






Análisis estructural de baches viales mediante redes neuronales profundas

Structural Analysis of Road Potholes Using Deep Neural Networks

 Angie-P. Rique-Sabogal¹;  Miguel-A. Guatame-Medina¹;  Cristhian-F. Moreno-Manrique¹;
 Fabián-R. Jiménez-López²;  Andrés-F. Jiménez-López³

¹Fundación Universitaria Compensar, Bogotá-Colombia

²Universidad Pedagógica y Tecnológica de Colombia, Tunja-Colombia

³Universidad de los Llanos, Villavicencio-Colombia

Correspondencia: ajimenez@unillanos.edu.co

Recibido: 1 Octubre 2024

Aceptado: 14 Mayo 2025

Disponible: 20 Junio 2025

Cómo citar / How to cite

A.-P. Rique-Sabogal, M.-A. Guatame-Medina, C.-F. Moreno-Manrique, F.-R. Jiménez-López, and A.-F. Jiménez-López, "Análisis estructural de baches viales mediante redes neuronales profundas," *TecnoLógicas*, vol. 28, no. 63, e3266, 2025. <https://doi.org/10.22430/22565337.3266>



Resumen

El aumento acelerado de la población mundial ha intensificado el tráfico vehicular, lo que plantea un desafío significativo para su gestión. En países en desarrollo como Colombia, los accidentes de tránsito presentan una alta tasa de mortalidad, atribuida en gran medida a defectos en las carreteras, como grietas y baches. Ante esta problemática, el objetivo de este estudio fue desarrollar un sistema automatizado para la detección de defectos en pavimentos urbanos, utilizando redes neuronales convolucionales (CNN, por sus siglas en inglés) para clasificar 11 tipos de fallas en las vías. La metodología empleada consistió en la creación de una base de datos a partir de imágenes de carreteras con defectos, la cual fue utilizada para entrenar los modelos de redes neuronales profundas. Se evaluaron dos optimizadores, SGDM y ADAM, aplicados sobre imágenes a color y en escala de grises, procesadas en MATLAB® y validadas por expertos en ingeniería civil. Los resultados mostraron que el optimizador SGDM alcanzó una precisión del 74.67 % con imágenes a color, mientras que ADAM obtuvo un desempeño del 52.51 % con imágenes en escala de grises. En general, los hallazgos confirmaron la viabilidad de las CNN y las técnicas de procesamiento digital de imágenes para automatizar el proceso de inspección de pavimentos y mejorar la eficiencia y precisión de la evaluación de la infraestructura vial. Finalmente, se concluye que el uso de redes neuronales profundas constituye una alternativa confiable para el desarrollo de sistemas inteligentes de gestión de pavimentos y respalda la implementación de soluciones basadas en datos para optimizar el mantenimiento de las vías urbanas.

Palabras clave

Infraestructura vial, mantenimiento predictivo, modelos de aprendizaje profundo, optimización de algoritmos, procesamiento digital de imágenes.

Abstract

The rapid growth of the global population has intensified vehicular traffic, posing a significant challenge to its management. In developing countries such as Colombia, traffic accidents exhibit high mortality rates, largely attributed to road defects, such as cracks and potholes. Given this problem, the objective of this study was to develop an automated system for detecting defects in urban pavements using Convolutional Neural Networks (CNNs) to classify 11 types of road surface failure. The methodology involved creating a dataset from images of defective roads, which was used to train deep neural network models. Two optimizers, SGDM and ADAM, were evaluated using color and grayscale pictures, processed in MATLAB® and validated by civil engineering experts. The results showed that the SGDM optimizer achieved an accuracy of 74.67 % with color images, while ADAM achieved a performance of 52.51 % with grayscale images. These findings demonstrated the potential of CNNs and digital image processing techniques to automate pavement inspection, increasing both efficiency and accuracy in evaluating road infrastructure. Finally, it is concluded that the use of deep neural networks represents a viable alternative for developing intelligent pavement management systems and supports the implementation of data-driven solutions to optimize urban road maintenance.

Keywords

Road infrastructure, predictive maintenance, deep learning models, algorithm optimization, digital image processing.

1. INTRODUCCIÓN

Las vías públicas constituyen un componente esencial de la infraestructura de transporte a nivel global, y su adecuado mantenimiento resulta crucial para garantizar la seguridad, eficiencia y comodidad en el tránsito de vehículos, mercancías y peatones [1], [2]. En países en desarrollo como Colombia, la presencia de baches y grietas en las carreteras no solo representan un riesgo para la seguridad de conductores y peatones, sino que también genera impactos económicos significativos debido a los daños vehiculares y los retrasos en la distribución de bienes y servicios [3], [4].

Las labores de mantenimiento y reparación de este tipo de imperfecciones son imprescindibles, no obstante, ello a la vez representa un reto de tipo logístico y económico para las agencias de transporte locales encargadas de la gestión del tráfico rodante. De ahí que el análisis estructural de los baches haya surgido como una tendencia de gran relevancia en el ámbito de la ingeniería civil, dado que tales estudios hacen posible que especialistas y técnicos dispongan de una explicación más profunda de las causas de la degradación de la red vial.

En un nivel más elevado, el enfoque de este estudio permite la implementación de estrategias de reparación y prevención más precisas y eficientes, maximizando de este modo el uso de los recursos relacionados disponibles desde el gobierno. Al mismo tiempo, los recientes avances en la inteligencia artificial en general y en particular la investigación de las redes neuronales convolucionales (CNN) han permitido abrir nuevas vías para el análisis automatizado de imágenes en los distintos campos de la ingeniería. Estas últimas han sido fuente de consolidación de herramientas útiles para la detección, clasificación y evaluación de defectos presentes en los pavimentos, aumentando así la eficiencia y efectividad de los procesos de inspección en vías pavimentadas [5]–[10].

Se han realizado distintos trabajos experimentales que han estudiado el uso y la capacidad de las CNN para detectar e identificar de manera automática defectos superficiales, ya sean grietas, fisuras o desprendimientos del pavimento, y cuyo resultado ha sido significativo para tareas de clasificación de imágenes [11]–[15]. Por el contrario, la mayoría de estos trabajos de investigación se han centrado sobre todo en identificar fallas superficiales y por consiguiente se ha menospreciado el análisis estructural profundo de los baches. Este tipo de análisis es clave para comprender la formación, la propagación y el alcance de los daños que sufren las vías, y así poder programar correctamente las intervenciones correctivas y preventivas y extender la vida útil de las infraestructuras de carreteras.

Por otro lado, recientes trabajos demuestran que la combinación de plataformas móviles y participación ciudadana trae consigo una gran mejora en la recolección de datos e incluso en la actividad de la gestión vial mediante la captura de imágenes en campo [16]–[22]. De igual

forma, la integración de tecnologías con algoritmos de aprendizaje de máquina ofrece la posibilidad de establecer un cambio en el procedimiento de realizar la inspección y el mantenimiento de las carreteras, tomando decisiones basadas en datos [23]–[30].

En ese contexto, en la presente investigación se propuso desarrollar un sistema basado en el modelo de redes neuronales profundas para la automatización del análisis estructural de baches existentes en pavimentos urbanos mediante imágenes capturadas en campo e incorporando una plataforma híbrida que involucra la participación ciudadana mediante el uso de una aplicación móvil. Esta plataforma permitirá a los usuarios reportar defectos viales, mientras que las autoridades recibirán un análisis detallado que respalde la toma de decisiones para la gestión y priorización del mantenimiento vial [30]–[35].

A continuación, en la sección 2 del documento se describe la metodología utilizada en el desarrollo del estudio, incluyendo los recursos de hardware y software utilizados, la forma en que se recolectaron y normalizaron las imágenes, la arquitectura de red CNN utilizada, su entrenamiento y evaluación. En la sección 3 se discuten los resultados y el análisis de las métricas de desempeño usadas para evaluar la eficacia del sistema de detección y clasificación automática de imperfecciones viales implementado. Finalmente, en la sección 4 se abordan las conclusiones y viabilidad del uso de estas tecnologías para mejorar los procesos de inspección y el mantenimiento de infraestructuras viales.

2. METODOLOGÍA

En este estudio, se utilizó un enfoque basado en la aplicación de modelos de redes neuronales convolucionales profundas (CNN) para el análisis estructural de baches e imperfecciones viales a partir de imágenes reales de carreteras urbanas. Las imágenes recolectadas como datos, fueron procesadas junto con el entrenamiento de los modelos de aprendizaje automático usando MATLAB® 2024a, en un equipo con las siguientes características: procesador AMD RYZEN 7 7730U, tarjeta gráfica Radeon Graphics de 2 GB, 16 GB de RAM y sistema operativo Windows 11 Pro (versión 23H2, 64 bits), lo que permitió aprovechar las capacidades del procesamiento paralelo, algo fundamental para el entrenamiento eficiente de la arquitectura de la red neuronal convolucional.

2.1 Recolección y preprocesamiento de datos

Se recolectaron 5000 imágenes de imperfecciones viales en la zona urbana de la ciudad de Bogotá, llevando a cabo la captura mediante estudiantes del semillero Bochica Software de la Fundación Universitaria Compensar. Las imágenes fueron tomadas con cámaras digitales de alta resolución (mínimo 12 MP) en diversas condiciones de iluminación (día y noche) y clima (seco, húmedo, lluvioso) para garantizar la viabilidad del conjunto de datos. Posteriormente, se realizó un proceso de filtrado de datos que excluyó las imágenes con más del 25 % de interferencias externas, como sombras pronunciadas, objetos ajenos a la vía (vehículos, personas) o vegetación que obstaculizara la visibilidad del pavimento.

Tras este proceso, se obtuvo una base de datos de 1001 imágenes aptas para el análisis. Cada imagen se normalizó a un tamaño estándar de 800 x 800 píxeles y se almacenó en formato PNG con resolución de 240 dpi para garantizar homogeneidad del conjunto de datos y facilitar el procesamiento. Las imágenes se clasificaron manualmente en 11 categorías de defectos viales siguiendo la nomenclatura propuesta en [11], [12], [32], [36] como: baches, pérdida de agregado, fisura transversal, fisura por reflexión, fisuras longitudinales y transversales, hundimientos, parches y "piel de cocodrilo" entre otras.

Con el fin de obtener muestras más detalladas para el entrenamiento de la red, cada imagen se dividió en parches de 100 x 100 píxeles, lo que generó un total de 199100 parches. Esta segmentación permitió capturar características locales de cada tipo de defecto vial.

2.2 Arquitectura de la red neuronal convolucional

La red neuronal convolucional profunda (CNN, por sus siglas en inglés) diseñada en este estudio estuvo compuesta por cuatro capas convolucionales y una capa completamente conectada para la salida. Se seleccionó este número de capas dado que, para detectar defectos viales, se buscó un equilibrio entre la extracción de características con el costo computacional sin incurrir en complejidad excesiva. Arquitecturas más complejas como ResNet o EfficientNet se descartaron por no ser necesarias para el tamaño y tipo de datos, logrando la CNN propuesta resultados satisfactorios en la clasificación de defectos en las carreteras.

La configuración de filtros por cada capa fue de 8, 16, 32, 64 y 128 filtros, respectivamente. En cada capa convolucional se aplicó la función de activación ReLU (Rectified Linear Unit, por sus siglas en inglés), que contribuye a mejorar la convergencia y el rendimiento del modelo [8], [10], [12], [19], [25], [31], [37].

El conjunto de 199100 parches se dividió aleatoriamente en tres subconjuntos, según la metodología estándar de entrenamiento supervisado así: 80 % para entrenamiento, 10 % para validación, y 10 % para prueba. Cada subconjunto incluyó imágenes que representaron las 11 categorías de defectos viales, garantizando un equilibrio en la distribución de clases y mejorar el aprendizaje del modelo [9], [19], [27], [35]–[39]. La Figura 1 ilustra el diagrama general de la arquitectura de la CNN implementada.

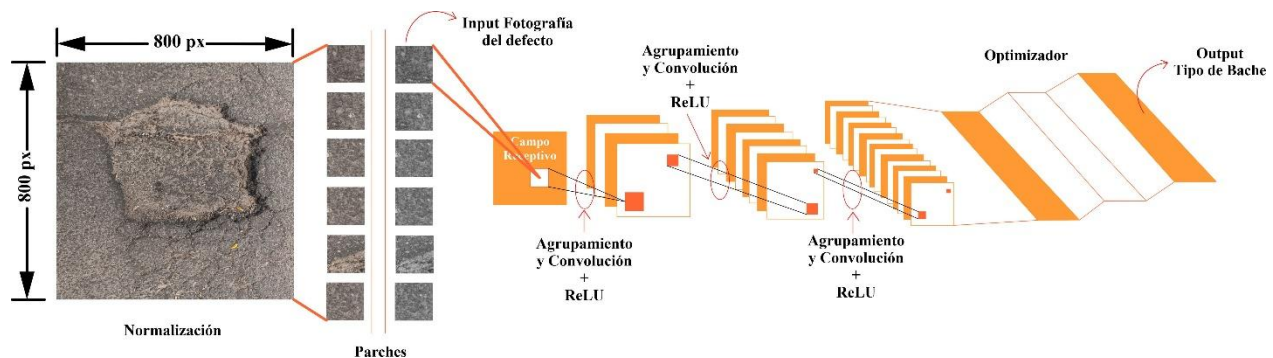


Figura 1. Diagrama general del modelo de Red Neuronal Convolucional implementado.

Fuente: elaboración propia.

2.3 Entrenamiento del modelo de la CNN

El modelo CNN se entrenó utilizando dos optimizadores: descenso de gradiente estocástico con momento (SGDM, por sus siglas en inglés) y la estimación adaptativa de momento (ADAM, por sus siglas en inglés) [20], [39]–[41].

Tras realizar pruebas experimentales del modelo de la CNN se eligieron tres épocas de entrenamiento con una frecuencia de validación cada 30 iteraciones para evitar el sobreajuste y optimizar el tiempo de cómputo requeridos, ya que se observó que el rendimiento del modelo se estabilizaba después de este punto, sin obtener mejoras significativas en su rendimiento con más épocas.

Se aplicaron técnicas de aumento de datos, como rotaciones aleatorias ($\pm 30^\circ$) y ajustes de brillo ($\pm 20^\circ$) y contraste ($\pm 15^\circ$), para incrementar la variabilidad del conjunto de datos y mejorar la robustez y capacidad de generalización del modelo. El conjunto de imágenes se procesó en formato RGB y en escala de grises, lo que permite investigar el efecto que tiene la representación de color sobre la precisión del modelo.

Durante el proceso de entrenamiento, se evaluó el rendimiento del modelo a partir de distintas métricas de evaluación habituales, como precisión, sensibilidad y especificidad; y se construyó una matriz de confusión por cada uno de los formatos (RGB y escala de grises), lo que permitió saber cuáles eran las categorías de defectos con los mayores errores de

clasificación, así como establecer un reajuste de los hiperparámetros de la red CNN a partir de los resultados obtenidos [14], [20], [34], [37], [39]–[42].

Los resultados se validaron estadísticamente llevando a cabo una prueba de hipótesis teniendo en cuenta la curva característica operativa del receptor (ROC, por sus siglas en inglés) y el cálculo del área bajo la curva (AUC, por sus siglas en inglés) para cada clase, lo cual permitió cuantificar de este modo el rendimiento del modelo y asegurar su reproducibilidad en ensayos posteriores.

2.4 Evaluación de desempeño del modelo de red CNN

El desempeño del modelo se evaluó con los subconjuntos de validación y prueba, en los cuales se compararon los resultados alcanzados con ambos optimizadores SGDM y ADAM. Para cada optimizador se estudiaron las métricas de precisión como los niveles de generalización sobre el conjunto de prueba que se presenta en [29]–[34], [42]–[46] tal como se comenta con posterioridad. También se contrastó el uso de imágenes en color RGB frente a imágenes en escala de gris, para calcular cuál de las dos modalidades era más eficiente con respecto a la precisión de clasificación del defecto vial.

2.5 Técnicas de aumento de datos

Con el objeto de mejorar el rendimiento de los modelos evaluados y reducir la probabilidad de ajustarse demasiado a los datos de entrenamiento, se implementaron durante el entrenamiento técnicas de aumento de datos. Se aplicaron las transformaciones para realizar aumento de datos tales como rotaciones aleatorias, ajustes en el brillo y cambios de contraste, lo que llevó a la generación de nuevas variantes de imágenes a partir de las imágenes originales, simulando situaciones más realistas y heterogéneas que las reales obtenidas durante el proceso de campo, lo que contribuyó a robustecer el propio modelo de cara a situaciones donde se presentaron variaciones iluminación y del entorno [42]–[46].

2.6 Métricas de desempeño

La evaluación del modelo desde el punto de vista de su rendimiento cuantitativo se complementó con la inclusión de pruebas estadísticas, calculando la precisión en términos generales y el error cuadrático medio (MSE, por sus siglas en inglés); esta métrica introduce un indicador de rendimiento que mide diferencia promedio de las predicciones del modelo con respecto a los valores reales. También, se llevó a cabo la construcción y análisis de una matriz de confusión que sirvió para poder ir calculando las tasas de aciertos y errores para cada clase de defecto vial.

De igual manera, se determinó cuál optimizador fue el más adecuado, realizando un análisis comparativo del desempeño entre SGDM y ADAM. Para saber cuál de los dos optimizadores estudiados era el más eficiente y apropiado para clasificar los defectos viales, adicional a la precisión lograda, se verificó la estabilidad y la capacidad de generalizar los resultados en diferentes subconjuntos de datos [29]–[34].

Para determinar el potencial del modelo para clasificar correctamente el tipo de defecto, se calculó la precisión por clase o sensibilidad, la cual identifica la proporción del número de verdaderos positivos (VP) bien clasificados frente a la cantidad del total de elementos reales de cada clase. Este tipo de medida fue de gran utilidad, ya que permitió encontrar los posibles sesgos en la clasificación de defectos concretos, y poder así ajustar los distintos hiperparámetros. También, se realizó igualmente un análisis de la AUC-ROC para cada categoría de defecto vial.

La metodología implementada en este trabajo permitió desarrollar un modelo de clasificación y análisis de defectos en carreteras con muy buena eficiencia y eficacia en la detección. Además, por medio del uso de técnicas de procesamiento de imágenes y de la

implementación de redes neuronales convolucionales, junto con la evaluación comparativa entre los optimizadores SGDM y ADAM, se obtuvo como resultado un procedimiento sistemático satisfactorio para el análisis automatizado de pavimentos.

3. RESULTADOS Y DISCUSIÓN

Los resultados alcanzados en este desarrollo comprueban el potencial que pueden llegar a tener las redes neuronales convolucionales para la clasificación automática de los defectos en vías urbanas, lo que establece un buen aporte en el campo de detección de imperfecciones en carreteras, como soporte a los sistemas de gestión de pavimentos y el mantenimiento de infraestructura vial urbana. En este sentido, este estudio agrega valor tecnológico con respecto a otras investigaciones que se han realizado, ya que se abordó el desempeño del comportamiento de los dos optimizadores evaluados (SGDM y ADAM) y los formatos de imagen utilizados (RGB y escala de grises), aportando conocimiento novedoso sobre el efecto que cada una de estas variables ha tenido sobre la precisión y robustez de los modelos.

Con respecto al rendimiento general y el consumo de recursos se pone de manifiesto que el modelo entrenado con el optimizador SGDM alcanzó una precisión del 74.67 % que utilizó imágenes en formato RGB y el modelo entrenado con ADAM proporcionó un 51.76 % de precisión, el cual utilizó imágenes en escala de grises. El modelo hecho con el optimizador ADAM alcanzó en imágenes RGB un 73.96 % y en imágenes en escala de grises un 52.51 % de precisión. A pesar del hecho de que la precisión entre estos optimizadores no dista mucho entre ellos, son muy diferentes para mejorar la clasificación automática de defectos viales. La comparación de las curvas de precisión de cada uno de los algoritmos de optimización ADAM y SGDM aplicados a imágenes RGB y en escala de grises se describen en la Figura 2.

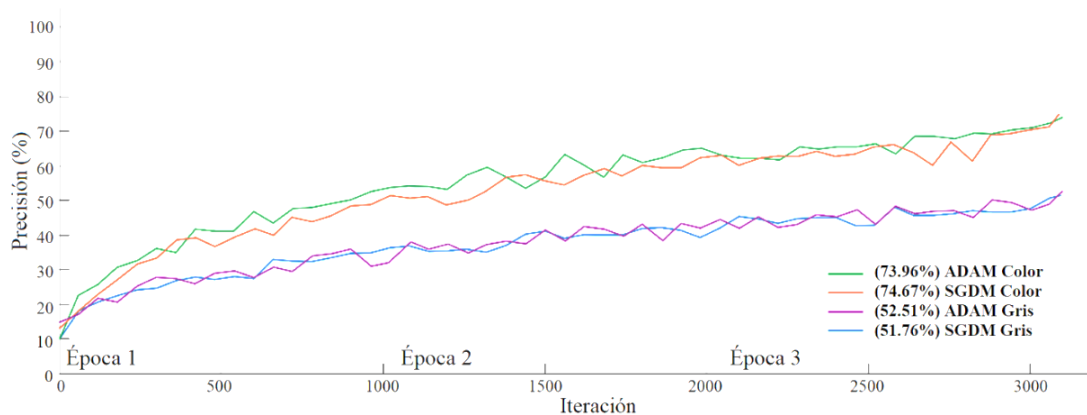


Figura 2. Curvas de precisión que comparan los optimizadores SGDM y ADAM aplicados.
Fuente: elaboración propia.

El optimizador SGDM se mostró más efectivo en el desempeño sobre imágenes RGB, lo cual es comprensible dado que este formato presenta una mayor cantidad de información para poder distinguir diferentes modos de fallas, en contrapartida, ADAM obtuvo mejores resultados usando imágenes en escala de grises, quizás porque este optimizador puede adaptarse mejor a la información que suele concentrarse en las texturas y bordes, las cuales son características principales de este tipo de imágenes.

En cuanto a consumo de recursos computacionales y tiempo de entrenamiento, ambos algoritmos de optimización reportaron un rendimiento similar, el consumo de CPU fue del 53 % al 58 %, la RAM del 64 %, y el de la GPU entre el 0 % y el 3 %. En cuanto a los tiempos de entrenamiento, variaron de los 80 a los 92 minutos, tomar en cuenta el optimizador y el tipo de imagen usado, el optimizador SGDM usando imágenes en escala de grises reportó un consumo

ligeramente inferior en promedio. Los resultados de consumo de recurso computacional de los optimizadores usados quedan resumidos en la Tabla 1.

Tabla 1. Consumo de recursos computacionales por variaciones de optimizadores.
Fuente: elaboración propia.

Optimizador	Imágenes	CPU	RAM	GPU	Tiempo
SGDM	Grises	53 %	63 %	1 %	81 min 43 s
SGDM	RGB	58 %	64 %	3 %	92 min 04 s
ADAM	Grises	55 %	64 %	0 %	80 min 13 s
ADAM	RGB	54 %	64 %	0 %	88 min 46 s

El análisis detallado de los resultados revela que el modelo entrenado con imágenes en escala de grises obtuvo una precisión en la validación del 79.91 %, coincidiendo con la precisión obtenida durante el entrenamiento, como se ilustra en la Figura 3.

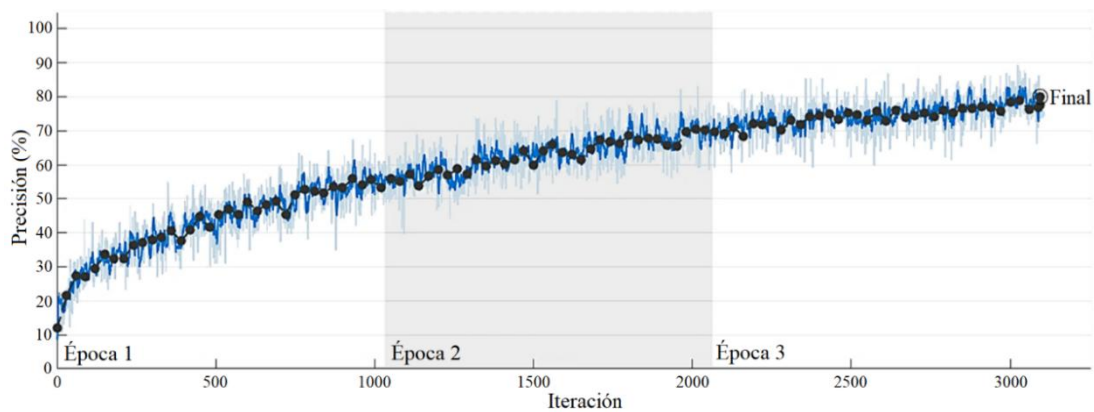


Figura 3. Precisión en entrenamiento y validación de CNN con imágenes en escala de grises.
Fuente: elaboración propia.

En contraste, el modelo entrenado con imágenes a color logró una precisión del 60.23 % en el entrenamiento y del 55 % en la validación, como se muestra en la Figura 4. Estos resultados sugieren que, a pesar de que el color aporta información adicional, en este caso particular, la información relacionada con texturas y bordes de las imágenes en escala de grises resulta más relevante para la clasificación precisa de los defectos viales.

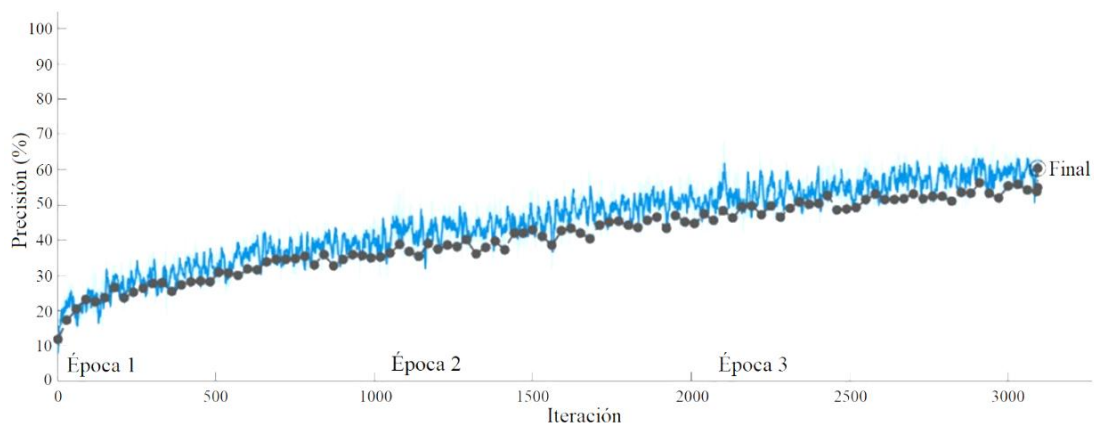


Figura 4. Precisión en entrenamiento y validación de CNN con imágenes RGB.
Fuente: elaboración propia.

Durante el proceso de entrenamiento del modelo en el mismo intervalo, se observó una tendencia interesante en la que la precisión iba aumentando de forma logarítmica mientras que la pérdida o el error del modelo empezaba a disminuir de forma logarítmica. Así fue como finalmente, tras todo el proceso de entrenamiento el modelo cargado con imágenes en escala de grises obtenía una pérdida de aproximadamente 0.7 que se puede observar en la Figura 5, y el que se entrenó con imágenes en RGB también logró una pérdida de 0.5 comprobable en la Figura 6. El valor de aproximación a cero proporciona un claro mensaje de que fue un buen ajuste y se trataba de un buen modelo con una buena capacidad de generalización.

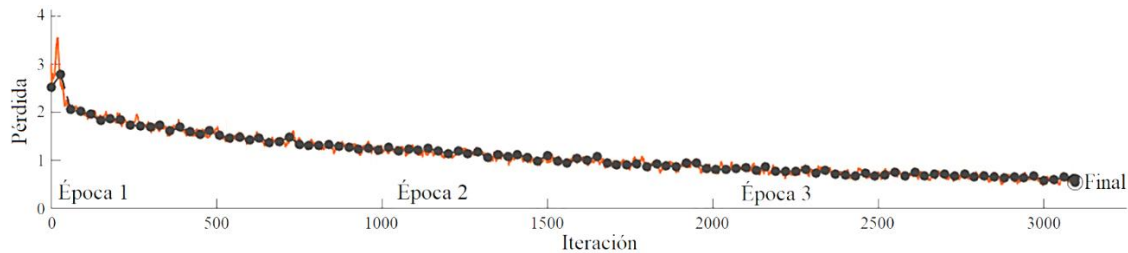


Figura 5. Pérdida en entrenamiento y validación de CNN con imágenes en escala de grises.
Fuente: elaboración propia.

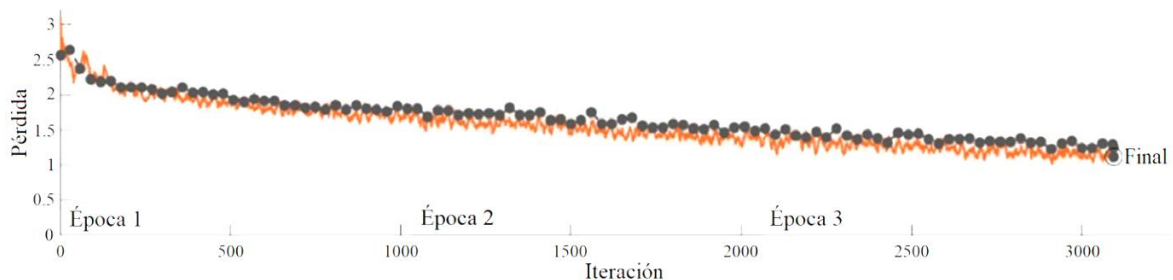


Figura 6. Pérdida en entrenamiento y validación de CNN con imágenes en RGB.
Fuente: elaboración propia.

La matriz de confusión del modelo entrenado con imágenes en escala de grises muestra que las categorías "bache", "descascaramiento", "fisura en bloque", "fisura por deslizamiento" y "fisura por reflexión" fueron correctamente clasificadas en más del 80 % de los casos con imágenes en escala de grises, como se representa la Figura 7.

Las clases de "fisuras longitudinales y transversales" y "parche" tienen una sensibilidad por debajo de la media, siendo del 59.6 % y 93.6 %, respectivamente. Estos resultados reflejan que el modelo tiene dificultades para discriminar las imperfecciones entre clasificaciones superiores como consecuencia de la existencia de superposiciones de patrones visuales.

Al estudiar los valores de la matriz de confusión que se encuentran fuera de la diagonal principal a la que se le asigna un 40.4 % de "fisuras longitudinales y transversales" como "hundimiento", lo que puede ser una consecuencia de la similitud visual que existe entre la fisura y los hundimientos.

Al calcular las curvas AUC-ROC por clase se obtuvo un promedio de 0.82 con las clases mejor identificadas, mientras que "fisuras longitudinales y transversales" descendió a 0.68 lo que evidencia la complejidad de identificar este tipo de fallas dado la superposición visual con otros defectos, como el "hundimiento".

Por otra parte, el modelo ajustado a imágenes RGB dio lugar a tasas de clasificación correcta elevadas en categorías como las de "bache", "descascaramiento", "fisura en bloque", "fisura por deslizamiento" y "fisura por reflexión", donde el valor superó el 65 %, tal como demuestra la Figura 8. Las categorías "fisuras longitudinales y transversales" y "parche" presentaron tasas de clasificación notablemente más bajas, un 44.9 % y un 47.0 %, respectivamente.

Clase Verdadera

Figura 7. Matriz de confusión de prueba de CNN con imágenes en escala de grises.
Fuente: elaboración propia.

Figura 8. Matriz de confusión de prueba CNN con imágenes RGB.
Fuente: elaboración propia.

Página 9 | 14

La comparación de estos resultados con estudios previos, como los que fueron presentados en [20]-[37] que mostraron precisiones por encima del 80 % en condiciones controladas, evidencia que los valores obtenidos en el presente estudio son competitivos entendiendo el enfoque orientado a condiciones urbanas reales, donde las variaciones en el carácter de las imágenes y de las características de captura influyen sobre la tipología de imperfección.

Además, el trabajo evaluó cómo el modelo se comporta ante cambios abruptos de contraste y de condiciones de iluminación, a partir del uso de técnicas de aumento de datos. El modelo muestra un mejor desempeño con imágenes en escala de grises frente a cambios en las condiciones climáticas y condiciones de iluminación, sugiriendo que el modelo ha sido diseñado para tener una buena capacidad de funcionamiento aumentando la estabilidad de un entorno urbano con iluminación no controlada. Esta capacidad es una característica destacada para aplicar en sistemas reales de inspección de vías, donde las condiciones de captura varían considerablemente.

Un importante punto débil es la confusión que puede inducir la exposición que provoca ambientes con patrones visuales parecidos; con ello, se puede concluir que hay que explorar el uso de arquitecturas más avanzadas como ResNet o EfficientNet o también la adición de variables contextuales para hacer una buena discriminación entre las clases.

Desde la perspectiva práctica, las conclusiones indican que es posible la implementación en sistemas de monitorización vial basados en visión artificial de manera que se puede automatizar la detección y clasificación de los defectos bajo un consumo razonable de recursos, de manera que sean favorables en el ámbito del montaje de plataformas de mantenimiento predictivo.

Finalmente, este trabajo ha mostrado que las CNN son prometedoras, pero aún tienen limitaciones respecto de la calidad de las imágenes y de los optimizadores que han sido utilizados: el SGDM ha dado mejores resultados en comparación con imágenes en RGB, mientras que el ADAM ha presentado un mejor desempeño con imágenes en escala de grises.

Este estudio pone de manifiesto la necesidad de adaptar métodos de clasificación a los datos. Los resultados obtenidos son similares a los que ya han sido documentados en otros trabajos, aunque se destaca el balance entre precisión y eficiencia computacional de los métodos propuestos. Los trabajos futuros deben centrarse en mejorar la clasificación de clases complejas, así como en la valoración del rendimiento del modelo en condiciones más complejas, como por ejemplo inclemencias meteorológicas o las peculiaridades de distintas tipologías de vías urbanas y rurales.

4. CONCLUSIONES

El presente estudio presenta la efectividad de las CNN para la automatización de la clasificación de defectos viales. Al analizar los resultados se puede observar que los mejores resultados se logran procesando las imágenes en RGB con los modelos entrenados con el optimizador SGDM alcanzando una mayor precisión (74.67 %) en contraste con los modelos entrenados con ADAM que alcanzan una menor precisión (52.51 %) pero utilizando para ello imágenes en escala de grises. Por otro lado, ambos optimizadores presentan un consumo de recursos computacionales similar, aunque el optimizador ADAM presenta tiempos de entrenamiento ligeramente menores. Igualmente, la arquitectura de CNN de 4 capas y con 3 épocas de entrenamiento muestra su efectividad ofreciendo una relación de complejidad y rendimiento equilibrada.

Dichos resultados confirman su viabilidad de entrar la técnica de aprendizaje automático dentro del ámbito de las aplicaciones para la gestión de la infraestructura vial o aplicaciones para móviles capaces de identificar y reportar defectos viales. Este trabajo también sienta las bases para investigaciones futuras en optimización de algoritmos de aprendizaje profundo con relación a la ingeniería civil y el mantenimiento de carreteras.

Los resultados obtenidos demuestran la utilidad de las CNN para el análisis estructural de baches utilizando imágenes del mundo real. Los modelos alcanzaron una precisión aceptable, sobre todo cuando se usaron imágenes de escala de grises, lo que indica que la información de textura contenida en ellas mejora notablemente la capacidad del modelo para distinguir y clasificar las diferentes categorías de las imperfecciones de las carreteras.

Así, la red convolucional entrenada con imágenes en escala de grises alcanzó una precisión del 79.91 %, tanto en el entrenamiento como en la validación, cifra superior a la obtenida con el modelo de imágenes RGB, que sólo obtuvo un 60.23 % en el entrenamiento y un 55 % en el proceso de validación. Este hecho hace cada vez más evidente que las texturas que se pueden observar en las imágenes en escalas de grises son más ricas en características que las que se encuentran en las imágenes de escala de colores RGB, y que estas son más relevantes en el proceso de clasificación.

Aun así, algunas categorías de las imperfecciones, como las "fisuras longitudinales y transversales" o los "parches", presentaron tasas de clasificación correcta más bajas, lo que sugiere que estas imperfecciones tienen características visuales similares a otras categorías, lo que provoca que el modelo no pueda hacer una correcta diferenciación.

El estudio realizado a través de las matrices de confusión denota que las confusiones más usuales se presentaron en los casos de "hundimiento" frente al "bache", ya que presentan similitudes entre ellos que a su vez reflejan sus patrones visuales.

Positivamente, los resultados obtenidos indican el potencial que tienen las redes neuronales convolucionales en el análisis estructural de los baches viales, después de haberse detectado los errores y lograr mejorar su exactitud y por lo tanto su capacidad de generalización.

Las investigaciones futuras podrían centrarse en ampliar el conjunto de datos y diversificarlo, explorar arquitecturas más complejas de redes neuronales o aplicar técnicas avanzadas de procesamiento de imagen y extracción de características con el fin de obtener mejoras en el rendimiento o en la calidad del sistema.

5. AGRADECIMIENTO Y FINANCIACIÓN

Los autores agradecen al Grupo de Investigación en Ingenierías – GIIS de la Facultad de Ingeniería de la Fundación Universitaria Compensar, por su soporte técnico, financiero y de personal de apoyo académico para el desarrollo de este estudio. De igual manera se hace extensivo agradecimiento al Grupo de Investigación en Ingeniería Electrónica – I2E de la Escuela de Ingeniería Electrónica de la Universidad Pedagógica y Tecnológica de Colombia y al Grupo de Investigación MACRYPT – FarmTechnology de la Facultad de Ciencias Básicas e Ingeniería de la Universidad de los Llanos, por su valioso trabajo, asesoría y aporte de cooperación investigativa en el desarrollo de este estudio.

6. REFERENCIAS

- [1] J. A. Lupano, *La infraestructura de transporte sostenible y su contribución a la igualdad en América Latina y el Caribe*, Santiago de Chile, Chile: Comisión Económica para América Latina y el Caribe (CEPAL), 2013. <https://hdl.handle.net/11362/35883>
- [2] J. D. Madroñero Urcuqui, and Y. C. Valencia López, "Metodología para la identificación automática del deterioro en pavimento flexible, por medio de fotografías aéreas tomadas desde vehículos no tripulados," Tesis de grado, Universidad del Valle, Santiago de Cali, Colombia, 2019. <https://bibliotecadigital.univalle.edu.co/handle/10893/15476>
- [3] J. F. Olarte Bustinza, and C. M. Soto Mallqui, "inspección de seguridad vial y propuesta de mejora en la intersección cuádruple de las avenidas el sol, tullumayo, pardo paseo de los héroes y la alameda pachacutec, ubicadas en el centro histórico de la ciudad del cusco," Tesis de grado, Pontificia Universidad Católica del Perú, Lima, Perú, 2023. <http://hdl.handle.net/20.500.12404/26106>

- [4] S. M. Martínez Niño, and G. A. Rojas Sánchez, "Inventario Vial y Alternativas de Intervención para el Tramo Vial que une los Municipios de Lourdes hasta Sardinata Desde el K10+00 al K17+00, Departamento de Norte de Santander," Tesis de grado, Universidad Francisco de Paula Santander, Cucutá, Colombia, 2023. <https://repositorio.ufps.edu.co/handle/ufps/6916>
- [5] X. Chen, C. Liu, L. Chen, X. Zhu, Y. Zhang, and C. Wang, "A Pavement Crack Detection and Evaluation Framework for a UAV Inspection System Based on Deep Learning," *Appl. Sci.*, vol. 14, no. 3, p. 1157, Feb. 2024. <https://doi.org/10.3390/app14031157>
- [6] D. Arya et al., "Deep learning-based road damage detection and classification for multiple countries," *Automat. Constr.*, vol. 132, p. 103935, Dec. 2021. <https://doi.org/10.1016/j.autcon.2021.103935>
- [7] M. Anis Benallal, and M. Si Tayeb, "An image-based convolutional neural network system for road defects detection," *IAES Int. J. Artif. Intell.*, vol. 12, no. 2, pp. 577–584, Jun. 2023. <http://doi.org/10.11591/ijai.v12.i2.pp577-584>
- [8] M. Mythili, D. Janani, D. S. Madhumitha, V. Mangala Madhumita, and R. Nandhini, "A Framework for Precise Road Damage Detection Using Deep Learning," in *2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA)*, Namakkal, India, 2024, pp. 1–5. <http://dx.doi.org/10.1109/aimla59606.2024.10531516>
- [9] X. Yang et al., "Research and applications of artificial neural network in pavement engineering: A state-of-the-art review," *J. Traffic Transp. Engin.*, vol. 8, no. 6, pp. 1000–1021, Dec. 2021. <https://doi.org/10.1016/j.jtte.2021.03.005>
- [10] A. Ashraf, A. Sophian, A. Akramin shafie, T. Surya Gunawan, and N. Nadia Ismail, "Machine learning-based pavement crack detection, classification, and characterization: a review," *Bul. Electr. Eng. Inform.*, vol. 12, no. 6, pp. 3601–3619, Dec. 2023. <https://doi.org/10.11591/eei.v12i6.5345>
- [11] J.-W. Baek, and K. Chung, "Pothole classification model using edge detection in road image," *Appl. Sci.*, vol. 10, no. 19, p. 6662, Oct. 2020. <https://doi.org/10.3390/app10196662>
- [12] N. Ma et al., "Computer vision for road imaging and pothole detection: a state-of-the-art review of systems and algorithms," *Transp. Saf. Environ.*, vol. 4, no. 4, pp. 1–16, Dec. 2022. <https://doi.org/10.1093/tse/tdac026>
- [13] L. Manoni, S. Orcioni, and M. Conti, "Recent Advancements in Deep Learning Techniques for Road Condition Monitoring: A Comprehensive Review," *IEEE Acc.*, vol. 12, pp. 154271–154293, Oct. 2024. <https://doi.org/10.1109/ACCESS.2024.3481649>
- [14] J. Li, T. Liu, X. Wang, and J. Yu, "Automated asphalt pavement damage rate detection based on optimized GA-CNN," *Automat. Constr.*, vol. 136, p. 104180, Apr. 2022. <https://doi.org/10.1016/j.autcon.2022.104180>
- [15] B. Bučko, E. Lieskovská, K. Zábovská, and M. Zábovský, "Computer Vision Based Pothole Detection under Challenging Condition," *Sensors*, vol. 22, no. 22, p. 8878, Nov. 2022. <https://doi.org/10.3390/s22228878>
- [16] Y. Sang, Q. Yu, Y. Fang, V. Vo, and R. Wix, "Smartphone-Based IRI Estimation for Pavement Roughness Monitoring A Data-Driven Approach," *IEEE Internet Things J.*, vol. 11, no. 11, pp. 19708–19720, Jun. 2024. <https://ieeexplore.ieee.org/document/10444016>
- [17] K. Zhao, S. Xu, J. Loney, A. Visentin, and Z. Li, "Road pavement health monitoring system using smartphone sensing with a two-stage machine learning model," *Automat. Constr.*, vol. 167, p. 105664, Nov. 2024. <https://doi.org/10.1016/j.autcon.2024.105664>
- [18] M. Staniek, "Road pavement condition diagnostics using smartphone-based data crowdsourcing in smart cities," *J. Traffic Transp. Eng.*, vol. 8, no. 4, pp. 554–567, Aug. 2021. <https://doi.org/10.1016/j.jtte.2020.09.004>
- [19] S. Matarneh, F. Elghaish, D. J. Edwards, F. Pour Rahimian, E. Abdellatef, and O. Ejohwomu, "Automatic Crack Classification on Asphalt Pavement Surfaces using Convolutional Neural Networks and Transfer Learning," *J. Informat. Technol. Constr.*, vol. 29, pp. 1239–1256, Dec. 2024. <https://doi.org/10.36680/j.itcon.2024.055>
- [20] G. Ochoa-Ruiz, A. A. Angulo-Murillo, A. Ochoa-Zezzatti, L. M. Aguilar-Lobo, J. A. Vega-Fernández, and S. Natraj, "An Asphalt Damage Dataset and Detection System Based on Retinanet for Road Conditions Assessment," *Appl. Sci.*, vol. 10, no. 11, p. 3974, Jun. 2020. <https://doi.org/10.3390/app10113974>
- [21] F. Jalili, S. Morsal Chavami, and H. Afsharnia, "An Artificial Neural Network approach to assess road roughness using smartphone-based crowdsourcing data," *Eng. Appl. Artif. Intell.*, vol. 138, no. Part A, p. 109308, Dec. 2024. <https://doi.org/10.1016/j.engappai.2024.109308>
- [22] A. M. Al-Sabaei, M. I. Souliman, and A. Jagadeesh, "Smartphone applications for pavement condition monitoring: A review," *Constr. Build. Mater.*, vol. 410, p. 134207, Jan. 2024.

- <https://doi.org/10.1016/j.conbuildmat.2023.134207>
- [23] L. Tello-Cifuentes, M. Aguirre-Sánchez, J. P. Díaz-Díaz, and F. Hernández, "Evaluación de daños en pavimento flexible usando fotogrametría terrestre y redes neuronales," *Tecnol.*, vol. 24, no. 50, p. e1686, Jan. 2021. <https://doi.org/10.22430/22565337.1686>
- [24] W. Song, G. Jia, H. Zhu, D. Jia, and L. Gao, "Automated Pavement Crack Damage Detection Using Deep Multiscale Convolutional Features," *J. Adv. Transport.*, vol. 2020, p. 6412562, Jan. 2020. <https://doi.org/10.1155/2020/6412562>
- [25] Y. Safyari, M. Mahdianpari, and H. Shiri, "A Review of Vision-Based Pothole Detection Methods Using Computer Vision and Machine Learning," *Sensors*, vol. 24, no. 17, p. 5652, Aug. 2024. <https://doi.org/10.3390/s24175652>
- [26] C.-C. Hsieh, H.-W. Jia, W.-H. Huang, and M.-H. Hsieh, "Deep Learning-Based Road Pavement Inspection by Integrating Visual Information and IMU," *Information*, vol. 15, no. 4, p. 239, Apr. 2024. <https://doi.org/10.3390/info15040239>
- [27] F. Liu, J. Liu, and L. Wang, "Asphalt pavement fatigue crack severity classification by infrared thermography and deep learning," *Automat. Constr.*, vol. 143, p. 104575, Nov. 2022. <https://doi.org/10.1016/j.autcon.2022.104575>
- [28] H. Garita-Durán, J. Philipp Stöcker, and M. Kaliske, "Deep learning-based system for automated damage detection and quantification in concrete pavement," *Results Eng.*, vol. 25, p. 104546, Mar. 2025. <https://doi.org/10.1016/j.rineng.2025.104546>
- [29] Y. Zhang, and L. Zhang, "Detection of Pavement Cracks by Deep Learning Models of Transformer and UNet," *IEEE Trans. Intell. Transp. Syst.*, vol. 25, no. 11, pp. 15791–15808, Nov. 2024. <https://doi.org/10.1109/TITS.2024.3420763>
- [30] W. Guo, L. Zhong, D. Zhang, and Q. Li, "Pavement Crack Detection using Fractal Dimension and Semi-Supervised Learning," *Fractal Fract.*, vol. 8, no. 8, p. 468, Aug. 2024. <https://doi.org/10.3390/fractalfract8080468>
- [31] J. Nnamdi Opara, A. B. Bo Thein, S. Izumi, H. Yasuhara, and P.-J. Chun, "Defect Detection on Asphalt Pavement by Deep Learning," *Int. J. GEOMATE*, vol. 21, no. 83, pp. 87–94, Jul. 2021. <https://doi.org/10.21660/2021.83.6153>
- [32] W. Cao, Q. Liu, and Z. He, "Review of Pavement Defect Detection Methods," *IEEE Acc.*, vol. 8, no. 1, pp. 14531–14544, Jan. 2020. <https://doi.org/10.1109/aACCESS.2020.2966881>
- [33] R. Tao, R. Peng, Y. Jin, F. Gong, and B. Li, "Automatic Detection of Asphalt Pavement Crack Width Based on Machine Vision," *IEEE Trans. Intell. Transport. Syst.*, vol. 26, no. 1, pp. 484–496, Jan. 2025. <https://doi.org/10.1109/TITS.2024.3492731>
- [34] S. Matarneh, F. Elghaish, F. Pour Rahimian, E. Abdellatef, and S. Abrishami, "Evaluation and optimisation of pre-trained CNN models for asphalt pavement crack detection and classification," *Automat. Constr.*, vol. 160, p. 105297, Apr. 2024. <https://doi.org/10.1016/j.autcon.2024.105297>
- [35] A. Rasheed Rababaah, "A Deep Learning based Process Model for Crack Detection in Pavement Structures," in *9th International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi, India, 2022, pp. 1–6. <https://doi.org/10.23919/INDIACom54597.2022.9763286>
- [36] J. Lan, H. Wang, Z. Zhu, and Q. Zhang, "Computer Vision Based Pothole Road Detection and Recognition," in *2024 5th Int. Conf. Comp. Vision, Image Deep Learn., (CVIDL)*, Zhuhai, China, 2024, pp. 505–509. <https://doi.org/10.1109/CVIDL62147.2024.10603934>
- [37] D. Shu et al., "Research on Methods of Pavement Distress Detection using Convolutional Neural Network based on Highway Rapid Inspection Images," in *2024 9th Int. Conf. Signal Image Process. (ICSIP)*, Nanjing, China, 2024, pp. 623–627. <https://doi.org/10.1109/ICSIP61881.2024.10671436>
- [38] E. Ranyal, A. Sadhu, and K. Jain, "Road Condition Monitoring using Smart Sensing and Artificial Intelligence: A Review," *Sensors*, vol. 22, no. 8, p. 3044, Apr. 2022. <https://doi.org/10.3390/s22083044>
- [39] M. Zeeshan, S. M. Adnan, W. Ahmad, and F. Zeeshan Khan, "Structural Crack Detection and Classification using Deep Convolutional Neural Network," *Pakistan J. Eng. Technol.*, vol. 4, no. 4, pp. 50–56, Dec. 2021. <https://doi.org/10.51846/vol4iss4pp50-56>
- [40] H. Jing, X. Zhang, and Z. Zhang, "Research on Automatic Identification of Asphalt Payment Defects Based on Deep Learning," in *2024 7th Int. Conf. Adv. Algor. Contr. Eng. (ICAACE)*, Shanghai, China, 2024, pp. 1407–1412. <https://doi.org/10.1109/ICAACE61206.2024.10548739>
- [41] A. As Sami, S. Sakib, K. Deb, and I. H. Sarker, "Improved YOLOv5-Based Real-Time Road Pavement Damage Detection in Road Infrastructure Management," *Algorithms*, vol. 16, no. 9, p. 452, Sep. 2023. <https://doi.org/10.3390/a16090452>
- [42] P. Li et al., "CNN-based pavement defects detection using grey and depth images," *Automat. Constr.*, vol. 158, p. 105192, Feb. 2024. <https://doi.org/10.1016/j.autcon.2023.105192>
- [43] T. Zhang, D. Wang, and Y. Lu, "ECSNet: An Accelerated Real-Time Image Segmentation CNN

- Architecture for Pavement Crack Detection," *IEEE Trans. Intell. Transport. Syst.*, vol. 24, no. 12, pp. 15105–15112, Dec. 2023. <https://doi.org/10.1109/TITS.2023.3300312>
- [44] D. Li, Z. Duan, X. Hu, D. Zhang, and Y. Zhang, "Automated classification and detection of multiple pavement distress images based on deep learning," *J. Traffic Transport. Eng.*, vol. 10, no. 2, pp. 276–290, Apr. 2023. <https://doi.org/10.1016/j.jtte.2021.04.008>
- [45] H. Chu et al., "Deep Learning Method to Detect the Road Cracks and Potholes for Smart Cities," *Comput., Mater. Continua*, vol. 75, no. 1, pp. 1863–1881, Feb. 2023. <https://doi.org/10.32604/cmc.2023.035287>
- [46] R. Kothai, N. Prabakaran, Y. V. Srinivasa, L. Reddy Cenkeramaddi, and V. Kakani, "Pavement Distress Detection, Classification, and Analysis Using Machine Learning Algorithms: A Survey," *IEEE Acc.*, vol. 12, pp. 126943–126960, Sep. 2024. <https://doi.org/10.1109/ACCESS.2024.3455093>

CONFLICTO DE INTERÉS

Los autores de este trabajo declaran de manera explícita que no existen conflictos de interés, sean estos financieros, profesionales o personales que hayan afectado de manera inapropiada los resultados obtenidos o bien las interpretaciones propuestas en este trabajo de investigación. Todos los autores han participado de manera equitativa en la investigación que ha llevado a la obtención de este artículo, y declaran que los datos y conclusiones aquí obtenidas son fruto de un análisis objetivo y riguroso.

CONTRIBUCIÓN DE AUTORÍA

Angie-P. Rique-Sabogal: supervisión del proyecto, conceptualización, diseño de la investigación, análisis de datos, redacción del borrador original.

Miguel-A. Guatame-Medina y Cristhian-F. Moreno-Manrique: recopilación de datos, metodología, validación de resultados, análisis estadístico, interpretación de resultados.

Fabián-R. Jiménez-López y Andrés-F. Jiménez-López: revisión crítica, análisis de resultados, redacción y revisión final del manuscrito.

Todos los autores han realizado una contribución sustancial a la investigación y al manuscrito, y han aprobado la versión final para su publicación. Se confirma que no existe autoría por cortesía ni autoría fantasma en este trabajo.

DECLARACIÓN DE IA

Durante la realización de esta investigación, los autores utilizaron la herramienta Grammarly, con el propósito de realizar corrección de errores de redacción, gramática y puntuación. Luego de emplear esta herramienta, los autores revisaron y editaron cuidadosamente el contenido según sea necesario y asumen total responsabilidad por el contenido de la publicación.