

Detección inteligente de baches y estimación de su profundidad con aprendizaje profundo para la conservación de carreteras

Intelligent Pothole Detection and Depth Estimation Using Deep Learning for Road Conservation

Mario Alberto Román Garay*

Tecnológico Nacional de México Campus Culiacán, México

mario.rg@culiacan.tecnm.mx

 <https://orcid.org/0000-0002-4328-4651>

Recepción: 23 de febrero de 2024

Aprobación: 14 de octubre de 2024



Carlos Alberto Hernández Beltrán

Tecnológico Nacional de México Campus Culiacán, México

carlos.hb@culiacan.tecnm.mx

 <https://orcid.org/0000-0002-5317-2506>

Luis Morales Rosales

Universidad Michoacana de San Nicolás de Hidalgo, México

luis.morales@umich.mx

 <https://orcid.org/0000-0002-4753-9375>

Héctor Rodríguez Rangel

Tecnológico Nacional de México Campus Culiacán, México

hector.rr@culiacan.tecnm.mx

 <https://orcid.org/0000-0003-4999-3472>

Miguel Antonio Villa Camacho

Tecnológico Nacional de México Campus Culiacán, México

19170703@itculiacan.edu.mx

 <https://orcid.org/0009-0003-2086-9619>

Jorge Luis Soto Audelo

Tecnológico Nacional de México Campus Culiacán, México

19170677@itculiacan.edu.mx

 <https://orcid.org/0009-0004-6704-6756>

Peter Lepej

Vistion d.o.o, Eslovenia

peter@vistion.si

 <https://orcid.org/0000-0002-7693-1966>

RESUMEN

Se aborda la detección inteligente de baches y la estimación precisa de su profundidad a partir de técnicas de aprendizaje profundo para la conservación de carreteras en entornos urbanos. Para ello, se emplea una cámara de profundidad y la red convolucional YOLOv8, entrenada con 273 imágenes, con la que se obtiene mAP de 84.7%, precisión de 85.7% y *recall* de 76.3%. Los resultados demuestran un error promedio de 5 mm en la estimación de profundidad. En este sentido, se ofrece una herramienta que apoya la gestión de infraestructura vial al proporcionar un inventario de la profundidad de los baches que facilite la planificación de tareas de mantenimiento. El deterioro de las carreteras, visible en forma de baches, conlleva altos costos de reparación y aumenta el riesgo de accidentes.

PALABRAS CLAVE: baches, cámara de profundidad, nube de puntos, estimación de profundidad.

*AUTOR PARA CORRESPONDENCIA

mario.rg@culiacan.tecnm.mx

ABSTRACT

This study addresses the intelligent detection of potholes and the precise estimation of their depth using deep learning techniques for road maintenance in urban environments. To achieve this, a depth camera and the convolutional network YOLOv8 were employed, trained with 273 images, obtaining a mAP of 84.7%, precision of 85.7%, and recall of 76.3%. The results demonstrate an average error of 5 mm in depth estimation. This work offers a tool that supports road infrastructure management by providing a detailed inventory of pothole depths, facilitating the planning of maintenance tasks. Road deterioration, visible in the form of potholes, entails high repair costs and increases the risk of accidents.

KEYWORDS: pothole, depth camera, point cloud, Depth estimation.

INTRODUCCIÓN

Las carreteras no son simplemente vías de tránsito, sino que son infraestructuras cruciales que impulsan la dinámica de la sociedad moderna al promover la conectividad y el flujo económico. Sin embargo, a pesar de su importancia, estas infraestructuras enfrentan desafíos significativos de mantenimiento que afectan tanto la seguridad de los usuarios como la eficiencia logística. El deterioro constante, visible en forma de baches, conlleva altos costos de reparación y aumenta el riesgo de accidentes. Por esta razón, son esenciales inspecciones periódicas para evaluar el estado de las carreteras y planificar el mantenimiento preventivo o correctivo.

La forma tradicional de evaluar las condiciones de las carreteras se basa en identificar y clasificar los daños según su severidad con el uso de manuales de inspección. La Secretaría de Infraestructura, Comunicaciones y Transportes (SICT) cuenta con el manual *M·CSV·CAR·1-03-008/23* (SICT, 2023), donde se describen las características de distintos daños, como agrietamientos y baches, para su identificación. Según este manual, la forma de identificar y evaluar la severidad de un bache es mediante su profundidad y su diámetro en las categorías baja, media o alta. Para medir su profundidad, los inspectores utilizan herramientas manuales como reglas, verniers y flexómetros. Sin embargo, este proceso de inspección y evaluación consume mucho tiempo y recursos, además de exponer a los trabajadores a condiciones de tráfico peligrosas.

Con la intención de agilizar las condiciones de recolección de las medidas de los deterioros, particularmente baches, con las cuales se obtiene la profundidad se han implementado diferentes tecnologías, acelerómetros y giroscopio (Mednis *et al.*, 2011), escáner laser (She *et al.*, 2021) o sensores LiDAR (Kang y Choi, 2017). El uso de acelerómetros y giroscopio se implementa con celulares inteligentes dentro de vehículos captando las vibraciones. La principal limitación de este enfoque es que es necesario sobreponerse sobre la zona dañada con las llantas del vehículo, además de que la toma de datos se ve influenciada por las condiciones de la suspensión del vehículo. En los escáneres laser y sensores LiDAR, a pesar de estimar la medición con un error relativamente pequeño la profundidad de los baches, el costo es su principal limitante para la implementación de este tipo de dispositivos, ya que llega a costar hasta 200 000 MXN, además de requerir una alta capacidad computacional para el procesamiento de los datos.

Ante estas dificultades mencionadas, el objetivo del artículo es desarrollar y evaluar una metodología para la detección y estimación de la profundidad de baches en carreteras con un bajo costo por medio de una cámara de profundidad Real Sense D435i y el modelo de red convolucional YOLOv8 (Jocher *et al.*, 2023). Este artículo está orientado a ofrecer una herramienta para apoyar en la inspección de pavimentos flexibles en entornos urbanos bajo condiciones ambientales no controladas y ofrecer una cuantificación que ayude en la planificación de las tareas de mantenimiento.

Para lograr estos objetivos, la cámara Real Sense D435i es esencial para la recolección de los datos que permitirán identificar y estimar la profundidad de los baches. Esta cámara, que funciona mediante la emisión de puntos láser, hace posible obtener referencias de la escena y medir la distancia desde el sensor hasta la superficie de la carretera, que generan así imágenes y nubes de puntos de la superficie afectada. Además, gracias a la cámara digital, se observan detalles de colores en la escena que, junto a una nube de puntos generada con el sensor láser, es capaz de extraer coloración en cada punto añadiendo más detalles.

Para la detección, se emplea YOLOv8, un modelo de red convolucional entrenado con 273 imágenes de baches. Este modelo identifica la zona dañada y proporciona coordenadas de una cuadro o caja delimitadora alrededor del bache. Los resultados del entrenamiento son los siguientes mAP 84.7%, precisión 85.7% y *recall* 76.3%.

Para la estimación de la profundidad, se utiliza la nube de puntos generada por la cámara de profundidad, que corresponde a la escena que se procesó con YOLOv8, y las coordenadas del bache detectado previamente. Con estas coordenadas se delimita el área de la nube de puntos que corresponde al bache, de manera que solo se consideren para la estimación de profundidad los puntos dentro del área delimitada. Si hay múltiples baches (indicados por varias cajas delimitadoras), el algoritmo se aplica individualmente a cada uno para obtener la mayor profundidad en cada caso.

En este trabajo, se presenta una estructura en cinco secciones. *a)* Estado del arte: revisa las investigaciones recientes sobre detección y estimación de profundidad de baches. *b)* Metodología: detalla el procedimiento seguido en el estudio. *c)* Resultados: muestra los datos obtenidos. *d)* Discusión: analiza estos resultados. *e)* Conclusión: resume los hallazgos y propone futuras líneas de investigación.

1. ESTADO DEL ARTE

La detección automática de deterioros en pavimentos ha avanzado significativamente, dado que se ha apoyado en la tecnología de cámaras digitales y procesamiento computacional. Inicialmente, se distinguía la superficie dañada de la sana por su tonalidad más oscura a partir de técnicas de umbralización como el método de Otsu para resaltar el contraste y detectores de bordes para delimitar las áreas afectadas (Han *et al.*, 2021). No obstante, estos métodos enfrentaban limitaciones al confundir deterioros con elementos ajenos en la carretera, como basura, manchas, sombras o agua.

Para superar estos desafíos, se recurrió a estrategias estadísticas combinadas con aprendizaje profundo para reducir así los errores de falsos positivos (Xu *et al.*, 2013). Las arquitecturas de redes neuronales convolucionales, especializadas en el reconocimiento de patrones en imágenes, han elevado la eficacia de detección. Aunque ha incrementado la carga de datos necesarios para el entrenamiento de modelos y el tiempo de preprocesamiento, su precisión justifica su uso (Ma *et al.*, 2020).

Las cámaras estéreo se han utilizado en otras aplicaciones para estimar la profundidad creando mapas de disparidad. No obstante, su aplicación en la detección de deterioros en pavimentos es limitada, ya que la ausencia de referencias de distancia en las áreas que van a inspeccionarse impide estimaciones precisas de profundidad (Guan *et al.*, 2021).

Recientemente, ha cobrado relevancia el uso de tecnologías como LiDAR, escáneres láser y cámaras de profundidad para la percepción más detallada del terreno al incluir magnitudes físicas a los datos de procesamiento (Cui *et al.*, 2021; Elamin y El-Rabbany, 2023). Estos dispositivos capturan información detallada del entorno que permiten la creación de nubes de puntos para la extracción de datos de dimensiones físicas de los deterioros. La integración de estas tecnologías con cámaras digitales posibilita una caracterización más precisa de los daños en pavimentos.

Los estudios coinciden en que la combinación de tecnologías de cámaras y LiDAR mejora notablemente la percepción ambiental y la detección de irregularidades en superficies pavimentadas. Los métodos de fusión de datos se han enfocado en tareas como la detección de deterioros y la estimación de distancias del automóvil hacia estos para la conducción autónoma que ofrecen un rendimiento superior en comparación con sistemas que dependen de un solo sensor. Sin embargo, no se han enfocado en la estimación precisa de la profundidad de los daños para evaluar las condiciones de la carretera.

A pesar de los avances, los autores de los artículos subrayan desafíos pendientes como la generalización de los sistemas para su implementación práctica y la gestión de la creciente demanda de procesamiento de datos en grandes y complejos conjuntos de nubes de puntos. Además, coinciden en que la investigación futura deberá

enfocarse en la mejora de la robustez de los modelos en variadas condiciones ambientales y en la exploración de técnicas que puedan aprovechar la rica información proporcionada por la fusión de nubes de puntos e imágenes para una caracterización detallada y precisa de los deterioros del pavimento.

2. DESCRIPCIÓN DE METODOLOGÍA

Esta sección describe detalladamente los pasos para estimar la profundidad de los baches. En la figura 1 se ilustra un diagrama general de la metodología empleada, que comienza con la recolección de muestras en carreteras, la extracción de datos de imágenes y nubes de puntos. Se procede a detectar baches en las imágenes y se aplica un preprocesamiento a las nubes de puntos para reducción de ruido y nivelación de la superficie. Las regiones de interés identificadas en las imágenes sirven para delimitar los deterioros en las nubes de puntos, sobre las cuales se aplica el algoritmo de estimación de profundidad. Este proceso culmina con la obtención de la profundidad específica de cada bache detectado.

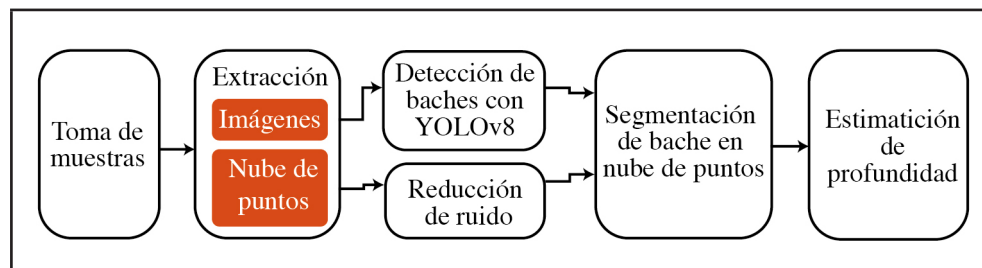


FIGURA 1

Metodología implementada para la estimación de profundidad de baches

Fuente: elaboración propia.

2. 1. Etapa 1: toma de muestras

Para la toma de muestras se empleó un sensor modelo Real Sense 435i de intel, el cual cuenta con una cámara estéreo y un sensor infrarrojo. Las especificaciones de este dispositivo son las siguientes:

- a) Intervalo de operación (mínimo-máximo): $\sim 3\text{m}-3\text{m}$
- b) Resolución de imagen: 1280×720
- c) Campo de visión profundo: 85.2×58
- d) Distancia de profundidad mínima a resolución máxima: 28 cm
- e) Precisión de profundidad: $<2\%$ a 2m de altura

Las muestras de pavimentos se llevaron a cabo con el apoyo de un trípode y con el recorrido de distintas calles de la ciudad de Culiacán, Sinaloa.

En la etapa inicial se condujeron pruebas de repetibilidad para evaluar la precisión y exactitud de la profundidad registrada por la cámara de profundidad para determinar la altura con menor desviación en la estimación. Para ello, se obtuvieron seis muestras de un bache con el sensor desde una posición fija. Tras cada conjunto de mediciones, se reposicionaba el sensor para capturar nuevamente la escena seis veces (repitiendo este proceso en tres ocasiones). Este procedimiento se efectuó a diferentes alturas: 85 cm, 90 cm y 95 cm.

Para comparar la estimación con las medidas reales se realizaron mediciones manuales en siete baches usando un vernier, tal y como se muestra en la figura 2. Se identifica el punto más profundo, se apoya una regla en los bordes del bache para determinar el nivel de la superficie y se utiliza la sonda del vernier que mide profundidad desde el punto más bajo dentro del bache hasta la superficie del pavimento. Lo anterior siguiendo la normativa

M.CSV.CAR.1.03.008.23. Determinación de los Deterioros Superficiales de los Pavimentos (DET) de la SICT, que es un manual en donde se describe la clasificación e interpretación detallada de los deterioros o fallas más comunes que se presentan en la superficie de los pavimentos flexibles y rígidos.

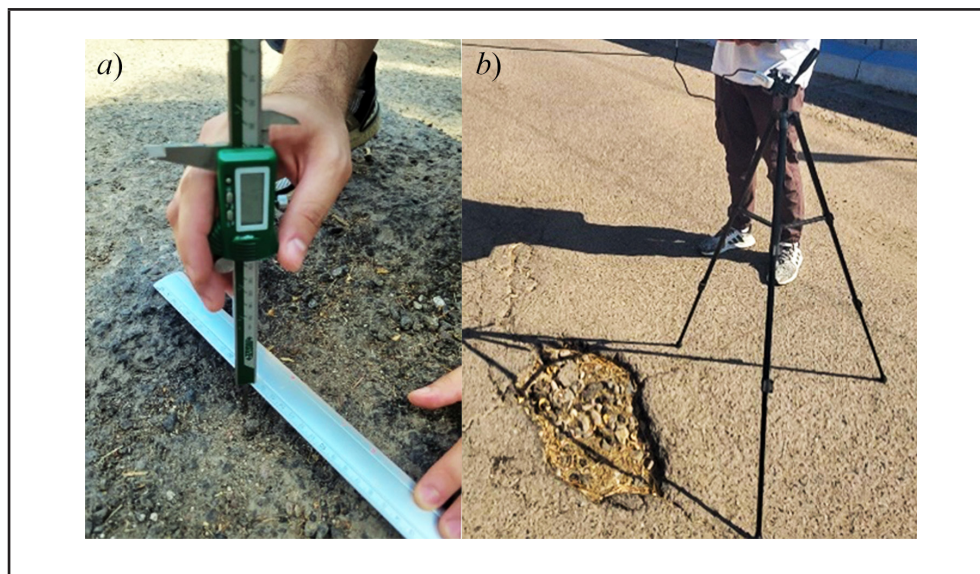


FIGURA 2

a) Medición de profundidad, b) toma de muestras con camara de profundidad

Fuente: elaboración propia.

Las pruebas de repetibilidad se evalúan con la desviación estándar y el error estándar de la media. La desviación estándar (SD) mide la cantidad de variación o dispersión de un conjunto de valores, en este caso las mediciones de profundidad de las distintas muestras. Una desviación estándar baja indica que los valores tienden a estar cerca del promedio del conjunto, mientras que una desviación estándar alta indica que los valores están esparcidos en un rango más amplio (ecuación 1):

$$SD = \frac{\sum (x_i - \bar{x})}{n - 1} \quad (1)$$

Donde x_i son las mediciones individuales, $\bar{x}$ es el promedio de las mediciones y n es el número total de mediciones.

El error estándar de la medida (EEM) se utiliza para estimar cuán cerca está el promedio de las mediciones del “verdadero valor” teórico, en este caso qué tan cerca está la estimación de la profundidad realizada en comparación con la profundidad real. Se calcula como la desviación estándar de las mediciones dividida por la raíz cuadrada del número de mediciones (ecuación 2):

$$EEM = \frac{\sigma}{\sqrt{n}} \quad (2)$$

Donde σ es la desviación estándar de las mediciones y n es el número total de mediciones.

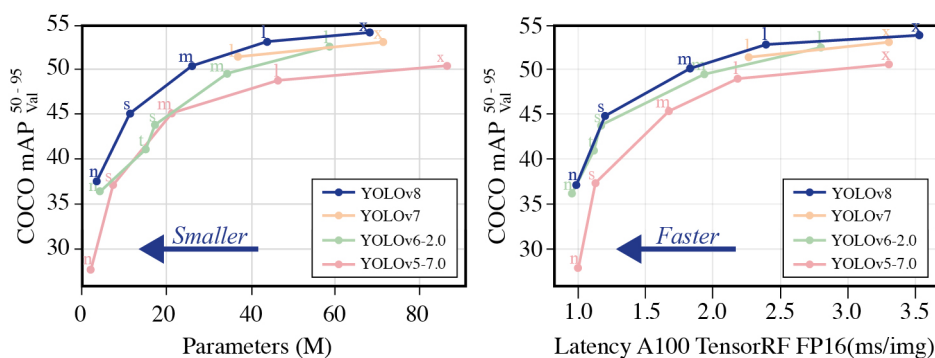
Por ejemplo, si mide la profundidad de un bache 30 veces ($n = 30$) y la desviación estándar de estas mediciones es de 2 mm, el error estándar será $230 \approx 0.36530 \approx 0.365$ mm.

2. 2. Etapa 2: extracción de imágenes y nubes de puntos

En este proceso se empleó la biblioteca `pyrealsense2` para la interacción con los archivos generados de la cámara de profundidad, los cuales contienen datos secuenciados de video y profundidad capturados por estos dispositivos. Mediante un *script* se extrajeron cuadros de color (RGB) y de profundidad de los mencionados archivos, con cada cuadro de profundidad ofreciendo una matriz de valores que cuantifica la distancia desde los píxeles individuales hasta la cámara que proporciona así un detalle dimensional crítico para la reconstrucción de la escena tridimensional.

2. 3. Etapa 3: detección de baches con YOLOv8

La red YOLOv8 ha sido elegida para la detección de baches debido al equilibrio que ofrece entre las métricas precisión y eficiencia en comparación con modelos, tales como YOLOv7, YOLOv6-2.0 y YOLOv5-7.0. En la gráfica 1 se observa que YOLOv8 alcanza un mayor MAP (Mean Average Precision) con un menor número de parámetros, lo cual indica una mayor eficiencia en términos de tamaño del modelo. Además, YOLOv8 ofrece una latencia menor, lo que significa que es más rápido en la inferencia mientras mantiene una alta precisión. Estas características hacen que YOLOv8 sea ideal para aplicaciones que requieren tanto alta precisión como rapidez, como lo es el caso de la detección de baches.



GRÁFICA 1

Comparación de rendimiento entre los modelos YOLOv5, YOLOv6, YOLOv7 y YOLOv8

Fuente: Jocher *et al.*, 2023.

En la figura 3 se muestra el proceso para obtener el modelo de inferencia de baches. A partir de una secuencia de video, se extrajeron fotogramas individuales, con los cuales se separaron manualmente las imágenes de baches. Estas imágenes se dividieron en tres conjuntos de datos (entrenamiento, validación y pruebas) en una proporción de 70%, 20% y 10%. Debido a que YOLOv8 realiza un entrenamiento supervisado, es necesario delimitar el área de interés. La delimitación de estas zonas resulta en una etiqueta de imagen que contiene las coordenadas del rectángulo que marca la zona de interés. Estas coordenadas se utilizan como entrada junto con la imagen para la red neuronal que le permiten identificar y enfocarse en la parte específica de la imagen donde se encuentra el deterioro por analizar. Este etiquetado se hizo en la aplicación web Roboflow (2022).

Aunque los modelos de aprendizaje profundo pueden entrenarse con datos sin preprocesar, incluir una fase de preprocesamiento antes del entrenamiento mejora significativamente la calidad de los datos y, por ende, la precisión del modelo (Ranganathan *et al.*, 2021). Las imágenes en su estado crudo a menudo contienen ruido, como sombras o iluminación inconsistente que obstaculizan la capacidad del modelo para identificar y segmentar correctamente los objetos de interés. La implementación de técnicas de preprocesamiento, tales como ajustes de contraste o la conversión a escala de grises, optimiza la calidad de las imágenes y facilita la detección precisa de los objetos relevantes.

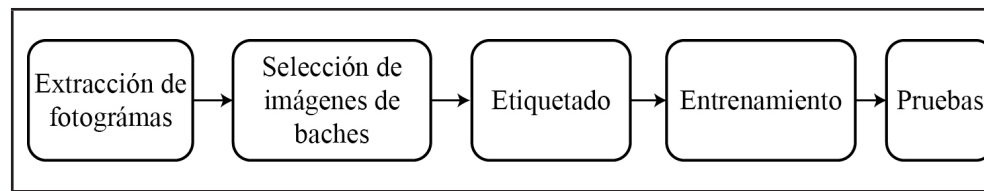


FIGURA 3

Proceso para obtener modelo de inferencia

Fuente: elaboración propia.

La herramienta utilizada permite generar varios conjuntos de datos a partir de las mismas imágenes con la aplicación de diferentes técnicas de preprocesamiento y aumento de datos. Esto incluye el uso de varios filtros, ajustes de contraste, rotaciones y recortes que resultan en conjuntos de datos que, aunque basados en las mismas imágenes, presentan variaciones en su tratamiento. Este enfoque se empleó específicamente para acentuar las características de los baches en relación con la superficie del pavimento. Los filtros aplicados incluyen:

a) Escala de grises: esto aumenta la eficiencia en el procesamiento de datos y reduce la cantidad de información a procesar para la red neuronal. Al eliminar la información de color, se disminuye el tamaño de los archivos y se eliminan redundancias, lo que acelera el entrenamiento y simplifica el modelo, facilitando una ejecución más rápida y eficiente.

b) Rotación: este proceso implica girar las imágenes originales un ángulo específico alrededor de su centro. Para este estudio, cada imagen se rotó 90° en sentido horario y se repitió el procedimiento hasta alcanzar un giro total de 270° . Así, se generaron tres imágenes rotadas adicionales a partir de cada imagen original.

Los parámetros del entrenamiento se muestran en la tabla 1 y la tabla 2. La primera muestra los componentes de *hardware* y la segunda los hiperparámetros utilizados para la configuración del entrenamiento.

TABLA 1

Componentes de *hardware*

Componente	Descripción
Procesador	Intel i7 12000k
Grafica	2 RTX 4070 ti
Memoria RAM	64 GB

Fuente: elaboración propia.

TABLA 2

Parámetros de entrenamiento

Hiperparámetros	Descripción
Épocas	130
Taza de aprendizaje	0.001
Batch	16
Tamaño imágenes	640×640
Parámetros	3151904
Capas	168
Optimizador	SDG (Descenso del gradiente estocástico)

Fuente: elaboración propia.

Una vez terminado el entrenamiento, el modelo es capaz de realizar inferencias sobre imágenes al identificar las regiones donde se ubican baches. Como resultado, proporciona coordenadas de la caja o cajas delimitadoras que marcan las zonas específicas de cada bache detectado. La figura 4 muestra un ejemplo de lo descrito; se observa la imagen que el modelo procesó, la caja delimitadora identificando al bache y en la parte derecha las coordenadas que utiliza para dibujar dicha caja.

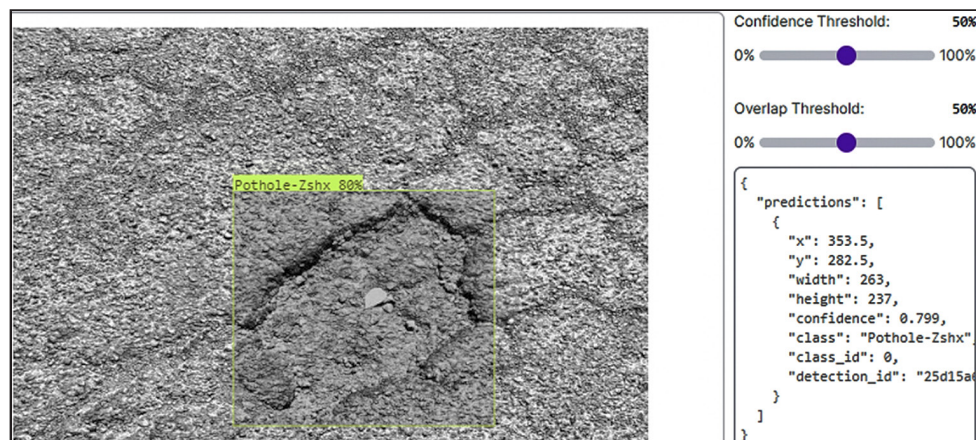


FIGURA 4
Ejemplo de inferencia realizada por el modelo
Fuente: elaboración propia.

2. 4. Etapa 4:

Reducción de ruido

En la fase preliminar de nuestro estudio, el preprocesamiento de los datos recogidos por la cámara de profundidad es crucial para asegurar la fiabilidad y precisión de nuestros análisis posteriores. Este preprocesamiento aborda dos desafíos principales inherentes a la adquisición dinámica de datos tridimensionales en entornos reales:

1. La presencia de ruido y puntos atípicos en la nube de puntos 3D, generados por condiciones de iluminación variable, reflejos superficiales y limitaciones propias del sensor de profundidad.
2. La variabilidad en la orientación e inclinación de la superficie del pavimento, ya que las carreteras no son perfectamente horizontales y pueden presentar pendientes longitudinales o transversales.

Eliminación de puntos atípicos

la cámara de profundidad, en su proceso de captura de datos, puede incluir erróneamente puntos atípicos debido a la iluminación o reflejos en superficies. Estos puntos atípicos se registran por encima de la superficie, por lo que si no se elimina, llegan a distorsionar el análisis. Para solucionar esto se emplea un enfoque estadístico para purificar la nube de puntos 3D de datos atípicos. Basándose en la distancia entre un punto y sus vecinos cercanos, los puntos se clasifican como atípicos si se desvían significativamente de la media.

Ya que las superficies de las carreteras no siempre son horizontales, es esencial hacer una rotación de estas superficies para alinearlas con un sistema de coordenadas común. Este proceso asegura la consistencia en todas las muestras al facilitar análisis comparativos y precisos. La nivelación permite que las mediciones de características como la profundidad de los baches sean coherentes y confiables, independientemente de la inclinación original de la carretera. Para abordar este problema, se utiliza el algoritmo RANSAC para ajustar la orientación de la nube de puntos, de modo que refleje una superficie nivelada. El algoritmo RANSAC (Random Sample Consensus) se

emplea para identificar un plano que mejor se ajuste a la superficie del terreno en la nube de puntos (Fischler y Bolles, 1987). El proceso se desarrolla en los siguientes pasos:

a) Estimación del plano: se selecciona un subconjunto aleatorio de puntos de la nube para estimar un plano. Este plano en el espacio tridimensional se define mediante la ecuación (A, B, C) , donde (A, B, C) es el vector normal al plano y D la distancia del plano al origen (ecuación 3).

$$Ax + By + Cz + D = 0 \quad (3)$$

Donde (A, B, C) es el vector normal al plano y D la distancia del plano al origen.

b) Determinación de *inliers*: calculamos la distancia de cada punto de la nube al plano estimado. Aquellos puntos cuya distancia es inferior a un umbral predefinido, denominado *distancia_thresh*, se clasifican como *inliers*, es decir, forman parte del plano.

c) Iteración y selección del mejor plano: el algoritmo RANSAC repite este proceso múltiples veces con diferentes conjuntos aleatorios de puntos. El plano que acumula el mayor número de *inliers* a lo largo de estas iteraciones es seleccionado como el plano óptimo.

2. 5. Etapa 5: segmentación de bache en nube de puntos

En esta etapa se utilizan los datos obtenidos de dos procesos anteriores: las nubes de puntos depuradas mediante filtros de reducción de ruido y las detecciones de baches realizadas en las imágenes. Cada imagen está vinculada a una nube de puntos, lo cual proporciona así una representación tridimensional de la superficie donde se ha detectado el bache. Durante la detección con YOLOv8, se almacenan las coordenadas del área donde se identificó el deterioro. Estas coordenadas son esenciales para localizar la zona específica en la nube de puntos.

2. 6. Etapa 6: estimación de profundidad

Para estimar la profundidad de los baches, se emplea la nube de puntos ya segmentada y nivelada del bache, una vez que se ha aplicado previamente un proceso de filtrado para eliminar puntos atípicos. Como se ilustra en la figura 5, el procedimiento se inicia con la estimación de la distancia del sensor a la superficie del terreno. Este paso se realiza calculando la mediana de las distancias en Z de los puntos en la nube de puntos, que representa así la distancia a la que se encuentra la superficie del pavimento. Luego, se identifica el punto más alejado de esta superficie en la nube de puntos. La profundidad del bache se determina obteniendo la diferencia entre la distancia del sensor a la superficie del pavimento y la distancia al punto más alejado del nivel de la superficie (ecuación 4).

$$\text{Profundidad} = Z_{\text{superficie}} - Z_{\text{punto más bajo}} \quad (4)$$

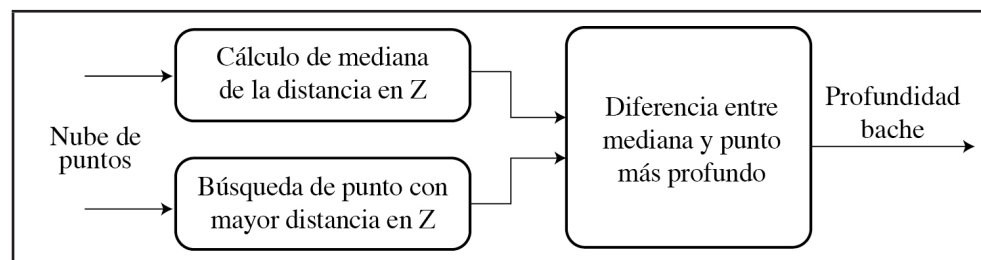


FIGURA 5

Proceso 3. Proceso de estimación de bache a partir de nube de puntos

Fuente: elaboración propia.

3. RESULTADOS

3. 1. Toma de muestras y extracción de nube de puntos e imágenes

La toma de muestras se llevó a cabo en calles de la ciudad de Culiacán, Sinaloa, entre las 8 y las 11 de la mañana, que presentó un clima soleado. El pavimento captado en las imágenes incluye tanto superficies flexibles como rígidas, además de la presencia de sombras, manchas en las calles y cambios de iluminación, lo que indica que el ambiente no era controlado. Estas condiciones contribuyen a que el conjunto de datos generado sea robusto, pues proporcionan al modelo de inferencia varios entornos donde identificar las características de las zonas de interés. Una vez hechos los recorridos de las calles para capturar imágenes del pavimento, se recopilaron 273 muestras, divididas de la siguiente forma: 173 para entrenamiento, 50 para validación y 50 para pruebas. La figura 6 exhibe muestras de imágenes que pertenecen al conjunto de datos.



FIGURA 6
Muestras de imágenes de conjunto de datos
Fuente: elaboración propia.

Para la prueba de repetibilidad se utilizaron seis muestras tomadas a un bache para determinar la altura en la cual obtener la mayor exactitud de profundidad (cuadro 1, cuadro 2 y cuadro 3). Cada cuadro registra las mediciones a distintas alturas 85, 90 y 95 cm respectivamente; asimismo, se tomaron seis mediciones en tres posiciones distintas.

CUADRO 1
Resultados de estimación de profundidad de 85cm

85 cm	Toma 1	Toma 2	Toma 3	Toma 4	Toma 5	Toma 6	Moda	Media	Desviación estándar
Posición 1	58.2	58.2	60.2	58.2	61.1	61.1	58.2	58.2	1.379130161
Posición 2	58.2	58.2	56.3	56.3	55.3	58.2	58.2	57.25	1.276583983
Posición 3	63.1	59.2	59.2	58.2	58.2	58.2	58.2	58.7	1.901315334

Fuente: elaboración propia.

CUADRO 2
Resultados de estimación de profundidad a 90 cm

90 cm	Toma 1	Toma 2	Toma 3	Toma 4	Toma 5	Toma 6	Moda	Media	Desviación estándar
Posición 1	59.5	61.4	59.5	62.4	60.4	60.4	60.4	60.4	1.1296017
Posición 2	61.4	61.4	61.4	62.4	60.4	57.5	61.4	61.4	1.882020191
Posición 3	65.3	64.4	65.3	60.4	64.4	60.4	65.3	64.4	2.33295235

Fuente: elaboración propia.

CUADRO 3
Resultados de estimación de profundidad a 95 cm

95 cm	Toma 1	Toma 2	Toma 3	Toma 4	Toma 5	Toma 6	Moda	Media	Desviación estándar
Posición 1	60.7	60.7	60.7	61.7	58.8	60.6	60.7	60.7	1.21209185
Posición 2	62.7	62.7	63.6	60.7	62.7	59.8	62.7	62.7	1.45281336
Posición 3	59.8	58.8	59.8	60.7	60.7	59.8	59.8	59.8	1.901315334

Fuente: elaboración propia.

3. 2. Detección de baches con YOLOv8

Para probar el modelo de inferencia, se utilizó el conjunto de datos de prueba. Este conjunto contenía imágenes no exhibidas en el entrenamiento. La figura 7 condensa seis distintas muestras del conjunto de pruebas y el resultado de procesarlas por el modelo de inferencia, además se observa el comportamiento del modelo ante distintas situaciones de textura, iluminación y tamaño en los baches.

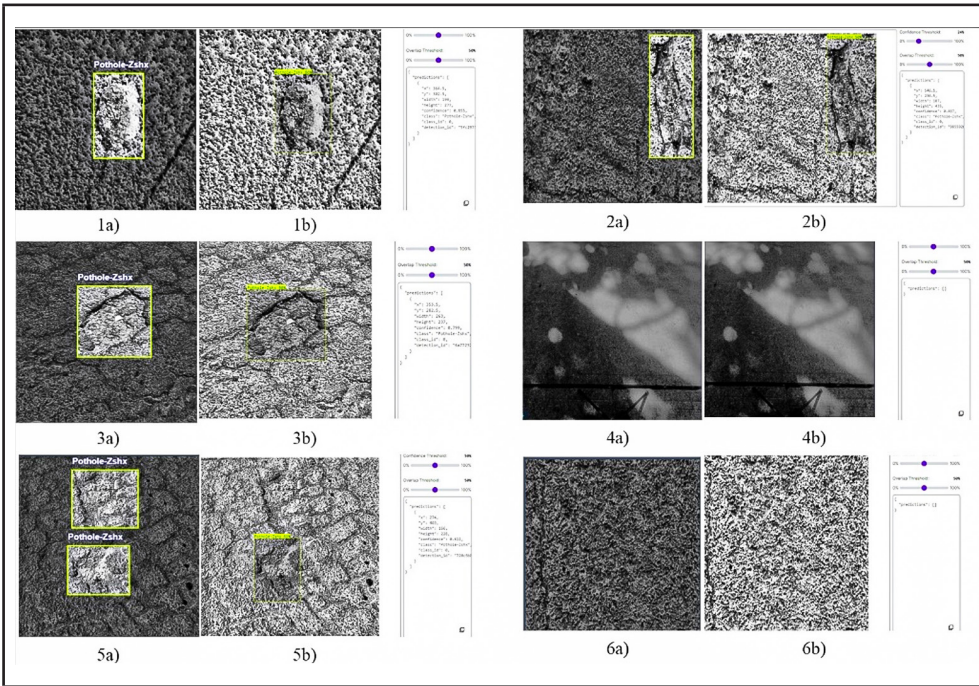


FIGURA 7
Detección de baches con YOLOv8
Fuente: elaboración propia.

Nota: las muestras a) corresponden a la imagen con su etiqueta manual y las muestras b) la inferencia del modelo.

Se incluyeron imágenes de baches, así como una imagen de ejemplo de las muestras nulas. Con este conjunto se obtuvieron los siguientes resultados: mAP 84.7%, precisión 85.7% y *recall* 76.3%.

3. 3. Reducción de ruido, nivelación y recorte de nube de puntos

Después de aplicar filtros para eliminar puntos atípicos, nivelar el terreno con RANSAC y el recorte con las coordenadas obtenidas en la detección a las nubes de puntos, se notó un cambio significativo en la distribución de estos. La figura 8 muestra una comparación visual de la nube de puntos antes y después del preprocesamiento. En esta nube los colores representan diferentes niveles de profundidad: los tonos rojizos indican mayor profundidad, mientras que los tonos azules señalan áreas menos profundas. Se destaca un bache, marcado en rojo, que fue capturado desde un trípode. Por ello, hay una zona sin color que corresponde a la superficie bajo los apoyos del trípode.

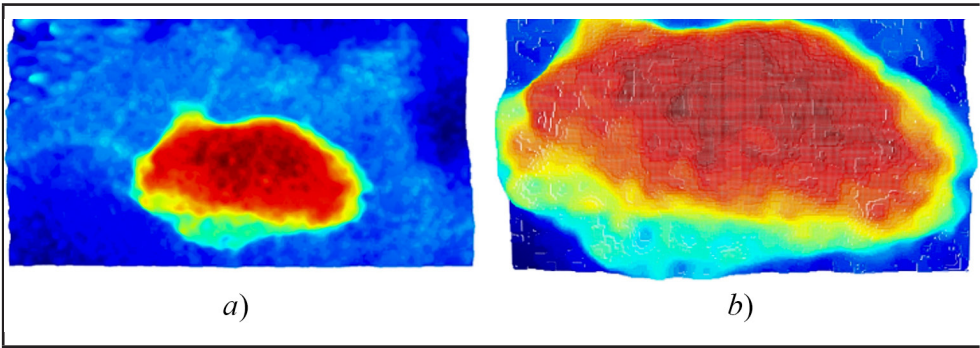


FIGURA 8
Comparativa antes y después de eliminación de puntos atípicos y recorte
Fuente: elaboración propia.
Nota: a) antes, b) después.

3. 4. Estimación de profundidad

El cuadro 4 registra mediciones de profundidad para ocho baches distintos que revelan tanto consistencias como discrepancias importantes en los datos recolectados. Cada bache tiene un número variado de mediciones, con un rango de 31 a 147. En general, los errores estándar son bajos, lo cual indica que las medias son representativas de los conjuntos de datos.

CUADRO 4
Resultados de estimación de profundidad

Bache	Muestra	Estimación de profundidad promedio (milímetros)	Moda estimada	Medición real [mm]	Error = estimación-medición real [mm]	Error estándar (EEM) [mm]	Dispersión (desviación estándar) [mm]
1	42	36.2	36.2	33	-3.2	0.0978	0.5691
2	74	32.2	32.2	28.3	-3.9	0.0442	0.3800
3	55	29.6	29.6	45.3	15.7	0.0866	0.6421
4	50	25.1	25.1	32.3	7.2	0.0595	0.4209
5	147	45.0	45.0	54.1	9.1	0.1870	2.2669
6	44	31.0	26.6	32	1	0.0587	0.3897
7	31	17.9	17.9	22.9	5	0.1230	0.6847

Fuente: elaboración propia.

Para los baches uno a cuatro, hay una notable precisión entre las estimaciones promedio y las mediciones reales, con la moda coincidiendo exactamente con la estimación promedio, lo que sugiere una distribución de mediciones concentrada y consistente. Sin embargo, el bache cinco muestra una desviación significativa entre la estimación y la medición real; esto podría indicar un error sistemático o una variabilidad en el proceso de medición.

La dispersión en las mediciones, medida como desviación estándar, varía significativamente; en este caso, el bache cinco es el más variable. Esto plantea preguntas sobre la consistencia del proceso de medición o sobre la naturaleza misma de este bache específico, que parece tener una variabilidad mucho mayor en comparación con los demás.

La cercanía entre la moda y el promedio significa que las estimaciones de profundidad son precisas y repetibles. En otras palabras, la mayoría de las mediciones coinciden estrechamente con el valor promedio, lo cual sugiere que nuestro método de estimación produce resultados fiables.

En el contexto de la estimación de profundidades de baches, esta precisión es crucial. Las estimaciones precisas y consistentes son esenciales para la planificación de reparaciones y mantenimiento, ya que permiten una evaluación fiable del estado de la carretera.

TABLA 3
Comparativa entre trabajos del estado del arte y este trabajo

Autor	Dispositivo	Costo (USD)	Error en profundidad
Wu <i>et al.</i> , 2019	SICK LMS 511	10 000	15 a 27 mm
	HiPRoSS (High Performance Road Survey System)	Aprox. 1 000 000	1 mm
De Blasiis <i>et al.</i> , 2020	2 rieg1 VQ-450 laser	No especificado	5 mm
Peralta-López <i>et al.</i> , 2023	ZED Camera	449	No estima profundidad
Este trabajo	Intel D435i	400	5 mm

Fuente: elaboración propia.

4. DISCUSIÓN DE RESULTADOS

La metodología propuesta ofrece resultados competitivos en comparación con trabajos del estado del arte incluidos en la tabla 3 al considerarse la identificación de baches, la cual supera 85% en la precisión, 84% en el mAP y 75% en *recall*. La precisión combinada con un *recall* mayor a 75% sugiere que el modelo es adecuado en detectar baches sin generar un alto número de falsos positivos, aunque todavía hay espacio para mejorar en la detección completa de todos los baches presentes, como se constata en la figura 7.

Sin embargo, donde se observa el valor de este trabajo es en la estimación de la profundidad debido a que esta información es necesaria para determinar la severidad de los baches. Con las pruebas de repetibilidad se encontró que el error promedio de las mediciones de la cámara de profundidad fue de 5 mm. Esta medida no representa una desviación significativa tomando en cuenta que para considerar que un desperfecto en la carretera sea un bache su profundidad debe ser mayor a 10 mm, como lo dice el manual *M.CSV.CAR.1-03.008/23* de la SICT.

PROSPECTIVA

La estimación de profundidad ha sido un tema estudiado durante los últimos años en el área de detecciones automáticas de deterioros en pavimentos, dada su importancia para determinar la severidad del daño para la estructura del firme. Por ello, se ha buscado estimar la profundidad por medio de distintos dispositivos, como cámaras, sensores LiDAR, sensores láser o infrarrojos, cada uno con sus ventajas y desventajas. Sin embargo, la

fusión de datos de distintas fuentes permite hacer una caracterización más fiable a la realidad de los daños, además de analizarlos de distintas perspectivas, como lo es una imagen en 2D o una reconstrucción en 3D. Esto hace posible la combinación de distintas técnicas en una misma metodología que soluciones problemas desde distintos enfoques. Sin embargo, aún se deben hacer investigaciones para fortalecer las debilidades existentes, como es el caso de la cantidad de datos que se requieren para todo el proceso y también su almacenamiento.

CONCLUSIONES

Este estudio aporta un enfoque para la detección automática de baches y en la estimación de su profundidad mediante el uso de la cámara de profundidad Real Sense D435i y tecnologías de aprendizaje profundo para su uso en el auscultamiento de carreteras. La metodología empleada ha demostrado la capacidad de identificar baches y estimar su profundidad en condiciones ambientales no controladas. Esta información es indispensable para la evaluación de tramos carreteros dentro de las actividades de mantenimiento de infraestructuras viales. El uso de esta metodología ayudará en la planificación de mantenimientos, puesto que detecta los daños en etapas tempranas donde su severidad no representa un gran peligro para las condiciones de la carretera ni un riesgo para los conductores. De esta forma, es posible que se hagan las reparaciones pertinentes y prolongar el ciclo de vida de las carreteras, además de reducir el riesgo al personal encargado de las inspecciones.

Este enfoque plantea un camino hacia la automatización en la evaluación de la integridad del pavimento, con potencial para su integración en sistemas de monitoreo y mantenimiento de carreteras. La eficiencia y capacidad de generalización de nuestro modelo presenta nuevas posibilidades en la gestión de infraestructuras y la seguridad vial. Para futuros trabajos, se sugiere investigar la implementación de este sistema ante una mayor diversidad de condiciones ambientales y tipos de carreteras, así como su integración con otras tecnologías como vehículos autónomos y sistemas de gestión de tráfico. La expansión del conjunto de datos y la mejora continua de los algoritmos de aprendizaje profundo contribuirán a la robustez del sistema en la gestión de infraestructuras.

AGRADECIMIENTOS

Agradecemos a los árbitros por sus valiosas sugerencias y comentarios que contribuyeron significativamente a mejorar la estructura y claridad de este artículo.

Este trabajo ha sido posible gracias al financiamiento de la beca 1082905 por parte de CONAHCYT, así como del proyecto de Tecnológico Nacional de México *Sistema de identificación y clasificación automatizada de deterioros superficiales de pavimentos flexibles basado en técnicas de aprendizaje profundo y visión artificial*, clave 17982.23-P y del Proyecto 613 “Sistemas Ciber-físicos para el Desarrollo de Sistemas Inteligentes de Transporte” por parte de CONAHCYT.

REFERENCIAS

- Cui, Y., Chen, R., Chu, W., Chen, L., Tian, D., Li, Y., & Cao, D. (2021). Deep learning for image and point cloud fusion in autonomous driving: A review. *IEEE Transactions on Intelligent Transportation Systems*, 23 (2), 722-739.
- De Blasiis, M. R., Di Benedetto, A., & Fiani, M. (2020). Mobile laser scanning data for the evaluation of pavement surface distress. *Remote Sensing*, 12 (6), 942.
- Roboflow. (2022). (Version 1.0) [Software]. <https://roboflow.com.computer vision>.
- Elamin, A., & El-Rabbany, A. (2023). UAV-Based Image and LiDAR Fusion for Pavement Crack Segmentation. *Sensors*, 23(23), 9315.

- Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), 381-395.
- Guan, J., Yang, X., Ding, L., Cheng, X., Lee, V. C., & Jin, C. (2021). Automated pixel-level pavement distress detection based on stereo vision and deep learning. *Automation in Construction*, 129, 103788.
- Han, H., Deng, H., Dong, Q., Gu, X., Zhang, T., & Wang, Y. (2021). An advanced Otsu method integrated with edge detection and decision tree for crack detection in highway transportation infrastructure. *Advances in Materials Science and Engineering*, 2021(1), 9205509.
- Jocher, G., Chaurasia, A., & Qiu, J. (2023). YOLO by Ultralytics (Version 8.0.0) [Computer software]. <https://github.com/ultralytics/ultralytics>
- Kang, B. H., & Choi, S. I. (2017, July). Pothole detection system using 2D LiDAR and camera. In *2017 ninth international conference on ubiquitous and future networks (ICUFN)* (pp. 744-746). IEEE.
- Ma, D., Fang, H., Xue, B., Wang, F., Msekh, M. A., & Chan, C. L. (2020). Intelligent detection model based on a fully convolutional neural network for pavement cracks. *CMES-Computer Modeling in Engineering & Sciences*, 123 (3).
- Mednis, A., Strazdins, G., Zviedris, R., Kanonirs, G., & Selavo, L. (2011, June). Real time pothole detection using android smartphones with accelerometers. In *2011 International conference on distributed computing in sensor systems and workshops (DCOSS)* (pp. 1-6). IEEE.
- Peralta-López, J. E., Morales-Viscaya, J. A., Lázaro-Mata, D., Villaseñor-Aguilar, M. J., Prado-Olivarez, J., Pérez-Pinal, F. J., ... & Barranco-Gutiérrez, A. I. (2023). Speed bump and pothole detection using deep neural network with images captured through zed camera. *Applied Sciences*, 13(14), 8349.
- Ranganathan, D. G. (2021). A study to find facts behind preprocessing on deep learning algorithms. *Journal of Innovative Image Processing*, 3(1), 66-74.
- SICT (Secretaría de Infraestructura, Comunicaciones y Transporte). (2023). CSV. CONSERVACIÓN. CAR. Carreteras. EVALUACIÓN. Evaluación de pavimentos. 008. Determinación de los Deterioros Superficiales de los Pavimentos (DET). *M.CSV.CAR.1-03-008/23*. Gobierno de México. <https://es.scribd.com/document/715449972/M-CSV-CAR-1-03-008-23#page=1>
- Wu, H., Yao, L., Xu, Z., Li, Y., Ao, X., Chen, Q. ..., & Meng, B. (2019). Road pothole extraction and safety evaluation by integration of point cloud and images derived from mobile mapping sensors. *Advanced Engineering Informatics*, 42, 100936.
- Xu, W., Tang, Z., Zhou, J., & Ding, J. (2013, September). *Pavement crack detection based on saliency and statistical features*. In 2013 IEEE international conference on image processing (pp. 4093-4097). IEEE.

CC BY-NC-ND



Disponible en: <https://cienciaergosum.uaemex.mx/article/view/22933>

Cómo citar el artículo:

Román Garay, M. A., Hernández Beltrán, C. A., Morales Rosales, L., Rodríguez Rangel, H., Villa Camacho, M. A., Soto Audelo, J. L. y Lepej, P., (2025). Detección inteligente de baches y estimación de su profundidad con aprendizaje profundo para la conservación de carreteras. *CIENCIA ergo-sum*, 32. <https://doi.org/10.30878/ces.v32n0a37>

DOI: <https://doi.org/10.30878/ces.v32n0a37>



Licencia Creative Commons Atribución-No-Comercial-SinDerivar 4.0 Internacional.

Declaración de autoría (CRediT):

Mario Alberto Román Garay: investigación, conceptualización, metodología, *software*, redacción-borrador original, redacción-revisión y edición.

Carlos Alberto Hernández Beltrán: conceptualización, *software*.

Luis Morales Rosales: supervisión, redacción-revisión y edición.

Héctor Rodríguez Rangel: redacción-revisión y edición.

Villa Camacho: investigación, curación de datos, visualización.

José A. Soto Audelo: investigación, curación de datos, visualización.

Peter Lepej: conceptualización, visualización.