmos de procesamiento de imágenes de reciente aparición [4]. En el Deep learning se utilizan funciones que permiten la representación jerárquica de la información entre varias capas de profundidad para realizar reconocimiento y clasificación [15].

Las aplicaciones de las redes neuronales convolucionales en la agricultura cubren una amplia variedad de tareas. La principal ventaja que brinda este tipo de algoritmos es su invarianza a datos con gran cantidad de ruido. En [16] se muestra la aplicación de un modelo de red convolucional llamado ConNet, se usa para clasificar varios tipos de hojas en ambientes no controlados de iluminación.

La segmentación realizada por algoritmos de deep learning se denomina segmentación semántica, este tipo de segmentación consiste en asignar cada píxel dentro de una imagen a una clase con la que fue entrenado el modelo [20]. Los algoritmos de deep learning requieren de un poder de cálculo computacional que no es accesible para todos, lo que limita su uso.

En La búsqueda bibliográfica realizada, no se encontró referencia a la posible segmentación que se pueda llevar por medio de la comparación directa de los canales del RGB, para separar los píxeles de la imagen en las clases de hojas, frutos y fondo en imágenes con cultivos de jitomate en ellas.

### III. DESCRIPCIÓN DEL MÉTODO PROPUESTO PARA SEGMENTACIÓN POR SUPERIORIDAD DE COLOR

Los métodos o algoritmos de segmentación de imágenes, evalúan los píxeles que componen una imagen en busca de

Las características que se deben de tomar en cuenta a la hora de trabajar con imágenes, al realizar un proceso de segmentación, entre otras están las siguientes:

- El formato de archivo, es la manera en la que se almacena información de la imagen en un dispositivo de almacenamiento.
- Una imagen se puede representar como una función bidimensional  $f(x,y)$ , donde el valor de la función representa el color en la posición  $x,y$ .
- Dimensiones de la imagen,  $N \times M$  en donde  $N$  y  $M$  representan el ancho y alto de la imagen.
- Formato o modelo de color, que es la forma representación del espectro de color visible.
- El formato RGB, crea los colores basándose en una combinación lineal de los valores de los colores primarios rojo, verde y azul, los valores están en el rango de 0 - 255.

#### A. Segmentación basada en superioridad de canales de color.

En la Figura 1 se observa una imagen de ejemplo con las que el método de segmentación que se propone será desarrollado y probado. Las imágenes fueron adquiridas en invernaderos con plantas de jitomate cultivándose en su interior, dentro de la toma se encuentran las tres clases que nos interesa segmentar: hojas, frutos y fondo.



Fig. 1: Ejemplo de imágenes a segmentar.

Es importante hacer mención que las imágenes fueron tomadas sin considerar cuestiones de iluminación o mayor cuidado o preparación al momento de realizar la toma de las mismas, esto con la finalidad de imitar las condiciones en las que un agricultor tomaría una fotografía, con un dispositivo a su que tenga a su alcance durante su jornada de trabajo.

En la Figura 2 se muestran las etapas del método de segmentado propuesto, este se basa en la superioridad de una las bandas de color del formato RGB, para segmentar las hojas y frutos de cultivos de jitomate en las imágenes a que serán procesadas.

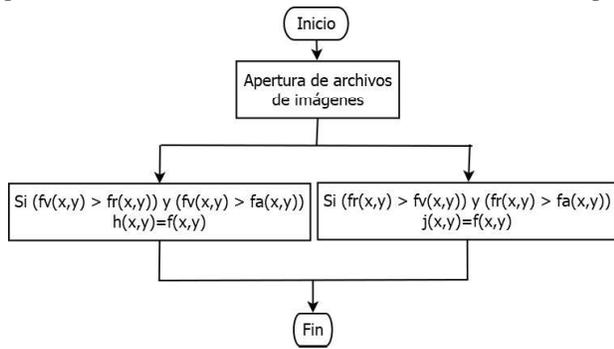


Fig. 2: Método de segmentado por superioridad de color de RGB.

Las variables involucradas en la Figura 2, que son usadas en el proceso de segmentación de las hojas y frutos se describen a continuación:

- La función bidimensional  $f(x,y)$  representa la imagen a ser segmentada.
- $x,y$  representan la posición de un píxel en la imagen.
- $f_r(x,y), f_v(x,y), f_a(x,y)$  contienen la información con los valores numéricos de los colores primarios rojo, verde y azul respectivamente de los píxeles de la imagen, los valores están en el rango de 0 a 255.
- Las funciones bidimensionales  $h(x,y)$  y  $j(x,y)$ , son funciones resultantes de los procesos de segmentación para las hojas y frutos respectivamente.

El conjunto de imágenes con las que se prueba el método de segmentación descrito en la Figura 2, tienen las siguientes características:

- Formato PNG.
- Dimensiones de las imágenes ( $N \times M$ ):
  - o 400 x 500 píxeles.
  - o 500 x 400 píxeles.
- 10 archivos de imágenes de prueba.

La lectura de los valores numéricos de los píxeles de las imágenes de prueba en sistema de representación de color RGB, se realiza utilizando la librería de “OpenCV” [23], la cual es una librería ampliamente usada y probada para el desarrollo de sistemas de visión por computadora. Es pertinente hacer mención de que no se realiza ningún procesamiento con funciones propias de “OpenCV”, salvo la lectura de los datos numéricos que forman la imagen a segmentar.

El objetivo del método de segmentación propuesto, es separar los píxeles que conforman las hojas y los frutos en la imagen. La separación de los píxeles que forman parte de las hojas representada por la función  $f(x,y)$ , se lleva a cabo aplicando la ecuación 1.

$$h(x,y) = \begin{cases} f(x,y), & \text{Si } f_v(x,y) > f_r(x,y) \text{ y } f_v(x,y) > f_a(x,y) \\ 0, & \text{en otro caso} \end{cases} \quad (1)$$

Para segmentar los píxeles de las hojas, es de esperar que el componente  $f_v(x,y)$  posean un valor numérico mayor en comparación con  $f_r(x,y)$  y  $f_a(x,y)$ , esta característica de

superioridad es expresada en la ecuación 1. Los píxeles que cumplan con la condición expresada en la ecuación 1, serán etiquetados como hojas.

Los resultados de aplicar la ecuación 1 sobre la función  $f(x,y)$  que representa la imagen de la Figura 1 se observan en la Figura 3.



Fig. 3: Segmentación resultante de aplicar la ecuación 2 a la Figura 1.

La segmentación de los frutos de los cultivos jitomate del resto de la imagen, se basa en la premisa de que el canal  $f_r(x,y)$  tendrá un valor numérico mayor en comparación con los canales  $f_v(x,y)$  y  $f_a(x,y)$ , esta superioridad se expresa en la ecuación 2.

$$j(x,y) = \begin{cases} f(x,y), & \text{Si } f_r(x,y) > f_v(x,y) \text{ y } f_r(x,y) > f_a(x,y) \\ 0, & \text{en otro caso} \end{cases} \quad (2)$$

Al aplicar la ecuación 2 en la función  $f(x,y)$  que representa la Figura 1, se genera el etiquetado de los píxeles que conforman los frutos, observe en la Figura 4.



Fig. 4: Segmentación resultante de aplicar la ecuación 2 a la Figura 1.

Con los elementos del método descritos y la aplicación de las ecuaciones 1 y 2, se puede llevar a cabo una segmentación de

las hojas y frutos de cultivos jitomates los cuales estén presentes en la imagen representada por la función  $f(x,y)$ .

IV. RESULTADOS

El método de segmentado descrito para etiquetar los píxeles de las hojas y frutos de cultivos de jitomate, fue aplicado un conjunto de imágenes de prueba, en las Figuras 5, 6 y 7 muestran los resultados de la aplicar las ecuaciones 1 y 2 en las funciones  $f(x,y)$  de cada imagen.

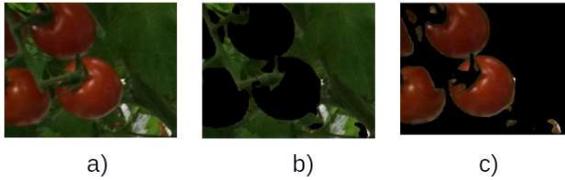


Fig. 5: a) Original b) Segmentado de hojas c) Segmentado de frutos.

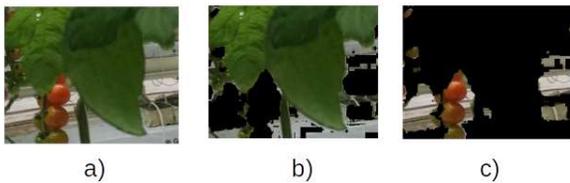


Fig. 6: a) Original b) Segmentado de hojas c) Segmentado de frutos.

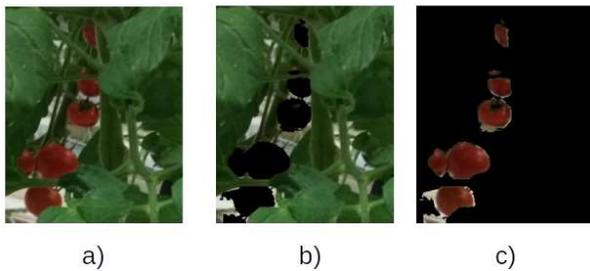


Fig. 7: a) Original b) Segmentado de hojas c) Segmentado de frutos.

Es importante el poder definir de manera cuantitativa la efectividad del método de segmentado propuesto, las métricas seleccionadas para esto son: “Intersección sobre Unión” y “Precisión”. Para poder utilizar las métricas mencionadas es necesario contar con una imagen en la cual se etiquetará de forma manual los píxeles que corresponden a las clases de hojas, frutos y fondo.

La máscara correspondiente a la Figura 1, se observa en la Figura 5, las hojas son etiquetadas de color verde, mientras los frutos son etiquetados con color rojo.

Las máscaras de las imágenes del conjunto de prueba fueron construidas usando la aplicación “Computer Vision Annotation Tool” (<https://www.cvat.ai/>), la cual es una herramienta web gratuita.

La métrica de “Intersección sobre Unión” (IoU) [24], es expresada en la ecuación 3.

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (3)$$



Fig. 8: Máscara de la Figura 1 etiquetada en CVAT.

En la ecuación 3, “A” representa la imagen generada por el método de segmentado, ya de las hojas o frutos, “B” representa la imagen de la máscara con el etiquetado manual de la imagen “A”.  $|A \cap B|$  es la cardinalidad del conjunto formado por intersección entre las imágenes “A” y “B”, es decir la cantidad píxeles que fueron etiquetados como hoja o frutos por el método de segmentado según corresponda y de igual modo el píxel en la máscara fueron etiquetados como hojas o frutos.  $|A \cup B|$  es la cardinalidad del conjunto formado por la unión de las imágenes “A” y “B”, es decir los píxeles que fueron etiquetados como hoja o frutos por el método de segmentado y en la máscara fueron etiquetados como hojas o frutos en la máscara.

La métrica de “Precisión” se calcula de por medio de la ecuación 4.

$$Precisión = \frac{VP}{VP + FP} \quad (4)$$

En la ecuación 4, se usan los términos VP y FP para determinar la eficacia del método. VP corresponde a los verdaderos positivos, es decir, los píxeles que el método de segmentado los etiqueto como hoja o fruto y de igual manera en la máscara de la imagen. FP son los falsos positivos, es decir, los píxeles que el método de segmentado etiqueto erróneamente como hoja o fruto en comparación con lo que se etiqueto en la máscara.

El método de segmentado propuesto es aplicado a un conjunto de diez imágenes de prueba. Los resultados de la segmentación son medidos con las métricas IoU y Precisión, la Tabla 1 contiene los resultados obtenidos al segmentar las hojas y mientras que el la Tabla 2 se encuentran los resultados de la segmentación de los frutos.

Tabla 1 Métricas de desempeño del método de segmentado para las hojas.

No	Imagen	Máscara	Hojas	IoU	Psc
1				0.82	0.82

2				0.88	0.92
3				0.96	0.96
4				0.65	0.65
5				0.90	0.90
6				0.80	0.81
7				0.87	0.88
8				0.58	0.59
9				0.66	0.66
10				0.72	0.72

Tabla 2 Métricas de desempeño del método de segmentado para los frutos.

No	Imagen	Máscara	Frutos	IoU	Pcs
1				0.69	0.69
2				0.85	0.85
3				0.32	0.32
4				0.63	0.63
5				0.74	0.74
6				0.78	0.78

7				0.69	0.69
8				0.84	0.85
9				0.76	0.76
10				0.74	0.74

Los experimentos de segmentado con las imágenes del conjunto de prueba fueron ejecutados una computadora con las siguientes características de hardware:

- Intel(R) Core(TM) i7-10750H CPU @ 2.60GHz
- 16 GB de memoria RAM

Los tiempos de segmentado por imagen están alrededor de los 0.15 segundos.

#### V. CONCLUSIONES

El método de segmentado propuesto para separar las hojas y frutos en cultivos de jitomate, poseen un desempeño aceptable, de acuerdo con los datos mostrados en la Tabla 1 y Tabla 2, con las diez imágenes del conjunto de prueba, es importante mencionar que se generan errores al etiquetar píxeles cercanos al color blanco.

Si bien los resultados en términos generales pueden considerarse como buenos, en las imágenes 4, 8 y 9 en la Tabla 1 y en las imágenes 1, 3, y 4 en la Tabla 2, muestran una marcada disminución en las métricas de desempeño. El decremento en el desempeño es debido a la presencia de píxeles similares al color blanco presentes en el fondo de la toma; el color blanco es formado con la combinación lineal de valores en los canales rojo, verde y azul cercanos al máximo numérico de 255, este tipo de combinaciones pueden llevar generara errores al etiquetar píxeles como hojas o frutos cuando no lo son en realidad, debido a que la diferencia entre los canales de color no es muy grande, lo representa en una clara limitante del método propuesto.

Las métricas *IoU* y *Precisión* pueden reportar resultados con bajos niveles de eficiencia, no únicamente debido al resultado de la segmentación del hecha por método propuesto; también un deficiente etiquetado puede afectar considerablemente los resultados de las métricas desempeño.

El método de segmentado propuesto posee algunas ventajas evidentes, la primera de ellas es la simplicidad de implementación, ya que se puede implementarse con unas cuantas líneas de código en cualquier lenguaje de programación de propósito general. La segunda ventaja se deriva de la simplicidad de implementación, que es la velocidad de ejecución al momento de segmentar las imágenes. Otra ventaja es que es un método directo, al no requerir de realizar etiquetados previos para llevar a cabo un proceso de entrenamiento o necesitar un hardware especializado.

Si bien el desempeño del método de segmentado presenta buenos resultados en imágenes con muchas hojas, existen otras imágenes en las cuales se etiquetan píxeles como hojas o frutos que evidentemente no pertenecen a esta clase, por lo tanto, es necesario continuar desarrollando el método de segmentado implementando alguna característica que permita un mejor filtrado.

Una opción para complementar el método de segmentado propuesto, pasa por el uso de la estadística. Se usará como una segunda característica distintiva, la desviación estándar de las diferencias entre los canales dominantes  $f_R(x,y)$  para los frutos y  $f_V(x,y)$  para las hojas, contra los dos canales restantes respectivamente; las desviaciones estándar se usaran como un posible umbral, para una segunda etapa de segmentado.

En otra etapa del proyecto se implementará un método de segmentación semántica, que use algún modelo de red neuronal convolucional. Con los resultados de los métodos de segmentado por umbrales y los de segmentado semántico, se realizara una comparativa entre los resultados de ambos, permitiendo ver las ventajas y desventajas de cada uno de los métodos.

Mejorar el etiquetado de las imágenes que componen el conjunto de prueba permitiría tener evaluaciones más confiables, buscando disminuir el error generado por un mal etiquetado de las imágenes.

#### VII. AGRADECIMIENTOS

Es importante dar reconocimiento a las Instituciones y Empresas que hicieron posible la realización de esta investigación al brindar apoyo financiero y acceso a sus instalaciones.

- Consejo Nacional de Ciencia y Tecnología (CONACYT).
- Centro de Investigaciones en Óptica A.C.
- Instituto Tecnológico de Estudios Superiores de Zamora,
- Empresa Opus Farms.

#### REFERENCIAS

[1] Y. Awasthi, "Press 'a' for artificial intelligence in agriculture: A review," *Int. J. Informatics Vis.*, vol. 4, no. 3, pp. 112–116, 2020, doi: 10.30630/ijov.4.3.387.

[2] P. P. Ray, "Internet of things for smart agriculture: Technologies, practices and future direction," *J. Ambient Intell. Smart Environ.*, vol. 9, no. 4, pp. 395–420, 2017, doi: 10.3233/AIS-170440.

[3] FAOSTAT, "FAOSTAT." <http://www.fao.org/faostat/es/#data/QC>.

[4] A. Kamilaris and F. X. Prenafeta-Boldú, "Deep learning in agriculture: A survey," *Comput. Electron. Agric.*, vol. 147, no. July 2017, pp. 70–90, Apr. 2018, doi: 10.1016/j.compag.2018.02.016.

[5] D. Story, M. Kacira, C. Kubota, A. Akoglu, and L. An, "Lettuce calcium deficiency detection with machine vision computed plant features in controlled environments," *Comput. Electron. Agric.*, vol. 74, no. 2, pp. 238–243, 2010, doi: 10.1016/j.compag.2010.08.010.

[6] L. Chen *et al.*, "Identification of nitrogen, phosphorus, and

potassium deficiencies in rice based on static scanning technology and hierarchical identification method," *PLoS One*, vol. 9, no. 11, pp. 1–17, 2014, doi: 10.1371/journal.pone.0113200.

[7] R. González Perea, E. Camacho Poyato, P. Montesinos, and J. A. Rodríguez Díaz, "Optimization of Irrigation Scheduling Using Soil Water Balance and Genetic Algorithms," *Water Resour. Manag.*, vol. 30, no. 8, pp. 2815–2830, 2016, doi: 10.1007/s11269-016-1325-7.

[8] Y. Wang, D. Wang, G. Zhang, and J. Wang, "Estimating nitrogen status of rice using the image segmentation of G-R thresholding method," *F. Crop. Res.*, vol. 149, pp. 33–39, 2013, doi: 10.1016/j.fcr.2013.04.007.

[9] M. Quemada, J. L. Gabriel, and P. Zarco-Tejada, "Airborne hyperspectral images and ground-level optical sensors as assessment tools for maize nitrogen fertilization," *Remote Sens.*, vol. 6, no. 4, pp. 2940–2962, 2014, doi: 10.3390/rs6042940.

[10] K. Kirk, H. J. Andersen, A. G. Thomsen, J. R. Jørgensen, and R. N. Jørgensen, "Estimation of leaf area index in cereal crops using red-green images," *Biosyst. Eng.*, vol. 104, no. 3, pp. 308–317, 2009, doi: 10.1016/j.biosystemseng.2009.07.001.

[11] J. RASMUSSEN, M. NØRREMARK, and B. M. BIBBY, "Assessment of leaf cover and crop soil cover in weed harrowing research using digital images," *Weed Res.*, vol. 47, no. 4, pp. 299–310, Aug. 2007, doi: 10.1111/j.1365-3180.2007.00565.x.

[12] H. Y. Jeon, L. F. Tian, and H. Zhu, "Robust crop and weed segmentation under uncontrolled outdoor illumination," *Sensors*, vol. 11, no. 6, pp. 6270–6283, 2011, doi: 10.3390/s110606270.

[13] S. P. Yadav, Y. Ibaraki, and S. D. Gupta, "Estimation of the chlorophyll content of micropropagated potato plants using RGB based image analysis," *Plant Cell. Tissue Organ Cult.*, vol. 100, no. 2, pp. 183–188, 2010, doi: 10.1007/s11240-009-9635-6.

[14] I. Philipp and T. Rath, "Improving plant discrimination in image processing by use of different colour space transformations," *Comput. Electron. Agric.*, vol. 35, no. 1, pp. 1–15, 2002, doi: 10.1016/S0168-1699(02)00050-9.

[15] J. Schmidhuber, "Deep Learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015, doi: 10.1016/j.neunet.2014.09.003.

[16] D. Hall, C. McCool, F. Dayoub, N. Sünderhauf, and B. Upcroft, "Evaluation of features for leaf classification in challenging conditions," *Proc. - 2015 IEEE Winter Conf. Appl. Comput. Vision, WACV 2015*, pp. 797–804, 2015, doi: 10.1109/WACV.2015.111.

[17] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification," *Comput. Intell. Neurosci.*, vol. 2016, 2016, doi: 10.1155/2016/3289801.

[18] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Front. Plant Sci.*, vol. 7, no. September, pp. 1–10, 2016, doi: 10.3389/fpls.2016.01419.

[19] T. T. Tran, J. W. Choi, T. T. H. Le, and J. W. Kim, "A comparative study of deep CNN in forecasting and classifying the macronutrient deficiencies on development of tomato plant," *Appl. Sci.*, vol. 9, no. 8, 2019, doi: 10.3390/app9081601.

[20] P. Wang *et al.*, "Understanding Convolution for Semantic Segmentation," *Proc. - 2018 IEEE Winter Conf. Appl. Comput. Vision, WACV 2018*, vol. 2018-Janua, pp. 1451–1460, 2018, doi: 10.1109/WACV.2018.00163.

[21] J. Zhuang, J. Yang, L. Gu, and N. Dvornek, "Fully Convolutional Networks for Semantic Segmentation," *Proc. - 2019 Int. Conf. Comput. Vis. Work. ICCVW 2019*, pp. 847–856, 2019, doi: 10.1109/ICCVW.2019.00113.

[22] H. Noh, S. Hong, and B. Han, "Learning deconvolution network for semantic segmentation," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. 2015 Inter, pp. 1520–1528, 2015, doi: 10.1109/ICCV.2015.178.

[23] OpenCV, "Home - OpenCV," 2022. <https://opencv.org/> (accessed Sep. 11, 2022).

[24] J. Yu *et al.*, "Learning Generalized Intersection Over Union for Dense Pixelwise Prediction," in *Proceedings of the 38th International Conference on Machine Learning*, 2021, vol. 139, pp. 12198–12207, [Online]. Available: <https://proceedings.mlr.press/v139/you21e.html>.