



## Inteligencia Artificial para la solución de problemas sociales 2

Alianza UNAM-Huawei



Septiembre 2025 • DOI: 10.22201/dgtic.26832968e.2025.13 • ISSN: 2683-2968.

TIES es una revista de acceso abierto bajo la licencia Creative Commons Atribución-No Comercial 4.0 (CC BY-NC 4.0).

© 2025 TIES, Revista de Tecnología e Innovación en Educación Superior es editada por la Universidad Nacional Autónoma de México (UNAM) a través de la Dirección General de Cómputo y de Tecnologías de Información y Comunicación (DGTIC). Circuito exterior s/n, Ciudad Universitaria, Alcaldía Coyoacán, C.P. 04510, Ciudad de México, México. Número de reserva de Derechos otorgado por INDAUTOR: 04-2019-011816190900-203.

El contenido de los artículos es responsabilidad de los autores y no refleja el punto de vista del Comité editorial, del Editor o de la UNAM.



## Índice

---

Editorial	I
Inteligencia artificial y su aplicación en la seguridad vial	1
Exploración de la Comunicación Visual del Cambio Climático mediante la Clasificación Automática de Imágenes	17
Optimización de la vigilancia volcánica: visión computacional para la detección de emisiones de ceniza del volcán Popocatepetl	31
Predicción de incendios forestales mediante índices espectrales y <i>Deep Learning</i> : un estudio preliminar	47
Experiencia en un proyecto de IA educativo para la detección de dificultades matemáticas, basado en emociones y desempeño estudiantil	60
<i>Heart beating</i> : una aplicación para el diagnóstico prenatal inteligente	77
Extracción de interacciones regulatorias transcripcionales de bacterias a partir de literatura biomédica utilizando inteligencia artificial	91

## Editorial

---

### Inteligencia Artificial para la solución de problemas sociales 2 | Alianza UNAM-Huawei

En este número de la revista TIES, presentamos siete artículos de investigación en inteligencia artificial, abordando los campos de visión computacional, análisis de señales y extracción de conocimiento en textos digitales.

Los siguientes artículos muestran cómo la implementación de soluciones de visión computacional, mediante técnicas de inteligencia artificial, puede aplicarse efectivamente en una amplia variedad de actividades humanas: “Inteligencia artificial y su aplicación en la seguridad vial”; “Exploración de la comunicación visual del cambio climático mediante la Clasificación Automática de Imágenes”; “Optimización de la vigilancia volcánica: visión computacional para la detección de emisiones de ceniza del volcán Popocatepetl”; “Predicción de incendios forestales mediante índices espectrales: *Deep learning* como herramienta innovadora para la gestión forestal”; “Experiencia en un proyecto de IA educativo para la detección de dificultades matemáticas, basado en emociones y desempeño estudiantil”.

Por otra parte, en el artículo titulado “*Heart beating*: una aplicación para el diagnóstico prenatal inteligente”, los autores reportan una aplicación de técnicas de inteligencia artificial en el análisis de señales de electrocardiografía materno-fetal. Finalmente, en “Extracción de interacciones regulatorias transcripcionales de bacterias a partir de literatura biomédica utilizando inteligencia artificial”, se demuestra cómo es posible extraer una gran cantidad de información significativa utilizando dichas técnicas para el análisis de textos digitales.

Todos los trabajos son desarrollados por grupos de investigación con sede en México e ilustran tanto los sólidos conocimientos nacionales en inteligencia artificial como nuestras capacidades para desarrollar soluciones avanzadas a problemas difíciles.

***Fernando Arámbula***



# Inteligencia artificial y su aplicación en la seguridad vial

## **José Anibal Arias Aguilar**

Universidad Tecnológica de la Mixteca, División de Estudios de Posgrado, Heroica Ciudad de Huajuapán de León, Oaxaca, México.

ORCID: 0000-0002-8838-877X

## **Eduardo Aragón López**

Universidad Tecnológica de la Mixteca, División de Estudios de Posgrado, Heroica Ciudad de Huajuapán de León, Oaxaca, México.

ORCID: 0009-0009-2787-9282

Recepción: 11 de noviembre de 2024.

Aceptación: 28 de marzo de 2025.

Septiembre 2025 • número de revista 13 • DOI: 110.22201/dgtic.26832968e.2025.13.53

## Inteligencia artificial y su aplicación en la seguridad vial

### Resumen

Debido a los altos índices de mortandad derivados de la conducción distraída, la cual es ocasionada por el desvío de la atención de la conducción segura, de la carretera y del entorno del conductor, es que surgen los Sistemas de Asistencia a la Conducción (ADAS, por sus siglas en inglés) como una forma de prevenir estos siniestros y asistir al conductor durante su trayecto. En el presente trabajo, se detectan y clasifican la conducción segura y nueve actividades distractoras en los conductores de automóviles mediante el uso de redes neuronales convolucionales, aplicadas primero a la base de datos "State Farm Distracted Driver Detection" y también a una base de datos propia. Además, con el uso de la biblioteca "TensorFlow Lite", el modelo se implementa en una aplicación móvil para su uso en tiempo real. El modelo resultante logra una exactitud del 91% en entrenamiento y del 95% en su validación. Con la implementación móvil, la exactitud del modelo se reduce en un 5%, sin embargo, se logran detectar y diferenciar correctamente las distracciones y se demostró la posibilidad de adaptar un ADAS en un dispositivo móvil.

**Palabras Clave:** ADAS, inteligencia artificial, redes neuronales, accidentes viales, dispositivos móviles.

## Artificial intelligence and its application in road safety

### Abstract

*Due to the high mortality rates resulting from distracted driving, which is caused by the diversion of attention from safe driving, the road and the driver's environment, Advanced Driver Assistance Systems (ADAS) emerge to prevent these accidents and assist the driver during their journey. In this work, safe driving and nine distracting activities in car drivers are detected and classified by using convolutional neural networks, applied first to the "State Farm Distracted Driver Detection" database and also to a self-created database. In addition, with the use of the "TensorFlow Lite" library, the model is implemented in a mobile application for real-time use. The resulting model achieves an accuracy of 91% in training and 95% in validation. In the mobile implementation, the accuracy of the model is reduced 5%, however, distractions are correctly detected and differentiated and the possibility of adapting an ADAS in a mobile device was demonstrated.*

**Keywords:** ADAS, artificial intelligence, neural networks, traffic accidents, mobile devices.

### Introducción

De acuerdo con la Organización Mundial de la Salud [1], en su informe sobre la situación de la seguridad vial 2023, se registraron alrededor de 1.19 millones de muertes por año desde el 2010 hasta la actualidad, siendo una de las principales causas de muerte de niños y jóvenes de entre los 5 y 29 años de edad. Debido al incremento del parque vehicular, se han desarrollado medidas para prevenir los altos índices de mortandad causados por los accidentes viales, entre ellos, encontramos a los Sistemas de Asistencia a la Conducción capaces de auxiliar a los conductores en la tarea compleja que conlleva la conducción segura. Sin embargo, los avances que han tenido estos sistemas por parte de la industria se han centrado únicamente en los modelos más recientes y caros, es decir, entre más funciones de seguridad contenga mayor será el costo del vehículo. Por otro lado, las investigaciones desarrolladas por parte de la comunidad científica se han basado principalmente en sistemas de visión artificial, los cuales buscan asistir al conductor mediante su monitorización con cámaras.

Liu et al.[2] utilizan redes neuronales convolucionales aplicadas a la visión por computadora; dentro de su marco propuesto, incluyen, en primer lugar, una red convolucional maestra (red preentrenada) que le transmitirá su conocimiento a una red estudiante; en conjunto, son capaces de detectar 9 distracciones con una exactitud del 95.64%. Por otro lado, Mbouna et al. [3] presentaron un análisis basado en el estado de apertura de los ojos y la postura de la cabeza para detectar signos de somnolencia, utilizando Máquinas de Soporte Vectorial (SVM) para hacer la clasificación entre distracción o conducción. Similarmente, en Fan y Shangbing [4], se trabaja con una red de fusión de características locales multiescala basada en imágenes para detectar 9 distracciones, obteniendo un 90.25% de exactitud. Por otra parte, en Xing et al. [5], se realizó un modelo que clasifica 3 actividades distractoras y otras 4 actividades de conducción normal; mediante la red preentrenada Alexnet, alcanzaron un 79% de exactitud en la clasificación.

## Desarrollo

### Materiales y métodos

En el presente estudio, se detalla la metodología de investigación llevada a cabo para la implementación de un Sistema de Asistencia a la Conducción (ADAS, por sus siglas en inglés) en un dispositivo móvil con sistema operativo Android, con el objetivo de monitorear al conductor y prevenirlo de actividades distractoras y potencialmente peligrosas. De esta manera, el sistema busca prevenir el riesgo de sufrir accidentes viales activando una alerta sonora cuando éste mismo, por medio del monitoreo en tiempo real, logre detectar una de las distracciones propuestas.

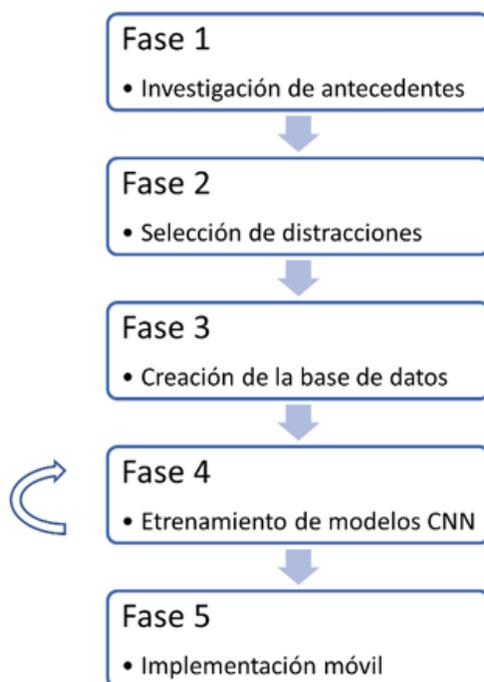
Para la captura de las fotografías con las que se construyó la base de datos, se trabajó con conductores de automóviles en un ambiente controlado a la luz de día, mientras que, para realizar los entrenamientos de las redes neuronales convolucionales, se utilizó un laboratorio computacional.

Para poder implementar el modelo de red neuronal convolucional en un dispositivo móvil, se requirió de un dispositivo con sistema operativo Android (versión indistinta) con el modo de depuración y desarrollador activados, necesarios para poner en funcionamiento el sistema por medio de una aplicación desarrollada en Android Studio.

## Estrategias de desarrollo

De acuerdo con Aguirre Barrera y Aguirre Barrera [6], la metodología en cascada es uno de los modelos que incluye ciertas ventajas en la realización de proyectos de software y de aplicaciones móviles, debido a que se organiza por medio de fases y cada una de ellas cuenta con entregables específicos; además, Velásquez et al [7] menciona que, con este tipo de metodología, resulta más sencillo realizar una planificación general del proyecto dado que, desde la primera fase, se acuerdan tanto los requerimientos como los recursos necesarios del mismo y, en las siguientes fases, se procede únicamente a llevar a cabo las actividades de desarrollo y prueba correspondientes con los recursos previamente planteados.

En este contexto, el ADAS se desarrolló siguiendo la metodología en cascada, destacando que esta metodología propone un orden y seguimiento riguroso de sus fases. Por tal motivo, fue necesario llevar a cabo una modificación en la Fase 4 (ver Fig. 1), en la cual, se lleva a cabo un proceso iterable correspondiente a los entrenamientos de los modelos de redes neuronales convolucionales y la búsqueda de los hiperparámetros del modelo óptimo.



**Fig. 1.** Metodología utilizada para el desarrollo del proyecto.

Tomando en consideración la investigación realizada durante la Fase 1, se llegó a la conclusión de utilizar técnicas de aprendizaje profundo, más concretamente, utilizar redes neuronales convolucionales debido a su uso particularmente eficaz en las tareas de extracción de características y clasificación de imágenes. De acuerdo con Artola Moreno y Pérez Carrasco [8], este tipo de redes neuronales artificiales optimizan las tareas de análisis de las imágenes por medio de la aplicación de filtros convolucionales encargados de extraer las características más representativas de cada imagen para su interpretación y clasificación.

En las Fases 2 y 3, se analizaron bases de datos con imágenes de conductores de automóvil realizando actividades distractoras. Como demuestran los estudios presentados en los trabajos de la introducción, las actividades distractoras más peligrosas fueron aquellas que involucran el uso de dispositivos móviles y las que incluyen apartar una o ambas manos del volante; motivo por lo cual, y a modo de comparación de resultados, se optó por trabajar con las distracciones presentadas en la Tabla I.

**Tabla I**  
Selección de distracciones

<b>Código</b>	<b>Nombre</b>
Clase 0	Conducción segura
Clase 1	Enviando SMS con mano derecha
Clase 2	Hablando por teléfono con mano derecha
Clase 3	Enviando SMS con mano izquierda
Clase 4	Hablando por teléfono con mano izquierda
Clase 5	Manipulando la radio/GPS
Clase 6	Ingiriendo bebidas
Clase 7	Buscando algo en la parte de atrás
Clase 8	Maquillándose/Arreglándose el cabello
Clase 9	Hablando con el pasajero

Para las primeras pruebas, se trabajó con la base de datos que se encuentra disponible en la página de Kaggle [9], la cual contiene 102,150 imágenes dentro de 10 clases diferentes; las imágenes cuentan con un tamaño de 640x480 y, durante el entrenamiento de los modelos de red, se redimensionaron a un tamaño de 48x48 para reducir el costo computacional de los cálculos matemáticos; en cada uno de los entrenamientos, se modificaron los hiperparámetros de los modelos hasta obtener los óptimos (ver Tabla II).

**Tabla II**

Mejores hiperparámetros con la base de datos de prueba.

Hiperparámetros	
Tasa de aprendizaje	0.0001
Tamaño del lote	200
Número de épocas	25

Durante la Fase 4, se capturaron fotografías con imágenes de conductores con la finalidad de desarrollar una base de datos propia y, de esta manera, adaptar el modelo para su uso en un contexto local, en conductores mexicanos (ver Fig. 2). Ésta cuenta con 10,000 imágenes que fueron etiquetadas y redimensionadas a un tamaño de 384x384 en formato RGB.



**Fig. 2.** Base de datos generada de las distracciones a detectar en conductores mexicanos.

Para los entrenamientos de los modelos de red, se utilizó la “transferencia de aprendizaje” con la finalidad de solventar la disminución de la cantidad de datos. Durante las pruebas con otros modelos, el modelo de red preentrenado que obtuvo mejores resultados fue la arquitectura de red “EfficientNet b1”. Éste es un modelo de red neuronal convolucional ligero en tamaño de almacenamiento (30.1 Mb) y en cantidad de parámetros entrenables (7,794,184); con éste, se alcanzó una mayor exactitud dentro de los entrenamientos, motivos por los cuales, se optó por utilizarlo debido a que uno de los requerimientos principales es que el modelo resultante sea lo más ligero posible en tamaño de almacenamiento debido a que estará almacenado dentro de una aplicación Android.

Este modelo realiza un método de escalamiento que ajusta uniformemente todas las dimensiones de profundidad, ancho y resolución de las imágenes de entrada, fue entrenado para clasificar 1,000 clases distintas a una resolución de 240x240 y está disponible para su uso libre por investigadores en el repositorio de redes preentrenadas de la librería de TensorFlow. Su arquitectura se muestra en la Figura 3.

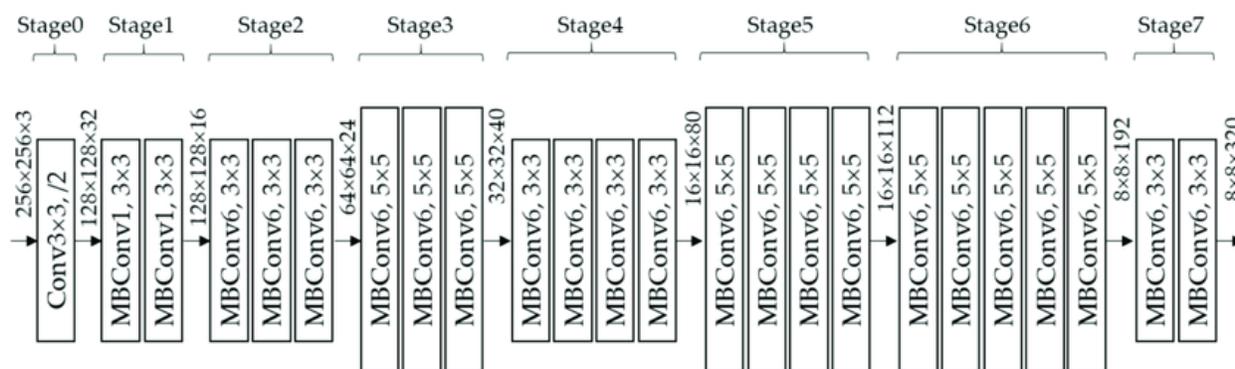
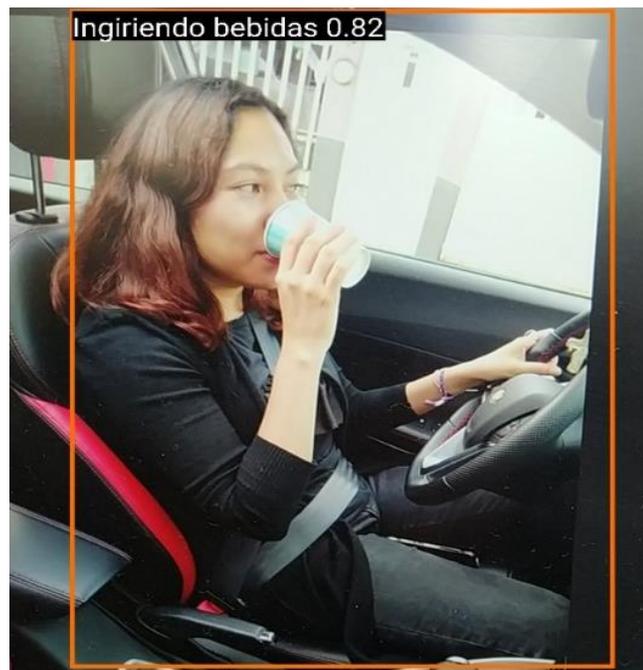


Fig. 3. Arquitectura del modelo "EfficientNet b1".

En la Fase 5, a pesar de contar con un modelo relativamente ligero, aún es necesario realizar un proceso de conversión con la herramienta “TensorFlow Lite”, la cual permite implementar los modelos de aprendizaje de forma eficiente en dispositivos con recursos de procesamiento y memoria limitados. Así pues, en comparación con el formato del modelo original, éste cuenta con un tamaño de almacenamiento aún más reducido y ayuda a lograr una inferencia más rápida, lo que logra reducir el consumo de energía para su uso en este

tipo de dispositivos. Esta misma aplicación cumple la función de detectar y realizar la predicción de clasificar, en tiempo real, las acciones del conductor; en busca de una de las 10 clases con las que fue entrenado el modelo, mediante un cuadro delimitador, enmarca al conductor al mismo tiempo que indica la etiqueta con la clase detectada junto con su grado de confianza (ver Fig. 4).



**Fig. 4.** ADAS en funcionamiento.

La aplicación incluye una interfaz de configuración en la que se recomienda que sólo una persona capacitada pueda manipularlo, debido a que, en ella, se puede incrementar o disminuir el umbral de detección del sistema. Lo anterior nos puede ayudar tanto a clasificar con menos errores el tipo de distracción y la clase a la que pertenece la detección, como a poder detectar, de manera menos precisa pero más rápida, una distracción. Los valores predeterminados y los recomendados de la interfaz de configuración se muestran en la Figura 5.

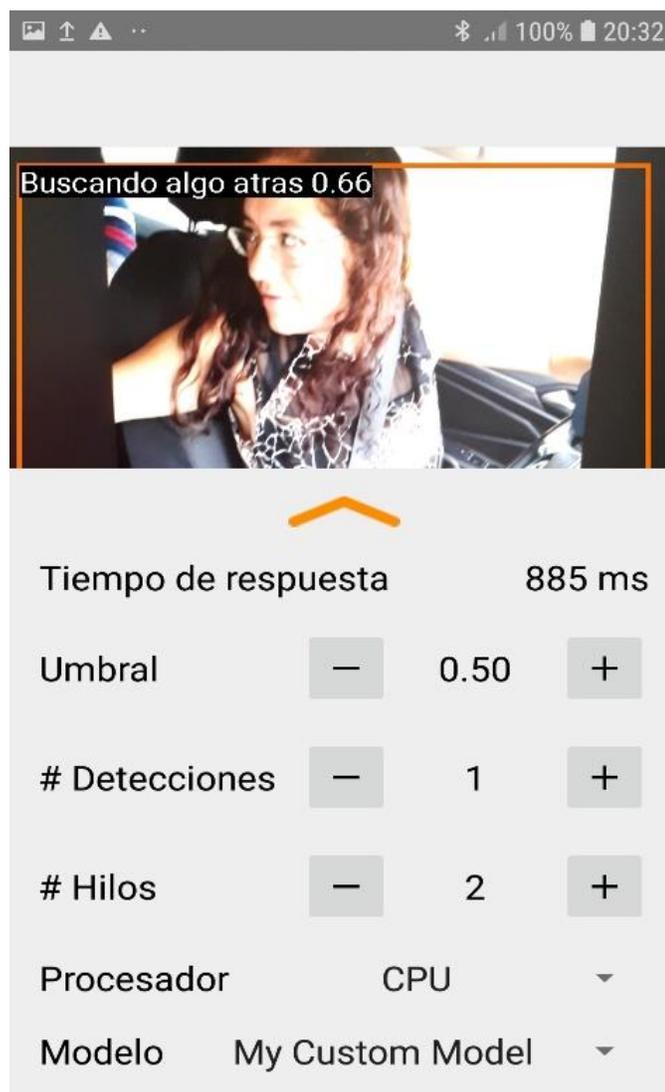


Fig. 5. Interfaz de configuración de la aplicación.

El ADAS también incorpora un sistema de advertencia, el cual se activa una vez que detecta una de las actividades distractoras y mediante un audio le indica al conductor la recomendación “Por favor, no te distraigas”.

Finalmente, para evaluar los resultados obtenidos por los modelos, se llevó a cabo una comparación entre el porcentaje de exactitud obtenido en cada modelo, el desempeño en sus gráficas de exactitud y pérdida, y el análisis de la matriz de confusión para problemas de clasificación multiclase.

## Resultados

El primer paso para evaluar el desempeño del ADAS se dio mediante el cálculo de la exactitud y pérdida de los modelos de aprendizaje; una vez comparados, se seleccionó el modelo con el mejor resultado en exactitud. La Tabla III contiene los resultados obtenidos del modelo entrenado con el conjunto de datos [9], mientras que la Tabla IV contiene los resultados que se obtuvieron con el conjunto de datos que fue creada utilizando la transferencia de aprendizaje con el modelo preentrenado "Efficient Net b1".

**Tabla III**

Resultados obtenidos con el conjunto de datos de prueba.

Entrenamiento	Exactitud	0.93
	Pérdida	0.21
Validación	Exactitud	0.94
	Pérdida	0.095

**Tabla IV**

Resultados obtenidos con el conjunto de datos creada y con transferencia de aprendizaje.

Entrenamiento	Exactitud	0.91
	Pérdida	0.41
Validación	Exactitud	0.95
	Pérdida	0.25

Como se puede apreciar en la Tabla IV, el modelo óptimo que se obtuvo fue al utilizar el modelo preentrenado "EfficientNet b1" con la base de datos creada; dicho modelo alcanza una exactitud de 95%.

Las gráficas de la Figura 6 y la Figura 7 corresponden a las curvas de exactitud y pérdida generadas durante el aprendizaje del modelo óptimo. Es importante utilizarlas como método de evaluación debido a que son un indicador de un correcto o mal aprendizaje por parte del modelo: si las curvas de entrenamiento y validación son similares entre sí, como en la Figura 6 y la Figura 7, se demuestra un correcto aprendizaje y se evita de problemas como el “sobreajuste”.

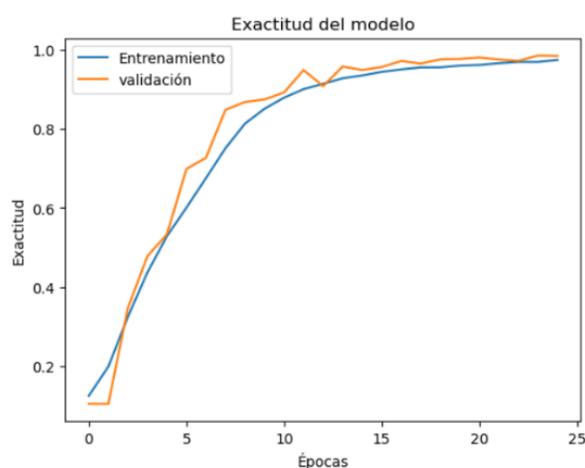


Fig. 6. Gráfica de exactitud del modelo.

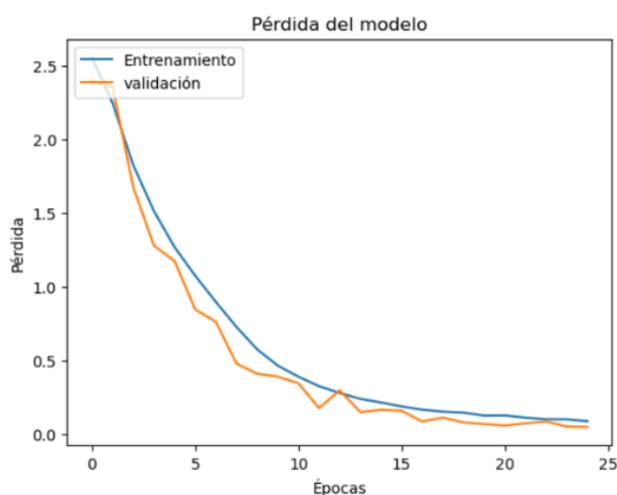
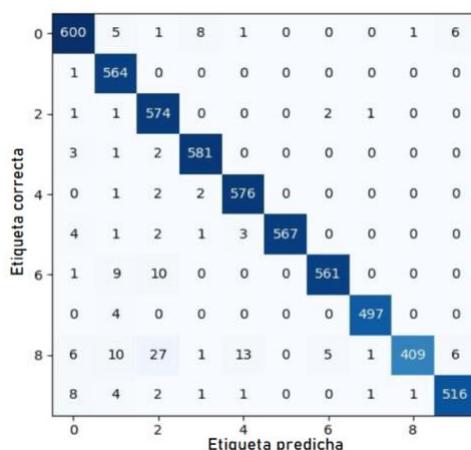


Fig. 7. Gráfica de pérdida del modelo.

El modelo resultante es capaz de clasificar 10 clases diferentes, enlistadas en la Tabla 1. La clase 0 “Conducción segura” corresponde a la clase por defecto a la cual deben de ir todos los conductores de automóvil y, por consiguiente, es la clase en la cual el sistema no emitirá ninguna alerta; por otro lado, las demás clases corresponden a distracciones potencialmente peligrosas.

Finalmente, como parte del análisis del desempeño del modelo, se utilizó una matriz de confusión (ver Fig. 8) en la que se pudo obtener que la clase que mejor es clasificada por el sistema corresponde a la clase 1 “Enviando SMS con mano derecha”, mientras que la clase más difícil de diferenciar entre las demás corresponde a la clase 8 “Maquillándose/Arreglándose el cabello”.



**Fig. 8.** Matriz de confusión obtenida del modelo óptimo en donde se aprecia un óptimo desempeño de clasificación.

Es importante destacar que, basado en la evaluación de los resultados, se obtuvo un modelo óptimo equiparable con los presentados en el estado del arte, los cuales reportan un aproximado del 90% en la exactitud de sus modelos. Aunado a esto, se cumplió el propósito de innovación, que consiste en implementar un modelo de aprendizaje profundo en un dispositivo móvil. A modo de resumen, la Tabla V muestra una comparación entre los modelos del estado del arte y la propuesta desarrollada en este artículo. En este sentido, los modelos del estado del arte son capaces de clasificar dos o más distracciones, sin embargo, carecen de una implementación real (en campo) del modelo.

De esta manera, los resultados obtenidos con la implementación revelan que el modelo presentado es equiparable con las soluciones ya propuestas. Finalmente, cabe aclarar que el sistema se ve influenciado por diversos factores que afectan su desempeño, tales como los cambios de iluminación y el movimiento constante de la cámara dentro del vehículo.

**Tabla V**

Comparación del ADAS con el estado del arte

Proyecto	Distracciones analizadas	Tecnología empleada	Evaluación del modelo
[2]	Detección de 9 distracciones	Redes neuronales convolucionales	Exactitud del 95.64%, no implementado
[3]	Detección de somnolencia	Máquinas de Soporte Vectorial	Alta precisión
[4]	Detección de 9 distracciones	Redes neuronales convolucionales	Exactitud del 90.25%, no implementado
[5]	Detección de 3 distracciones	Redes neuronales convolucionales	Exactitud del 79%, no implementado
ADAS	Detección de 9 distracciones	Redes neuronales convolucionales	Exactitud del 95%, implementado en un dispositivo móvil

## Conclusiones y trabajo a futuro

El alto índice de accidentes viales derivados de una distracción por parte del conductor se ha convertido en un tema de gran relevancia para la industria y para la comunidad científica, quienes se han dedicado a investigar posibles soluciones que puedan aminorar esta problemática. La mayoría de los sistemas propuestos se centran en analizar sólo alguna distracción en específico, mientras que otras conllevan técnicas más complejas y difíciles de

reproducir, mismas que implican un aumento en los costos económicos como computacionales. Este artículo abordó la implementación de un ADAS en un dispositivo móvil mediante la metodología en cascada, cuyo objetivo es poder contribuir a la detección y alerta de distracciones en conductores, buscando así ayudar a disminuir el riesgo de sufrir un accidente de tránsito. Los resultados obtenidos en la implementación del modelo demuestran haber obtenido resultados equiparables con los trabajos mencionados en el estado del arte, los cuales se centraron en analizar una o dos distracciones, además de que es destacable que en ninguno de los trabajos propuestos se llevó a cabo una implementación (en un dispositivo móvil) como la propuesta en este trabajo.

Como parte del trabajo a futuro, se tiene contemplado trabajar mediante un análisis temporal de las distracciones con el uso de redes neuronales recurrentes, además de incorporar distracciones más complejas tales como la somnolencia y el estado de ebriedad.

## Referencias

- [1] Informe sobre la situación mundial de la seguridad vial 2023 [Global status report on road safety 2023]. Ginebra: Organización Mundial de la Salud; 2024. Licencia: CC BY-NC-SA 3.0 IGO. [En línea]. Disponible: <https://www.who.int/es/publications/i/item/9789240086517>
- [2] D. Liu, T. Yamasaki, Y. Wang, K. Mase y J. Kato, "Toward Extremely Lightweight Distracted Driver Recognition With Distillation-Based Neural Architecture Search and Knowledge Transfer," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 1, pp. 764-777, Jan. 2023, doi: 10.1109/TITS.2022.3217342
- [3] R. Oyini Mbouna, S. G. Kong y M. G. Chun, "Visual Analysis of Eye State and Head Pose for Driver Alertness Monitoring," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 3, pp. 1462-1469, Sept. 2013, doi: 10.1109/TITS.2013.2262098.
- [4] L. S. Fan y G. Shangbing, "Optimally-Weighted Multi-Scale Local Feature Fusion Network for Driver Distraction Recognition," in *IEEE Access*, vol. 10, pp. 128554-128561, 2022, doi: 10.1109/ACCESS.2022.3224585.

- [5] Y. Xing et al., "End-to-End Driving Activities and Secondary Tasks Recognition Using Deep Convolutional Neural Network and Transfer Learning," 2018 IEEE Intelligent Vehicles Symposium (IV), Changshu, China, 2018, pp. 1626-1631, doi: 10.1109/IVS.2018.8500548.
- [6] J. Aguirre Barrera y S. Aguirre Barrera, "Metodologías para el desarrollo de proyectos", Unicatólica, 2021. [En línea]. Disponible: <https://repository.unicatolica.edu.co/handle/20.500.12237/2037>
- [7] S. M. Velásquez, J. D. Vahos Montoya, M. E. Gómez Adasme, E. J. Restrepo Zapata, A. A. Pino, y S. Londoño Marín, "Una revisión comparativa de la literatura acerca de metodologías tradicionales y modernas de desarrollo de software", in Rev. Cintex, vol. 24, n.º 2, pp. 13–23, dic. 2019.
- [8] A. Artola Moreno y J. A. Pérez Carrasco. "Clasificación de imágenes usando redes neuronales convolucionales en Python". Trabajo de fin de grado en Ingeniería de las Tecnologías de Telecomunicación, Universidad de Sevilla, Sevilla, 2019. [En línea]. Disponible: <https://biblus.us.es/bibing/proyectos/abreproy/92402/fichero/TFG-2402-ARTOLA.pdf>
- [9] Kaggle. "State Farm Distracted Driver Detection". 2024 [En línea]. Disponible: <https://www.kaggle.com/competitions/state-farm-distracted-driver-detection/data>



# Exploración de la Comunicación Visual del Cambio Climático mediante la Clasificación Automática de Imágenes

## **Gabriela Elisa Sued**

Universidad Nacional Autónoma de México, Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas (IIMAS), Postdoctorante SECHITI, Ciudad de México, México.  
ORCID: 0000-0002-4516-678X

## **Wilmer Efrén Pereira-González**

Instituto Tecnológico Autónomo de México (ITAM) Departamento de Computación y Ciencia de Datos; Universidad Internacional de la Rioja, Facultad de Ingeniería, Ciudad de México, México.  
ORCID: 0000-0001-9699-7121

## **Judith Zubieta-García**

Universidad Nacional Autónoma de México, Instituto de Investigaciones Sociales, Ciudad de México, México.  
ORCID: 0000-0001-5983-987X

Recepción: 11 de noviembre de 2024.

Aceptación: 10 de junio de 2025.

Septiembre 2025 • número de revista 13 • DOI: 10.22201/dgtic.26832968e.2025.13.52

## Exploración de la Comunicación Visual del Cambio Climático mediante la Clasificación Automática de Imágenes

---

### Resumen

Uno de los múltiples desafíos que presenta el cambio climático es el de su adecuada difusión entre la población, pues eso contribuye a crear conciencia, modificar hábitos y divulgar acciones de mitigación. A fin de obtener un diagnóstico exploratorio de las imágenes con las que actualmente se comunica este fenómeno, así como sus impactos devastadores, este trabajo presenta un ejercicio exploratorio para la clasificación automática de imágenes mediante una red neuronal preentrenada a partir de etiquetas que describen causas, consecuencias y acciones de mitigación. Entre los hallazgos, encontramos la presencia de una mayor representación de acciones de adaptación y mitigación que de impactos y actividades antropogénicas relacionadas con el aumento de la temperatura del planeta. Respecto a la técnica empleada, describimos cuáles fueron las etiquetas con las que se logró una mejor clasificación y cuáles tuvieron una precisión baja, al tiempo que argumentamos las posibles causas de este desempeño desigual.

**Palabras Clave:** cambio climático, redes neuronales, comunicación visual, mitigación ambiental.

### *Exploring Climate Change Visual Communication through Automatic Image Classification*

---

#### **Abstract**

*One significant challenge posed by climate change is effectively introducing this trend to society to raise awareness and encourage behavioral change. This also involves disseminating mitigation*

*strategies and proposals that can foster greater commitment to environmental stewardship. To provide an initial assessment of the images used to communicate this phenomenon and its severe impacts, this paper presents an exploratory analysis in which images were automatically classified with a pre-trained neural network. The classification was based on labels describing the causes, consequences, and mitigation actions related to climate change. Our analysis revealed a more substantial representation of adaptation and mitigation measures compared to those documenting impacts and anthropogenic activities contributing to global temperature increases. Regarding the technique employed, labels utilized to achieve better classification outcomes are described as well as those rendering lower accuracy, while discussing potential factors contributing to this performance disparity.*

**Keywords:** *climate change, neural networks, visual communication, environmental mitigation.*

## Introducción: la comunicación visual del cambio climático

El cambio climático se define como variaciones meteorológicas atribuidas directa o indirectamente a la actividad humana que alteran la composición de la atmósfera y que se suman a la variabilidad natural del ambiente observada durante períodos de tiempo comparables. Los gases producidos por la actividad industrial, conocidos también como “gases de efecto invernadero”, producen calentamiento de la Tierra, con consecuencias devastadoras para el futuro del planeta [1]. En México, las temperaturas promedio han aumentado aproximadamente 0.85 grados centígrados y se espera que, entre 2015 y 2039, la temperatura promedio aumente entre 1.5 y 2°C en todo el territorio. Este incremento ya empieza a tener efectos sobre la biodiversidad y los sistemas productivos [2].

El cambio climático comienza a ser abordado en la academia como un conjunto de hechos empíricos y observaciones basadas en datos que se traduce en la opinión pública como un conjunto de valores, representaciones y guías para la acción, tanto comunitarias como individuales, preferentemente derivados de investigaciones científicas rigurosas, pero en formas y contenidos accesibles a públicos diversos [3]. Si bien el cambio climático también ha entrado en la agenda de actores institucionales y políticos, el involucramiento cotidiano

de los ciudadanos sigue representando un reto [4]. En este contexto, la comunicación mediática adquiere una mayor importancia para reconocer tanto el origen antropogénico del cambio climático como sus desafíos.

Las redes sociales constituyen una vía destacada para la comunicación de información, especialmente entre los segmentos más jóvenes de la sociedad. Además, éstos se han convertido en agentes de gran relevancia en la configuración del discurso público sobre el cambio climático, llegando incluso a relevar a los medios de comunicación tradicionales en el debate [5]. En los medios digitales, la expresión pública de la ciencia es impulsada por muy diversos actores y presenta características heterogéneas. La que corresponde al cambio climático, por ejemplo, incluye entidades de gobierno, activistas, medios de comunicación, portales educativos, prousuarios —usuarios que producen a la vez que consumen contenidos [6]— comprometidos con su remediación, instituciones científicas, entre otros. Los contenidos también son variables: anuncian desde cumbres de gobierno hasta ecosistemas dañados, pasando por las causas del daño ambiental, las iniciativas para la definición de políticas locales y regionales que ayuden a su mitigación y conferencias de expertos.

El proyecto que aquí se presenta aprovechó los recursos ofrecidos por la Alianza UNAM-Huawei para analizar un conjunto de 11,846 imágenes publicadas en las plataformas Google, Instagram, Twitter y Facebook con el propósito de identificar las temáticas principales a las que éstas aluden, así como las diversas formas de representación del cambio climático. Este trabajo se complementó con un cuestionario aplicado a población mayoritariamente universitaria ( $n=351$ ), a quienes se solicitó evaluar una muestra de 20 imágenes que se consideraron representativas de toda la base de datos recopilada, en cuanto a su efectividad de representación del cambio climático, los sentimientos que les evocaban y si les invitaban a modificar actitudes o hábitos. Por razones de espacio, aquí sólo se comunican los resultados de la clasificación automática. En un trabajo previo se presentó el proceso de iteración seguido para obtener las etiquetas de la clasificación automática, antes de la revisión manual [7].

Mediante las frecuencias obtenidas para las imágenes de la base, se pretende, por un lado, identificar qué tipo de contenidos predominan en la representación gráfica del cambio

climático y, por el otro, evaluar el uso de técnicas de reconocimiento automático de imágenes en las ciencias sociales, aplicación poco explorada hasta el momento.

A continuación, se desarrolla la metodología empleada; posteriormente, se exponen los hallazgos y, finalmente, se plantean algunas conclusiones en las que también se identifican alcances y limitaciones del trabajo desarrollado.

## Metodología

La investigación se realizó en tres fases: recolección, procesamiento y análisis de los datos. En la primera, se empleó software de recolección automática de datos digitales para identificar imágenes publicadas en las plataformas Google, Instagram, Twitter y Facebook entre 2020 y 2022. La búsqueda se ejecutó con las claves 'cambio climático' y 'calentamiento global'. De acuerdo con las características y limitaciones de cada plataforma, se intentó restringir la búsqueda a datos producidos en México o accedidos desde este país. Luego de eliminar imágenes repetidas, se obtuvo una base compuesta por 11,846 imágenes.

Para facilitar una primera exploración del contenido de la base, se utilizó la aplicación ImageX [8], que permite obtener visualizaciones agrupadas por patrones cromáticos y temáticos, compuestas por miniaturas de las imágenes originales. Con el apoyo de un experto en medio ambiente, se determinaron 43 categorías de contenido que, después de ser analizadas detalladamente, se redujeron a 29 y, finalmente, a 23, descontando las que no mostraban una clara relación con el tema de interés. Esta tipología de contenidos sirvió para calcular las frecuencias de aquellas imágenes que se consideraron altamente relacionadas con el cambio climático.

El etiquetado automático de las imágenes según las categorías previamente determinadas fue realizado mediante CLIP, una red neuronal preentrenada que aprende a partir de imágenes y textos tomados de Internet [9]. Esta herramienta de deep learning usa la estrategia Zero-Shot Learning, que consiste en generar secuencias cortas de texto o etiquetas asociadas a cada imagen que no forman parte de los datos de entrenamiento; es decir, la imagen dada no necesariamente formó parte del conjunto de entrenamiento. En nuestro trabajo, la utilizamos en varias secuencias de texto o etiquetas relevantes para el

cambio climático y así determinar la probabilidad de pertenencia de cada imagen [10]. Cada imagen se asoció a la etiqueta cuyos rasgos reportaron el mayor porcentaje de pertenencia a esa clase o etiqueta en todas las pruebas que se realizaron.

Para lograr una clasificación que tuviera una precisión cercana al 75%, se realizaron cinco iteraciones y varias operaciones sobre el nombre de las categorías, con el propósito de simplificar la semántica, precisar su denominación, y adaptarlas a las decisiones que tomaba el algoritmo. Para mejorar la categorización automática en las iteraciones sucesivas, se tomaron las siguientes decisiones:

Suprimir las categorías con significados amplios; por ejemplo, ‘no relacionado con el cambio climático’, o ‘sin definir’, porque allí se depositó más del 30 % de las imágenes. De este modo, se forzó la herramienta a que realizara un mayor esfuerzo para colocarlas en categorías específicas (como ‘activismo’, o ‘polución’).

Adecuar las etiquetas para facilitar el proceso de clasificación mediante el uso de CLIP. Por ejemplo, en la categoría ‘agricultura’, definida inicialmente, CLIP juntaba imágenes con otras que referían a alimentos de granja y a ferias de productores. Por ello, en vez de catalogar nuevamente esas imágenes, se reformuló su denominación a ‘*agriculture farm food*’. Así se minimizó el traslape con la intención de obtener clases “casi” mutuamente excluyentes.

Sustituir los sustantivos y conceptos abstractos por denominaciones concretas: por ejemplo, ‘children’ se cambió por ‘*group of children*’, o ‘garbage’ por ‘*piles of trash*’, y ‘*political deliberations*’ por ‘*politicians deliberating*’.

Reclasificar manualmente objetos simples y otros complejos. Dentro de los primeros, por ejemplo, el sistema preentrenado no reconoció a las aves como animales, sino que las clasificó como ‘naturaleza’. En el caso de los segundos, no reconoció imágenes humorísticas, tales como caricaturas o memes, ni imágenes diseñadas por computadora, evidenciando dificultades para clasificar imágenes acompañadas de texto. En este proceso, se crearon dos nuevas etiquetas: ‘imágenes no representativas y sólo texto’, y ‘conciencia sobre cambio climático’, que se sumaron a las existentes y se les agregaron las imágenes manualmente.

Corroborar con una estrategia de clustering, si había agrupamientos con poco traslape (a conservar) y/o objetos fuera de cualquier clase o outliers (a eliminar). Si el número de agrupamientos mutuamente excluyentes y los demás presentaban poco traslape o coincidencia con otras clases, se detenía el proceso. Las categorías con menor coincidencia con otras clases fueron cuatro: 'ártico', 'sequía', 'planeta Tierra' y 'eficiencia energética', que constituyen el 18% del total de categorías. Por otro lado, seis parejas de clases no tuvieron coincidencia con otras categorías, constituyendo éstas el 27% de la muestra. Por lo tanto, al llegar al 45% de etiquetas con mínimo traslape, se decidió detener el procesamiento.

Posteriormente, se analizaron las frecuencias obtenidas en dos direcciones. Primero, se realizó un análisis de contenido en el que se agruparon las categorías según se vincularan con causas, consecuencias o con la mitigación del cambio climático. Luego, se determinó la efectividad alcanzada por la categorización automática, comparando los resultados iniciales de la misma con los obtenidos después de la recategorización manual.

## Hallazgos

Según el Grupo Intergubernamental de Expertos sobre el Cambio Climático (IPCC) [11], el cambio climático es impulsado principalmente por actividades humanas, especialmente la quema de combustibles fósiles, la deforestación y ciertos procesos industriales, que aumentan las concentraciones de gases de efecto invernadero (GEI) y elevan las temperaturas globales. Estas transformaciones se representaron en la muestra analizada mediante las etiquetas 'contaminación', 'plástico', 'deforestación', 'petróleo', 'autos', 'ganado' y 'montones de basura', las cuales representaron el 17% del total de imágenes en la muestra.

Estos factores generan graves impactos en ecosistemas, salud, agricultura y comunidades vulnerables. Las imágenes pertenecientes a estas categorías incluyen al planeta Tierra envuelto en llamas, animales en peligro de extinción, sequías, paisajes de los polos ártico y antártico, afectaciones a la salud y a la infancia, inundaciones, incendios forestales y escasez de agua, y equivalieron al 37% del total de imágenes.

Frente a estos impulsos humanos y sus impactos, ciertas acciones se enfocan en crear infraestructuras y prácticas de adaptación para reducir la vulnerabilidad a eventos concretos,

como la creación de conciencia sobre el cambio climático, las acciones para efficientar la energía o proteger la naturaleza, como actitud positiva hacia el cuidado del planeta. Por su parte, la mitigación implica la reducción de gases de efecto invernadero mediante cambios energéticos, en el transporte y la industria, y la promoción de tecnologías sostenibles y políticas climáticas [12]. Las acciones de adaptación y mitigación corresponden al 14% y el 25% de la muestra, respectivamente. Finalmente, los textos con formato de imagen y las imágenes no relacionadas con el cambio climático abarcaron el 7% de la muestra.

La Tabla I muestra la frecuencia encontrada para las etiquetas utilizadas en el proceso de clasificación. La de mayor frecuencia es la correspondiente al planeta Tierra. Esta etiqueta incluye imágenes intervenidas o diseñadas por computadora que representan al planeta quemándose o derritiéndose. En segundo lugar, aparecen imágenes vinculadas a prácticas de adaptación relacionadas con el ahorro de energía, el uso de energías sustentables como paneles solares y molinos eólicos, y diversas recomendaciones para el ahorro de energía en el hogar. Estas acciones de mitigación se encuentran más presentes en la muestra que las acciones de reciclaje, que están más al alcance de la ciudadanía. El número relativamente alto de imágenes sobre políticos deliberando en diferentes contextos, ya sea en recintos legislativos, cumbres de cambio climático o incluso en campañas electorales, muestra que el tema se ubica en las agendas políticas, a pesar de que se minusvaloran los daños que las diferentes regiones sufren debido al cambio climático antropogénico.

**Tabla I.** Frecuencia de las etiquetas utilizadas.

Categoría	Etiqueta	Frecuencia %
<b>Impulso por actividad humana</b>	Contaminación	5
	Plástico	4
	Deforestación	3
	Ganado	1
	Montones de basura	1
	Petróleo	2
	Automóviles	1
	Subtotal	17
<b>Impacto</b>	Planeta Tierra	15
	Animales (en extinción)	4
	Sequía	4
	Ártico	4
	Salud	3
	Inundación	2
	Grupo de niños	2
	Agua	2
	Incendios forestales	1
	Subtotal	22
<b>Adaptación</b>	Eficiencia energética	9
	Alimentos de agricultura y granja	3
	Conciencia sobre cc	2
Subtotal	14	
<b>Mitigación</b>	Políticos deliberando	8
	Activismo	7
	Reciclaje	5
	Naturaleza	3
	Bicicletas	2
	Subtotal	25

Un segundo nivel de hallazgos corresponde a la precisión con la que el método empleado para el etiquetado categorizó las imágenes. Dicho porcentaje es variable, aunque sea independiente del número de imágenes que se hayan agrupado en cada categoría. La Tabla II muestra las categorías con las que CLIP logró una precisión igual o superior al 90%.

**Tabla II.** Precisión mayor al 90% en el etiquetado automático.

<b>Etiqueta</b>	<b>CLIP</b>	<b>Manual</b>	<b>Definitiva</b>	<b>Precisión %</b>
Plástico	496	1	497	100
Petróleo	209	1	210	100
Montones de basura	68	1	69	99
Deforestación	301	5	306	98
Salud	364	8	372	98
Ganado	90	3	93	97
Bicicletas	188	7	195	96
Grupo de niños	206	8	214	96
Autos	157	9	166	95
Reciclaje	537	39	576	93
Ártico	411	44	455	90
Alimentos de agricultura y granja	351	39	390	90

Algunas de las etiquetas que tuvieron precisiones elevadas fueron renombradas, como se describió en la metodología, a partir de diferentes estrategias. Por ejemplo, la sustitución de conceptos abstractos por expresiones concretas, como es el caso de ‘basura’ por ‘montones de basura’ o infancia por ‘grupos de niños’. En otros casos, funcionó bien ajustar las etiquetas a las clasificaciones que realizaba CLIP; por ejemplo, sustituyendo ‘enfermedad’ por ‘salud’ y ‘agricultura’ por ‘alimentos de agricultura y granja’.

Hay otros casos en los que las etiquetas corresponden a objetos muy concretos, como ‘bicicletas’, ‘autos’ o ‘ganado’. Marcadores un poco más abstractos como ‘reciclaje’ o ‘ártico’ muestran un buen desempeño del entrenamiento de CLIP en esos tópicos. El segundo, especialmente, presenta patrones de colores y formas asociadas a hielos y cielos que seguramente hacen que los objetos incluidos en las imágenes sean reconocidos fácilmente.

La Tabla III presenta las etiquetas que no obtuvieron precisión alguna o en las que ésta fue muy baja.

**Tabla III.** Etiquetas con menor precisión.

Etiqueta	CLIP	Manual	Definitiva	Precisión %
No representativas y texto	0	789	789	0
Conciencia sobre cc	0	259	259	0
Humor	0	29	29	0
Eficiencia energética	3545	-2510	1035	29
Políticos deliberando	484	514	998	48
Activismo	577	275	852	68
Naturaleza	553	-152	401	73
Sequía	340	128	468	73
Incendio	108	22	130	83

La precisión nula se debió a la existencia de imágenes poco literales, como aquellas relacionadas con ‘humor’, las que resultaron poco concretas, como el caso de ‘activismo’, o también por falta de precisión de la red al clasificar a las ‘aves’ como ‘naturaleza’ y no como ‘animales’. La herramienta tampoco fue capaz de distinguir archivos con formato de imágenes que contenían solamente texto, diseños realizados por computadora, ni imágenes no representativas del cambio climático.

En el caso de ‘eficiencia energética’, se vio que la red tiende a usar una etiqueta como contenedora de todas las imágenes que no logró distribuir en el resto. Esta categoría contenía, además, imágenes semánticamente complejas, ya que combinaban fotografías, diseño digital de imagen y textos; según se ha advertido, esto dificultó la clasificación con CLIP. En ese marcador, estaban depositadas las imágenes que luego fueron clasificadas manualmente como ‘conciencia sobre cambio climático’ y que consistían, en general, en combinaciones entre texto, dibujo digital e imagen diseñada por computadora.

## Conclusiones

Esta investigación permitió explorar el etiquetado automático de una muestra de 11,846 imágenes recolectadas de plataformas digitales y relacionadas con el cambio climático, a

través de un algoritmo que usó CLIP como red neuronal preentrenada; éstas fueron clasificadas temáticamente en 23 categorías, construidas a partir de una exploración previa de la muestra.

Las categorías relacionadas con las actividades humanas causantes del aumento en la temperatura obtuvieron una frecuencia menor que los impactos, mientras que las acciones de adaptación y mitigación presentaron, en conjunto, frecuencias mayores que las de los impactos. En un trabajo anterior, donde se analizan los resultados de una encuesta a personas residentes en México [7], se determinó que las representaciones del cambio climático se asocian directamente con causas antropogénicas, que en esta muestra son las que poseen menor frecuencia de aparición, por lo que el incremento de la exposición a este tipo de imágenes generaría mayor conciencia en la ciudadanía sobre los daños ambientales de origen antropogénico.

Cabe mencionar que, durante las cinco iteraciones realizadas con CLIP, se observó un proceso de reducción, simplificación y ajuste semántico en las etiquetas que resultó satisfactorio, en términos generales, puesto que mejoró el proceso de clasificación.

También se observó que CLIP funciona mejor si las etiquetas tienden a ser semánticamente concretas y léxicamente simples. No obstante, no se pudieron resolver las limitaciones que surgieron en el proceso de clasificar imágenes que combinan elementos, tales como texto, dibujo, fotografía o diseño gráfico.

## Agradecimientos

Los autores agradecen al Mtro. Uriel Martínez-Sánchez el apoyo técnico en el procesamiento de imágenes. Este artículo fue realizado con apoyo de la Alianza para promover el desarrollo de las capacidades digitales en México, convocatoria 2022. La Dra. Sued agradece también el apoyo del programa de estancias postdoctorales SECIHTI (2023-2025) y del Instituto de investigaciones Sociales de la UNAM.

## Referencias

- [1] Organización de las Naciones Unidas (ONU), "Protocolo de Kioto de la Convención Marco de las Naciones Unidas sobre el Cambio Climático", 1998. [En línea]. Disponible: <https://unfccc.int/resource/docs/convkp/kpspan.pdf>
- [2] INCC-Instituto Nacional del cambio climático, "Impactos del Cambio Climático en México [México ante el cambio climático]", 2022. [En línea]. Disponible: <https://cambioclimatico.gob.mx/impactos-del-cambioclimatico-en-mexico/>
- [3] C. Callison, *How Climate Change Comes to Matter: The Communal Life of Facts*. Durham, NC: Duke University Press, 2014.
- [4] ] A. Giddens, *The politics of climate change*. London: Polity Press, 2011.
- [5] ] B. León y M. del C. Erviti, "La comunicación del cambio climático en redes sociales: fortalezas y debilidades", en Daniel Rodrigo Cano, Rocío Mancinas-Chávez y Rogelio Fernández-Reyes (eds.) *La comunicación del cambio climático: una herramienta ante el gran desafío* pp.181-221 Madrid: Dykinson, 2021.
- [6] A. Bruns, "The Future Is User-Led: The Path towards Widespread Producers", *Fibreculture Journal*, núm. 11, 2008, consultado el 04 de noviembre de 2024. [En línea] <https://doaj.org/article/2c46e8b335b54104bc419ae882911170>
- [7] G. Sued, J. Zubieta-García, W. Pereira, J. L. Hernández, A. Rodríguez, R. Toledo, A. Carmona, y U. Martínez, "Las imágenes digitales del cambio climático en México: de problema científico a tema de opinión pública", en *Memorias de las Segundas Jornadas de Estudios Sociales de Ciencia y Tecnología*, 29–47, 2023. Mérida: Universidad Nacional Autónoma de México. [https://www.academia.edu/110313314/Memorias\\_de\\_las\\_Segundas\\_Jornadas\\_de\\_Estudios\\_Sociales\\_de\\_Ciencia\\_y\\_Tecnolog%C3%ADa?source=swp\\_share](https://www.academia.edu/110313314/Memorias_de_las_Segundas_Jornadas_de_Estudios_Sociales_de_Ciencia_y_Tecnolog%C3%ADa?source=swp_share)
- [8] N. Hezel y K. Werthel, ImageX. (2018). *Visual Computing*. [software] <https://visual-computing.com/project/imagex/>

- [9] Openai, CLIP. *Connecting Text and Images* (2021). [software]  
<https://openai.com/research/clip>
- [10] A. Radford et al., "Learning Transferable Visual Models from Natural Language Supervision", *Proceedings of the 38th International Conference on Machine Learning, PMLR*, pp. 8748–8763, jul. 2021. Consultado el 4 de noviembre de 2024 [En línea]. Disponible: <https://proceedings.mlr.press/v139/radford21a.html>
- [11] Grupo Intergubernamental de Expertos sobre el Cambio Climático, "Cambio Climático 2021. Bases físicas. Resumen para responsables de políticas", 2021. [En línea]. Disponible: [https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC\\_AR6\\_WG1\\_SPM\\_Spanish.pdf](https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WG1_SPM_Spanish.pdf)
- [12] Grupo Intergubernamental de Expertos sobre el Cambio Climático, "Climate Change 2022: Impacts, Adaptation and Vulnerability", 2022. [En línea]. Disponible: <https://www.ipcc.ch/report/sixth-assessment-report-working-group-ii/>



# Optimización de la vigilancia volcánica: visión computacional para la detección de emisiones de ceniza del volcán Popocatepetl

**Eric B. Tellez Ugalde**

Universidad Nacional Autónoma de México, Instituto de Geofísica, Ciudad de México, México.  
ORCID: 0009-0002-5654-7518

**Miguel Ángel Vela Rosas**

Universidad Nacional Autónoma de México, Instituto de Geofísica, Ciudad de México, México.  
ORCID: 0000-0003-3621-1769

**Yadir Emmanuel Sánchez Tafolla**

Universidad Nacional Autónoma de México, Instituto de Geofísica, Ciudad de México, México.  
ORCID: 0009-0006-8054-4391

**Angélica Fernández Pineda**

Universidad Nacional Autónoma de México, Instituto de Geofísica, Ciudad de México, México.  
ORCID: 0009-0000-1616-1464

**Fátima Teresa García Fierros**

Universidad Nacional Autónoma de México, Instituto de Geofísica, Ciudad de México, México.  
ORCID: 0009-0007-8236-9651

**Humberto Jesús Abaffy Castillo**

Universidad de Bergen, Bergen, Noruega.  
ORCID: 0009-0005-5237-974X

**Hugo Delgado Granados**

Universidad Nacional Autónoma de México, Instituto de Geofísica, Ciudad de México, México.  
ORCID: 0000-0001-5263-7968

Recepción: 12 de noviembre de 2024.

Aceptación: 21 de agosto de 2025.

Septiembre 2025 • número de revista 13 • DOI: 10.22201/dgtic.26832968e.2025.13.55

## Optimización de la vigilancia volcánica: visión computacional para la detección de emisiones de ceniza del volcán Popocatepetl

---

### Resumen

El monitoreo del volcán Popocatepetl busca proporcionar herramientas a las autoridades de protección civil para mitigar los riesgos asociados con la actividad volcánica. En México, el Cenapred y la UNAM llevan a cabo esta vigilancia para reducir los riesgos asociados a las emisiones de ceniza, los cuales pueden afectar la salud de la población y causar interrupciones en los servicios de aeronavegación. El monitoreo visual se realiza mediante cámaras web que transmiten al Cenapred, donde se requiere la presencia de personal las 24 horas del día, por lo que un sistema automatizado sería crucial para optimizar los recursos humanos disponibles. Para abordar esta necesidad, se propone un modelo de visión computacional que reconozca las emisiones de ceniza. Los datos para el entrenamiento del modelo se obtuvieron gracias a la red de cámaras operadas por Cenapred y Webcams de México, y la programación del mismo se realizó con TensorFlow en equipos Atlas de Huawei. Se obtuvo una precisión de 98% en el modelo de clasificación de imágenes en dos categorías: con actividad y sin actividad; y 84% con tres categorías: sin actividad, gases y cenizas.

**Palabras Clave:** inteligencia artificial, visión computacional, clasificación de imágenes, detección de objetos, monitoreo volcánico, emisión de ceniza, Popocatepetl.

## Optimizing volcanic monitoring: computer vision for detecting ash emissions from Popocatepetl volcano

### Abstract

*Monitoring the Popocatepetl volcano aims to provide civil protection authorities with tools to mitigate the risks associated with volcanic activity. In Mexico, the National Center for Disaster Prevention (Cenapred) and the National Autonomous University of Mexico (UNAM) conduct surveillance of the volcano to reduce the risks associated with ash emissions, which can affect public health and disrupt air navigation services. Visual monitoring is carried out using webcams transmitting to Cenapred, where personnel is needed 24/7, making an automated system crucial to optimize available human resources. To address this need, a computer vision model is proposed to recognize ash emissions. The dataset for training the model was obtained from the cameras operated by Cenapred and Webcams de Mexico, and the coding was made with TensorFlow on a Huawei's Atlas server. A precision of 98% was obtained in an image classification model with two categories: activity and no activity; and 84% in three categories: no activity, gas and ashes.*

**Keywords:** *artificial intelligence, computer vision, image classification, object detection, volcanic monitoring, ash emission, Popocatepetl.*

### Introducción

El volcán Popocatepetl es uno de los volcanes más monitoreados del mundo debido al riesgo constante que representa para más de 25 millones de personas que habitan en un radio de 100 km alrededor de su cráter, incluyendo una extensa área urbana donde convergen poblaciones de los estados de Puebla, Morelos, Tlaxcala y la Ciudad de México [1]. Los peligros asociados a la segunda montaña más alta de México incluyen flujos piroclásticos, avalanchas y colapsos del edificio volcánico, así como la emisión de nubes de ceniza capaces de cubrir amplias zonas del centro del país [2]; fenómenos que, en suma, tienen el potencial de causar graves efectos como pérdida de vidas humanas, destrucción de infraestructuras, daños a la agricultura y la ganadería, así como desplazamiento de población.

La única manera de percibir y evaluar la actividad de un volcán es mediante una vigilancia sistemática, usando diversos métodos visuales e instrumentales. Si se aplican en las fases previas a una erupción, se pueden detectar cambios cualitativos y cuantitativos en la actividad, lo que podría permitir predecir una erupción inminente a corto plazo. Por lo tanto, es crucial implementar sistemas de vigilancia y monitoreo en todos los volcanes potencialmente peligrosos para establecer programas preventivos ante posibles desastres [3]. En conjunto con el monitoreo del volcán, la comunicación oportuna a la sociedad es esencial para mitigar riesgos; ambas permiten alertar a la población y facilitar la implementación de medidas de prevención y respuesta ante posibles amenazas volcánicas [4].

Estudios han demostrado que la exposición a ceniza volcánica puede provocar afecciones respiratorias y oculares [5]. Las erupciones volcánicas explosivas tienden a liberar grandes cantidades de ceniza y gases corrosivos, que ascienden hasta la troposfera superior y la estratosfera baja, zonas que coinciden con las altitudes utilizadas por las aeronaves comerciales. En las últimas décadas, más de 60 aviones han sufrido daños debido a nubes errantes de ceniza volcánica que han afectado tanto las rutas aéreas como las infraestructuras aeroportuarias [3]. Las emisiones de ceniza suponen un riesgo considerable para la navegación aérea, principalmente debido a los daños que estas partículas pueden causar en las turbinas de los aviones. Un caso emblemático es el incidente del vuelo 9 de British Airways en 1982, en el que la aeronave perdió potencia en sus cuatro motores tras atravesar una nube de ceniza del monte Galunggung en Indonesia. La tripulación logró recuperar el control de la aeronave después de descender varios miles de pies [6].

En la actualidad, muchos observatorios volcánicos emiten alertas dirigidas al tráfico aéreo. Una de las herramientas más empleadas en el monitoreo es la observación visual mediante cámaras en tiempo real, complementada por otros métodos como el monitoreo sísmico, geoquímico y geodésico. A través de estas cámaras, un equipo de analistas vigila la actividad volcánica para emitir avisos a las autoridades de protección civil y aeronáutica. Sin embargo, este enfoque requiere la presencia de personal especializado las 24 horas del día, los 365 días del año, lo que representa un desafío logístico y financiero significativo, especialmente en regiones con varios volcanes activos que demandan atención simultánea y en las que los observatorios cuentan con recursos humanos limitados. La dependencia en el monitoreo visual y manual no siempre es viable, lo que resalta la necesidad de

implementar herramientas automatizadas que complementen o incluso reemplacen algunas de estas tareas.

## Aplicación de la inteligencia artificial al monitoreo volcánico

En los últimos años, las tecnologías de visión computacional han avanzado significativamente y se han extendido a diversas áreas del conocimiento, desde la medicina y la detección de huesos mediante las imágenes de rayos X [7] hasta la clasificación de especies nativas de frijol [8]. Su aplicación en el monitoreo volcánico puede resultar en soluciones de software valiosas para observatorios con alta actividad eruptiva simultánea, lo cual es útil en todo el mundo. Uno de los primeros ejemplos de monitoreo de volcanes con visión computacional se encuentra en el artículo de Bertuccio *et al.* [9] de 1999, una publicación en la que se hace uso de una arquitectura de red neuronal celular (CNN) para procesar imágenes del volcán Etna y del Stromboli en tiempo real. Ahí, se expone una nueva metodología (para aquel momento) con la que se pueden extraer parámetros físicos de los eventos, aunque este sistema se limita a actividad no paroxísmica.

Existen casos de éxito recientes del uso de este método en el área de monitoreo volcánico visual alrededor del mundo, por ejemplo [10], [11], [12], [13] y [14].

El presente artículo tiene como objetivo explicar el desarrollo de un sistema automatizado de detección de emisiones de ceniza en el volcán Popocatepetl bajo diversas condiciones atmosféricas y de actividad volcánica. Este sistema representa una innovación en el monitoreo del Popocatepetl; sin embargo, se limita a la actividad suficientemente visible por las cámaras, por lo que está diseñado como una herramienta complementaria para instituciones de protección civil, como el Cenapred, y de seguridad aérea, como los Servicios a la Navegación en el Espacio Aéreo Mexicano (SENEAM) mediante su Centro de Análisis y Pronósticos (CAPMA).

## Datos

Los datos (imágenes) para alimentar los modelos son generados por las cámaras web de la red de monitoreo del Cenapred, disponibles para el público general en las plataformas de Webcams de México [15] y del propio Cenapred [16]. Estas imágenes fueron recopiladas, en su mayoría, a través de Observación Volcánica [17], un proyecto que, desde 2020, se dedica a generar boletines en redes sociales en donde se resume la actividad diaria del volcán. De forma adicional, se hicieron solicitudes de imágenes al Cenapred para incluir eventos que no se recopilaron en Observación Volcánica.

Las imágenes provienen principalmente de las cámaras de Tlamacas, Altzomoni, Tianguismanalco y San Nicolás de los Ranchos, las cuales se encuentran aproximadamente a 5, 11, 19 y 15 km del cráter, respectivamente [16]. Todas las cámaras son de alta definición y cuentan además con visión nocturna. También se utilizaron imágenes de cámaras antiguas las cuales fueron obtenidas de los reportes diarios que Cenapred publica desde los años 90. Algunas de las cámaras que se incluyeron fueron Canario, Juncos, Chipiquixtle y las cámaras antiguas de Tlamacas y Altzomoni. Asimismo, se incluyeron cámaras propias de Webcams de México, las cuales son Puebla y Amecameca.

El primer set de datos consistió de aproximadamente 2500 imágenes fijas. Se incluyeron imágenes de diferentes fases de actividad volcánica, que van desde la calma total hasta emisiones fumarólicas y explosiones con emisiones de ceniza con diferentes rangos intensidad. Para clasificar estas imágenes, se crearon diferentes categorías acordes al tipo de actividad del volcán. En un primer momento, se propusieron 18 categorías de acuerdo a todos los tipos de actividad volcánica combinada con las condiciones meteorológicas y hora del día de las observaciones, pero rápidamente fue obvio que esto sería innecesario y demasiado complicado. Debido a que el principal interés del monitoreo visual está en la ceniza, se decidió enfocar el modelo en ese fenómeno, por lo que una clasificación menos ambiciosa se propuso de la siguiente forma:

- Sin actividad: Volcán en calma.
- Emisiones de gas: Emisiones de gases volcánicos, principalmente vapor de agua, sin presencia significativa de ceniza.
- Emisiones de ceniza y gas: Aquellas emisiones con contenido de ceniza. Todas las emisiones de ceniza están acompañadas de gases volcánicos.

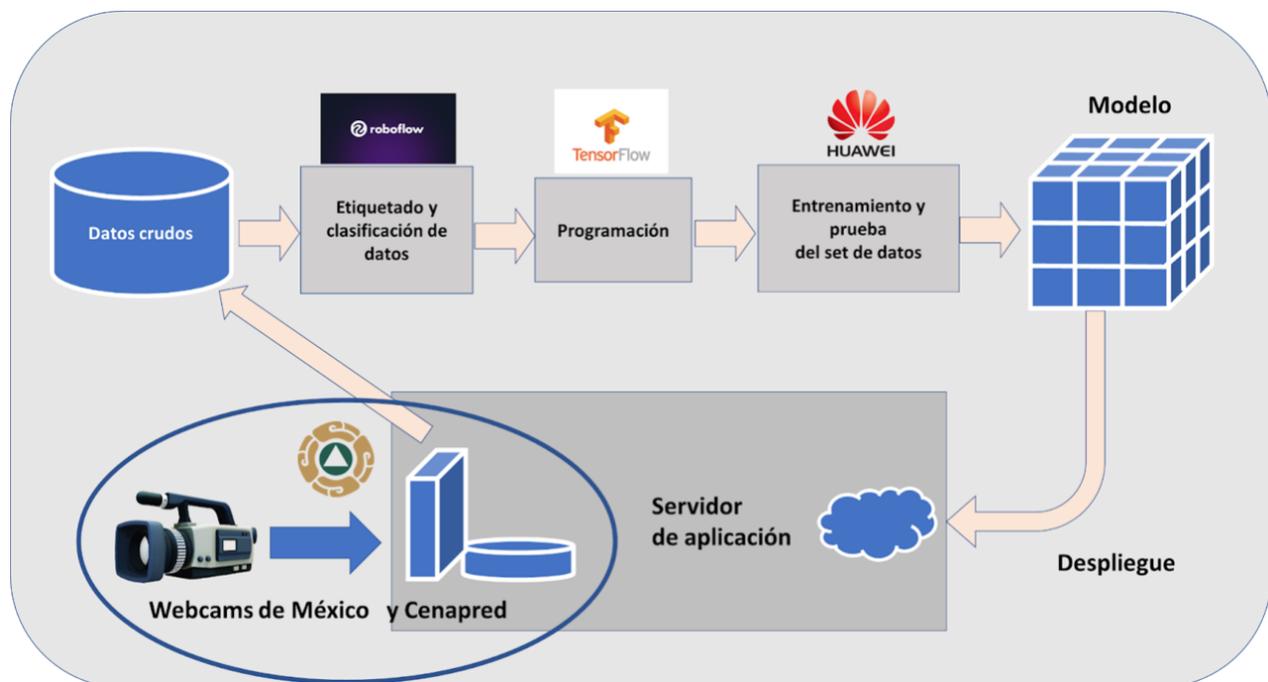
Las tres categorías abarcan imágenes tomadas tanto en condiciones de buena visibilidad como en condiciones de visibilidad limitada. Esta clasificación en tres fue una meta a mediano plazo del proyecto, puesto que se decidió realizar primero un código piloto de solo dos categorías, volcán con actividad y volcán sin actividad. Este código más simple también permitiría ver si habría la necesidad de recopilar imágenes diferentes, que, por sus características especiales, permitieran mejorar el modelo en caso necesario.

## Clasificación de imágenes

La clasificación de imágenes es una técnica de visión computacional en la cual se le aplica una etiqueta alfabética (palabra o palabras) a una imagen. Es decir, el dato de entrada es una imagen y el de salida es una cadena de texto [18]. Con las imágenes clasificadas, se comenzó la escritura de un programa piloto con clasificación de imágenes en Jupyter Notebook [19] que nos diera información exploratoria acerca del comportamiento general del aprendizaje con este set de datos en particular. Para realizar esta tarea, se empleó la técnica de transferencia de aprendizaje que consiste en tomar una red neuronal ya entrenada y usarla para resolver un problema similar [20]; en este caso se usó MobileNetv2 [21]. El código base, a partir del cual se trabajó, se tomó de [22]. Al principio, se obtuvo un sobreajuste, es decir, las imágenes no eran suficientemente variadas y el modelo era incapaz de hacer predicciones o conclusiones precisas a partir de ningún otro dato que fuera de los de entrenamiento [23]. Para corregir esto, se aumentó la base de datos con más imágenes reales pero también se multiplicó ésta de manera artificial con la técnica de aumento de datos [24]. Este código escrito en Jupyter Notebook se ejecutó en Google Colab; sin embargo, pronto se cambió a un formato no interactivo de Python para poder hacer la ejecución directamente en el servidor Atlas de Huawei [25].

Con el modelo de clasificación de imágenes en dos categorías suficientemente comprendido, se procedió a ampliarlo a tres categorías, lo cual redujo significativamente su precisión. Para mejorarla, se aumentó nuevamente de manera artificial el conjunto de datos y se probaron ajustes en los optimizadores y en otros dos parámetros. Sin embargo, estos cambios no ofrecieron los resultados esperados, por lo que se optó por modificar el código en TensorFlow mediante la actualización de la instrucción de importación de imágenes a una versión mejorada y de ejecución más rápida [26].

A la par que se trataba de mejorar la precisión del programa para tres categorías, se intentó poner el modelo de clasificación de imágenes de dos categorías en un servidor de inferencia para la clasificación de imágenes y de video en tiempo real. A este paso del proyecto se le conoce como despliegue. En la Figura 1, se puede observar el flujo de trabajo completo, desde la obtención de las imágenes, pasando por la clasificación, la programación y el entrenamiento. En este proyecto, se pudo llegar hasta la obtención del modelo y, actualmente, se trabaja para realizar el despliegue.

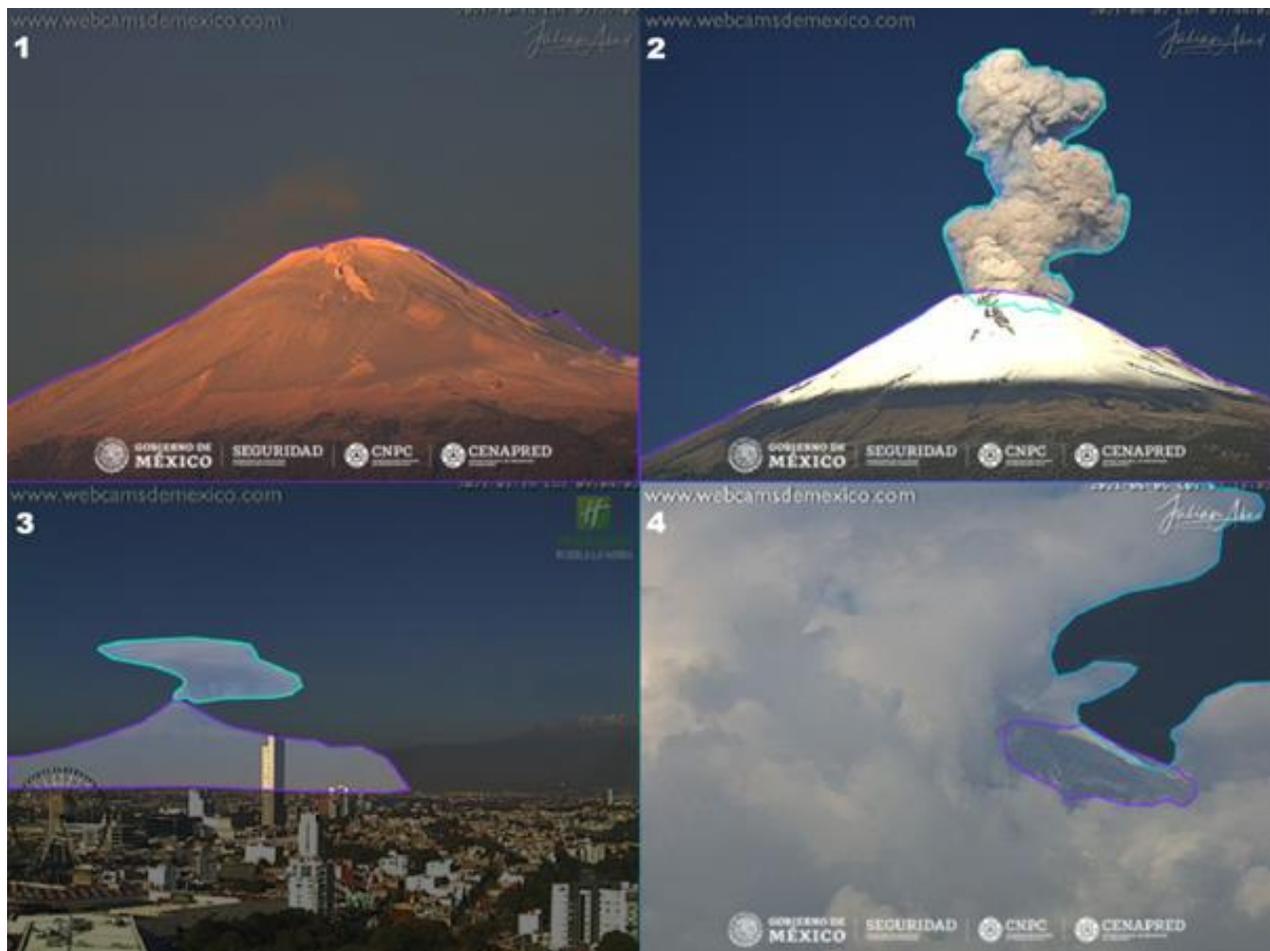


**Fig. 1.** Flujo de trabajo de este proyecto, desde la obtención de las imágenes hasta el despliegue en tiempo real.

## Detección de objetos

Mientras se avanzaba en el despliegue del modelo de clasificación de imágenes en tiempo real, a la par, se utilizó el mismo set de imágenes para entrenar un modelo de detección de objetos. La detección de objetos es una técnica de visión computacional que tiene como

objetivo localizar objetos en imágenes digitales [27]. Para poder usar este modelo, primero es necesario etiquetar las imágenes de tal forma que, en cada imagen, se resalte el objeto u objetos de interés. En la Figura 2, podemos ver una imagen etiquetada. En color verde, tenemos la nube de ceniza volcánica, en color morado está el volcán y en color azul están las nubes meteorológicas. El etiquetado de 1523 imágenes se realizó en la página de Roboflow [28].



**Fig. 2.** Ejemplo de imágenes etiquetadas con Roboflow. Se observa la categoría del volcán en calma (1), emisión en condiciones despejadas (2 y 3) y condiciones adversas de nubosidad (4).

Para entrenar este modelo de detección de objetos, se utilizó You Only Look Once versión 8 (YOLOv8) [29]. YOLO es un algoritmo de detección de objetos que conceptualiza el problema como una regresión y busca reducir la complejidad de los procesos al realizarlos a la vez. Funciona de la siguiente manera: una única red convolucional predice simultáneamente varios recuadros delimitadores y las probabilidades de clase de esos recuadros. YOLO se entrena en imágenes completas y optimiza directamente el rendimiento de detección [30]. El entrenamiento se realizó en el servidor de Huawei, pero también se usó la plataforma de Roboflow para entrenar un modelo. Debido a que se usó la versión gratuita, el entrenamiento se pudo hacer únicamente con 400 imágenes. En la Figura 3, se puede apreciar un fotograma de la salida de video con el entrenamiento en Roboflow.



**Fig. 3.** Fotograma de un video de salida con el modelo de Roboflow. En esta imagen, se muestran tres clases y la probabilidad de que la clase reconocida sea correcta, es decir, la computadora tiene una certeza de 89% de reconocer la nube de ceniza de la imagen.

## Resultados

En el caso de clasificación de imágenes de dos categorías (sin actividad, con actividad), se obtuvo una precisión de 98% con los parámetros “activación sigmoïdal”, “optimizador Adam” y con “pérdida entropía cruzada binaria”; para este caso, se tienen dos programas: uno de ejecución lenta y otro de ejecución rápida. En el caso de clasificación de imágenes con tres categorías (sin actividad, gases, ceniza), con activación sigmoïdal, optimizador Adamax y con pérdida entropía cruzada categórica, se obtuvo una precisión de 84%; para este caso, también se tienen dos programas: uno de ejecución lenta y otro de ejecución rápida. En el caso de detección de objetos con YOLOv8, la precisión estuvo apenas por encima de 70%; sin embargo, aunque se tienen 1517 imágenes etiquetadas, este modelo se entrenó con apenas 400 en la plataforma Roboflow de forma gratuita, pero se espera poder hacer el entrenamiento con el conjunto ampliado de datos dentro del servidor de Huawei. Todos estos programas se encuentran en [19].

Aunque la precisión con dos categorías es del 98%, es menor (aunque no se cuantificó) cuando se usa un video en tiempo real, puesto que no se tomó en cuenta el caso donde no hay visibilidad debido a las nubes. En éste, el programa de clasificación de imágenes falla en reconocer si hay actividad o no, dependiendo de qué tan grande sea la nube que cruce por la cámara. Esta variable debe ser tomada en cuenta como una categoría por sí misma al momento de entrenar el modelo para que el cálculo de la imagen arroje que no hay visibilidad. No se ha cuantificado este error en el algoritmo de clasificación de imágenes, pero, en el caso de detección de objetos, sí se ha hecho el etiquetado de nubes que tapan parcial o totalmente la cámara para incorporar esta variable en ese algoritmo.

## Conclusiones

La visión computacional ofrece un recurso efectivo para el monitoreo de la actividad volcánica en zonas de difícil acceso o con recursos humanos limitados, además de ser una solución relativamente económica. Actualmente, se ha completado la primera etapa de esta investigación, en la cual se desarrollaron modelos de clasificación con un nivel de precisión adecuado y modelos de detección aún con posibilidades de mejora. De hecho, el error producido intrínsecamente por los modelos se puede reducir de dos formas: mediante un

promedio espacial y uno temporal. En primer lugar, cuando hay nubosidad y poca visibilidad se puede hacer uso de múltiples cámaras: si una cámara no tiene línea de vista, se puede confirmar con otra e incluso se puede hacer un promedio pesado que descarte el resultado arrojado por una o varias de las cámaras (esto ya se está implementando). Además, si la imagen es muy complicada y el modelo tiene una precisión baja, de tal forma que no sabe si es una nube volcánica, se puede esperar a que la nube cambie de forma y volver a cuantificar la imagen unas cuantas veces más algunos segundos más tarde; de esta manera, si el resultado es siempre positivo, se tiene más certeza de lo que está viendo el modelo. En una segunda etapa, se proyecta implementar estos modelos en un entorno de monitoreo en tiempo real accesible al público a través de una plataforma web. Aunque esta aplicación presenta limitaciones inherentes al tipo de fenómeno monitoreado, sus alcances están claramente definidos y representa una herramienta complementaria valiosa para los métodos de monitoreo volcánico existentes.

## Agradecimientos

Los autores agradecen al Grupo Especial de Innovación de la DGTIC Alianza-Huawei por la ayuda prestada para la elaboración de este trabajo.

## Referencias

- [1] R. Espinasa Pereña, *Historia de la actividad del volcán Popocatepetl, 17 años de erupciones*, 1ra ed. México, D.F.: Secretaría de Gobernación, Cenapred, 2012. [En línea]. Disponible: <https://www.cenapred.unam.mx/es/Publicaciones/archivos/225-HISTORIADELAACTIVIDADELVOLCNPOPOCATPETL-17AOSDEERUPCIONES.PDF>
- [2] A. L. Martin Del Pozzo, et al., "Memoria técnica del mapa de peligros del volcán Popocatepetl, 1ra ed." en *Monografías del Instituto de Geofísica*, no. 22. México, D.F.: Instituto de Geofísica, UNAM, 2017. [En línea]. Disponible: <https://www.geofisica.unam.mx/assets/monografias22.pdf>

- [3] R. Quass Weppen y S. M. Alcocer Martínez de Castro (coords.), *Las cenizas volcánicas del Popocatepetl y sus efectos para la aeronavegación e infraestructura aeroportuaria*, 1era ed. México, D.F.: Secretaría de Gobernación, Cenapred, 2001. [En línea]. Disponible: <https://www.cenapred.unam.mx/es/Publicaciones/archivos/111-LASCENZASVOLCNICASDELPOPOCATPETLYSUSEFECTOSPARALAAERONAVEGACINEINFRAESTRUCTURAAEROPORTUARIA.PDF>
- [4] H. U. Schmincke, *Volcanism*. Berlin Heidelberg: Springer Verlag, 2004.
- [5] M. Rojas Ramos, H. Noa Ortega, J. Pérez Neria, y J. Villalba Caloca, "Efecto sobre la función pulmonar en personas expuestas a cenizas del volcán Popocatepetl: diciembre 1994-enero 1995", *Rev. Inst. Nac. Enfermedades Respiratorias*, pp. 112–8, 1995.
- [6] R. J. Clarkson, E. J. Majewicz, y P. Mack, "A re-evaluation of the 2010 quantitative understanding of the effects volcanic ash has on gas turbine engines", *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, vol. 230, núm. 12, pp. 2274–2291, oct. 2016, doi: 10.1177/0954410015623372.
- [7] Z. Fanti Gutiérrez, A. Gastelum-Strozzi, E. Hazan-Lasri, y F. Arámbula Cosío, "Aprendizaje computacional aplicado a la detección de huesos, en cirugía ortopédica asistida por computadora", *TIES, Revista de Tecnología e Innovación en Educación Superior*, núm. 4, pp. 1–11, may 2021, doi: 10.22201/dgtic.26832968e.2021.4.5.
- [8] J. L. Morales Reyes, H. G. Acosta Mesa, E. N. Aquino Bolaños, M. D. S. Herrera Meza, N. Cruz Ramírez, y J. L. Chávez-Servia, "Clasificación de poblaciones nativas de frijol utilizando visión artificial", *TIES, Revista de Tecnología e Innovación en Educación Superior*, núm. 4, pp. 1–11, oct. 2021, doi: 10.22201/dgtic.26832968e.2021.4.2.
- [9] L. Bertucco, M. Coltelli, G. Nunnari, y L. Occhipinti, "Cellular neural networks for real-time monitoring of volcanic activity", *Computers & Geosciences*, vol. 25, núm. 2, pp. 101–117, mar. 1999, doi: 10.1016/S0098-3004(98)00114-9.
- [10] C. D. Martínez, R. C. Uribe, C. Ospina, y R. Vargas Cañas, "Clasificación de imágenes de cámaras de vigilancia volcánica usando visión artificial", *RISTI: Revista Ibérica de Sistemas e Tecnologías de Informação*, núm. Extra 38, pp. 185–198, 2020.

- [11] R. Centeno, V. Gómez-Salcedo, I. Lazarte, J. Vilca-Nina, S. Osoro, y E. Mayhua-Lopez, "Near-real-time multiparametric seismic and visual monitoring of explosive activity at Sabancaya volcano, Peru", *Journal of Volcanology and Geothermal Research*, vol. 451, p. 108097, jul. 2024, doi: 10.1016/j.jvolgeores.2024.108097.
- [12] J. F. Guerrero Tello, M. Coltelli, M. Marsella, A. Celauro, y J. A. Palenzuela Baena, "Convolutional Neural Network Algorithms for Semantic Segmentation of Volcanic Ash Plumes Using Visible Camera Imagery", *Remote Sensing*, vol. 14, núm. 18, Art. núm. 18, ene. 2022, doi: 10.3390/rs14184477.
- [13] S. Korolev, A. Sorokin, I. Urmanov, A. Kamaev, y O. Girina, "Classification of Video Observation Data for Volcanic Activity Monitoring Using Computer Vision and Modern Neural Networks (on Klyuchevskoy Volcano Example)", *Remote Sensing*, vol. 13, núm. 23, Art. núm. 23, ene. 2021, doi: 10.3390/rs13234747.
- [14] T. C. Wilkes, T. D. Pering, y A. J. S. McGonigle, "Semantic segmentation of explosive volcanic plumes through deep learning", *Computers & Geosciences*, vol. 168, p. 105216, nov. 2022, doi: 10.1016/j.cageo.2022.105216.
- [15] Volcanes, webcams en vivo, Webcams de México. Consultado: el 6 de noviembre de 2024. [En línea]. Disponible: <https://webcamsdemexico.com/webcams/volcanes/>
- [16] *Reporte del monitoreo del Cenapred al volcán Popocatepetl*, Cenapred. Consultado: el 6 de noviembre de 2024. [En línea]. Disponible: <https://www.cenapred.unam.mx/reportesVolcanesMX/>
- [17] Observación Volcánica. Consultado: el 6 de noviembre de 2024. [En línea]. Disponible: <https://www.facebook.com/ObervacionVolcanica>
- [18] S. Wang y Z. Su, "Metamorphic Testing for Object Detection Systems", el 19 de diciembre de 2019, *arXiv*: arXiv:1912.12162. doi: 10.48550/arXiv.1912.12162.
- [19] E. Tellez, *erictellez/Popocatepetl\_CV*. (el 20 de septiembre de 2024). Jupyter Notebook. Consultado: el 6 de noviembre de 2024. [En línea]. Disponible: [https://github.com/erictellez/Popocatepetl\\_CV](https://github.com/erictellez/Popocatepetl_CV)

- [20] “¿Qué es el aprendizaje por transferencia?”, IBM, Consultado: el 7 de noviembre de 2024. [En línea]. Disponible: <https://www.ibm.com/es-es/topics/transfer-learning>
- [21] “Keras documentation: MobileNet, MobileNetV2, and MobileNetV3”, Keras Team, Consultado: el 7 de noviembre de 2024. [En línea]. Disponible: <https://keras.io/api/applications/mobilenet/>
- [22] Ringa Tech, *¿Pocos datos de entrenamiento? Prueba esta técnica*, (el 30 de noviembre de 2021). Consultado: el 7 de noviembre de 2024. [Video en línea]. Disponible: [https://www.youtube.com/watch?v=9Dur\\_oUMGG8](https://www.youtube.com/watch?v=9Dur_oUMGG8)
- [23] “¿Qué es el sobreajuste?”, IBM. Consultado: el 6 de noviembre de 2024. [En línea]. Disponible: <https://www.ibm.com/mx-es/topics/overfitting>
- [24] “Aumento de datos | TensorFlow Core”, TensorFlow. Consultado: el 8 de noviembre de 2024. [En línea]. Disponible: [https://www.tensorflow.org/tutorials/images/data\\_augmentation?hl=es-419](https://www.tensorflow.org/tutorials/images/data_augmentation?hl=es-419)
- [25] “Recursos – Alianza”. UNAM. Consultado: el 30 de junio de 2025. [En línea]. Disponible: <https://alianza.unam.mx/recursos/>
- [26] [26] “tf.keras.preprocessing.image\_dataset\_from\_directory | TensorFlow v2.16.1”, TensorFlow. Consultado: el 8 de noviembre de 2024. [En línea]. Disponible: [https://www.tensorflow.org/api\\_docs/python/tf/keras/preprocessing/image\\_dataset\\_from\\_directory](https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/image_dataset_from_directory)
- [27] “¿Qué es la detección de objetos?”, IBM. Consultado: el 9 de noviembre de 2024. [En línea]. Disponible: <https://www.ibm.com/mx-es/topics/object-detection>
- [28] “IAPOPO Dataset > Overview”, Roboflow. Consultado: el 9 de noviembre de 2024. [En línea]. Disponible: <https://universe.roboflow.com/ia-popo/ia-popo>
- [29] J. Torres, “YOLOv8 Documentation: A Deep Dive into the Documentation - YOLOv8”. Consultado: el 4 de octubre de 2024. [En línea]. Disponible: <https://yolov8.org/yolov8-documentation/>

- [30] J. Redmon, S. Divvala, R. Girshick, y A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection", el 9 de mayo de 2016, *arXiv*: arXiv:1506.02640. doi: 10.48550/arXiv.1506.02640.



# 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):

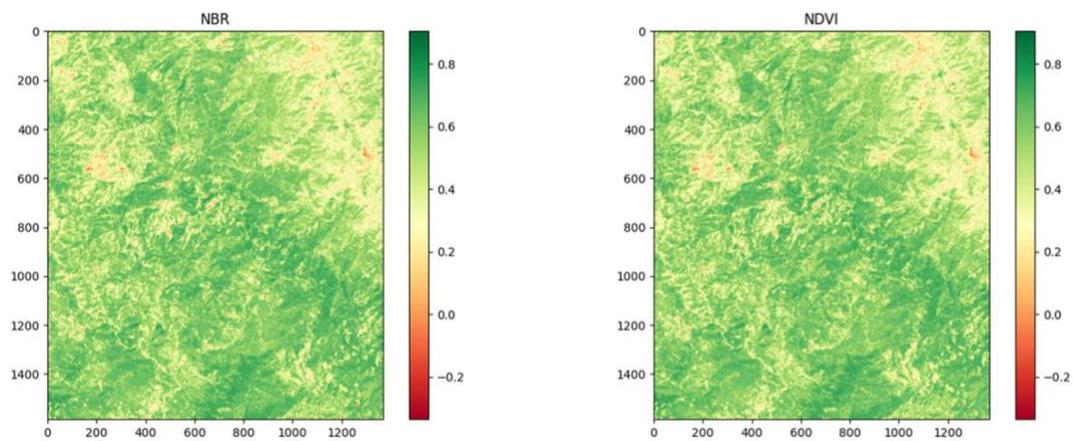
NDVI: Para evaluar la salud de la vegetación:

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

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)

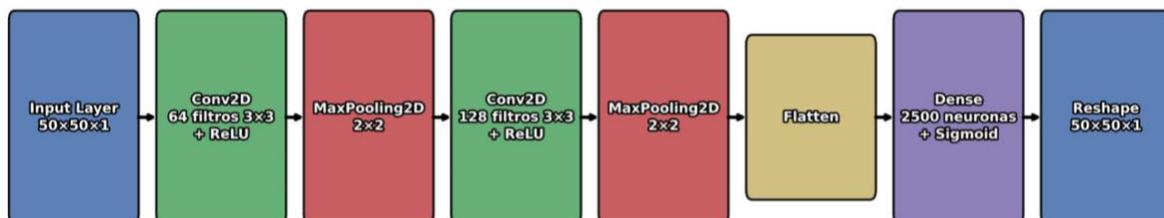


Fig. 2. Diagrama de la arquitectura CNN implementada (red neuronal)

### 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](#).

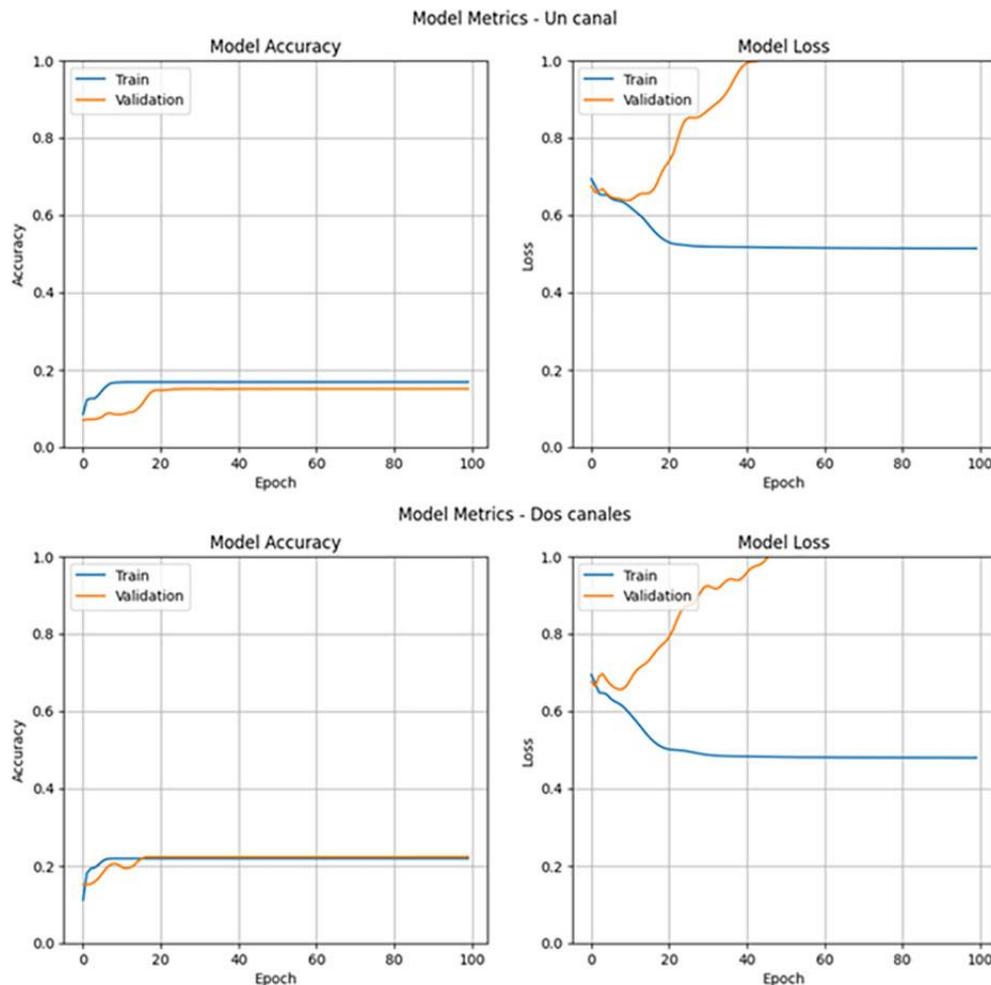
## 2.2 Resultados del Modelo

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*).

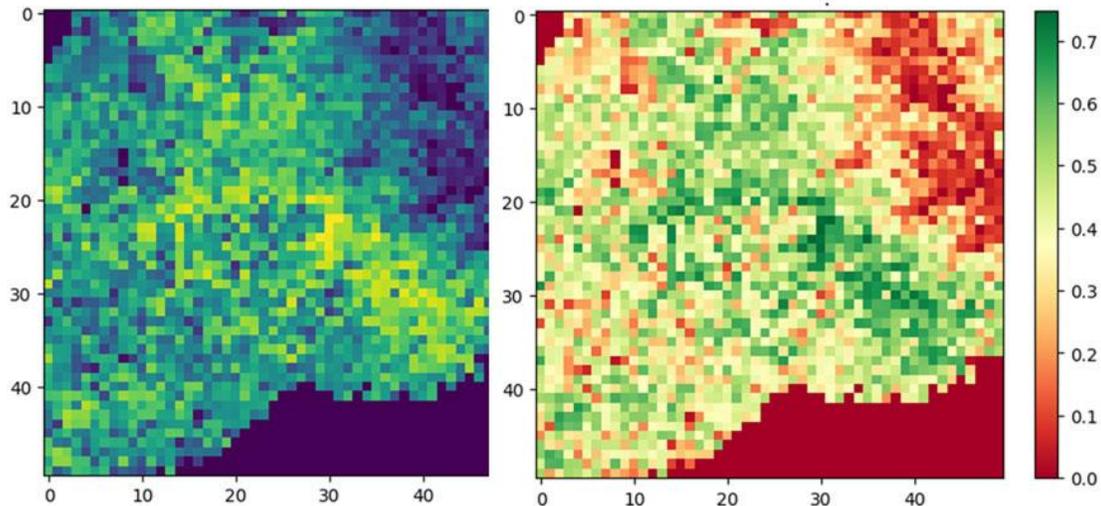


**Fig. 3.** Curvas de entrenamiento y validación para el modelo de un canal (arriba) y dos canales (abajo) (métricas).

### 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.



# Experiencia en un proyecto de IA educativo para la detección de dificultades matemáticas, basado en emociones y desempeño estudiantil

**Nelly Rigaud Téllez**

Universidad Nacional Autónoma de México, Facultad de Estudios Superiores Aragón,  
Estado de México, México.

ORCID: 0000-0002-2719-5464

**Arturo Rodríguez García**

Universidad Nacional Autónoma de México, Facultad de Estudios Superiores Aragón,  
Estado de México, México.

ORCID: 0000-0001-8523-3214

**Judith Mayte Flores Pérez**

Universidad Nacional Autónoma de México, Facultad de Estudios Superiores Cuautitlán,  
Estado de México, México.

ORCID: 0009-0003-3610-9737

**Viviana Flores Herrera**

Universidad Autónoma de Baja California Sur, La Paz, Baja California Sur, México.

ORCID: 0009-0008-1705-9232

**Marcelo Daniel Pérez Maceda**

Consultor Psicológico, Estado de México, México.

ORCID: 0009-0003-9068-7263

**Roberto Blanco Bautista**

Universidad Nacional Autónoma de México, Facultad de Estudios Superiores Aragón,  
Estado de México, México.

ORCID: 0009-0003-2875-6678

Recepción: 4 de noviembre de 2024.

Aceptación: 22 de agosto de 2025.

Septiembre 2025 • número de revista 13 • DOI: 10.22201/dgtic.26832968e.2025.13.50

## Experiencia en un proyecto de IA educativo para la detección de dificultades matemáticas, basado en emociones y desempeño estudiantil

---

### Resumen

El éxito de un proyecto de IA educativa depende, en gran medida, de su gestión, colaboradores y talento, así como de la infraestructura tecnológica disponible. Realizar un proyecto de IA es una tarea compleja y multifacética que implica varios factores, en la que los riesgos y retos persisten a lo largo de todo su ciclo de vida. El objetivo de este artículo es compartir la experiencia del desarrollo de una solución tecnológica orientada a detectar dificultades de comprensión matemática en estudiantes de carreras de ingeniería y tecnología, mediante el uso de técnicas de IA aplicadas al análisis del desempeño, emociones y expresiones faciales. Para ello, se usaron lineamientos de la norma ISO 21050:2020 en dirección de proyectos con un esquema de trabajo basado en SCRUM. En el procesamiento de imágenes de estudiantes, se usaron técnicas de visión por computadora para el análisis emocional y, posteriormente, redes neuronales convolucionales, asociando patrones emocionales y desempeño, durante momentos de aprendizaje matemático. En cada etapa se describen desafíos, aprendizajes prácticos y experiencias derivadas de su implementación, lo que permite extraer una serie de lecciones aprendidas a partir de esta experiencia.

**Palabras Clave:** proyecto de IA educativa, clases virtuales de matemáticas, ciclo de vida de proyectos de IA, detección de emociones, experiencia interdisciplinaria, desafíos en gestión de proyectos, Alianza UNAM-HUAWEI.

## ***Experience in an educational AI project for detecting mathematical learning difficulties, based on emotions and student performance***

---

### ***Abstract***

*The success of an educational AI project depends largely on its management, collaborators, and talent, as well as the available technological infrastructure. Developing an AI project is a complex and multifaceted task, involving multiple factors, with persistent risks and challenges throughout its lifecycle. The objective is to share the experience of developing a technological solution aimed at detecting mathematical comprehension difficulties among engineering and technology students, through the use of AI techniques applied to performance, emotions, and facial expressions analysis. To fulfill this objective, we use ISO21050:2020 guidelines for project management with SCRUM-based work framework. In processing student images, we use computer vision techniques for emotional analysis, followed by convolutional neural networks, associating emotional patterns and performance during moments of mathematical learning. Each phase describes challenges, practical learning, and experiences derived from its implementation. These reflections provide a set of lessons learned for future initiatives in similar contexts.*

**Keywords:** *educational AI project, virtual mathematics classes, AI project lifecycle, emotion detection, interdisciplinary experience, project management challenges, UNAM-HUAWEI Alliance.*

### **Introducción**

La ejecución exitosa de un proyecto de inteligencia artificial (IA) educativa representa un conjunto de desafíos únicos y oportunidades que van más allá de la ideación tecnológica; requiere de una combinación de estrategia, talento y experiencia técnica, recursos y gestión organizada.

Si bien la literatura especializada reporta aplicaciones de IA en educación, persiste una escasez de estudios que documenten la experiencia de implementación tecnológica, donde colaboran universidad-industria [1]. Particularmente, trabajos existentes sobre iniciativas de IA aplicada a problemáticas educativas concretas no siempre ofrecen una cobertura completa y presentan limitaciones en cuanto al nivel de detalle organizativo, personal y responsabilidades administrativas [2].

El presente trabajo se enmarca en el proyecto “La inteligencia artificial como apoyo a las estrategias de aprendizaje de las matemáticas”, una iniciativa impulsada por la Alianza UNAM- HUAWEI, cuya primera etapa concluyó en enero de 2024. El objetivo de este artículo es compartir la experiencia del desarrollo de una solución tecnológica orientada a detectar dificultades de comprensión matemática en estudiantes de carreras de ingeniería y tecnología, mediante el uso de técnicas de IA aplicadas al análisis del desempeño, emociones y expresiones faciales. Lo anterior busca mejorar el aprendizaje de las matemáticas en las Ciencias Físico-Matemáticas e Ingenierías entre alumnos en situación de vulnerabilidad académica de Geometría Analítica, asignatura clave en la formación universitaria. En ésta, se tienen ubicados los temas de cónicas en un rango de temas difíciles de acuerdo con el examen de diagnóstico de la UNAM [3].

Para la comprensión de esta asignatura, además del dominio técnico, es común el lenguaje corporal tanto en docentes como en estudiantes, quienes, conforme se sumergen en el entorno de aprendizaje, expresan ideas matemáticas y experimentan una amplia gama de emociones que influyen en el desempeño. Sin embargo, en una clase en línea, la comunicación no verbal y la interpretación de emociones se ve limitada, más cuando se tratan de captar, de manera inmediata, reacciones emocionales y expresiones de confusión, frustración o desinterés [4]. En este sentido, la disminución de señales emocionales visibles, como expresiones faciales o gestos de confusión, puede limitar la capacidad del docente para ajustar sus estrategias pedagógicas en tiempo real, afectando el vínculo educativo y el desempeño académico [5][6]. Este vínculo de expresiones faciales, emociones y desempeño matemático fue trabajado como un proyecto de IA, lo cual representó retos diversos durante su implementación.

El artículo está organizado por etapas del proyecto, en donde se describe la experiencia vivida en cada una de ellas. En las conclusiones, se presentan reflexiones y lecciones

prácticas. El valor del presente trabajo se encuentra en presentar la experiencia de innovación educativa centrada en el proceso, la gestión, y los aprendizajes prácticos derivados de su implementación.

## Desarrollo

El proyecto se centró en el diseño, implementación y validación de una solución tecnológica orientada a detectar dificultades en el aprendizaje de las matemáticas en estudiantes de los primeros semestres de licenciaturas de ingeniería y tecnología. Este apartado describe las experiencias a lo largo de sus principales fases: (1) Conformación del equipo interdisciplinario, (2) Planeación y diseño del proyecto, (3) Recolección de datos, (4) Procesamiento, análisis emocional y etiquetado, (5) Entrenamiento y validación del modelo, y (6) Difusión de resultados. A continuación, se destacan los aprendizajes obtenidos y desafíos en cada una de ellas.

### Etapa 1: Conformación del equipo interdisciplinario

En el presente caso, se siguió una organización de colaboración interdisciplinaria y dinámica, cuyos perfiles especializados son en matemáticas, gestión de proyectos, diseño instruccional, metodología de la investigación, computación e IA, tecnología y psicología (Figura 1).



**Fig. 1.** El equipo estratégico e interdisciplinario.

Los especialistas de alto nivel de la Figura 1 están adscritos a las Facultades de Estudios Superiores Aragón y Cuautitlán de la UNAM, Universidad Autónoma de Baja California Sur y un especialista en Psicología, quienes se integraron para abordar problemas multifacéticos, favoreciendo soluciones sistémicas y contextualizadas. La colaboración no sólo funcionó para alcanzar metas, sino también se convirtió en un indicador de identidad y cohesión del equipo de un entorno tecnológico. Los miembros del equipo se organizaron en estructuras especializadas en cinco áreas estratégicas: gestión de proyectos, didáctica de matemáticas, reconocimiento de emociones, red neuronal y soporte técnico operativo.

## **Etapas 2: Planeación y diseño del proyecto.**

Se realizó un asesoramiento con el personal de la Oficina Jurídica de la Facultad de Estudios Superiores Aragón y revisión de avisos de la UNAM para temas legales como el consentimiento informado, privacidad en el uso de grabaciones de estudiantes y aspectos

interinstitucionales (bases de colaboración en investigación, desarrollo y transferencia de conocimiento y tecnología).

Además, se propuso una forma de trabajo, bajo los lineamientos de la norma ISO 21050:2020 en dirección de proyectos [7], donde se establece tener siempre presentes los intereses involucrados alrededor del proyecto. Por ejemplo, para los patrocinadores de la infraestructura tecnológica, Alianza UNAM-HUAWEI, el énfasis puede estar en hacer el seguimiento de necesidades, verificar el cumplimiento de metas, así como confirmar resultados sostenibles y de impacto.

Conforme el marco rector de la ISO, también se identificaron perfiles, aunque fue necesario una intensa dedicación a la construcción de una visión común, una filosofía de autoorganización en el equipo y un compromiso ético con el problema educativo abordado. Adicionalmente, el esquema de trabajo, basado en SCRUM, para realizar actividades, garantizar la calidad, minimizar riesgos y hacer entregas con valor, no sólo se alineó a la norma ISO, también facilitó la gestión de conflictos mediante el conocimiento y la comunicación compartida.

Con respecto la implementación y ejecución, los integrantes del equipo de soporte técnico y operativo proporcionaron asistencia técnica, lo que garantizó la resolución de problemas de compatibilidad, formato y rendimiento mediante procesos de prueba y error, análisis detallado de ejecución y depuración del código; por ejemplo, el desarrollo de scripts para el conteo automático de imágenes etiquetadas o lograr la conexión con el servidor de la Alianza UNAM-HUAWEI.

Un reto encontrado junto con la colaboración Alianza UNAM-HUAWEI fue la planeación operativa bajo condiciones reales, que exigían una curva de aprendizaje acelerada para el manejo de la infraestructura tecnológica. Las herramientas proporcionadas eran altamente especializadas y, en su mayoría, desconocidas para los integrantes del equipo. Como parte del proceso de formación, se ofrecieron diversos cursos teóricos por parte de los patrocinadores, que sentaron las bases para adquirir una comprensión de los requerimientos esenciales y llevar a cabo el proyecto.

Sin embargo, ante la ausencia de material práctico y soporte técnico detallado, el equipo adoptó una estrategia de autoformación. Esto implicó una exploración autónoma y

proactiva del funcionamiento de las herramientas, la resolución colaborativa de problemas técnicos y un constante intercambio de aprendizajes entre los miembros. Esta experiencia puso en evidencia la importancia crítica de la adaptabilidad y una gestión eficaz del conocimiento en el desarrollo de proyectos de innovación tecnológica.

### **Etapa 3: Recolección de datos**

El diseño instruccional se estructuró con base en el enfoque de comprensión y resolución de problemas en tres sesiones: parábola, elipse y una sesión de repaso; éstas incluyeron técnicas didácticas de acuerdo con el enfoque propuesto.

Fue un desafío lograr que los grupos estudiantiles se involucraran de manera efectiva en las actividades de clase. Aunque se proporcionaron instrucciones claras para la grabación de video, especialmente para la toma de imágenes, en la práctica, muchos estudiantes se grabaron en espacios no controlados, como el transporte público, automóviles, habitaciones con mala calidad de luz o con intervenciones inesperadas (familiares o mascotas), lo que afectó tanto la calidad de los datos como la dinámica instruccional, considerando las limitaciones propias del entorno virtual, así como las brechas en su preparación matemática. Adicionalmente, la motivación basada en calificación funciona de manera parcial, lo que refleja la complejidad de lograr un compromiso sostenido en ambientes virtuales con estudiantes vulnerables.

Asimismo, con el fin de evaluar el aprendizaje de los estudiantes se aplicaron dos evaluaciones por tema (al inicio y al final de las clases), diseñadas con Formularios de Google. De igual forma, se aplicó un cuestionario sobre las percepciones emocionales de los estudiantes. Cada una de las clases tuvo una duración máxima de 1 hora con 15 minutos. Con las pruebas de matemáticas, se generó un primer archivo tabular con el registro de estudiantes, los resultados de sus pruebas y la decisión de si tuvo o no dificultades.

#### **Etapa 4: Procesamiento y análisis emocional**

Esta etapa tuvo por objetivo principal crear un archivo tabular de datos etiquetados sobre el estado emocional de los estudiantes participantes. El archivo constituyó otro insumo esencial para el entrenamiento y validación del modelo de IA.

En esta línea, se realizó la descomposición de videos de estudiantes en fotogramas, empleando herramientas de visión por computadora. Las herramientas empleadas para la detección de emociones y su clasificación se realizaron mediante bibliotecas especializadas en Python, específicamente OpenCV y DeepFace [8], que, en conjunto, proporcionaron un flujo robusto de detección facial y clasificación emocional. A partir de cada rostro detectado, se clasificaron expresiones faciales en siete categorías emocionales primarias: felicidad, tristeza, sorpresa, enojo, miedo, disgusto y neutralidad.

Cada fotograma procesado generó vectores de características emocionales en secuencia, es decir, un perfil detallado (representación numérica de emociones en cada imagen), donde la confiabilidad de modelos de aprendizaje automático subyacentes en DeepFace garantizó la validez de los resultados. Esta labor fue especialmente desafiante debido a las condiciones reales en las que los estudiantes tomaron sus clases. Frecuentemente, las condiciones del entorno dificultaron la captura precisa de expresiones, impidiendo la obtención de suficientes fotogramas para construir una base de datos robusta. Asimismo, la prevalencia de expresiones faciales neutras por parte de los estudiantes complicó el reconocimiento preciso de las emociones.

El etiquetado automático se realizó con los resultados de DeepFace, pero fue complementado por un proceso de validación manual por parte de expertos en educación y psicología, quienes, además de revisar la coherencia entre emociones detectadas y el contexto de las sesiones virtuales, también establecieron la regla de decisión sobre si hubo o no dificultad.

Más allá del reconocimiento facial automatizado, esta etapa tuvo un valor pedagógico significativo al revelar barreras y facilitadores en el aprendizaje. La persistencia y secuencia de ciertas emociones permitió inferir patrones que abren posibilidades para intervenciones didácticas más oportunas y personalizadas: las emociones negativas (enojo, tristeza) indicaron posibles dificultades con el contenido, frustración o desmotivación; por otro lado,

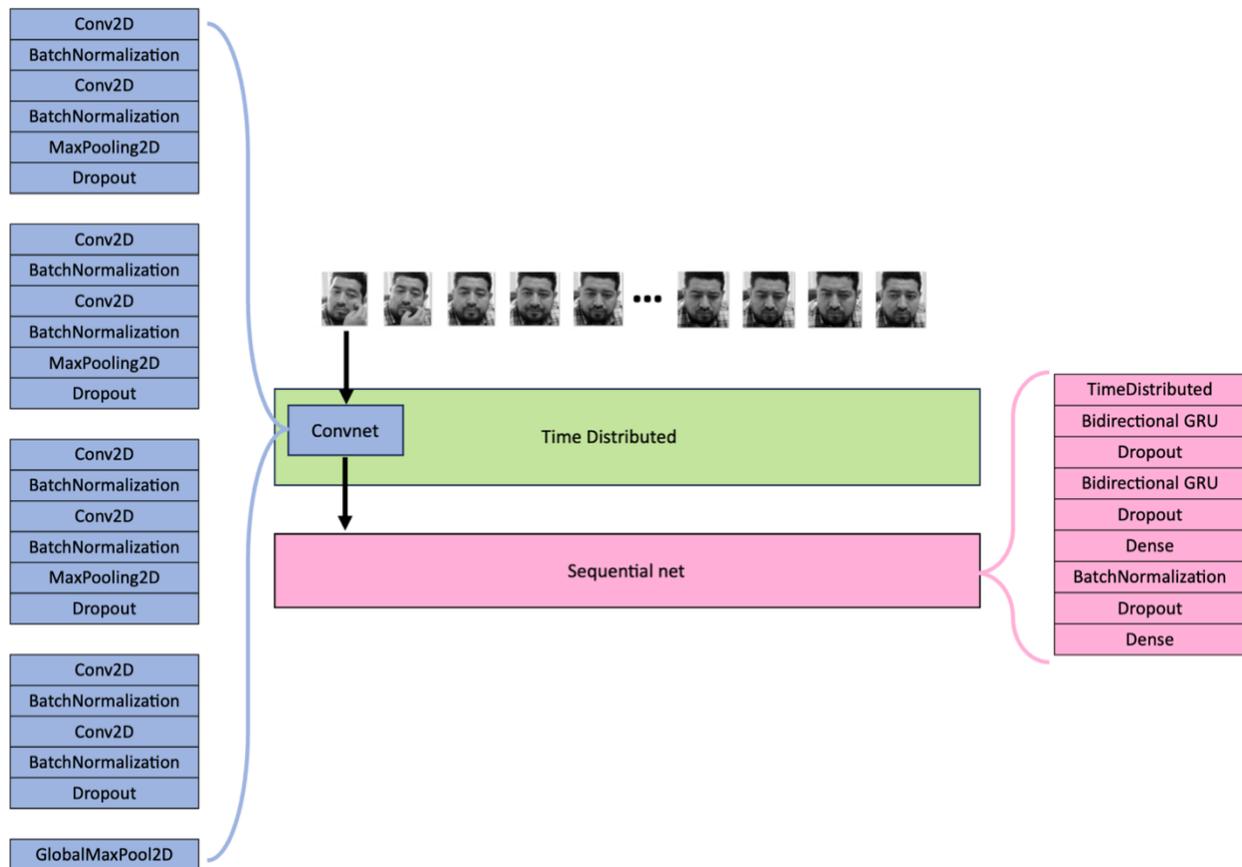
las emociones positivas (felicidad, sorpresa) se vincularon con momentos de comprensión, interés o disfrute del entorno de aprendizaje. Además, desde una perspectiva psicológica, se observó que las emociones detectadas no siempre se presentan de forma aislada [4], sino en secuencias dinámicas, lo que podría interpretarse como “shifting emocional”, que puede reflejar momentos de aprendizaje significativo.

### **Etapa 5: Entrenamiento y validación del modelo**

Se revisó literatura especializada, particularmente, para la selección de la arquitectura, bajo un análisis comparativo desde las perspectivas computacionales, pedagógicas y psicológicas, donde se destaca la estrategia propuesta por Liang *et al.* [9], quienes implementan una red profunda de fusión convolucional BiLSTM para el reconocimiento de emociones a partir de la expresión facial.

Se utilizó la propuesta de Sharma [10], quien implementa una red convolucional bidireccional LSTM en Keras para la clasificación de emociones básicas a partir de secuencias de fotogramas y que entrena con una versión recortada de la base CK+48, formada por 250 secuencias de 3 fotogramas cada una, en la que cada secuencia está etiquetada con una emoción básica. Para el diseño, se adaptó dicha arquitectura a que pudiera trabajar con la base de datos propia, esto es, secuencias de 100 fotogramas etiquetadas como “no” y “sí”, las cuales, de manera resumida, corresponden a “el alumno no muestra dificultades de aprendizaje” y “el alumno muestra dificultades de aprendizaje”.

Asimismo, se sustituyeron las capas LSTM por capas GRU, por lo que la arquitectura modificada corresponde a una red convolucional bidireccional GRU. La red extrae características faciales con una capa convolucional y las analiza secuencialmente en capas GRU bidireccionales. En la Figura 2, se muestra la arquitectura de la red.



**Fig. 2.** Arquitectura de la red neuronal convolucional bidireccional GRU [10].

A pesar de que los datos habían sido normalizados, existía una gran cantidad de fotogramas con expresión de neutralidad. Esto dificultó que la red neuronal pudiese diferenciar claramente si había o no dificultades de aprendizaje. Se requirió del apoyo del especialista en psicología para hacer una interpretación contextual de expresiones y reetiquetación de imágenes. Esta actividad resultó una tarea laboriosa, pero le permitió al equipo reinterpretar gestos estudiantiles de bloqueo cognitivo, más aún, le ayudó a comprender mejor la naturaleza emocional del aprendizaje en línea.

Aunado a lo anterior, debido a las restricciones de tiempo y personal, no se pudo realizar una experimentación exhaustiva; por ello, los resultados presentados corresponden a un experimento piloto. El mejor desempeño se obtuvo con 37 secuencias de 100

fotogramas (32x32x1 píxeles). Se utilizaron 31 secuencias para entrenamiento (11 de la clase “no” y 20 de la clase “sí”) y 6 para evaluación. La Figura 3 muestra un ejemplo de una secuencia de datos cuya etiqueta es “no”.

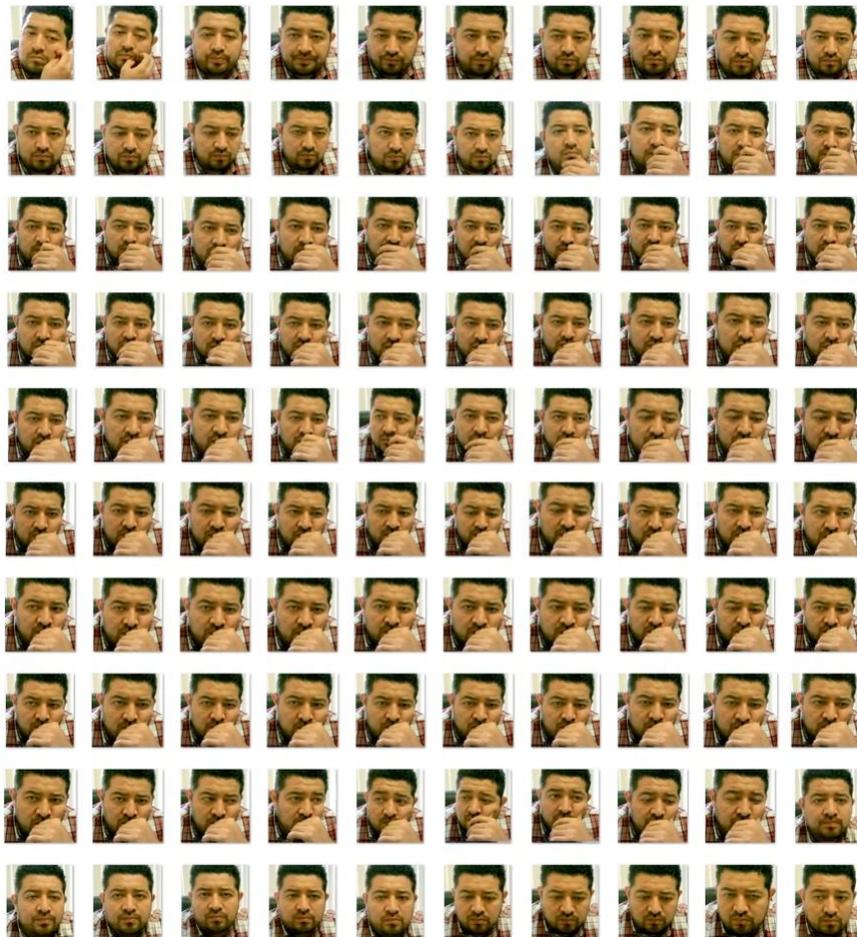
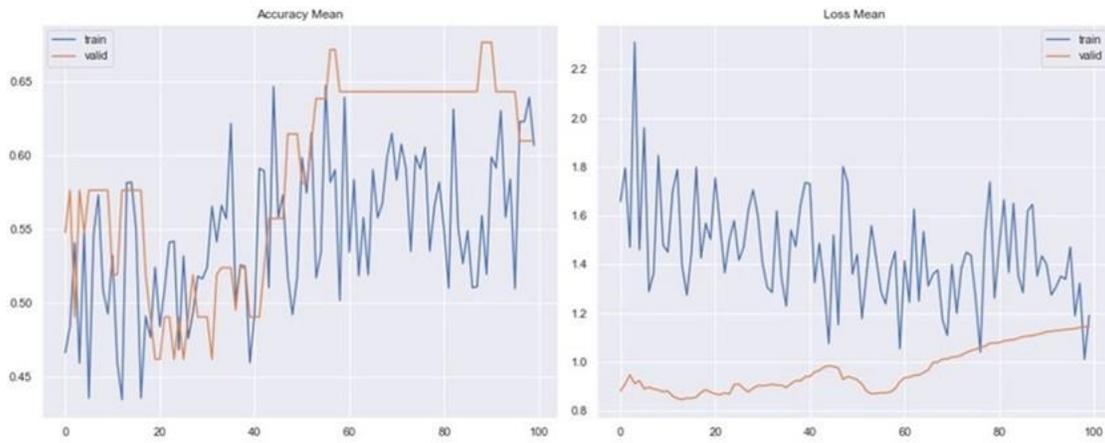


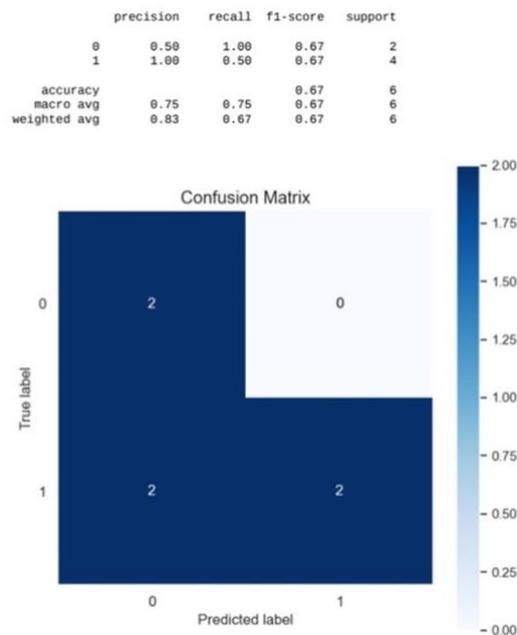
Fig. 3. Expresiones de no muestra dificultades.

Se realizó una validación cruzada de 5 iteraciones ( $k=5$ ). El análisis de las curvas de desempeño (Figura 4), mostró que la *accuracy* de validación alcanzó dos picos máximos en las épocas 56 y 57 (aproximadamente 67%), y en las épocas 88-90 (aproximadamente 68%). Sin embargo, dado que el primer pico coincide con un mínimo local de pérdida, se determinó que el entrenamiento óptimo se alcanza en la época 56.



**Fig. 4.** Resultados promedio obtenidos a través de la validación cruzada con k=5 durante las primeras 100 épocas.

Se procedió a entrenar la red con 57 épocas y se procedió con la evaluación final del modelo usando los datos reservados para la prueba (n=6). Los resultados obtenidos se muestran en la Figura 5. Se observa que se obtiene una capacidad predictiva de 67%.



**Fig. 5.** Representación de capacidad predictiva.

## Etapa 6: Difusión de resultados

Esta etapa tuvo como propósito garantizar un impacto académico mediante la difusión estratégica de resultados, lecciones aprendidas y experiencia adquirida, a pesar de no contar con un presupuesto asignado y de asumir voluntariamente con los costos relacionados. Se planificaron acciones específicas orientadas a distintos públicos y canales, lo que permitió participar en conferencias de divulgación, seminarios, talleres y reuniones que coincidían con hitos relevantes del proyecto. Asimismo, se generaron publicaciones en revistas científicas, conferencias nacionales e internacionales en temas de IA, educación y análisis afectivo. En cuanto a la evaluación general del proyecto, éste fue calificado como satisfactorio por los patrocinadores, quienes apreciaron el compromiso y desempeño destacado de los integrantes.

## Conclusiones

Durante ese año, se completó un ciclo de desarrollo de un sistema basado en IA, subrayando la integración eficaz, así como se consideraron distintos paradigmas y enfoques que permitieron enfrentar restricciones típicas como limitaciones de tiempo, recursos humanos y financiamiento. Se identifica lo siguiente:

(1) Planificación y diseño. La definición clara de objetivos se extiende a todos los participantes mediante una comunicación y retroalimentación continua, transparente, y bajo una visión sistemática, lo cual debe estar alineado a los resultados o efectos potenciales del desarrollo de una solución tecnológica.

(2) Desarrollo. Se describieron etapas y desafíos, en donde los autores concluimos que la transparencia fomenta la aceptación del sistema, dada la sensibilidad de datos y el riesgo de sesgos inherentes; por ello, comunicar métodos empleados para datos de entrenamiento, construcción e interpretación del modelo de IA o validación de algoritmos mediante métricas relevantes, como el sesgo y los falsos positivos, son acciones de accesibilidad al público. Los autores estamos conscientes de que la red neuronal entrenada que se obtuvo constituye un primer acercamiento para decidir si un alumno está o no presentando dificultades de comprensión. Sin embargo, se requieren de mayores recursos humanos y económicos.

(3) Infraestructura. El equipo tomó consciencia de que contar con la infraestructura de la UNAM- HUAWEI representó una ventaja significativa al proveer capacitación. Este apoyo también es indispensable para enfrentar exigencias computacionales asociadas a soportar grandes cargas de datos y su entrenamiento mediante procesos de redes neuronales profundas.

(4) Consecuencias. La realización de un proyecto de IA genera impactos significativos en las organizaciones, personas y grupos implicados. Para las facultades participantes, representa el fortalecimiento institucional. Para los que contribuyeron en este proyecto, se mejoró su marca personal a nivel académico y profesional. Para los estudiantes, representó curiosidad e interés, una mayor conciencia emocional y metacognitiva, pero, especialmente, expectativas de mejora pedagógica.

(5) Recomendaciones. Los autores estamos convencidos de que todo desarrollo en IA debe incorporar consideraciones éticas, en todas sus etapas, que se reflejen en la equidad, confiabilidad, transparencia, explicabilidad de los modelos o la rendición de cuentas, lo cual coadyuva a evaluar el impacto potencial, particularmente en contextos sensibles, cuando se trabaja con estudiantes y su aprendizaje de matemáticas. Para escalar el proyecto y mejorar sus resultados, es indispensable contar con un segundo ciclo de trabajo, respaldado por mayores recursos económicos que permitan incorporar recursos humanos dedicados a incrementar la cantidad y calidad de los datos de entrenamiento, realizar ajustes finos en los hiperparámetros y optimizar el desempeño del modelo, automatizar la recolección, organización de datos y etiquetado para reducir la carga operativa, así como diseñar una interfaz de usuario que responda a propósitos de aprendizaje matemático.

Otras recomendaciones son: (a) extender el periodo de la convocatoria para sumar más estudiantes a través de programas de servicio social, prácticas profesionales o becas, lo cual permite fortalecer el proyecto mientras se forman nuevos talentos; (b) fortalecer al proyecto en su vinculación con programas de posgrado, invitando estudiantes de maestría o doctorado; y (c) articular el proyecto con otras fuentes institucionales de financiamiento.

## Agradecimientos

Este artículo ha sido posible gracias al apoyo recibido de la UNAM y Proyectos de Investigación en Inteligencia Artificial en el Espacio de Innovación UNAM – HUAWEI, Proyecto No. 4 (antes No. 8). Agradecemos a todos los alumnos que tomaron clases y se grabaron en Zoom. Asimismo, a los prestadores de servicio social Ivan Gallegos Zamora, Santiago Nieva y Cruz, Israel Villa Olvera, Cesar Haziél Pineda Pacheco y Cristian Omar Rubio Ceja, quienes contribuyeron significativamente en la realización del proyecto.

## Referencias

- [1] G. J. Miller, "Framework for Managing Artificial Intelligence (AI) Projects: Avoiding Harms, Losses, and Damages", in International Program and Project Management - Best Practices in Selected Industries, K. D. Strang and N. R. Vajjhala, Eds., vol. 31, Cham, Switzerland: Springer, 2025, pp. 135-162, doi: 10.1007/978-3-031-80275-1\_7.
- [2] J. Friedrich, A. Brückner, J. Mayan et al., "Human-centered AI development in practice -insights from a multidisciplinary approach" Zeitschrift für Arbeitswissenschaft, vol. 78, pp. 359-376, 2024. doi: 10.1007/s41449-024-00434-5.
- [3] UNAM, "Estadísticas de exámenes", (publicación privada), 2024.
- [4] R. Pekrun, S. Lichtenfeld, H. W. Marsh, K. Murayama, and T. Goetz, "Achievement emotions and academic performance: longitudinal models of reciprocal effects", Child Development, vol. 88, no. 5, pp. 1653-1670, 2017.
- [5] U. Eligio, "An overview of the growth and trends of current research on emotions and mathematics", in Understanding Emotions in Mathematical Thinking and learning. U. Eligio, Ed. Mexico City: Academic Press, pp. 3-41, 2017.
- [6] B. Woolf, M. Betke, H. Yu, H., S. Bargal, I. Arroyo, J. Magee, D. Allesio, and W. Rebelsky, W., "FACE READERS: The Frontier of Computer Vision and Math Learning", in 2023 Workshop: Towards the Future of AI -Augmented Human Tutoring in Math Learning, CEUR Workshop, pp. 1-13, 2023.

- [7] International Standards Organization, iso 21502:2020-12 Project, Programme and Portfolio Management- Guidance on Project Management, 1st ed. Geneva, Switzerland: ISO, 2020.
- [8] S. Serengil, DeepFace (Versión 0.0.68) [Software de reconocimiento facial en Python]. GitHub. <https://github.com/serengil/deepface> (Accedido: octubre 15, 2024).
- [9] D., Liang, H. Liang, Z. Yu, and Y. Zhang, "Deep convolutional BiLSTM fusion network for facial expression recognition", *The Visual Computer*, vol. 36, pp. 499-508, 2020.
- [10] G. Sharma, "FER using Convolutional Bidirectional LSTM," Kaggle, 2020. [En línea]. Disponible en: <https://www.kaggle.com/code/gauravsharma99/fer-using-convolutional-bidirectional-lstm>. [Accedido: 21-oct-2024].



# *Heart beating*: una aplicación para el diagnóstico prenatal inteligente

## **Julio César Pérez Sansalvador**

Secretaría de Ciencia, Humanidades, Tecnología e Innovación, Ciudad de México, México.  
Instituto Nacional de Astrofísica, Óptica y Electrónica, Puebla, México.  
ORCID: 0000-0002-2656-3659

## **Arelly Ornelas Vargas**

Secretaría de Ciencia, Humanidades, Tecnología e Innovación, Ciudad de México, México.  
Instituto Politécnico Nacional, Centro interdisciplinario de Ciencias Marinas (CICIMAR-IPN),  
Baja California Sur, México.  
ORCID: 0000-0002-8371-699X

## **Alejandro Barreiro Valdez**

Estudiante de Ingeniería en Computación en la Universidad Nacional Autónoma de México,  
Facultad de Ingeniería, Ciudad de México, México.  
ORCID: 0009-0000-4262-5912

## **Kernel Enrique Prieto Moreno**

Realiza una estancia posdoctoral en The University of British Columbia, Centre for Disease Control,  
Vancouver, Canadá.  
ORCID: 0000-0002-3195-964X

## **Socrates Romero Reyes**

Estudiante doctoral en el Instituto Nacional de Astrofísica, Óptica y Electrónica, Puebla, México.  
ORCID: 0000-0002-4153-4258

Recepción: 12 de noviembre de 2024.

Aceptación: 20 de agosto de 2025.

Septiembre 2025 • número de revista 13 • DOI: 10.22201/dgtic.26832968e.2025.13.56

## **Heart beating: una aplicación para el diagnóstico prenatal inteligente**

---

### **Resumen**

La electrocardiografía fetal no invasiva (NI-FECG, por sus siglas en inglés) es una técnica que captura la señal combinada del ritmo cardíaco materno y fetal. Esta señal es analizada por especialistas en señales y cardiólogos pediátricos en busca de anomalías cardíacas en el feto. En este trabajo, presentamos *Heart beating*, una aplicación para el análisis de señales NI-FECG que utiliza redes neuronales profundas para estimar el ritmo cardíaco del feto e identificar anomalías como la bradicardia o la taquicardia. A corto plazo, esta tecnología busca servir como una herramienta para el entrenamiento de cardiólogos pediátricos y en el diagnóstico médico asistido por computadora.

**Palabras Clave:** electrocardiografía fetal no invasiva, redes neuronales profundas, procesamiento de señales biomédicas, anomalías del ritmo cardíaco fetal, inteligencia artificial en medicina.

### **Heart beating: An Intelligent Prenatal Diagnosis Application**

---

#### **Abstract**

*Non-invasive fetal electrocardiography (NI-FECG) is a technique that captures the combined signal of both maternal and fetal heart rates. This signal is analyzed by signal processing specialists and pediatric cardiologists to detect cardiac anomalies in the fetus. In this article, we present Heart beating, an application for analyzing NI-FECG signals using deep neural networks to estimate the fetus heart rate and identify anomalies like the bradycardia or tachycardia. In the short term, this technology aims to serve as a tool for training pediatric cardiologists and for computer-assisted medical diagnosis.*

**Keywords:** *non-invasive fetal electrocardiography, deep neural networks, biomedical signal processing, fetal heart rhythm anomalies, artificial intelligence in medicine.*

## Introducción

En el año 2023, en México, se registraron alrededor de 23,000 muertes fetales, lo que corresponde a una tasa de 67.5 muertes por cada 100 mil mujeres de entre 15 a 49 años, de las cuales el 15% declaró no haber recibido atención prenatal [1, 2].

La detección de anomalías cardíacas en fetos requiere de especialistas en ecocardiografía obstétrica y cardiólogos pediátricos [3]. Con base en datos oficiales, en México, contamos con 107 especialistas por cada 100 mil habitantes, cifras por debajo del mínimo recomendado por la ONU, que es de 230 especialistas por cada 100 mil habitantes [4]. Esto es preocupante pues, a nivel global, se ha identificado que el 98% de las muertes fetales ocurren en poblaciones con acceso limitado a especialistas médicos, con escasos recursos o en vías de desarrollo [5].

Por otro lado, existe evidencia que la atención oportuna de anomalías en el ritmo cardíaco en fetos logra reducir la probabilidad de muerte y mejorar la calidad de vida de los neonatos [6, 7]. Los patrones anormales como aceleración o desaceleración del ritmo cardíaco son indicativos de alteraciones cardíacas como la bradicardia o la taquicardia. La manera de identificar estos patrones es a partir de técnicas de observación basadas en ecocardiografía (ECOs), electrocardiogramas (ECGs), estudio Doppler y magnetocardiografía [4].

El ECG es el procedimiento para medir la actividad eléctrica del corazón y ayuda a identificar alteraciones en el ritmo cardíaco [6]. Al practicar un ECG a una mujer embarazada, colocando los electrodos del sistema de monitoreo alrededor del vientre de la mujer, es posible obtener un ECG compuesto, es decir, con el ritmo cardíaco materno y fetal. A esta técnica, se le conoce como electrocardiograma fetal no invasivo (NI-FECG) [8].

El análisis de un ECG compuesto requiere de una colaboración multidisciplinaria: primero, especialistas en señales biomédicas separan la señal cardíaca fetal (fECG) de la materna (mECG); posteriormente, cardiólogos pediátricos y especialistas en ecocardiografía obstétrica analizan la fECG en busca de anomalías [7, 9, 10].

## **Heart beating**

El desarrollo de tecnologías en el ámbito médico para automatizar, administrar y procesar grandes cantidades de datos son parte de los retos del *Healthcare* 4.0 [11, 12]. Para abordar estos retos, utilizamos tecnologías como el aprendizaje computacional profundo, que requiere de grandes cantidades de datos para generar resultados confiables. Con ayuda de éstas, desarrollamos *Heart beating*, una aplicación para la identificación de anomalías cardíacas en el feto a partir del análisis de señales NI-FECG.

La aplicación *Heart beating* busca reducir la probabilidad de muerte y mejorar la calidad de vida de neonatos con padecimientos cardíacos al reducir los tiempos en el proceso de detección de anomalías en el ritmo cardíaco de fetos.

## **Metodología para el desarrollo de *Heart beating***

El proceso de desarrollo contempla tres etapas principales:

1. Generación de señales NI-FECG.
2. Diseño de los modelos de redes neuronales.
3. Entrenamiento y evaluación de los modelos computacionales.

## **Generación de señales NI-FECG**

De formar análoga al aprendizaje humano, el modelo de *Heart beating* requiere analizar un gran volumen de señales NI-FECG para identificar patrones y asociarlos al ritmo cardíaco fetal previamente anotado por un especialista.

Los modelos de redes neuronales de esta primera versión son entrenados con señales NI-FECG sintéticas, es decir, señales simuladas con anotaciones. Estas señales son generadas por el simulador *fecgsyn* de *PhysioNet* [13, 14, 15], el cual puede generar señales NI-FECG con una gran variedad de características. La selección de este simulador se basa en la variedad de parámetros que nos permite configurar como el tiempo de simulación para la señal, la

frecuencia de muestreo, el número de derivaciones, el ritmo cardíaco de la madre y del feto, movimiento fetal, contracciones uterinas en la madre y diferentes tipos de ruido en la señal.

Con ayuda del simulador, generamos más de ocho mil señales NI-FECG sintéticas con las siguientes características: tiempo de simulación de 60 segundos, frecuencia de muestreo de 1Mhz, treinta y cuatro derivaciones, ritmos cardíacos maternos en el rango de 65 a 120 *bpm*, ritmos cardíacos fetales en el rango de 60 a 200 *bpm*, sin movimiento fetal, sin contracciones uterinas de la madre y sin ruido.

Con las características anteriores, cada señal generada es representada como una matriz de 34 x 60,000 elementos. Las filas representan las lecturas obtenidas por cada uno de los 34 electrodos simulados colocados alrededor del vientre de la madre. Las columnas corresponden al número de lecturas realizadas en cada electrodo a la frecuencia de muestreo especificada. Los rangos para los ritmos cardíacos fetales corresponden a ritmos cardíacos en fetos con más de doce semanas de gestación.

## Diseño de los modelos de redes neuronales

Luego de realizar un análisis de la literatura científica relacionada, identificamos varios trabajos que procesan señales NI-FECG para automatizar la extracción del complejo QRS materno y fetal a partir de un proceso de dos etapas. En la primera etapa, se elimina el QRS materno utilizando técnicas de procesamiento de señales como descomposición, reconstrucción, filtros adaptativos y eliminación de ruido [16, 17]. En la segunda etapa, se extrae el QRS fetal a partir de técnicas de filtrado de señales, entropía, extracción del espectro de la señal, aprendizaje computacional y combinaciones de éstas [18, 19]. A partir del QRS fetal, especialistas en cardiología pediátrica analizan la señal en busca de patologías del ritmo cardíaco del feto.

Durante el diseño de *Heart beating*, revisamos modelos clásicos de redes neuronales para señales en 2D como *U-Net Medical*, *FasterRCNN Resnet101* e *Inception v4*, sin embargo, ninguno de ellos se ha aplicado al procesamiento de señales de ECGs, lo que sugiere una ventana de oportunidad para esto. En esta primera versión de *Heart beating*, adaptamos la arquitectura de la red neuronal convolucional (CNN) presentada por Zhong *et. al.* en [20] para

la detección del factor QRS en señales NI-FECG. En la Tabla I, presentamos las diferencias con *Heart beating*.

**Tabla I.** Diferencias entre el trabajo de Zhong et. al. [20] y *Heart beating*.

<b>Característica</b>	<b>Zhong et. al. [20]</b>	<b><i>Heart beating</i></b>
Número de señales	75	>8,000
Número de derivaciones	4	34
Preprocesamiento de la señal de entrada	Sí, se descartan señales que no cumplen con criterios de calidad	No es necesario
Longitud de señal de entrada	100 ms	60,000 ms
Tipo de clasificación	Binaria (existe/no existe complejo QRS en la señal de entrada)	Multivariable (identificación del rango del ritmo cardíaco fetal)

Además, *Heart beating* explota la relación espacio-temporal de las señales generadas por el simulador. Esta relación surge de utilizar treinta y cuatro electrodos que capturan al mismo tiempo el ritmo cardíaco materno y fetal en diferentes posiciones del vientre de la madre. En la Tabla II, se describe la arquitectura de la CNN implementada.

**Tabla II.** Capas de la CNN implementada en *Heart beating*

<b>Secuencia de capas</b>	<b>Descripción</b>
Convolución 2D	Operación de convolución en la señal 2D de entrada
Batch normalization	Normalización de datos
ReLU	Función de activación ReLU
Convolución 2D	Operación de convolución en la señal 2D

Secuencia de capas	Descripción
Dropout	Eliminación de datos en la señal para evitar sobre ajuste
Maxpool	Extracción de valores máximos en la señal 2D
ReLU	Función de activación ReLU
Convolución 2D	Operación de convolución en la señal 2D
Dropout	Eliminación de datos en la señal para evitar sobre ajuste
Maxpool	Extracción de valores máximos en la señal
ReLU	Función de activación ReLU
Dense	Capa densa para preparar datos para la capa de salida
Softmax	Cómputo de pertenencia de la señal a las clases objetivo

### Entrenamiento y evaluación del modelo neuronal *Heart beating*

La frecuencia cardíaca del feto permite identificar alteraciones cardíacas como la bradicardia, con frecuencia cardíaca menor a 100bpm, y la taquicardia, mayor a 160bpm [21]. Debido a que las señales NI-FECG fueron generadas por un simulador, las dividimos en siete grupos que nos permiten identificar estas alteraciones, ver Tabla III.

**Tabla III.** Correspondencia entre frecuencias cardíaca fetal y patología cardíaca

Rango de frecuencia cardíaca fetal	Patología cardíaca
60 – 79 bpm	Bradicardia severa
80 – 99 bpm	Bradicardia leve

Rango de frecuencia cardíaca fetal	Patología cardíaca
100 – 119 bpm	Sin patología
120 – 139 bpm	Sin patología
140 – 159 bpm	Sin patología
160 – 179 bpm	Taquicardia leve
180 – 200 bpm	Taquicardia severa

Los grupos de rangos de frecuencia cardíaca fetal son las clases objetivo del modelo entrenado. Para evitar el desbalance de clases, se generó un número similar de instancias para cada uno de los grupos propuestos.

Para el entrenamiento y la evaluación del modelo neuronal presentado, creamos tres grupos: “entrenamiento”, “validación” y “pruebas”, con el 70%, 20% y 10%, respectivamente, del total de señales NI-FECG generadas por el simulador. Los grupos “entrenamiento” y “validación” son utilizados durante la etapa de entrenamiento para monitorear el progreso de aprendizaje del modelo y evitar el sobreajuste. Al término de ésta, evaluamos el modelo neuronal con el grupo de “pruebas” para medir sus capacidades de generalización.

Los valores para los hiperparámetros de entrenamiento se presentan en la Tabla IV; observe que se utiliza un factor de aprendizaje bajo para asegurar la convergencia estable del modelo.

**Tabla IV.** Valores para los hiperparámetros de entrenamiento del modelo

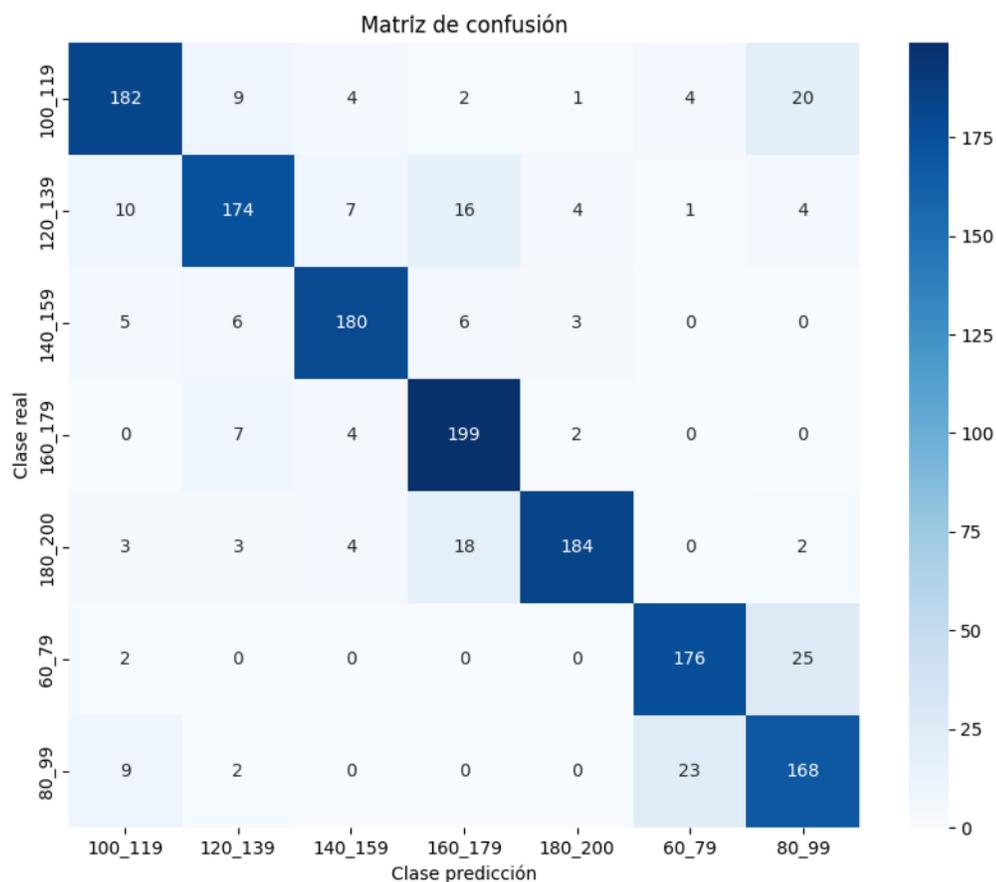
Número de señales	8,806
Dimensiones de cada señal	34 x 60,000

Tamaño del lote ( <i>batch size</i> )	64
Número de derivaciones	34
Número de épocas ( <i>epochs</i> )	10
Factor de aprendizaje ( <i>learning rate</i> )	1.0e-4
Factor de regularización ( <i>weight decay</i> )	1.0e-3

Los resultados de la evaluación del modelo neuronal en el grupo de “pruebas” se presentan en la Tabla V. Los resultados son superiores a los reportados en la literatura que utilizan el complejo QRS para la predicción del ritmo cardíaco fetal. En la Figura 1, se presenta la matriz de confusión para el grupo “pruebas”, el eje horizontal presenta la clase predicha y el eje vertical la clase real. En esta matriz, observamos el número de señales NI-FECG clasificadas de manera correcta en la diagonal de ésta. Los elementos fuera de la diagonal representan a los falsos positivos y falsos negativos.

**Tabla V.** Resultados de evaluación del modelo neuronal

Exactitud	8,86%
AUROC	0.98
Sensibilidad	0.86
Especificidad	0.97



**Fig. 1.** Matriz de confusión para el grupo “pruebas”.

### Infraestructura para el desarrollo de la aplicación *Heart beating*

Con el objetivo de apoyar e incentivar el desarrollo de proyectos de innovación tecnológica basados en inteligencia artificial, la Universidad Nacional Autónoma de México, con la participación de Huawei Technologies Mexico, la Secretaría de Relaciones Exteriores y la Cooperación Alemana para el Desarrollo Sostenible en México, invitó a la comunidad científica y sociedad en general a participar en la Convocatoria “Proyectos de investigación en Inteligencia Artificial en el Espacio de Innovación UNAM - HUAWEI” para obtener acceso a recursos de cómputo de alto desempeño con el fin de entrenar modelos de inteligencia artificial [22].

Entre los recursos a los que se tuvieron acceso durante el proyecto de desarrollo de la aplicación *Heart beating*, se encuentran servidores Huawei Atlas 800, códigos de ejemplo, modelos de redes neuronales Huawei-Ascend y cursos de capacitación.

Este tipo de esta infraestructura es fundamental en el desarrollo de aplicaciones basadas en inteligencia artificial, pues el entrenamiento de los modelos neuronales requiere de una gran cantidad de recursos computacionales, en capacidades tanto de almacenamiento como de procesamiento.

## Conclusiones

La aplicación *Heart beating* permite la identificación de ritmos cardíacos fetales a partir de señales NI-FECG sin requerir preprocesamiento, esto permite la identificación de anomalías cardíacas como la bradicardia y la taquicardia. Los resultados de la evaluación del modelo superan a los reportados por los trabajos relacionados, sin embargo, aún se realizan mejoras para extender sus capacidades de aplicación para su uso clínico.

La primera versión de *Heart beating* ha sido desarrollada como una prueba de concepto por lo que, para versiones posteriores, se espera contar con señales NI-FECG más complejas, que incluyan movimientos del feto y contracciones uterinas de la madre. Además, para evaluar la robustez de *Heart beating*, se espera aplicarla en la predicción de ritmos cardíacos fetales en señales clínicas y, de ser necesario, ajustar tanto los parámetros como el diseño de los modelos neuronales con base en los resultados de predicción.

En particular, el desarrollo de la aplicación *Heart beating* es un primer paso para reducir los tiempos de diagnóstico y tratamiento, cruciales en el caso de anomalías cardíacas en fetos. La aplicación de la inteligencia artificial como apoyo en el diagnóstico médico se visualiza como un área con grandes oportunidades, pero también con grandes retos que, para superarlos, se requiere de la colaboración de especialistas en varios campos de la ciencia como la medicina, ciencias computacionales, matemáticas, física, biología, química, entre otras.

Por otro lado, es también importante dar crédito a las estrategias de colaboración y acceso a infraestructura de cómputo de alto desempeño como la propuesta por la Alianza UNAM-HUAWEI mediante la convocatoria "Proyectos de investigación en Inteligencia Artificial

en el Espacio de Innovación UNAM - HUAWEI", pues resultan esenciales para el desarrollo de tecnologías y proyectos de investigación que utilicen la inteligencia artificial como apoyo en el diagnóstico médico.

## Referencias

- [1] INEGI "Estadísticas de defunciones fetales (EDF) 2023", Comunicado de prensa número 499/24. [online] 24/10/2024  
Disponible:<https://www.inegi.org.mx/contenidos/saladeprensa/boletines/2024/EDF/EDF2023.pdf>
- [2] CONEVAL "Estimaciones de pobreza multidimensional 2018 y 2020, CONEVAL", Comunicado de prensa Num. 09. [online] 30/08/2024. Disponible:  
[https://www.coneval.org.mx/SalaPrensa/Comunicadosprensa/Documents/2021/COMUNICADO\\_009\\_MEDICION\\_POBREZA\\_2020.pdf](https://www.coneval.org.mx/SalaPrensa/Comunicadosprensa/Documents/2021/COMUNICADO_009_MEDICION_POBREZA_2020.pdf)
- [3] C. García-Guevara, J. Pérez-Pedregosa, M. Cazzaniga, "El diagnóstico en el feto cardiópata", *An Pediatr Contin*. Vol. 10, No. 6, pp. 324-333, 2012.
- [4] Secretaría de salud, Médicos especialistas en México, May, 2022. [online] 30/10/2024. Disponible:  
[https://www.gob.mx/cms/uploads/attachment/file/729437/CPM\\_Salud\\_Me\\_dicos\\_especialistas\\_\\_31may22.pdf](https://www.gob.mx/cms/uploads/attachment/file/729437/CPM_Salud_Me_dicos_especialistas__31may22.pdf)
- [5] Z. A. Bhutta, M.Y. Yakoob, J. E. Lawn, A. Rizvin, I. K. Friberg, E. Weissman, E. Buchmann, R. L. Goldenberg "Stillbirths Series steering committee. Stillbirths: what difference can we make and at what cost?", *Lancet*, Vol. 377, No. 9776, pp. 1523-1538, Apr 2011.
- [6] L. M. Garrido-García, M. G. Delgado-Onofre, "Trastornos del ritmo en el recién nacido", *Acta Pediat Mex*. Vol. 35, No. 2, pp. 148-158, Jan 2014.
- [7] R. Vullings, "Fetal Electrocardiography and Deep Learning for Prenatal Detection of Congenital Heart Disease", 2019 Computing in Cardiology CinC, Singapore, 2019, pp. 1-4.

- [8] G. Aggarwal, Y. Wei, "Non-Invasive Fetal Electrocardiogram Monitoring Techniques: Potential and Future Research Opportunities in Smart Textiles", *Signals* Vol. 2, No. 3. pp. 392-412. 2021.
- [9] J. Behar, J. Oster, G. D. Clifford. "Non-invasive FECG extraction from a set of abdominal sensors", *Computing in Cardiology 2013*, Zaragoza, Spain, 2013, pp. 297-300.
- [10] W. Lo, P. Y. Tsai "Deep Learning for Detection of Fetal ECG from MultiChannel Abdominal Leads", 2018 Asia-Pacific Signal and Information Processing, Association Annual Summit and Conference (APSIPA ASC), Honolulu, HI, USA, 2018, pp. 1397-1401.
- [11] A. Gupta, A. Singh, "Healthcare 4.0: recent advancements and futuristic research directions". *Wirel Pers Commun*. Vol. 29, No. 2, pp. 933-952. 2023.
- [12] J. N. Muñoz-Montes-de-Oca, H. Romero-Morales, J. J. Reyes-Lagos, "De estetoscopios a dispositivos inteligentes: una revolución del monitoreo materno-fetal", *Ciencia*, Vol. 75. No. 4, pp. 85-89, 2024
- [13] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. C. Ivanov, R. Mark, J. E. Mietus, G. B. Moody, C. K. Peng, H. E. Stanley. "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals", 2000, [online] 07/11/2024. Disponible: <https://physionet.org/content/challenge-2013/1.0.0>
- [14] F. Andreotti, J. Behar, S. Zaunseder, J. Oster, G.D. Clifford, "An open-source framework for stress-testing non-invasive foetal ECG extraction algorithms", *Physiological Measurement*, Vol. 37, No. 5, pp. 627-648, 2016.
- [15] J. Behar, F. Andreotti, S. Zaunseder, Q. Li, J. Oster, G. D. Clifford, "An ECG imulator for generating maternal-foetal activity mixtures on abdominal ECG recordings", *Physiological Measurement*, Vol. 35, No. 8. pp. 1537-1550, 2014.
- [16] I. Silva, J. Behar, R. Sameni, T. Zhu, J. Oster, G.D. Clifford, G. B. Moody, "Noninvasive Fetal ECG: the PhysioNet/Computing in Cardiology Challenge 2013". *Computing in cardiology 2013*, Zaragoza, Spain, 2013, Vol. 40, pp. 149-152.

- [17] A. Georgieva, P. Abry, M. G. Frasc, I. M. Nunes, eds. (2022), *Fetal-Maternal Monitoring in the Age of Artificial Intelligence and Computer-Aided Decision Support: A Multidisciplinary Perspective*, Lausanne: Frontiers Media SA.
- [18] X. Li, J. Wan, & X. Peng, "Review of Non-Invasive Fetal Electrocardiography Monitoring Techniques", *Sensors*, 25(5), 1412, 2025.
- [19] F. Eleni, V. Rik, "Multi-Channel Fetal ECG Denoising With Deep Convolutional Neural Networks", *Frontiers in Pediatrics*, Vol. 8, 2020.
- [20] W. Zhong, L. Liao, X. Guo, G. Wang "A deep learning approach for fetal QRS complex detection", *Physiological Measurement*, Vol. 39, pp. 1-9. 2018.
- [21] "What is fetal heart monitoring", *Fetal Heart Monitoring*, Johns Hopkins, Medicine, [online] 13/08/2025. Available: <https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/fetal-heart-monitoring>
- [22] Alianza "Proyectos de Investigación en Inteligencia Artificial en el Espacio de Innovación UNAM-HUAWEI", [online] 13/08/2025 Disponible: <https://alianza.unam.mx/convocatoria-2022/>



# Extracción de interacciones regulatorias transcripcionales de bacterias a partir de literatura biomédica utilizando inteligencia artificial.

## **Ali-Berenice Posada-Reyes**

Universidad Nacional Autónoma de México, Facultad de Estudios Superiores Cuautitlán, Laboratorio de Microbiología, Inmunología y Salud Pública, Estado de México, México.  
ORCID: 0000-0003-3165-6099

## **Alfredo Varela-Vega**

Universidad Nacional Autónoma de México, Centro de Ciencias Genómicas, Programa de Genómica Computacional, Morelos, México.  
ORCID: 0000-0003-4857-4997

## **Sara Berenice Martínez-Luna**

Universidad Nacional Autónoma de México, Centro de Ciencias Genómicas, Programa de Genómica Computacional, Morelos, México.  
ORCID: 0009-0001-1329-6344

## **Carlos-Francisco Méndez-Cruz**

Universidad Nacional Autónoma de México, Centro de Ciencias Genómicas, Programa de Genómica Computacional, Morelos, México.  
ORCID: 0000-0002-2549-1614

Recepción: 29 de octubre de 2024.

Aceptación: 21 de marzo de 2025.

Septiembre 2025 • número de revista 13 • DOI: 10.22201/dgtic.26832968e.2025.13.49

## Extracción de interacciones regulatorias transcripcionales de bacterias a partir de literatura biomédica utilizando inteligencia artificial

---

### Resumen

Las redes de regulación transcripcional (TRNs) de bacterias brindan una visión global de los mecanismos de respuesta de éstas a los cambios en su ambiente. El estudio de estas redes permite ampliar el conocimiento biológico y derivar en investigaciones con implicaciones clínicas o farmacéuticas. Sin embargo, la reconstrucción de TRNs se hace tradicionalmente de forma manual mediante un proceso demandante y costoso de curación de artículos científicos. En este trabajo, describimos la aplicación de aproximaciones de inteligencia artificial (IA), específicamente la afinación de *transformers* pre-entrenados BERT, para la extracción automática de TRNs de la literatura. Con 1562 oraciones de entrenamiento de la bacteria *Escherichia coli*, comparamos seis arquitecturas tipo BERT. El mejor modelo afinado obtuvo resultados significativos (F1-score: 0.8685, coeficiente de correlación de Matthews: 0.8163). Con este modelo, extrajimos correctamente el 82% de una TRN de la bacteria *Salmonella* utilizando 264 artículos completos. El factor de transcripción PhoP fue relevante en la red por obtener el mayor valor de conexiones (*degree*=180) por lo que analizamos biológicamente su comunidad de genes. Este trabajo muestra el uso de la IA para facilitar la extracción de conocimiento biológico que podría ser utilizado por estudios futuros en el área biomédica.

**Palabras Clave:** Inteligencia artificial, procesamiento de lenguaje natural, extracción de información, extracción de interacciones regulatorias, regulación transcripcional.

## ***Extraction of transcriptional regulatory interactions of bacteria from biomedical literature using artificial intelligence***

### **Abstract**

*Transcriptional regulatory networks (TRNs) from bacteria provide a global view of their mechanisms of response to changes in the environment. The study of these networks allows to expand biological knowledge and leads to research with clinical or pharmaceutical implications. However, the reconstruction of TRNs is traditionally done manually through a demanding and expensive process of curation of scientific articles. In this work, we describe the application of Artificial Intelligence (AI) approaches, specifically the fine-tuning of BERT pre-trained transformers, for the automatic extraction of TRNs from biomedical literature. With 1562 training sentences from the Escherichia coli bacterium, we compared six BERT architectures. The best fine-tuned model obtained significant results (F1-score: 0.8685, Matthews correlation coefficient: 0.8163). With this model, we correctly extracted 82% of a TRN from the Salmonella bacterium using 264 full articles. The PhoP transcription factor was relevant in the network because it obtained the highest value of connections (degree=180) so we biologically analyzed its gene community. This work shows the use of AI to facilitate the extraction of biological knowledge that could be used for future studies in the biomedical field.*

**Keywords:** Artificial intelligence, Natural language processing, Information extraction, Regulatory interaction extraction, Transcriptional regulation.

### **1. Importancia de las redes de regulación transcripcional de bacterias**

El genoma de las bacterias es controlado por diferentes mecanismos de regulación que permiten o inhiben la expresión de sus genes en un proceso denominado regulación de la transcripción. Así, con la expresión de sus genes, las bacterias pueden responder a señales internas o ambientales [1]. Un ejemplo de lo anterior es lo que ocurre en la bacteria *Escherichia coli* (*E. coli*), en la que altas concentraciones de oxígeno funcionan como una señal para activar la respiración aeróbica mediante la activación o represión de ciertos genes [2]. Estos mecanismos reguladores pueden ser representados por las denominadas redes de

regulación transcripcional (*transcriptional regulatory networks*, TRNs), que describen la organización global de la regulación transcripcional de un organismo y muestran características jerárquicas y modulares gracias a la representación en red [3].

Las bacterias se adaptan a los cambios de su entorno principalmente a través de la acción de proteínas denominadas factores de transcripción (*transcription factors*, TFs), que activan o reprimen la expresión de genes blanco para inducir mecanismos dentro de la bacteria y así adaptarse a su entorno [3]. Precisamente, el conjunto de interacciones regulatorias (TF-gen) de la transcripción en una bacteria conforman una TRN bacteriana [4]. El estudio de TRNs de bacterias es relevante para el entendimiento de diversos mecanismos, por ejemplo, se ha descrito que la TRN de *E. coli* desempeña un papel importante en la supervivencia de esta bacteria para colonizar el intestino grueso dentro de sus hospederos [5].

Reportar el conocimiento derivado del estudio de las TRNs es crucial, ya que podemos descubrir las adaptaciones que tienen las bacterias a los desafíos ambientales. Por ejemplo, desde la mirada de la evolución, Dalldorf y colaboradores [6] descubrieron en experimentos de laboratorio que *E. coli* puede afrontar desafíos ambientales a corto plazo a través de la modulación genética de las TRNs a pesar de mutaciones aplicadas a los TFs [6]. Una utilidad relevante del estudio de TRNs de bacterias en el área de la biomedicina es la identificación de mutaciones reguladoras en cepas de bacterias con asociación a resistencia a fármacos. Tal es el caso del estudio de Praski y colaboradores del 2017 [7], que se encargaron de identificar mutaciones en el gen *marR* en aislados clínicos de *E. coli* resistentes a ciprofloxacina. El gen *marR* codifica la proteína MarR, que es un TF represor que regula la transcripción de diversos genes encargados del mecanismo de resistencia a fármacos. En su estudio, reportan las correlaciones de las mutaciones en *marR* y la susceptibilidad al fármaco ciprofloxacina [7].

Otros estudios han descrito las modificaciones en la regulación de genes de bacterias de importancia clínica, responsables de infecciones en humanos o animales, y su asociación con implicaciones importantes en la gravedad de una enfermedad. Un ejemplo de lo anterior es el reportado por Shah en el 2013 [8], quien describió el transcriptoma global de la bacteria *Salmonella enterica* serovar Enteritidis en cepas de aves de corral. La *Salmonella* es una bacteria causante de gastroenteritis bacteriana en todo el mundo, transmitida por alimentos. En su trabajo, Shah estudió 252 genes expresados de manera diferencial en cepas con baja

patogenicidad en comparación con su expresión en cepas con alta patogenicidad y comprobó que los genes con baja patogenicidad mostraron un perfil transcripcional único que se caracterizó por la expresión reducida de genes involucrados en la virulencia (genes fimbriales, de motilidad y asociados al estrés). Caracterizar este tipo de patrones de regulación ayudaría a diversos investigadores a identificar cepas de virulencia relevantes, que posteriormente podrían ser atenuadas y ser candidatas de vacunas contra la bacteria [8].

Todo lo anterior demuestra que reportar redes de regulación transcripcional de bacterias tiene implicaciones importantes para el conocimiento biológico, biomédico y clínico. Además, pueden tener aplicación en el área farmacéutica con el posible desarrollo de vacunas o la identificación de genes con funciones en virulencia que posteriormente sean blanco de fármacos.

## 2. Bases de datos y curación de redes de regulación transcripcional

Hoy en día, para tener acceso a TRNs, existen bases de datos de conocimiento biológico en plataformas digitales. Algunos ejemplos son: RegulonDB [9] y EcoCyc [10], que reportan la TRN de *E. coli* K-12, organismo modelo para estudios en microbiología; RegulomePA, que reporta la TRN de *Pseudomonas aeruginosa* PAO1 [11]; y SalmoNet2 [12], que reporta TRNs de varias cepas de *Salmonella*; entre otras bases de datos. Su objetivo es integrar el conocimiento de los diferentes mecanismos reguladores de las bacterias en redes cada día más completas. Sin embargo, a pesar de estos y otros esfuerzos, las TRNs publicadas hasta hoy aún están incompletas o son inexistentes para muchas bacterias [13]. Lo anterior puede deberse a que aún no se han identificado experimentalmente las interacciones de regulación para una gran cantidad de bacterias o, por otro lado, a que los esfuerzos sostenidos de curación de la literatura que extraiga sus TRNs, las organice y brinde acceso en base de datos se ha hecho en pocas bacterias [14].

Es así que uno de los retos más importantes para la publicación de TRN es su creación manual. A partir de enormes bases de datos de literatura biomédica, como PubMed [15], tradicionalmente se realiza una extracción manual de información de esa literatura, tarea que se define como *curación de literatura biomédica*. Ésta consiste en la identificación de las publicaciones pertinentes y en la extracción de la información, de manera selectiva, para su

despliegue en forma organizada en bases de datos [16]. Dicha curación manual permite obtener la información de forma muy precisa, pero presenta la desventaja de ser demandante y requerir mucho tiempo [17]-[18]. Además, enfrenta el reto de una acelerada producción científica [19]. Dadas estas limitaciones y retos, desde hace tiempo, ha sido indispensable la adopción de aproximaciones automáticas para facilitar e impulsar el trabajo de curación de literatura biomédica [20]-[21].

El presente trabajo va en esa dirección y busca facilitar la creación de TRNs de bacterias usando aproximaciones novedosas de IA, como son los *transformers* pre-entrenados bidireccionales, llamados BERT (*Bidirectional Encoder Representations from Transformers*). Por ello, hemos participado por dos años en los *Proyectos de investigación en Inteligencia Artificial en el Espacio de Innovación UNAM-HUAWEI*, desarrollando modelos para aplicarlos en la extracción de conocimiento de literatura biomédica. Gracias al apoyo del *Espacio de Innovación UNAM-HUAWEI* de la *Alianza para Promover el Desarrollo de Capacidades Digitales en México*, en su convocatoria 2022, logramos obtener un modelo de IA para la extracción de TRNs usando artículos científicos. Esto se logró mediante la afinación (*fine-tuning*) y evaluación de modelos tipo BERT.

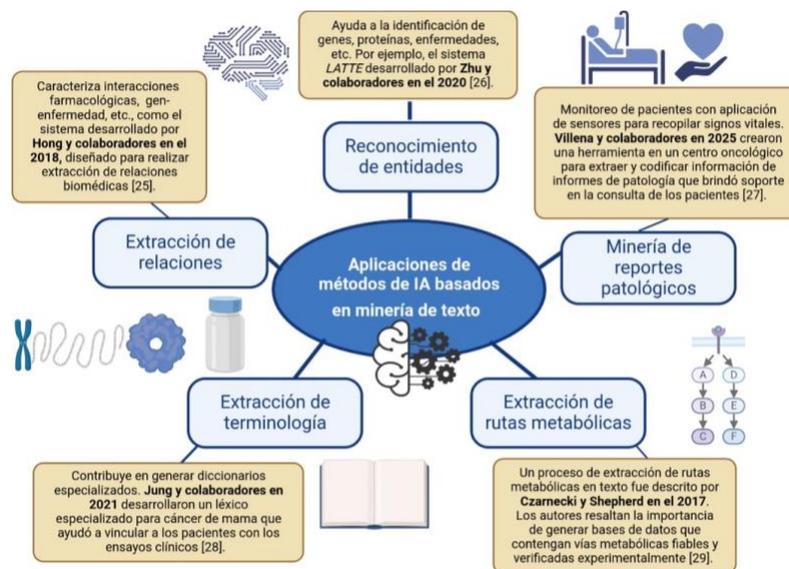
### 3. La IA en la extracción de conocimiento de literatura biomédica

Nuevas oportunidades se han creado para la curación de literatura biomédica gracias a las tecnologías de Inteligencia artificial (IA), en particular, aquellas basadas en métodos de aprendizaje profundo (*deep learning*). Un ejemplo de lo anterior es el chatbot de la empresa OpenAI, ChatGPT (*Chat Generative Pretrained Transformer*), que tiene la capacidad de escribir ensayos, responder preguntas e imitar una conversación humana, aunque presenta la desventaja de que, a menudo, la información que ofrece es incorrecta o irreproducible. Es posible que las actualizaciones continuas y el mejoramiento de las técnicas basadas en órdenes (*prompts*) ofrezcan el potencial para un análisis de conocimiento biomédico más preciso y confiable en el futuro [22].

Las herramientas de minería de texto basadas en IA aprovechan las metodologías del Procesamiento del lenguaje natural (*Natural Language Processing*, NLP) para facilitar el reconocimiento de entidades y la extracción de relaciones que se describen en la literatura biomédica. Por ejemplo, en la literatura publicada al respecto, el gen "BRCA1" es descrito

como una entidad relacionada con la enfermedad “cáncer” y es asociado con un mayor riesgo de la misma [23]. Mediante el NLP, se podrían extraer automáticamente de la literatura estas entidades (genes y enfermedades) y sus relaciones para aumentar el número de éstas relacionadas entre sí, reportadas en bases de datos de asociaciones gen-enfermedad.

Dos tareas típicas del NLP para la extracción de conocimiento de literatura biomédica son el reconocimiento de entidades nombradas (*Named Entity Recognition*, NER) y la extracción de relaciones (*Relation Extraction*, RE) [24]. La primera tarea consiste en identificar y etiquetar nombres de entidades biológicas o biomédicas, como proteínas, genes, fármacos o enfermedades, en un corpus de literatura biomédica. El NER tiene aplicación en la extracción de conceptos biológicos clave, lo que ayuda a construir ontologías y bases de conocimiento. La segunda tarea detecta si entidades presentes en la literatura tienen alguna relación, como las interacciones proteína-proteína, asociaciones gen-enfermedad, relaciones genotipo-fenotipo o interacciones farmacológicas (fármaco-enfermedad, fármaco-efecto secundario). Debe notarse que, además de texto, la literatura biomédica también contiene conocimiento valioso en forma de figuras, diagramas y rutas metabólicas sobre eventos moleculares que conducen a un proceso biológico o enfermedad [22]. En la figura 1, se pueden observar algunas aplicaciones de los métodos de IA basados en minería de texto.



**Fig. 1.** Aplicaciones de los métodos de IA basados en minería de texto. En color azul, se muestran las áreas en las se puede aplicar la IA y, en color naranja, se describe un ejemplo desarrollado en la biología o medicina. Imagen creada con Biorender.com [30].

## 4. Desarrollo

### 4.1. Materiales y métodos

#### 4.1.1. Transformers pre-entrenados tipo BERT

Las arquitecturas de redes de neuronas artificiales para NLP, entre ellas los *transformers* pre-entrenados bidireccionales BERT [31] y los generativos GPT (*Generative Pre-trained Transformer*) [32], están en pleno auge en la industria de la IA. Aunque los modelos GPT se han convertido en los más populares, los modelos BERT siguen siendo su principal competidor. La arquitectura BERT fue desarrollada en 2018 por Google [31]. En comparación con otras arquitecturas, presenta la ventaja de leer simultáneamente ambas direcciones del texto (bidireccionalidad). Leer de izquierda a derecha o de derecha a izquierda es importante porque puede ayudar al modelo a captar mejor el contexto de una palabra objetivo en una oración. Los modelos BERT son de acceso abierto, contrario a los GPT de la empresa OpenAI. El comprender las arquitecturas de los modelos y las características de cada uno puede ayudar a utilizarlos de mejor manera, por ejemplo, los modelos GPT son aprovechados para interactuar por medio de chatbots o bien utilizarse eficientemente para traducción. Por su parte, los modelos BERT se recomiendan para el entendimiento de lenguaje natural (*natural language understanding*), al igual que son recomendados para interpretar emociones u opiniones en texto (*sentiment analysis*) [33].

En su momento, un modelo BERT pre-entrenado con 800 millones de palabras del BookCorpus y 2500 millones de palabras de la Wikipedia en inglés fue lanzado a la comunidad [31]. Después, gracias a que fue una arquitectura abierta, diversos grupos de investigación entrenaron y publicaron variantes de este modelo, ya sea cambiando algún mecanismo de la arquitectura o aumentando los textos de entrenamiento. Sin embargo, la característica de estos *transformers* que más llamó nuestra atención fue su capacidad para ser afinado (*fine-tuning*) con pocos datos y para tareas específicas de NER o RE. El *fine-tuning* consiste entonces en aprovechar los pesos ya ajustados del modelo BERT durante la fase de pre-entrenamiento y sólo ajustar una parte de la arquitectura encargada de clasificar palabras (*tokens*) u oraciones de entrada, utilizando sólo pocos datos.

En este trabajo, comparamos la afinación de modelos de seis arquitecturas tipo BERT: BERT-base [31], BioBERT [34], BioLinkBERT [35], BioMegatron [36], BioRoBERTa [37] y LUKE [38]. Los modelos fueron afinados con búsqueda por *grid* de dos hiperparámetros:

*batch size*: 10, 16, 32, 64.

*learning rate*: 1e-5, 3e-5, 5e-5.

Para obtener los mejores hiperparámetros, utilizamos una estrategia de parada anticipada (*early stopping*) cuando la pérdida de entropía cruzada en el conjunto de datos de validación no mejorara en dos épocas (*epochs*). El modelo final fue creado uniendo los conjuntos de datos de entrenamiento y validación; luego fue evaluado con los datos de prueba y las métricas más usadas en problemas de clasificación: *F1-score*, *Precision*, *Recall* y el coeficiente de correlación de Matthews (*Matthews correlation coefficient*, MCC).

#### 4.1.2. Datos de *E. coli* para entrenamiento, validación y prueba

Un conjunto de 1562 oraciones de literatura de *E. coli* K-12, en las que se describen interacciones de regulación transcripcional entre TFs y genes, se clasificó manualmente en 4 categorías de acuerdo al tipo de efecto regulatorio del TF sobre el gen: activación, represión, regulación y sin relación. Este conjunto de datos fue dividido en conjuntos para entrenamiento (999 oraciones), validación (250 oraciones) y prueba (313 oraciones). Estas oraciones fueron proporcionadas para nuestro estudio, sin clasificación en las cuatro categorías, por el equipo de curación de RegulonDB, la base de datos con la principal TRN disponible públicamente para *E. coli* [39]. RegulonDB comprende una gran colección de interacciones regulatorias seleccionadas manualmente durante décadas a partir de la literatura biomédica y, en su momento, su equipo de curación recopiló el conjunto de 1562 oraciones.

#### 4.1.3. Datos de *Salmonella* para la extracción de una TRN

El equipo de curación de RegulonDB también nos proporcionó un conjunto de 3005 oraciones con interacciones de regulación de la bacteria *Salmonella enterica* serovar Typhimurium (*Salmonella*), que curó manualmente de 264 artículos de literatura biomédica. *Salmonella* es uno de los principales patógenos, a nivel mundial, que infectan humanos y animales [40] y su regulación transcripcional ha sido estudiada para enfrentar problemas de resistencia a antimicrobianos [41]. Este conjunto de oraciones fue utilizado para evaluar la

extracción automática de nuestro mejor modelo BERT, usando los mismos 264 artículos curados manualmente y comparando las interacciones curadas con las interacciones extraídas. Estos artículos fueron seleccionados por el equipo de RegulonDB, nosotros sólo los recopilamos y procesamos (ver sección *Evaluación de la extracción de una TRN de Salmonella*). Un análisis desde la perspectiva biológica de la TRN extraída también fue realizado.

## 4.2. Resultados y Discusión

### 4.2.1. Mejor modelo BERT afinado

Encontramos que el mejor modelo BERT afinado con las oraciones de *E. coli* fue de tipo LUKE con una entropía cruzada de 0.4024 y un *F1-Score Macro* de 0.9107, ambos medidos en el conjunto de validación. Es interesante que el modelo LUKE superara a los modelos BERT de dominio biomédico; LUKE es especializado en tarea (*task-specific*) y no en dominio (*domain-specific*). Consideramos que los mecanismos de auto-atención enfocados en entidades (*entity-aware self-attention*) del modelo LUKE fueron relevantes para lograr un mejor aprendizaje [38].

Este modelo logró un rendimiento significativo en el conjunto de datos de prueba con *Precision* de 0.8601 y *Recall* de 0.8788 (*F1-Score*: 0.8685), lo que indica que, cuando el modelo predice un tipo de regulación para una interacción, el 86% de las veces es correcto y que el 87% de las oraciones totales tendrá asignada correctamente su tipo de regulación. El modelo obtuvo un MCC de 0.8163. Este coeficiente indica la correlación entre los tipos de regulación de las oraciones del conjunto de prueba y los tipos predichos por el modelo; dado que este coeficiente va de -1 a +1, nuestro mejor modelo obtuvo una correlación sobresaliente.

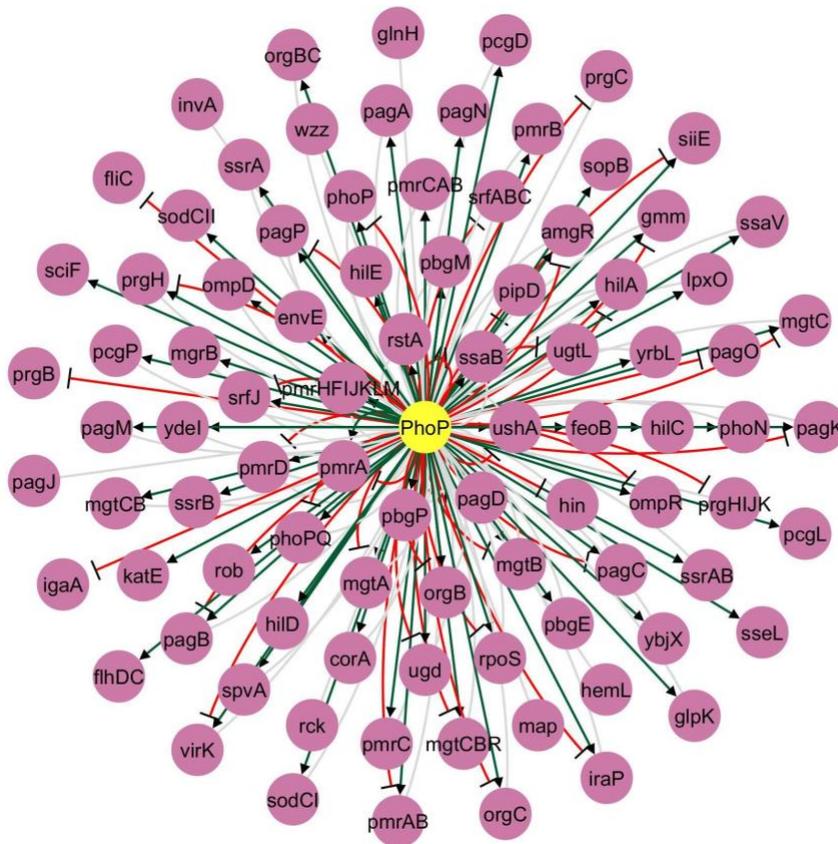
Finalmente, resaltamos que el rendimiento de nuestro mejor modelo fue significativo a pesar del desbalance de datos en las categorías de regulación de las oraciones del conjunto de prueba (activación: 113, represión: 56, regulación: 44 y sin relación: 100). De hecho, la segunda categoría mejor predicha fue regulación (*F1-Score*: 0.8915), no obstante, era la categoría con menos oraciones. Todas las categorías obtuvieron un *F1-Score* arriba de 0.82.

#### 4.2.2. Evaluación de la extracción de una TRN de *Salmonella*

Como mencionamos, el estudio nos permitió extraer automáticamente una TRN de la bacteria *Salmonella*, utilizando 264 artículos completos correspondientes a los datos curados que nos fueron proporcionados. Estos artículos fueron descargados en archivos PDF aprovechando el acceso abierto (*open-access*) o el acceso dado por la UNAM. Luego, los archivos PDF fueron convertidos en archivos de texto plano con una herramienta desarrollada por el Programa de Genómica Computacional del CCG-UNAM y separados en oraciones con la herramienta *Stanford CoreNLP* [42]. El reconocimiento de genes y TFs en estas oraciones fue realizado mediante una lista de ambas entidades, proporcionada por el equipo de RegulonDB. Después, el conjunto de oraciones fue ingresado al modelo LUKE para su clasificación en alguno de los tipos de interacción (activación, represión o regulación) o sin relación. La comparación de esta TRN con la curada manualmente mostró que nuestro modelo pudo extraer correctamente el 82% de la red curada (*Recall*: 0.8217). Además, en una revisión de 158 oraciones seleccionadas aleatoriamente de aquellas predichas erróneamente por el modelo, pudimos constatar que 40% de estas oraciones fueron en realidad bien predichas y pudimos extraer 25 interacciones ausentes en la curación manual. Nuestro proyecto es uno de los pioneros en obtener un enfoque exitoso basado en modelos BERT para extraer este tipo específico de interacciones [43].

#### 4.2.3. Análisis biológico de la TRN de *Salmonella*

El estudio también nos permitió examinar, desde una mirada biológica, la TRN de *Salmonella* extraída por nuestro modelo. Utilizamos las 1826 interacciones regulatorias extraídas automáticamente para construir una red y visualizarla con Cytoscape [44]. Con este software, se hizo además un análisis que calculó la centralidad y el número de conexiones asociadas a los nodos de la red (*degree*). Observamos que el TF PhoP presentó el quinto valor mayor de centralidad (0.2642) y el valor más alto de *degree* (180), que indicó que PhoP presentó el mayor número de enlaces relevantes (*edges*) entre nodos en la red. Biológicamente, nos interesó seleccionar una comunidad relevante de la TRN para un análisis más profundo, por lo que la red del TF PhoP fue seleccionada. Esta comunidad (subred) contiene 90 nodos con 180 enlaces (Figura 2).



**Fig. 2.** Red de la comunidad del TF PhoP. Los nodos describen al factor de transcripción (amarillo) y genes regulados (rosa). Por otro lado, los enlaces (edges) se caracterizan por su efecto regulador: activación (verde), represión (rojo) y regulación (gris).

Ahora bien, la lista de genes de la comunidad de PhoP fueron los datos de entrada para el sistema PANTHER [45], con el fin de identificar sus procesos biológicos y características de función biológica mediante un análisis de sobrerrepresentación de elementos con la prueba exacta de Fisher. Al analizar los resultados arrojados por el sistema PANTHER, se encontraron dos procesos biológicos relevantes e interesantes: el primero asociado a la función de transportadores de membrana de iones de Magnesio, con un *P-value* asociado a la prueba de Fisher de  $8.31e^{-05}$  y FDR de  $3.88e^{-02}$ ; el segundo, a procesos de sistemas de transducción de señales de fósforo con un *P-value* de  $1.20e^{-05}$  y FDR de  $1.43e^{-02}$ .

Lo siguiente que decidimos realizar fue describir la función de PhoP. Lo primero que identificamos es que esta proteína regula el 3% de los genes de *Salmonella* y pertenece al

sistema de dos componentes PhoP/PhoQ [46]. PhoP tiene influencia en la virulencia junto con la proteína sensora PhoQ y se activa por múltiples señales o variaciones ambientales, incluyendo niveles bajos de  $Mg^{2+}$ , ciertos péptidos antimicrobianos y ácidos grasos insaturados de cadena larga. PhoQ promueve el estado fosforilado de PhoP, en consecuencia, la fosforilación de PhoP puede tener efectos transcripcionales sobre los genes y puede modificar la transcripción en *Salmonella* [46]. En futuras aplicaciones, los bioinformáticos, biólogos y médicos podrían generar estudios directos en este TF en el área de farmacología, agronomía y clínica, entre otros.

#### 4.2.4. Sigüientes pasos en el estudio

Como vimos, los avances recientes en las tecnologías de IA ofrecen el potencial de ayudar al proceso de curación e integración del conocimiento extraído de la literatura biomédica. Por ello, para la convocatoria 2023 de los *Proyectos de investigación en Inteligencia Artificial en el Espacio de Innovación UNAM-HUAWEI*, propusimos dar un avance más en el acceso y organización de conocimiento biomédico sobre la TRNs. El proyecto que estamos desarrollando ahora consiste en obtener un modelo BERT afinado para extraer automáticamente las condiciones en las que se dan las interacciones regulatorias bacterianas, usando miles de registros textuales tomados de la *Gene Expression Omnibus* (GEO) [47]. La GEO es un repositorio a gran escala de datos genómicos funcionales de tecnologías de alto rendimiento. Nuestro objetivo final es el reconocimiento y extracción de condiciones experimentales de crecimiento celular de miles de registros de la GEO usando el mejor modelo afinado. Recuperar las condiciones de crecimiento en las que se manejan los experimentos biológicos depositados en la GEO es crucial para contextualizar, reproducir y comparar la información de análisis posteriores.

Actualmente contamos con resultados preliminares en la extracción de las siguientes condiciones utilizadas en experimentos microbianos: Aireación, agitación, medio de cultivo, antecedentes genéticos, fases de crecimiento, suplementos del medio, densidad óptica, pH, temperatura y recipiente. Para llegar a estos resultados, hemos comparado tres arquitecturas tipo BERT: BERT-base, BioBERT y BioLinkBERT. Un modelo BERT-base afinado superó a los otros dos modelos de dominio biomédico (BioBERT y BioLinkBERT) logrando un *F1-score* relevante de 0.9424 (*Precision*: 0.9449, *Recall*: 0.9676) y un alto MCC de 0.9544 en un conjunto de datos de prueba. El valor obtenido del MCC mostró un rendimiento relevante a

pesar de la distribución desequilibrada de instancias en los diez diferentes tipos de condiciones. Nuestro enfoque se podrá eventualmente aplicar para identificar y anotar las condiciones de crecimiento de miles de experimentos de la GEO para respaldar los análisis posteriores de datos genómicos que realicen otros grupos de investigación [48].

## 5. Conclusión

Este trabajo presentó la extracción automática de interacciones regulatorias transcripcionales de dos bacterias de importancia biológica (*E. coli* y *Salmonella*) con el uso de aproximaciones de IA y mediante el procesamiento de literatura biomédica en el marco de los *Proyectos de investigación en Inteligencia Artificial en el Espacio de Innovación UNAM-HUAWEI*. Ya que nuestro trabajo resaltó la importancia del TF PhoP y sus interacciones regulatorias en *Salmonella*, nuestra aproximación puede ser de utilidad para el desarrollo futuro de nuevas aplicaciones en el área farmacéutica con estudios de resistencia farmacológica o en aplicaciones en el área industrial. Además, este trabajo deja claro el potencial de las aproximaciones actuales de IA, como los *transformers* pre-entrenados BERT, para la extracción de interacciones regulatorias transcripcionales de bacterias que aún carecen de una TRN curada y publicada. Nuestro estudio se suma a muchos otros que aprovechan las tecnologías asociadas a la IA para extraer, organizar y expandir el conocimiento biológico.

## Financiamiento

Investigación realizada gracias al Programa UNAM-PAPIIT IN219523.

## Agradecimientos

Nuestra gratitud por el soporte computacional, servicio y becas que otorgó el *Espacio de Innovación UNAM-HUAWEI* de la *Alianza para Promover el Desarrollo de Capacidades Digitales en México*. Nuestro agradecimiento al equipo de curación de RegulonDB y Sara Berenice Martínez Luna por los conjuntos de oraciones curadas de *E. coli* y *Salmonella*, respectivamente. Agradecemos enormemente a los revisores del manuscrito por sus comentarios, que ayudaron a mejorar la calidad y brindar robustez al artículo. También agradecemos a los Técnicos Académicos Víctor Del Moral y Alfredo Hernández por el soporte informático para realizar este estudio.

## Referencias

- [1] X. Fang, A. Sastry, N. Mih, D. Kim, J. Tan, J.T. Yurkovich, C.J. Lloyd, Y. Gao, L. Yang, and B.O. Palsson, "Global transcriptional regulatory network for Escherichia coli robustly connects gene expression to transcription factor activities," in Proc. Natl. Acad. Sci., vol. 114, no. 38, pp. 10286-10291 Sep. 2017. Available: <https://doi.org/10.1073/pnas.1702581114>
- [2] J.A. Freyre-Gonzalez and L.G. Trevino-Quintanilla, "Analyzing regulatory networks in bacteria," Nat. Educ., vol. 3, no. 24, 2010. [Online]. Available: <https://www.nature.com/scitable/topicpage/analyzing-regulatory-networks-in-bacteria-14426192/> [Accessed Feb. 28, 2025].
- [3] E. Balleza, L.N. López-Bojorquez, A. Martínez-Antonio, O. Resendis-Antonio, I. Lozada-Chávez, Y.I. Balderas-Martínez, S. Encarnación, and J. Collado-Vides, "Regulation by transcription factors in bacteria: beyond description," FEMS Microbiol. Rev., vol. 33, no.1, pp. 133-151, Jan. 2009. Available: <https://doi.org/10.1111/j.1574-6976.2008.00145.x>
- [4] M.M. Babu and S. A. Teichmann, "Evolution of transcription factors and the gene regulatory network in Escherichia coli," Nucleic Acids. Res., vol. 31, no. 4, pp. 1234-1244, Feb. 2003. Available: <https://doi.org/10.1093/nar/gkg210>

- [5] K. Yamamoto, "The hierarchic network of metal-response transcription factors in *Escherichia coli*," *Biosci. Biotechnol. Biochem.*, vol. 78, no. 5, pp. 737-747, May 2014. Available: <https://doi.org/10.1080/09168451.2014.915731>
- [6] C. Dalldorf, Y. Hefner, R. Szubin, J. Johnsen, E. Mohamed, G. Li, J. Krishnan, A.M. Feist, B.O. Palsson, and D.C. Zielinski, "Diversity of Transcriptional Regulatory Adaptation in *E. coli*," *Mol. Biol. Evol.*, vol. 41, no. 11, p. msae240, Nov. 2024. Available: <https://doi.org/10.1093/molbev/msae240>
- [7] L. Praski Alzrigat, D.L. Huseby, G. Brandis, and D. Hughes, "Fitness cost constrains the spectrum of *marR* mutations in ciprofloxacin-resistant *Escherichia coli*," *J. Antimicrob. Chemother.*, vol. 72, no. 11, pp. 3016-3024, Aug. 2017. Available: <https://doi.org/10.1093/jac/dkx270>
- [8] D.H. Shah, "RNA sequencing reveals differences between the global transcriptomes of *Salmonella enterica* serovar *enteritidis* strains with high and low pathogenicities," *Appl. Environ. Microbiol.*, vol. 80, no. 3, Feb. 2014. Available: <https://doi.org/10.1128/AEM.02740-13>
- [9] H. Salgado, S. Gama-Castro, P. Lara, C. Mejia-Almonte, G. Alarcón-Carranza, A.G. López-Almazo, F. Betancourt-Figueroa, P. Peña-Loredo, S. Alquicira-Hernández, D. Ledezma-Tejeida, L. Arizmendi-Zagal, F. Mendez-Hernandez, A.K. Diaz-Gomez, E. Ochoa-Praxedis, L.J. Muñiz-Rascado, J.S. García-Sotelo, F.A. Flores-Gallegos, L. Gómez, C. Bonavides-Martínez, V.M. Del Moral-Chávez, A.J. Hernández-Alvarez, A. Santos-Zavaleta, S. Capella-Gutierrez, J. L. Gelpi, and J. Collado-Vides, "RegulonDB v12.0: a comprehensive resource of transcriptional regulation in *E. coli* K-12," *Nucleic Acids Res.*, vol. 52, no. D1, p. D255-D264, Jan. 2024. Available: <https://doi.org/10.1093/nar/gkad1072>
- [10] P. D. Karp, S. Paley, R. Caspi, A. Kothari, M. Krummenacker, P. Midford, L.R. Moore, P. Subhraveti, S. Gama-Castro, V. Tierrafria, P. Lara, L. Muñiz-Rascado, C. Bonavides-Martinez, A. Santos-Zavaleta, A. Mackie, G. Sun, T.A. Ahn-Horst, H. Choi, M.W. Covert, J. Collado-Vides, and I. Paulsen "The EcoCyc Database (2023)," *EcoSal Plus*, vol. 11, no. 1, pp. 1-22, Dec. 2023. Available: <https://doi.org/10.1128/ecosalplus.esp-0002-2023>

- [11] E. Galán-Vásquez, B.C. Luna-Olivera, M. Ramírez-Ibáñez, and A. Martínez-Antonio, "RegulomePA: a database of transcriptional regulatory interactions in *Pseudomonas aeruginosa* PAO1," *Database (Oxford)*, vol. 2020, p. baaa106, Dec. 2020. Available: <https://doi.org/10.1093/database/baaa106>
- [12] M. Olbei, B. Bohar, D. Fazekas, M. Madgwick, P. Sudhakar, I. Hautefort, A. Métris, J. Baranyi, R.A. Kingsley, and T. Korcsmaros, "Multilayered Networks of SalmoNet2 Enable Strain Comparisons of the *Salmonella* Genus on a Molecular Level," *MSystems*, vol.7, no.4, p. e01493-21, Aug. 2022. Available: <https://doi.org/10.1128/msystems.01493-21>
- [13] J.M. Escorcía-Rodríguez, A. Tauch, and J.A. Freyre-González. "Abasy atlas v2.2: The most comprehensive and up-to-date inventory of meta-curated, historical, bacterial regulatory networks, their completeness and system-level characterization," *Comput. Struct. Biotechnol. J.*, vol. 18, pp. 1228-1237, Jan. 2020. Available: <https://doi.org/10.1016/j.csbj.2020.05.015>
- [14] S.R. Davies, "Working in biocuration: contemporary experiences and perspectives," *Database (Oxford)*, vol. 2025, p. baaf003, Feb. 2025. Available: <https://doi.org/10.1093/database/baaf003>
- [15] National Center for Biotechnology Information (NCBI), "Pubmed Central," National Library of Medicine (US). [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/>. [Accessed Feb. 28, 2025].
- [16] S. Burge, T.K. Attwood, A. Bateman, T.Z. Berardini, M. Cherry, C. O'Donovan, L. Xenarios, and P. Gaudet. , "Biocurators and biocuration: surveying the 21st century challenges," *Database (Oxford)*, vol. 2012, p. bar059, Mar. 2012. Available: <https://doi.org/10.1093/database/baaf003>
- [17] J. Guallar, L. Codina and E. Abadal, "La investigación sobre curación de contenidos: análisis de la producción académica," *Ibersid.*, vol. 14, no. 1., pp. 13-22, Jun. 2020. Available: <https://repositori.upf.edu/handle/10230/52168>
- [18] I.S. for Biocuration: "Distilling data into knowledge," *PLoS Biol.*, vol. 16, no. 4, p. 2002846, Apr. 2018. Available: <https://doi.org/10.1371/journal.pbio.2002846>

- [19] L. Bornmann, R. Haunschild, and R. Mutz, "Growth rates of modern science: a latent piecewise growth curve approach to model publication numbers from established and new literature databases," *Humanit. and Soc. Sci. Commun.*, vol. 8, no. 1, pp. 1-15, Oct. 2021. Available: <https://doi.org/10.1057/s41599-021-00903-w>
- [20] S. Ananiadou, P. Thompson, R. Nawaz, J. McNaught, and D. B. Kell, "Event-based text mining for biology and functional genomics," *Briefings Funct. Genomics*, vol. 14, no. 3, pp. 213-230, May. 2014. Available: <https://doi.org/10.1093/bfgp/elu015>
- [21] C.C. Huang and Z. Lu, "Community challenges in biomedical text mining over 10 years: success, failure and the future," *Briefings Bioinf.*, vol. 17, no. 1, pp. 132-144, Jan. 2016. Available: <https://doi.org/10.1093/bib/bbv024>
- [22] F. He, K. Liu, Z. Yang, M. Hannink, R.D. Hammer, M. Popescu, and D. Xu, "Applications of cutting-edge artificial intelligence technologies in biomedical literature and document mining," *Med. Rev.*, vol. 3, no. 3, pp. 200-204, Jun. 2023. Available: <https://doi.org/10.1515/mr-2023-0011>
- [23] M.A. Thorat and R. Balasubramanian, "Breast cancer prevention in high-risk women," *Best Pract. Res. Clin. Obstet. Gynaecol.*, vol. 65, pp. 18-31, May 2020. Available: <https://doi.org/10.1016/j.bpobgyn.2019.11.006>
- [24] S. Zhao, C. Su, Z. Lu, and F. Wang, "Recent advances in biomedical literature mining," *Briefings Bioinf.*, vol. 22, no. 3, May 2020. Available: <https://doi.org/10.1093/bib/bbaa057>
- [25] L. Hong, J. Lin, S. Li, F. Wan, H. Yang, T. Jiang, D. Zhao, and J. Zeng. "A novel machine learning framework for automated biomedical relation extraction from large-scale literature repositories". *Nat. Mach. Intell.*, vol. 2, no. 6, pp. 347-355, Jun. 2020. Available: <https://doi.org/10.1038/s42256-020-0189-y>
- [26] M. Zhun, B. Celikkaya, P. Bhatia, and C.K. Reddy, "Latte: Latent type modeling for biomedical entity linking," In *Proc. of the AAAI Conf. on Artificial Intelligence*, vol. 34, no. 05, pp. 9757-9764, Apr. 2020. Available: <https://doi.org/10.1609/aaai.v34i05.6526>
- [27] F. Villena, P. Báez, S. Peñafiel, M. Rojas, I. Paredes, and J. Dunstan, "Developing and Validating an Automatic Support System for Tumor Coding in Pathology Reports in

- Spanish," *JCO Clin. Cancer Inform.*, vol. 9, p. e2400124, 2025. Available: <https://doi.org/10.1200/CCI.24.00124>
- [28] E. Jung, H. Jain, A.P. Sinha, and C. Gaudioso, "Building a specialized lexicon for breast cancer clinical trial subject eligibility analysis," *Health Informatics J.*, vol. 27, no. 1, Jan. 2021. Available: <https://doi.org/10.1177/1460458221989392>
- [29] J.M. Czarnecki and A.J. Shepherd, "Metabolic Pathway Mining," *Methods Mol. Biol.*, vol. 2, pp. 139-158, Nov. 2016. Available: [https://doi.org/10.1007/978-1-4939-6613-4\\_8](https://doi.org/10.1007/978-1-4939-6613-4_8)
- [30] Created in BioRender. Available: <https://www.BioRender.com/>
- [31] J. Devlin, M.W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," In *Proc. of the 2019 Conf. of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Minneapolis, Minnesota, Association for Computational Linguistics, vol. 1, pp. 4171-4186, May. 2019. Available: <https://doi.org/10.18653/v1/N19-1423>
- [32] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving language understanding by generative pre-training," [Online]. 2018. Available: [https://cdn.openai.com/researchcovers/languageunsupervised/language\\_understanding\\_paper.pdf](https://cdn.openai.com/researchcovers/languageunsupervised/language_understanding_paper.pdf) [Accessed Feb. 28, 2025].
- [33] C. Mottes, "BERT vs GPT: Comparing the two most popular language models", 29 August, 2024. [Online]. Available: <https://blog.invgate.com/gpt-3-vs-bert> [Accessed Feb. 28, 2025].
- [34] J. Lee, W. Yoon, S. Kim, D. Kim, S. Kim, C.H. So, and J. Kang, "BioBERT: a pre-trained biomedical language representation model for biomedical text mining," *Bioinformatics*, vol. 36, no. 4, pp. 1234-1240, Feb. 2020. Available: <https://doi.org/10.1093/bioinformatics/btz682>
- [35] M. Yasunaga, J. Leskovec, and P. Liang, "LinkBERT: Pretraining Language Models with Document Links," In S. Muresan, P. Nakov, A. Villavicencio, Eds. *Proc. of the 60th Annual Meeting of the Association for Computational Linguistics*, Dublin, Ireland:

- Association for Computational Linguistics, vol. 1, pp. 8003–8016, Mar. 2022. Available: <https://aclanthology.org/2022.acl-long.551>
- [36] H.C. Shin, Y. Zhang, E. Bakhturina, R. Puri, M. Patwary, M. Shoeybi, and R. Mani, "BioMega-tron: Larger Biomedical Domain Language Model," In B. Webber, T. Cohn, Y. He, Y. Liu, Eds. *Proc. of the 2020 Conf. on Empirical Methods in Natural Language Processing (EMNLP)*, Online: Association for Computational Linguistics, pp. 4700–4706, Oct. 2020. Available: <https://aclanthology.org/2020.emnlp-main.379>
- [37] S. Gururangan, A. Marasovic, S. Swayamdipta, K. Lo, I. Beltagy, D. Downey, and N.A. Smith, "Don't Stop Pretraining: Adapt Language Models to Domains and Tasks," In D. Jurafsky, J. Chai, N. Schlueter, J. Tetreault, Eds. *Proc. of the 58th Annual Meeting of the Association for Computational Linguistics*, Online: Association for Computational Linguistics, pp. 8342–8360, Jul. 2020. Available: <https://aclanthology.org/2020.acl-main.740>
- [38] I. Yamada, A. Asai, H. Shindo, H. Takeda, and Y. Matsumoto, "LUKE: Deep Contextualized Entity Representations with Entity-aware Self-attention," In B. Webber, T. Cohn, Y. He, Y. Liu, Eds. *Proc. of the 2020 Conf. on Empirical Methods in Natural Language Processing (EMNLP)*, Online: Association for Computational Linguistics, pp. 6442–6454, Nov. 2020. Available: <https://aclanthology.org/2020.emnlp-main.523>
- [39] V.H. Tierrafría, C. Rioualen, H. Salgado, P. Lara, S. Gama-Castro, P. Lally, L. Gómez-Romero, P. Peña-Loredo, A.G. López-Almazo, G. Alarcón-Carranza, F. Betancourt-Figueroa, S. Alquicira-Hernández, J.E. Polanco-Morelos, J. García-Sotelo, E. Gaytan-Núñez, C.F. Méndez-Cruz, L.J. Muñiz, C. Bonavides-Martínez, G. Moreno-Hagelsieb, J.E. Galagan, J.T. Wade, and J. Collado-Vides, "RegulonDB 11.0: Comprehensive high-throughput datasets on transcriptional regulation in *Escherichia coli* K-12," *Microb. Genom.*, vol. 8, no. 5, p. 000833, May 2022. Available: <https://doi.org/10.1099/mgen.0.000833>
- [40] A. Fàbrega and J. Vila, "Salmonella enterica serovar Typhimurium skills to succeed in the host: virulence and regulation," *Clin. Microbiol. Rev.*, vol. 26, no. 2, pp. 308-341, Apr. 2013. Available: <https://doi.org/10.1128/cmr.00066-12>

- [41] J.S., Gunn and S.I. Miller, "PhoP-PhoQ activates transcription of pmrAB, encoding a two-component regulatory system involved in *Salmonella typhimurium* antimicrobial peptide resistance," *J. Bacteriol.* vol. 178, no 23, pp. 6857–6864, Dec. 1996. Available: <https://doi.org/10.1128/jb.178.23.6857-6864.1996>
- [42] C.D. Manning, M. Surdeanu, J. Bauer, J.R. Finkel, S. Bethard, and D. McClosky, "The Stanford CoreNLP Natural Language Processing Toolkit," in K. Bontcheva and J. Zhu, Eds. *Proc. of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, Baltimore, Maryland: Association for Computational Linguistics, pp. 55-60, Jun. 2014. Available: <https://aclanthology.org/P14-5010/>
- [43] A. Varela-Vega, A.B. Posada-Reyes, and C.F. Méndez-Cruz, "Automatic extraction of transcriptional regulatory interactions of bacteria from biomedical literature using a BERT-based approach," *Database (Oxford)*, vol. 2024, p. baae094, Aug. 2024. Available: <https://doi.org/10.1093/database/baae094>
- [44] D. Otasek, J.H. Morris, J. Boucas, A.R. Pico, and B. Demchank, "Cytoscape automation: empowering workflow-based network analysis," *Genome Biol.*, vol. 20, pp.1-15, Sep. 2019. Available: <https://cytoscape.org/>. [Accessed Feb. 28, 2025].
- [45] H. Mi, A. Muruganujan, J.X. Huang, D. Ebert, C. Mills, X. Gou, and P.D. Thomas, "Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0)," *Nat. Protoc.*, vol. 14, no. 3, pp. 703-721, Feb. 2019. Available: <https://doi.org/10.1038/s41596-019-0128-8>
- [46] E.A. Groisman, A. Duprey, and J. Choi, "How the PhoP/PhoQ system controls virulence and Mg<sup>2+</sup> homeostasis: lessons in signal transduction, pathogenesis, physiology, and evolution," *Microbiol. Mol. Biol. Rev.*, vol. 85, no. 3, p. e0017620, Aug. 2021. Available: <https://doi.org/https://doi.org/10.1128/mubr.00176-20>
- [47] T. Barrett, S.E. Wilhite, P. Ledoux, C. Evangelista, I.F. Kim, M. Tomashevsky, K.A. Marshall, K.H. Phillippy, P.M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C.L. Robertson, N. Serova, S. Davis, and A. Soboleva, "NCBI GEO: archive for functional genomics data sets–update," *Nucleic. Acids. Res.*, vol. 41, pp. D991-D995, Nov. 2012. Available: <https://doi.org/10.1093/nar/gks1193>

- [48] A. Varela-Vega, M. Ávila-García, A. León Mendoza, A.B. Posada-Reyes, Víctor H. Tierrafría, C.F. Méndez-Cruz, and J. Collado-Vides, "Automatic annotation of growth conditions from *Escherichia coli* experiments deposited in the Gene Expression Omnibus," Presented at the 19th *Conf. on Computational Intelligence methods for Bioinformatics and Biostatistics (CIBB 2024)*, Benevento, Italy, 2024.