

ARTÍCULO ACEPTADO

Reconocimiento de brócoli mediante la anotación automática en imágenes RGB-D

Ricardo Ángel Álvarez Macedo, José Luis López Díaz, Héctor Alejandro Montes Venegas, Rosa María Valdovinos Rosas y José Raymundo Marcial Romero

El etiquetado o la anotación de datos para entrenar modelos de clasificación implica el manejo de grandes cantidades de información, además es una tarea ardua y tediosa que representa uno de los cuellos de botella en el entrenamiento de sistemas de aprendizaje automático. En este trabajo se propone mejorar el etiquetado o la anotación de plantas de brócoli de una manera semiautomática, lo que incrementará la disponibilidad de datos para la experimentación y evaluación de sistemas de detección y localización automáticos. Para ello, se utilizó un algoritmo de extracción de grupos basado en distancia Euclidiana y se realizó la estimación de características para cada nuevo grupo, lo que permite determinar la similitud entre grupos ya detectados y grupos extraídos de observaciones posteriores. Las pruebas realizadas en conjuntos de datos tomados en campos de cultivo de brócoli del Reino Unido mostraron que la detección y anotación semiautomática de brócoli permite reducir el tiempo que se requiere en el etiquetado de datos.

Introducción

La agricultura desempeña un papel importante en la economía mundial, al ser el sustento imprescindible de las personas, proporcionar alimentos, materias primas y generar empleo en todo el mundo, especialmente en las zonas rurales. Sin embargo, la agricultura también enfrenta muchos desafíos, como el cambio climático, la degradación del suelo, la escasez de agua y la pérdida de biodiversidad.

De forma particular, el cultivo de hortalizas juega un papel fundamental en la alimentación humana y dentro de ellas se encuentra el brócoli o *Brassica oleracea var. italica*, que ha ganado popularidad en los últimos años debido a sus propiedades nutricionales. Su consumo frecuente ha sido recomendado, ya que podría reducir los riesgos de diabetes, anemia y algunos tipos de cáncer debido a algunos compuestos químicos anticancerígenos conocidos como *glucosinolatos*.

En su constitución, el brócoli está formado por una parte comestible (cabeza floral), una masa densa de yemas florales, que es una agrupación de flores inmaduras.

Dependiendo del país de producción, las pérdidas de cultivo ascienden a un 10 % y un 40 %, que pueden ocurrir en la manipulación, almacenamiento, empaque, envío, comercialización, consumo y por cosecha tardía, siendo ésta última la principal causa [1]. En el caso de la

recolección de pellas maduras de brócoli, se debe efectuar en un plazo muy breve de tiempo, ya que las inflorescencias mantienen su estado óptimo de corte en tan solo dos días. Después de este periodo, la calidad se reduce, por lo que el personal de la cosecha debe estar preparado y entrenado para reconocer visualmente las plantas listas para cosechar y realizar el corte adecuadamente.

Últimamente, se ha observado una creciente automatización en la agricultura a través de Inteligencia Artificial (IA) y Sistemas Robóticos. Esto se debe principalmente a los avances que los algoritmos *Machine Learning* tienen en la detección y clasificación de elementos.

Lo anterior representa un cuello de botella ya que la recolección de datos es casi automática, mientras que la anotación de datos es mucho más lenta al realizarse comúnmente de forma manual [2].

Algunas aplicaciones en tecnologías agrícolas como la detección y recolección automática de frutos o cultivos utilizan sensores que capturan el color y su mapa de profundidad (conocido como imágenes RGB-D, Red-Green-Blue-Depth, por sus siglas en inglés).

Es común que una imagen RGB-D sea convertida y almacenada como una nube de puntos para su posterior procesamiento por medio de herramientas como PCL (*Point Cloud Library*, por sus siglas en inglés,) o Open3D, que ofrecen una variedad de algoritmos para la percepción 3D [3].

Blok et al. [4] implementaron un algoritmo de aprendizaje que puede estimar el tamaño de las cabezas de brócoli bajo condiciones de oclusión en imágenes RGB-D usando ORCNN (*Occlusion Region-based Convolutional Neural Network*, por sus siglas en inglés). Estos resultados fueron comparados con una Mask R-CNN (*Mask Regional Convolutional Neural Network*, por sus siglas en inglés), donde detectaron y estimaron el tamaño de 229 de las 232 cabezas de brócoli cosechables de tres variedades diferentes [5]. De igual modo, Kusumam et al. [6] describen un sistema de visión para la detección de brócoli utilizando sensores RGB-D, aprendizaje automático y métodos de filtrado temporal para la detección de cabezas de brócoli.

Por otro lado, Louedec et al. [7] hacen uso de la información 3D para detectar cabezas de brócoli basado en CNN (*Convolutional Neural Network*, por sus siglas en inglés). Así mismo, el trabajo realizado por Montes et al. [8] proponen dos métodos para procesar datos 3D

que detectan de manera eficiente cabezas de brócoli en tiempo real.

En este artículo se presenta un método basado en la anotación interactiva de datos en imágenes RGB-D para el reconocimiento del brócoli con la mínima intervención humana posible, mejorando el proceso de selección de los puntos en la notación manual.

Metodología

Los experimentos presentados en este artículo fueron implementados utilizando PCL sobre una serie de datos RGB-D tomadas por un Kinect 2 (1920×1080 de resolución RGB y 512×424 de resolución de profundidad) disponibles públicamente*. El método propuesto consta de 5 etapas (Figura 1).

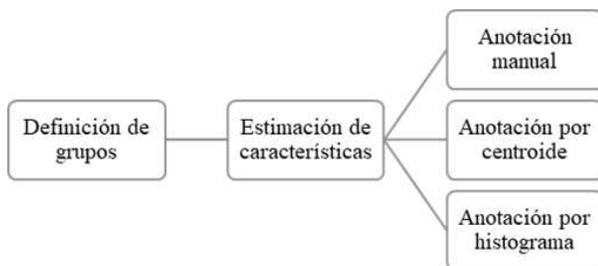


Figura 1. Diagrama del método para anotación de brócoli. Fuente: Elaboración propia.

Etapa de agrupamiento

```

Algorithm 1: EuclideanClusterExtraction


---


Data: points, tolerance, minSize, maxSize
Result: clusters
clusters ← [];
foreach point in points do
  closestCluster ← None;
  foreach cluster in clusters do
    if distance(point, cluster[0]) < tolerance then
      closestCluster ← cluster;
      break;
    end
  end
  if closestCluster is None then
    clusters.append([point]);
  end
  else
    closestCluster.append(point);
  end
end
foreach cluster in clusters do
  if len(cluster) < minSize or len(cluster) > maxSize then
    clusters.remove(cluster);
  end
end
return clusters;


---


  
```

Figura 2. Algoritmo de extracción de grupo basado en la distancia euclidiana. Fuente: Elaboración Propia.

*https://1cas.lincoln.ac.uk/nextcloud/shared/agritech-datasets/broccoli/broccoli_datasets.html

El *clustering* busca agrupar y recolectar datos relevantes, centrándose en identificar las cabezas de brócoli presentes en la nube de puntos. Para lograr esto, se utiliza la extracción de grupos basado en la distancia euclidiana (ver Figura 2).

Donde *points* son los puntos de la nube de puntos, *tolerance* es la distancia máxima que dos puntos pueden tener para considerarse parte del mismo grupo, *minSize* y *maxSize* donde indican el tamaño mínimo y máximo de puntos que puede tener un grupo y si no se cumple es descartado, este valor también se determinó empíricamente (Figura 3).

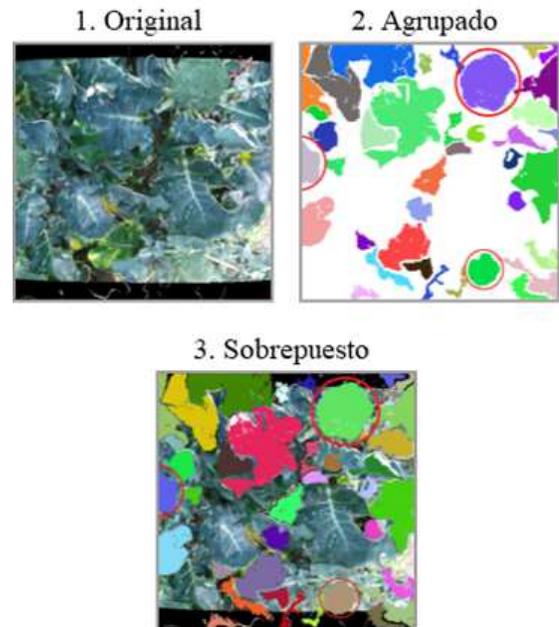


Figura 3. Identificación de grupos en una nube de puntos. (1) Vista frontal de la imagen original, (2) Grupos extraídos con cada brócoli dentro de un círculo, (3) Grupos sobrepuestos en la imagen original.

Etapa de estimación de características

En esta etapa se determinan las propiedades de interés de un grupo. Se considerarán dos formas: mediante el cálculo del centroide (ecuación 1) que se utiliza para representar el punto central del grupo, y mediante el cálculo del histograma utilizando el descriptor VFH (*Viewpoint Feature Histogram*). Éste último se construye usando las normales de los puntos del grupo y la comparación de ángulos de inclinación entre las normales [9].

$$(x, y, z) = \left(\frac{1}{n}\right) \sum_{i=0}^n (x_i, y_i, z_i) \tag{1}$$

Donde: (x, y, z) representa el promedio de todos los puntos de un grupo como un punto en el espacio tridimensional. La variable n es el número total de puntos que tiene el grupo, por lo que (x_i, y_i, z_i) es un punto dentro del grupo en la posición i .

Etapa de anotación manual

La anotación de datos es un proceso que generalmente se realiza de forma manual, lo que implica que una persona utiliza su criterio para seleccionar los datos relevantes. En el caso específico del brócoli, el objetivo es que una persona pueda visualizar una nube de puntos mediante una interfaz interactiva y, así, identificar y seleccionar un grupo que represente una cabeza del brócoli.

La anotación se lleva a cabo en la misma interfaz, donde se escogen los puntos que pertenecen a un brócoli. Si alguno de los puntos seleccionados forma parte de un grupo que ya fue extraído previamente, los demás puntos de ese grupo se seleccionan automáticamente para asegurar que todos los puntos relevantes sean anotados.

Etapa de anotación por centroide

La anotación por centroide tiene como objetivo la identificación y etiquetado automático de grupos extraídos que puedan corresponderse a un brócoli. Se requiere de grupos extraídos y un conjunto de datos de referencia (grupos previamente anotados).

En cada uno de estos grupos, se aplica el proceso de estimación de características, centrado en el cálculo de los centroides. Posteriormente, se calculan las distancias euclidianas entre los centroides anotados y los centroides extraídos (ecuación 2).

$$d = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2} \tag{2}$$

Con la definición de un umbral de cercanía se determina si los grupos extraídos son lo suficientemente cercanos a los grupos anotados para entonces realizar la notación (ver Figura 4).

Etapa de anotación por histograma

La anotación por histograma busca identificar y etiquetar de forma automática grupos extraídos que puedan corresponder a un brócoli. Para ello, se requiere de grupos extraídos y un conjunto de datos de referencia (grupos previamente anotados).

Una vez calculados los histogramas, se utiliza el coeficiente de correlación de Pearson, para encontrar la similitud entre dos histogramas y determinar qué grupos se anotarán como brócoli (ver Figura 4).

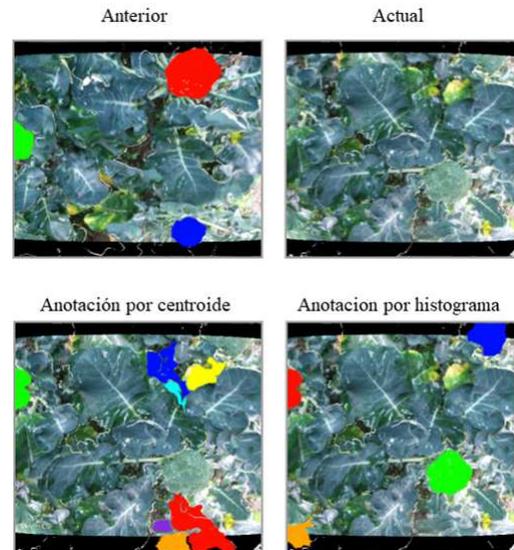


Figura 4. Anotación automática por estimación de centroides e histogramas. Fuente: Elaboración propia.

En la Figura 4, se muestran los grupos anotados en diferentes colores sobre una nube de puntos. La imagen etiquetada como ‘Actual’ muestra otra nube de puntos adyacente a la que se ve en la imagen ‘Anterior’. En la imagen ‘Actual’, se aplica el proceso de agrupamiento para después aplicar el proceso de anotación por centroide y en otra prueba se cambia a anotación por histograma.

Large Experimentación y resultados

La experimentación consideró el uso de 299 muestras previamente etiquetadas para asegurar la precisión en los resultados obtenidos. Este conjunto es conocido como *ground truth*. (ver Figura 5).

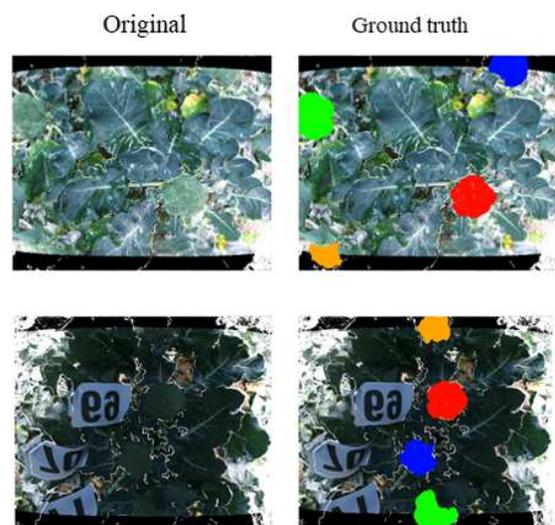


Figura 5. Muestras de campos sin ground truth y con ground truth. Fuente: Elaboración propia.

La idea es solo usar una muestra de referencia para los procesos de anotación por centroide e histograma (Figura 6).

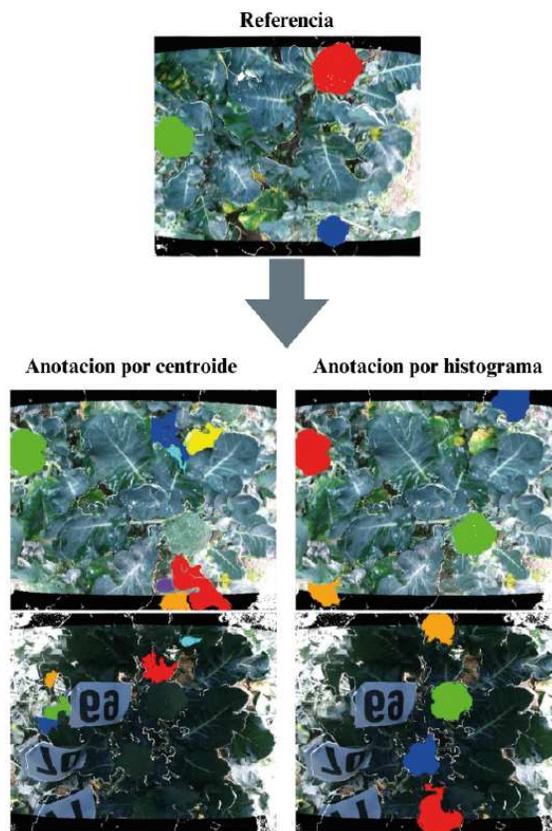


Figura 6. Anotación de brócoli comparando la anotación por centroide y la anotación por histograma usando solo una muestra de referencia. Fuente: Elaboración Propia

Con las 299 muestras se experimentó con la anotación por histograma y se compararon los resultados con el *ground truth*, calculando así la efectividad del método en un 77 % (ver Tabla 1).

Tabla 1. Resultados con la experimentación de anotación por histograma.

| Método utilizado | | Etiquetados correctamente |
|---------------------|------------|---------------------------|
| <i>Ground truth</i> | histograma | |
| 1058 grupos | 850 grupos | 819 grupos |

Conclusiones

forma automática cabezas de brócoli utilizando imágenes RGB-D. El método considera la obtención de grupos que serán utilizados como guía en el proceso de anotación manual. Con esto se busca tener una selección más rápida y precisa de las cabezas de brócoli que serán utilizadas en el proceso de estimación de características. Por

último, estas estimaciones son usadas para los procesos de anotación por centroide y anotación por histograma.

Los resultados obtenidos reportan un 77 % de efectividad en el etiquetado automático de brócoli, lo que, en el proceso de cultivo permite disminuir el trabajo arduo que implica la anotación de datos, para después revisar los datos que se anotaron automáticamente y que una persona los pueda considerar válidos o que pueda seguir completando la anotación usando el proceso de anotación manual.

En un futuro se proponer mejoras en el proceso de agrupamiento para que se logre discriminar elementos que no son de interés (hojas, tallos o suelo), mejorando así la precisión y evitando la pérdida de los datos o el etiquetado de estos.

Otra dirección posible de estudio es ampliar el proceso de estimación de características y, consecuentemente, los procesos de anotación.*

REFERENCIAS

- Erkan, M. y Dogan, A. (2019) Harvesting of Horticultural Commodities,» de Postharvest Technology of Perishable Horticultural Commodities. Woodhead Publishing, 129-159.
- Telus International. Five common data annotation challenges and how to solve them. Accedido el 4 de Noviembre de 2022. Disponible en: <https://www.telusinternational.com/insights/ai-data/article/data-annotation-challenges>.
- Li, L., Wang, R. y Zhang, X. (2021). A Tutorial Review on Point Cloud Registrations: Principle, Classification, Comparison, and Technology Challenges. *Mathematical Problems in Engineering*, 2021, 32.
- Blok, P., van Henten, E., van Evert, F. y Kootstra, G. (2021). Image-based size estimation of broccoli heads under varying degrees of occlusion. *Biosystems Engineering*, 208, 2013-233.
- Blok, P., van Evert, F., Tielen, A., van Henten, E. y Kootstra, G. (2021). The effect of data augmentation and network simplification on the image-based detection of broccoli heads with Mask R-CNN. *Journal of Field Robotics*, 38, 85-104.
- Kusumam, K., Krajcnik, T., Pearson, S., Duckett, T. y Cielniak, G. (2017). 3D-vision based detection, localization, and sizing of broccoli heads in the field. *Journal of Field Robotics*, 34, 1505-1518.
- Louedec, J., Montes, H., Duckett, T. y Cielniak, G. (2020). Segmentation and detection from organised 3D point clouds: a case study in broccoli head detection. En *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 285-293.
- Montes, H., Le Louedec, J., Cielniak, G. y Duckett, T. (2020). Real-time detection of broccoli crops in 3D point clouds for autonomous robotic harvesting. En *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 10483-10488.
- Rusu, R., Bradski, G., Thibaux, R. y Hsu, J. (2010). Fast 3D recognition and pose using the Viewpoint Feature Histogram. *IEEE*.
- Kootstra, G., Wang, X., Blok, P., Hemming, J. y van Henten, E. (2021). Selective Harvesting Robotics: Current Research, Trends, and Future Directions. *Current Robotics Reports*, 2, 95-104.

SOBRE LOS AUTORES



Ricardo Ángel Álvarez Macedo. es estudiante de Ingeniería en computación participante en proyectos de investigación relacionados a tratamiento de imágenes digitales, ciencia y analítica de datos aplicada.



José Luis López Díaz. es estudiante de Ingeniería en computación participante en proyectos de investigación relacionados a tratamiento de imágenes digitales, ciencia y analítica de datos aplicada.



Héctor Alejandro Montes Benegas. es Doctor en Ciencias Computacionales por la Universidad de Lincoln, UK. Cuenta con varios años de experiencia en empresas del sector público y privado como asesor externo en desarrollo de software, capacitación de personal y asesoría en sistemas de computo. Es Profesor de Tiempo Completo en la Facultad de Ingeniería en la UAEMex. Sus intereses de investigación incluyen algoritmos de Computo Evolutivo aplicados a problemas de optimización combinatoria y Percepción Robótica que involucra la adquisición, procesamiento, análisis e interpretación de los datos que producen para tareas de robótica móvil y aprendizaje máquina. Ha dirigido y codirigido tesis de Licenciatura y de Maestría.



Rosa María Valdovinos Rosas. es Doctora en ciencias computacionales, miembro del SNI nivel II. Con productos que inciden en varios PRONACES: proyectos de investigación, artículos en revistas indexadas y congresos, capítulos de libro, libros y 1 licencia de patente. Ha formado estudiantes de nivel licenciatura, Maestría y Doctorado. Colabora en la organización de eventos académicos para la divulgación de la ciencia, acciones orientadas a niñas en la computación y vocaciones tempranas. Con más de 1000 citas y un índice H14, el *International AD Scientific index 2023* la considera una de las 10,000 científicas más influyentes a nivel Nacional.



José Raymundo Marcial Romero. es Doctor en Ciencias de la Computación por The University of Birmingham, UK. Profesor de tiempo completo en la Facultad de Ingeniería de la UAEMex. Miembro del SNI nivel 1. Miembro de la academia Mexicana de Ciencias de la Computación, con aportaciones en las áreas de teoría de grafos, semántica de lenguajes de programación y la lógica computacional. De acuerdo con Google Scholar, ha publicado más de 90 artículos de investigación para revista y capítulos de libro. Ha impulsado proyectos estratégicos para la Universidad en la que labora desde la Secretaría de Planeación y la Secretaría de Docencia.



Imagen tomada de:
https://cvg.cit.tum.de/research/rgb-d_sensors_kinect.