Revista Internacional de Investigación e Innovación Tecnológica

Página principal: www.riiit.com.mx

Microplásticos en el ambiente: una revisión sobre su identificación mediante inteligencia artificial

Microplastics in the environment: a review of their identification using artificial intelligence

Huichapa-Tinoco, A.P.[a], Mejía-Estrella, I.A.[b], Recio-Colmenares, R.B.[a], Sulbarán-Rangel, B.[a]*, Zuñiga-Grajeda, V.[c]

[a] Departamento de Estudios del Agua y la Energía. Centro Universitario de Tonalá. Universidad de Guadalajara C. P. 45425 México. (Orcid: https://orcid.org/0009-0007-4557-4787, https://orcid.org/0000-0002-3649-0334, https://orcid.org/0000-0002-5758-6140*).

[b] Departamento de Sustentabilidad y Ciencias del Territorio. Centro Universitario de Tlajomulco. Universidad de Guadalajara C. P. 45641 México. (Orcid: https://orcid.org/0000-0003-4443-5957).

[c] Departamento de Ciencias de la Información y Desarrollos Tecnológicos. Centro Universitario de Tonalá. Universidad de Guadalajara C. P. 45425 México. (Orcid: https://orcid.org/0000-0002-8248-0604).

alejandra.huichapa0629@alumnos.udg.mx; ixchel.simone@gmail.com; roxana.recio@academicos.udg.mx; belkis.sulbaran@academicos.udg.mx*; virgilio.zuniga@academicos.udg.mx

Innovación tecnológica: Procesamiento de datos por aprendizaje automático y profundo.
Área de aplicación industrial: Contaminación ambiental.

Recibido: 13 octubre 2025
Aceptado: 31 diciembre 2025

Abstract

The objective of this study was to conduct a systematic review to collect, analyze, and examine the use of artificial intelligence (AI) in the identification and characterization of microplastics (MPs) in the environment. This area has recently experienced substantial growth because these investigations generate large amounts of data, and AI significantly facilitates the rapid and accurate analysis of such data. The accurate identification and classification of MPs are essential for understanding their distribution, assessing their environmental impact, and developing effective mitigation strategies. Therefore, an updated review of research on MP identification and classification is necessary. The review revealed that machine learning (ML) and deep learning (DL) algorithms applied to spectroscopic and synthetic polymer characterization data achieve accuracies greater than 90% in classifying different types of polymers, with algorithms such as support vector machines, random forests, and convolutional neural networks standing out. The combination of artificial intelligence (AI) with spectroscopic techniques is emerging as an essential tool for the

efficient identification and classification of microplastics (MPs), whose contamination has increased significantly due to poor plastic waste management and their accumulation in the environment. MP pollution poses a critical threat to ecosystems and human health due to its persistence, ability to transport contaminants, and ease of entry into food chains.

Keywords: Machine Learning, Deep Learning, Microplastic Pollution, Artificial Intelligence, Synthetic Polymers,

Resumen

El objetivo de esta investigación fue realizar una revisión sistemática para recopilar, analizar y explorar el uso de inteligencia artificial (IA) en la identificación y caracterización de microplásticos (MP) en el ambiente. Esto debido a que la IA es un campo que recientemente ha tenido un auge en esta área debido a que estas investigaciones generan grandes cantidades de datos, y la IA facilita significativamente el análisis rápido y preciso de estos volúmenes de datos. La identificación y clasificación precisas de los MP son esenciales para comprender su distribución, evaluar su impacto ambiental y desarrollar estrategias de mitigación eficaces. Es por ello que se hace necesario una revisión actualizada sobre investigaciones relacionadas con la identificación y clasificación de MP. La revisión permitió identificar que los algoritmos de aprendizaje automático (ML) y aprendizaje profundo (DL) aplicados a datos de espectroscopia y caracterización de polímeros sintéticos muestran precisiones mayores al 90% en la clasificación de diferentes tipos de polímeros destacando los algoritmos como las máquinas de vectores de soporte, los bosques aleatorios y las redes neuronales convolucionales. La combinación de la inteligencia artificial (IA) con técnicas espectroscópicas se perfila como una herramienta esencial para la identificación y clasificación eficiente de microplásticos (MP), cuya contaminación ha aumentado de manera significativa debido a la deficiente gestión de los residuos plásticos y a su acumulación en el medio ambiente. La contaminación por MP representa una amenaza crítica para los ecosistemas y la salud humana debido a su persistencia, capacidad para transportar contaminantes y facilidad de entrada en las cadenas alimentarias.

Palabras clave: Aprendizaje Automático, Aprendizaje profundo, Contaminación por microplásticos, Inteligencia Artificial, Polímeros sintéticos.

Introducción

Los plásticos son fracciones de hidrocarburos polimerizados que frecuentemente se complementan con aditivos que se utilizan para mejorar la flexibilidad y la durabilidad (resistencia a la oxidación/al fuego y no degradabilidad general) (Falandysz et al. 2024). Estos se encuentran en todos los ámbitos de nuestra vida, ya que es un material increíblemente versátil y económico (Kane & Clare 2019). Sin embargo, la mala gestión de los residuos plásticos es la principal causa de contaminación del entorno natural (Kumar et al. 2021).

Los desechos plásticos se degradan y se desintegran en pequeños fragmentos conocidos como microplásticos (MP). Los microplásticos son fragmentos plásticos con un tamaño inferior a los 5 mm (Kumar et al. 2021). Debido a eso hoy en día, los microplásticos están presentes en prácticamente todos los ambientes naturales, su presencia ha sido verificada en agua

(Castro-Jiménez et al. 2019; Kane & Clare 2019) en ecosistemas terrestres (B. Xu et al. 2020), en sedimentos marinos (Mejía-Estrella et al. 2023) y aire (Gasperi et al. 2018). Lo anterior se debe a que por su pequeño tamaño tienen una gran capacidad de migración, representando una amenaza potencial. Se ha confirmado que los MP son un vector accesible para transportar contaminantes orgánicos, metales pesados (como cadmio, aluminio, cobre, plomo, manganeso), bifenilos policlorados, plaguicidas, sustancias químicas disruptoras endocrinas, e hidrocarburos aromáticos policíclicos haciéndolos llegar a organismos vivos a través de las cadenas alimenticias (Auta et al. 2017; Zou et al. 2025).

Los MP pueden ingresar a los organismos mediante la absorción de partículas en el aire, a través de las cámaras branquiales de organismos acuáticos o por ingesta directa. Estas partículas pueden transferirse a niveles tróficos superiores, llegando incluso a los seres humanos. Además, los MP pueden tener un efecto tóxico debido a la liberación de aditivos plásticos, por estrés por ingestión (por ejemplo, bloqueo físico o gasto de energía para la ingestión) y por la exposición a contaminantes adheridos a ellos (Anderson et al. 2016; Auta et al. 2017). Por lo tanto, es evidente el riesgo potencial que los MP representan para los ecosistemas y la salud de los seres vivos.

Actualmente, las técnicas más utilizadas para la detección y clasificación de MP incluyen la microscopía óptica, la microscopía electrónica de barrido, la cromatografía de gases acoplada a espectrometría de masas, la espectroscopia infrarroja por transformada de Fourier (FTIR) y la espectroscopia Raman. De estas, FTIR es una de las más empleadas debido a que es una técnica rápida, no destructiva, de alto rendimiento y que no requiere preparación previa de muestras, al combinar espectroscopía infrarroja con microscopía óptica (Qin et al. 2024).

Recientemente, la combinación de técnicas espectroscópicas con modelos de inteligencia artificial (IA) ha comenzado a aplicarse en la detección y clasificación de MP (Qin et al. 2024; Su et al. 2023; Zou et al. 2025). Por ejemplo, la integración de datos espectroscópicos Raman con algoritmos de aprendizaje automático (ML) ha permitido clasificar hasta ocho tipos diferentes de polímeros, logrando precisiones superiores al 80% (Qin et al. 2024). De igual forma, se han explorado combinaciones de IA con espectroscopía de infrarrojo cercano (NIR) (Zou et al. 2025), espectroscopía infrarroja por transformada de Fourier de reflectancia total atenuada (ATR-FTIR) (Enyoh & Wang 2024), e incluso con microscopía electrónica de barrido (SEM) (Yurtsever (Yurtsever & Yurtsever 2019). El análisis quimiométrico de los datos espectrales mediante IA permite procesar grandes volúmenes de información de manera más rápida y eficiente (Qin et al. 2024).

En este contexto, el objetivo de la presente revisión es analizar las principales investigaciones publicadas durante los últimos diez años sobre el uso de algoritmos de inteligencia artificial en la identificación y clasificación de microplásticos, así como identificar los factores que influyen en la mejora del desempeño de estas herramientas. Se ha reportado revisiones sobre el uso y aplicación de IA que abordan de manera general la integración de esta con la identificación de MP y se enfocan más en el monitoreo y control de la contaminación ambiental con MP (Jin et al. 2024), esta revisión se enfoca específicamente en los aspectos metodológicos asociados a la identificación y clasificación de MP mediante IA y las técnicas de caracterización de polímeros. En particular, se realiza un análisis comparativo del desempeño de algoritmos de

aprendizaje automático y aprendizaje profundo en función del tipo de dato de entrada y de la técnica analítica de caracterización empleada, destacando la importancia del preprocesamiento de datos como un factor determinante en la exactitud y robustez de los modelos. También, se analizan de forma critica las ventajas, limitaciones y desafíos prácticos de cada enfoque, primordialmente en el análisis de muestras de MP recolectadas en el ambiente y que ha sido afectadas el envejecimiento, contaminación superficial y heterogeneidad de composición. Desde esta perspectiva, la revisión aporta criterios metodológicos que pueden orientar el diseño de estudios futuros y la selección informada de estrategias basadas en IA para el análisis de microplásticos.

Metodología

Esta revisión se realizó mediante la búsqueda sistemática e integradora de investigaciones en bases de datos científicas de amplio uso y acceso como son Web of Science y ScienceDirect, con el objetivo de recopilar, analizar y sintetizar artículos recientes sobre el uso de IA en la clasificación e identificación de MP en el ambiente. Se consideraron publicaciones correspondientes a los últimos diez años y la búsqueda bibliográfica se realizó utilizando operadores booleanos y comillas, mediante combinaciones de términos asociados a microplásticos e inteligencia artificial, entre ellos se destacan: “microplastics” AND/OR “artificial intelligence”, “machine learning” OR “deep learning”, así como “FTIR” AND “Raman”. Las consultas se efectuaron de forma manual, aplicando filtros por año de publicación (2015–2025), tipo de documento (artículos de investigación y revisiones) e idioma (inglés). Únicamente se consideraron estudios que abordaran de manera explícita la aplicación de algoritmos de inteligencia artificial en la identificación y clasificación de microplásticos, mientras que se descartaron aquellos trabajos enfocados en aspectos generales de la contaminación plástica o que no emplearan IA de forma directa. La selección definitiva se llevó a cabo tras la revisión de títulos, resúmenes y textos completos, priorizando estudios con descripciones metodológicas claras y resultados cuantitativos relevantes.

Generalidades de la Inteligencia Artificial

La simulación de la inteligencia humana mediante máquinas, en particular a través de sistemas computacionales, se conoce comúnmente como Inteligencia Artificial (IA). Esta disciplina integra conocimientos computacionales y grandes conjuntos de datos para resolver problemas complejos y apoyar la toma de decisiones (Thangagiri & Sivakumar 2024). Los antecedentes de la inteligencia artificial pueden rastrearse hasta 1936, cuando Alan Turing propuso el concepto de la máquina de Turing como un modelo teórico de computación. Posteriormente, John McCarthy en 1956, describió la inteligencia artificial como un campo de la informática orientado al desarrollo de sistemas capaces de reproducir, de manera computacional, ciertos procesos asociados al razonamiento humano formalizaron así el término de inteligencia artificial. En el ámbito de las aplicaciones prácticas, uno de los primeros hitos fue la implementación de Unimate, considerado el primer robot industrial, desarrollado por George Devol e incorporado en la línea de ensamblaje de General Motors en Nueva Jersey durante la década de 1950 (Sanbar 2025).

Estos desarrollos han impulsado la inteligencia artificial y esta ha evolucionado de forma significativa, un ejemplo es el surgimiento de agentes inteligentes, como asistentes digitales y los bots, hasta los desarrollos en donde el enfoque es más

avanzado y de análisis de datos. Entre estos destacan el aprendizaje automático o machine learning (Por sus siglas en ingles ML) y el aprendizaje profundo o deep learning (Por sus siglas en ingles DL), cuyas principales ramas y relaciones se muestran de manera esquemática en la Figura 1.

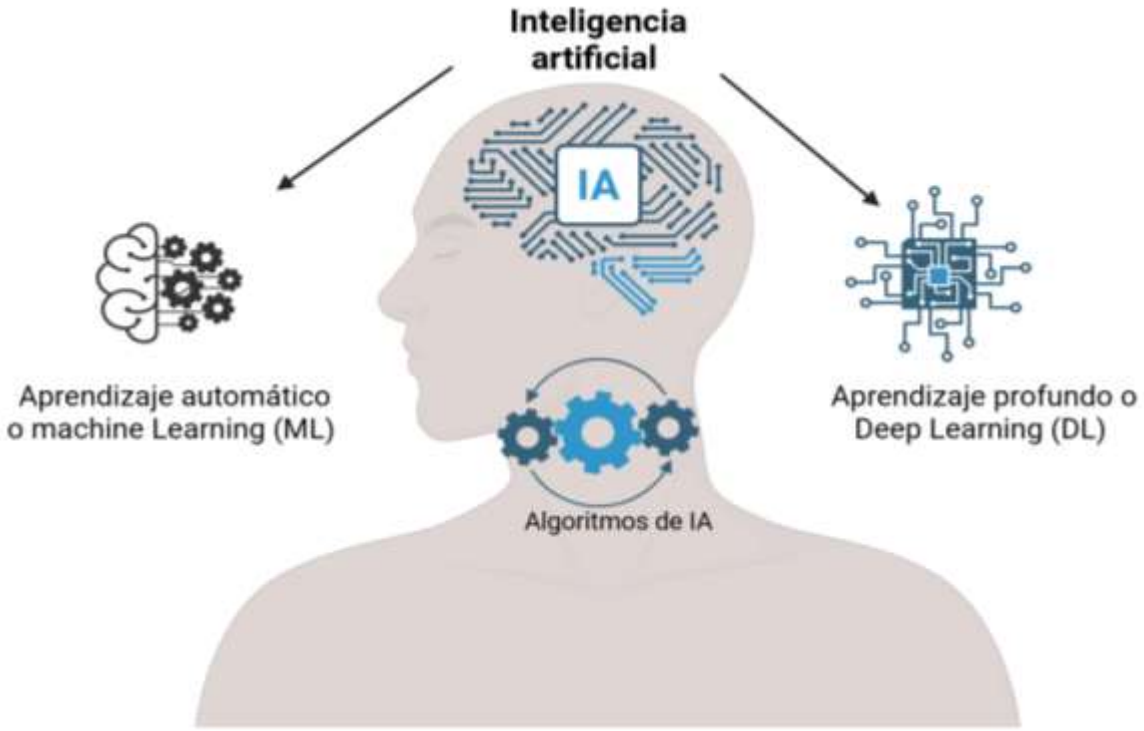


Figura 1. Ramas de la Inteligencia artificial.
Elaboración propia.

Uno de los componentes centrales de la inteligencia artificial es el aprendizaje automático ya que permite analizar de forma eficiente grandes cantidades de datos, tanto lineales como no lineales aplicando diferentes algoritmos. El ML es capaz de establecer relaciones entre variables, identificar patrones y extraer características relevantes, lo que facilita la realización de tareas como la predicción de comportamientos o propiedades y la clasificación automática (Su et al. 2023). Por su parte el aprendizaje profundo, se considera como una extensión del ML, ya que se basa en el uso de redes neuronales con múltiples capas ocultas. Estas arquitecturas permiten aprender representaciones complejas de los datos mediante procesos iterativos de entrenamiento, lo que ha favorecido su adopción en diversas áreas de la ciencia y la ingeniería (Qin et al. 2024).

Inteligencia artificial para la identificación y clasificación de MP

El uso de la IA para la identificación y clasificación de MP tiene pocos años, el primer artículo científico relacionado con ellos se publicó en 2018 (Lorenzo-Navarro et al. 2018). En dicha investigación, se presentó un método para el recuento y clasificación automática de MP mediante el análisis de imágenes de las muestras y algoritmos de ML. Los resultados mostraron que, utilizando características basadas en color y forma, junto con el algoritmo de ML Random Forest (RF), se logró una precisión del 96.6% en la identificación de cuatro tipos de partículas:

pellets, fragmentos, alquitrán y línea (fibras) (Lorenzo-Navarro et al. 2018). En la Figura 2 se ilustra el proceso general para la identificación y clasificación de MP mediante IA.

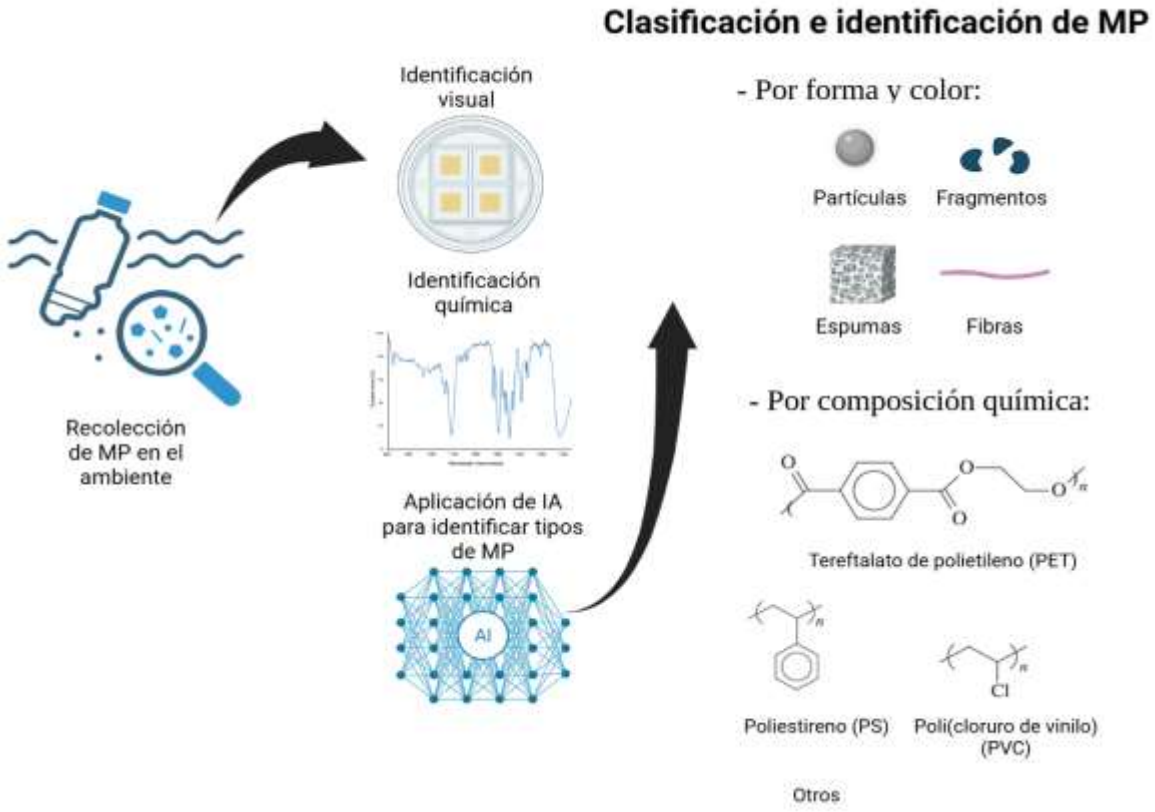


Figura 2. Proceso general de identificación y clasificación de MP mediante IA.
Elaboración propia.

Utilizando espectros o imágenes de los MP, el ML puede lograr una identificación y clasificación en función de las diferentes características de los MP como picos característicos, las huellas espectrales, su forma o color (Su et al. 2023). Además, el ML ha sido utilizado para la predicción de la citotoxicidad de los MP (J.-L. Xu et al. 2022) y para predecir su capacidad de adsorción de iones de metales pesados, incluidos Cd, Pb, Cr, Cu y Zn, en MP en los entornos acuáticos (Guo & Wang 2021).

Diversos investigadores han reportado el uso de algoritmos de ML y DL para la detección y clasificación de MP. En un estudio, se compararon varios modelos de clasificación de ML y DL utilizando espectros FTIR de MP provenientes del mar Mediterráneo, con el objetivo de identificar 13 tipos de polímeros. El algoritmo de ML, *Máquina de vectores de soporte* (SVM), fue seleccionado como clasificador final debido a su adecuada combinación entre simplicidad y rendimiento, proporcionando una caracterización automática rápida y precisa de MP con una exactitud del 94% (De Medeiros-Back et al. 2022). En otra investigación, se alcanzó la mayor precisión en la clasificación de MP mediante la utilización del modelo de DL, *red neuronal convolucional unidimensional* (1D-CNN), utilizando datos de espectroscopia Raman, el modelo alcanzó una precisión promedio del

95.8% para datos sin procesar y del 95.5% para datos preprocesados, identificando 10 tipos de polímeros (W. Zhang et al. 2023). En la Tabla 1 se presentan los algoritmos que han sido aplicados recientemente a la detección y clasificación de microplásticos, empleando distintos tipos de datos de entrada obtenidos mediante técnicas analíticas de caracterización de materiales. Entre estas se incluyen la espectroscopia de infrarrojo cercano (NIR), la espectroscopia Raman, la espectroscopia infrarroja por transformada de Fourier con reflectancia total atenuada (ATR-FTIR), la espectroscopia infrarroja por transformada de Fourier (FTIR), la combinación de espectroscopia NIR con imagen hiperespectral (NIR-HSI) y la microscopía electrónica de barrido (SEM).

Tabla 1. Algoritmos de IA para la detección y clasificación de MP.

Datos de entrada	Algoritmo utilizado	Referencia
NIR	Aprendizaje automático: • Partial Least Squares-Discriminant Analisis (PLS-DA) • Extreme Gradient Boost (XGBOOST) • Máquina de vectores de soporte o Support Vector Machine (SVM) • Bosque aleatorio o Random Forest (RF)	(Zou et al. 2025)
Raman	Aprendizaje automático: • Análisis discriminatorio lineal (LDA) • Árbol de decisiones o Decision Tree (DT) • Máquina de vectores de soporte o Support Vector Machine (SVM) Aprendizaje profundo: • One-dimensional convolutional neural network (1D-CNN)	(Qin et al. 2024)
ATR-FTIR	Aprendizaje automático: • Bosque aleatorio o Random Forest (RF) • Gerchberg-Saxton (GB) • Árbol de decisiones o Decision Tree (DT) • Vecinos más cercanos K-NN • De izquierda a derecho (LR) • Máquina de vectores de soporte o Support Vector Machine (SVM) • Multi-Layer Perceptron (MLP)	(Enyoh & Wang 2024)
Raman	Aprendizaje automático: • Vecinos más cercanos (K-NN) • Máquina de vectores de soporte o Support Vector Machine (SVM) • Perceptrón multicapa (MLP) • AlexNet Aprendizaje profundo: • One-dimensional convolutional neural network 1D-CNN	(W. Zhang et al. 2023)
NIR-HSI	Aprendizaje automático: • Modelado Independiente Suave por Analogía de Clases (SIMCA) • Análisis discriminante por mínimos cuadrados parciales (PLS-DA) • Máquina de vectores de soporte o Support Vector Machine (SVM) Aprendizaje profundo: • Redes Neuronales Artificiales (ANN)	(Goyetche et al. 2023)

ATR-FTIR	Aprendizaje automático: • Máquina de vectores de soporte o Support Vector Machine (SVM) • Vecinos más cercanos (K-NN) • Bosque aleatorio o Random Forest (RF) • Sistema de alerta de cambio de carril (LDA) • Multi-Layer Perceptron (MLP) • Análisis discriminante por mínimos cuadrados parciales (PLSDA) • Modelado Independiente Suave por Analogía de Clases (SIMCA) Aprendizaje profundo: • Redes neuronales (NN)	(Yan et al. 2022)
SEM	Red neuronal convolucional (CNN)	(Yurtsever & Yurtsever 2019)
ATR-FTIR	Vecinos más cercanos (K-NN)	(Kedzierski et al. 2019)
FTIR	Aprendizaje automático: • Árbol de decisiones o Decision Tree (DT) • Probabilística Gaussian Naive Bayes GNB • Vecinos más cercanos (K-NN) • Bosque aleatorio o Random Forest (RF) • De izquierda a derecho (LR) • Máquina de vectores de soporte o Support Vector Machine (SVC) Aprendizaje profundo: • Redes Neuronales (NN)	(De Medeiros-Back et al. 2022)
µ-FTIR y rojo de Nilo	Bosque aleatorio o Random Forest (RF)	(Meyers et al. 2024)
Raman	• Un solo modelo • Un modelo por clase	(Weber et al. 2023)

Los autores Zou et al. (2023) utilizaron espectroscopia NIR con varios algoritmos de ML, destacando que PLS-DA en donde obtuvo los resultados más robustos frente a variables como el color y tipo del polímero. En contraste, Qin et al. (2024) integraron espectroscopia Raman con modelos ML y DL, encontrando que el modelo 1D-CNN alcanzó una precisión del 97%, superando a los modelos ML tradicionales. Asimismo, Enyoh y Wang (2024) emplearon ATR-FTIR y lograron una exactitud del 99% en la clasificación de PET no degradado y envejecido utilizando algoritmos como RF y k-NN. De manera similar, Zhang et al. (2023) utilizaron Raman y DL para identificar diez polímeros, alcanzando una precisión cercana al 96%. Sin embargo, Goyetche et al. (2023), al aplicar NIR-HSI, reportaron limitaciones por falsos positivos y negativos, aunque el algoritmo SIMCA fue el más eficaz dentro de su estudio.

Otros estudios compararon el rendimiento de varios clasificadores. Yan et al. (2022), usando ATR-FTIR, concluyeron que el modelo SVM con reducción de dimensionalidad (PCA_SVM) fue el más eficiente para clasificar once polímeros. De Medeiros-Back et al. (2022), por su parte, evaluaron múltiples modelos con FTIR, seleccionando SVC como el más equilibrado en precisión y simplicidad (94% de exactitud). Técnicas alternativas como imágenes SEM (Yurtsever & Yurtsever 2019) y tinción con rojo de Nilo (Meyers et al. 2024) también demostraron resultados prometedores al combinarse con redes

neuronales y algoritmos como RF, logrando eficiencias superiores al 80% en ambientes biológicos. Por último, Weber et al. (2023) propusieron un modelo por clase con Raman, alcanzando precisiones superiores al 97% en la identificación de cinco polímeros comunes. En conjunto, los estudios evidencian que los modelos basados en DL, especialmente CNN y 1D-CNN, ofrecen un rendimiento superior en la clasificación de MP, aunque su implementación requiere mayor capacidad computacional. Mientras tanto, algoritmos de ML como SVM, PLS-DA y RF siguen siendo opciones robustas y confiables cuando se cuenta con bases de datos espectrales bien estructuradas.

También se han utilizado diversos parámetros fisicoquímicos, como el pH, la conductividad eléctrica, el carbono orgánico total (TOC), la salinidad, y las características de sedimentos de turberas (como las proporciones de arcilla, arena y limo), para combinarlos con modelos de ML en un intento de predecir la cantidad de MP presentes. Los resultados indicaron que la cantidad de MP estuvo influenciada principalmente por el pH, el TOC y la salinidad. Sin embargo, los modelos no fueron muy eficientes para predecir cantidades relativamente altas de MP presentes en las muestras de prueba (Tran et al. 2023). Además, se han llegado a utilizar fotografías tomadas con teléfonos móviles para contar y clasificar MP en clases dimensionales y morfológicas. El método utilizado hace uso tanto de técnicas de visión por computadora como de un algoritmo de ML, K-Nearest Neighbors (K-NN). Aunque la clasificación supervisada mostró un gran potencial, también presentó algunas deficiencias, especialmente en la clasificación correcta de algunos tipos específicos de partículas microplásticas, como los pellets (Massarelli et al. 2021).

A partir de los estudios revisados, y en concordancia con revisiones recientes de alcance general sobre la aplicación de la inteligencia artificial al análisis de microplásticos (Jin et al. 2024), es posible identificar fortalezas y limitaciones comunes en los distintos enfoques basados en aprendizaje automático y aprendizaje profundo desde una perspectiva metodológica centrada en la identificación y clasificación de microplásticos. Si bien numerosos trabajos reportan altas precisiones en estos procesos, el desempeño de los modelos depende de manera significativa tanto del tipo de dato de entrada como del contexto experimental. En este sentido, algoritmos de aprendizaje automático como la máquina de vectores de soporte, Random Forest y PLS-DA han mostrado un comportamiento robusto cuando se aplican a datos espectrales bien estructurados, particularmente en estudios basados en FTIR, Raman y NIR, con la ventaja adicional de requerir menor capacidad computacional y ofrecer una mayor interpretabilidad de los resultados (Back (De Medeiros-Back et al. 2022; Goyetche et al. 2023; Zou et al. 2025). En estos reportes sobre el uso de la IA se hace hincapié en que el desempeño de estos los algoritmos puede verse condicionado por factores propios de los datos analizados. Por ejemplo, si los espectros analizados tienen alto nivel de ruido o si los MP analizados estaban sometidos a procesos de envejecimiento, estos factores podrían afectar la capacidad de análisis de los modelos, sobre todo cuando el entrenamiento se basa en bibliotecas espectrales construidas a partir de polímeros vírgenes. Estas alteraciones suelen traducirse en una reducción de la exactitud y en un incremento de los errores de clasificación cuando los algoritmos se aplican directamente a muestras ambientales (Goyetche et al. 2023; Liu et al. 2023; Yan et al. 2022). Estos factores podrían transformarse en limitaciones relevantes que afectarían la aplicación de estos modelos de IA de condiciones de laboratorio a aplicaciones ambientales reales.

Por tanto, la elección del enfoque más adecuado de la aplicación de la IA requiere considerar no solo la exactitud reportada por los algoritmos, sino también la naturaleza de los datos disponibles, la complejidad de la muestra y los recursos computacionales accesibles. En el caso de los métodos de ML, las redes neuronales convolucionales, han mostrado una mayor flexibilidad de los datos de entrada (Qin et al. 2024; Yurtsever & Yurtsever 2019; Y. Zhang et al. 2023). Este comportamiento se asocia principalmente a su capacidad para aprender representaciones complejas de los datos de manera automática, reduciendo la dependencia de etapas extensas de preprocesamiento manual. Sin embargo, esta flexibilidad de estos modelos puede tener como desafíos la necesidad de mayor costo computacional, mayor conjunto de datos para el entrenamiento y una menor interpretabilidad de los resultados. Estos aspectos podrían limitar la adopción del ML en estudios ambientales caracterizados por matrices complejas y alta heterogeneidad (Enyoh & Wang 2024; Liu et al. 2023; Renner et al. 2019).

Otro aspecto importante en la clasificación e identificación es la presencia de biofilms microbianos y de contaminantes orgánicos e inorgánicos adheridos a la superficie de los microplásticos, esto constituye una fuente significativa de interferencia, particularmente en técnicas espectroscópicas como FTIR y Raman, donde dichos recubrimientos pueden distorsionar o enmascarar las bandas características de los polímeros, afectando tanto el preprocesamiento de los datos como la capacidad de los modelos para identificar correctamente (Enyoh & Wang 2024; Renner et al. 2019). Estas limitaciones se incrementan en muestras de MP procedentes de sedimentos, aguas residuales y muestras biológicas, donde la heterogeneidad y el ruido experimental son altos. En este contexto, distintos autores coinciden en la necesidad de desarrollar bases de datos más representativas de microplásticos degradados y ambientalmente contaminados, así como de incorporar estrategias de preprocesamiento y entrenamiento que permitan mejorar la robustez y transferibilidad de los modelos de inteligencia artificial en escenarios reales.

Preprocesamiento de datos

Antes de aplicar el algoritmo seleccionado para el análisis de datos, el primer paso es el preprocesamiento de estos datos, cuyo objetivo es eliminar elementos que interfieran con la propiedad química real de la muestra. En los datos espectrales y de imagen utilizados para el análisis de microplásticos es común encontrar alteraciones no deseadas, las cuales pueden agruparse, de manera general, en ruido, efectos multiplicativos, cambios en la línea base, presencia de datos faltantes y variaciones en la intensidad o forma de los picos. Estas distorsiones suelen estar asociadas a distintos factores, entre ellos el modo de medición empleado, las condiciones químicas y ambientales, así como el estado físico de la muestra analizada (Liu et al. 2023).

Frente a estas limitaciones, el preprocesamiento de los datos constituye una etapa fundamental y suele incluir, de forma combinada, la corrección de la línea base, la compensación de efectos multiplicativos, la reducción de ruido y el escalamiento de las variables. La normalización se utiliza con frecuencia para corregir variaciones en la línea base, mientras que técnicas como la eliminación de tendencia (DE), la variable normal estándar (SNV), la normalización por clasificación de variables (VSN) y la corrección de dispersión multiplicativa (MSC) se emplean para mitigar efectos multiplicativos. Asimismo, para la atenuación del ruido presente en espectros o imágenes químicas, se recurre habitualmente a métodos de suavizado como el filtro de Savitzky–

Golay (SG) (Liu et al. 2023). En algunos casos, los datos pueden tener diferentes escalas, lo que podría generar sesgos; para evitarlo, se pueden normalizar los datos utilizando escalamientos mínimo-máximo (Enyoh & Wang 2024). Desafortunadamente, en la mayoría de los estudios de análisis de MP, no se presta suficiente atención al proceso de preprocesamiento de espectros o imágenes (Liu et al. 2023; Renner et al. 2019).

El preprocesamiento de los datos constituye una etapa crítica en la aplicación de modelos de inteligencia artificial para la identificación y clasificación de microplásticos, ya que influye directamente en la calidad de la información utilizada para el entrenamiento y la generalización de los modelos. En técnicas espectroscópicas como FTIR y Raman, las etapas de corrección de la línea base y reducción de ruido resultan esenciales para minimizar interferencias asociadas a fenómenos como la fluorescencia, la dispersión de la señal y las variaciones instrumentales. Para ello se ha reportado que existen métodos de normalización como el de corrección de dispersión multiplicativa (MSC), variable normal estándar (SNV) y suavizado de Savitzky–Golay que contribuyen a mejorar la separación entre clases e incrementar la estabilidad de los modelos de aprendizaje automático, particularmente en algoritmos como SVM, Random Forest y PLS-DA (Enyoh & Wang 2024; Liu et al. 2023; Renner et al. 2019).

Es importante señalar que estos métodos de normalización cobran especial relevancia en datos provenientes de espectroscopia Raman, ya que la fluorescencia y el ruido de fondo suelen ser más pronunciados. Por tanto, la corrección de la línea base y un filtrado espectral permiten resaltar de manera más efectiva las bandas características de los polímeros. Un preprocesamiento adecuado de los datos de entrada no solo reduce la variabilidad no asociada al material, sino que también mejora el desempeño de modelos de ML y de redes neuronales convolucionales unidimensionales (Qin et al. 2024; W. Zhang et al. 2023). De manera similar, en espectros NIR, la aplicación de técnicas de escalamiento y normalización contribuye a mitigar efectos multiplicativos y variaciones relacionadas con el tamaño de partícula o el color, favoreciendo una clasificación más robusta mediante modelos como PLS-DA y SVM (Goyetche et al. 2023; Zou et al. 2025).

En el contexto de modelos de aprendizaje profundo, si bien estas arquitecturas presentan una mayor capacidad para extraer características relevantes de forma automática, distintos autores coinciden en que un preprocesamiento mínimo pero adecuado sigue siendo necesario para evitar el aprendizaje de ruido o patrones irrelevantes, especialmente cuando se trabaja con conjuntos de datos limitados o altamente heterogéneos (Liu et al. 2023; Renner et al. 2019). En conjunto, estos estudios evidencian que la selección de técnicas de preprocesamiento debe adaptarse tanto al tipo de dato espectral como al modelo de inteligencia artificial empleado, ya que no existe una estrategia universalmente óptima para todos los escenarios.

Ventajas y desventajas de los algoritmos de ML y DL

El estudio de la identificación y clasificación de MP mediante IA se encuentra todavía en sus primeras etapas de desarrollo (Jin et al. 2024; Thangagiri & Sivakumar 2024). Antes de aplicar un modelo de aprendizaje automático, es necesario contar con datos adecuados, seleccionar el algoritmo más conveniente, entrenar el modelo y validarlo correctamente. Dentro de este proceso, la elección del algoritmo es un punto clave (Zhu et al. 2022). A continuación, se describen los algoritmos más utilizados y algunas de sus

principales ventajas y desventajas, considerando su aplicación en la clasificación de MP:

- Máquina de vectores de soporte o Support Vector Machine (SVM): Es un algoritmo versátil compatible con datos numéricos, categóricos y de lenguaje natural. Presenta alta precisión, robustez y eficiencia de memoria. Su principal limitación es el alto coste computacional, debido al manejo de grandes cantidades de datos, así como la dificultad de ajustar sus parámetros (Qin et al. 2024; Zou et al. 2025).
- Bosque aleatorio o Random Forest, (RF): es un algoritmo de uso común, que combina el resultado de múltiples árboles de decisión para llegar a un resultado único. Ofrece una alta precisión y una buena resistencia al sobreajuste, lo que lo convierte en una alternativa adecuada para el análisis de grandes volúmenes de datos, incluso cuando el preprocesamiento es limitado. Su implementación puede resultar computacionalmente demandante, ya que requiere de un ajuste de hiperparámetros y puede generar tiempos de predicción altos cuando se emplean muchos árboles (Enyoh & Wang 2024; Zou et al. 2025).
- Vecinos más cercanos (K-Nearest Neighbors, K-NN): algoritmo que se destaca por su simplicidad conceptual y facilidad de implementación, dado que no requiere una fase explícita de entrenamiento del modelo. Aunque puede mostrar un comportamiento robusto frente a ciertos valores atípicos, su desempeño es sensible al ruido y a la presencia de variables irrelevantes. Además, su complejidad computacional aumenta de manera significativa cuando se aplica a conjuntos de datos de gran tamaño (Enyoh & Wang 2024; Y. Zhang et al. 2023).
- Árbol de decisiones o Decision Tree (DT): es fácil de entender e interpretar gracias a su estructura visual. Es útil para resolver problemas con múltiples salidas. No obstante, es propenso al sobreajuste y puede volverse inestable con pequeños cambios en los datos, además de requerir cálculos complejos (De Medeiros-Back et al. 2022; Enyoh & Wang 2024).
- Perceptrón multicapa o Multi-Layer Perceptron (MLP): una red neuronal clásica, es adecuada para modelar relaciones no lineales complejas. Aunque es muy flexible, puede sobreajustarse fácilmente, necesita un entrenamiento lento y requiere un cuidadoso ajuste de hiperparámetros para funcionar correctamente (Yan et al. 2022; W. Zhang et al. 2023).
- Análisis discriminante de mínimos cuadrados parciales o Partial Least Squares-Discriminant Analysis (PLS-DA): funciona bien con datos de alta dimensión y permite tareas tanto de regresión como clasificación, incorporando variables de respuesta. Sin embargo, puede presentar dificultades para interpretar los componentes obtenidos y, a menudo, necesita validación cruzada (Goyetche et al. 2023).
- Regresión logística o Logistic Regression (LR): es un modelo lineal sencillo, que consume pocos recursos computacionales y es fácil de implementar. Su desempeño disminuye con datos ruidosos o de alta dimensionalidad, y no es aplicable en problemas no lineales (Enyoh & Wang 2024).
- Aumento de gradiente extremo o Extreme Gradient Boosting

(XGBOOST): destaca por su alta precisión, flexibilidad y capacidad de regularización. Aun así, su principal desventaja es su complejidad computacional y el hecho de ser un modelo tipo "caja negra", dificultando la interpretación de sus predicciones (Zou et al. 2025).

- Redes neuronales artificiales o Artificial Neural Networks (ANN): son modelos versátiles que aceptan diversos tipos de datos, incluyendo texto. Su principal fortaleza es la alta flexibilidad en la resolución de distintos problemas, aunque son propensas al sobreajuste y requieren muchos recursos computacionales (Goyetche et al. 2023; Kedzierski et al. 2019).
- Redes neuronales convulocionadas o Convolutional Neural Networks (CNN): son ideales para trabajar con imágenes y funciones complejas. Si bien tienen un alto potencial en tareas de reconocimiento visual, requieren gran cantidad de datos de entrenamiento, tienen alta complejidad de cálculo y su interpretabilidad es limitada (Yurtsever & Yurtsever 2019; W. Zhang et al. 2023).

La elección del algoritmo más adecuado depende del tipo de dato disponible, la complejidad del problema y los recursos computacionales con los que se cuente. Modelos de aprendizaje automático como SVM, RF y PLS-DA han mostrado ser eficaces y eficientes para tareas con datos estructurados, especialmente cuando se busca interpretabilidad y rapidez de implementación. Por otro lado, los modelos de aprendizaje profundo como ANN y CNN son más adecuados para datos complejos como imágenes, y aunque requieren mayor capacidad de cómputo y entrenamiento, pueden alcanzar desempeños superiores. En general, no existe un algoritmo universalmente mejor; por ello, evaluar cuidadosamente el contexto experimental y los objetivos específicos del análisis es fundamental para lograr una clasificación confiable de microplásticos.

Evaluación del rendimiento de los modelos de ML

Para evaluar el rendimiento de los modelos de ML o DL, generalmente se emplean métricas de clasificación como la matriz de confusión, la exactitud, la precisión, la recuperación y el puntaje F1. Estas métricas proporcionan una comprensión integral de las fortalezas y debilidades de un modelo al clasificar con precisión los MP analizados (Enyoh & Wang 2024). La matriz de confusión es una herramienta ampliamente utilizada para resumir el desempeño de un modelo de clasificación, ya que organiza los resultados en cuatro categorías fundamentales: verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos, como se muestra en la Figura 3. Esta representación permite evaluar de manera directa cómo se comporta el modelo en cada una de las clases consideradas y sirve como base para el cálculo de distintas métricas de desempeño, entre ellas la exactitud, la precisión, la recuperación y el puntaje F1 (Enyoh & Wang 2024; Su et al. 2023; Yan et al. 2022; W. Zhang et al. 2023).

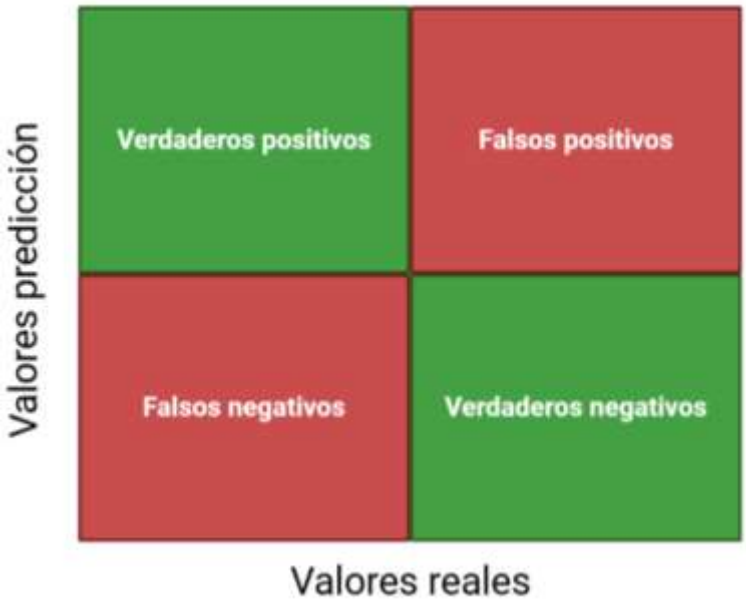


Figura 3. Representación de una Matriz de Confusión.
Elaboración propia.

Una vez que se obtiene la matriz de confusión, es posible calcular la exactitud (Ec. 1), definida como la proporción de observaciones correctamente clasificadas respecto al total de datos analizados. Por su parte, la precisión (Ec. 2) refleja la capacidad del modelo para identificar correctamente los casos positivos entre aquellos predichos como positivos, poniendo énfasis en la reducción de falsos positivos. Asimismo, la recuperación (Ec. 3) evalúa qué tan eficaz es el modelo para detectar los casos positivos reales dentro del conjunto de datos, siendo especialmente sensible a la presencia de falsos negativos. Finalmente, el puntaje F1 (Ec. 4) ofrece una medida equilibrada del rendimiento global del modelo, ya que integra la información aportada por la precisión y la recuperación mediante su media armónica (Enyoh & Wang 2024). La expresión matemática de cada una de estas métricas se muestra a continuación:

$$Exactitud = \frac{TP + TN}{TP + FP + TN + FN} \qquad (\boldsymbol{Ec.\,1})$$

$$Precisión = \frac{TP}{TP + FP} \qquad (\boldsymbol{Ec.\,2})$$

$$Recuperación = \frac{TP}{TP + FN} \qquad (\boldsymbol{Ec.\,3})$$

$$Puntaje\ F1 = 2 * \frac{Precisión * Recuperación}{Precisión + Recuperación} \qquad (\boldsymbol{Ec.\,4})$$

Donde:
TP: Verdaderos Positivos, TN: Verdaderos Negativos, FP: Falsos Positivos y FN: Falsos Negativos.

Perspectivas futuras del uso de IA en la identificación y clasificación de MP

La aplicación de la IA en la identificación y clasificación de MP presenta limitaciones ya que se encuentra aún en una fase de desarrollo, sin embargo, a pesar de ello existe un gran margen de oportunidad para mejorar su implementación y consolidarla como una herramienta confiable en estudios ambientales.

Una de las limitaciones más relevantes señaladas en la revisión realizada es la escasa disponibilidad de espectros en las bases de datos utilizadas para su identificación y clasificación representativos de MP en condiciones reales. Debido a que esas muestras presentan procesos de envejecimiento y degradación, así como la presencia de suciedad, biofouling o materia orgánica adherida a la superficie de las partículas, que pueden modificar de manera sustancial las bandas espectrales de los microplásticos. Como consecuencia, la firma química resultante puede diferir de forma significativa de la correspondiente a polímeros vírgenes, lo que reduce la eficacia de los métodos de detección automatizada basados en inteligencia artificial (De Medeiros-Back et al. 2022; Goyetche et al. 2023; Liu et al. 2023). En este sentido, varios autores advierten sobre los riesgos de emplear bibliotecas espectrales construidas exclusivamente a partir de polímeros puros, ya que estas pueden definir límites de clase excesivamente estrechos y comprometer la correcta identificación de microplásticos en muestras ambientales reales (Enyoh & Wang 2024; Goyetche et al. 2023; Y. Zhang et al. 2023). Por tanto, en concordancia con esto una de las principales líneas futuras de investigación se relaciona con el desarrollo y la disponibilidad de bases de datos abiertas y estandarizadas que incluyan espectros y datos de microplásticos envejecidos, contaminados y provenientes de distintas matrices ambientales (Goyetche et al. 2023; Liu et al. 2023).

Otro punto importante, seria lograr una automatización de los flujos de análisis, es decir que mediante herramientas de IA se pueda integrar la preparación de las muestras, la adquisición de datos espectroscópicos (FTIR, Raman y NIR) y el procesamiento de forma que se disminuya la intervención manual y por tanto se reduca el sesgo experimental y se acorte los tiempos de análisis. Esto podría ser posible si se realiza el desarrollo de protocolos reproducibles y escalables, capaces de adaptarse a la variabilidad inherente de las muestras ambientales. En este sentido, diversos estudios destacan que la combinación de técnicas espectroscópicas con modelos de ML y DL puede facilitar el monitoreo sistemático de microplásticos, siempre que se consideren estas limitaciones metodológicas desde la etapa de diseño experimental (Enyoh & Wang 2024; Renner et al. 2019; Zou et al. 2025).

Es importante señalar que para que se dé un uso más amplio de la IA en la clasificación e identificación de MP se debe mejorar el desarrollo de los modelos ya que, aunque los enfoques de ML han demostrado un alto desempeño en tareas de clasificación, su naturaleza de “caja negra” dificulta la interpretación fisicoquímica de los resultados y puede limitar su aceptación en contextos regulatorios o de monitoreo ambiental. En respuesta a estos desafíos, varios autores proponen que se debe mejorar y avanzar hacia modelos que logren un equilibrio entre capacidad predictiva e interpretabilidad, de modo que sea posible identificar qué características espectrales o morfológicas influyen en la clasificación de los MP (Jin et al. 2024; Liu et al. 2023; Renner et al. 2019). En conjunto, el fortalecimiento de estas líneas de investigación podría contribuir a consolidar el uso de la IA como una herramienta más robusta, transparente y confiable para el análisis de MP en escenarios reales.

Conclusión

En esta revisión se encontró que la integración de herramientas de inteligencia artificial como el ML, DL y redes neuronales junto con técnicas de espectroscopias como NIR, Raman, ATR-FTIR y FTIR pueden ser aplicadas de manera significativa para detectar e identificar microplásticos con alta

precisión. Debido a que los modelos de ML y DL permiten procesar grandes volúmenes de información espectral, lo que facilita la clasificación eficiente de distintos tipos de polímeros y mejora la interpretación de conjuntos de datos complejos.

A pesar de los avances alcanzados en las investigaciones recientes sobre la aplicación de IA en la identificación de MP en el ambiente, persisten retos relevantes que condicionan el desempeño de estos modelos Entre ellos, es el acceso a bases de datos abiertas y estandarizadas que incluyan espectros y datos de microplásticos envejecidos, contaminados y provenientes de distintas matrices ambientales y el preprocesamiento de los datos, estos aspectos constituyen etapas críticas, ya que una selección inadecuada de datos de entrada puede introducir interferencias que no reflejan las propiedades químicas reales de las muestras de MP y, en consecuencia, afectar negativamente la exactitud de la clasificación de la IA.

Otro aspecto importante, es que resulta necesario continuar perfeccionando los algoritmos de IA con un enfoque específico en la optimización de la identificación y clasificación de microplásticos. Diversos estudios han evidenciado que algoritmos como la máquina de vectores de soporte, el bosque aleatorio y el método de vecinos cercanos ofrecen un desempeño consistente, destacando por su robustez, exactitud y versatilidad en diferentes escenarios experimentales. A medida que se avance en la optimización de los algoritmos de ML y DL, la combinación de inteligencia artificial con técnicas espectroscópicas se consolida como una herramienta con alto potencial para abordar de manera más eficaz la problemática ambiental asociada a la creciente presencia de microplásticos.

Referencias

Anderson, J. C., Park, B. J., & Palace, V. P. (2016). Microplastics in aquatic environments: Implications for Canadian ecosystems. Environmental Pollution, 218, 269-280. https://doi.org/10.1016/j.envpol.2016.06.074

Auta, H. S., Emenike, C. U., & Fauziah, S. H. (2017). Distribution and importance of microplastics in the marine environment: A review of the sources, fate, effects, and potential solutions. Environment International, 102, 165-176. https://doi.org/10.1016/j.envint.2017.02.013

Castro-Jiménez, J., González-Fernández, D., Fornier, M., Schmidt, N., & Sempéré, R. (2019). Macro-litter in surface waters from the Rhone River: Plastic pollution and loading to the NW Mediterranean Sea. Marine Pollution Bulletin, 146, 60-66. https://doi.org/10.1016/j.marpolbul.2019.05.067

De Medeiros-Back, H., Vargas Junior, E. C., Alarcon, O. E., & Pottmaier, D. (2022). Training and evaluating machine learning algorithms for ocean microplastics classification through vibrational spectroscopy. Chemosphere, 287, 131903. https://doi.org/10.1016/j.chemosphere.2021.131903

Enyoh, C. E., & Wang, Q. (2024). Automated Classification of Undegraded and Aged Polyethylene Terephthalate Microplastics from ATR-FTIR Spectroscopy using Machine Learning Algorithms. Journal of Polymers and the Environment, 32(9), 4143-4158.

https://doi.org/10.1007/s10924-024-03199-4

Falandysz, J., Liu, G., & Rutkowska, M. (2024). Analytical progress on emerging pollutants in the environment: An overview of the topics. TrAC Trends in Analytical Chemistry, 175, 117719. https://doi.org/10.1016/j.trac.2024.117719

Gasperi, J., Wright, S. L., Dris, R., Collard, F., Mandin, C., Guerrouache, M., Langlois, V., Kelly, F. J., & Tassin, B. (2018). Microplastics in air: Are we breathing it in? Current Opinion in Environmental Science & Health, 1, 1-5. https://doi.org/10.1016/j.coesh.2017.10.002

Goyetche, R., Kortazar, L., & Amigo, J. M. (2023). Issues with the detection and classification of microplastics in marine sediments with chemical imaging and machine learning. TrAC Trends in Analytical Chemistry, 166, 117221. https://doi.org/10.1016/j.trac.2023.117221

Guo, X., & Wang, J. (2021). Projecting the sorption capacity of heavy metal ions onto microplastics in global aquatic environments using artificial neural networks. Journal of Hazardous Materials, 402, 123709. https://doi.org/10.1016/j.jhazmat.2020.123709

Jin, H., Kong, F., Li, X., & Shen, J. (2024). Artificial intelligence in microplastic detection and pollution control. Environmental Research, 262, 119812. https://doi.org/10.1016/j.envres.2024.119812

Kane, I. A., & Clare, M. A. (2019). Dispersion, Accumulation, and the Ultimate Fate of Microplastics in Deep-Marine Environments: A Review and Future Directions. Frontiers in Earth Science, 7 1-28. https://doi.org/10.3389/feart.2019.00080

Kedzierski, M., Falcou-Préfol, M., Kerros, M. E., Henry, M., Pedrotti, M. L., & Bruzaud, S. (2019). A machine learning algorithm for high throughput identification of FTIR spectra: Application on microplastics collected in the Mediterranean Sea. Chemosphere, 234, 242-251. https://doi.org/10.1016/j.chemosphere.2019.05.113

Kumar, M., Chen, H., Sarsaiya, S., Qin, S., Liu, H., Awasthi, M. K., Kumar, S., Singh, L., Zhang, Z., Bolan, N. S., Pandey, A., Varjani, S., & Taherzadeh, M. J. (2021). Current research trends on micro- and nano-plastics as an emerging threat to global environment: A review. Journal of Hazardous Materials, 409, 124967. https://doi.org/10.1016/j.jhazmat.2020.124967

Liu, Z., Wang, W., & Liu, X. (2023). Automated characterization and identification of microplastics through spectroscopy and chemical imaging in combination with chemometric: Latest developments and future prospects. TrAC Trends in Analytical Chemistry, 160, 116956. https://doi.org/10.1016/j.trac.2023.116956

Lorenzo-Navarro, J., Castrillón-Santana, M., Gómez, M., Herrera, A., & Marín-Reyes, P. A. (2018). Automatic counting and classification of microplastic particles. ICPRAM 2018 Proceedings of the 7th International Conference on Pattern Recognition Applications and Methods,, 646–652. https://doi.org/10.5220/0006725006460652

Massarelli, C., Campanale, C., & Uricchio, V. F. (2021). A Handy Open-Source Application Based on Computer Vision and Machine Learning Algorithms to Count and Classify Microplastics. Water, 13(15), 2104. https://doi.org/10.3390/w13152104

Mejía-Estrella, I. A., Peña-Montes, C., Peralta-Peláez, L. A., Del Real Olvera, J., & Sulbarán-Rangel, B. (2023). Microplastics in Sandy Beaches of Puerto Vallarta in the Pacific Coast of Mexico. Sustainability, 15(21), 15259. doi:https://doi.org/10.3390/su152115259

Meyers, N., Everaert, G., Hostens, K., Schmidt, N., Herzke, D., Fuda, J.-L., Janssen, C. R., & De Witte, B. (2024). Towards reliable data: Validation of a machine learning-based approach for microplastics analysis in marine organisms using Nile red staining. Marine Pollution Bulletin, 207, 116804. https://doi.org/10.1016/j.marpolbul.2024.116804

Qin, Y., Qiu, J., Tang, N., He, Y., & Fan, L. (2024). Deep learning analysis for rapid detection and classification of household plastics based on Raman spectroscopy. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 309, 123854. https://doi.org/10.1016/j.saa.2024.123854

Renner, G., Nellessen, A., Schwiers, A., Wenzel, M., Schmidt, T. C., & Schram, J. (2019). Data preprocessing & evaluation used in the microplastics identification process: A critical review & practical guide. TrAC Trends in Analytical Chemistry, 111, 229-238. https://doi.org/10.1016/j.trac.2018.12.004

Sanbar, S. S. (2025). 4 - History of Artificial Intelligence, Machine Learning, and Deep Learning. In M. American College of Legal & M. American Board of Legal (Eds.), Legal Medicine (Eighth Edition) (pp. 22-27.e21). New Delhi: Elsevier. https://doi.org/10.1016/B978-0-443-12128-9.00004-8

Su, J., Zhang, F., Yu, C., Zhang, Y., Wang, J., Wang, C., Wang, H., & Jiang, H. (2023). Machine learning: Next promising trend for microplastics study. Journal of Environmental Management, 344, 118756. https://doi.org/10.1016/j.jenvman.2023.118756

Thangagiri, B., & Sivakumar, R. (2024). Prospective Application of Artificial Intelligence Towards the Detection, and Classifications of Microplastics with Bibliometric Analysis. Water, Air, & Soil Pollution, 235(6), 362. https://doi.org/10.1007/s11270-024-07151-z

Tran, H.-T., Hadi, M., Nguyen, T. T. H., Hoang, H. G., Nguyen, M.-K., Nguyen, K. N., & Vo, D.-V. N. (2023). Machine learning approaches for

predicting microplastic pollution in peatland areas. Marine Pollution Bulletin, 194, 115417. https://doi.org/10.1016/j.marpolbul.2023.115417

Weber, F., Zinnen, A., & Kerpen, J. (2023). Development of a machine learning-based method for the analysis of microplastics in environmental samples using µ-Raman spectroscopy. Microplastics and Nanoplastics, 3(1), 9. https://doi.org/10.1186/s43591-023-00057-3

Xu, B., Liu, F., Cryder, Z., Huang, D., Lu, Z., He, Y., Wang, H., Lu, Z., Brookes, P. C., Tang, C., Gan, J., & Xu, J. (2020). Microplastics in the soil environment: Occurrence, risks, interactions and fate – A review. Critical Reviews in Environmental Science and Technology, 50(21), 2175-2222. https://doi.org/10.1080/10643389.2019.1694822

Xu, J.-L., Lin, X., & Gowen, A. A. (2022). Combining machine learning with meta-analysis for predicting cytotoxicity of micro- and nanoplastics. Journal of Hazardous Materials Advances, 8, 100175. https://doi.org/10.1016/j.hazadv.2022.100175

Yan, X., Cao, Z., Murphy, A., & Qiao, Y. (2022). An ensemble machine learning method for microplastics identification with FTIR spectrum. Journal of Environmental Chemical Engineering, 10(4), 108130. https://doi.org/10.1016/j.jece.2022.108130

Yurtsever, M., & Yurtsever, U. (2019). Use of a convolutional neural network for the classification of microbeads in urban wastewater. Chemosphere, 216, 271-280. https://doi.org/10.1016/j.chemosphere.2018.10.084

Zhang, W., Feng, W., Cai, Z., Wang, H., Yan, Q., & Wang, Q. (2023). A deep one-dimensional convolutional neural network for microplastics classification using Raman spectroscopy. Vibrational Spectroscopy, 124, 103487. https://doi.org/10.1016/j.vibspec.2022.103487

Zhang, Y., Zhang, D., & Zhang, Z. (2023). A Critical Review on Artificial Intelligence—Based Microplastics Imaging Technology: Recent Advances, Hot-Spots and Challenges. International Journal of Environmental Research and Public Health, 20(2), 1150. doi:https://doi.org/10.3390/ijerph20021150

Zhu, M., Wang, J., Yang, X., Zhang, Y., Zhang, L., Ren, H., Wu, B., & Ye, L. (2022). A review of the application of machine learning in water quality evaluation. Eco-Environment & Health, 1(2), 107-116. https://doi.org/10.1016/j.eehl.2022.06.001

Zou, H.-H., He, P.-J., Peng, W., Lan, D.-Y., Xian, H.-Y., Lü, F., & Zhang, H. (2025). Rapid detection of colored and colorless macro- and micro-plastics in complex environment via near-infrared spectroscopy and machine learning. Journal of Environmental Sciences, 147, 512-522. https://doi.org/10.1016/j.jes.2023.12.004