



# Predicción de incendios forestales mediante índices espectrales y *Deep Learning*: un estudio preliminar

**Kevin Manuel Galván Lara**

Universidad de Sonora, Departamento de Investigaciones Científicas y Tecnológicas,  
Posgrado en Biociencias, Hermosillo, Sonora, México.

ORCID: 0009-0001-6124-4549

**Yosune Miquelajauregui Graf**

Universidad Nacional Autónoma de México, Instituto de Ecología,  
Laboratorio Nacional de Ciencias de la Sostenibilidad, Ciudad de México, México

ORCID: 0000-0001-7084-7782

Recepción: 12 de noviembre de 2024.

Aceptación: 26 de agosto 2025.

Septiembre 2025 • número de revista 13 • DOI: 10.22201/dgtic.26832968e.2025.13.54

## Predicción de incendios forestales mediante índices espectrales y *Deep Learning*: un estudio preliminar

---

### Resumen

La creciente incidencia de incendios forestales representa una amenaza ecológica y económica significativa. La detección temprana de zonas de alto riesgo es fundamental para una gestión forestal proactiva. Este artículo presenta los resultados preliminares de un sistema para la identificación de áreas con riesgo de incendio utilizando imágenes satelitales de acceso libre (Landsat-8) y un modelo de Red Neuronal Convolutiva (CNN). Se definieron tres clases de riesgo a partir del análisis de índices espectrales (NDVI y NBR). El modelo fue entrenado y validado utilizando los recursos del Espacio de Innovación UNAM-HUAWEI, específicamente servidores Huawei Atlas 800. El modelo CNN alcanzó una exactitud global preliminar del 22% en el conjunto de validación para el Bosque de la Primavera (Jalisco, México). Si bien la exactitud es baja, el estudio identifica desafíos clave como el sobreajuste del modelo y la necesidad de incorporar variables adicionales, sentando las bases para futuras investigaciones.

**Palabras Clave:** *Deep Learning*, incendios forestales, índices espectrales, Landsat-8, redes neuronales convolucionales (CNN).

### *Prediction of Wildfires Using Spectral Indices and Deep Learning: A Preliminary Study*

---

#### Abstract

*The increasing incidence of wildfires poses a significant ecological and economic threat. Early detection of high-risk areas is essential for proactive forest management. This article presents preliminary results of a system designed to identify fire risk areas using freely available satellite*

*imagery (Landsat-8) and a convolutional neural network (CNN) model. Three risk classes were defined based on the analysis of spectral indices (NDVI and NBR). The model was trained and validated using resources from the Espacio de Innovación UNAM-HUAWEI, specifically Huawei Atlas 800 servers. The CNN model achieved a preliminary overall accuracy of 22% on the validation set for the Bosque de la Primavera (Jalisco, México). Although the accuracy is low, the study identifies key challenges such as model overfitting and the need to incorporate additional variables, laying the groundwork for future research.*

**Keywords:** *Deep learning, wildfires, spectral indices, Landsat-8, convolutional neural networks (CNN).*

## 1. Introducción

Los incendios forestales son uno de los desastres naturales con mayor impacto a nivel global, provocando pérdidas ecológicas, económicas y sociales [1][2]. México, por su diversidad de ecosistemas, es particularmente vulnerable [3]. La gestión tradicional de incendios se ha centrado en la supresión, un enfoque reactivo y costoso [4]. Sin embargo, el paradigma está cambiando hacia la prevención y la gestión proactiva, donde la tecnología juega un papel crucial [4][3].

La teledetección, mediante el uso de imágenes satelitales, ofrece una capacidad sin precedentes para monitorear grandes extensiones de territorio de manera continua y eficiente [5][6][4]. Una de las herramientas más poderosas en este campo son los índices espectrales [7]. Estos son combinaciones matemáticas de diferentes bandas de reflectancia de una imagen satelital, diseñadas para realzar características específicas de la superficie terrestre que no son visibles a simple vista [8]. Por ejemplo, el Índice de Vegetación de Diferencia Normalizada (NDVI) es sensible a la salud y densidad de la vegetación [9], mientras que el Índice Normalizado de Quemadura (NBR) es eficaz para detectar estrés hídrico en la vegetación y áreas previamente quemadas [10][5][6]. Ambos son indicadores indirectos clave del riesgo de incendio [11], y su aplicación en modelos predictivos puede mejorar significativamente la identificación de zonas vulnerables. La integración de datos espectrales y técnicas de aprendizaje profundo, como se ha explorado en estudios anteriores, es esencial

para avanzar en la mitigación de incendios forestales en áreas críticas como el Bosque de la Primavera [12][13].

En los últimos años, la combinación de la teledetección con técnicas de inteligencia artificial y, en particular, el Deep Learning ha abierto nuevas fronteras en el análisis geoespacial [7], [8]. Las redes neuronales convolucionales (CNN) han demostrado una gran eficacia en la clasificación y segmentación de imágenes, convirtiéndose en un estándar en el análisis de imágenes de teledetección [9]. El uso de imágenes satelitales, como las de Landsat, permite una evaluación continua y precisa de los cambios en la cobertura del suelo, fundamental para la gestión de incendios forestales [14].

A pesar del auge de estas tecnologías, su aplicación en el contexto mexicano para la predicción de incendios, utilizando datos de libre acceso y plataformas de nube, aún es un área con un gran potencial de desarrollo. Este artículo no pretende ser una revisión bibliográfica exhaustiva, sino que reporta una experiencia concreta de integración tecnológica. El objetivo de este trabajo es presentar el desarrollo y los resultados preliminares de un modelo de Deep Learning, basado en una CNN, para la clasificación de zonas de riesgo de incendio en el Bosque de la Primavera, Jalisco, utilizando como datos de entrada los índices espectrales NDVI y NBR derivados de imágenes Landsat-8.

## 2. Desarrollo

### 2.1 Metodología del estudio

El flujo de trabajo se dividió en cinco etapas principales: definición del área de estudio, adquisición y preprocesamiento de datos, definición de clases de riesgo, diseño y entrenamiento del modelo, y evaluación.

#### 2.1.1 Área de estudio

Se seleccionó como área de estudio el Área de Protección de Flora y Fauna La Primavera, comúnmente conocida como el Bosque de la Primavera, ubicada en el estado de Jalisco,

México. Esta zona, con una superficie aproximada de 305 km<sup>2</sup>, fue elegida por su alto valor ecológico, su proximidad a la Zona Metropolitana de Guadalajara y su historial recurrente de incendios forestales, lo que representa un caso de estudio relevante y de alto impacto para la clasificación de riesgo.

### 2.1.2 Adquisición y preprocesamiento de datos

Se utilizaron imágenes del satélite Landsat 8, de la NASA, por su resolución espacial de 10 m, resolución temporal de 15 días y su política de datos abiertos. Se descargaron imágenes corregidas atmosféricamente correspondientes a la temporada de estiaje. El preprocesamiento se realizó con bibliotecas de Python como Rasterio y GDAL. Se calcularon dos índices para cada imagen (Figura 1):

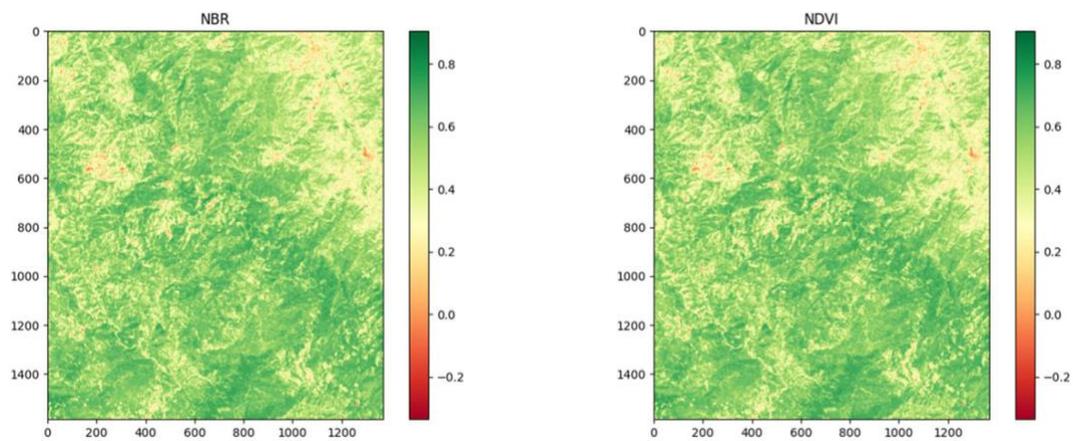
NDVI: Para evaluar la salud de la vegetación:

$$\text{NDVI} = \frac{\text{NIR} - \text{Red}}{\text{NIR} + \text{Red}}$$

NBR: tasa de quema normalizada:

$$\text{NBR} = \frac{\text{NIR} - \text{SWIR}}{\text{NIR} + \text{SWIR}}$$

Donde NIR es la banda del Infrarrojo Cercano, Red es la Roja y SWIR es la del Infrarrojo de Onda Corta.



**Fig. 1.** Ejemplos visuales de los índices NBR (izquierda) y NDVI (derecha) calculados para el área de estudio.

### 2.1.3 Definición de clases de riesgo

Para entrenar el modelo supervisado, se generó un mapa de verdad (*ground truth*) con tres clases de riesgo:

Alto Riesgo: Zonas con vegetación densa y seca (NDVI alto, NBR bajo).

Riesgo Moderado: Áreas de vegetación mixta o de matorral.

Bajo Riesgo / Sin Riesgo: Cuerpos de agua, zonas urbanas, roca expuesta. La etiquetación se realizó de forma semi-automática utilizando el software QGIS.

### 2.1.4 Arquitectura del modelo de *Deep Learning*

Se implementó una arquitectura de CNN utilizando Keras con TensorFlow. La arquitectura secuencial (Figura 2) consistió en: 1) Un bloque Conv2D (64 filtros, 3x3, ReLU) y MaxPooling2D (2x2); 2) Un segundo bloque Conv2D (128 filtros, 3x3, ReLU) y MaxPooling2D (2x2); 3) Una capa Flatten; 4) Una capa Dense (2500 neuronas, Sigmoid); 5) Una capa Reshape para transformar la salida a las dimensiones del parche (50x50). Se compiló con el optimizador 'Adam' y la función de pérdida 'binary\_crossentropy'. Se probaron dos configuraciones: una con un solo índice como entrada (un canal) y otra combinando NDVI y NBR (dos canales).

Arquitectura de la Red Neuronal Convolutiva (CNN)

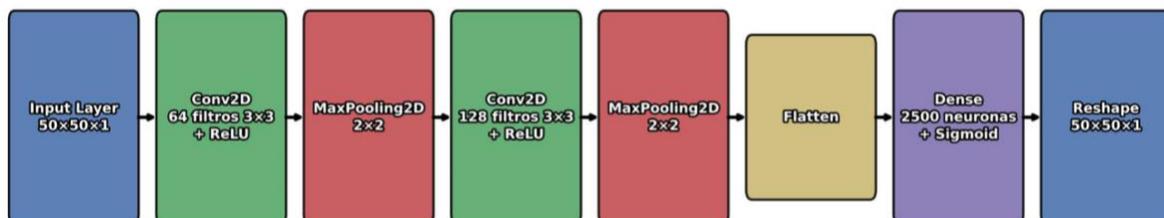


Fig. 2. Diagrama de la arquitectura CNN implementada (red neuronal)

### 2.1.5 Entrenamiento y Entorno Computacional (Espacio de Innovación UNAM-HUAWEI)

El *dataset* se construyó dividiendo las imágenes en parches de 50x50 píxeles, con una división de 70% para entrenamiento, 15% para validación y 15% para prueba.

El entrenamiento del modelo, una tarea computacionalmente intensiva, se llevó a cabo gracias a los recursos proporcionados por la convocatoria 2023 del Espacio de Innovación UNAM-HUAWEI. Específicamente, se utilizó el poder de cómputo de los servidores Huawei Atlas 800 (Modelo 9000). El acceso a esta infraestructura de alto rendimiento, que incluye no solo el hardware sino también recursos lógicos y de capacitación en plataformas como MindSpore, fue fundamental para acelerar los ciclos de experimentación y procesar los datos geoespaciales de manera eficiente.

### 2.1.6 Disponibilidad de Código

Con el objetivo de fomentar la transparencia, la colaboración y la reproducibilidad de la investigación, el *notebook* de Jupyter, con el código utilizado para el procesamiento de datos y el entrenamiento del modelo, se encuentra disponible públicamente en el siguiente repositorio de [GitHub](#).

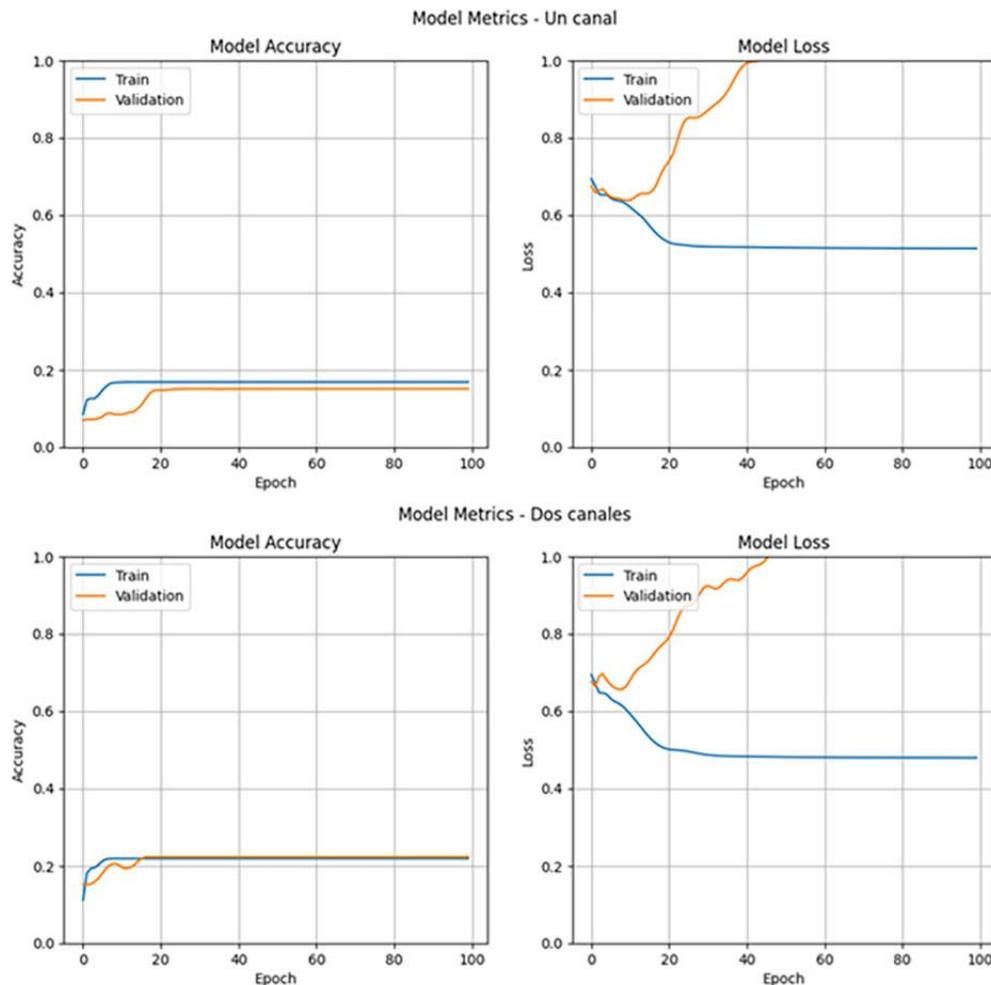
## 2.2 Resultados del Modelo

La evaluación del modelo se realizó tanto cuantitativa como cualitativamente.

### 2.2.1 Métricas de Desempeño

Las curvas de entrenamiento y validación se muestran en la Figura 3. El modelo que utilizó dos canales (NDVI y NBR) obtuvo un desempeño ligeramente superior, alcanzando una exactitud máxima en el conjunto de validación de aproximadamente 22%. Como se observa

en las gráficas de pérdida, ambos modelos muestran una clara tendencia al sobreajuste (*overfitting*).

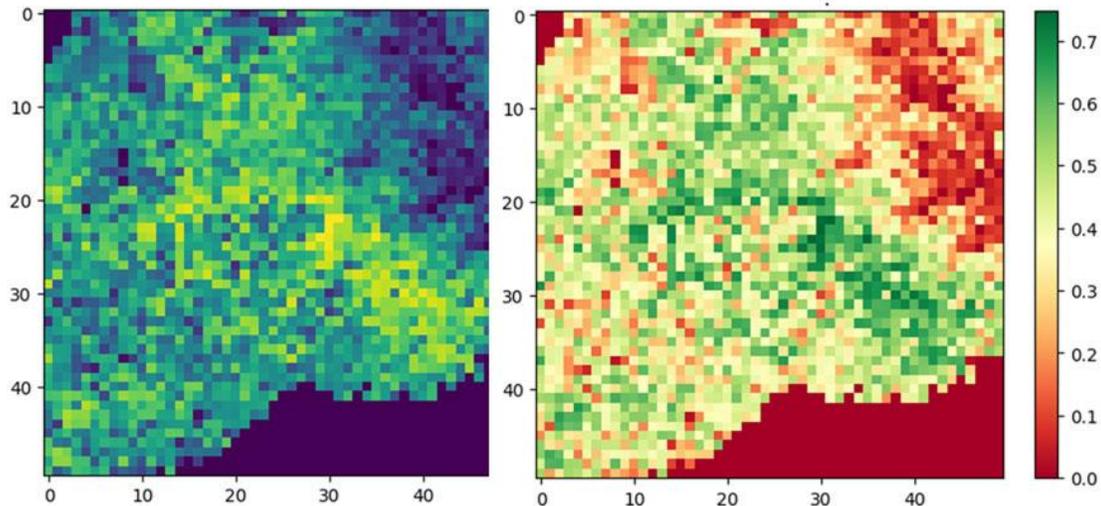


**Fig. 3.** Curvas de entrenamiento y validación para el modelo de un canal (arriba) y dos canales (abajo) (métricas).

### 2.2.2 Predicción Visual

La Figura 4 muestra un ejemplo de la salida del modelo para un parche de imagen no visto durante el entrenamiento. La salida visual muestra la capacidad del modelo para generar un

mapa de predicción, aunque su baja exactitud cuantitativa sugiere que la correspondencia con las zonas de riesgo reales es limitada.



**Fig. 4.** Resultado de predicción del modelo. A la izquierda, el parche de entrada (datos reales). A la derecha, la predicción de riesgo generada por el modelo (*output*).

### 2.3 Discusión

Los resultados de este estudio preliminar, con una exactitud de validación del 22%, deben interpretarse en el contexto de un proyecto de investigación más amplio. El objetivo de esta fase inicial era determinar la viabilidad de un modelo simple basado exclusivamente en índices espectrales. La baja exactitud y el severo sobreajuste observado son hallazgos clave que validan la hipótesis central del proyecto doctoral: la predicción de incendios es una tarea de alta complejidad que no puede resolverse únicamente con variables de vegetación [8][15], lo que sugiere la necesidad de integrar otros factores ambientales y climáticos para mejorar la precisión del modelo [16][17].

El sobreajuste indica que el modelo, si bien es capaz de memorizar los datos de entrenamiento, no logra generalizar. Esto confirma que los datos de entrada (NDVI y NBR) son insuficientes para capturar la compleja interacción de factores que determinan el riesgo

real. Este resultado, lejos de ser un fracaso, es un paso metodológico necesario que justifica científicamente la necesidad de avanzar hacia modelos más integrales. Demuestra que, para mejorar el desempeño, es imperativo incorporar las variables adicionales que el proyecto general contempla [8], tales como datos topográficos (pendiente, orientación), climáticos (temperatura, humedad) y socioeconómicos (proximidad a infraestructura humana, actividad agrícola), tal como lo señala la literatura [4][18].

### 3. Conclusión

Este trabajo ha presentado el desarrollo y la evaluación de un modelo de CNN, cumpliendo con uno de los objetivos específicos de una investigación doctoral en curso: implementar y determinar la capacidad de un algoritmo de aprendizaje automatizado. La metodología, desde la obtención de datos hasta el entrenamiento en una plataforma de cómputo en la nube, ha sido detallada para garantizar su reproducibilidad.

La principal contribución de este estudio no radica en la alta precisión del modelo, sino en la validación de una hipótesis fundamental: el uso exclusivo de índices espectrales es insuficiente para una predicción robusta de incendios forestales [19]. Los resultados, con una exactitud del 22% y un claro sobreajuste, son un hallazgo crucial que orienta y justifica los siguientes pasos de la investigación. Este hallazgo es fundamental para el objetivo general del proyecto, que busca construir una base de datos histórica y multivariable para México y, a partir de ella, desarrollar una herramienta de apoyo para la toma de decisiones.

La experiencia reportada subraya el valor de la colaboración entre la academia y la industria tecnológica para explorar soluciones a problemas complejos, y sienta una base sólida y honesta sobre la cual construir un sistema de predicción más completo y eficaz.

### Agradecimientos

Se agradece al Espacio de Innovación UNAM-HUAWEI por proporcionar los recursos de cómputo de alto rendimiento (servidores Atlas 800), los cuales fueron fundamentales para el desarrollo de este proyecto en el marco de su convocatoria 2023.

## Referencias

- [1] V. M. Cvetković, J. Gačić, and V. Jakovljević, "Geospatial and temporal distribution of forest fires as natural disasters," *Vojno Delo*, vol. 68, no. 2, pp. 108–127, Jan. 2016, doi: 10.5937/vojdelo1602108c.
- [2] M. Gill, S. L. Stephens, and G. J. Cary, "The worldwide 'wildfire' problem," *Ecological Applications*, vol. 23, no. 2, pp. 438–454, Sep. 2012, doi: 10.1890/10-2213.1.
- [3] M. Ponce and A. Virgen, "Vulnerability of biodiversity to land use change and climate change in Mexico," 2016. [Online]. Available: <https://era.ed.ac.uk/handle/1842/21701>
- [4] R. Kanwal, W. Rafaqat, M. Iqbal, and S. Weiguo, "Data-Driven Approaches for wildfire mapping and prediction Assessment using a Convolutional Neural Network (CNN)," *Remote Sensing*, vol. 15, no. 21, p. 5099, Oct. 2023, doi: 10.3390/rs15215099
- [5] J. R. Dymond, A. Bégue, and D. Loseen, "Monitoring land at regional and national scales and the role of remote sensing," *International Journal of Applied Earth Observation and Geoinformation*, vol. 3, no. 2, pp. 162–175, Jan. 2001, doi: 10.1016/s0303-2434(01)85008-x.
- [6] Y. Wang, Z. Lu, Y. Sheng, and Y. Zhou, "Remote sensing applications in monitoring of protected areas," *Remote Sensing*, vol. 12, no. 9, p. 1370, Apr. 2020, doi: 10.3390/rs12091370.
- [7] D. Montero, C. Aybar, M. D. Mahecha, F. Martinuzzi, M. Söchting, and S. Wieneke, "A standardized catalogue of spectral indices to advance the use of remote sensing in Earth system research," *Scientific Data*, vol. 10, no. 1, Apr. 2023, doi: 10.1038/s41597-023-02096-0.
- [8] M. İ. Gürsoy, O. Orhan, and S. Tekin, "Creation of wildfire susceptibility maps in the Mediterranean Region (Turkey) using convolutional neural networks and multilayer perceptron techniques," *Forest Ecology and Management*, vol. 538, p. 121006, May 2023, doi: 10.1016/j.foreco.2023.121006.

- [9] G. Zhang, M. Wang, and K. Liu, "Deep neural networks for global wildfire susceptibility modelling," *Ecological Indicators*, vol. 127, p. 107735, Apr. 2021, doi: 10.1016/j.ecolind.2021.107735.
- [10] "NBR Index-Based Fire detection using Sentinel-2 images and GIS: A case study in Mosul Park, Iraq," *International Journal of Geoinformatics*, pp. 67–74, May 2023, doi: 10.52939/ijg.v19i3.2607.
- [11] Santopaolo, S. S. Saif, A. Pietrabissa, and A. Giuseppi, "Forest Fire Risk Prediction from Satellite Data with Convolutional Neural Networks," 2022 30th Mediterranean Conference on Control and Automation (MED), pp. 360–367, Jun. 2021, doi: 10.1109/med51440.2021.9480226.
- [12] J. L. Ibarra-Montoya and F. M. Huerta-Martínez, "Modelado espacial de incendios: una herramienta predictiva para el Bosque La Primavera, Jalisco México," *Ambiente E Agua - an Interdisciplinary Journal of Applied Science*, vol. 11, no. 1, Jan. 2016, doi: 10.4136/ambi-agua.1536.
- [13] N. P. A. M and N. P. Ankalkoti, "Forest Wildfire Detection from Satellite Images using Deep Learning," *International Journal of Advanced Research in Science Communication and Technology*, pp. 578–583, Jul. 2023, doi: 10.48175/ijarsct-12086.
- [14] S. A. Azzouzi, "Técnicas de detección de cambios mediante teledetección para el desarrollo sostenible y la desertificación," 2019. doi: 10.4995/thesis/10251/117994.
- [15] M. Naderpour, H. M. Rizeei, and F. Ramezani, "Forest fire Risk prediction: A spatial deep neural Network-Based framework," *Remote Sensing*, vol. 13, no. 13, p. 2513, Jun. 2021, doi: 10.3390/rs13132513.
- [16] L. Knopp, "Development of a burned area processor based on Sentinel-2 data using deep learning," *PFG - Journal of Photogrammetry Remote Sensing and Geoinformation Science*, vol. 89, no. 4, pp. 357–358, Aug. 2021, doi: 10.1007/s41064-021-00177-6.
- [17] Y. Michael, D. Helman, O. Glickman, D. Gabay, S. Brenner, and I. M. Lensky, "Forecasting fire risk with machine learning and dynamic information derived from

satellite vegetation index time-series," *The Science of the Total Environment*, vol. 764, p. 142844, Oct. 2020, doi: 10.1016/j.scitotenv.2020.142844.

- [18] M. Gonçalves, T. Brandão, and J. C. Ferreira, "Wildfire Detection with Deep Learning—A Case study for the CICLOPE Project," *IEEE Access*, vol. 12, pp. 82095–82110, Jan. 2024, doi: 10.1109/access.2024.3406215.
- [19] Karimi et al., "Evaluating Models and Effective Factors Obtained from Remote Sensing (RS) and Geographic Information System (GIS) in the Prediction of Forest Fire Risk, Structured Review," *Journal of Geography and Cartography*, vol. 1, no. 4, Sep. 2018, doi: 10.24294/jgc.v1i4.618.