



Fenotipado de plantas de frijol (*Phaseolus Vulgaris*) usando técnicas de visión por computadora

María José Herrera Hernández

Tecnológico Nacional de México / ITS de
Irapuato,
Irapuato, Guanajuato, México

Gilberto Muñoz Moreno

Tecnológico Nacional de México / ITS de
Irapuato,
Irapuato, Guanajuato, México

Adolfo Rafael López Núñez

Tecnológico Nacional de México / ITS de
Irapuato,
Irapuato, Guanajuato, México

Marcos Jesús Villaseñor Aguilar *

Tecnológico Nacional de México / ITS de
Irapuato,
Irapuato, Guanajuato, México

* Autor de correspondencia: marcos.vac@irapuato.tecnm.mx

Resumen: El uso de los sistemas de visión computacional (SVC) en la agricultura es un área de investigación que ha crecido en los últimos años. Se destaca la importancia del desarrollo de herramientas tecnológicas para aumentar la confiabilidad y reducir los costos en la producción agrícola. En este trabajo se describe un método para predecir los días de crecimiento de la planta de frijol empleando el área foliar. Este método relaciona el área foliar de la imagen capturada y predice la fecha con un error absoluto medio de 3.323 Días y una correlación estadística de 0.7720. Para su implementación se utilizó el software Matlab y una cámara comercial de bajo costo.

Palabras clave: agricultura, imágenes, producción, visión computacional.

1. Introducción: lo que debemos saber de inicio

El frijol (*Phaseolus Vulgaris*) es una leguminosa originaria de México, perteneciente al género *Phaseolus* de la familia *Fabaceae* (Flores, 2023). Es un alimento fundamental en la dieta diaria de la población mexicana debido a su alto valor nutricional. Aporta proteínas,

fibra, vitaminas, minerales y antioxidantes, lo que contribuye a una alimentación equilibrada y saludable. En este sentido, diversos estudios han demostrado que el consumo regular de frijoles puede ayudar a prevenir enfermedades cardiovasculares, diabetes tipo 2, cáncer de colon y obesidad, gracias a su bajo índice glucémico y su riqueza en fibra y fitoquímicos (Fernández y Sánchez, 2017).

En términos de producción, México se encuentra entre los principales productores de frijol a nivel mundial. Según datos de la Secretaría de Agricultura y Desarrollo Rural (SADER), en 2023 se cultivaron aproximadamente 723,642 toneladas en el país. Conviene señalar que los principales estados productores a nivel nacional son Zacatecas, Sinaloa, Durango, Chihuahua y Nayarit. Sin embargo, en 2023, la producción nacional de esta leguminosa cayó a 724 mil toneladas, una disminución del 25% respecto al año anterior, debido principalmente a condiciones climatológicas adversas como la sequía (Morales, R, 2024).

Actualmente, los productores utilizan como criterio los rasgos visibles de las plantas como son la forma de las hojas, las dimensiones de la planta, la resistencia a enfermedades y la capacidad de adaptación al medio ambiente, esta técnica es subjetiva y no considera las diferentes variables como las condiciones ambientales y la reproducibilidad (Bautista-Lozada et al., 2012). Con el estudio de estas características se ayuda a establecer estrategias que incrementen la producción. Por ejemplo, se han incorporado los Sistemas de Visión por Computadora (SVC) debido a que facilitan automatizar procesos de producción agrícola. Estos sistemas ofrecen ventajas como son bajo costo, exactitud, incremento de precisión y buena repetibilidad (Villaseñor-Aguilar et al., 2019). Este proyecto consta de seis etapas. La primera fue la siembra y cultivo de la semilla. La segunda consistió en acondicionar el sistema de captura de imágenes. En la tercera se capturo la imagen de la planta. En la cuarta y quinta se emplea el algoritmo que segmenta la plántula y se identifican las diferentes regiones de interés. En la última etapa se elimina el ruido y se determinan las características del cultivo mediante el uso de una regresión lineal simple.

En el presente trabajo se tiene como pregunta de investigación, ¿Cómo se relaciona la cantidad de píxeles correspondientes a la planta de frijol en imágenes capturadas con su tiempo de crecimiento mediante el uso de una regresión lineal simple? Así mismo, con la metodología propuesta se busca estimar los días de crecimiento de la planta considerando su área foliar mediante técnicas de procesamiento digital de imágenes en consecuencia este trabajo contribuirá a reducir los errores agronómicos y pérdidas económicas.

2. Fundamentos Teóricos: reglas y principios científicos importantes

Una imagen digital está formada por una matriz de píxeles (elementos de imagen), dependiendo la posición del píxel le corresponde un valor numérico el cual corresponde a la intensidad o color en una posición específica. Así, mediante el uso de algoritmos computacionales las imágenes digitales pueden ser modificadas. Esta representación permite que las imágenes sean procesadas y analizadas mediante algoritmos computacionales. La segmentación por umbrales de color RGB, según Rafael C. González (2018), es una técnica utilizada en el procesamiento de imágenes para identificar y separar regiones de interés basadas en los valores de los canales Rojo (R), Verde (G) y Azul (B). Este método se basa en establecer umbrales para cada componente de color, permitiendo clasificar los píxeles según su intensidad y generar una imagen segmentada (Correa et al., 2019).

Para este proceso se siguieron estos pasos:

- Binarización de la imagen a su representación en el espacio RGB mostrado en la Figura 1.
- Determinación de los valores de umbral correcto por medio del análisis de histogramas.
- Utilización de umbrales a los canales de color por separado para generar una máscara binaria como se observa en la Figura 2.
- Obtención de una segmentación final por medio de la mezcla de máscaras.



Figura 1. Binarización de la imagen en escala RGB.

Obtenida de: elaboración propia.

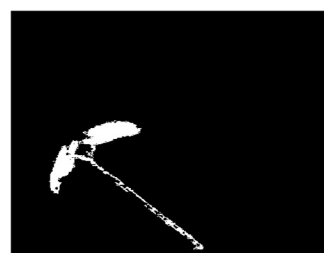


Figura 2. Intersección del umbral de los canales RGB.

Obtenida de: elaboración propia.

Este enfoque es útil en aplicaciones como la detección de objetos en imágenes digitales y la clasificación de regiones con colores específicos. El algoritmo de etiquetado de componentes conectados (*Connected Component Labeling*, CCL) fue fundamental en el procesamiento de las imágenes, ya que permitió identificar las regiones contiguas de píxeles que comparten ciertas características, como lo son la intensidad o el color, tal como se observa en la Figura 3. Este proceso fue esencial para segmentar los objetos y realizar el análisis de las estructuras dentro de una imagen binaria, como se aprecia en la Figura 4. El algoritmo de etiquetado se constituye de dos fases principales. La primera realiza un recorrido en la imagen de izquierda a derecha y de arriba hacia abajo, para cada píxel de interés (por ejemplo, con valor de 1 en una imagen binaria), se examinan sus vecinos ya procesados y si no hay vecinos etiquetados, se asigna una nueva etiqueta al píxel, si no hay vecinos etiquetados, se asigna una etiqueta más baja entre ellos, por último, en esta fase si los vecinos tienen etiquetas distintas, se registra una equivalencia entre estas etiquetas. La segunda fase es la resolución de equivalencias, se crea una tabla de equivalencias para unificar etiquetas que representan la misma región y se realiza un segundo recorrido de la imagen digital, se remplazan las etiquetas temporales por las etiquetas unificadas correspondientes. Para este método se puede usar conectividad 4 (utilizando los vecinos de arriba, abajo, izquierda a derecha) o conectividad 8 (incluyendo las diagonales), lo que varía la segmentación (Gonzalez y Woods, 2018).



Figura 3. Identificación de las regiones de interés.

Obtenida de: elaboración propia.

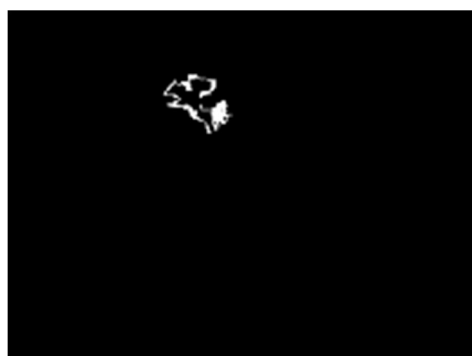


Figura 4. Imagen sin ruido.

Obtenida de: elaboración propia.

Se utilizó la función *regionprops* en MATLAB, la cual es una herramienta fundamental en el procesamiento de imágenes, especialmente útil para analizar imágenes binarias que

permite extraer diversas características geométricas y espaciales de las regiones conectadas dentro de una imagen como son, el área en una región de píxeles, el centroide que son las coordenadas del centro de masa de la región (MathWorks, 2024).

3. Desarrollo del Trabajo: *aplicando las reglas y principios científicos*

La metodología propuesta para el fenotipado fue integrada por seis fases como se muestra en la Figura 5.

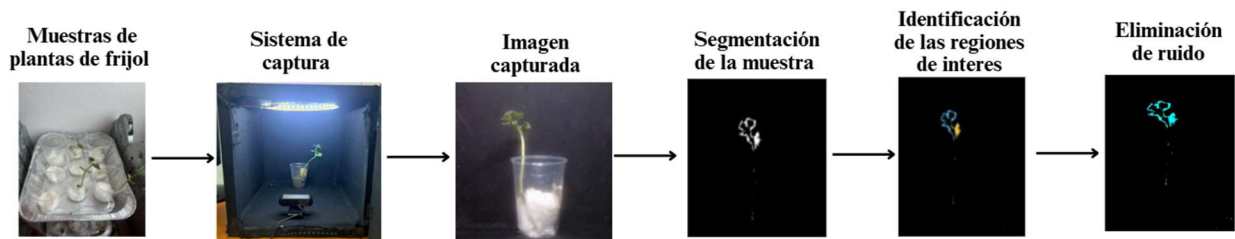


Figura 5. Esquema general del fenotipado de la planta de frijol.

Obtenida de: elaboración propia.

La primera fue el cultivo de las cuarenta muestras de frijol con la finalidad de explorar el impacto de variables como la temperatura, humedad y nutrientes en la germinación y crecimiento de las muestras, con el fin de estandarizar condiciones óptimas. En la segunda se desarrolló el sistema de adquisición de imágenes aislado representado en la Figura 6 con las medidas 37cm de ancho, 35 cm de alto y 37cm de profundidad, se aplicó iluminación multispectral para mejorar la calidad de las imágenes, permitiendo la detección precisa de características clave. La tercera fase consistió en capturar imágenes una cámara digital, utilizando el modelo Logitech C920 HD Pro Webcam (Full HD 1080p a 30 fps), de las muestras desde su germinación hasta los catorce días de crecimiento. Para asegurar la consistencia de los datos y reducir el margen de error humano, se obtuvieron imágenes con una calidad de 480×640 píxeles. La cuarta fase fue la segmentación de las muestras, separándolas y reduciendo la interferencia de elementos no deseados. Para la segmentación de las muestras se utilizó una conversión del espacio de color RGB a HSV. Los umbrales utilizados fueron Hue ($0.053 \leq H \leq 0.524$), Saturación ($0 \leq S \leq 0.399$) y Valor ($0 \leq V \leq 1$). La quinta fue la identificación de las regiones de interés y la sexta fue la eliminación del ruido de fondo

para extraer características específicas de las muestras, facilitando la comparación entre diferentes condiciones de cultivo.

$$\text{Días pronosticados} = 0.01041 \frac{\text{Días}}{\text{Área}} \times \text{Área} + 24.26 \text{ Días} \quad (1)$$

En la Tabla 1 se presentan los resultados obtenidos tras la segmentación e identificación de las regiones de interés en ocho muestras de plantas. La primera columna muestra el área foliar de cada muestra, mientras que las columnas dos y tres reflejan los valores de los ejes mayor y menor, los cuales están relacionados con la forma y tamaño de las hojas. La cuarta columna indica los días reales de crecimiento, y la quinta muestra los valores estimados por el modelo de regresión lineal descrito en la Ecuación (1). En la sexta columna se reporta el error entre los días estimados y los días reales. Por otro lado, en la Figura 7 se observa la comparación entre el área foliar y los días de crecimiento. De acuerdo con los resultados, el modelo alcanzó una correlación estadística de 0.7720 y un error absoluto medio de 3.323 días, lo que indica una buena predicción en la estimación del crecimiento a partir del área.

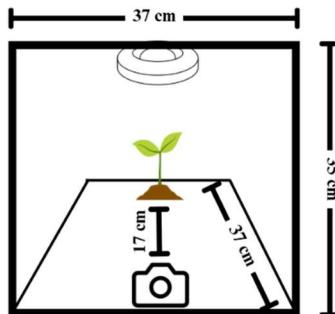


Figura 6. Sistema de captura del fenotipado de la planta de frijol.

Obtenida de: elaboración propia.

4. Conclusiones: lo que podemos aprender de este artículo

El presente estudio destaca la aplicación de sistemas de visión computacional en la agricultura, específicamente en el fenotipado de plantas de frijol. De acuerdo con los resultados obtenidos se logró encontrar un modelo matemático de regresión lineal que cuenta con un factor de sensibilidad de 0.01041(días/área). Además, se observó que el modelo presentó el mayor error en el cambio del eje mayor en el rango 58.085 a 100.417. Este método puede ser utilizado en aplicaciones de campo con la finalidad de reducir el impacto

ambiental en el manejo de fertilizantes y el uso de recursos. Como trabajo a futuro se tiene contemplado incrementar el número de muestras y realizar pruebas de campo.

Tabla 1. Datos experimentales y evaluación del modelo matemático.

Área	Longitud eje mayor	Longitud eje menor	Días de crecimiento	Días pronosticados	Error de días
518	44.801	22.196	28	29.6524	1.6524
552	155.821	9.231	28	30.0063	2.0063
609	58.031	23.067	28	30.5997	2.5997
688	58.085	26.890	28	31.4221	3.4221
867	100.417	19.778	40	33.2855	-6.7145
966	127.335	13.344	40	34.3161	-5.6839
1857	237.333	72.624	40	43.5914	3.5914
1424	102.477	60355	40	39.0838	-0.9162

Obtenida de: elaboración propia.

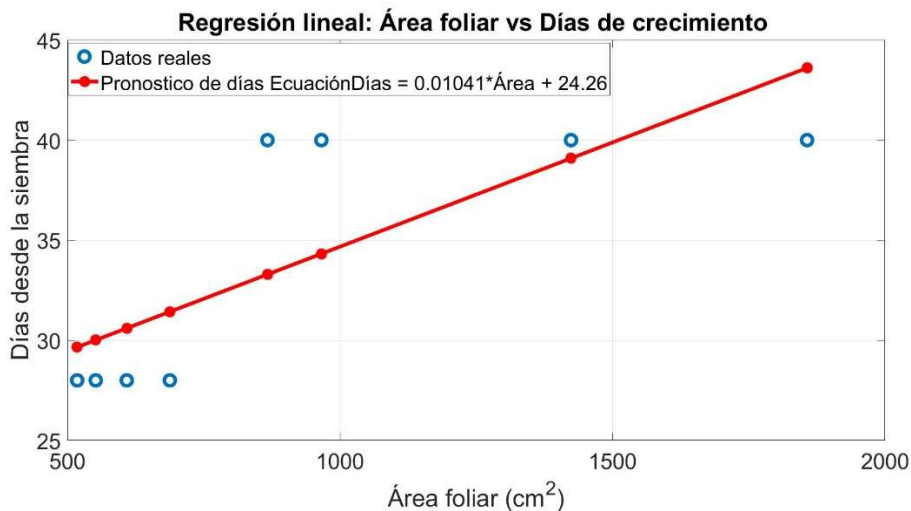


Figura 7. Respuesta del modelo propuesto. Obtenida de: elaboración propia.

5. Referencias: *por si quieres seguir conociendo más*

- Bautista-Lozada, A., Parra, R. F., & Espinosa-García, F. J. (2012). Efectos de la domesticación de plantas en la diversidad fitoquímica. *Temas Selectos de Ecología Química de los Insectos*. México: Colegio de la Frontera Sur, 253-264.
- Correa, P., Bautista, P., López, M., Villaseñor, M., García, C., Rostro-González, H., & Pérez-Pinal, F. J. (2019). An approach of beans plant development classification using fuzzy logic. *Journal of Sensors*, 2019(1), 4232536.
- Flores, J. D. L. C. (2023). *Respuesta de dos variedades de frijol a la aplicación de diferentes fuentes nutrimentales* (Disertación Doctoral, Universidad Nacional Autónoma De México).
- Fernández Valenciano, A. F., & Sánchez Chávez, E. (2017). Estudio de las propiedades fisicoquímicas y calidad nutricional en distintas variedades de frijol consumidas en México. *Nova scientia*, 9(18), 133-148.
- Gonzalez, R. C., & Woods, R. E. (2018). *Digital image processing* (4ta. Edición). Pearson.
- MATLAB. (s.f.) *Conncomp: Connected graph components*.
<https://www.mathworks.com/help/matlab/ref/graph.conncomp.html>
- Morales, R. (2024, 30 agosto). Sequía disparó 156% importaciones mexicanas de frijol en enero-junio. *El Economista*. <https://www.eleconomista.com.mx/empresas/Sequia-disparo-156-importaciones-mexicanas-de-frijol-en-enero-junio-20240829-0140.html>
- Villaseñor-Aguilar, M. J., Botello-Álvarez, J. E., Pérez-Pinal, F. J., Cano-Lara, M., León-Galván, M. F., Bravo-Sánchez, M. G., & Barranco-Gutiérrez, A. I. (2019). Fuzzy classification of the maturity of the tomato using a vision system. *Journal of Sensors*, 2019(1), 3175848.