

# Detección automática de brócoli en imágenes RGB-D

## Automatic broccoli detection in RGB-D images

### Resumen

El brócoli es una hortaliza considerada de alto valor a nivel mundial. Sin embargo, en varios países su producción continuamente se ve impactada por una alta inestabilidad en la mano de obra causada por una amplia gama de fluctuaciones económicas, políticas y de fenómenos migratorios. Por este motivo, se han buscado alternativas de recolección autónomas y semiautónomas que faciliten este trabajo, incrementen la producción y reduzcan costos. En este artículo, se revisan algunas estrategias propuestas para la detección automática en el proceso de recolección de brócoli basadas en la segmentación y el filtrado de imágenes 3D, además de mostrar la eficacia de cada estrategia comparando los resultados obtenidos entre ellas.

Palabras clave

**Segmentación, filtrado de imágenes, Circulo de ajuste**

### Abstract

Broccoli is a vegetable considered to be of high value worldwide. However, in several countries its production is continually impacted by high labor instability caused by a wide range of economic and political fluctuations and migratory phenomena. For this reason, autonomous and semi-autonomous harvesting alternatives have been sought to facilitate this work, increase production and reduce costs. In this article, we review some proposed strategies for automatic detection in the broccoli harvesting process based on segmentation and 3D image filtering and show the effectiveness of each strategy by comparing the results obtained between them.

Keywords

**Segmentation, Image filtering, Circle fit**

## 1. Introducción

La recolección de hortalizas de alto valor como el brócoli es una tarea que consume tiempo a los agricultores por ser un trabajo aún se realiza enteramente de forma manual en todo el mundo.

Para cosechar el brócoli, los jornaleros deben identificar la cabeza del producto, determinar visualmente si ya tiene el tamaño adecuado y, finalmente, realizar un corte de forma inclinada en el tallo de hasta unos 12 cm. Un ejemplo de este proceso se puede ver en la Figura 1.

Este trabajo de cosecha manual es físicamente exigente y requiere que los trabajadores, en cada corte, se agachen, se inclinen y realicen una serie repetitiva de movimientos que, en el mediano y largo plazo, pueden resultar en lesiones de diversa índole. Aunado a esto, en varios países su producción constantemente se ha visto impactada por una alta inestabilidad en la mano de obra, causada por diversos factores que van desde condiciones económicas, dinámicas sociales, situaciones políticas y todas las complejidades involucradas en los fenómenos migratorios [1].



Figura 1. Proceso de recolección manual de brócoli realizado por cuadrillas de trabajadores.

## 2. Trabajos previos

Se han realizado varias investigaciones encaminadas a apoyar a los agricultores en la tarea de identificación y recolección del brócoli. En este sentido, el desarrollo de sistemas dotados de diversos grados de autonomía que permitan cosechar de forma continua ha recibido una atención considerable en las últimas décadas.

Ramírez [2] desarrolló uno de los primeros intentos de localizar las cabezas de brócoli utilizando un pequeño conjunto de 13 imágenes RGB. Este método detectó el brócoli basándose en el contraste y el análisis estadístico de la textura de la imagen. Mientras que, Block *et al.* [3] describieron un sistema que también usaba imágenes RGB para detectar brócoli de dos tipos de variedades. El método incluía un filtro de textura y un análisis de color de la apariencia de la cabeza.

Kusumam *et al.* [4] detectó cabezas de brócoli en datos tomados con un sensor RGB-D. Éste método combina la segmentación basada en las distancias entre puntos, la estimación de un vector de características compuesto por los valores de las normales de los puntos de cada segmento, un clasificador *Support Vector Machines* (SVM) y un filtro temporal para detectar y rastrear las cabezas de brócoli de dos variedades diferentes en una secuencia de tomas recolectadas a diferentes números de cuadros por segundo. Posteriormente, un enfoque similar aplicado a un conjunto mayor de datos RGB-D fue extendido a cuatro diferentes métodos de segmentación de mayor precisión que procesaban un número elevado de cuadros por segundo [5] y rastreaban de manera única cada cabeza detectada utilizando Filtros de Partículas [6].

De manera similar, por varios años las técnicas de aprendizaje profundo o *deep learning* (DL) se han aplicado con éxito en varios campos incluyendo la agricultura [7]. Las técnicas de DL han logrado un alto grado de clasificación y una ejecución muy eficiente. Sin embargo, la aplicación de DL requiere una gran cantidad de datos de entrenamiento debido a la complejidad de las tareas y un trabajo considerable de etiquetado no siempre disponibles debido al costo y tiempo de adquisición necesarios.

Por otro lado, Bender *et al.* [8] recolectaron un conjunto de datos de brócoli y coliflor en diferentes etapas de crecimiento y demostraron su detección usando una *Red Neuronal Convolutiva* (CNN, por sus siglas en inglés). Estos resultados, sin embargo, fueron obtenidos para toda la planta de brócoli visible en las imágenes y no únicamente para la cabeza de la hortaliza.

Trabajos más recientes propuestos por Block *et al.* [9] emplearon varias técnicas de DL en imágenes RGB para la detección de brócoli con altos grados de precisión, aún en condiciones de oclusión que ocurrían de manera natural o inducidas de manera artificial por los mismos investigadores [10]. Louedec *et al.* [11], utilizaron un modelo de CNN adaptado a los datos RGB-D representados por los valores de las normales de los puntos de cada cuadro adquirido por el sensor. Este enfoque arrojó una mayor generalización entre los diferentes conjuntos de datos utilizados en su experimentación.

De los estudios antes mencionados, el reconocimiento visual en 2D ha sido, por lo general, un paso previo para el funcionamiento de los sistemas autónomos reportados. Sin embargo, el procesamiento de información 3D es una fuente útil de información en diversas etapas del desarrollo de un producto agrícola para determinar características como su forma, tamaño y ubicación en el espacio.

Adicionalmente, se han creado herramientas de software que facilitan el procesamiento de datos 3D, tales como *Point Cloud Library*<sup>[2]</sup> y *Open3D*<sup>[3]</sup> que incluyen una amplia gama de funcionalidades específicas para procesar datos de nubes de puntos.

Un problema por considerar al procesar datos RGB-D es la existencia de objetos no deseados, tales como hojas o partes del mismo suelo donde se ha plantado la hortaliza (ver Figura 2). Esto dificulta el procesamiento de cada cuadro y demanda la utilización de algoritmos de filtrado y limpieza debido a que, al momento de segmentar los puntos, los elementos que no pertenecen a una cabeza de brócoli frecuentemente se agregan al mismo segmento.

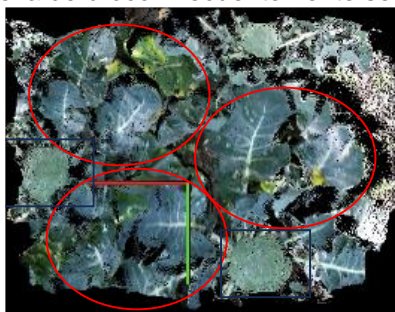


Figura 2. Elementos no deseados en la planta de brócoli tal como se encuentran en los campos de cultivo

En este artículo, se presenta un estudio sobre métodos de procesamiento de segmentos extraídos de imágenes RGB-D con el fin de depurar y filtrar el conjunto de clústers producidos en cada cuadro y, así, mejorar la detección y localización de las cabezas de brócoli. El desarrollo de pruebas se realiza utilizando datos obtenidos con un sensor Kinect en cultivos de Reino Unido y de España.

### 3. Herramientas y métodos

Para validar los métodos de filtrado, se utilizó un conjunto de datos de imágenes de brócoli adquiridas en campos de cultivo bajo diferentes condiciones meteorológicas con el sensor Kinect 2, el cual tiene una resolución de 1920x1080 RGB y 512x424 de resolución de profundidad [8].

El sensor se fijó dentro de una caja cerrada montada en un tractor agrícola con el fin de tener una iluminación uniforme para protegerlo de las condiciones externas.

El conjunto de datos se obtuvo de dos campos de cultivo de brócoli. El primero de ellos, obtenido de campos en Reino Unido, consta de 600 *cuadros* o imágenes RGB-D de una variante denominada *Ironman*. En tanto que las imágenes del segundo conjunto son procedentes de un campo de cultivo en España e incluye 300 *frames* o imágenes RGB-D de una variante de brócoli denominada *Titanium*. Una muestra de estos datos se puede ver en la Figura 3.

Las imágenes obtenidas por el sensor se encuentran en formato PCD (*point cloud data*), por lo que para su manipulación se hace uso de la biblioteca PCL (*point cloud library*) implementada en C++.

### 4. Segmentación de Nubes de Puntos

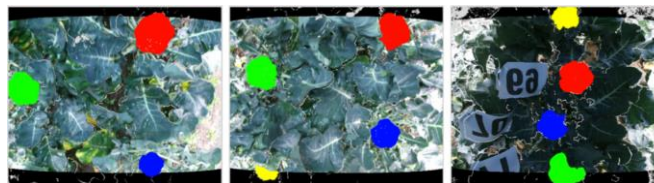
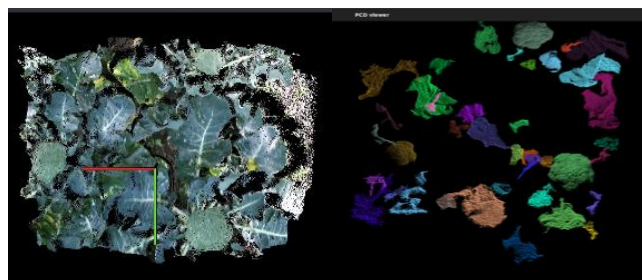


Figura 3. Imágenes RGB-D de cultivos de brócoli tomadas de campos del Reino Unido -dos primeras imágenes- y de España -tercera imagen. Las partes en color representan los puntos que han sido etiquetados como pertenecientes a cabezas de brócoli.

El primer paso del proceso de filtrado es la extracción de *clústers* o segmentación de los puntos en 3D. Para los experimentos presentados en este trabajo, se ha decidido utilizar un método de segmentación o *clusterización* basado en la distancia euclidiana entre los puntos de cada una de las imágenes RGB-D tomadas por el sensor [6].

La segmentación euclidiana comprueba la distancia entre dos puntos, si esta es inferior a un umbral, se considera que ambos pertenecen al mismo clúster. En este algoritmo un punto inicial cualquiera se incluye al *clúster* actual y luego se incluyen también todos los demás puntos que estén lo suficientemente cerca y, a partir de éstos, aún más puntos se agregan hasta que no se puede añadir ningún punto nuevo. Luego, se inicializa un nuevo *clúster* y el procedimiento se repite con el resto de los puntos que todavía no han sido procesados. El resultado de este proceso puede verse en las imágenes de la Figura 4.



a) Imagen Original

b) Segmentado

Figura 4. Resultado obtenido al aplicar el algoritmo de segmentación a un sólo cuadro tomado del conjunto de datos del Reino Unido. Los segmentos o clústers se visualizan cada uno en un color diferente.

## 5. Segmentación de Nubes de Puntos

Una vez obtenidos los *clústers* en cada imagen, en algunos de los segmentos que pueden ser identificados como brócoli, frecuentemente se ven acompañados de porciones de la imagen que realmente no son parte de una cabeza de la hortaliza. Un ejemplo de esto se puede observar en la Figura 5.

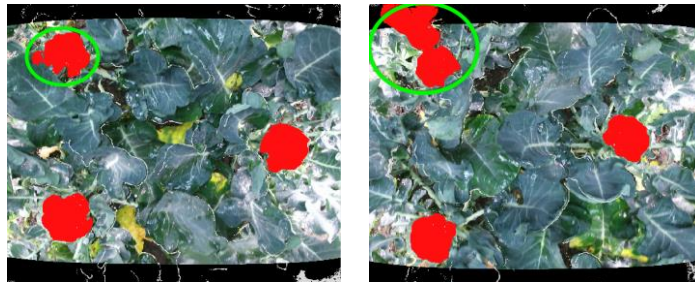
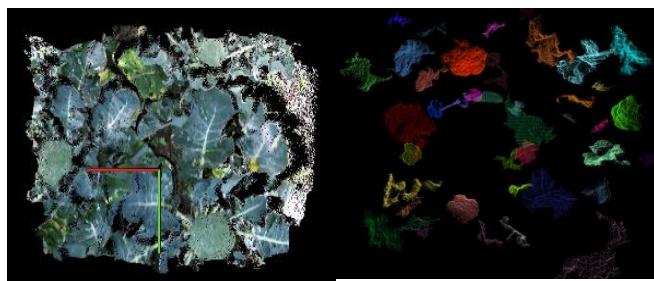


Figura 5. Algunos clústers, marcados con un círculo verde, se han segmentado con puntos que no pertenecen a una cabeza de brócoli.

Debido a este problema de *sobre-segmentado*, una alternativa es eliminar o filtrar los puntos adicionales con el fin de mejorar la identificación de aquellos que representan sólo las cabezas de brócoli. Para lograrlo, en este trabajo, se ha implementado un algoritmo denominado *Ajuste de Círculos* o *Circle Fit*. Este método encuentra el círculo que más se aproxima a los puntos de datos de un clúster.

Matemáticamente, un algoritmo de este tipo es un cálculo de optimización que maximiza o minimiza un parámetro determinado dentro de ciertos grados de libertad y que está sujeto a restricciones. Para el caso de ajuste de círculos, el tamaño y el punto central son las variables *flotantes* cuyos valores se ajustan y las desviaciones de los puntos [12]. Una muestra de este proceso de filtrado se muestra en la Figura. 6.



a) Imagen RGB-D Original

b) Clústers después de haber sido filtrados.

Figura. 6. Una muestra del proceso de filtración sobre los clusters segmentados de una imagen RGB-D.

El funcionamiento del algoritmo de filtrado es el siguiente: el primer paso es la integración y lectura de los *clústers* segmentados. Para cada *clúster*, se realiza lo siguiente:

- a) Ajuste del círculo al *clúster*. Primero, se aplica un algoritmo de ajuste de círculo a los puntos del *clúster*. Dado que PCL no proporciona un método directo de ajuste de círculo, se pueden utilizar otras bibliotecas o algoritmos para este propósito, como el método de ajuste de mínimos cuadrados (*least squares fitting*).
- b) Determinar los parámetros del círculo. Se calculan los parámetros del círculo ajustado, como las coordenadas del centro y el radio. Estos parámetros definen la forma y posición del círculo.
- c) Evaluar la calidad del ajuste del círculo. Se determina la calidad o bondad del ajuste del círculo al *clúster*. Esta evaluación puede basarse en varios criterios, como los errores residuales entre el círculo y los puntos o medidas estadísticas elementales.
- d) Filtrar el *clúster* basado en la calidad del ajuste del círculo. Para esto, se debe definir un umbral o criterio para determinar si la calidad del ajuste del círculo cumple con los criterios deseados. Si la calidad del ajuste del círculo está por debajo del umbral, se considera el *clúster* como un valor atípico y se filtra. En caso contrario se conserva el *clúster*.
- e) Guardar los *clústers* filtrados. Se guardan los *clústers* filtrados que pasaron el filtro de ajuste de círculo para su posterior procesamiento o análisis.

El esquema completo del proceso de filtrado por *circle fit* se muestra en la Figura 7.

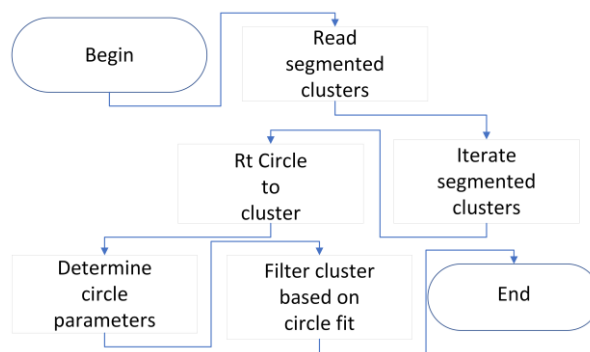


Figura 7: Esquema del proceso completo de filtrado de clústers.

## 6. Resultados

Después del proceso de filtración, se realizó una comparación de los clústeres resultantes con respecto a los datos identificados y etiquetados como cabezas de brócoli (*ground truth*).

En la Tabla 1 se muestra la precisión obtenida tanto en el proceso de segmentación como el de segmentación + filtrado.

Frames	Segmentación	Filtración	Porcentaje de mejora	Desviación Estandar
UK f1	0.95	0.97	0.02	0.01
UK f2	0.95	0.98	0.03	0.015
SP f1	0.88	0.94	0.06	0.03
SP f2	0.82	0.93	0.11	0.055

Tabla 1: Comparativo de porcentaje de precisión del proceso de filtrado respecto al proceso de segmentación aplicado en todos los conjuntos de datos.

En la tabla se puede observar que el proceso de filtrado mejora entre 2-3% la precisión de la segmentación para el conjunto de datos del Reino Unido, mientras que, para el conjunto de datos de España, se incrementa en más del 5%. Esto resulta en una identificación más homogénea en los conjuntos de datos utilizados en nuestros experimentos, aún y cuando las plantas de la hortaliza pertenecen a diferentes variedades y fueron tomadas en diferentes países y condiciones de cultivo.

## 7. Conclusiones

En este artículo, se presenta un estudio sobre métodos de procesamiento de segmentos extraídos de imágenes RGB-D con el fin de depurar y filtrar el conjunto de clústers producidos en cada cuadro con el objetivo de mejorar la detección de cabezas de brócoli en cultivos de dos países diferentes.

En nuestros experimentos se empleó una estrategia que utilizó el filtrado como postproceso de la segmentación y fue posible identificar de forma automática las cabezas de brócoli con una mayor precisión con respecto a los clústers segmentados. Se observó que el proceso de filtrado ayuda a mejorar la detección de los brócolis y hace más homogénea su identificación.

Este algoritmo puede ser usado para la generación de un modelo de reconocimiento y recolección de brócoli, puesto que es adaptativo según a la variedad de brócoli a reconocer en un cultivo.

## 8. Referencias

1. Duckett, T., Pearson, S., Blackmore, S., Grieve, B., Wen-Hua, C., Cielniak, G., Cleaversmith, J., Dai, J., Davis, S., Fox, C., From, P. J., Georgilas, I., Gill, H. S., Gould, I., Hanheide, M., Hunter, A. K., Iida, F., Mihalyova, L., Nefti-Meziani, S., . . . Yang, G. (2018). Agricultural Robotics: The Future of Robotic agriculture. arXiv (Cornell University). <https://doi.org/10.48550/arxiv.1806.06762>
2. Ramirez, R. A. (2006). Computer vision based analysis of broccoli for application in a selective autonomous harvester. ResearchGate. [https://www.researchgate.net/publication/279691048\\_Computer\\_Vision\\_Based\\_Analysis\\_of\\_Broccoli\\_for\\_Application\\_in\\_a\\_Selective\\_Autonomous\\_Harvester](https://www.researchgate.net/publication/279691048_Computer_Vision_Based_Analysis_of_Broccoli_for_Application_in_a_Selective_Autonomous_Harvester)
3. Blok, P. M., Barth, R., & Van Den Berg, W. (2016). Machine vision for a selective broccoli harvesting robot. IFAC-PapersOnLine, 49(16), 66-71. <https://doi.org/10.1016/j.ifacol.2016.10.013>.
4. Kusumam, K., Krajník, T., Pearson, S., Duckett, T., & Cielniak, G. (2017). 3D-Vision based detection, localization, and sizing of broccoli heads in the field. Journal of Field Robotics, 34(8), 1505-1518. <https://doi.org/10.1002/rob.21726>.
5. Montes, H. A., Louède, J. L., Cielniak, G., & Duckett, T. (2020) Montes, H. A., Le Louedec, J., Cielniak, G., and Duckett, T. Real-time detection of broccoli crops in 3D point clouds for autonomous robotic harvesting. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 10483–10488. <https://doi.org/10.1109/iros45743.2020.9341381>.
6. Montes, H. A. and Cielniak, G. (2022). Multiple broccoli head detection and tracking in 3D point clouds for autonomous harvesting. In Proceedings of the 36th AAAI Conference on Artificial Intelligence, workshop on AI for Agriculture and Food Systems, 2022.
7. Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in Agriculture: A survey. Computers and Electronics in Agriculture, 147, 70-90. <https://doi.org/10.1016/j.compag.2018.02.016>
8. Bender, A., Whelan, B., & Sukkarieh, S. (2019b). A high-resolution, multimodal data set for agricultural robotics: a Ladybird 's-eye view of Brassica. Journal of Field Robotics, 37(1), 73-96. <https://doi.org/10.1002/rob.21877>
9. Blok, P. M., van Evert, F. K., Tielen, A. P. M., van Henten, E. J., & 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(1), 85-104. Advance online publication. <https://doi.org/10.1002/rob.21975>
10. Blok, P. M., Van Henten, E., Van Evert, F., & Kootstra, G. (2021b). Image-based size estimation of broccoli heads under varying degrees of occlusion. Biosystems Engineering, 208, 213-233. <https://doi.org/10.1016/j.biosystemseng.2021.06.001>
11. Louedec, J. L., Montes, H. A. (2020). *et al.*, Segmentation and detection from organised 3D point clouds: A case study in broccoli head detection. In IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 285–293, <https://doi.org/10.1109/cvprw50498.2020.00040>.
12. Janeshewski, E. (s. f.). Calculations for circle fitting in CMM software. <http://www.mbccmm.com/cmmq/index.php/training/general/236-calculations-for-circle-fitting-in-cmm-software#:~:text=Calculations%20for%20Circle%20Fitting%20in%20CMM%20Software%201,is%20minimized%20C%20with%20no%20other%20constraints.%20M%C3%A1s%20elementos>



## Sobre los autores



Luis Fernando Talavera Ramírez estudiante de Ingeniería en computación participante en proyectos de investigación relacionados a tratamiento de imágenes digitales, ciencia y analítica de datos aplicada.



Luis Antonio Quiroz Mercado 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 Venegas es Doctor en Ciencias Computacionales por la Universidad de Lincoln en el Reino Unido. Cuenta con siete años de experiencia en diferentes empresas del sector público y privado donde trabajó principalmente como asesor externo dando servicios en desarrollo de software, capacitación de personal y asesoría general en sistemas de computo. Actualmente, es profesor de tiempo completo en la División de Computo de la Facultad de Ingeniería en la Universidad Autónoma del Estado de México. Sus intereses de investigación incluyen varias ramas de la Inteligencia Artificial, en particular algoritmos de Computo Evolutivo aplicados a problemas de optimización combinatoria y Percepción Robótica que involucra la adquisición de información mediante sensores (principalmente cámaras) y los consecuentes procesamiento, análisis e interpretación de los datos que producen (ya sean imágenes en 2D y 3D) para diversas tareas de robótica móvil y aprendizaje máquina. Ha dirigido y codirigido varias tesis de Licenciatura y de Maestría relacionadas con estas líneas de investigación.



Rosa María Valdovinos Rosas. Dra. en ciencias computacionales con doble grado, miembro del SNI nivel II. Ha participado en más de 15 proyectos de investigación con impacto social y científico, además cuenta con más de 50 artículos en revistas indizadas, más de 20 capítulos de libro, más de 20 publicaciones en congresos, 3 libros, más de 100 conferencias dictadas a nivel nacional e internacional y 1 licencia de patente otorgada. Ha formado más de 40 estudiantes de nivel licenciatura, Maestría y Doctorado, de estos últimos todos con reconocimiento por el SNI. Colabora en la organización de eventos de alcance nacional e internacional para la divulgación de la ciencia, así como acciones orientadas a niñas en la computación y vocaciones tempranas a nivel básico, secundaria y superior. La producción científica reporta más de 1000 citas, con un índice H14 y de acuerdo con el *International AD Scientific index 2023* es una las 10,000 científicas más influyentes a nivel Nacional.



José Raymundo Marcial Romero. Doctor en Ciencias de la Computación por The University of Birmingham, UK, en The School of Computer Science. Profesor investigador de tiempo completo en la Facultad de Ingeniería de la Universidad Autónoma del Estado de México. Miembro del Sistema Nacional de Investigadores del CONACYT nivel 1. Miembro de la academia Mexicana de Ciencias de la Computación. Sus aportaciones más relevantes se han dado en las áreas de aplicación 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, capítulos de libro. Ha impulsado proyectos estratégicos para la Universidad en la que labora desde la Secretaría de Planeación 2017-2022 y la Secretaría de Docencia 2022 a la fecha.



This work is under a Creative Commons

Attribution-NonCommercial-ShareAlike 2.5 Mexico license .