PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

ESCUELA DE POSGRADO

TEMPORAL ANALYSIS OF MACROPLASTIC WASTE CONCENTRATION IN PERUVIAN COASTAL RIVERS USING A DEEP LEARNING APPROACH

Tesis para optar el grado académico de Doctor en Ingeniería, que presenta:

Miguel Angel Astorayme Valenzuela, Mg.

Asesor:

Ian Vázquez Rowe, PhD.

Co-Asesor:

Ramzy Francis Kahhat Abedrabbo, PhD.

Lima, 2025

Informe de Similitud

Yo, Ian Vázquez Rowe, docente de la Escuela de Posgrado de la Pontificia Universidad Católica del Perú, asesor de la tesis de doctorado titulada ***Temporal analisis of macroplastic waste concentration in Peruvian coastal rivers using a Deep learning approach***, del autor **Miguel Angel Astorayme Valenzuela**, dejo constancia de lo siguiente:

- El mencionado documento tiene un índice de puntuación de similitud de 15%. Así lo consigna el reporte de similitud emitido por el software *Turnitin* el 22 de octubre de 2025.
- He revisado con detalle dicho reporte y la Tesis o Trabajo de investigación, y no se advierte indicios de plagio.
- Las citas a otros autores y sus respectivas referencias cumplen con las pautas académicas.

Lugar y fecha:

Lima, a 22 de octubre de 2025

Apellidos y nombres del asesor / de la asesora: Vázquez Rowe, Ian	
CE: 001084152	Firma Ian
ORCID: 0000-0002-7469-2033	

Dedication

To my mother, Victoria V., whose immense love cannot be repaid in a single lifetime.

Acknowledgements

One of the most significant milestones in academic life is the completion of doctoral research. Therefore, having financial support and an environment conducive to the development of ideas is essential. In this context, I would like to express my deepest gratitude for Agreement No. 174-202 between ProCiencia and PUCP, which provided the necessary financial support. I also thank the PELCAN research group for welcoming me and offering invaluable human, logistical, technical, and financial assistance.

I am equally grateful to my advisors for their constant guidance, patience, and academic rigor in reviewing the scientific output generated throughout this research. I also extend my thanks to BSc. E. Muñoz for his logistical support during fieldwork, to the PUCP Artificial Intelligence Laboratory, and especially MSc. F. Pineda for his technical assistance during the experimental phase.

I wish to express my special gratitude to SENAMHI, and particularly to the Department of Hydrology, for planting in me the seed of scientific curiosity. I can now say that it was worth taking my first steps at your institution, where the camaraderie and enthusiasm for learning helped guide me to this point. This marked an important first step toward a new stage: doing science.

Finally, I would like to thank my alma mater, UNMSM, for supporting my academic journey by granting me a leave of absence from work during my doctoral studies.

Abstract

Rivers are the main source of plastic waste entering the oceans, primarily due to inadequate solid waste management. The presence of polymers in water bodies creates various problems, such as affecting the mobility of species, sometimes even causing their death. Furthermore, over time, the polymers fragment, resulting in the suspension of microplastics and nanoplastics, and in some cases, releasing toxic additives that disrupt the natural balance of aquatic ecosystems. In this context, monitoring macroplastics (MPs) in waterways can contribute to more efficient solid waste management. Based on this premise, this thesis proposed the development of a methodological strategy for adapting a deep learning-based artificial intelligence (AI) model for the classification and quantification of MPs and other elements derived from solid waste. This thesis also sought to understand the behavior of MPs concentrations in relation to the temporality of waterways. A 1.6-km stretch of the Rímac River, which flows through the city of Lima and empties into the Pacific Ocean, was defined as the area of interest. The results of this thesis are divided into two parts: on the one hand, a scientific analysis of current trends in the use of artificial intelligence-based tools; on the other, the application of the proposed model. First, it was observed that the introduction of AI into the detection of MPs in water bodies is recent, beginning in 2018 with machine learning-based techniques. The current trend is toward the use of deep learning approaches, particularly neural networks and transfer learning. In this context, the YOLOv11 model was selected for its single-shot architecture, offering advantages for potential real-time monitoring systems. The results suggest that the model is capable of detecting and classifying plastic bags (mAP@$50 = 0.75$) and tires (mAP@$59 = 0.82$) with acceptable accuracy. Additionally, it was found that black plastic bags were the most frequently observed items throughout the year, followed by colored bags and raffia bags. The findings of this research enhance our understanding of how plastic waste interacts with aquatic environments through observed concentration patterns. Moreover, the results contribute to bridging the knowledge gap regarding the role of rivers in transporting solid waste to the ocean, particularly in the context of the Global South.

Keywords: *Circular economy, Deep learning, Industrial ecology, Machine learning, Marine litter, Plastic waste, Artificial intelligence, Transfer learning, Waste management, YOLOv11*

Resumen

Los ríos constituyen la principal fuente de ingreso de desechos plásticos a los océanos, principalmente debido a una gestión inadecuada de los residuos sólidos. La presencia de polímeros en los cuerpos de agua genera diversos problemas, como la afectación de la movilidad de las especies, llegando incluso a causar su muerte. Con el tiempo, los polímeros se fragmentan, dando lugar a microplásticos y nanoplásticos, y en ciertos casos, liberan aditivos tóxicos que alteran el equilibrio natural de los ecosistemas acuáticos. En este contexto, el monitoreo de macroplásticos (MPs) en los cursos de agua puede contribuir a una gestión más eficiente de los residuos sólidos. Bajo esta premisa, la presente tesis propuso una estrategia metodológica para adecuar un modelo de inteligencia artificial (IA) basado en deep learning, orientado a la clasificación y cuantificación de MPs y otros elementos derivados de los residuos sólidos. Asimismo, se buscó comprender el comportamiento de la concentración de MPs en relación con la temporalidad de los cursos de agua. El área de estudio corresponde a un tramo de 1.6 km del río Rímac, el cual atraviesa la ciudad de Lima y desemboca en el océano Pacífico. Los resultados se estructuran en dos partes: un análisis científico de las tendencias actuales en el uso de herramientas basadas en IA y la aplicación del modelo propuesto. Se observó que la incursión de la IA en la detección de MPs es reciente, iniciándose en 2018 con técnicas de machine learning. La tendencia actual se orienta hacia enfoques de deep learning, especialmente redes neuronales y transferencia de aprendizaje. En este marco, se seleccionó el modelo YOLOv11 por su arquitectura de una sola pasada (single shot), que representa una ventaja para futuros sistemas de monitoreo en tiempo real. Los resultados indican que el modelo puede detectar y clasificar bolsas plásticas (mAP$@50 = 0.75$) y neumáticos (mAP$@59 = 0.82$) con una precisión aceptable. Además, las bolsas negras fueron las más frecuentes durante el año, seguidas por las de color y las de rafia. Los hallazgos de esta investigación permiten comprender cómo interactúan los desechos plásticos en el medio acuático, a partir de la variación en su concentración. Además, los resultados obtenidos contribuyen a una mejor comprensión del rol de los cursos de agua en el transporte de residuos sólidos hacia el océano, ayudando a cerrar la brecha de información existente en el Sur Global.

Palabras claves: *Economía circular, Aprendizaje profundo, Ecología industrial, Aprendizaje automático, Basura marina, Residuos plásticos, Inteligencia artificial, Aprendizaje por transferencia, Gestión de residuos, YOLOv11.*

Contents

Chapter 1

Overview of the Research

“The greatest threat to our planet is the belief that someone else will save it.”
—Robert Swan

1.1 Introducction

1.1.1 Plastic pollution in the aquatic environment: a growing environmental concern

With the industrial revolution, urban expansion and the need for a better quality of life, large-scale anthropogenic pollution began to impact natural ecosystems, affecting water quality, air, and climate (Edo et al., 2024). For instance, agricultural waste containing pesticides, wastewater tributaries carrying plastic fragments, and industrial waste loaded with chemical compounds have contributed to the contamination of natural environments, particularly aquatic systems, affecting the quality of life of multiple aquatic species since the 20th century (Brander and Betjemann, 2023). Among them, emerging pollutants, included those derived from petroleum-based products, such as plastic polymer emissions (i.e., nanoplastics, microplastics and macroplastics MPs), have shown to be non-biodegradable and persistent in natural systems (Bautista-Zamudio et al., 2023), leading to their accumulation in different natural compartments, such as lakes, oceans or in sediments (Nawaz et al., 2025).

According to Hu et al. (2024), large-scale plastic production began in the 1950s, increasing from 1.5 million metric tons to 400 million metric tons by 2019. Furthermore, plastic production is expected to reach 1,124 million metric tons by 2050 (PlasticsEurope, 2019). In this context, although the use of plastics includes beneficial applications for society, due to inadequate solid waste management (N. et al., 2025), polymers have accumulated over the years, disrupting the balance of natural ecosystems (Padha et al., 2022). The aquatic environment (AE) has contributed to the mobility of plastic emissions, making rivers located in urban environments, particularly large cities, a critical source of pollution, transporting inorganic waste to the oceans and spreading plastic pollution (Lebreton et al., 2017) to other coastlines such as islands (Sánchez-García and Sanz-Lázaro, 2023) and even certain remote vaguely populated areas of the world (Aliani et al., 2020).

As discussed by Ballengee et al. (2021), for a long time the accumulation of plastic polymers in the AE went unnoticed until their discovery in the late 1990s, as in the case of the Great Pacific Garbage Patch (Tamoria and Schofield, 2024). By then, little had been done to prevent plastic pollution, on the one hand, or to mitigate its effects through cleanup efforts (Moore and Phillips, 2011), on the other, leading to multiple negative impacts on the development of marine species. For instance, plastic accumulation has shown to be ingested by marine animals, in some cases as microplastics (e.g., most fish species), but also as MPs, such as macrofauna like turtles (Ramon-Gomez et al., 2024), dolphins (Coram et al., 2021) or sea lions (Gómez et al., 2024). Other processes that have originated different types of damage on marine animals include suffocation (Choudhury et al., 2022), and entanglement (Hong et al., 2024), as well as episodes of plastic polymers enhancing invasive species (Pinochet et al., 2024). All this has also generated effects on human health, as direct and indirect human marine biota consumption has led to these particles entering the human food chain (HUE et al., 2021). According to Dullah et al. (2025) microplastics exhibit a variety of physical and chemical characteristics that complicate human health risk assessment, partly due to the lack of standardized methods for evaluating microplastic concentrations in biota. However, findings by Ho et al. (2025) suggest that ingesting marine biota contaminated with microplastics may lead to adverse effects such as respiratory problems, organ damage, metabolic disorders, inflammation, among others. As a result, pollution of the AE is now recognized as one of the greatest environmental threats to our planet (Mattsson et al., 2024).
This can be analyzed from three complementary perspectives. First, plastic accumulation in water bodies over the years has grown gradually, to the point that, based on data estimates from 2023, 50% of the worlds plastic production (i.e., 322 million metric tons), ended up in the AE, including ocean, rivers, or lakes (Flores-Díaz et al., 2025). Second, the slow degradation rate of plastic polymers exacerbates the problem of plastic accumulation in natural systems. For instance, polyethylene, the most prevalent among marine debris, accounts for over one-third of plastic waste (Wang et al., 2021) and can take up to 1000 year to completely degrade. Third, as these plastics degrade over time, fragmentation leads to the formation of microplastic and nanoplastics, which not only disrupt marine ecosystems but also release toxic additives into the AE, such as antioxidants or pigments (Hahladakis et al., 2018), altering their natural balance. Furthermore, according to Campanale et al. (2020), nanoparticles can adsorb hydrophobic organic contaminants from AEs.

By the end of the 20th century, oxo-biodegradable plastics were introduced and promoted as a potential solution to plastic degradation. However, studies as the one conducted by Padermshoke et al. (2025) indicate that their biodegradability in the AE remains poorly understood, raising global concerns as their fragmentation is known to harm marine ecosystems. Consequently, compostable bioplastics have emerged in recent decades as a promising alternative for replacing certain types of packaging, such as carrier bags (Maes, 2022). Although their production was initially costly and accounted for only 1% to 2% of total plastic production in the past decade (Pagliaro, 2019). Nevertheless, a recent study by Boldrini et al. (2024) suggests that the accumulation of biodegradable plastic bags waste is harmful to the ecosystem in ocean environments, because they release toxic as phthalic acid esters, and they can also generate similar physical effects as conventional polymers when ingested by biota.

1.1.2 Plastic pollution and public policy

In recent years awareness regarding plastic pollution has increased worldwide and an array of different actions have been taken to regulate plastic consumption. One of the first efforts to alert and recognize plastic pollution as an emerging issue was introduced at the Stockholm Conference in the 1970s (Paglia, 2021). In the following decades, obligations were established to prevent, mitigate and control pollution in the marine environment e.g., Part XII of the United Nations Convention on the Law of the Sea (Scott, 2024), which addressed regulations related to the discharge of plastic into the sea from ships. However, this initiative must only be considered as a set of general provisions for nations regarding the protection of oceans against pollution, including plastic waste, as there was no specific binding treaty at the time.

Thereafter, in 1992 the first treaties and regulations were established at the Earth Summit in Rio de Janeiro, Brazil (Bell, 1998), leading to the creation of Agenda 21, a global action plan to promote sustainable development (UNCED, 1992). Although not primarily focused on plastics, Agenda 21, in Chapter 17 of Section II (i.e., Conservation and Management of Resources for Development) emphasized the importance of preventing ocean pollution, including enclosed and semi-enclosed seas and coastal zones, addressing plastic waste and debris from both land and vessels. Additionally, Chapter 21 underscored the need for proper solid waste management to mitigate environmental impacts.

In the first decade of the 21st century, the Stockholm Convention on Persistent Organic Pollutants recognized and warned that certain plastic polymers may contain or absorb persistent organic pollutants POPs (e.g., biphenyls, dioxins), which are highly toxic and can be released into water upon contact, contaminating the environment (Glüge et al., 2016). Furthermore, by the end of the same decade, many countries began banning single-use plastic bags and promoting biodegradable alternatives (Ljubomir et al., 2009). By 2015, efforts to address the reduction of marine pollution were concentrated in the UN Agenda 2030 within the Sustainable Development Goal "Life Underwater" (UN, 2015).

At a regional level, and aligned with the 2030 Agenda, the European Union approved

Directive (EU) 2019/904 (UE, 2019), which bans certain single-use plastics to reduce their environmental impact (e.g., cutlery, plates, expanded polystyrene containers, and cups for food and beverages) by 2021. This directive came into effect in 2021 setting an important precedent as one of the first large-scale laws to ban single-use plastic products. It imposed restrictions of specific disposable plastic items aiming to reduce marine plastic pollution (UE, 2019).

Thereafter, on March 2nd, 2022, during the 5th United Nations Environment Assembly (UNEA-5.2), a historic resolution called *5/14 End plastic pollution: Towards an international legally binding instrument* was adopted by 175 nations (UNEA, 2022), which introduces a mandate to develop lobal treaty to address plastic pollution at different levels, with the aim of regulating the life cycle of plastic (Senathirajah et al., 2023). It stands out as the first global agreement that could potentially develop a legally binding international treaty, including obligations at the country level to improve the management of plastic throughout the supply chain. According to Johnson et al. (2025) the current agreement also calls on states to implement policies aimed at minimizing the use of unsafe chemicals in plastic packaging and reducing the production of certain problematic types of plastics. The resolution established by UNEA-5.2 promoted the creation of the Intergovernmental Negotiating Committee (INC) to develop a legally binding international agreement addressing plastic pollution from design to production and disposal (UN, 2022).

The INC, which began its work in the second half of 2022, held its first session (INC-1) between November and December in Uruguay, establishing a roadmap for the negotiations (UNEP, 2022). The second session (INC-2), held between May and June 2023 in France, focused on debating measures to reduce plastic production and control its life cycle (UNEP, 2023). In November of the same year, the third session (INC-3) took place in Kenya (UNEP, 2023), where disagreements among participating countries became evident regarding production limits. Additionally, plastic pollution was recognized as a public health issue, with concerns raised about the impacts of microplastics on human health (RAP-AL, 2023).

In April 2024, the fourth session (INC-4) was held in Canada, where a revised draft of a global legally binding agreement on plastic pollution was presented and analyzed by country delegations (UNEP, 2024a). In December of the same year, the first part of the fifth session (INC-5.1) took place in the Republic of Korea (UNEP, 2024b). However, a global agreement on plastic pollution was not reached due to opposition from plastic-producing countries, such as Saudi Arabia and Russia, and others despite nearly 100 countries advocating for manufacture limits Anderson (2024). Consequently, a second part of the fifth session (INC-5.2) is scheduled for 5-14 August 2025 in Geneva, Switzerland, with the aim of reaching a final agreement.

In Latin America, one of the first efforts to legislate on plastic pollution was made at the provincial level in Buenos Aires, Argentina, with the enactment of Law 13.868 in 2009. This law prohibits the use of polyethylene bags and other plastic materials in supermarkets and commercial establishments across its jurisdiction, mandating their progressive replacement with biodegradable alternatives to minimize environmental impact (Buenos-Aires, 2009). In 2016, Resolution No. 341/APRA/16 was enacted, introducing a total ban on plastic bags, with exceptions only for justified sanitary purposes (Buenos-Aires, 2016). This measure set a precedent for other provinces to implement similar regulations.

In 2016, Colombia legislation took an important step forward with the enactment of Resolution 668, which regulates the use of plastic bags and establishes measures to reduce their consumption (MADS, 2016). Thereafter in 2019, it was replaced by Resolution 2184, which established an obligation for plastic bags distribution (MADS, 2022). However, it was not until 2022 that Law 2232, establishing measures for the gradual reduction of the production and consumption of single-use plastic products, was passed, setting deadlines for their phasing out, and promoting their replacement with biodegradable products (MADS, 2022).

Chile is another country in the region that has legislated in terms of plastic pollution, by enacting, in 2018, Law 21.100, which bans the use of plastic shopping bags throughout its territory, becoming the first country in Latin America to eliminate plastic bags, although it is also the country with the highest generation of plastic waste per capita (Circulaelplastico, 2018).

Regarding Peru, it was not until late 2018 that Law 30884 was passed. In particular, the law was centered on single-use plastics and disposable containers or packaging, aiming to protect the environment and public health (Peruano, 2018). Furthermore, similarly to Argentina, and to discourage the use of single-use plastic bags, a 0.1 of Peruvian sol tax was imposed in 2019 on plastic bags handed out in commercial establishments such as supermarkets, department stores and warehouses (OCEANA, 2019). This tax doubled each year until reaching 0.5 Peruvian soles (i.e., PEN) by 2023 and beyond, (SUNAT, 2023). The law prohibits the purchase and sale of single-use plastics on beaches, cultural heritage sites, and protected areas, as well as promoting awareness-raising programs (por Naturaleza, 2019).

1.1.3 Strategies for estimating plastic emissions in environmental compartments

Given that current efforts worldwide focus on estimating the number of plastic emissions that exist around the planet, their mobility and their interaction with natural ecosystems, a range of techniques have been developed to grasp the magnitude of the environmental hazard. For instance, in terms of MPs, approaches range from manual counting methods (Teng et al., 2022) to those using geographic information systems for larger areas (Moy et al., 2018), as well as holistic approaches focusing on regional (ItaNagy et al., 2022) or global scales (Jambeck et al., 2015). While these methods help fill gaps in information concerning plastic accumulation in terrestrial and AEs, they may require significant monetary and human resources, time, and may involve errors in the process (Teng et al., 2022) (Angelini et al., 2019).

In this context, geographic information systems (GIS) and satellite imagery were explored for plastic detection. While GIS enabled the coverage of larger areas, the quantification process still required semi-automated methods and relied on human operators for accurate results (Moy et al., 2018). In parallel, the use of artificial intelligence (AI) regained popularity in the early 21st century after the so-called AI winter between the late 1970s and early 1990s, a period when research in this field significantly declined (He et al., 2021). Advances in terms of more powerful CPUs and, especially, GPUs facilitated the development and expansion of Machine Learning (ML), a subfield of AI, (LeCun et al., 2015). Consequently, by 2018, the first studies on plastic detection and quantification in AE were directed by (Martin et al., 2018), exploring Random Forest (RF), a common ML technique.

From that point onward, research expanded to include analysis comparing different techniques, such as RF and Support Vector Machine (SVM) (Cortesi et al., 2021); or Artificial Neural Networks (ANNs) and Decision Tree (Lavender, 2022). Additionally, ensemble methods combining two or more ML models were explored, e.g., ensemble of light gradient-boosting machine + K-Means methods for detecting floating plastic in the ocean (Taggio et al., 2022), along with innovative modifications to classical ML approaches (Hernández-González et al., 2019).

It was not until 2021 that deep learning (DL), a subfield of ML (LeCun et al., 2015), gained attention, leading to the testing of complex convolutional neuronal network (CNN) based architectures with the first research conducted by Freitas et al. (2021), who used CNN to detect marine litter.

Some studies, like the one conducted by Sannigrahi and colleagues (2022), propose a system for detecting floating plastic particles in intercoastal water near beaches. However, the study is limited by the resolution of satellite images (Sentinel-2), which restricts the detection to larger plastic particles, mainly MPs. Furthermore, the use of aerial images captured by unmanned aerial vehicles (UAVs) with conventional and multispectral cameras has also been explored, improving MPs detection in river systems (Cortesi et al., 2022).

According to Wolf et al. (2020), DL is a viable ML tool for processing high-resolution images using CNN, as well as for image segmentation and MPs classification (Lorenzo-Navarro et al., 2021). However, the use of CNNs in this area of study has been less explored. The application of Transfer Learning (TL) is emerging, particularly for estimating MPs from images and video recordings using object detection models like YOLOv5 (Teng et al., 2022). Additionally, TL has been explored MPs detection combined with holographic techniques for classification in water (Zhu et al., 2021).Therefore, the use of AI in the field of ML, specifically the DL approach, is presented as a new paradigm (Kylili et al., 2021) adding to the efforts to develop automatic detection tools that can be used to classify plastic and quantify its presence in natural ecosystems. Furthermore, in line with increasing legislation worldwide banning single-use plastic, these tools open the possibility of combining efforts in regulation at the local and even regional level, including regions in the Global South, where plastic pollution emissions are more prevalent, and emission sources are more varied.

In this context, the objective of this doctoral thesis is to monitor and analyze the current state of the art in the automatic detection of MPs in AEs. Additionally, it explores the application of DL approaches in quantifying plastic in specific watercourses. For this, a case study was carried out in one of the main rivers along the Peruvian coast. The novelty of this research lies in the exploration of DL techniques for the automatic detection and classification of MPs in an unfavorable context, characteristic of Global South countries, where solid waste accumulates densely due to inadequate waste management and a lower allocation of resources in the sector compared to more developed countries (Lerpiniere et al., 2025). It also explores the seasonal variations in the concentrations of common plastic types influenced by the AE.
Given the complexity of manually counting MPs, the implementation of automatic counting techniques could serve as a viable alternative; however, prior processing is required to ensure adequate accuracy.

The thesis is structured into four chapters. Chapter 1 provides an overview of the general aspects of the research. Chapter 2 addresses the issue of plastic waste through a bibliometric and literature review focused on the detection, classification, and quantification of MPs in AE using AI tools. It identifies the most widely used techniques, current trends, global approaches, and the growing role of DL. Additionally, it examines the geographical distribution of scientific contributions worldwide. The critical analysis is divided into two sections: the first discusses traditional ML techniques, while the second highlights the emerging dominance of DL. Each section is further categorized based on technique types, input data, and advanced methodologies for clarity. To aid understanding, this chapter begins with a concise overview of the evolution of AI, key approaches, and classifications, which may be useful reference for new readers. Chapter 3 explores and validates the methodology based on a DL approach for the estimation of the MPs present in natural water courses. The proposed methodology required a one-year hydrological field study, during which aerial images were collected using drones. The purpose was to capture both spatial and temporal variations within the watercourse from aerial images. A subsequent image processing phase was conducted alongside an experimental process to refine the You Only Look Once (YOLO) model version 11 (Sapkota et al., 2024), a CNN-based approach, utilizing TL. The model was trained to detect, classify, and quantify MPs in a section of the river Rímac (Lima, Peru) as a case study to demonstrate the proposed methodology. In addition, the concentrations of the most common types of plastics, such as polyethylene bags and tires, are examined as a function of the seasonality of the river flow, which exhibits a well-defined bimodal behavior.

Moreover, the proposed methodology is expected to automate MPs counting, reducing uncertainty, processing time, and human resource requirements compared to manual counting. Additionally, the findings can potentially be extrapolated to other sections of the Rímac River and similar river systems in the region, with adjustments based on the type of plastic waste present. Furthermore, this research lays the foundation for developing an automated, near real-time detection system for plastic elements in water courses, contributing to improved solid waste management. Finally, Chapter 4 presents the overall conclusions of the research and outlines recommendations for future work.

1.2 Objectives

1.2.1 General Objective

Evaluate the concentration of macroplastic elements and their relationship with the hydrological regime in riverine systems using artificial intelligence techniques.

1.2.2 Specific Objectives

- Identify the most widely applied machine learning and deep learning architectures for the detection of macroplastics in aquatic environments, assessing their performance, adoption, and degree of innovation in both global and regional contexts.
- Delimit the area of interest based on its geographic location, environmental significance, and accessibility, to enable field campaigns throughout a full hydrological year.
- Fine-tune a pre-trained YOLOv11 model for the detection and classification of macroplastic elements in the area of interest.
- Analyze the spatiotemporal variation in the concentration of macroplastic elements in relation to the bimodal hydrological regime of the watercourse.

1.3 Research Problem, Justification, and Hypothesis

The presence of inorganic elements such as polymers in water bodies is a direct consequence of inadequate solid waste management practices worldwide. This issue is particularly evident in the Global South, where information is scarce and limited data exists regarding the volume of polymers entering rivers system alone the oceans. In this context, the present research proposes the exploration of a new paradigm: the use of artificial intelligence as a tool to classify and quantify these polymers, with a focus on macroplastics (MPs) and other solid waste derivatives. The importance of this work is justified by the automation of detection and estimation processesin the automation of detection and estimation processes, allowing for the identification of areas with higher contamination levels, the analysis of spatial and temporal trends, and the examination of the relationship between MPs concentrations and the bimodal hydrological of the river system. The proposed methodology is expected to reduce uncertainty in estimating MPs volumes in waterways and to improve our understanding of their dynamics in natural river systems. Validating this approach may serve as a basis for implementing future monitoring systems in the Rímac and other coastal rivers, with potential for replication in similar contexts. The central hypothesis of this research is that DL techniques can be effectively utilized for counting MPs in the waterways of the Peruvian coast.

Chapter 2

Manuscript one: The use of artificial intelligence algorithms to detect macroplastics in aquatic environments: A critical review

> *“The future of life on Earth depends on our ability to take action. Many people are doing what they can, but true success can only come if there is a change in our societies, our economy, and in our politics.”*
>
> —Sir David Attenborough

This manuscript was published in the Science of The Total Environment.

Astorayme, M. A., Vázquez-Rowe, I., & Kahhat, R. (2024). The use of artificial intelligence algorithms to detect macroplastics in aquatic environments: A critical review. Science of The Total Environment, 945, 173843. DOI: 10.1016/j.scitotenv.2024.173843

In this chapter, we present a scientific review of the application of AI in the detection of MPs in aquatic environments, with an emphasis on the methods used, current trends, and the future direction of AI in this field. We also examine the global progress of related research and the most commonly used AI-based strategies and architectures, highlighting the results obtained and the contributions of the most relevant studies.

Science of the Total Environment 945 (2024) 173843

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Review

The use of artificial intelligence algorithms to detect macroplastics in aquatic environments: A critical review

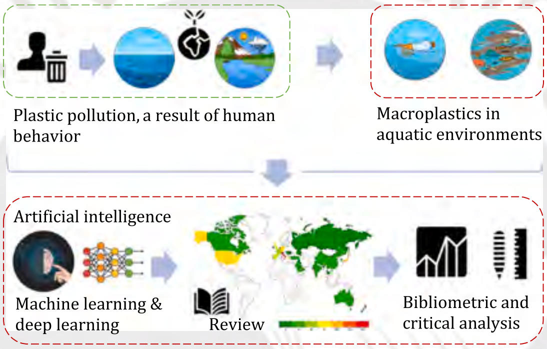
Miguel Angel Astorayme[a,b,*], Ian Vázquez-Rowe[a], Ramzy Kahhat[a]

[a] *Peruvian Life Cycle Assessment & Industrial Ecology Network (PELCAN), Department of Engineering, Pontificia Universidad Católica del Perú, Av. Universitaria 1801, San Miguel 15074, Lima, Peru*
[b] *Dept. of Fluid Mechanics Engineering, Universidad Nacional Mayor de San Marcos, Av. Universitaria/Av. Germán Amézaga s/n., Lima 1508, Lima, Peru.*

HIGHLIGHTS

- The use of AI in the quantification of macroplastics has gained strength since 2018.
- Conventional ML models can be used for a binary classification of macroplastics.
- DL models perform better to detect and multiclass classification of macroplastics.
- CNN-based architectures are currently trending in macroplastic detection.
- Transformer-based architectures are yet to be applied to detect macroplastics.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Dimitra A Lambropoulou

Keywords:
Circular economy
Deep learning
Industrial ecology
Machine learning
Marine litter
Plastic waste

ABSTRACT

The presence of macroplastic (MP) is having serious consequences on natural ecosystems, directly affecting biota and human wellbeing. Given this scenario, estimating MPs' abundance is crucial for assessing the issue and formulating effective waste management strategies. In this context, the main objective of this critical review is to analyze the use of machine learning (ML) techniques, with a particular interest in deep learning (DL) approaches, to detect, classify and quantify MPs in aquatic environments, supported by datasets such as satellite or aerial images and video recordings taken by unmanned aerial vehicles. This article provides a concise overview of artificial intelligence concepts, followed by a bibliometric analysis and a critical review. The search methodology aimed to categorize the scientific contributions through temporal and spatial criteria for bibliometric analysis, whereas the critical review was based on generating homogeneous groups according to the complexity of ML and DL methods, as well as the type of dataset. In light of the review carried out, classical ML techniques, such as random forest or support vector machines, showed robustness in MPs detection. However, it seems that achieving optimal efficiencies in multiclass classification is a limitation for these methods. Consequently, more advanced techniques such as DL approaches are taking the lead for the detection and multiclass classification of MPs. A series of architectures based on convolutional neural networks, and the use of complex pre-trained models through the transfer learning, are currently being explored (e.g., VGG16 and YOLO models), although currently

---

[*] Corresponding author at: Peruvian Life Cycle Assessment & Industrial Ecology Network (PELCAN), Department of Engineering, Pontificia Universidad Católica del Perú, Av. Universitaria 1801, San Miguel 15074, Lima, Peru.
E-mail address: m.astoraymev@pucp.edu.pe (M.A. Astorayme).

https://doi.org/10.1016/j.scitotenv.2024.173843
Received 30 January 2024; Received in revised form 8 May 2024; Accepted 5 June 2024

the computational expense is high due to the need for processing large volumes of data. Additionally, there seems to be a trend towards detecting smaller plastic, which need higher resolution images. Finally, it is important to stress that since 2020 there has been a significant increase in scientific research focusing on transformer-based architectures for object detection. Although this can be considered the current state of the art, no studies have been identified that utilize these architectures for MP detection.

1. Introduction

Plastic accumulation in aquatic environments (AEs) has shown to generate multiple environmental hazards (Jindal and Kumar, 2020), such as the ingestion of microplastic particles by plankton and fish, or the entanglement of marine macrofauna with macroplastics - MPs (So et al., 2022), especially fishing nets (Deville et al., 2023). Considering that plastic debris cannot be processed by the ocean, the accumulation of this inorganic material is rapidly increasing (Jamali and Mahdianpari, 2021). In fact, recent studies reveal that 13 million metric tons of plastic are released into the ocean annually (Reddy and Lau, 2020) and a report published by UNEP suggests this value could double by 2030 if rapid and robust actions are not enforced soon (UNEP, 2021). As a consequence, the scientific literature has already identified that over 1400 marine species have been affected by marine plastic (Claro et al., 2019; Roman et al., 2021), and there is increasing evidence that humans are exposed to an ever-growing cocktail of micro- and nanoplastics that are either being inhaled or ingested in their daily lives (Rodrigues et al., 2022; Trevisan et al., 2022).

Rivers have a significant contribution in terms of plastic release into the oceans (Meijer et al., 2021). Inadequate municipal solid waste (MSW) management in cities, towns and rural areas (Ita-Nagy et al., 2022), among other factors, such as natural disasters (Semernya et al., 2017) or littering (Newbould et al., 2021), are the main responsible sources that deliver plastic emissions to fluvial systems. In fact, Meijer et al. (2021), in line with previous research conducted by Lebreton et al. (2017), estimated that 0.8–2.7 million metric tons are discharged into the ocean by rivers worldwide each year. The study highlighted that those rivers located in certain urban areas, especially in the Global South, pollute significantly more than others. Despite this important role of rivers in marine plastic accumulation, Ita-Nagy et al. (2022) have identified a set of barriers that may retain microplastics and MPs, which may eventually accumulate in river sediments or fluvial plains, rather than being released into the ocean.

Research on plastic waste in AEs became a mainstream research line in the past decade. For instance, Barnes et al. (2009) established some of the first comprehensive studies related to the production, and accumulation of plastic debris worldwide. Likewise, a few years later, Jambeck et al. (2015) estimated the flow of plastic debris entering the ocean from terrestrial sources. Ever since, other studies have aimed at updating and improving the metrics and methodologies to estimate plastic debris amounts in the ocean (Lebreton et al., 2017; Mai et al., 2023; Ita-Nagy et al., 2022) and other water bodies (Hoffman and Hittinger, 2017). These and other studies focus on linking global data associated with solid waste (e.g., the state of the local economy and population density), to estimate the volume of plastic release to the environment (Margallo et al., 2019). However, the data generated in many cases cannot be validated, and, in other cases, the estimates are gross, increasing the epistemic uncertainty of the quantifications. In fact, the full understanding of plastic debris trajectories and behavior in AEs, as well as the final amounts of plastic debris entering the ocean at various scales (e.g., MP, microplastic or nanoplastic) is poorly understood (Garello et al., 2019). Thus, more reliable methods that consider spatial and temporal components are likely essential to validate and reduce the uncertainty in estimates. Furthermore, climate change and global events such as COVID-19 affect the sustainability of ecosystems, markets and human behavior (Karmaoui et al., 2023). Consequently, the estimation of plastic waste concentrations may be influenced by these factors, requiring a more complex analysis in this context.

To close the gaps regarding data quality and availability for the quantification of MPs in natural water bodies, artificial intelligence (AI) has emerged as a new paradigm in plastic debris monitoring and quantification (Kylili et al., 2021). AI has acquired great relevance since the development of microprocessors (i.e., CPU and GPU). In fact, in recent years, AI has been explored in various areas of science (Hu et al., 2019), for the benefit of the technical and scientific community, including new approaches for malware detection (Shaukat et al., 2024), detection of diseases from medical images (Wang et al., 2021), the search for new deposits in the mining industry (Lindsay et al., 2022), or developing models to classify sediments in natural canals (Himanshu et al., 2017), among others.

The potential of AI to quantify plastic debris in AEs has started to emerge in the scientific literature as an incipient field of study in recent years (Topouzelis et al., 2021), as an on-going process in which scientists are moving from manual detection to the use of remote sensors and automation processes that involve the use of machine learning (ML) approaches. In fact, the use of AI and geographic information systems (GIS) appears as an attractive solution to address data gaps, as well as aid in the qualitative and quantitative analysis, regarding plastic debris. For instance, as analyzed in the review by Gnann et al. (2022), a range of ML and deep learning (DL) techniques have already been applied to detect (Hidaka et al., 2022) or predict (Hernández-González et al., 2019) solid waste on beaches, locate and map the shape of plastic debris (Kylili et al., 2021), quantify marine litter abundance (Franceschini et al., 2019) or detect plastic concentrations along rivers (Cortesi et al., 2021).

Consequently, the two research questions that are answered in this critical review are the following:

- What are the most relevant AI techniques and the current trends in MP detection?
- How has the use of AI influenced the reduction of data gaps related to the detection and quantification of MPs in AEs?

Therefore, the objective of this critical review is to focus on showing the connection between the use of ML techniques, with a particular interest in DL approaches, to detect, classify and quantify MPs in AEs, supported by satellite, aerial images or video recordings taken by unmanned aerial vehicles (UAVs). Hence, this article starts with a general review of AI, describing its general importance and the different approaches that can be applied. Thereafter, a bibliometric analysis is conducted on the state-of-the-art regarding the use of AI methods to detect plastic debris in AEs. Lastly, the advantages and disadvantages of new AI approaches are discussed, as well as highlighting potential future challenges and recommendations.

Overall, the novelty of the review is centered on the critical discussion and classification of the most commonly used methods and new advances in this field, highlighting the main advantages and drawbacks of using different methods. The target readership is mainly linked to researchers or consultants that are familiar with AI and wish to generate more accurate metrics and potential solutions for the problem of plastic accumulation in AEs, but also for researchers that are considering the use of AI models in the near future as applied to plastic debris. This review differs from other previous works, e.g., Topouzelis et al. (2021) and Andriolo et al. (2022), in the focus and extension of the discussion related to the deep learning (DL) approach, which is a new trend in the field of plastic debris detection. Although other studies, such as Gnann et al. (2022) and Tamin et al. (2023), included critical reviews that discussed research associated with the use of the DL approach, this review goes a step further by updating and analyzing the discussion surrounding the DL approach, with particular emphasis on the transfer learning (TL) methodology.

2. Artificial intelligence: methodological framework and definitions

AI, which commenced with Alan Turing's work linked to the concept of the "universal machine" in the 1940s and 1950s, as well as with the introduction of neural networks – NNs (McCulloch and Pitts, 1943; Rosenblatt, 1958), only started to gain popularity in the 1980s with the development of expert systems, e.g., applied to medical diagnoses (Buchanan, 1982). Despite this steady development in its early life, AI can currently perform complex tasks that would require, if viable, human intervention, such as image recognition, natural language processing or autonomous driving (Just, 2024). AI is strongly related to other disciplines such as big data, computer science and statistics. ML, which is considered a subfield of AI, can be defined as a set of techniques (i.e., algorithms) that enable computers to learn from data through experience, without requiring explicit programming for a particular task (Murphy, 2013).

According to Jordan and Mitchell (2015), and later explained by Geron (2019), ML techniques can be classified in supervised learning (SL), unsupervised learning (UL), semi-supervised learning (SSL), and reinforcement learning (RL). The main difference between UL and SL is that SL approaches need to study the previously labeled input data (i.e., the classification of the information is known). Consequently, this information is used to train the model (i.e., the training and testing process). In contrast, UL approaches are based on studying the input data (Flach, 2012), which is unlabeled (i.e., the machine does not receive initial information on data classification). Therefore, the algorithm begins to identify patterns and correlations.

Currently, the SSL approach combines elements of SL (with the help of labeled data) and UL (with the help of unlabeled data). For instance, according to Rani and Kant (2020), in the SL approach, the algorithm of ML receives a set of input examples along with the associated desired responses, named tags. The goal of this procedure is to enable the algorithm to learn a relationship between inputs and outputs, allowing it to make precise predictions on unlabeled data. Finally, in reinforcement learning (RL) approaches, which are inspired by the learning of the sequential behavior of animals, based on learning by error (Ishii and Yoshida, 2006), the machine learns by itself, which is the best strategy to be applied, penalizing the error, and valuing the success, causing the behavior to be modified. According to Eßer et al. (2023), RL has shown significant advancements in recent years across a range of demanding tasks, including traditional strategy and real-time computer games, applications in the field of robotics, maintenance of engineering systems (Marugán, 2023), reliable transmissions in internet of things (Alipio and Bures, 2023), or as an effective method to interpret graphs (Tang et al., 2024). Although it can be challenging to make a classification of all the existing ML methods, Table 1 classifies the main ML and DL methods, following the two classical approaches: supervised and unsupervised learning.

2.1. Deep learning - DL

DL is a subfield of ML which was proposed due to the limitations found in processing natural data in classical ML techniques (LeCun et al., 2015). It differs from the latter by using neurons or nodes, which in turn form layers and multilayers, giving rise to artificial neural networks (ANNs). This architecture is able to learn on its own, inspired by the structure and function of the human brain (Gospodinov, 1991). ANNs are organized as follows: input layer, one or more hidden layers, and one output layer. However, it should be noted that shallow ANNs (i.e., one hidden layer) belong to the field of ML, whereas with the increase in computational capacity (i.e., GPU and TPU) and the successful use of algorithms like *gradient descent* and *backpropagation* (LeCun et al., 1998), more complex architectures were explored, giving rise to deep neural networks (i.e., two or more hidden layers, multilayers, among other complex arrangements), creating the field of DL.

Table 1
List of supervised and unsupervised machine learning (ML) and deep learning (DL) techniques (adapted from Geron, 2019).

Approach	Technique	Related bibliography
ML Supervised	k-Nearest Neighbors (KNN)	Andriolo et al. (2022)
	Linear Regression model (LM)	Hernández-González et al. (2019)
	Support Vector Regression (SVR)	Kantanantha and Pattaraumpornchai (2023)
	Logistic Regression (LR)	Nanehkaran et al. (2023)
	Decision Tree (DT)	Lavender (2022)
	Light Gradient Boosting Machine (LightGBM)	Taggio et al. (2022)
	Binary classifiers: Support Vector Machines (SVM)	Cortesi et al. (2021)
	Multiclass classifiers:	
	Naive Bayes (NB)	Nanehkaran et al. (2023)
	SVM	Freitas et al. (2021a)
	Random Forests (RF)	Kaandorp et al. (2022)
	Artificial Neural Networks* (ANN)	Kako et al. (2020)
ML Unsupervised	**Clustering**:	
	K-Mean	Taggio et al. (2022)
	Density-based spatial clustering of applications with noise (DBSCAN)	Tunukovic et al. (2024)
	Hierarchical Cluster Analysis (HCA)	Kiruba-Sankar et al. (2023)
	Self-organizing map-(SOM)	Franceschini et al. (2019)
	Deep belief network (DBN)	Shaukat et al. (2020b) Shaukat et al. (2020c)
	Dimensionality reduction:	
	Principal Component Analysis (PCA)	Shaikh et al. (2022)
	Kernel PCA (K-PCA)	Lu et al. (2020)
	Locally Linear Embedding (LLE)	Roh et al. (2017)
	t-distributed Stochastic Neighbour Embedding (t-SNE)	Maliks and Kadikis (2021)
DL Supervised	Convolutional Neural Networks (CNNs)	Moorton et al. (2022)
	Recurrent Neural Networks (RNNs)	Tur (2023)
	Long short memory Networks* (LSTMs)	Tur (2023)
	Transformer*	Thanh et al. (2024)
DL Unsupervised	Generative Adversarial Networks** (GANs)	Jamali and Mahdianpari (2021)
	Autoencoder**	Burr et al. (2023)
	Guide Deep Decoder (GDD)	Chakma et al. (2023)

[*] They can also operate in an unsupervised manner.
[**] They can also operate in a supervised manner.

According to Glorot and Bengio (2010), before 2006, the strategy for training deep multilayer ANNs had limited success, until the development of various methods. For instance, Hinton et al. (2006) proposed a successful training strategy for a specific type of deep ANNs. Thereafter, Goodfellow et al. (2016) found that this strategy could be used to train many other types of deep ANNs, giving rise to the term "deep learning" implying that it was now possible to train much deeper neural networks.

Together with the contribution of Hinton and colleagues (2006) it is important to highlight a study by Glorot and Bengio (2010), who provided a systematic strategy for the initial assignment of weights based on an algorithm, named *Glorot* or *Xavier initialization*. According to the results obtained in their research, they improve the stability and efficiency of training deep ANNs, which is crucial to mitigate *gradient fading* or *gradient explosion*. These were probably the most important contributions to improving the effectiveness and stability of training in deep neural networks. Nowadays, DL focuses on resolving complex tasks successfully, such as speech recognition (Mehrish et al., 2023), natural language processing (Jahan and Oussalah, 2023), and object detection (Kamath and Renuka, 2023), the latter being relevant in the field of this review.

A limitation that was encountered when applying ANNs is linked to the need to address pattern recognition problems in images (e.g., detection of plastic debris). In response to this constraint, convolutional neural networks (CNNs), which derived from ANNs, were proposed (see Table 1). CNNs add convolutional layers within their architecture, allowing the disaggregation of the image from a filter operation with the aim of generating feature maps, which take some characteristics of the initial image, connected with a fully connected layer. CNN models can be divided in three main components: convolutional, pooling, and fully connected layers (Marin et al., 2021). Hence, CNNs are a deep architecture focused on learning with images and videos.

CNNs have developed further with increasing levels of complexity, especially with the aim of enhancing object detection. For instance, the region-based convolutional neural network (R-CNN) was probably the first two-stage model created with the ability to detect objects, showing an advantage over conventional methods (Girshick et al., 2013; Pal et al., 2021). However, with the aim of solving the execution time issue that R-CNNs present, alternative architectures were presented, such as Fast R-CNN (Girshick, 2015) or Faster R-CNN as an improved version of the former (Ren et al., 2015). Thereafter, with the aim of overcoming constraints found in previous versions, such as computational efficiency and the ability to handle objects of different sizes, architectures such as R-FCN (Dai et al., 2016) and Mask R-CNN (He et al., 2017) have been proposed.

A set of models with single-pass architectures (i.e., one-stage detectors) have been developed in recent years aimed at detecting in real time, such as *You Only Look Once* – YOLO (Redmon et al., 2016), single-shot detector – SSD (Liu et al., 2015), de-convolutional single-shot detector – DSSD (Fu et al., 2017), or RetinaNet (Zhang et al., 2017b). YOLO models, which can be considered as part of the current state-of-the-art in object detection, have been updated through the years (Jocher et al., 2020), gaining stability in their architecture (Reis et al., 2023). Finally, the use of a transformer, a DL architecture based on the multi-head attention mechanism (Vaswani et al., 2017), has gained relevance in recent years in various areas of AI, such as natural language processing (Kumar and Kumar, 2022), speech recognition (Tang, 2023), and object detection (Amjoud and Amrouch, 2023).

3. Materials and methods

Data collection for the critical review comprised an exhaustive analysis of scientific literature published from January 2018 to December 2022, considering the detection, classification and quantification of the six most common types of polymers, as described by Geyer et al. (2017): polyethylene (PE); high density PE (HDPE); low density PE (LDPE); polypropylene (PP); polyvinyl chloride (PVC); polyurethane (PUR); polystyrene (PS); and, polyethylene terephthalate (PET). The remaining types of polymers, such as Teflon, polycarbonate, acrylonitrile butadiene styrene, etc., were included in the group “Others”. The time frame was chosen considering that the search conducted in scientific databases uncovered publications related to the topic starting in 2018. To obtain the scientific articles, two abstract and citation databases accessible to the scientific community were used: Scopus and Web of Science (Elsevier, 2023; Clarivate Analytics, 2023).

The review focused solely on MPs (>5 mm). The rationale behind this decision is linked primarily to the fact that scientific literature related to the use of AI to detect and quantify microplastics and nanoplastics is made up of a different methodology that focuses on field work, sampling, and laboratory analysis (Andrady, 2011). Moreover, MPs may be more likely to end up being plastic waste in freshwater and/or marine ecosystems or buried in the ground (e.g., riverbed or fluvial plain sediments), whereas microplastics end up not only in the natural environment but may be ingested or inhaled by fauna (Hartmann et al., 2019).

The geographical boundaries of the review were limited to aquatic systems (i.e., marine, fluvial or lacustric), excluding other environmental compartments. For each study identified, special attention was given to the way in which different types of images were obtained (i.e., obtained by satellite, UAVs, terrestrial camera, or video recording) and on how field data (i.e., slope, area, discharge, etc.) were collected and processed for their subsequent use as input in the ML techniques.

For the bibliometric analysis, an advanced search was applied, considering historical research related solely to AI or only to MPs, following the Booleans for plastics mentioned in Table 2. Additionally, a filter was applied to assess the scientific contributions per country based on the Booleans in Table 2. In this case, MPs and microplastics are included separately and together only with the objective of showing their distribution (heat map), considering the time frame 2018–2022.

For the critical review, to identify scientific articles related to MPs, an advanced search was performed on February 20th, 2023, using Booleans related to plastic in AEs and the use of AI (see Table 2). The search was performed considering the occurrence in titles, abstracts, and keywords. The search was supported through code programming in Python, for the evaluation of the most used words in the scientific articles, with the aim of optimizing the searches.

Fig. 1 presents the methodological process followed, including the workflow carried out to develop this research. Step 1 lists the research questions that aimed to be answered in the review. Thereafter, Step 2 shows the scientific databases in which articles were selected from, and the number of articles found and selected following the advanced search criteria, considering the Booleans listed in Table 2. In addition, a second round of advanced search criteria, adding new Booleans, allowed reducing the number of articles from 82 to 60 articles. Finally, a total of 53 were finally considered following the criteria for ruling out duplicates, relevance, and use of AI. Finally, Step 3 shows the criteria considered for filtering the selected research.

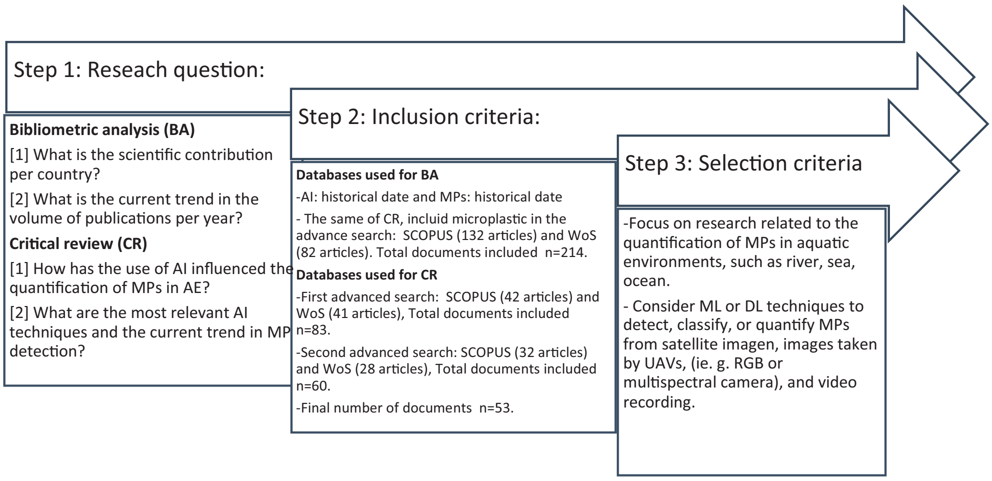
Once the final set of scientific articles were selected and analyzed, two different reviews were conducted. In the first place, a bibliometric analysis was conducted to comprehend the volume and historical trend of research related to the topics assessed in this article (AI and MPs), as well as with their contributions during the time frame assessed (2018–2022). Furthermore, a spatial analysis of the contributions per country was executed.

Secondly, a comprehensive critical review was undertaken, encompassing two thematic axes: research linked to classical ML techniques, and research related to advanced ML techniques (i.e., DL). Whenever applicable, research studies were grouped and discussed jointly based on the use of common imagery elements, such as cameras, drones, and satellites. Thereafter, the criteria for capturing aerial images, as well as the correlation between the efficacy of ML models and their capability to

Table 2

Booleans considered for the advanced search in the critical review. The format for the Scopus and Web Science database is shown.

Criterion	Database	Booleans
First advanced search	Scopus	TITLE-ABS-KEY([“Machine Learning” OR “Deep Learning” OR “Artificial Intelligence”] AND [“Plastic * debris” OR “Marine Plastic*” OR “Plastic litter*” OR “Macroplastic*” OR “Macroplastic*”])
	Web Science	TS = ((“Machine Learning” OR “Deep Learning” OR “Artificial Intelligence”) AND (“Plastic * debris” OR “Marine Plastic” OR “Plastic litter “OR “Macroplastic*”))
Second advanced search	Scopus	... AND [“Detection” OR “classification” OR “quantification”])
	Web Science	…AND (“Detection” OR “classification” OR “quantification”))


Step 1: Reseach question:

Bibliometric analysis (BA)

[1] What is the scientific contribution per country?

[2] What is the current trend in the volume of publications per year?

Critical review (CR)

[1] How has the use of AI influenced the quantification of MPs in AE?

[2] What are the most relevant AI techniques and the current trend in MP detection?

Step 2: Inclusion criteria:

Databases used for BA

-AI: historical date and MPs: historical date

- The same of CR, incluid microplastic in the advance search: SCOPUS (132 articles) and WoS (82 articles). Total documents included n=214.

Databases used for CR

-First advanced search: SCOPUS (42 articles) and WoS (41 articles), Total documents included n=83.

-Second advanced search: SCOPUS (32 articles) and WoS (28 articles), Total documents included n=60.

-Final number of documents n=53.

Step 3: Selection criteria

-Focus on research related to the quantification of MPs in aquatic environments, such as river, sea, ocean.

- Consider ML or DL techniques to detect, classify, or quantify MPs from satellite imagen, images taken by UAVs, (ie. g. RGB or multispectral camera), and video recording.

Fig. 1. Graphical representation of the methodological flow of the semi-systematic review conducted for the period 2018–2022, specifying the number of documents in the field selected.

classify MPs, is discussed. Lastly, the spotlight is directed towards the current state-of-the-art and prospects in the field of detection, classification, and quantification of MPs in AEs.

4. Results from bibliometric review

The historical data for both AI and MPs were examined independently and jointly. The analysis revealed the inception of AI-related studies dates in the early 1960s, whereas research focused on MPs started tenuously in the early 1970s. However, it was not until 2008, as shown in Fig. 2a, that the volume of research related to MPs exhibited a noticeable increasing trend. When AI and MPs are analyzed jointly, the first two studies are identified in 2018, experiencing an important surge to 43 studies in 2022. Fig. 2b presents the total numbers of publications per year in the period 2018–2022. Publications in all three groupings are experiencing exponential growth, although studies in which AI is used to analyze MPs in AEs only represented 2.5 % of total studies in 2022. Hence, it is plausible to assert that the field under study is youthful, yet rapidly evolving.

To assess the impact of the most frequently utilized terms, a word

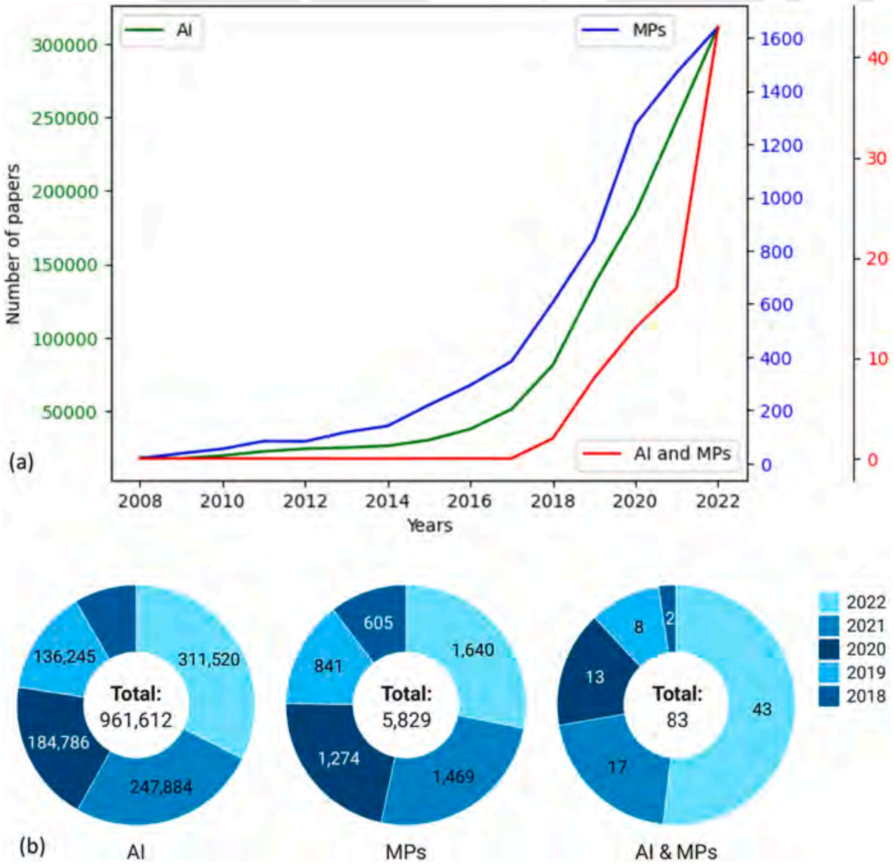


Fig. 2. (a) Time series: Historical trend of the yearly published article count; (b) Pie chart: Article distribution since 2018, marking the initiation of research on artificial intelligence (AI) and macroplastics (MPs).

cloud was generated (see Fig. 3). Words were illustrated in varying sizes, with an increasing size corresponding to a higher frequency of occurrence. When applied to the Scopus database (Fig. 3a), a notable cohesion is observed between the words "machine learning" and "deep learning" with terms such as "object detection," "remote sensing," "macroplastic," and "marine pollution". In contrast, when analyzing the WoS database (Fig. 3b), it is evident that, while considering the same number of occurrences for the map creation, there appears to be less coherence among keywords. For instance, "plastic waste" and "beaches" are close to "DL," whereas "ML" exhibits a stronger correlation with terms like "marine debris," "remote sensing," and "beach litter." Finally, when examining both word clouds, a notable trend emerges: the term "ML" appears more frequently and exhibits a stronger cohesion with other words as compared to "DL." This observation suggests the notion that the utilization DL has been comparatively less explored in comparison to ML.

In the word cloud formed by the research found in Scopus (Fig. 3a), the red group associates related research plus research associated with ML, and, to a lesser extent, with DL, the axis being marine pollution. The green group aggregates research associated with plastics in marine environments, highlighting RF and Support Vector Machine (SVM) techniques. Finally, the blue group relates research that concentrates plastic pollution in marine environments. Certain differences are observed with respect to the groups formed by the research found in WoS (Fig. 3b). The blue group apparently associates research that concentrates plastic pollution in marine environments, including those that use satellite images, while the red group focuses more on techniques associated with DL.

The global distribution of the scientific papers included was visualized in Fig. 4 through the creation of a heat map. The main affiliation of the first author of each publication was considered to determine the country each paper would be assigned to. Fig. 4a represents the distribution of publication related to the use of AI and all types of plastic litter, whereas Fig. 4b represents only those that focus on MPs.

An initial spatial analysis was conducted, encompassing nanoplastics and microplastics. The results revealed that the highest concentration of research efforts is centered in China (15 %), Italy (9 %), the USA (7 %), Germany (6 %), Canada (4 %), Hong Kong (5 %), and The Netherlands (4 %), collectively accounting for 50 % of the total investigations conducted (Fig. 4a). The utilization of AI is notably scarce in South America and Africa, with only 4 publications developed by Brazilian researchers.

Conversely, when focusing solely on MPs (see Fig. 4b), Italy (16 %) and The Netherlands (14 %) lead in terms of number of publications, followed by Japan (8 %), Germany and the US (7 % each), and then UK and Spain (5 % each). No research has been recorded in South America or Africa for MPs.

It is observed that the largest number of articles is concentrated in Europe, with Italy being the country that has produced the most publications (13 papers). This can be directly attributed to the availability of high-resolution images from the Hyperspectral Precursor of the Application Mission (PRISMA) satellite, launched in 2019 by the Italian Space Agency (Loizzo et al., 2018). While its resolution (30 m) is lower than that of Sentinel-2 (10 m), PRISMA offers a greater number of bands within the visible range (Belgiu et al., 2023). As a result, these images have facilitated the exploration of a wide range of ML techniques. Hence, it is likely that motivated by this enhanced access to satellite images for experimentations, Italian researchers have also explored the registration of RGB and spectral aerial images for object detection (Masiero et al., 2022), including the identification of plastic waste (Cortesi et al., 2022).

5. Critical review and discussion

The critical review has been developed following the structure shown in Fig. 5. Consequently, Section 5.1 discusses how remote sensing (RS) and AI allow automating the counting processes for MPs and their advantages over traditional methods. Section 5.2 examines research related to the detection of MPs using classic ML techniques (Section 5.2.1), as well as the use of aerial (Section 5.2.2) and satellite images (Section 5.2.3). Finally, Section 5.3 focuses on the discussion of research related DL techniques, starting the discussion with research related to deep ANNs (Section 5.3.1), followed by CNNs (Section 5.3.2), and the use of pre-trained models based on TL (Section 5.3.3).

5.1. *How is artificial intelligence (AI) connected with the detection of macroplastics (MPs)?*

Over the past few decades, the estimation of marine plastic has been analyzed through the application of data collection criteria (Jambeck et al., 2015), including data linked to production, recycling, and end-of-life (Ita-Nagy et al., 2022). Despite facilitating comprehension of the worldwide extent of plastic pollution—exemplified by its presence in oceans— these methods yield rough estimations. In fact, manual methods to estimate plastic debris in specific areas (e.g., shoreline on the beach), which consist of the use of human eyesight to spot and hands to sort plastic debris items, are predominant in the scientific literature

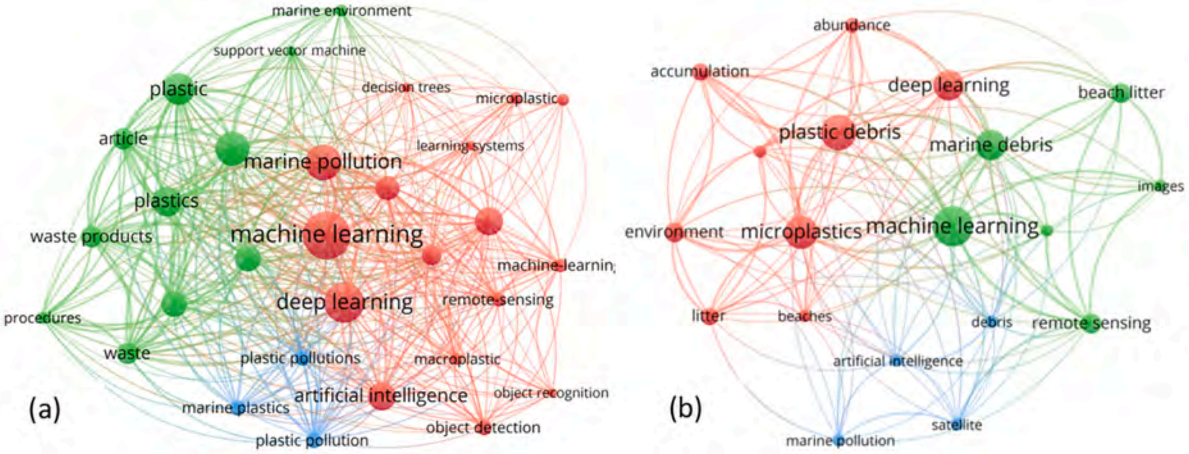


Fig. 3. Graphical representation of a word cloud for studies in which machine learning or deep learning have been applied to detect, classify and or quantify macroplastics (MPs) in aquatic environments (AEs). A minimum of 5 occurrences are considered for the word to appear in the cloud. Aggregations were created represented with 3 distinct groups, which are demarcated with color coding: red, blue, and green. Fig. 3a represents the word cloud derived from Scopus, whereas Fig. 3b represents the word cloud developed with WoS.

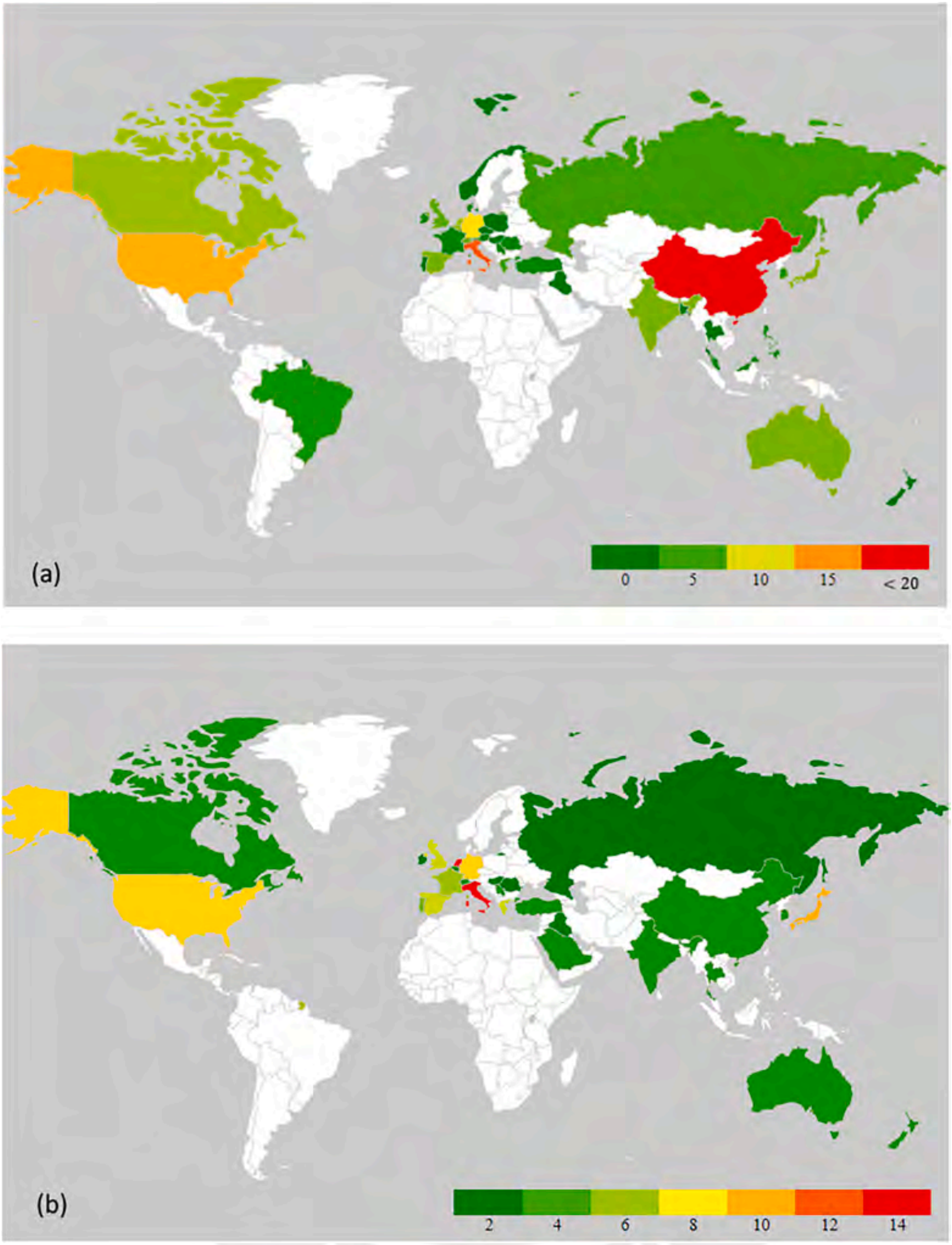


Fig. 4. Heat map representing the number of scientific publications published per country based on the affiliation of the first author in the period 2018–2022 related to the use of artificial intelligence (AI) to measure plastic waste in aquatic environments (AEs). Fig. 4a includes any type of polymers (214 papers). Fig. 4b only includes macroplastics - MPs (83 papers).

(Teng et al., 2022). However, these methods are limited to local areas (Aleem et al., 2022) and require human resources and time (Angelini et al., 2019), depending on weather conditions and observation distance (Teng et al., 2022). Furthermore, according to Song et al. (2022), traditional monitoring techniques pose challenges when applied to the examination of remote and challenging-to-access locations, such as uninhabited islands, rugged coastlines, or coastal cliffs.

The use of GIS and RS facilitates a spatial analysis of plastic debris through the utilization of aerial imagery, captured via UAV (Geraeds et al., 2019) or satellites (Sannigrahi et al., 2022). According to Moy et al. (2018), the integration of these methodologies presents a viable framework for pinpointing the location, distribution, and composition of MPs, avoiding double counting. Notably, while these techniques have been extensively employed (e.g., in shoreline and riverbank contexts), thereby highlighting the potential of GIS and RS, it is crucial to note that the methodology needs a physical inspection within the study areas (Moy et al., 2018). Consequently, it can be inferred that within this framework, the estimation of plastic debris continues to be a non-automated process.

Within this context, and propelled by the advancements in computational power, AI has emerged as an avenue of exploration, aimed at expediting non-automated processes. According to Quetglas et al. (2011), in recent decades, ML models have shown successful results in marine sciences. This affirmation was corroborated by Song et al. (2022), who indicate that within the realm of marine sciences, the use of ML and DL technologies have shown promising traits for their application in marine pollution. In this regard, research related to plastic pollution employing ML/DL has seen notable advancements. For

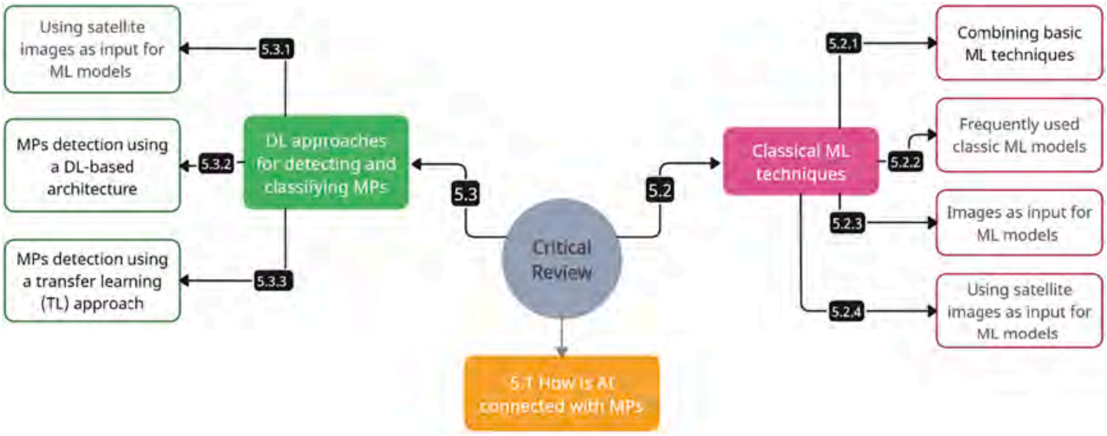


Fig. 5. Conceptual map illustrating the hierarchical structure established within the critical review.

instance, ML/DL has been applied to plastic waste object detection (Fulton et al., 2019; Maharjan et al., 2022), images segmentation (Singh and Valdenegro-Toro, 2021; Song et al., 2022), classification (Armitage et al., 2022; Aleem et al., 2022), and quantification (Wolf et al., 2020; De Vries et al., 2021). In fact, AI has the potential to establish automated processes even leading to early warning systems, as proposed by Garello et al. (2019). In addition, according to Gonçalves et al. (2020a), the use of UAV-derived orthophotos alongside the use of ML requires much less human effort and allows for broader area coverage. However, based on the findings presented in Section 4, it should be noted that AI to detect MPs in AEs has shifted gradually from more classic ML methods (Cortesi et al., 2021) to more robust DL methods (e.g., Lieshout et al., 2020; Amjoud and Amrouch, 2023).

Hence, the potential advantages of AI in MPs detection together with GIS and RS as compared to non-automated methods seems to lie in reducing both time and human resources, while covering extensive land or marine areas efficiently (Martin et al., 2018). However, it is contingent upon familiarity with these techniques and access to computational power, as well as the availability of extensive datasets to provide improved levels of accuracy (Goswami and Goswami, 2022), often requiring fieldwork. Furthermore, according to Moy et al. (2018), once implemented, these methods can readily be extended for application to future large-scale and long-term coastal monitoring datasets. In the automatic detection of objects, the use of images as input for ML models is recurring. The images are commonly aerial photographs recorded by conventional or higher band cameras (i.e., spectral and multispectral), and the use of satellite images is also a valid option.

Fig. 6 shows the steps to follow to adapt an ML model for the detection or segmentation of objects. For instance, in data collection, fieldwork involves gathering datasets, including images and features,

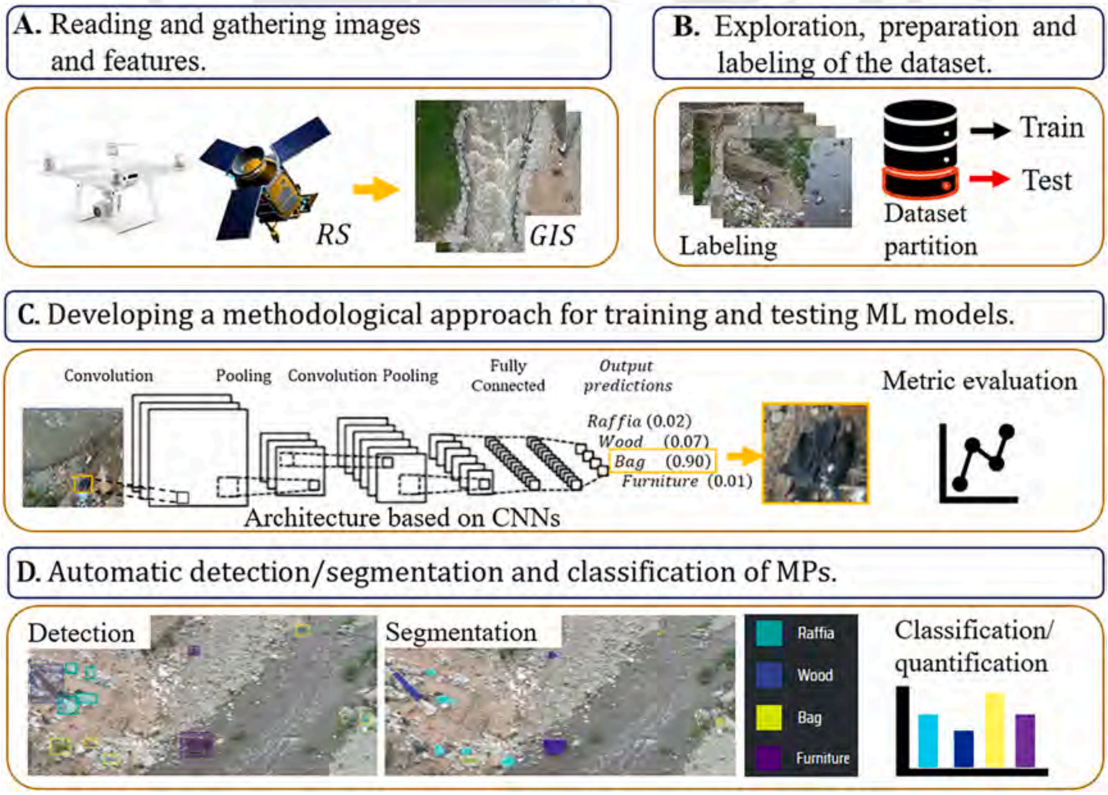


Fig. 6. Graphical representation of a simplified artificial intelligence (AI) integration workflow for detecting macroplastics (MPs). Each cloud represents a potential example in each part of the workflow.

using RS technologies like UAVs, cameras, or satellites, together with GIS. In terms of data preparation, for supervised ML techniques, data exploration and correction are performed, and labeling may be required. Thereafter, establishing an effective training strategy is of paramount importance to ensure the development of a robust model and proceed with the testing. Finally, model validation can be conducted once the model exhibits reasonable efficiency, and can be applied to other datasets.

Since the development of AI, research increasingly points towards the use of more advanced techniques for the detection of MPs. In this sense, synergies are discussed between the use of classical ML techniques (see Section. 5.2), and, thereafter, the most advanced techniques commonly used such as DL approach are reviewed (see Section 5.3). In addition, in the latter section the commonly used datasets and methods are discussed.

5.2. *Classical machine learning (ML) techniques for detecting and classifying macroplastics (MPs)*

While the state-of-the-art in object detection is primarily rooted in the use of CNN architecture (Aggarwal et al., 2023), which belong to the realm of DL, traditional ML techniques continue to be explored in this novel field. However, based on the findings of the scientific review, these studies tend to utilize a combination of classical ML techniques (as shown in Table 3) and, to a greater extent, methods such as SVM and RF. In this context, Section 5.2 is divided in four blocks, in which different levels of ML architecture complexity are analyzed.

5.2.1. *Combining basic machine learning (ML) techniques*

Table 3 presents the contributions made in the exploration of using two or more classical ML methods, with a particular emphasis on the exploration of model ensembles. For instance, the study conducted by Taggio et al. (2022) explores potential combinations of two ML methods, namely K-means and LGBM, for the detection of 12 categories of marine floating plastic. For this, they used satellite images from PRISMA to capture their shapes. The results suggest that the combination of these two unsupervised (K-means) and supervised (LGBM) ML methods can recognize different forms of plastic in water bodies. Although the experiment conducted was a controlled test, with the plastics arranged as sheets in the ocean, it opens the possibility to further explore the ensemble of ML models for the detection of MPs.

Another example of such ensembles has been explored by Hernández-González et al. (2019), who propose a novel approach that uses an aggregated outputs (AOs) framework integrated with linear regression models (LMs), Support Vector Regression (SVR), and K-Nearest Neighbors (KNN) to predict marine litter beaching. While the LM may appear simplistic, their results reinforce the notion that combining a traditional ML model with other frameworks (Martin et al., 2018; Sannigrahi et al., 2022) or methods (Cortesi et al., 2022) enhances training efficiency, reducing prediction errors.

In the effort to dig deeper into the complexities of ML methods, research related to ANNs, which are renowned for being one of the pioneering algorithms and have been extensively investigated (Goodfellow et al., 2016), was examined. According to Goodfellow and colleagues (2016), despite limited success in the 1980s, ANNs are now widely applied across various fields, primarily due to advancements in computational capacity.

In this context, one of the first studies prepared by Franceschini et al. (2019) is highlighted, in which an unsupervised ANN called Self-Organizing Map (Kohonen, 1982) is explored. They focused on estimating the regional quantity of seabed marine litter in the Central Mediterranean Sea in Italy. Their dataset considered environmental (e.g., depth, slope, coastal debris, etc.) and anthropogenic (e.g., distance from rivers and cities, maritime traffic, fishing gear, etc.) variables. The model was trained to differentiate between the first 5 marine litter categories according to the MEDITS protocol (AAVV, 2017). Even though the metrics were not as expected ($R^2 = 0.55$), the contribution revolves around the experimentation of which variables influence the results. Thus, it was concluded that the critical variables were depth, and passive and active fishing gear, while those that had no effect were certain environmental variables, such as coastal terrain, mud, and slope.

Following the exploration of ANNs in the detection and classification

Table 3
Scientific articles related to the combination of use of two or more classical machine learning (ML) methods and/or novel contributions.

Reference	Country (First author)	ML Approach	Dataset collected by			Target	Class	Metric
			Cam/ sensor	UAV	Satellite			(Testing)
Martin et al. (2018)	Saudi Arabia	RF with HOG		x		Beach litter	3	Overall Accuracy = 0.375
Hernández-González et al. (2019)	Spain	[LM, SVR and KNN] + AO	x			Beach litter prediction	–	RMSE = [1.5–2.5]
Franceschini et al. (2019)	Italy	A set of ANNs called SOM	x			Marine litter abundance	5	R2 = 0.55 MSE = 0.86
Gonçalves et al. (2020a)	Portugal	RF and multidiscipli-nary approach		x		Litter density map	2	F1-Score = 0.75
Pinto et al. (2021)	Portugal	Multi-class ANN		x		Marine Litter detection	2 5	F1-score = 0.73 F1-score = [0.37-0.61]
Cortesi et al. (2021)	Italy	RF and SVM	x(M)			River plastic detection (Artificial plastic targets)	2	Accuracy = [0.98–0.92] Recall = [0.84–0.49]
Freitas et al. (2021a)	Portugal	RF and SVM		x(H)		Marine Litter (Artificial plastic targets)	4	Accuracy : RF = [0.973–0.981] SVM = [0.966–0.98]
Taggio et al. (2022)	Italy	LightGBM + K-Means			PRISMA	MFP detection	4	Overall Accuracy = 0.729
Sannigrahi et al. (2022)	Ireland	[RF and SVM] + KNDVI -index			Sentinel-2	MFP detection	2	Accuracy = [0.8–0.9]
Cortesi et al. (2022)	Italy	Two RF classifiers		x (M)		River plastic detection (Artificial scenario)	2	Accuracy and Recall >0.98
Lavender (2022)	UK	ANNs + Decision Tree			Sentinel-1-2	Waste plastics (Terrestrial and aquatic)	4	F1-score = [0.83–1] Recall = [0.85–1]

ANNs: Artificial Neuronal Networks; AO: Aggregated output; H: Hyperspectral; HOG: Histogram of Oriented Gradient; KNDVI: Kernel Vegetation Indices; LightGBM: Light Gradient Boosting Machine; M: Multispectral; MFP: Marine floating plastic, RF: Random Forest; SOM: Self-organizing map; SVM: Support Vector Machine.

of MPs in AEs, Pinto et al. (2021) assessed the ability of ANNs to establish a multiclass structure (i.e., three layers). They considered 5 categories: plastic bottles, fishing ropes, octopus pots and fragments. Using images recorded by UAV, they experimented with RGB color bands and their combinations. The model was trained considering the multiclass and a binary approach. Although the results were not as expected when considering the multiclass classification compared to a binary classification (see Table 3), it represents an advance in comparison with previous studies which consider a binary approach. Interestingly, this study suggests that the altitude at which the images are recorded could influence the ability of ANNs to be able to generate a correct classification, based on the idea that at lower altitudes the images would have a higher resolution, and thus, a better classification. This conclusion converges with the findings mentioned by Martin et al. (2018).

5.2.2. *Frequently used classic machine learning (ML) models*

As SVMs and RFs are widely recognized as the most popular methods among classical ML techniques (Zhang et al., 2017a), they have been extensively applied for the detection and classification of MPs. Although it is not possible to generalize that one method is better than the other in the classification task, Cortes and Vapnik (1995) suggested that SVM is more commonly used for binary classification tasks. In contrast, according to Martin et al. (2018), RF appears to excel in terms of classification accuracy and is more popular when it comes to implementing ensemble learning methods. This observation was further supported by Cortesi et al. (2021), who found that the SVM classifier yielded less satisfactory results compared to RF in MP detection, indicating approximately half the number of false positives as compared to SVM. However, it is important to clarify that SVM has shown efficacy when handling the classification of high spatial dimensions (Tristan Fletcher, 2009), but struggles in MP detection (Martin et al., 2018). In contrast, Freitas et al. (2021a) suggested that the use of both methods shows similar results in terms of commonly used metrics in ML to assess the efficiency of the models (e.g., Accuracy, F1-Score, Precision, etc.) when detecting and classifying MPs on beaches. However, also according to Freitas et al. (2021a), the results obtained when using SVM show a non-uniform detection pattern compared to RF, which provides more stable results when detecting a target.

Gonçalves et al. (2020a) also obtained favorable results when using RF. In this case, the authors focused on the estimation of MPs using aerial images (i.e., multispectral camera), which were obtained by UAV. The method proved capable of successfully detecting and performing multiclass MPs classification. This framework incorporated photogrammetric methods, enabling a geomorphological analysis of the study area, and included hydrodynamic modeling. Furthermore, Cortesi et al. (2022) reported encouraging results to detect MPs when using a combination of two RF classifiers in fluvial environments. Although the efficiency of the training model was high, the results were not as expected in terms of precision and quality. However, the efficiency of assembling two RF classifiers stands out, being a novel contribution in terms of the successful use of classical ML techniques.

5.2.3. *Using aerial images as input for machine learning (ML) models*

The use of aerial images taken from traditional cameras (RGB color band) or multispectral camera, as well as video recordings taken as a video framework (Armitage et al., 2022), play an important role as input for ML model training. In addition, the use of images from hyperspectral cameras may potentially be favorable due to the greater number of bands. However, access to hyperspectral or multispectral cameras is still limited due to high costs, meaning that, many times, access to this technology only allows lower dynamic ranges (DRs) than required in the scenes. In this context, Shaikh et al. (2022) proposed a multi-exposure method to increase the DR of hyperspectral imaging using an InGaAs-based short-wave infrared hyperspectral line camera (SWIR). Employing PCA and two RF models the study showed an improved efficiency of the models. Furthermore, an increase in accuracy from 90 to 98 % was observed as compared to not using SWIR to detect polymers. The SWIR method appears as an interesting alternative to explore since the Shaikh et al. (2022) show that this method reduces the number of errors in the classification of polymers.

According to the study carried out by Martin et al. (2018), the altitude and the angle of inclination of the camera, mounted in UAVs, with which the aerial images are taken, influence the efficiency of the trained models. Tests carried out at various heights show that the best results are given at an altitude of 10 m with a 90-degree gimbal angle. The study highlights how 90° allowed a better recognition of the shape of the MPs in AEs. In relation to the study above mentioned, Cortesi et al. (2022) found that there is a better efficiency of the models when, using the same equipment, they are trained with images taken at an altitude of 30 m than with images recorded at an altitude of 80 m. Gonçalves et al. (2020a), instead, chose an altitude of 20 m rather than 30 m based on previous studies (Deidun et al., 2018). Anyhow, it is probable that more tests are still required to reinforce these studies, as altitude seems to play a preponderant role in the recording of datasets.

Camera inclination or the effects linked to the position of the sun are also other parameters of interest (Sekrecka et al., 2020), as they may play a key role when recording terrain with significant slopes. Moreover, in flat areas there are also areas that could be prone to being affected by the inclination of the camera (e.g., river slopes). Ultimately, limitations in image collection are intertwined with weather conditions, wherein wind represents a climatic variable that can impact the autonomy of multirotor drones (Andriolo et al., 2022).

5.2.4. *Using satellite images as input for machine learning (ML) models*

The use of an adequate dataset, with abundant and high-quality information, is crucial to guarantee high quality results and a valid interpretation (Goswami and Goswami, 2022). In this context, the use of satellite images in ML has been identified as an alternative trend among researchers, since it significantly reduces the need for fieldwork, such as planning UAV flights.

Sannigrahi et al. (2022) and Lavender (2022) performed image analysis from the Copernicus Sentinel 1–2 mission (European Space Agency, 2020), a product widely employed by the scientific community in this field, to detect and classify MPs taking advantage of the various spectral bands offered by Sentinel. While there are no prior studies specifically addressing satellite images from WorldView (NASA, 2013), it is worth noting the research conducted by Kremezi et al. (2022), which employs a DL approach (see Section 5.3). These products have been more frequently utilized on extensive surfaces, particularly on beaches (Taggio et al., 2022), coastlines (Kremezi et al., 2022) and the ocean surface (Jamali and Mahdianpari, 2021). In contrast, less research has been conducted related to rivers. For example, Lavender (2022) and Lieshout et al. (2020) focus on monitoring floating MPs using DL. It is hypothesized that this lower applicability to fluvial systems may be attributable to the resolution of the images in relation to the width of the watercourses. For instance, the resolution of the Copernicus Sentinel-2 missions, which is 10 × 10 m, might hinder its ability to analyze a watercourse if the width of its channel were not sufficiently wide.

This opens the discussion about the resolution of the images and how it influences the target to be detected, as it is clear that the target must have a dimension greater than the resolution of the dataset. However, the advantage of satellite images over traditional images is the number of bands available. According to Cortesi et al. (2022), the use of more bands would potentially allow better detection of plastic in the infrared spectrum, making plastics more distinguishable from the sunlight, although future studies are still required. In this context, a novel contribution was made by Sannigrahi et al. (2022), in which in addition to using the RF and SVM classifiers, the authors added a newly developed index, the kernel Normalized Difference Vegetation Index (kNDVI), to detect and classify marine floating plastic. As expected, RF showed a better performance in the testing stage. It should be noted that unlike other authors, Sannigrahi et al. (2022) used a dataset of images from Sentinel – 2, establishing various models from the bands.

Overall, it appears that classical ML techniques seem to be favorable in the detection of MPs in AEs. However, most of the research reviewed so far focused on a binary classification process, e.g., 'plastic' or 'others', (Cortesi et al., 2021, 2022), or 'plastic' or 'water' (Sannigrahi et al., 2022). Although other authors proposed a more complex classification framework such as water and land, orange plastic, white plastic and ropes (Freitas et al., 2021a, 2021b), or land, shallow water and deep water (Taggio et al., 2022), in essence, the objective was still to distinguish plastic from other elements. An outlier is the study by Martin et al. (2018), who, unlike other authors, differentiated between a wider range of microplastics, including drink containers, bottle caps and plastic bags. However, a relatively low accuracy was obtained (44 %) using a RF model.

5.3. Deep learning (DL) approaches for detecting and classifying macroplastics (MPs)

Considering the requirement not only to detect MPs but also to have the ability to classify the most common polymers present in AEs (Geyer et al., 2017), the scientific community has begun to explore more robust algorithms, such as those related to the DL approach. Based on the bibliography reviewed, it appears that DL approaches cannot only detect MPs but can also classify the most common polymers (Freitas et al., 2021b). This subsection focuses on discussing the deployed methodologies, the novel contributions, the classification level established for the MPs, and the results obtained using DL approaches. Among the methods in greatest demand, the use of CNN, and the use of pre-trained models, known as TL, are the most abundant in the literature. In this context, the remainder of Section 5.3 is divided in three blocks, in which different types of DL architectures are presented and discussed.

5.3.1. Deep artificial neural networks (ANNs) as deep learning (DL) approaches

While ANNs are traditionally associated with classic ML techniques, they can be categorized under DL when they incorporate a multilayer architecture, e.g., more than one hidden layer (LeCun et al., 2015). A notable contribution utilizing deep ANNs was presented by Kako et al. (2020). They introduced a novel approach for estimating the volume of plastic marine debris by employing a multilayer ANN with two hidden layers, unlike the approaches presented by Franceschini et al. (2019) and Pinto et al. (2021), which incorporated a hidden layer into their architecture. This ANN was used for the binary classification of plastic waste using a 3D model generated from a photogrammetric process of aerial images captured by UAV. The results (error < 5 % and accuracy $= 98$ %) suggest the approach has the potential to estimate the volume of plastic waste by considering the Hue Saturation Value (HSV) color space on the beach. However, the authors identified that the method may not provide good results in areas where the color of the plastic waste is similar to the color of the terrain.

Another study related to the utilization of deep ANNs was conducted by Lavender (2022). In contrast to Kako et al. (2020), Lavender presented an ANN architecture in which *Dropout* was applied at the output of each hidden layer, consisting of two hidden layers. *Dropout* is a regularization method used in ANNs to mitigate overfitting and enhance model performance (Srivastava et al., 2014). In their methodology, Lavender (2022) introduced an additional post-processing step using a decision tree to enhance the classification process, complementing the use of ANNs. The classification task involved the consideration of 4 distinct subcategories, namely water, land, urban, and plastic, resulting in a total of 20 classes. It is important to note that only four of these classes pertain to different types of plastic materials. The study employed Sentinel 1–2 mission images, encompassing a total of 17 bands, across various AEs. This contribution can be viewed from two significant perspectives. Firstly, it involves the collaborative exploration of architectures based on DL, deep ANNs, i.e., multilayer, in conjunction with ML through the application of a decision tree. Secondly, it addresses the research challenge of experimenting with a dataset containing 17 bands, which had undergone prior preprocessing steps.

5.3.2. Macroplastics (MPs) detection using a DL-based architecture

CNNs are regarded as a pivotal advancement in DL for computer vision due to their significant learning capabilities (Goodfellow et al., 2016). CNNs enable them to handle complex tasks such as object recognition (Mahdianpari et al., 2018), showing improved results as compared to ANNs (Krizhevsky et al., 2017). In this context, as shown in Table 4, the exploration of DL-based architecture, particularly CNNs, in the domain of plastic waste detection has gained momentum in recent years. Therefore, this subsection is dedicated to a critical review of DL-based architectures and their synergy with conventional ML techniques as supplementary or comparative models.

In the literature review, it becomes evident that there has been limited exploration of the utilization of hyperspectral images. This lack of exploration is likely attributable to the significant financial investment needed for the acquisition of such equipment. Nonetheless, among the scarce studies performed, those conducted by Freitas et al. (2021b, 2022) are highlighted. Their initial contribution lies in the development of a novel architecture termed CNN-3D, wherein NNs incorporate both spatial and spectral information (i.e., bands) to categorize marine litter into 4 distinct classes. The categories of marine litter targeted for detection were artificially defined within the ocean, using a 10 × 10 grid system, a method derived from Freitas et al. (2021a). These procedures, such as those described in Kako et al. (2020) and, subsequently, in Cortesi et al. (2022), have been previously documented.

It is essential to observe that the accuracy achieved in Freitas et al. (2021b), which uses CNN-3D, is lower than the results obtained when employing SVM and RF in a prior study conducted by Freitas et al. (2021a). However, according to Freitas et al. (2021b), the reduction in efficiency is due to making the process run faster as well as manual annotation of the data. Furthermore, according to El Zaar et al. (2022), a substantial dataset is required for DL-based architectures to be efficient. Another case where a classical ML method performed relatively better than a DL-based method was done by Gonçalves et al. (2020b) who obtained better performance in classifying marine little using RF compared to CNN (see Table 4). In fact, Gonçalves et al. (2020b) based CNN on an architecture known as *densely connected networks* (DenseNet) (Huang et al., 2017), whose complexity was reduced to adapt it to the purpose of the research.

Studies like those conducted by Wolf et al. (2020) and Jamali and Mahdianpari (2021) employed DL-based architectures, comparing their outcomes with those achieved using classical ML methods such as RF and SVM. It is crucial to underscore that both methods demonstrated higher efficiency when adopting the DL approach, in line with the observations of El Zaar et al. (2022). In the case of Wolf et al. (2020) their main achievement lies in establishing 18 distinct classes for the detection and classification of marine debris, 12 of which corresponded to plastic litter, constituting the highest number of classes found in the literature assessed. In contrast to the results obtained by Jamali and Mahdianpari (2021), who designed a system with 4 classes; however, it is conceivable that the effectiveness of the proposed architectures may be inversely proportional to the number of classes targeted for detection (see Table 4).

The contributions by Wolf et al. (2020) and Freitas et al. (2021b) are characterized by developing innovative architectures based on CNNs. Conversely, the study by Jamali and Mahdianpari (2021) stands out for exploring Generative Adversarial Networks (GANs), as proposed by Goodfellow et al. (2014), and their ensemble with RF (i.e., a classical ML method). Unlike Wolf et al. (2020), Jamali and Mahdianpari (2021) used Sentinel-2 images as their dataset. In parallel, Freitas et al. (2022) employed a novel unsupervised architecture known as Zero-Shot Learning for the detection and classification of marine litter. While the

Table 4
Scientific articles related to the utilization of DL-based architecture.

Reference	Country (First author)	DL Approach	Dataset collected by		Target	Class	Metric	
			UAV	Satellite			Training	Testing
Wolf et al. (2020)	Germany	APLASTIC-Q (based on CNN), 03-SVM and RF	x		Plastic litter (Detection and quantification)	18	precision, recall, F1-score > 0.81	Accuracy: APL-Q = 0.83 SVM = [0.78–0.47] RF = 0.71
Gonçalves et al. (2020b)		CNN and RF	x		Marine litter (Density map)	9	N/A	F1-score: RF = 0.70 CNN = 0.60
Freitas et al. (2021b)	Portugal	CNN-3D	x(H)		Marine litter (Artificial plastic targets)	4	N/A	Accuracy = [0.84–0.917]
Kremezi et al. (2021)	Greece	PCA and 03 CNNs: PNN, CAE and GDD.		PRISMA	Marine floating plastic	4	N/A	R = [0.94–0.99]
Jamali and Mahdianpari (2021)	Turkey	GAN-RF, RF, SVM		Sentinel-2	Large scale marine plastic (Detection)	2	Accuracy: GAN-RF > 96 % RF > 88 % SVM > 84 %	Accuracy: GAN-RF = 96 RF = 0.88 SVM = 0.84
Freitas et al. (2022)	Portugal	Zero-Short Learning	x(H)		Marine Litter (Artificial plastic targets)	3		Accuracy = 0.987
Kremezi et al. (2022)	Greece	06 CNNs: PNN, SRGAN, RCAN, Fusion-PNN-Siames, Fusion-ResNet, Fusion-GAN		Fusion: Sentinel-2 and WorldView-2/3	Marine plastic (Artificial plastic targets)	5	N/A	Accuracy = [0.60–0.99]
Song et al. (2022)	Republic of Korea	U-Net model based on FCN	x		Marine Debris	6	Average F1-score = 0.74	F1-score = [0.26–0.97]

CAE: Convolutional Autoencoder; FCN: Fully Convolutional Network; GAN: Generative Adversarial networks; GDD: Guided deep decoder; H : Hyperspectral imaging; M: Multispectral imaging; PNN: Pansharpening Neuronal Networks; RCA: Residual Channel Attention network; SRGAN: Super-resolution GAN.

research was constrained to classifying three types of marine litter, the efficiencies achieved surpass those attained in prior studies (see Tables 4–5).

Based on the literature used in this critical review, it is evident that the utilization of imagery from the Sentinel-1 and Sentinel-2 missions has gained widespread popularity in scientific communication. However, its application had been confined to the detection and classification of MPs while considering a low resolution (e.g., 10x10m). In this regard, the contribution by Kremezi et al. (2022) proposes DL-based architectures aimed at enhancing the resolution of Sentinel-2 imagery by integrating it with WorldView images. More specifically, they proposed the fusion of the 13 bands of Sentinel-2 images with band from WorldView-2/3, a high satellite resolution, 2 × 2 m, to detect marine plastic using image fusion techniques (Yokoya et al., 2017). Experiments were controlled by setting artificial plastic targets over the ocean considering 4 plastic materials as targets: HDPE, PET, PS and others. Furthermore, they implemented and compared 6 DL approaches. The first three were employed in accordance with current state-of-the-art in DL super-resolution methods, while the subsequent ones were specifically developed for the study (see Table 4). In addition, based on the use of CNNs, they proposed other architectures such as GAN (Goodfellow et al., 2014) and ResNet (Ronneberger et al., 2015). Their results suggest that the use of image fusion is a valid alternative to improve the quality of marine plastic detection.

In Table 4 it can be seen that the most complex models (i.e., CNNs, PNN and U-net) are associated with a greater number of classes, e.g., 18, 9 and 6 classes types of waste that are not necessarily plastic. Furthermore, the metrics obtained by Wolf et al. (2020), who consider a total of 18 classes in the detection, present lower efficiency compared to other research, such as Jamali and Mahdianpari (2021), with 2 classes, or Freitas et al. (2022), with 4 classes. More specifically, Wolf et al. (2020) reached a highest accuracy value of 83 %, whereas Jamali and Mahdianpari (2021) and Freitas et al. (2022) were able to attain values above 95 %.

5.3.3. *Macroplastics (MPs) detection using a transfer learning (TL) approach*

According to Kylili et al. (2019) and, in accordance with the scientific review carried out, their study represents the first associated with the use of pre-trained models based on CNNs, a technique known as TL, for the detection and classification of MPs in AEs (see Table 5). In this context, the first model of this category they explored was VGG16 (Simonyan and Zisserman, 2014), which has been pre-trained on the large-scale *ImageNet* dataset (Krizhevsky et al., 2012). This TL technique relies on the parameters established during pre-training and, subsequently, employs them as a foundation for training the research dataset. Kylili et al. (2019) proposed a prototype device mounted onboard a marine vessel to collect the dataset (1200 images of a beach in Cyprus), as well as the training strategy through *data augmentation* (van Dyk and Meng, 2001), which increases the set of images of the dataset by generating versions slightly modified from the original images; and *L1* and *L2 regularization*, a strategy used to force the parameters not to increase. The efficiency computed, as shown in Table 5, in the classification of 3 types of floating marine plastic (e.g., plastic bottles, plastic buckets and plastic straws) suggests that the use of the TL technique “is not only capable of improving detection accuracy but can also accelerate the plastic trash identification process”. Furthermore, the authors refer to the number of training stages and the number of images, to be used as a preponderant factor for obtaining acceptable metrics.

The bulk of the research associated with TL techniques is found in countries in Europe (see Table 5). In addition, the hyperparameters associated with the models used are shown (i.e., parameters that require calibration for a correct prediction), with the Learning rate (Lr) parameter being the one that presents the greatest variation, probably as a consequence of the type of model and size of the dataset, the number of classes, among others. It should be noted that when analyzing the nature of the models, differences in the metrics can be noticed. For example, when the VGG16–19 architecture (Simonyan and Zisserman, 2014), which employs multi-stage classification processes, and YOLO (Redmon et al., 2016), which performs object detection more efficiently by considering the entire image field in a single operation, are compared, it

Table 5
Scientific articles related to the exploration of one or more models based on TL.

Autor	Country	Model	Computational	Hyper	Dataset collected by			Class	Metric	
	(First author)		Power	Parameters	Other	Cam/ video	UAV		Training	Testing
Kylili et al. (2019)	Cyprus	VGG16	CPU: Xenon; GPU: NVIDIA Quadro k4200 (28.6 GB)	Epoch = 50 Batch size = N/A Lr = None	Prototype device mounted in vessel			3	Accuracy ≈ 1.00	Accuracy = 0.86
Kylili et al. (2020)	Cyprus	VGG16 with BM structure	From Kylili et al. (2019).	Epoch = 50 Batch size = 5 Lr = 0.001	From Kylili et al. (2019).			8	Accuracy = 0. 98	Accuracy = 0.90
Kylili et al. (2021)	Cyprus	YOLOv5 and YOLACT++	Google-Colab platform: NIVIDIA Tesla K80 GPU	Epoch = 300 Batch size = 10 Lr = 0.01	From Kylili et al. (2019).			7	Accuracy: YOLOv5 ≈ 0.995 YOLACT++ ≈ 0.945	Accuracy: YOLOv5 = 0.995 YOLACT++ = 0.945
(Marin et al., 2021)	Croatia	VGG19, InceptionV3, Inception-ResNetV2, ResNet50, DenseNet121 MobileNetV2	N/A	Epoch = 100 Batch size = 16 Lr = [10^{-3} -10^{-4}]	Deep-sea Debris the Japan Agency JAMSTEC			6	N/A	Accuracy = [0.77–0.91] F1 score = [0.76–0.92]
Papakonstantinou et al. (2021)	Greece	VVG16; VVG19; DenseNet121; DenseNet169 and DenseNet201	CPU:Intel i7 8700 GPU: NVIDIA GeForce RTX2070 (8Gb)	Epoch = 40 Batch size = 64 Lr = 0.000001			x	2	Overall accuracy >0.80	Accuracy: VVG16 = 0.68 VVG19 = 0.77 DenseNet~0.61
Song et al. (2021)	Republic of Korea	YOLOv5	N/A	N/A			x	7	Overall mAP = 0.87	mAP = [0.80–0.93]
Aleem et al. (2022)	Pakistan	faster–RCNN +: VGG16; ResNet50 and ResNet50+ augmentation (AU) + pre-processing (PP)	Google-Colab platform:16Gb Ram and 12Gb GPU memory.	Epoch = 150 Batch size = 64 Lr = 0.001	Marine Debris Dataset FLS			10	N/A	Accuracy: Faster-RCNN +: VGG16 = 0.91 ResNet50 = 0.93 ResNet50 + AU + PP = 0.96
Almeida et al. (2022)	Portugal	MobileNetV2	Google-Colab platform: NIVIDIA Tesla P100 PCI-E 16Gb GPU	Epoch = 200 Batch size = 12 Lr = 0.004			x	9	N/A	F1-score = [0.56–0.66] Precision = [0.74–0.78]
Armitage et al. (2022)	UK	YOLOv5	MAGEO Cluster, provided through the NEODAAS	Epoch = 50 Batch size = 32, 128 Lr = 0.001		x		4	N/A	Accuracy =0.95 F1-Score = 0.64 Precision = 0.68
Luo et al. (2022)	China	YOLOv3 and 5 Faster – CNN SSD	GPU of NVIDIA Pascal™ Architecture with 256 CUDA cores	N/A			x	3	mAP >0.80	mAP = [0.73–0.86] Precision = [0.73–0.88] Recall = [0.65–0.89]
Maharjan et al. (2022)	Thailand;	YOLO family models	Intel i7, 16Gb Ram and GPU NVIDIA RTX 2060 (6Gb)	Epoch = 100 Batch size = 16, 32, 64 and 128 Lr = 0.01–0.001				6	N/A	mAP for family YOLO: =[0.37–0.83] F1-score = [0.32–0.78] Precision = [0.51–0.83]
Moorton et al. (2022)	UK	CNN y VGG-16	CPU: i7 7700HQ, 16Gb Ram; GPU: NVIDIA GTX	N/A	The Japan Agency JAMSTEC.	x		2	VGG-16 Accuracy≈ 1.0 CNN Accuracy[0.90 – 0.93]	Accuracy:VGG-16 = 0.95 CNN = 0.89
Takaya et al. (2022)	Japan	RetinaNET	N/A	Epoch = 60 Batch size = 6 Lr = 0.01–0.001		x		2	N/A	Average: Precision = 0.76 F1-score = 0.61
Teng et al. (2022)	USA	YOLOv5	Google-Colab platform: NIVIDIA Tesla K4 CPU	Epoch = 300 Batch size = 10 Lr = 0.01		x	x	9	Accuracy = 0.995	Accuracy = 0.80

(*continued on next page*)

Table 5 (*continued*)

Autor	Country (First author)	Model	Computational Power	Hyper Parameters	Dataset collected by			Class	Metric	
					Other	Cam/video	UAV		Training	Testing
Veerasingam et al. (2022)	Qatar	YOLOv5	N/A	Epoch = 150 Batch size and lr = N/A		x		7	N/A	mAP = 0.91

BM : Bottleneck method; DenseNet: Densely Connected Networks; FLS: Forward Looking Sonar Image; MAGEO: Massive GPU for Earth Observation; MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications; NEODAAS: Earth Observation Data Acquisition and Analysis Service; R-CNN: Region-based Convolutional Neural Networks; ResNet: Residual Net; VGG: visual geometry group; YOLO: You Only Look Once.

can be noted that the metrics obtained by YOLO are slightly higher. Furthermore, this occurs in a similar way when reviewing the metrics of InceptionV3 (Szegedy et al., 2016a), which focuses on improving computational efficiency and performance through multi-sized convolutions and "inception" modules, which appears to be better than VGG16 but not more than the YOLO family. Although the statements made are based on the comparison of the scientific research given in Table 5, it must be highlighted that stating that one model is better than another requires a comparative analysis under equivalent methodological conditions. In the detection process, for instance, in the case of proposing early warning systems, it is likely that YOLO family models will be the most appropriate. On another note, as expected, the metrics in the training stage reach optimal values, outperforming the metrics in the testing stage. This is because in the training stage the model adjusts the parameters until the predictions are optimal.

A subsequent research was performed by Kylili et al. (2020) employing the Bottleneck Method (BM) algorithm in conjunction with VGG-16. This enabled them to enhance classification predictions, increasing accuracy from 86 % to 90 %, while also expanding the number of classes from 3 to 8, without compromising efficiency (see Table 5). However, as noted by the authors, the BM technique presents certain limitations, e.g., the fact that the BM classifier is specifically designed to identify marine debris items from static images. In other words, the BM can label the type of plastic waste that appears in a snapshot image (i.e., static image), so its use would not be possible when using recordings. Nevertheless, VGG-16 has proven to be a useful architecture for detecting and classifying MPs. In fact, Aleem et al. (2022) used it for increasing the number of classes to 10, whereas Moorton et al. (2022) trained it using images from Sentinel-2.

Other pre-trained models that were considered relevant include the contribution of Papakonstantinou et al. (2021), who compare 5 architectures based on CNNs. The first two refer to the VGG-16 and VGG-19, and the remaining three focus on the DenseNet family (Huang et al., 2017): DenseNet-121, DenseNet-169 and DenseNet-201. The authors discuss that one of the main limitations of very deep CNNs is the loss of information as the network progresses. However, this problem has been solved by making connections between all the layers, increasing the flow between them, in what has been named DenseNet. Thereafter, they generated a marine litter mapping considering a binary classification. They found that DenseNet was superior to VGG-16 although VGG-19, which is an improvement of VGG-16, presented the highest efficiency (see Table 5).

Marin et al. (2021) stands out with the highest number of comparisons and experimentation involving pre-trained models. They compared 6 architectures based on pre-trained CNNs: VGG19, InceptionV3 (Szegedy et al., 2016a), ResNet50 (He et al., 2016), Inception-ResNetV2 (Szegedy et al., 2016b), DenseNet121 (Huang et al., 2016) and MobileNetV2 (Sandler et al., 2018). Their results suggest that the inception-Resnetv2 (also known as InceptionV4) and InceptionV3 model obtained the best efficiencies surpassing VGG-19 in the classification of 6 classes of marine debris (see Table 5). It should be noted that the training process was possible by using a large, annotated dataset on underwater debris from the Japan Agency for Marine and Earth Sciences and Technology (JAMSTEC) and complemented with Google imaging.

Marin et al. (2021) performed combinations with classic ML methods (i.e., NN, RF, SVM, NB, LR and KNN), finding that the best combination was given by ResNetV2 and SVM. Along with the work of Jamali and Mahdianpari (2021), who combine DL-based architectures with conventional ML models, the results they found warn about the need to expand these comparisons. In contrast, they found that MobileNetV2, pre-trained with ImageNet, presented the lowest efficiency. However, they suggest that this architecture should not be easily dismissed as it has several advantages, such as faster performance due to reduced network size and low latency. It should be noted that MobileNetV2 was also explored by Almeida et al. (2022), who found efficiencies similar to those of Marin et al. (2021), although Almeida increased the number of classes to 9, a methodological choice that could also be a factor for low efficiency.

Another contribution was provided by Aleem et al. (2022), who assessed the Faster R-CNN model by taking advantage of the pre-trained ResNet and VGG-16 architectures, integrating them with Faster R-CNN (as baselines). Two key accomplishments in the state-of-the-art can be highlighted from this study. On the one hand, they achieved high performance (see Table 5), even when addressing the classification of 10 classes of marine debris, a notably larger classification than other studies related to DL models that incorporate TL techniques. On the other hand, they improve the training time compared to prior modeling conducted by Valdenegro-Toro (2019). This improvement was achieved by employing pre-trained models (i.e., VGG-16 and ResNet) as backbones for Faster R-CNN, thereby optimizing the training process. Furthermore, the obtained results demonstrate that ResNet yields superior outcomes compared to the utilization of VGG-16. The combination of Faster R-CNN with ResNet, along with the implementation of training strategies such as *data augmentation* (van Dyk and Meng, 2001), and *pre-processing* (Pizer et al., 1987), constitutes the most effective model.

Following the introduction of Faster R-CNN and pre-trained models previously described in the preceding paragraphs, Redmon et al. (2016) presented the YOLO model, which was pre-trained from a large-scale object detection, segmentation, and captioning called the *COCO* dataset (Lin et al., 2014). To date, this architecture has undergone several improvements and is currently in its 8th version – YOLOv8 (Reis et al., 2023). A total of 6 articles included in this review have assessed the capabilities of this model family, with version YOLOv5 (Jocher et al., 2020) being the latest one employed up to the cut-off performed in February 2023 linked to MPs in AEs in the framwork of the current review. In fact, YOLO family models are the most sought pre-trained models in the literature, surpassing VGG models. This can be explained because YOLO, according to Fulton et al. (2019), has the potential to produce results as high as other approaches by using high computational power, with the advantage of being faster.

In this context, Song et al. (2021) conducted a study in South Korea in which they assessed conventional and automated methods to predict the standing-stock of beach debris. They used YOLOv5, which allowed them to successfully carry out a classification of 7 categories of marine debris, such as *styrofoam*, metal rubber, glass, fishing gear and unspecified debris. The pre-trained model was trained to detect beach debris with images taken by conventional cameras mounted on a tripod (precision = 0.88 mAP). They obtained errors below 4 % in the prediction of the estimation of the standing-stock. In fact, Song et al. (2021) highlight the potential of their methodology for the development of effective systems and policies in the management of debris in littoral ecosystems.

Other similar studies include Luo et al. (2022), who focused on monitoring water surface plastic waste in real time using UAV, comparing the capacity of YOLOv3 model with Faster-CNN, or Kylili et al. (2021) who, besides using YOLOv5, explored YOLACT++ (Bolya et al., 2019). The authors determined that both models attained very high accuracy (> 90 %). However, Yolov5 was found to be better than YOLACT++ when using images containing plastic debris, while YOLACT++ showed slightly better performance when using video footage.

The use of the TL approach allows establishing a greater number of marine debris groups (i.e., classes) as objects to be detected. Thus, the study by Teng et al. (2022) proposes the detection and classification of 9 classes of beach debris, the largest identified in the literature reviewed, as well as automating the counting of suggested marine debris categories using video recordings. The methodology proposed, which is based on the modeling by Kylili et al. (2018), allowed generating a correct classification of marine plastic. In fact, despite having a high number of classes to identify and classify, the efficiency of the model was not compromised (accuracy = 80 % and precision = 89.4 %). Moreover, the experimentation carried out by Teng et al. (2022) corroborates that when expanding from 7 to 9 categories, the efficiency was not being affected as may have been expected in the training stage. Finally, their results exceeded the efficiency in counting marine debris compared to the study carried out by Papakonstantinou et al. (2021) with the VGG-19 pre-trained model (see Table 5).

Two studies by Armitage et al. (2022) and Veerasingam et al. (2022) assessed the performance of YOLOv5 when classifying marine debris from video recordings and images captured by UAVs. Armitage et al. (2022) used a camera mounted on a vessel, similarly to Kylili et al. (2019), to record videos of floating plastic litter on a beach in southwest England. Their results indicate that the metrics obtained in the performance of the YOLOv5 model can detect and classify marine plastic litter, although they only considered 3 classes of marine plastic litter. In contrast, Veerasingam et al. (2022) considered 7 classes. The latter study obtained a higher efficiency than Armitage et al. (2022), although this could be attributable to the larger set of images collected, as they established in a case study for an uninhabited island in the Gulf of Arabia. The results suggest that the YOLOv5 model has the potential to monitor ocean areas to conserve marine ecosystems and remote islands, where on-site access is difficult.

As mentioned above, the YOLO model has been constantly updated since its creation in 2016. Additionally, Maharjan et al. (2022) evaluated various versions within the YOLO model family. Their research, together with being the only study linked to MPs in AEs comparing the YOLO model family (i.e., versions 2–5), stands out as one of the few studies considering the detection and classification (6 classes) of MPs in rivers. The authors highlight the challenge of detecting MPs in rivers, as they lack a definite shape, size, or thickness in every river, contributing to the difficulty. Additionally, some MPs can be confused with the coloration of watercourses and land. As hypothesized by the authors, the YOLOv5 model was superior to the previous versions, although it is notable that the metrics obtained differ little from version 3 and 4. Given that rivers are significant contributors to plastic pollution in oceans, and considering the wide range in estimates of river contributions, there is an urgent need to gather more observations (Garello et al., 2019) beyond the study by Maharjan et al. (2022).

Finally, Takaya et al. (2022) conducted the only study associated with the evaluation of the RetinaNET pre-trained model (Lin et al., 2017). They focus on the detection and binary classification of marine debris from images taken by UAV. Their results show that RetinaNet presents many false negatives, which has an impact on the efficiency of the model. In fact, based on a study carried out by Daniel and Aleksandr (2023), YOLOv5 outperforms RetinaNet when comparing the metrics obtained with each model (Song et al., 2021; Armitage et al., 2022).

Fig. 7 shows the number of scientific articles that have used pre-trained models (marked in blue). The YOLOv5 model leads with 5 articles, followed by the VGG-16 model with 4 articles. In addition, the result of the metrics found in the testing stage is presented (red and green). In this regard, beyond the most explored models, the “Faster-RCNN + ResNet50” model stands out with an Accuracy >0.9. Finally, Fig. 7 shows the maximum number of classes considered when performing the detection or segmentation tasks, highlighting the VGG16, Faster-RCNN + VGG16, Faster-RCNN + ResNet50 and YOLOv5 models.

6. Conclusions and perspectives

The present critical review, to the best of our understanding, is the first to incorporate a bibliometric analysis in addition to the critical analysis in the topic of AI and MPs detection in AEs. For the former, it is important to highlight those articles related to MPs in AEs using AI are still limited in the scientific literature, with most contributions concentrating in a limited number of countries, such as China, the US, and Italy, as well as other European Union countries. In contrast, studies in most of the Global South are scarce or yet to be developed.

The critical analysis allowed corroborating that the use of AI in the detection and quantification of MPs in AEs has been evolving to explore an increasing number of modeling techniques. In this context, the scientific community is increasingly interested in the use of DL approaches. In fact, the research reviewed suggests that the ability to further classify MPs types goes hand-by-hand with the use of TL. Notwithstanding, available research suggests that both ML (e.g., SVM and RF) and DL (e.g., ANN, CNN and TL) approaches have sufficient capacity for MPs detection, considering a binary classification, while for a multiclass classification the use of DL is recommended. However, the research reviewed points towards the need to continue exploring how image registration, flight altitude, type of camera (i.e., the number of image bands) or registration form influence the efficiency of the models to accurately detect MPs, regardless of whether the ML or DL approaches are applied. As of the cut-off date of this review, the application of transformer-based architectures had not been explored for the detection of MPs in AEs.

The results found in the research reviewed aim to be able to classify an increasing number of types of MPs, including other types of materials such as stones, sand, vegetation, to reduce false positives in the detection of MPs. Hence, the demand for images with higher resolution and a greater number of bands as input for the exploration of increasingly complex models is required. Consequently, it is likely that the limitation in the future will be the use of multispectral and hyperspectral cameras for the generation of datasets, in addition to more thorough fieldwork.

The temporal factor in the complexity of DL models, especially pre-trained architectures, demands significant computational resources and large datasets, constituting a limitation for their exploration. However, it is plausible to assume that future trends may involve optimizing these architectures to reduce training times (e.g., by reducing hyperparameters and the need for extensive datasets to enhance effectiveness). Conversely, given the complexity of certain cases, there may be a need to propose specific architectures targeting a particular set of polymers, as suggested by Shaukat et al. (2020a). Undoubtedly, the temporal factor could play a crucial role in choosing between traditional ML models and DL models, particularly if the disparities in accuracy are minimal. In addition to aiming for improved accuracy in object detection, however, it is crucial to consider the potential threat of adversarial attacks (Shaukat et al., 2022) for almost real-time monitoring systems (Luo et al., 2022), relying on images or videos. Models should not only aim for high performance metrics, but also be equipped to defend

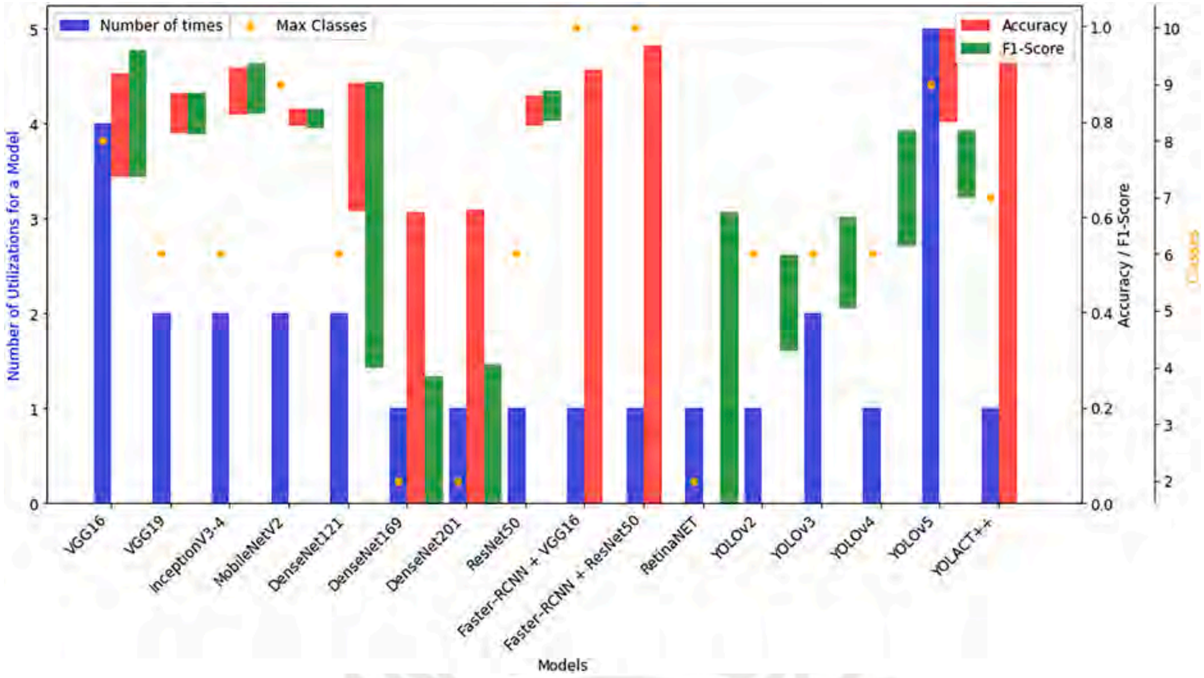


Fig. 7. Bar chart for each model based on the transfer learning (TL) approach. The left hand-side Y-axis represents the frequency of scans in blue. The first right hand-side Y-axis scale indicates the values of Accuracy (red bars) and F1-Score (green bar) metrics, displaying the range between the maximum and minimum values. The second right hand-side Y-axis shows the maximum number of classes detected or segmented, presented as yellow points.

against such threats.

Therefore, it is likely that in upcoming years the trend in the scientific literature will be to explore ensembles between ML and DL approaches, as well as how the number of epochs in the training stage and the size and quality of the dataset influence the capacity to reduce prediction errors. This tendency to combine ML and DL architectures has already been identified in other sectors (Shaukat et al., 2023, 2024). Furthermore, the advancement of computational capacity and new architectures could establish new strategies in the future in models based on increasingly robust assemblies that will require greater analysis and evaluation.

Finally, while efforts to estimate the amounts of MPs present in AEs supported by AI constitute a novel area of research, the bibliometric analysis reveals a notable surge in such studies in the period 2020-2022, indicating a positive trend. These efforts must go hand in hand with policy-making that allow the establishment of legislation and mitigation and waste management measures based on site-specific realities. However, for AI to reduce the gap in terms of data quality and availability, it is necessary to establish standards that are in line with the reality of the systems under study, as well as the creation of control stations strategically located around the globe that can serve as a basis for future monitoring centers at municipal or regional level. The absence of spatial and temporal data poses a significant limitation to progress in the use of these models to support in MPs detection in AEs, but it is expected that this gap will be gradually closed.

CRediT authorship contribution statement

Miguel Angel Astorayme: Writing – original draft, Methodology, Investigation, Formal analysis. **Ian Vázquez-Rowe:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Conceptualization. **Ramzy Kahhat:** Writing – review & editing, Validation, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare no conflict of interest.

Data availability

Data will be made available on request.

Acknowledgment

MSc Miguel Angel Astorayme wishes to thank ProCiencia for funding his PhD Programme through agreement N°174-2020. Profs. Ian Vázquez-Rowe and Ramzy Kahhat thank the Natural Environment Research Council (NERC) of the United Kingdom for financial support via the “Reducing the impacts of plastic waste in the Eastern Pacific Ocean” project (NERC reference: NE/V005448/1).

References

AAVV, 2017. MEDITS-handbook. Version n.9. MEDITS group 106p. http://www.sibm.it/MEDITS%202011/principale%20project.htm.

Aggarwal, A., Kumar, V., Gupta, R., 2023. Object detection based approaches in image classification: a brief overview. 2023 IEEE Guwahati Subsection Conference (GCON) 1–6. https://doi.org/10.1109/GCON58516.2023.10183609.

Aleem, A., Tehsin, S., Kausar, S., Jameel, A., 2022. Target classification of marine debris using deep learning. Intelligent Automation & Soft Computing 32 (1), 73–85. https://doi.org/10.32604/iasc.2022.021583.

Alipio, M., Bures, M., 2023. Deep reinforcement learning perspectives on improving reliable transmissions in IoT networks: problem formulation, parameter choices, challenges, and future directions. Internet of Things 23, 100846. https://doi.org/10.1016/j.iot.2023.100846.

Almeida, S., Radeta, M., Kataoka, T., Canning-Clode, J., Pessanha Pais, M., Freitas, R., Monteiro, J.G., 2022. Designing unmanned aerial survey monitoring program to assess floating litter contamination. Remote Sens. (Basel) 15 (1), 84. https://doi.org/10.3390/rs15010084.

Amjoud, A.B., Amrouch, M., 2023. Object detection using deep learning, CNNs and vision transformers: a review. IEEE Access 11, 35479–35516. https://doi.org/10.1109/ACCESS.2023.3266093.

Andrady, A.L., 2011. Microplastics in the marine environment. Mar. Pollut. Bull. 62 (8), 1596–1605. https://doi.org/10.1016/j.marpolbul.2011.05.030.

Andriolo, U., Garcia-Garin, O., Vighi, M., Borrell, A., Gonçalves, G., 2022. Beached and floating litter surveys by unmanned aerial vehicles: operational analogies and differences. Remote Sens. (Basel) 14 (6), 1336. https://doi.org/10.3390/rs14061336.

Angelini, Z., Kinner, N., Thibault, J., Ramsey, P., Fuld, K., 2019. Marine debris visual identification assessment. Mar. Pollut. Bull. 142, 69–75. https://doi.org/10.1016/j.marpolbul.2019.02.044.

Armitage, S., Awty-Carroll, K., Clewley, D., Martinez-Vicente, V., 2022. Detection and classification of floating plastic litter using a vessel-mounted video camera and deep learning. Remote Sens. (Basel) 14 (14), 3425. https://doi.org/10.3390/rs14143425.

Barnes, D.K.A., Galgani, F., Thompson, R.C., Barlaz, M., 2009. Accumulation and fragmentation of plastic debris in global environments. Philos. Trans. R. Soc., B 364 (1526), 1985–1998. https://doi.org/10.1098/rstb.2008.0205.

Belgiu, M., Marshall, M., Boschetti, M., Pepe, M., Stein, A., Nelson, A., 2023. PRISMA and Sentinel-2 spectral response to the nutrient composition of grains. Remote Sens. Environ. 292, 113567 https://doi.org/10.1016/j.rse.2023.113567.

Bolya, D., Zhou, C., Xiao, F., Lee, Y.J., 2019. YOLACT++: Better Real-Time Instance Segmentation. https://doi.org/10.1109/tpami.2020.3014297.

Buchanan, B.G., 1982. NEW RESEARCH ON EXPERT SYSTEMS. Machine Intelligence 10, 269–299.

Burr, J., Sarishvili, A., Just, D., Katsaouni, N., Moser, K., 2023. Plastic extrusion process optimization by inversion of stacked autoencoder classification machines. Chem. Ing. Tech. 95 (10), 1555–1562. https://doi.org/10.1002/cite.202200211.

Chakma, A., Alam, S.S., Rahman, M.H., Jang, Y.M., 2023. Deep decoder CsiNet for FDD massive MIMO system. IEEE Wireless Communications Letters 12 (12), 2073–2077. https://doi.org/10.1109/LWC.2023.3307164.

Clarivate Analytics. (n.d.). Web of Science. Https://Www.Webofscience.Com.

Claro, F., Fossi, M.C., Ioakeimidis, C., Baini, M., Lusher, A.L., Mc Fee, W., McIntosh, R.R., Pelamatti, T., Sorce, M., Galgani, F., Hardesty, B.D., 2019. Tools and constraints in monitoring interactions between marine litter and megafauna: insights from case studies around the world. Mar. Pollut. Bull. 141, 147–160. https://doi.org/10.1016/j.marpolbul.2019.01.018.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20 (3), 273–297. https://doi.org/10.1023/A:1022627411411.

Cortesi, I., Masiero, A., De Giglio, M., Tucci, G., Dubbini, M., 2021. RANDOM FOREST-BASED RIVER PLASTIC DETECTION WITH A HANDHELD MULTISPECTRAL CAMERA. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLIII-B1-2021, 9–14. https://doi.org/10.5194/isprs-archives-XLIII-B1-2021-9-2021.

Cortesi, I., Masiero, A., Tucci, G., Topouzelis, K., 2022. UAV-based river plastic detection with a multispectral camera. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives 43 (B3–2022), 855–861. https://doi.org/10.5194/isprs-archives-XLIII-B3-2022-855-2022.

Dai, J., Li, Y., He, K., Sun, J., 2016. R-FCN: Object detection via region-based fully convolutional networks. In: Lee, D., Sugiyama, M., Luxburg, U., Guyon, I., Garnett, R. (Eds.), Advances in Neural Information Processing Systems, vol. 29. Curran Associates, Inc.

Daniel, G., Aleksandr, S., 2023. Comparing Yolov5 and Retinanet object detection models for highway trash detection: a computer vision approach to mitigating environmental impact and promoting community health and safety. Comput. Sci. Inf. Technol. 13 (11), 57–67.

Deidun, A., Gauci, A., Lagorio, S., Galgani, F., 2018. Optimising beached litter monitoring protocols through aerial imagery. Mar. Pollut. Bull. 131, 212–217. https://doi.org/10.1016/j.marpolbul.2018.04.033.

Deville, A., Vazquez-Rowe, I., Ita-Nagy, D., Kahhat, R., 2023. Ocean-based sources of plastic pollution: an overview of the main marine activities in the Peruvian EEZ. Mar. Pollut. Bull. 189, 114785 https://doi.org/10.1016/j.marpolbul.2023.114785.

van Dyk, D.A., Meng, X.-L., 2001. The art of data augmentation. J. Comput. Graph. Stat. 10 (1), 1–50. https://doi.org/10.1198/10618600152418584.

El Zaar, A., Aoulalay, A., Benaya, N., El Mhouti, A., Massar, M., El Allati, A., 2022. A deep learning approach to manage and reduce plastic waste in the oceans. E3S Web of Conferences 336, 00065. https://doi.org/10.1051/e3sconf/202233600065.

Elsevier. (n.d.). Scopus. Https://Www.Scopus.Com.

Eßer, J., Bach, N., Jestel, C., Urbann, O., Kerner, S., 2023. Guided reinforcement learning: a review and evaluation for efficient and effective real-world robotics [survey]. IEEE Robotics & Automation Magazine 30 (2), 67–85.

European Space Agency, 2020. Copernicus open access hub. https://search.asf.alaska.edu/.

Flach, P., 2012. Machine learning: the art and science of algorithms that make sense of data. Cambridge university press. https://doi.org/10.1017/CBO9780511973000.

Fletcher, Tristan, 2009. Support Vector Machines Explained.

Franceschini, S., Mattei, F., D'Andrea, L., Di Nardi, A., Fiorentino, F., Garofalo, G., Scardi, M., Cataudella, S., Russo, T., 2019. Rummaging through the bin: modelling marine litter distribution using artificial neural networks. Mar. Pollut. Bull. 149, 110580 https://doi.org/10.1016/j.marpolbul.2019.110580.

Freitas, S., Silva, H., Silva, E., 2021a. Remote hyperspectral imaging acquisition and characterization for marine litter detection. Remote Sens. (Basel) 13 (13), 2536. https://doi.org/10.3390/rs13132536.

Freitas, S., Silva, H., Almeida, C., Viegas, D., Amaral, A., Santos, T., Dias, A., Jorge, P.A. S., Pham, C.K., Moutinho, J., Silva, E., 2021b. Hyperspectral imaging system for marine litter detection. In: OCEANS 2021: San Diego – Porto, pp. 1–6. https://doi.org/10.23919/OCEANS44145.2021.9705953.

Freitas, S., Silva, H., Silva, E., 2022. Hyperspectral imaging zero-shot learning for remote marine litter detection and classification. Remote Sens. (Basel) 14 (21), 5516. https://doi.org/10.3390/rs14215516.

Fu, C.-Y., Liu, W., Ranga, A., Tyagi, A., Berg, A.C., 2017. DSSD: Deconvolutional Single Shot Detector.

Fulton, M., Hong, J., Islam, M.J., Sattar, J., 2019. Robotic detection of marine litter using deep visual detection models. 2019 International Conference on Robotics and Automation (ICRA) 5752–5758. https://doi.org/10.1109/ICRA.2019.8793975.

Garello, R., Plag, H.-P., Shapiro, A., Martinez, S., Pearlman, J., Pendleton, L., 2019. Technologies for Observing and Monitoring Plastics in the oceans. In: OCEANS 2019 - Marseille, pp. 1–6. https://doi.org/10.1109/OCEANSE.2019.8867401.

Geraeds, M., van Emmerik, T., de Vries, R., bin Ab Razak, M.S., 2019. Riverine plastic litter monitoring using unmanned aerial vehicles (UAVs). Remote Sens. (Basel) 11 (17), 2045. https://doi.org/10.3390/rs11172045.

Geron, A., 2019. Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd ed. O'Reilly Media, Inc.

Geyer, R., Jambeck, J.R., Law, K.L., 2017. Production, use, and fate of all plastics ever made. Science Advances 3 (7). https://doi.org/10.1126/sciadv.1700782.

Girshick, R., 2015. Fast R-CNN.

Girshick, R., Donahue, J., Darrell, T., Malik, J., 2013. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation.

Glorot, X., Bengio, Y., 2010. Understanding the difficulty of training deep feedforward neural networks. J. Mach. Learn. Res. 9, 249–256.

Gnann, N., Baschek, B., Ternes, T.A., 2022. Close-range remote sensing-based detection and identification of macroplastics on water assisted by artificial intelligence: a review. Water Res. 222, 118902 https://doi.org/10.1016/j.watres.2022.118902.

Gonçalves, G., Andriolo, U., Pinto, L., Bessa, F., 2020a. Mapping marine litter using UAS on a beach-dune system: a multidisciplinary approach. Sci. Total Environ. 706, 135742 https://doi.org/10.1016/j.scitotenv.2019.135742.

Gonçalves, G., Andriolo, U., Pinto, L., Duarte, D., 2020b. Mapping marine litter with unmanned aerial systems: a showcase comparison among manual image screening and machine learning techniques. Mar. Pollut. Bull. 155, 111158 https://doi.org/10.1016/j.marpolbul.2020.111158.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press.

Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2014. Generative Adversarial Networks.

Gospodinov, M., 1991. Artificial intelligence and quantitative assessment of posture and human brain asymmetries. Proceedings of the Annual Conference on Engineering in Medicine and Biology 13 (pt 2), 867–868.

Goswami, P.K., Goswami, G., 2022. A comprehensive review on real time object detection using deep Learing model. In: 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART), pp. 1499–1502. https://doi.org/10.1109/SMART55829.2022.10046972.

Hartmann, N.B., Hüffer, T., Thompson, R.C., Hassellöv, M., Verschoor, A., Daugaard, A. E., Rist, S., Karlsson, T., Brennholt, N., Cole, M., Herrling, M.P., Hess, M.C., Ivleva, N. P., Lusher, A.L., Wagner, M., 2019. Are we speaking the same language? Recommendations for a definition and categorization framework for plastic debris. Environ. Sci. Technol. 53 (3), 1039–1047. https://doi.org/10.1021/acs.est.8b05297.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 770–778. https://doi.org/10.1109/CVPR.2016.90.

He, K., Gkioxari, G., Dollár, P., Girshick, R., 2017. Mask R-CNN.

Hernández-González, J., Inza, I., Granado, I., Basurko, O.C., Fernandes, J.A., Lozano, J. A., 2019. Aggregated outputs by linear models: an application on marine litter beaching prediction. Inform. Sci. 481, 381–393. https://doi.org/10.1016/j.ins.2018.12.083.

Hidaka, M., Matsuoka, D., Sugiyama, D., Murakami, K., Kako, S., 2022. Pixel-level image classification for detecting beach litter using a deep learning approach. Mar. Pollut. Bull. 175, 113371 https://doi.org/10.1016/j.marpolbul.2022.113371.

Himanshu, S.K., Pandey, A., Yadav, B., 2017. Ensemble wavelet-support vector machine approach for prediction of suspended sediment load using Hydrometeorological data. J. Hydrol. Eng. 22 (7) https://doi.org/10.1061/(ASCE)HE.1943-5584.0001516.

Hinton, G.E., Osindero, S., Teh, Y.-W., 2006. A fast learning algorithm for deep belief nets. Neural Comput. 18 (7), 1527–1554. https://doi.org/10.1162/neco.2006.18.7.1527.

Hoffman, M.J., Hittinger, E., 2017. Inventory and transport of plastic debris in the Laurentian Great Lakes. Mar. Pollut. Bull. 115 (1–2), 273–281. https://doi.org/10.1016/j.marpolbul.2016.11.061.

Hu, H., Wang, D., Huang, S., 2019. International collaboration in the field of artificial intelligence: Global trends and networks at the country level. In: 17th International Conference on Scientometrics and Informetrics, ISSI 2019 - Proceedings, 2, pp. 2501–2502.

Huang, G., Liu, Z., van der Maaten, L., Weinberger, K.Q., 2016. Densely connected convolutional networks. http://arxiv.org/abs/1608.06993.

Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2261–2269. https://doi.org/10.1109/CVPR.2017.243.

Ishii, S., Yoshida, W., 2006. Part 4: reinforcement learning: machine learning and natural learning. N. Gener. Comput. 24 (3), 325–350. https://doi.org/10.1007/BF03037338.

Ita-Nagy, D., Vázquez-Rowe, I., Kahhat, R., 2022. Developing a methodology to quantify mismanaged plastic waste entering the ocean in coastal countries. J. Ind. Ecol. 26 (6), 2108–2122. https://doi.org/10.1111/jiec.13349.

Jahan, M.S., Oussalah, M., 2023. A systematic review of hate speech automatic detection using natural language processing. Neurocomputing 546, 126232. https://doi.org/10.1016/j.neucom.2023.126232.

Jamali, A., Mahdianpari, M., 2021. A cloud-based framework for large-scale monitoring of ocean plastics using multi-spectral satellite imagery and generative adversarial network. Water 13 (18). https://doi.org/10.3390/w13182553.

Jambeck, J.R., Geyer, R., Wilcox, C., Siegler, T.R., Perryman, M., Andrady, A., Narayan, R., Law, K.L., 2015. Plastic waste inputs from land into the ocean. Science 347 (6223), 768–771. https://doi.org/10.1126/science.1260352.

Jindal, M., Kumar, M., 2020. Effect of marine pollution: review. Journal of Critical Reviews 7 (10), 938–941. https://doi.org/10.31838/jcr.07.10.184.

Jocher, G., Stoken, A., Borovec, J., NanoCode012, ChristopherSTAN, Changyu, L., Laughing, Tkianai, Hogan, A., Lorenzomammana, yxNONG, AlexWang1900, Diaconu, L., Marc, wanghaoyang0106, ml5ah, Doug, Ingham, F., Frederik, Rai, P., 2020. ultralytics/yolov5: v3.1 - Bug Fixes and Performance Improvements. Zenodo. https://doi.org/10.5281/zenodo.4154370.

Jordan, M.I., Mitchell, T.M., 2015. Machine learning: trends, perspectives, and prospects. Science 349 (6245), 255–260. https://doi.org/10.1126/science.aaa8415.

Just, J., 2024. Natural language processing for innovation search – reviewing an emerging non-human innovation intermediary. Technovation 129, 102883. https://doi.org/10.1016/j.technovation.2023.102883.

Kaandorp, M.L.A., Ypma, S.L., Boonstra, M., Dijkstra, H.A., van Sebille, E., 2022. Using machine learning and beach cleanup data to explain litter quantities along the Dutch North Sea coast. Ocean Sci. 18 (1), 269–293. https://doi.org/10.5194/os-18-269-2022.

Kako, S., Morita, S., Taneda, T., 2020. Estimation of plastic marine debris volumes on beaches using unmanned aerial vehicles and image processing based on deep learning. Mar. Pollut. Bull. 155, 111127 https://doi.org/10.1016/j.marpolbul.2020.111127.

Kamath, V., Renuka, A., 2023. Deep learning based object detection for resource constrained devices: systematic review, future trends and challenges ahead. Neurocomputing 531, 34–60. https://doi.org/10.1016/j.neucom.2023.02.006.

Kantanantha, N., Pattaraumpornchai, P., 2023. Machine Learning-Based Price Forecasting for Polypropylene Granules in Thailand. In: Proceedings of the 2023 6th International Conference on Computational Intelligence and Intelligent Systems, pp. 14–19. https://doi.org/10.1145/3638209.3638212.

Karmaoui, A., Yoganandan, G., Sereno, D., Shaukat, K., El Jaafari, S., Hajji, L., 2023. Global network analysis of links between business, climate change, and sustainability and setting up the interconnections framework. Environ. Dev. Sustain. https://doi.org/10.1007/s10668-023-03883-w.

Kiruba-Sankar, R., Saravanan, K., Adamala, S., Selvam, K., Kumar, K.L., Praveenraj, J., 2023. First report of marine debris in Car Nicobar, a remote oceanic Island in the Nicobar archipelago, Bay of Bengal. Reg. Stud. Mar. Sci. 61, 102845 https://doi.org/10.1016/j.rsma.2023.102845.

Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. Biol. Cybern. 43 (1), 59–69. https://doi.org/10.1007/BF00337288.

Kremezi, M., Kristollari, V., Karathanassi, V., Topouzelis, K., Kolokoussis, P., Taggio, N., Aiello, A., Ceriola, G., Barbone, E., Corradi, P., 2021. Pansharpening PRISMA data for marine plastic litter detection using plastic indexes. IEEE Access 9, 61955–61971. https://doi.org/10.1109/ACCESS.2021.3073903.

Kremezi, M., Kristollari, V., Karathanassi, V., Topouzelis, K., Kolokoussis, P., Taggio, N., Aiello, A., Ceriola, G., Barbone, E., Corradi, P., 2022. Increasing the Sentinel-2 potential for marine plastic litter monitoring through image fusion techniques. Mar. Pollut. Bull. 182, 113974 https://doi.org/10.1016/j.marpolbul.2022.113974.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2017. ImageNet classification with deep convolutional neural networks. Commun. ACM 60 (6), 84–90. https://doi.org/10.1145/3065386.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks. http://code.google.com/p/cuda-convnet/.

Kumar, C.H., Kumar, R.S., 2022. Natural language processing of movie reviews to detect the sentiments using novel bidirectional encoder representation-BERT for transformers over support vector machine. Journal of Pharmaceutical Negative Results 13, 619–628. https://doi.org/10.47750/pnr.2022.13.S04.069.

Kylili, K., Artusi, A., Kyriakides, I., & Hadjistassou, C. (2018). Tracking and identifying floating marine debris . 6th International Marine Debris Conference.

Kylili, K., Hadjistassou, C., Artusi, A., 2020. An intelligent way for discerning plastics at the shorelines and the seas. Environ. Sci. Pollut. Res. 27 (34), 42631–42643. https://doi.org/10.1007/s11356-020-10105-7.

Kylili, K., Kyriakides, I., Artusi, A., Hadjistassou, C., 2019. Identifying floating plastic marine debris using a deep learning approach. Environ. Sci. Pollut. Res. 26 (17), 17091–17099. https://doi.org/10.1007/s11356-019-05148-4.

Kylili, K., Artusi, A., Hadjistassou, C., 2021. A new paradigm for estimating the prevalence of plastic litter in the marine environment. Mar. Pollut. Bull. 173, 113127 https://doi.org/10.1016/j.marpolbul.2021.113127.

Lavender, S., 2022. Detection of waste plastics in the environment: application of Copernicus Earth observation data. Remote Sens. (Basel) 14 (19), 4772. https://doi.org/10.3390/rs14194772.

Lebreton, L.C.M., van der Zwet, J., Damsteeg, J.-W., Slat, B., Andrady, A., Reisser, J., 2017. River plastic emissions to the world's oceans. Nat. Commun. 8 (1), 15611. https://doi.org/10.1038/ncomms15611.

LeCun, Y., Bottou, L., Orr, G.B., Müller, K.-R., 1998. Efficient BackProp, pp. 9–50. https://doi.org/10.1007/3-540-49430-8_2.

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444. https://doi.org/10.1038/nature14539.

Lieshout, C., Oeveren, K., Emmerik, T., Postma, E., 2020. Automated river plastic monitoring using deep learning and cameras. Earth and Space Science 7 (8). https://doi.org/10.1029/2019EA000960.

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L., 2014. Microsoft COCO: Common Objects in Context, pp. 740–755. https://doi.org/10.1007/978-3-319-10602-1_48.

Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollár, P., 2017. Focal Loss for Dense Object Detection.

Lindsay, M.D., Piechocka, A.M., Jessell, M.W., Scalzo, R., Giraud, J., Pirot, G., Cripps, E., 2022. Assessing the impact of conceptual mineral systems uncertainty on prospectivity predictions. Geosci. Front. 13 (6), 101435 https://doi.org/10.1016/j.gsf.2022.101435.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., Berg, A.C., 2015. SSD: Single Shot MultiBox Detector. https://doi.org/10.1007/978-3-319-46448-0_2.

Loizzo, R., Guarini, R., Longo, F., Scopa, T., Formaro, R., Facchinetti, C., Varacalli, G., 2018. Prisma: The Italian Hyperspectral Mission. IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium, 175–178. https://doi.org/10.1109/IGARSS.2018.8518512.

Lu, X., Yi, Q., Tian, G.Y., 2020. A comparison of feature extraction techniques for delamination of CFRP using Eddy current pulse-compression thermography. IEEE Sensors Journal 20 (20), 12415–12422. https://doi.org/10.1109/JSEN.2020.2993154.

Luo, W., Han, W., Fu, P., Wang, H., Zhao, Y., Liu, K., Liu, Y., Zhao, Z., Zhu, M., Xu, R., Wei, G., 2022. A water surface contaminants monitoring method based on airborne depth reasoning. Processes 10 (1), 131. https://doi.org/10.3390/pr10010131.

Maharjan, N., Miyazaki, H., Pati, B.M., Dailey, M.N., Shrestha, S., Nakamura, T., 2022. Detection of river plastic using UAV sensor data and deep learning. Remote Sens. (Basel) 14 (13), 3049. https://doi.org/10.3390/rs14133049.

Mahdianpari, M., Salehi, B., Rezaee, M., Mohammadimanesh, F., Zhang, Y., 2018. Very deep convolutional neural networks for complex land cover mapping using multispectral remote sensing imagery. Remote Sens. (Basel) 10 (7), 1119. https://doi.org/10.3390/rs10071119.

Mai, L., Sun, X., Zeng, E.Y., 2023. Country-specific riverine contributions to marine plastic pollution. Sci. Total Environ. 874, 162552 https://doi.org/10.1016/j.scitotenv.2023.162552.

Maliks, R., Kadikis, R., 2021. Multispectral data classification with deep CNN for plastic bottle sorting. In: 2021 6th International Conference on Mechanical Engineering and Robotics Research (ICMERR), pp. 58–65. https://doi.org/10.1109/ICMERR54363.2021.9680850.

Margallo, M., Ziegler-Rodriguez, K., Vázquez-Rowe, I., Aldaco, R., Irabien, Á., Kahhat, R., 2019. Enhancing waste management strategies in Latin America under a holistic environmental assessment perspective: a review for policy support. Sci. Total Environ. 689, 1255–1275. https://doi.org/10.1016/j.scitotenv.2019.06.393.

Marin, I., Mladenović, S., Gotovac, S., Zaharija, G., 2021. Deep-feature-based approach to marine debris classification. Appl. Sci. 11 (12), 5644. https://doi.org/10.3390/app11125644.

Martin, C., Parkes, S., Zhang, Q., Zhang, X., McCabe, M.F., Duarte, C.M., 2018. Use of unmanned aerial vehicles for efficient beach litter monitoring. Mar. Pollut. Bull. 131, 662–673. https://doi.org/10.1016/j.marpolbul.2018.04.045.

Marugán, A.P., 2023. Applications of reinforcement learning for maintenance of engineering systems: a review. Adv. Eng. Softw. 183, 103487 https://doi.org/10.1016/j.advengsoft.2023.103487.

Masiero, A., Cortesi, I., Tucci, G., 2022. DEVELOPMENT OF A GRAPHICAL USER INTERFACE TO SUPPORT THE SEMI-AUTOMATIC SEMANTIC SEGMENTATION OF UAS-IMAGES. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 48(4/W1-2022), 295–300. https://doi.org/10.5194/isprs-archives-XLVIII-4-W1-2022-295-2022.

McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. 5 (4), 115–133. https://doi.org/10.1007/BF02478259.

Mehrish, A., Majumder, N., Bharadwaj, R., Mihalcea, R., Poria, S., 2023. A review of deep learning techniques for speech processing. Information Fusion 99, 101869. https://doi.org/10.1016/j.inffus.2023.101869.

Meijer, L.J.J., van Emmerik, T., van der Ent, R., Schmidt, C., Lebreton, L., 2021. More than 1000 rivers account for 80% of global riverine plastic emissions into the ocean. Sci. Adv. 7, 1–14.

Moorton, Z., Kurt, Z., Woo, W.L., 2022. Is the use of deep learning an appropriate means to locate debris in the ocean without harming aquatic wildlife? Mar. Pollut. Bull. 181, 113853 https://doi.org/10.1016/j.marpolbul.2022.113853.

Moy, K., Neilson, B., Chung, A., Meadows, A., Castrence, M., Ambagis, S., Davidson, K., 2018. Mapping coastal marine debris using aerial imagery and spatial analysis. Mar. Pollut. Bull. 132, 52–59. https://doi.org/10.1016/j.marpolbul.2017.11.045.

Murphy, K.P., 2013. Machine learning: a probabilistic perspective. MIT Press. https://www.amazon.com/Machine-Learning-Probabilistic-Perspective-Computation/dp/0262018020/ref=sr_1_2?ie=UTF8&qid=1336857747&sr=8-2.

Nanehkaran, Y.A., Licai, Z., Azarafza, M., Talaei, S., Jinxia, X., Chen, J., Derakhshani, R., 2023. The predictive model for COVID-19 pandemic plastic pollution by using deep learning method. Sci. Rep. 13 (1), 4126. https://doi.org/10.1038/s41598-023-31416-y.

NASA, 2013. Earthdata Worldview. https://worldview.earthdata.nasa.gov/.

Newbould, R.A., Powell, D.M., Whelan, M.J., 2021. Macroplastic debris transfer in rivers: a travel distance approach. Frontiers in Water 3. https://doi.org/10.3389/frwa.2021.724596.

Pal, S.K., Pramanik, A., Maiti, J., Mitra, P., 2021. Deep learning in multi-object detection and tracking: state of the art. Appl. Intell. 51 (9), 6400–6429. https://doi.org/10.1007/s10489-021-02293-7.

Papakonstantinou, A., Batsaris, M., Spondylidis, S., Topouzelis, K., 2021. A citizen science unmanned aerial system data acquisition protocol and deep learning techniques for the automatic detection and mapping of marine litter concentrations in the coastal zone. Drones 5 (1), 6. https://doi.org/10.3390/drones5010006.

Pinto, L., Andriolo, U., Gonçalves, G., 2021. Detecting stranded macro-litter categories on drone orthophoto by a multi-class neural network. Mar. Pollut. Bull. 169, 112594 https://doi.org/10.1016/j.marpolbul.2021.112594.

Pizer, S.M., Amburn, E.P., Austin, J.D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar Romeny, B., Zimmerman, J.B., Zuiderveld, K., 1987. Adaptive histogram equalization and its variations. Computer Vision, Graphics, and Image Processing 39 (3), 355–368. https://doi.org/10.1016/S0734-189X(87)80186-X.

Quetglas, A., Ordines, F., Guijarro, B., 2011. The Use of Artificial Neural Networks (ANNs) in Aquatic Ecology. In Artificial Neural Networks - Application. InTech. https://doi.org/10.5772/16092.

Rani, B., Kant, S., 2020. Semi-supervised learning approach to improve machine learning algorithms for churn analysis in telecommunication. International Journal of Computer Information Systems and Industrial Management Applications 12, 265–275. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85099162411&partnerID=40&md5=475bc9b35508b3b9c88e28d167b53f8f.

Reddy, Simon, Lau, Winnie, 2020. Preventing Ocean Plastics. Pew Charitable Trusts. August.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: unified, real-time object detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 779–788. https://doi.org/10.1109/CVPR.2016.91.

Reis, D., Kupec, J., Hong, J., Daoudi, A., 2023. Real-Time Flying Object Detection with YOLOv8.

Ren, S., He, K., Girshick, R., Sun, J., 2015. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.

Rodrigues, A.C.B., de Jesus, G.P., Waked, D., Gomes, G.L., Silva, T.M., Yariwake, V.Y., da Silva, M.P., Magaldi, A.J., Veras, M.M., 2022. Scientific evidence about the risks of micro and nanoplastics (MNPLs) to human health and their exposure routes through the environment. Toxics 10 (6), 308. https://doi.org/10.3390/toxics10060308.

Roh, S.-B., Oh, S.-K., Choi, W.Z., 2017. Design of identification system based on Gaussian mixture model for black plastic wastes. 14th international symposium on east Asian resources recycling technology, EARTH 2017. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85065994972&partnerID=40&md5=cf994d74d7daed2244ab3baa406ee3b3.

Roman, L., Schuyler, Q., Wilcox, C., Hardesty, B.D., 2021. Plastic pollution is killing marine megafauna, but how do we prioritize policies to reduce mortality? Conserv. Lett. 14 (2) https://doi.org/10.1111/conl.12781.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation.

Rosenblatt, F., 1958. The perceptron: a probabilistic model for information storage and organization in the brain. Psychol. Rev. 65 (6), 386–408. https://doi.org/10.1037/h0042519.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C., 2018. MobileNetV2: Inverted Residuals and Linear Bottlenecks.

Sannigrahi, S., Basu, B., Basu, A.S., Pilla, F., 2022. Development of automated marine floating plastic detection system using Sentinel-2 imagery and machine learning models. Mar. Pollut. Bull. 178, 113527 https://doi.org/10.1016/j.marpolbul.2022.113527.

Sekrecka, A., Wierzbicki, D., Kedzierski, M., 2020. Influence of the Sun position and platform orientation on the quality of imagery obtained from unmanned aerial vehicles. Remote Sens. (Basel) 12 (6), 1040. https://doi.org/10.3390/rs12061040.

Semernya, L., Ramola, A., Alfthan, B., Giacovelli, C., 2017. Waste management outlook for mountain regions: sources and solutions. Waste Management & Research: The Journal for a Sustainable Circular Economy 35 (9), 935–939. https://doi.org/10.1177/0734242X17709910.

Shaikh, M.S., Jaferzadeh, K., Thörnberg, B., 2022. Extending effective dynamic range of hyperspectral line cameras for short wave infrared imaging. Sensors 22 (5), 1817. https://doi.org/10.3390/s22051817.

Shaukat, K., Luo, S., Chen, S., Liu, D., 2020a. Cyber threat detection using machine learning techniques: a performance evaluation perspective. 2020 International Conference on Cyber Warfare and Security (ICCWS) 1–6. https://doi.org/10.1109/ICCWS48432.2020.9292388.

Shaukat, K., Luo, S., Varadharajan, V., Hameed, I.A., Xu, M., 2020b. A survey on machine learning techniques for cyber security in the last decade. IEEE Access 8, 222310–222354. https://doi.org/10.1109/ACCESS.2020.3041951.

Shaukat, K., Luo, S., Varadharajan, V., Hameed, I., Chen, S., Liu, D., Li, J., 2020c. Performance comparison and current challenges of using machine learning techniques in cybersecurity. Energies 13 (10), 2509. https://doi.org/10.3390/en13102509Simonyan, K., Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition.

Shaukat, K., Luo, S., Varadharajan, V., 2022. A novel method for improving the robustness of deep learning-based malware detectors against adversarial attacks. Eng. Appl. Artif. Intel. 116, 105461 https://doi.org/10.1016/j.engappai.2022.105461.

Shaukat, K., Luo, S., Varadharajan, V., 2023. A novel deep learning-based approach for malware detection. Eng. Appl. Artif. Intel. 122, 106030 https://doi.org/10.1016/j.engappai.2023.106030.

Shaukat, K., Luo, S., Varadharajan, V., 2024. A novel machine learning approach for detecting first-time-appeared malware. Eng. Appl. Artif. Intel. 131, 107801 https://doi.org/10.1016/j.engappai.2023.107801.

Simonyan, K., Zisserman, A., 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition.

Singh, D., Valdenegro-Toro, M., 2021. The Marine Debris Dataset for Forward-Looking Sonar Semantic Segmentation.

So, M.W.K., Vorsatz, L.D., Cannicci, S., Not, C., 2022. Fate of plastic in the environment: from macro to nano by macrofauna. Environ. Pollut. 300, 118920 https://doi.org/10.1016/j.envpol.2022.118920.

Song, K., Jung, J.-Y., Lee, S.H., Park, S., 2021. A comparative study of deep learning-based network model and conventional method to assess beach debris standing-stock. Mar. Pollut. Bull. 168, 112466 https://doi.org/10.1016/j.marpolbul.2021.112466.

Song, K., Jung, J.-Y., Lee, S.H., Park, S., Yang, Y., 2022. Assessment of marine debris on hard-to-reach places using unmanned aerial vehicles and segmentation models based on a deep learning approach. Sustainability 14 (14), 8311. https://doi.org/10.3390/su14148311.

Srivastava, N., Hinton, G.E., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. J. Mach. Learn. Res. 15, 1929–1958. https://api.semanticscholar.org/CorpusID:6844431.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z., 2016a. Rethinking the inception architecture for computer vision. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2818–2826. https://doi.org/10.1109/CVPR.2016.308.

Szegedy, C., Ioffe, S., Vanhoucke, V., Alemi, A., 2016b. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning.

Taggio, N., Aiello, A., Ceriola, G., Kremezi, M., Kristollari, V., Kolokoussis, P., Karathanassi, V., Barbone, E., 2022. A combination of machine learning algorithms for marine plastic litter detection exploiting hyperspectral PRISMA data. Remote Sens. (Basel) 14 (15), 3606. https://doi.org/10.3390/rs14153606.

Takaya, K., Shibata, A., Mizuno, Y., Ise, T., 2022. Unmanned aerial vehicles and deep learning for assessment of anthropogenic marine debris on beaches on an island in a semi-enclosed sea in Japan. Environmental Research Communications 4 (1), 015003. https://doi.org/10.1088/2515-7620/ac473b.

Tamin, O., Moung, E.G., Dargham, J.A., Yahya, F., Farzamnia, A., Sia, F., Naim, N.F.M., Angeline, L., 2023. On-shore plastic waste detection with YOLOv5 and RGB-near-infrared fusion: a state-of-the-art solution for accurate and efficient environmental monitoring. Big Data and Cognitive Computing 7 (2), 103. https://doi.org/10.3390/bdcc7020103.

Tang, L., 2023. A transformer-based network for speech recognition. International Journal of Speech Technology. https://doi.org/10.1007/s10772-023-10034-z.

Tang, Z., Li, T., Wu, D., Liu, J., Yang, Z., 2024. A systematic literature review of reinforcement learning-based knowledge graph research. Expert Syst. Appl. 238, 121880 https://doi.org/10.1016/j.eswa.2023.121880.

Teng, C., Kylili, K., Hadjistassou, C., 2022. Deploying deep learning to estimate the abundance of marine debris from video footage. Mar. Pollut. Bull. 183, 114049 https://doi.org/10.1016/j.marpolbul.2022.114049.

Thanh Nguyen, T., Tung Luu, T., Thanh An Tong, P., 2024. Fine-tuning DETR: toward holistic process in plastic waste sorting system. Waste Manag. 179, 154–162. https://doi.org/10.1016/j.wasman.2024.03.015.

Topouzelis, K., Papageorgiou, D., Suaria, G., Aliani, S., 2021. Floating marine litter detection algorithms and techniques using optical remote sensing data: a review. Mar. Pollut. Bull. 170, 112675 https://doi.org/10.1016/j.marpolbul.2021.112675.

Trevisan, R., Ranasinghe, P., Jayasundara, N., Di Giulio, R., 2022. Nanoplastics in aquatic environments: impacts on aquatic species and interactions with environmental factors and pollutants. Toxics 10 (6), 326. https://doi.org/10.3390/toxics10060326.

Tunukovic, V., McKnight, S., Pyle, R., Wang, Z., Mohseni, E., Gareth Pierce, S., Vithanage, R., K.W., Dobie, G., MacLeod, C.N., Cochran, S., O'Hare, T., 2024. Unsupervised machine learning for flaw detection in automated ultrasonic testing of carbon fibre reinforced plastic composites. Ultrasonics 140, 107313. https://doi.org/10.1016/j.ultras.2024.107313.

Tur, E., 2023. Applying Deep Learning for Automated Quality Control and Defect Detection in Multi-stage Plastic Extrusion Process. 2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 01–05. https://doi.org/10.1109/HORA58378.2023.10156750.

United Nations Environment Programme, 2021. From Pollution to Solution a Global Assessment of Marine Litter and Plastic Pollution.

Valdenegro-Toro, M., 2019. Learning Objectness from Sonar Images for Class-Independent Object Detection.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2017. Attention Is all you need.

Veerasingam, S., Chatting, M., Asim, F.S., Al-Khayat, J., Vethamony, P., 2022. Detection and assessment of marine litter in an uninhabited island, Arabian gulf: a case study with conventional and machine learning approaches. Sci. Total Environ. 838, 156064 https://doi.org/10.1016/j.scitotenv.2022.156064.

de Vries, R., Egger, M., Mani, T., Lebreton, L., 2021. Quantifying floating plastic debris at sea using vessel-based optical data and artificial intelligence. Remote Sens. (Basel) 13 (17), 3401. https://doi.org/10.3390/rs13173401.

Wang, Y., Zhang, L., Qi, L., Yi, X., Li, M., Zhou, M., Chen, D., Xiao, Q., Wang, C., Pang, Y., Xu, J., Deng, H., Liu, L., Guan, X. (2021). Machine learning: applications and advanced progresses of radiomics in endocrine neoplasms. J. Oncol., 2021, 1–17. doi:https://doi.org/10.1155/2021/8615450Weng, W., & Zhu, X. (2021). INet: Convolutional Networks for Biomedical Image Segmentation. IEEE Access, 9, 16591–16603. doi:https://doi.org/10.1109/ACCESS.2021.3053408.

Wolf, M., van den Berg, K., Garaba, S.P., Gnann, N., Sattler, K., Stahl, F., Zielinski, O., 2020. Machine learning for aquatic plastic litter detection, classification and quantification (APLASTIC-Q). Environ. Res. Lett. 15 (11), 114042 https://doi.org/10.1088/1748-9326/abbd01.

Yokoya, N., Grohnfeldt, C., Chanussot, J., 2017. Hyperspectral and multispectral data fusion: a comparative review of the recent literature. IEEE Geoscience and Remote Sensing Magazine 5 (2), 29–56. https://doi.org/10.1109/MGRS.2016.2637824.

Zhang, C., Liu, C., Zhang, X., Almpanidis, G., 2017a. An up-to-date comparison of state-of-the-art classification algorithms. Expert Systems with Applications 82, 128–150. https://doi.org/10.1016/j.eswa.2017.04.003.

Zhang, S., Wen, L., Bian, X., Lei, Z., Li, S.Z., 2017b. Single-Shot Refinement Neural Network for Object Detection.

Chapter 3

Manuscript two: Advanced Deep Learning strategies for detection and quantification of macroplastics in rivers along the Peruvian Coast

> *“The rise of powerful AI will be either the best or the worst thing ever to happen to humanity. We do not yet know which.”*
>
> —Stephen Hawking

This manuscript is currently under its second round of revision in Marine Pollution Bulletin
Astorayme, M. A., Vázquez-Rowe, I., Muñoz-Sovero, E. & Kahhat, R. (2025). Advanced Deep Learning strategies for detection and quantification of macroplastics in rivers along the Peruvian Coast. DOI: 10.1016/j.marpolbul.2025.118649

In this chapter, we propose a methodology based on a DL architecture for the detection of MP items in natural waterways, using a transfer learning approach. The methodology incorporates the use of UAVs to capture RGB imagery. We analyze the behavior of the most commonly detected MPs and their relationship with the hydrological dynamics of the watercourse, based on data collected over a full hydrological year. Finally, we present the results obtained, discuss the limitations of the methodology, and outline potential improvements and considerations for its replication in other Peruvian coastal rivers.

Marine Pollution Bulletin 222 (2026) 118649

Contents lists available at ScienceDirect

Marine Pollution Bulletin

journal homepage: www.elsevier.com/locate/marpolbul

Advanced deep learning strategies for detection and quantification of macroplastics in rivers along the Peruvian coast

Miguel Angel Astorayme [a,b,*], Ian Vázquez-Rowe [a], Eizo Muñoz-Sovero [a], Ramzy Kahhat [a]

[a] *Peruvian Life Cycle Assessment & Industrial Ecology Network (PELCAN), Department of Engineering, Pontificia Universidad Católica del Perú, Av. Universitaria 1801, San Miguel 15074, Lima, Peru*

[b] *Dept. of Fluid Mechanics Engineering, Universidad Nacional Mayor de San Marcos, Av. Universitaria/Av. Germán Amezaga s/n., Lima 15081, Lima, Peru*

ARTICLE INFO

Keywords:
Artificial intelligence
Industrial ecology
Plastic waste
Transfer learning
Waste management
YOLOv11

ABSTRACT

Rivers are the primary contributors to plastic waste pollution entering the oceans, largely due to inadequate solid waste management, especially in the Global South. Macroplastics become difficult to remove from water bodies, and eventually fragment into smaller polymers, affecting wildlife and human health. However, methods for estimating these flows still face significant limitations. This study develops a methodological framework that incorporates artificial intelligence, particularly Deep Learning, to detect and classify eight classes of mixed inorganic municipal solid waste (MSW), with a focus on macroplastics present in rivers. This approach considers the spatial and temporal dynamics of the watercourse under study by using YOLOv11, a convolutional neural network model, by training and validating images captured by drones. A section of the river Rímac (Lima, Peru) was examined for one year. Results suggest that the YOLOv11 model is suitable for the rapid counting of certain macroplastic classes, such as tires, and black and colored bags. The model showed very high accuracy for tires (mAP = 0.94) in the testing stage, whereas for plastic bags values were above 0.74. Lower precision was identified for other categories, such as furniture and PET bottles due to debris size, abundance or chromatic contrast. Temporal changes in abundance were analyzed, with relevant changes observable between dry and wet seasons. This research validates the potential for establishing fieldwork projects covering larger areas to capture images of MSW mixes in rivers along the Peruvian coast, enabling future development of an automatic monitoring system.

1. Introduction

Ever since the discovery of nylon in 1935, the production and use of plastic polymers has soared linked to the increasing demand in world markets aimed at satisfying human needs (Ramkumar et al., 2022). By 2019, according to Kumar et al. (2021), plastic production had surpassed 370 million metric tons per year. Unfortunately, when plastic is disposed of, its emissions to the environment have also increased at alarming rates in the past decades (Jambeck et al., 2015). This situation has resulted in a silent environmental hazard in which inefficiencies in waste disposal within the technosphere contribute to this problem (Browning et al., 2021). These include mismanaged waste, particularly in the Global South (Margallo et al., 2019), littering or the degradation of plastic polymers due to weathering (e.g., tire abrasion), as well as low recycling rates (Singh and Walker, 2024), which have shaped one of the main environmental concerns in the 21st century.

Plastic waste that enters rivers has shown potential for high mobility due to the relatively low density of most plastic polymers (Andrade-Muñoz et al., 2025). In fact, the dynamic nature of rivers has allowed the transportation of plastic particles of different sizes in a similar way to suspended solids that end up as sediments, with many of these particles either reaching the ocean or floodplains (Roebroek et al., 2022), increasing the scope of pollution on a global scale and altering the aquatic ecosystem (Ryan, 2015). Besides, according to Schreyers et al. (2024), transportation dynamics depend on the characteristics of plastic items. Hence, parameters such as polymer type, size, shape or density can influence the accumulation or release of plastic waste along rivers and other aquatic environments (Shamskhany et al., 2021; Ranjan et al., 2023).

Addressing the transport and effects of riverine plastic debris, while

---

[*] Corresponding author at: Peruvian Life Cycle Assessment & Industrial Ecology Network (PELCAN), Department of Engineering, Pontificia Universidad Católica del Perú, Av. Universitaria 1801, San Miguel 15074, Lima, Peru.

E-mail address: m.astoraymev@pucp.edu.pe (M.A. Astorayme).

https://doi.org/10.1016/j.marpolbul.2025.118649
Received 16 June 2025; Received in revised form 23 August 2025; Accepted 26 August 2025

considering the temporality of waterways, can help reduce uncertainties when estimating plastic emissions to the ocean (Chowdhury et al., 2021). Under this scenario, there is an urgent need to establish methodologies that allow quantifying the presence of plastic debris in aquatic environments. In fact, this necessity is particularly important in countries of the Global South, where mismanaged plastic waste is rampant (Cottom et al., 2024), but monitoring and quantification mechanisms are scarce (Alencar et al., 2023).

Unplanned urban sprawling in cities, especially in the Global South, has often led to vulnerable sectors of the population settling in certain areas in which the quality of services, housing, and critical infrastructure are poor (Chettry, 2023). These areas include, in some cases, the banks and flooding areas of rivers and streams (Dourojeanni and Jouravlev, 1999), where an increasing number of illegal waste disposal sites have been identified in recent years (Honingh et al., 2020). Furthermore, open dumping and population density have been identified as the main factors that most influence waste accumulation risk (Radhakrishnan et al., 2024).

Although in recent decades the scientific community has made efforts to establish methodologies that allow monitoring and quantifying microplastics and macroplastics (MPs) in water bodies, the location and mobility of plastic particles in the environment has proven to be an important limitation to building methods for quantification. González-Fernández and Hanke (2017) proposed a methodology based on visual observations for counting MPs in rivers (JRC, 2016). In contrast, Ita-Nagy et al. (2022) and Alencar et al. (2023) developed regionally oriented methodologies in Peru and Brazil, respectively, focusing on the estimation of plastic flows to the ocean based on population density and socioeconomic conditions, as well as the solid waste ratio at the municipal level. However, these methods require manual processes and holistic approaches that may lead to errors. In this context, a recent review identified that in recent years a new paradigm in the quantification of MPs has begun to be explored, using artificial intelligence (AI), as it has the potential to automate the processes of quantification of MPs.

Machine learning (ML), which is a subfield of AI, has started to be explored as an automatic method for counting MPs in water bodies (Hernández-González et al., 2019; Martin et al., 2018). The exploration of more sophisticated techniques with the aim of improving efficiency in multiclass classification of MPs led to the use of deep learning (DL), a subfield of ML (Wolf et al., 2020), and transfer learning (TL) approaches (Teng et al., 2022). Within the context of DL techniques, You Only Look Once (YOLO) models have emerged as a type of convolutional neural networks (CNNs), which use a single-stage detection algorithms (Lin et al., 2014) and have shown to be appropriate to detect and identify objects efficiently while reducing detection times (Fulton et al., 2019).

In this context, the main goal of the current study is to explore and assess the capacity of using a DL approach to detect, classify, and establish the occurrence of different categories of MPs and other MSW items in a river in an urban environment in Peru. The river Rímac, the main river that flows through the megacity of Lima, Peru, was selected as the main case study. This selection is linked to the existence of open dumping sites in the final part of the river in areas of informal urban sprawl. Although the application of a DL approach is relatively new in the detection of MPs in fluvial systems, there is evidence of its effectiveness in other geographical contexts, i.e., countries with more robust waste management systems (Astorayme et al., 2024). However, as far as the authors were able to ascertain, no studies have been conducted in Peru, where it is hypothesized that the presence of MPs is likely to be denser and more heterogeneous than in the Global North.

The novelty of this research is associated with the use of one of the most updated YOLO models, i.e., v11 (Jocher and Qiu, 2024; Redmon et al., 2016) in the detection of MPs in rivers. It is hypothesized that YOLOv11 has the potential to perform automatic detections in areas of plastic waste abundance, while allowing for reduced detection time due to its *single-pass* architecture, allowing the establishment of monitoring systems in the future (Zailan et al., 2021). We consider that the target audience of this study consists of two main groups. On the one hand, the scientific community involved in the use of ML techniques, especially those linked to evaluating environmental damage, could benefit from the methodological framework proposed. On the other hand, decision makers within municipal and regional governments, mainly related to waste management and urban planning, may encounter valuable findings that allow using these techniques to monitor and quantify plastic waste emissions in fluvial ecosystems.

2. Materials and methods

This section outlines the steps followed in the development of the research, starting with Section 2.1, which reviews the AI-based model used in this study, and then addresses the research methodology summarized in Fig. 1. Section 2.2 describes the case study and its relevance to the research. Section 2.3 covers the fieldwork conducted to collect data from the study area. In Section 2.4, we explain the criteria applied to process the data collected. Section 2.5 discusses the establishment of the MPs classes as the target for classification, along with the generation of labels before inputting them into the YOLOv11 model. Finally, Section 2.6 details the hardware specifications used (Section 2.6.1), the metrics employed to evaluate the model (Section 2.6.2), and the fine-tuning processes applied during model training.

2.1. YOLOv11 model

The YOLOv11 model was launched by Ultralytics in September 2024 and represents one of the most recent versions of the YOLO family, based on the foundations established by its predecessors, as shown in Fig. S1 in the Supplementary Material (Sapkota et al., 2025). The model is divided into 5 subversions based on the complexity of its architecture, which influences training time and computing resource consumption (see

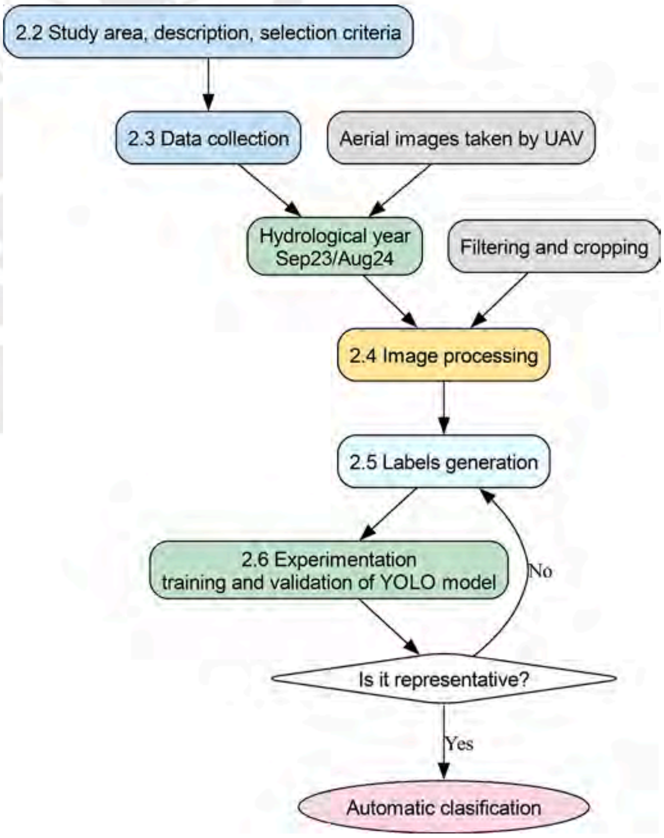


Fig. 1. Methodological workflow for automatically detecting macroplastics (MPs) using the You Only Look Once (YOLO) model.

Table 1
Comparison of YOLOv11 Model Versions and Key Features. Size refers to the scale of the architecture. Accuracy indicates the level of performance that the model is able to achieve. The final column, parameters (M), represents the number of parameters in millions (Jocher and Qiu, 2024).

Model	Size	Accuracy	Parameters (M)
Yolov11-n	Nano	Basic	2.6
Yolov11-s	Small	Good	9.4
Yolov11-m	Medium	High	20.1
Yolov11-l	Large	Very High	25.3
Yolov11-x	Extra-large	Maximum	56.9

Table 1). YOLOv11 redefines the architecture (see Fig. S1) by enhancing feature extraction and reducing the number of parameters, making the model more compact while still maintaining high precision (Jocher and Qiu, 2024). Like its predecessors, YOLOv11 supports both object detection and segmentation, and according to Sapkota et al. (2025) the model is approximately 2 % faster than YOLOv10 in the inference stage, making it more suitable for real-time applications (He et al., 2024).

The architecture of YOLOv11 is divided into three parts, as shown in Fig. S1: the backbone, the neck and the head. To enhance the efficiency of YOLOv11 feature extraction, the model introduces the Cross Stage Partial with kernel size 2 (C3K2 block) into its backbone, allowing better information integration between layers (Sapkota et al., 2024a). The neck fuses the features extracted at different scales by the backbone. On the one hand, it incorporates the Spatial Pyramid Pooling – Fast (SPPF) module, which facilitates multi-scale feature aggregation while preventing a significant increase in computational load (Sapkota et al., 2024b). On the other hand, it integrates the Convolutional block with Parallel Spatial Attention (C2PSA block), which enhances spatial attention mechanisms, enabling the model to focus on the most relevant regions of an image for more precise object detection (Sapkota et al., 2024b). Finally, similarly to YOLOv5 and later versions (He et al., 2024), the head produces predictions based on the processed features by defining bounding boxes around the detected objects, providing a probability for each detected object.

2.2. *Description of the area of interest*

The area of interest covers a ~ 1700 m stretch of the river Rímac situated in the city of Lima, along the central coast of Peru. It consists of two sections of roughly 800 and 900 m each, separated by road bridges and located 6 km away from the river mouth in the Pacific Ocean (see Fig. 2). The average flow rate of the rivers is 34.7 m^3/s, occasionally exceeding 100 m^3/s during the wet season (Astorayme and Gutiérrez, 2017). The river exhibits bimodal behavior, with peak water flows typically occurring from December to March. However, these flows are registered upstream from the area of interest. Hence, considering that the river acts as a vital source of water for the city (Vázquez-Rowe et al., 2017), the flow in the section analyzed is expected to be considerably lower throughout the year than it would naturally be, which inherently affects river dynamics and sediment and pollutant mobility. In fact, the water treatment plant of La Atarjea, which is located 10 km upstream from the area under study, captures on average 15 m^3/s to produce potable water for the city (Vázquez-Rowe et al., 2017).

The selection of the section of the river was done in two stages. Firstly, images from the Peruvian satellite PerúSAT (Chávez et al., 2024) were used for a general assessment of the watercourse up to its mouth, which enabled the identification of a potential area of study. Secondly, a field visit was carried out that allowed delimiting the area into two sections (see yellow and red boxes in Fig. 2), in which the presence of informal dumping along the riverbanks was observed. The first section extends from the *Bella Unión* Bridge (12°02′07″S; 77°04′40″W), connecting the districts of Cercado de Lima and San Martín de Porres, to 600 m downstream (see the yellow box in Fig. 3). This starting point was chosen given that road infrastructure upstream of the bridge restricts pedestrian movement and prevents the establishment of a safe area for drone flights. The second section extends from the *Belaúnde Terry* Bridge (12° 02′ 18″ S; 77° 05′ 26″ W), connecting the districts of San Martín de Porres (Lima) and Carmen de la Legua Reynoso (Callao), to the *Elmer Faucett* Bridge, 800 m downstream (see red box in Fig. 3). Beyond this point, the river enters a no-fly zone due to its proximity to the city's main airport and a military base which extends up to the river mouth in the Pacific Ocean.

2.3. *Data collection*

A full sampling campaign was conducted over a hydrological year from September 2023 to August 2024 with an unmanned aerial vehicle (UAV – i.e., a drone), although an additional campaign was carried out in March 2025. For this, a compact, ultra-lightweight drone (Mavic DJI Mini 3 Pro), equipped with a 48-megapixel camera, was used to collect images and videos (DJI, 2021). The remote-controlled UAV was complemented with two batteries, each providing 15 min of mission flight

Fig. 2. Map of the area under study, which comprises two sections of the river Rímac, red and yellow boxes, houses are observed on the right bank. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

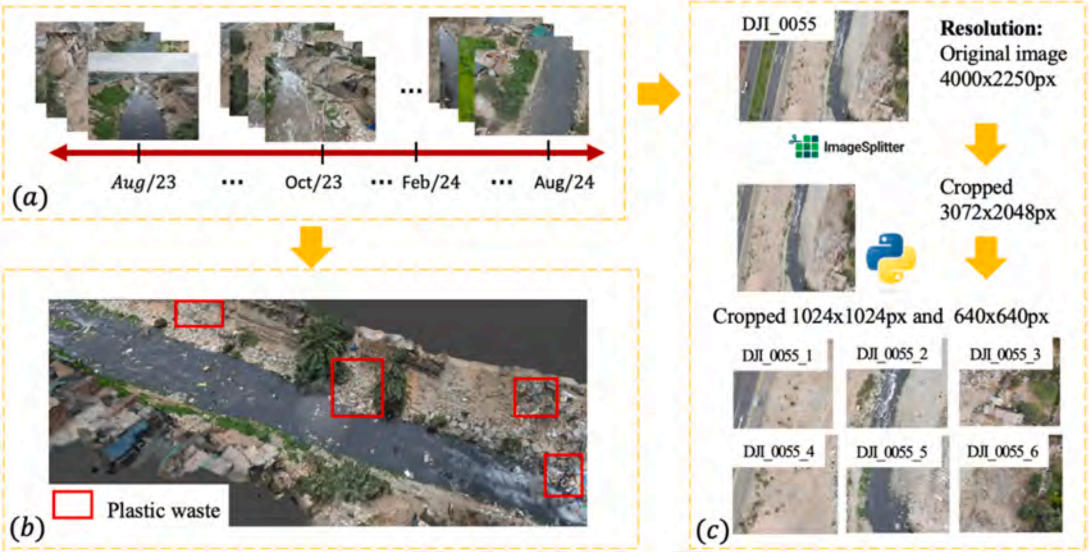


Fig. 3. Methodological scheme for image processing: (a) Review of aerial images for each month; (b) 3D model generated using photogrammetry techniques from aerial images, corresponding to the second section of the watercourse in October 2023, with areas of highest solid waste density marked in red; (c) To obtain aerial image inputs for the model, the original image was divided into six square patches of 1024 × 1024 and 640 × 640 pixels each. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

time (i.e., excluding departure and landing). The images collected were used to fine-tune the model, and the video recordings were subsequently employed to perform rapid counting of the target items using the trained architecture.

Initially, a few preliminary flights were conducted in August 2023 to validate the areas established for flights, identify constraints, and seek solutions to these. The video recordings were made considering the direction of flow, using as starting points the *Bella Unión* and *Belaúnde Terry* bridges. Fieldwork included 2 sampling periods per month, with an average of about 150 RGB images of 4032x3024px in each sample, as well as a Full HD video recording for each watercourse. Flights were generally conducted on Fridays between 9 am and 12 pm to avoid peak traffic hours. However, on some occasions, schedules were adjusted due to safety concerns.

In terms of seasonality, during the winter season (i.e., June to September), flights were avoided on foggy or rainy days to prevent image quality loss and reduce flight time due to wind resistance. In contrast, in the summer season (i.e., December to March), flights were scheduled earlier to avoid hours of intense solar radiation, as this could cause reflections or overexposure on the sensor, affecting the capture of images (Rodríguez et al., 2022). Additionally, high temperatures could overheat drone batteries, limiting flight duration.

The aerial images and video recordings from the first campaigns (August to September 2023) were captured at a height of 60 m. However, during the image processing, it was observed that this height did not allow clear differentiation of some of the objects that were being monitored. Therefore, the images from the first four flight campaigns (Aug-Sep 2023) were not used for training the DL model. Thereafter, the target flying height of the drone was reduced to 40 m. While a lower altitude could have improved image clarity of the video recording, 40 m was established as a safe height threshold, minimizing the risk of drone collision due to the shape of the river, settlements on the slopes, and obstacles such as power lines and bridge overhangs. Nevertheless, in some specific areas of the water sections, where a higher concentration of debris was found, additional images were collected at lower altitudes (between 20 m to 30 m) to help differentiate the objects that were classified in the study from other inorganic waste.

Videos recorded from September 2023 to August 2024 were included for the quantification of the polymers after the model training, allowing the continuity of information throughout the hydrological year. In contrast, images and videos from the final campaign that was carried out in March 2025 were used for the testing of the previously trained model.

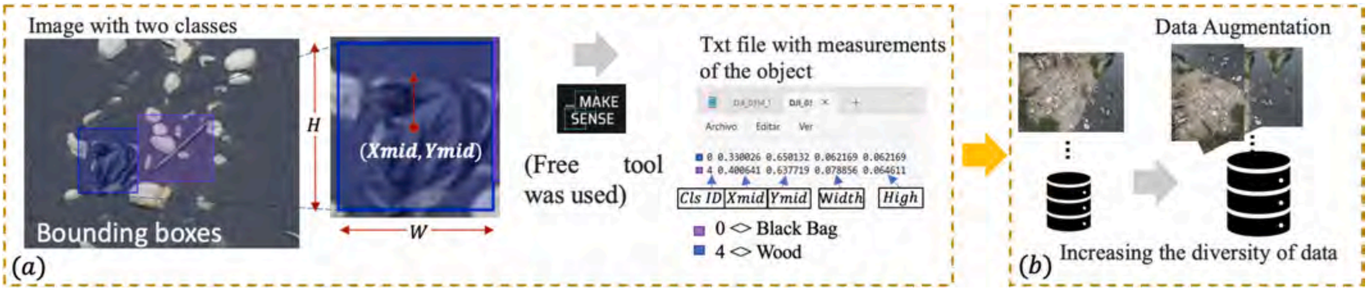
2.4. *Image processing*

A set of 3D models of the watercourse were created using photogrammetric techniques from aerial images using Agisoft PhotoScan (see Fig. 3b) with the aim of an in-depth understanding of the area under study (Agisoft, 2023). This process was carried out monthly, starting in September 2023. However, generating models for all months was not possible due to the orientation and low overlap between aerial images, which were avoided to optimize UAV battery usage. Despite this constraint, the 3D models served as a reference to getting familiarized with the two sections of the river and obtain a better reference of the location of the objects detected, facilitating the labeling task (see Section 2.5).

The images from each flight campaign were reviewed, selecting those with greater clarity, avoiding overlaps, and discarding those without a relevant density of objects that did not correspond to the sections and their slopes (see Fig. 3a). The images, with a native resolution of 4000 × 2250 pixels, were manually cropped to 3072 × 2024 using the free-use online portal *image-splitter* to establish an appropriate size (Postcron, 2023). Subsequently, a code was created in Python to divide into six 1024 × 1024 px images (see Fig. 4c), which went through a manual review and selection to establish the final dataset. The resolution chosen was determined based on the original image size, the size of the objects to be detected, and their concentrations in the area of interest. Additionally, the images in the dataset were cropped to a native resolution of 640 × 640 to reduce training time and enable further hyperparameter exploration for the YOLOv11 model in alignment with the GPU's capabilities during the training phase.

2.5. *Generation of labels*

Once the solid waste accumulated in the two river sections was manually evaluated, a total of 8 classes of objects were established for labeling (see Table 2). Given the heterogeneity of debris and waste present in the sections, it was decided to classify not only polymer-based objects, such as black bags, synthetic raffia bags, other bags, tires, and polyethylene terephthalate (PET) bottles, but also other objects that may



```
0 0.330026 0.650132 0.062169 0.062169
4 0.400641 0.637719 0.078856 0.064611
```

Fig. 4. Graphical representation of the methodological process that begins with image collection, followed by the identification of the target classes (*cls ID*), measurement of bounding boxes (H:*High* and W:*Width*), and determination of their center coordinates (X_{mid}, Y_{mid}). Additionally, data augmentation techniques, such as exposure adjustments and image rotation, are applied to generate new variations to enhance model robustness. (a) Image labeling involves generating a bounding box using makesense.ai, the annotations of which are then exported in YOLO format, where each row follows the structure: *Clsd ID*, X_{mid}, Y_{mid}, *W*, *H*. Each row represents the position of an object within the image, and the annotations are saved in a .txt file with the same name as the corresponding image. Each image is uniquely associated with a single .txt file in YOLO format, containing its object annotations. (b) Data augmentation techniques were applied on the training set and generated rotations, changes in coloration, and compositional techniques, from hyperparameters of the YOLO model.

Table 2
Classification and description of the debris and waste categories considered in the generation of labels.

Category	Category ID	Description
Black bags	0	Plastic garbage bags (mainly polyethylene – PE).
Other bags	1	Plastic bags of other colors (mainly PE).
Synthetic raffia bags	2	Raffia bags (braided polypropylene – PP).
Tires	3	Terrestrial vehicle tires (i.e., synthetic rubber).
Wood	4	Wooden bars or boxes.
Furniture	5	Sofas, chairs and cabinets.
PET bottles	6	Polyethylene terephthalate (PET) bottles.
Other objects	7	Clothes, shoes, buckets, or plastic containers.

be made partially out of plastic (other objects and furniture) or made exclusively of wood (i.e., wooden bars). The eight classes, consisting of five objects derived from MPs and three non-plastics categories, correspond to the most abundant and visually identifiable elements in the images.

Classes 0 to 3 correspond to different types of MPs, including black bags (Class 0), other types of bags (Class 1), synthetic raffia (Class 2), and tires (Class 3). Classes 0 and 1 are mainly bin bags or waste sacks made of polyethylene (PE). Class 6, which detects PET plastic bottles completes the list of plastic waste. However, it should be noted that classes 0 through 2 are carrier bags, so they may contain differing amounts of other waste fractions, including other plastic polymers. Class 4 corresponds to wood items that were identified in the riverbanks, which is the only full class that is plastic-free. Class 5, in contrast, detects large furniture objects, such as sofas, chairs or cabinets. It should be noted that most of these objects will most commonly include some plastic fractions, although this is not necessarily visible in the imaging. Finally, Class 7, named “others”, includes other objects that may or may not contain plastic fractions, depending on the characteristics of each object. These are not present with the same level of recurrence as the previous classes in the training dataset, which would have resulted in insufficient samples for reliable model training if subdivided into individual categories.

Considering YOLOv11 is a supervised learning model, a labeling process was performed on each previously processed image (see Section 2.5) to identify the 8 object classes. The goal of labeling was to generate a bounding box for each object present in the processed images. This information, along with the corresponding images, served as input for model training. The labels were made on the free portal called makesense.ai/ (Makesense.ai, 2025). Although labeling is a manual process, the free portal facilitates the process by exporting the labels with relevant information such as the class number (ranging from 0 to 7) and information on the measurements of the bounding box associated with the object in the image in meters (see Fig. 4a).

A total of 2553 images were labeled for the training and validation set and 300 images for the test set. The latter set corresponded to the last fieldwork conducted in March 2025. In addition, the dataset includes 46 images recorded at low altitudes (15–25 m) during new fieldwork conducted in July 2025. These images specifically targeted underrepresented classes, thereby contributing to reducing the imbalance across classes. The labeling process took place mainly between January 2024 and February 2025, which included manual quality control to avoid labeling errors. In the labeled images, a higher number of labeled objects were identified in classes 0–2 and 4, with 6472, 6010, 2935, and 3862 detections, respectively (see Fig. 5). However, the remaining classes had fewer than 1000 labeled objects, with classes 3 (i.e., tires) and 6 (i.e., PET bottles) having the lowest counts, at 158 and 307, respectively. This resulted in an imbalance among classes. To address this issue, the data augmentation technique was applied (Lecun et al., 1998), using the Python Albumentations library (Buslaev et al., 2020), as well as the model hyperparameters associated with data augmentation (see Table 1S in SM) that enable geometric transformations (*Fliplr*, *Scale and Translate*), color adjustments (Hsv_v, Hsv_s and Hsv_h), and compositional techniques (*Mosaic*) providing the model with diverse perspectives of the same object, enhancing its ability to generalize (Xu et al., 2023).

The dataset was split into 75 % for training and 25 % for validation (see Fig. 5). The dataset exhibits a certain degree of class imbalance. As the training and validation sets were selected randomly while ensuring a proportional distribution of each class in both stages, a python script was made for this purpose.

2.6. Experimentation

2.6.1. Hardware specifications, dependencies, and configuration

For the experimental phase, a desktop computer equipped with an Intel 12th Gen Core (TM) i7 -12700k processor (20 CPUs), with an Nvidia GeForce RTX4080 graphics card of 16 GB was used. For the experiments that included the entire dataset, a server from the Department of Engineering at PUCP was used, equipped with a 48 GB Nvidia GeForce RTX A6000 graphics card. The experiments were carried out in the Python programming language (Python, 2024), within the Visual Studio Code environment (Microsoft Corporation, 2024).

The coding of the model was supported by the framework provided by Ultralytics (Jocher and Qiu, 2024) and it was executed using PyTorch dependencies in Python 3.8. The experiments were conducted successively, gradually increasing the dataset size over time. The first

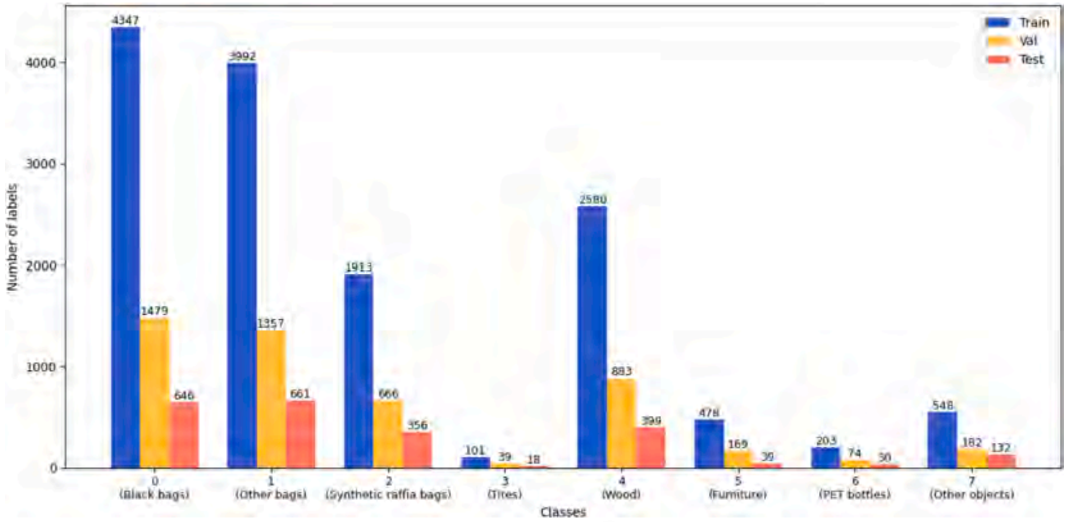


Fig. 5. Distribution of the total number of labels generated per class in the dataset. The training set (blue bars) was used to adjust the model's parameters by learning patterns from labeled images. The validation set (yellow bars) was used during training to monitor performance and help prevent overfitting. Finally, the testing set (red bars) was used to evaluate the model's performance on a separate set of labeled images that it had not seen before (i.e., a new dataset). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

experiments used the YOLOv8 model provided by Ultralytics, as it was the most recent version at the time. However, with the release of YOLOv11 in September 2024, this model was chosen for the experiments until the final dataset was established. In addition, TL was applied by importing the pre-trained weights of the YOLOv11 model using the universal COCO dataset (Lin et al., 2014), which facilitated convergence during training.

2.6.2. *Metrics and dataset used*

To evaluate the performance of the model, the metrics Precision (P), Recall (R), and mean average precision at 50 (mAP50) were used (see Table 3). P measures the accuracy of detections, indicating how many positive predictions of the model are correct (Kent et al., 1955); R measures the model's ability to correctly detect all positive instances in the data (Oksuz et al., 2018); and mAP50 is calculated by averaging the accuracy of all classes in detection, considering a prediction correct if the intersection over union (IoU) is greater than or equal to 50 % (Lin et al., 2014). Among the three metrics used, mAP50 is the most widely adopted and accepted metric for object classification (Mao et al., 2019). Therefore, it was selected as the target metric to be optimized throughout the experiments.

P and R are defined using true positives (TPs), false positives (FPs), and false negatives (FNs), serving as key performance evaluators. TPs occur when the model correctly identifies an object. FPs arise when the model incorrectly detects an object, and FNs occur when an actual object is missed and not detected by the model.

2.6.3. *Fine-tuning process: training, validation, and testing phases*

The hyperparameters considered for initial experiments included the learning rate (Lr), batch size and epochs (see Table S2 in the SM). Lr was initially determined automatically using the Adam optimizer (Tong et al., 2022), using command *optimizer* = *Auto,* which automatically selects the most appropriate Lr and weight-decay hyperparameters (see Table S2 in the SM). The training primarily focused on evaluating the performance of the model with different batch sizes and epochs (see Section S3 and S4 in the SM). The experiments were carried out when time labeled images became available. The initial experiments were conducted using square images with a resolution of 1024 × 1024, exploring initially YOLOv8 and the most robust architecture subversions of YOLOv11, namely: the large and extra-large subversions (see Section S3 in the SM). These were later migrated, along with their labels, to square images with a resolution of 640 × 640 (i.e., the native input size accepted by the model), as training did not allow for an increased batch size due to GPU limitations (see Section S4 in the SM). In addition, the effect of layer freezing was explored by progressively unfreezing the first ten layers until none remained frozen. A slight decrease in performance was observed (see tables S23-S28 in the SM), leading to the decision to discard layer freezing.

The hyperparameter values identified in previous experiments (i.e., Lr, batch size, epochs and weight-decay) were used to define search ranges for a random hyperparameter optimizations process. This process was carried out using the *Optuna* library (Akiba et al., 2019) and aimed to identify the set of hyperparameters that yielded the best performance based on the mAP50 metric. The optimization involved multiple training sessions with values sampled from predefined ranges, as detailed in Table S2 in the SM. In addition, training-related hyperparameters, such as *patience* (for early stopping) and *seed* (to ensure reproducibility by fixing the initial weights), were also configured to stabilize the training process. The remaining hyperparameters were used with their default values. To reduce computational overhead, the search of hyperparameters was conducted with 10 epochs, identifying two potential candidates with the best performance from a total of 8 models trained (see Table S33 in the SM). Consequently, the two models selected (see tables S34-S35 in the SM) were trained with the previously identified hyperparameters to improve generalization. Experiments with an increasing number of epochs revealed that the model avoided overfitting up to epoch 17, at which point training was intentionally stopped.

Additionally, a testing phase was conducted for the two best models found, which involved evaluating the ability of the model to predict

Table 3
Performance metrics used in the evaluation of the YOLOv11 model results for macroplastics (MPs) detection in the current study (Sapkota et al., 2025).

Metric	Symbol	Equation	Description
Precision	P	$P = \frac{TP}{TP + FP}$	Ratio of true positive detections to the total predicted positives.
Recall	R	$R = \frac{TP}{TP + FN}$	Ratio of true positive detections to the total actual positives.
Intersection over union	IoU	$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union}$	Measures the overlap between the predicted and actual bounding boxes.

TP = true positives; FP = false positives; FN = false negatives.

labels on a set of images that the model had not seen before. For this, a set of 300 images from the fieldwork carried out in March 2025 were used. The best model was selected based on a comparison of the metrics obtained during the validation and testing phase, whereas the results of the unselected model are shown in in the SM.

Thereafter, two modifications to the model were tested with the aim of obtaining improvements in the performance scores. On the one hand, the super-resolution (SR) technique was explored by using the SRTransGAN model (Baghel et al., 2024), based on transformers (Islam et al., 2024) within a generative adversarial network (Goodfellow et al., 2020) framework. According to Baghel and colleagues this model combines an encoder–decoder generator with a Vision Transformer-based discriminator to enhance the global reconstruction of images. It was applied to increase the resolution of the dataset images from 640 px to 1280 px. Due to GPU limitations, it was not possible to explore any other image sizes (i.e., 2480 px or its multiples). On the other hand, with the aim of mitigating class imbalance, the Python Albumentations library was applied to augment the underrepresented classes. Two data augmentation scenarios were considered: (i) a twofold increase in the least frequent classes, with the constraint that augmented classes would not exceed 40 % of the instance count of the most represented class, and (ii) a threefold increase in underrepresented classes, subject to the same 40 % threshold relative to the most frequent classes.

Finally, videos recorded along the area of interest were used to identify and evaluate the behavior of MPs throughout the hydrological year with the trained model. The recording of the highest quality of the two videos performed each month was selected based on the following criteria: better flight stability, i.e., controlled flight and without much turning or topping, and the longest length travel associated with the recorded section. However, in some cases, a manual review of the videos was necessary to prevent double-counting, as the UAV occasionally stopped mid-flight or changed direction for safety reasons, ensuring there were no obstacles nearby.

2.6.4. *Generalization of the model to other rivers*

The model was tested in a second river, the river Chillón, which passes through Lima's northern districts, providing water resources for agriculture and urban areas while facing pollution challenges from industrial and urban runoff, as well as informal waste dumping. While the main scope of the current study was not aimed at testing the model in a wide range of different rivers, it was considered a valuable input in the testing phase. For this, a ca. 400 m stretch near the river mouth was selected in which informal housing and waste dumping close to the riverbanks is predominant. From the aerial survey conducted in July 2025, 394 images were analyzed. Thereafter, to evaluate the detection performance of the model, manual and automatic counting were conducted along the 400 m stretch. A final step performed with the river Chillón was to compare the automated results obtained with the model with a classic manual counting model to determine the accuracy of the model. The reason for conducting this final step with the river Chillón and not the Rímac was linked to the fact that the model had not previously seen its characteristics, providing a more meaningful assessment.

3. Results and discussion

3.1. *Performance of the model*

The performance of the different training runs using default hyperparameters, as well as those observed when increasing the number of labeled images over time (i.e., a larger dataset), were computed with the aim of identifying the model with the best performance. In general, direct relationship is noted when observing the different models between dataset size and per-class metrics, along with an improvement in mAP50 performance when moving from the YOLOv8 to the YOLOv11 model (see sections S3 and S4 in the SM). In fact, Table S33 in the SM presents the metrics from 8 training runs carried out during the random hyperparameter search phase using the *Optuna* Python library. From this set of runs, YOLOv11-x (extra-large) was the model with the best performance detected in terms of the main metrics (e.g., P, R and mAP50), with higher consistency between the performance of the validation and test phases, as low variation during these stages is desirable to ensure good generalization (Roth et al., 2024).

Results, as shown in Table 4, show that classes 0 (black bags), 1 (other colored bags), and 3 (tires) achieved the best mAP50 scores. As a result, plastic objects in these categories are less likely to be misclassified (i.e., fewer false positives). Firstly, Class 3 achieved an mAP50 of 0.89 during the validation. Moreover, in the testing stage the same value of 0.94 was attained, indicating good consistency and suggesting that the model has the potential for rapid counting of these elements. The *P* value of this class is notably high, which can be attributed to the well-defined shape, size, and contrast with the environment, e.g., exposure and saturation of the tires, which likely facilitates detection (Redmon et al., 2016). In addition, P and R values in this class are similar during both the validation and testing stages, indicating good balance.

Secondly, classes 0 and 1 showed relatively high mAP50 values >0.74 for the validation and testing (0.74 and 0.63, respectively) stages. However, the *P* values obtained for these categories (in validation stages) are somewhat lower, probably due to their greater complexity, especially in terms of shape, which makes them more difficult to detect. Moreover, P and R values across these two classes are similar during the validation stage, which, as mentioned above, indicates good balance. The higher values of P as compared to R in the testing stage suggest that the model correctly classifies most of the objects it detects (i.e., low rate of false positives); however, the lower value for R with respect to P also indicates that the model struggled when detecting several objects that are actually present (i.e., false negatives) (Oksuz et al., 2018). These undetected objects could be associated with variations in shape and size of the plastic bags, which tend to lack a well-defined structure when dumped in riverbeds (Tharani et al., 2021) making it difficult for the model to recognize them if it has not previously encountered similar shapes.

PET bottles (i.e., Class 6) also show relatively low presence in the river and are less prominent in relation to the image size. Consequently, the dataset contained few labeled instances, which made it challenging for the model to generalize their detection. This limitation is also partly due to the altitude at which some images were captured, reducing their visibility and distinguishability. The low presence of PET in the area of interest is likely due to its economic value, as both informal and formal collectors gather and resell them (Ciudad saludable, n.d.), so that in

Table 4
Model performance on validation and testing using the best YOLOv11x configuration and hyperparameters.

Model:	Yolov11x						
Stage		Validation			Testing		
Classes	**Class ID**	**P**	**R**	**mAP50**	**P**	**R**	**mAP50**
Black bags	0	0.68	0.71	0.75	0.81	0.54	0.74
Other bags	1	0.67	0.72	0.74	0.64	0.58	0.63
Synthetic raffia	2	0.59	0.52	0.55	0.52	0.55	0.55
Tires	3	0.85	0.90	0.89	0.93	0.94	0.94
Wood	4	0.59	0.52	0.68	0.46	0.57	0.55
Furniture	5	0.68	0.57	0.66	0.62	0.42	0.49
PET	6	0.60	0.48	0.51	0.39	0.60	0.57
Others	7	0.50	0.31	0.31	0.52	0.33	0.35
Hyperparameter							
Epochs	17						
Batch size	16						
Lr	1.1716E-04						
Weight-decay	3.6325E-04						
Optimizer	AdamW						

Lr = learning rate; P = precision; R = recall; mAP = mean average precision; PET = polyethylene terephthalate.

many cases they reenter the plastic production system as flakes or pellets to produce recycled PET (Queiroz et al., 2021). Due to the distinct shape of PETs compared to other classes such as black bags, their performance metric could potentially be improved through additional flight campaigns aimed at increasing the number of annotations. As discussed in Section 2.5, this issue was partially addressed through complementary fieldwork, which successfully increased the mAP to 0.57 in the testing stage.

For classes 2 (synthetic raffia) and 4 (wood), an mAP50 value of 0.55 was obtained in the testing stage. In the case of class 2 the validation stage presented the same value (i.e., 0.55), while a notable decrease was observed in the testing stage, as the validation value was 0.68. These results suggest that the model is able to detect these objects, but with considerably lower accuracy than bags or tires. However, on a more positive note, the model also ensures good consistency by showing that its performance does not decrease significantly when moving to the testing stage. On the one hand, in the case of synthetic raffia bags ($P = 0.52$ and $R = 0.55$ in the testing stage), it was found that they are often partially buried, deteriorated, or torn, which could have influenced the difficulty when detecting these objects. As a result, the precision and recall metrics are considerably low. On the other hand, for wooden bars ($P = 0.46$ and $R = 0.57$), when many are stacked together, the model tends to ignore those that are incomplete. Additionally, their coloration can sometimes blend with that of the watercourse slopes, where vegetation is scarce, further hindering detection. Moreover, the density of solid waste present along the slopes contributes to the difficulty in detecting this class. In fact, in some cases bounding boxes overlapped by more than 50 %, leading to a labeling approach that prioritized wooden bars fully visible in the image. This limitation may have affected the ability of the model to detect wooden objects effectively.

For class 5 mAP50 values were not those desired (i.e., 0.66 in the validation stage and 0.49 in the testing stage), showing that the model is unable to efficiently detect these items. It was observed that many furniture objects were partially submerged in the watercourse or buried in the riverbanks, with the model failing to detect them in these cases. Moreover, the limited detection performance is likely due to the small number of labeled examples available for the training and testing stage. However, as shown in Fig. 5, tires also have a low number of labeled examples in the validation stage, but their distinct shape and color with respect to other residues allows for higher mAP50 values with low labeling rates. In any case, it is expected that the accuracy of the model for class 5 would improve with a larger training dataset.

Finally, Class 7 (i.e., other waste items) presented very low values for mAP50, P and R, as the model presented increased limitations to detect these objects, given their high variability in terms of materials, sizes, or shapes (i.e., clothing, pieces of polystyrene foam…) and the low number of labeled samples.

Overall, the model presented is useful for a first group of classes, which would allow for its use under a series of different circumstances as a tool for decision-making and effective removal of waste objects from watercourses. This is the case, for instance, with tires, as the model shows very high effectiveness in detecting these objects, proving to be a valuable tool to support tire removal from riverbeds in rivers with similar characteristics to those of the river Rímac. Waste bags, regardless of color, also presented high effectiveness in detection. This feature is very useful as a means of identifying and quantifying the impact of informal dumpsters along the riverbanks of rivers, a valuable trait for waste management authorities (e.g., municipalities) to enhance their waste collection and disposal systems (Cottom et al., 2024). Moreover, it should be noted that many of these bags are totally or partially filled with other waste, which can include a wide array of non-hazardous or hazardous municipal waste, demolition waste, etc. Hence, the use of a model of these characteristics to guide waste identification and removal from rivers would not only enhance the disappearance of plastic bags from riverbeds, but also a more complex array of waste items.

In contrast, a second group of classes show low efficiency rates in terms of being detected by the model, which hinders the utility of the latter as an aid in the removal of these objects. However, it is also true that the abundance of these items, as shown in Fig. 5, is significantly lower than those modeled in classes 0–3, except for the case of wood (i.e., Class 4). As regards wood it is important to note, however, that while many items arrive directly from natural sources (e.g., branches), other wood items are dumped along the riverbank (e.g., wooden bars and boxes). These latter items may have different types of coatings with toxic substances that are being emitted into the riverine ecosystem (Hedmark and Scholz, 2008) or that they are stacked or mixed (i.e., hetero-aggregation), complicating the detection task.

Two modifications to the model were tested with the aim of obtaining improvements in the performance scores. On the one hand, when the super-resolution technique was applied (see Table S37 in the SM), marginal improvements were identified in classes which had originally shown low scores, whereas slight reductions in mAP were observed for the classes with higher initial performance (i.e., tires or bags). These findings suggest that applying super-resolution may distort larger objects, which led to the decision to discard its use. On the other hand, when we compared applying data augmentation to the classes less represented in the sample as opposed to augmenting all classes through the YOLOv11 hyperparameters (see tables S38–39 in the SM), the latter was identified as the most appropriate strategy (see Table 4). We hypothesize that the lower effectiveness of augmenting underrepresented classes is likely due to the generation of an artificial imbalance, where these classes dominated through synthetic variations (Buda et al., 2018).

Finally, regarding the hyperparameters identified and shown in Table 4, these were obtained from a random search within a systematic training process and bounded by previously established ranges (see Table S2 in SM). Therefore, it cannot be ruled out that other hyperparameter combinations may yield similar results. However, it is important to note that hyperparameter optimization requires considerable computational resources, time, and careful prioritization of which parameters to include in the search (Feurer and Hutter, 2019). In fact, the more hyperparameters involved, the higher the computational cost.

In this study, priority was given to the weight-decay hyperparameter and the activation of the optimizer parameter, in addition to Lr, batch size, and epochs, which are known to have a strong influence on performance. The former helps control how quickly the model converges during training, while the latter allows for flexibility in determining the Lr rather than relying on a default value. The hyperparameter optimization process was guided by the goal of maximizing the mAP50 metric on the validation set throughout the training epochs (Fig. 6d).

As part of the evaluation of the results, Fig. 6 illustrates the evolution of the loss function values, divided into three components: *bounding box regression loss* (*box_loss*), *classification loss* (*cls_loss*), and *distribution focal loss* (*dfl_loss*). The decreasing trend across epochs suggests that the model is learning effectively. To mitigate overfitting, the epochs hyperparameter played a critical role, as it was observed that the model began to overfit after approximately 17 epochs (see Fig. 6c).

Box_loss, which evaluates the accuracy of the bounding box prediction compared to the manually annotated ones (labels) is shown in Fig. 6a. It is observed that the model consistently reduces the loss with each epoch, indicating that it improves its ability to localize objects correctly as training progresses. A similar behavior is seen with *dfl_loss* (Fig. 6c), which evaluates the accuracy at each edge of the predicted bounding box with respect to the one made manually. Towards the final epochs, the loss stabilizes again with a slight upward trend, indicating that the model has stopped improving and may no longer be learning.

Cls_loss (Fig. 6b) measures the performance of the model in terms of correct class identification. Of the three losses, it is the most relevant for this study, as correct object classification is prioritized over precise bounding box localization. In this context, small differences between predicted and manually annotated bounding boxes are acceptable. A consistent downward trend in Cls_loss is observed throughout the 17 epochs. Beyond this point, the loss stabilizes and no longer decreases,

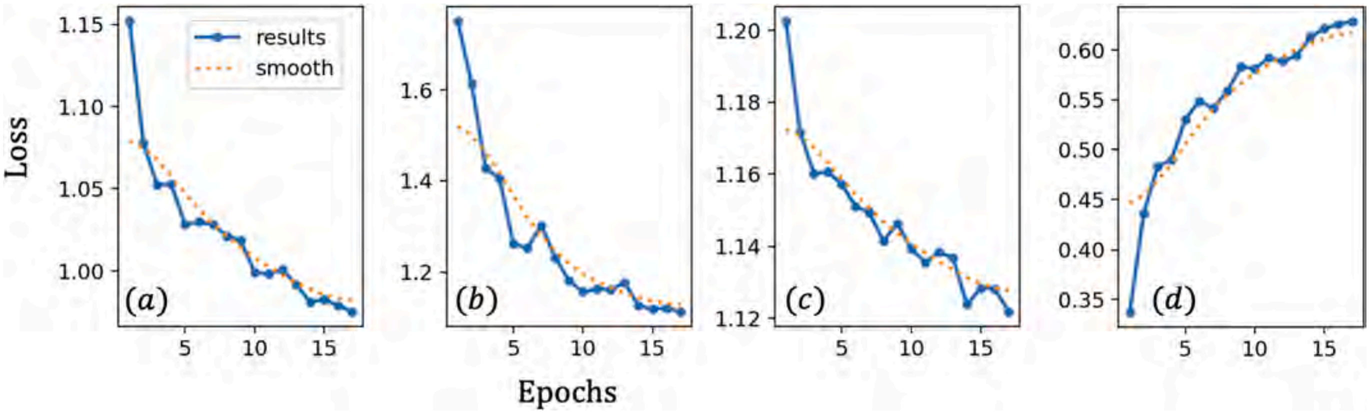


Fig. 6. Evolution of the loss function values and mAP50 throughout training epochs during the validation stage. Fig. 6a shows the bounding box regression loss (box_loss). Fig. 6b presents the classification loss (cls_loss). Fig. 6c shows the distribution focal loss (dfl_loss). Fig. 6d illustrates the evolution of the mean average precision (mAP50) metric.

showing a slight upward trend, reflecting overfitting.

3.2. Detection and classification

While the YOLO model selected can detect and classify a given class, whether correctly or incorrectly, this is limited to the probability (confidence measure), with which the model assigns the class type, which determines whether the classification is correct (Jocher and Qiu, 2024). In this sense, analyzing the classification probabilities across the image set can provide valuable insights into the strengths and weaknesses of the model, based on the different scenarios in which a given object appears. Consequently, object detection and classification were performed on each image in the test set (see Section S2 in the SM).

Fig. 7 shows five selected examples from this set, demonstrating how the model generates object detections (i.e., bounding boxes) and their corresponding classifications. For instance, it can be observed that the probability assigned to Class 3 (i.e., tires), which has the highest mAP50, is high in Fig. 7a, while its confidence decreases in Fig. 7e. This suggests that, in cases of stacked objects, as shown in Fig. 7e, the model loses confidence in the classification.

Similarly, for Class 4 (wood), Fig. 7a displays two objects of this class. It is evident that the probability of being classified as wood is higher when the color contrasts with the rocks and the watercourse in the background, providing a clearer distinction. In contrast, when the object is surrounded by stones, as in the second case, the color is less distinguishable. In the case of black bags, we observe that the

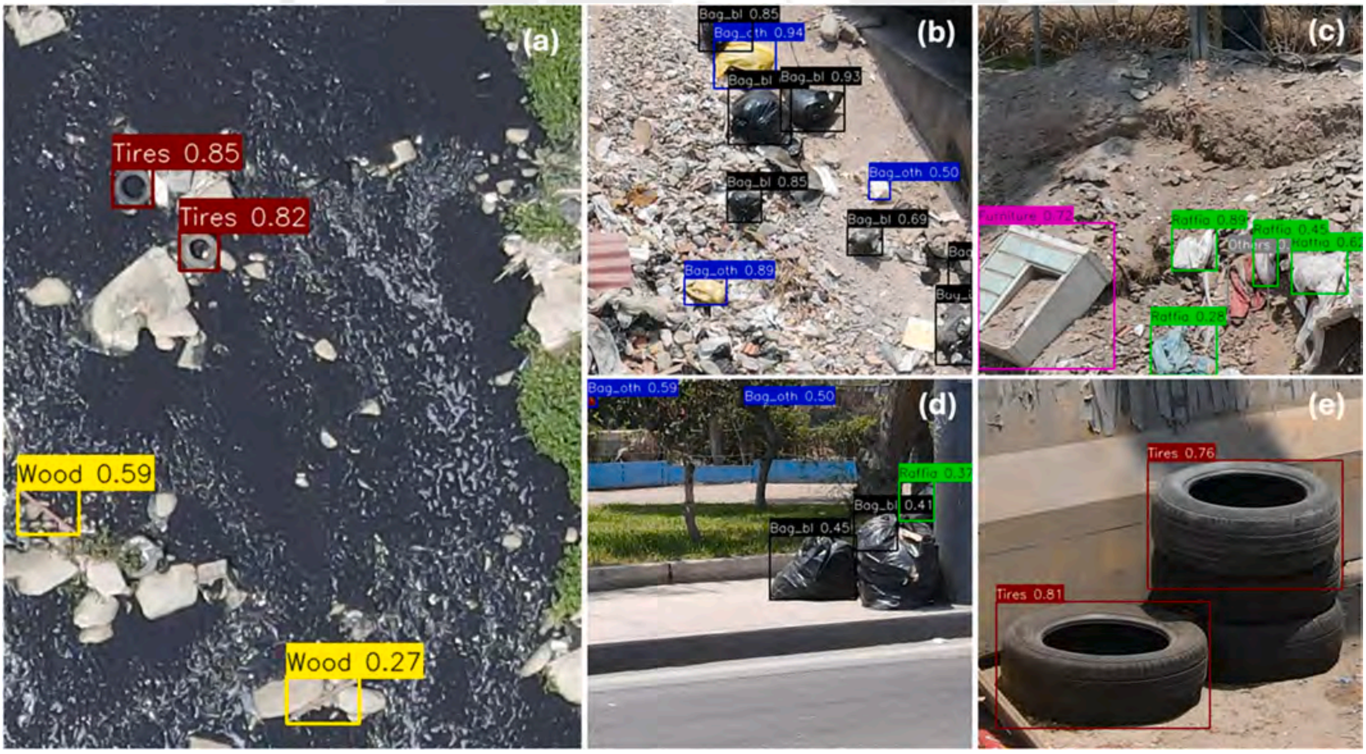


Fig. 7. Images of the testing set showing detected objects with their bounding boxes, class, and detection probability for classes 0 (black bags), 1 (other bags), 2 (synthetic raffia), and 3 (wood) and 4 (furniture), colored black, blue, green, and red, and magenta, respectively. Fig. 7a highlights objects from classes 2 and 4 detected along the watercourse. Fig. 7b shows objects from classes 0 and 1 along the riverbank. Fig. 7c shows the identification of objects from classes 2 and 4. Fig. 7d shows objects from classes 0, 1, and 2 outside the area of interest. Fig. 7e shows tires (Class 3) outside the area of interest. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

probability is high when the bags are round and more pronounced (Fig. 7b), while it decreases as the bags become smaller. A further decrease in confidence is observed in Fig. 7d, where the shape of the bag changes to a rectangular geometry due to the base and stacking, and the reflection of sunlight distorts its color. Finally, for the case of synthetic raffia (see Fig. 7c), we observe that when the raffia bag is deteriorated or open, forming an amorphous geometry, the probability of detection decreases.

Regarding the probability assigned to each object detected in the testing set, Fig. 8 shows a wide dispersion in confidence values. For black bags (i.e., Class 0), most detections fall within two probability ranges: 28 %–40 % and 70 %–95 %. The presence of detections in the lower range supports the idea discussed above, in which the model lacks certainty in those specific cases. For Class 1 (bags of other colors than black), the separation is more distinct, with two well-defined ranges: 25 %–50 % and 78 %–92 %, where most detections are concentrated. This could be attributed to variations in the coloration of the bags relative to their surrounding environment. In our study, the most common colors were yellow, red, and blue, which may have facilitated more confident detections. In contrast, the lower probability range could be primarily associated with small, deteriorated, or irregularly shaped bags. This applies also to raffia bags (i.e., Class 2), which were often in poor condition or partially buried, two factors that likely reduced the detection confidence of the model.

3.3. Evaluation of macroplastics (MPs) concentration over a hydrological year

When examining the concentrations of black bags (Class 0) in the area of interest, slight fluctuations were detected throughout the hydrological year, as shown in Fig. 9. The modest increase in their volume during the dry season may be explained by the combination of two main factors, the discharge of new waste along the riverbanks, on the one hand, and the fact that these bags may be carried by the current, especially during the rainy season, on the other. It is also possible that in the rainy season, some of the bags are partially or totally submerged under the water flow, while these will emerge with the sharp reduction of water flow in the dry season. Based on our own observations during flight campaigns, black bags, commonly used to store municipal solid waste, are frequently discarded, which contributes to their recurrent presence. Additionally, their mobility is influenced by their weight, with discarded bags more likely to fall down the riverbank into the water flow if they are heavier, while these same bags will be less likely to be moved by the water flow. The floatability of the materials inside the plastic bags is another parameter that will affect their mobility along the watercourse.

In contrast, the concentration of colored bags shows a more fluctuating behavior, with higher levels during the dry season (July to December) compared to the rainy season (January to June), reaching the lowest levels between April and May. Although both colored and black bags are commonly used for solid waste disposal, our field observations indicate that colored bags are generally smaller. Hence, it is likely that colored bags are more easily carried away as river flow increases, resulting in a reduced concentration during periods of flooding.

Regarding tires, a very low concentration is observed compared to other classes, with the lowest levels recorded during the flood season. This reduced concentration during floods could be related to the increased flow, which either completely submerges these items or allows their mobility under high flow conditions. In the case of furniture items (class 5), a behavior similar to that of the black bags class is observed, although on a different scale. Due to the weight of most furniture items, they are more likely to end up underwater or partially covered, making them undetectable by the modeling. Additionally, smaller furniture pieces may be carried away by the higher flow rates recorded during

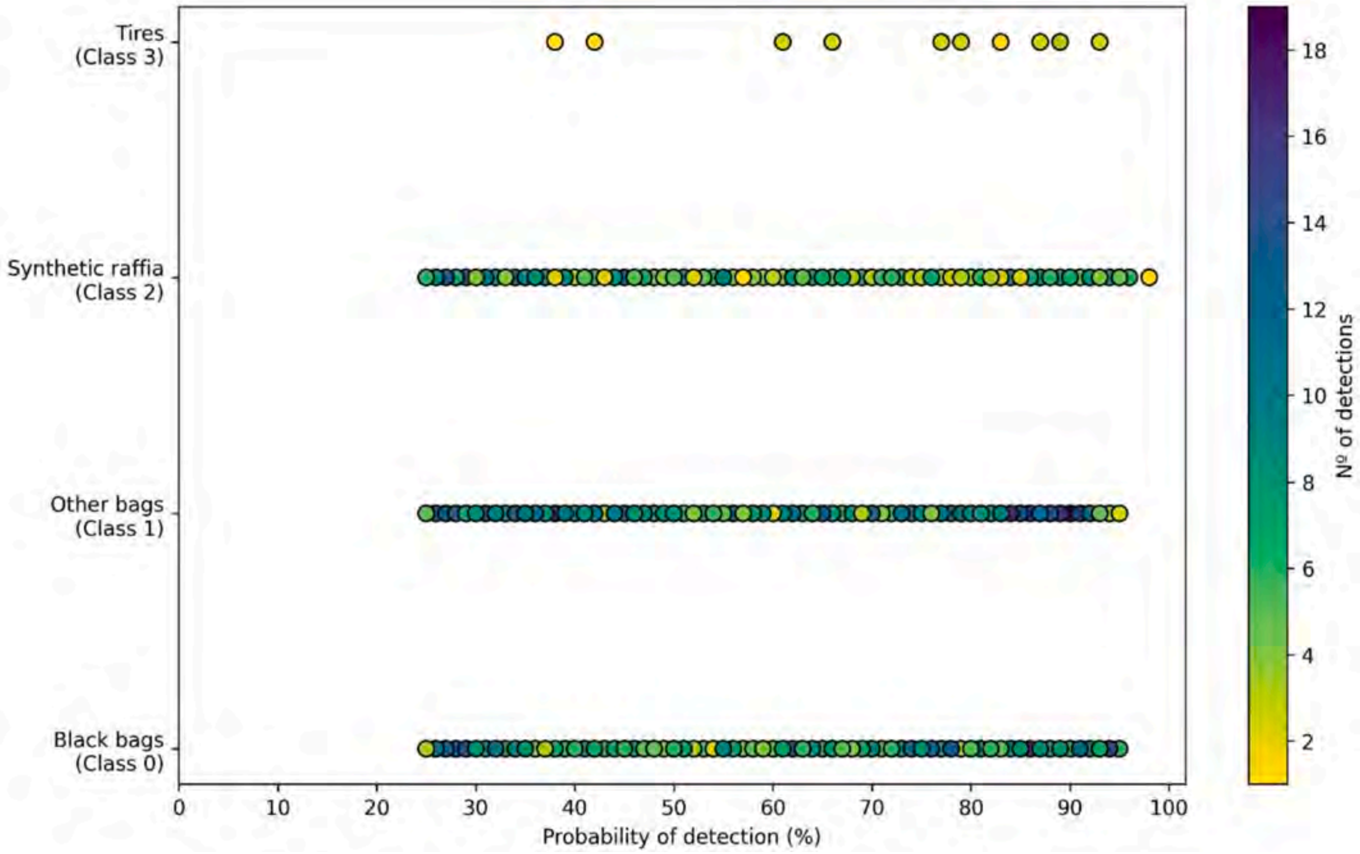


Fig. 8. Scatter plot of the probability assigned to each detected object for selected classes, assigning the number of detected objects with a color palette.

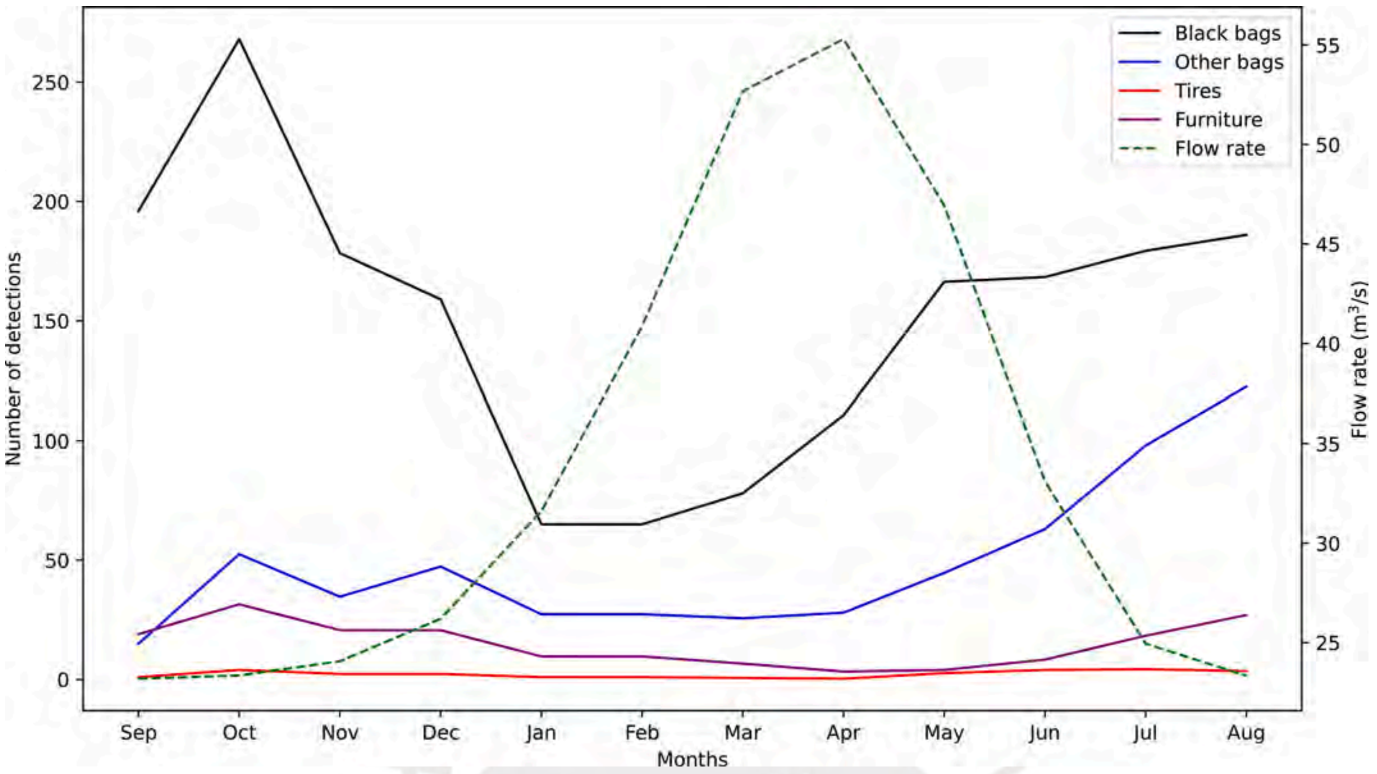


Fig. 9. Time series of macroplastic (MP) quantification per month for the final water section assessed in the river Rímac in the 2023/2024 hydrological year. The average monthly flow is represented by the dotted line.

floods. However, based on our observations throughout the year, it is also possible that people remove some furniture items for reuse. In fact, we detected that some homeless use pieces of furniture previously discarded in the river to build temporary structures for sleeping near the bridges that mark the area of interest, especially during the dry season. It is likely that the detection of the furniture class could be affected, as the items may be grouped in such a way that they lose their original shape.

3.4. *Performance of the model in an alternative river*

When the model was applied to the aerial images from the river Chillón, the testing results, which can be observed in Table 5, show an overall decrease in performance for the classes with lower performance, such as synthetic raffia, PET bottles or furniture. In contrast, tires maintained the highest performance (mAP = 0.74), although at lower rates than in the river Rímac (mAP = 0.94). The bag classes showed very similar performance results, with other bags (i.e., Class 1) showing slight improvements (mAP = 0.66). In addition, wood (Class 4) also showed a moderate increase (mAP = 0.61). Consequently, the model is still representative for the best-performing classes, and we hypothesize that its detection abilities could be extendable to other river with similar characteristics along the Peruvian coast.

Table 5
Performance metrics obtained in the testing stage for the river Chillón, as well as a comparison between automatic and classic manual counting methods.

Classes	Class ID	P	R	mAP50	Automatic counting	Manual counting
Black bags	0	0.73	0.6	0.72	724	604
Other bags	1	0.71	0.54	0.66	1071	935
Synthetic raffia	2	0.43	0.40	0.32	213	129
Tires	3	0.58	0.80	0.74	14	10
Wood	4	0.66	0.52	0.61	339	302
Furniture	5	0.42	0.43	0.32	20	14
PET	6	0.36	0.40	0.32	164	67
Others	7	0.30	0.26	0.17	223	136

However, future efforts to extend the use of this model to other rivers would benefit from obtaining additional data (i.e., images) from other rivers, allowing it to learn from slight differences in climatic characteristics (e.g., brightness), terrain, or other parameters. For instance, in the case of the river Chillón it was observed that some areas of the watercourse were somewhat brighter green vegetation, which could be influencing the detection ability of the model.

Finally, the automatic counting of the model in this watercourse was compared to manual counting. The results, shown in Table 5, indicate that the model generally overestimates object counts, with the best performance observed for black and colored bag classes, followed by the tire class, while the remaining classes showed lower performance. In the case of tires, 10 items were detected in the manual counting as compared to 14 in the automatic count. This discrepancy can be attributed to the limited number of images and annotations available for this class, which hinders its performance despite the relatively high mAP (0.74).

3.5. *Environmental impacts related to MPs*

The model detects macro-waste categories that pose significant environmental risks through microplastic emissions, chemical contamination, or flood hazards (Al-Zawaidah et al., 2021). For the former, mechanical forces (e.g., water flow and interactions with sediments or boulders) combine with biochemical forces (e.g., sunlight or microbial degradation) to increase the degradation rates of plastic waste into micro- and nanoplastics (Chamas et al., 2020). However, it should be noted that certain studies have demonstrated how these forces, together with polymer types, sizes or shapes, can generate heterogeneous degradation rates depending on the characteristics of the river (e.g., perennial vs. intermittent), water flow rates, or the location of the waste items in the inundated or non-inundated areas of the watercourse (Liro et al., 2023). In the case of the river Rímac, very high solar radiation in certain periods of the year (Cacciuttolo et al., 2024), and highly variable water flow rates between the rainy and dry seasons (Vázquez-Rowe et al., 2017) could suggest that degradation rates may be substantially higher than in average fluvial conditions. However, this hypothesis should be validated with experimental tests under the conditions described, as data on degradation rates and pathways in the natural environment is still sparse (Chamas et al., 2020).

For instance, plastic bags and synthetic raffia represent particularly hazardous categories due to their high surface area-to-volume ratios and fibrous structures, facilitating rapid fragmentation into microplastics (Chamas et al., 2020). Moreover, these fragments readily absorb persistent organic pollutants and heavy metals, creating toxic vectors that bioaccumulate through food chains (Wang et al., 2021). Although degradation rates in freshwater systems have been reported to be understudied as compared to what would occur in seawater or soil, differences in degradation rates between different types of plastic polymers are expected to be notable (Chamas et al., 2020). For example, polyethylene is known to be more resistant to photodegradation than polypropylene (Weber et al., 2011), a characteristic that could highly determine not only the emission rate of plastic bags, but also the rate of emission of the contained waste in those bags.

Similarly, tires, which the model effectively identified, is one of the largest sources of aquatic microplastics globally, also releasing zinc compounds and benzothiazoles that exceed toxicity thresholds for sensitive aquatic species (Chamas et al., 2020). While most studies on microplastic release from tires focus on the abrasion process when used in transportation (Wagner et al., 2018), their disposal as tire waste also creates persistent contamination in riverine sediments through continuous weathering and mechanical breakdown.

The co-occurrence of these diverse waste categories creates synergistic environmental effects, altering riverine biogeochemistry and providing surfaces for pathogenic bacterial biofilm formation (Wu et al., 2019), and an important vector for different size polymers to penetrate the food web through freshwater systems or by eventually making their way into the ocean. In fact, given the closeness of the section of the river analyzed to the ocean, an important fraction of macroplastic waste and its derived microplastic fragmentation should be expected (Lebreton et al., 2017). Nevertheless, the retention capabilities of riverine systems should not be underestimated (Ita-Nagy et al., 2022), as macroplastics may accumulate at river infrastructure and in-channel structures, potentially blocking drainage systems and reducing open cross-sections, thereby increasing flooding risks and backwater rise (Honingh et al., 2020).

Our automated detection results provide an essential baseline framework that would enable a comprehensive environmental impact assessment in the river Rímac by delivering accurate identification and spatial mapping of macroplastic hotspots, critical prerequisites for risk evaluation and management prioritization in Lima's most relevant watershed.

3.6. *Limitations of the study*

Despite the utility of the model presented in the identification of certain classes of objects along the river Rímac, a set of limitations of the methodology presented must be discussed. Firstly, the density of solid waste in the areas of interest introduced further challenges, as semi-buried, degraded plastic elements and waste mixed with other materials were frequently encountered. In fact, high macroplastic density can affect the detection efficiency of YOLOv11 through object occlusion (He et al., 2024), class confusion (Zhang et al., 2024), or by considering duplicate detections when objects are very crowded in a particular section (Nasir et al., 2025). This complexity, which hampers detection, is also enhanced by other limitations linked to the use of DL techniques, e.g., the detection of MPs in images with weather-induced distortion, such as sun glare (Marye et al., 2025). According to Jia et al. (2023), future efforts should focus on developing more robust automatic detection models based on DL, capable of enhancing performance where MPs are located in adverse geographical, environmental, and other challenging conditions, as is often the case in countries of the Global South. It is therefore likely that such challenges have not been thoroughly addressed in previous studies. Furthermore, as noted by Astorayme et al. (2024), there is a limited number of studies conducted in river systems. Therefore, this research contributes to ongoing efforts to explore the application of deep learning in MPs detection.

Secondly, as mentioned in Section 3.2, the UAV flight altitude was set at 40 m above the takeoff point, which corresponds to the highest part of the riverbank. Consequently, objects located near or within the riverbed were captured from a higher relative altitude, due to the river cross-section depression, which can be up to 20 m lower. This variation in altitude reduced the image clarity of smaller objects, lessening the ability of the model to detect them, as CNNs rely on pixel density and object sharpness to learn effectively (Yan et al., 2021). This limitation could be mitigated by using higher-resolution cameras and UAVs with greater flight endurance, allowing for extended operation at lower altitudes and slower flight speeds, thereby minimizing the risk of collisions with surrounding obstacles present in the area of interest.

A third limitation regards the accuracy of the model, which could likely be improved with a larger dataset containing more labeled images, thereby reducing class imbalance. This could have been achieved by incorporating an additional section of the Rímac River, increasing the variability of the objects to be detected. However, due to the location of the section selected, expanding the study area would have required more complex logistics, as the nearest viable takeoff base was far from the area of interest. Additionally, this would have increased resource demands, required more batteries, and raised operational costs for each flight campaign.

Including additional images in the dataset showing the objects to be detected in different scenarios and positions (e.g., simulated events) such as those located in waterways, near river-adjacent housing, areas beneath bridges, and on slopes with varying vegetation density, could enhance the performance of the model, especially for classes with fewer labeled items (e.g., "PET bottles" and "Others"). However, conducting such data collection would demand a greater investment in time and logistics resources, which was not feasible within the scope of the study.

Finally, the YOLOv11 model enables rapid object detection once trained. This advantage over other architectures (e.g., Transformers) is attributed to its *single-pass* model. This feature was considered valuable for exploration in comparison to other models, particularly in the context of future adaptations for MPs monitoring systems or as an early warning system (Arishi, 2025), which is especially relevant in regions with inadequate solid waste management. However, this *single-pass* nature also entails a trade-off: the model sacrifices some accuracy in exchange for faster detection speeds. Finally, it should be noted that due to the architecture of YOLO, the model may fail to detect all objects of the same class when they are positioned too closely together.

4. Conclusions and perspectives

In light of the results obtained, the YOLOv11 model proposed has proven to effectively detect certain waste items, such as waste bags, tires and, to a lesser extent, synthetic raffia containers, all of which are recurrently present in riverbanks in an urban context in the city of Lima, allowing for a quick count of these. It should be noted, however, that the ability of the model to detect these objects depends on a combination of factors. These include the abundance of these items in the area of interest validated through flight campaigns, and characteristics related to the visualization of these objects, such as size, shape, color, among others. Interestingly, tires, despite the low recurrence in the sample, were detected effectively by the model, probably linked to their very characteristic ring shape, relatively large size, and distinctive dark tonality.

In contrast, other waste types, such as furniture or wood, were not detected with the same level of precision and accuracy, suggesting that the experimental method is not as valid for less abundant waste items with heterogeneous shapes, colors, and sizes. While capturing images at a lower altitude could have improved the metrics of the latter waste items, it was not feasible to conduct this option in the area of interest due to UAV safety constraints.

We consider that the trained model can serve as a foundation for its application in the rivers Rímac and Chillón, as well as other rivers along the Peruvian coast, as a tool for rapid identification and management of certain waste fractions, such as plastic bags and tires. Our framework would allow for continuous and objective assessment of waste accumulation in these riverine systems, enabling authorities to optimize waste collection routes, allocate resources more efficiently during cleanup activities, or establish urban mining strategies. Furthermore, this detection system could also support regulatory compliance by providing visual documentation of areas where illegal dumping along riverbanks has occurred.

The use of the model would also be appropriate to aid in cleanup activities along rivers with similar characteristics throughout the Peruvian coast, as well as rivers in other areas of the world with certain similarities (e.g., scarce vegetation, informal urban sprawl along riverbanks). In fact, the main advantage of the model lies in addressing the challenges posed by limited accessibility for manual counting within river ecosystems. By leveraging this model, the time and effort required to quantify MPs can be significantly reduced, as has already been demonstrated in more accessible contexts. Moreover, the advantages of this type of architecture can be further exploited in the development of real-time counting systems, enabling continuous monitoring and more responsive waste management strategies, with minimal human resources as compared to conventional detection techniques.

Finally, to the best of our knowledge, this study represents the first research in South America utilizing a CNN-based model for the detection of waste fractions, mainly MPs, in water bodies. While the study applies an existing YOLOv11 architecture, the research contribution lies in its novel application to MSW monitoring in riverine environments along the Peruvian coast, addressing a critical environmental impact with a distinct methodological framework as compared to conventional object detection models. In future studies, the collection of a wider sample of images, the use of improved cameras, or the incorporation of additional river sections of the river Rímac or other rivers with similar characteristics and types of plastic waste, could enhance the effectiveness of the model, improve metrics, and generalize its detection capabilities.

CRediT authorship contribution statement

Miguel Angel Astorayme: Writing – original draft, Methodology, Investigation, Formal analysis, Conceptualization. **Ian Vázquez-Rowe:** Writing – original draft, Validation, Supervision, Project administration, Funding acquisition, Conceptualization. **Eizo Muñoz-Sovero:** Writing – review & editing, Resources. **Ramzy Kahhat:** Writing – review & editing, Validation, Supervision, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

MSc. Miguel Angel Astorayme would like to thank the Peruvian Government through its ProCiencia Program for funding his Ph.D. studies through Grant Agreement No. 174-2020. He also extends his gratitude to the Artificial Intelligence Laboratory at PUCP, led by Prof. César Beltrán, for providing access to its servers. Profs. Ian Vazquez-Rowe and Ramzy Kahhat thank the Natural Environment Research Council (NERC) of the United Kingdom for financial support via the "Reducing the impacts of plastic waste in the Eastern Pacific Ocean" project (NERC reference: NE/V005448/1). The authors also acknowledge funding from the European Research Executive Agency (REA) under the Horizon Europe Research and Innovation Programme through Grant Agreement No. 101081744 (RAINFOREST project). Special thanks are extended to MSc. Ferdinand Pineda for his support in both technical and logistical aspects and BSc. Paulo Cardenas for his support in office work.

Appendix A. Supplementary data

The Supplementary Material (SM) to this article provides a historical description of the YOLO series, explaining the methodological improvements from YOLOv1 through YOLOv11. In addition, the SM provides a list of production inputs, as well as additional validation results that are not included in the main manuscript. In addition, two Google Drive folders have been created with the following information:

A. The river Rímac folder:

https://drive.google.com/drive/folders/1ntSyiumR6QHe6sn61Pj9sXy_BdohNj0k?usp=sharing

It contains the following subfolders:

i. Final Dataset. It includes the dataset used in the research (images with 640 px resolution) organized into the "Train_640", "Val_640", and "Test_640" folders representing the stages used for fine-tuning the model.

ii. Data Augmentation. It holds two scenarios generated by augmenting the images.

iii. Super Resolution". It contains the set of images upscaled to 1280 px and the original images used.

B. The river Chillón folder:

https://drive.google.com/drive/folders/1GKsXHKsQexLpfQpu-cHiV2RsDsLDivcr?usp=sharing.

It contains the set of images used exclusively for testing the model. Supplementary data to this article can be found online at https://doi.org/10.1016/j.marpolbul.2025.118649.

Data availability

Additional data not presented in the main manuscript or the Supplementary Material will be made available on request.

References

Agisoft, L., 2023. Agisoft Metashape (versión 2.0.2). https://www.metashape-la.com/agisoft-metashape.
Akiba, T., Sano, S., Yanase, T., Ohta, T., Koyama, M., 2019. Optuna. In: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pp. 2623–2631.
Alencar, M.V., Gimenez, B.G., Sasahara, C., Elliff, C.I., Velis, C.A., Rodrigues, L.S., Conti, L.A., Gonçalves-Dias, S.L.F., Cetrulo, T.B., Scrich, V.M., Turra, A., 2023. Advancing plastic pollution hotspotting at the subnational level: Brazil as a case study in the Global South. Mar. Pollut. Bull. 194, 115382.
Al-Zawaidah, H., Ravazzolo, D., Friedrich, H., 2021. Macroplastics in rivers: present knowledge, issues and challenges. Environ Sci Process Impacts 23 (4), 535–552.
Andrade-Muñoz, A.S., Miserendino, M.L., Quinteros, C.P., Dromaz, W.M., Abrial, E., Di Prinzio, C.Y., 2025. Macroplastic pollution in riparian corridors of urban and pristine mountain streams in Patagonia (Argentina). Sci. Total Environ. 959, 178092.
Arishi, A., 2025. Real-time household waste detection and classification for sustainable recycling: a deep learning approach. Sustainability 17 (5), 1902.
Astorayme, M.A., Gutiérrez, R.R., 2017. Statistical downscaling based correction of the HadGEM2 family general circulation models: Rimac Basin, Peru. World Environmental and Water Resources Congress 2017, 545–556.
Astorayme, M.A., Vázquez-Rowe, I., Kahhat, R., 2024. The use of artificial intelligence algorithms to detect macroplastics in aquatic environments: a critical review. Sci. Total Environ. 945, 173843. https://doi.org/10.1016/j.scitotenv.2024.173843.
Baghel, N., Dubey, S.R., Singh, S.K., 2024. SRTransGAN: image super-resolution using transformer based generative adversarial network. In: International Conference on Computer Vision and Image Processing. Springer Nature Switzerland, Cham, pp. 308–322.

Browning, S., Beymer-Farris, B., Seay, J.R., 2021. Addressing the challenges associated with plastic waste disposal and management in developing countries. Curr. Opin. Chem. Eng. 32, 100682.

Buda, M., Maki, A., Mazurowski, M.A., 2018. A systematic study of the class imbalance problem in convolutional neural networks. Neural Netw. 106, 249–259.

Buslaev, A., Iglovikov, V.I., Khvedchenya, E., Parinov, A., Druzhinin, M., Kalinin, A.A., 2020. Albumentations: fast and flexible image augmentations. Information *11* (2), 125.

Cacciuttolo, C., Guardia, X., Villicaña, E., 2024. Implementation of renewable energy from solar photovoltaic (PV) facilities in Peru: a promising sustainable future. Sustainability 16 (11), 4388.

Chamas, A., Moon, H., Zheng, J., Qiu, Y., Tabassum, T., Jang, J.H., et al., 2020. Degradation rates of plastics in the environment. ACS Sustain. Chem. Eng. 8 (9), 3494–3511.

Chávez, T.W., Moreno, R., Pasapera, J., Perez, N., 2024. Comparison of vegetation indices of crops from satellite images PerúSAT-1 and UAV. In: International Conference on Mediterranean Geosciences Union. Springer Nature Switzerland, Cham, pp. 303–306.

Chettry, V., 2023. A critical review of urban sprawl studies. J. Geovis. Spat. Anal. 7 (2), 28.

Chowdhury, G.W., Koldewey, H.J., Duncan, E., Napper, I.E., Niloy, Md.N.H., Nelms, S.E., Sarker, S., Bhola, S., Nishat, B., 2021. Plastic pollution in aquatic systems in Bangladesh: a review of current knowledge. Sci. Total Environ. 761, 143285.

Ciudad saludable. (n.d.). Ciudad saludable. Retrieved April 23, 2025, from https://www.ciudadsaludable.org/?utm_medium=referral&utm_source=idealist.

Cottom, J.W., Cook, E., Velis, C.A., 2024. A local-to-global emissions inventory of macroplastic pollution. Nature 633 (8028), 101–108.

DJI, 2021. DJI-Perú: authorized retail store. In: SDI Products. https://dji.pe/drones/serie-mini.

Dourojeanni, A., Jouravlev, A., 1999. Gestión de cuencas y ríos vinculados con centros urbanos. Comisión Económica para América Latina y el Caribe (CEPAL).

Feurer, M., Hutter, F., 2019. Hyperparameter optimization. In: Automated Machine Learning: Methods, Systems, Challenges. Springer International Publishing, Cham, pp. 3–33.

Fulton, M., Hong, J., Islam, M.J., Sattar, J., 2019. Robotic detection of marine litter using deep visual detection models. In: 2019 International Conference on Robotics and Automation (ICRA), pp. 5752–5758.

González-Fernández, D., Hanke, G., 2017. Toward a harmonized approach for monitoring of riverine floating macro litter inputs to the marine environment. Front. Mar. Sci. 4. https://doi.org/10.3389/fmars.2017.00086.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y., 2020. Generative adversarial networks. Commun. ACM 63 (11), 139–144.

He, J., Chen, H., Liu, B., Luo, S., Liu, J., 2024. Enhancing YOLO for occluded vehicle detection with grouped orthogonal attention and dense object repulsion. Sci. Rep. 14 (1), 19650.

Hedmark, A., Scholz, M., 2008. Review of environmental effects and treatment of runoff from storage and handling of wood. Bioresour. Technol. 99 (14), 5997–6009.

Hernández-González, J., Inza, I., Granado, I., Basurko, O.C., Fernandes, J.A., Lozano, J.A., 2019. Aggregated outputs by linear models: an application on marine litter beaching prediction. Inf. Sci. 481, 381–393.

Honingh, D., van Emmerik, T., Uijttewaal, W., Kardhana, H., Hoes, O., van de Giesen, N., 2020. Urban River water level increase through plastic waste accumulation at a rack structure. Front. Earth Sci. 8, 28.

Islam, S., Elmekki, H., Elsebai, A., Bentahar, J., Drawel, N., Rjoub, G., Pedrycz, W., 2024. A comprehensive survey on applications of transformers for deep learning tasks. Expert Syst. Appl. 241, 122666.

Ita-Nagy, D., Vázquez-Rowe, I., Kahhat, R., 2022. Developing a methodology to quantify mismanaged plastic waste entering the ocean in coastal countries. J. Ind. Ecol. 26 (6), 2108–2122.

Jambeck, J.R., Geyer, R., Wilcox, C., Siegler, T.R., Perryman, M., Andrady, A., Narayan, R., Law, K.L., 2015. Plastic waste inputs from land into the ocean. Science 347 (6223), 768–771.

Jia, T., Kapelan, Z., de Vries, R., Vriend, P., Peereboom, E.C., Okkerman, I., Taormina, R., 2023. Deep learning for detecting macroplastic litter in water bodies: a review. Water Res. 231, 119632.

Jocher, G., Qiu, J., 2024. Ultralytics YOLO11 (Version 11.0.0) [Computer software]. GitHub. https://github.com/ultralytics/ultralytics.

JRC, 2016. European Commission, Joint Research Center. Exploratory Research Project: RIMMEL (Riverine and Marine Floating Macro Litter Quantification). https://mcc.jrc.ec.europa.eu/main/dev.py?N=simple&O=397&titre_chap=%C2%A0&titre_page=RIMMEL%20observation%20network%20call.

Kent, A., Berry, M.M., Luehrs, F.U., Perry, J.W., 1955. Machine literature searching VIII. Operational criteria for designing information retrieval systems. Am. Doc. 6 (2), 93–101.

Kumar, R., Verma, A., Shome, A., Sinha, R., Sinha, S., Jha, P.K., Kumar, R., Kumar, P., Shubham, Das, S., Sharma, P., Vara Prasad, P.V., 2021. Impacts of plastic pollution on ecosystem services, sustainable development goals, and need to focus on circular economy and policy interventions. Sustainability 13 (17), 9963.

Lebreton, L.C.M., Van Der Zwet, J., Damsteeg, J.W., Slat, B., Andrady, A., Reisser, J., 2017. River plastic emissions to the world's oceans. Nat. Commun. 8, 1–10.

Lecun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proc. IEEE 86 (11), 2278–2324.

Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., et al., 2014. Microsoft coco: common objects in context. In: European Conference on Computer Vision. Springer International Publishing, Cham, pp. 740–755.

Liro, M., Zielonka, A., van Emmerik, T.H., 2023. Macroplastic fragmentation in rivers. Environ. Int. 180, 108186.

Makesense.ai, 2025. Makesense: Artificial intelligence solutions. Https://Www.Makesense.Ai/.

Mao, H., Yang, X., Dally, W.J., 2019. A delay metric for video object detection: what average precision fails to tell. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 573–582.

Margallo, M., Ziegler-Rodriguez, K., Vázquez-Rowe, I., Aldaco, R., Irabien, Á., Kahhat, R., 2019. Enhancing waste management strategies in Latin America under a holistic environmental assessment perspective: a review for policy support. Sci. Total Environ. 689, 1255–1275.

Martin, C., Parkes, S., Zhang, Q., Zhang, X., McCabe, M.F., Duarte, C.M., 2018. Use of unmanned aerial vehicles for efficient beach litter monitoring. Mar. Pollut. Bull. 131, 662–673.

Marye, A.T., Caramiello, C., De Nardi, D., Miglino, D., Proietti, G., Saddi, K.C., Biscarini, C., Manfreda, S., Poggi, M., Tauro, F., 2025. Remote sensing for monitoring macroplastics in rivers: a review. WIREs Water 12 (2).

Microsoft Corporation, 2024. Visual Studio Code (1.94.2).

Nasir, R., Jalil, Z., Nasir, M., Alsubait, T., Ashraf, M., Saleem, S., 2025. An enhanced framework for real-time dense crowd abnormal behavior detection using YOLOv8. Artif. Intell. Rev. 58 (7), 202.

Oksuz, K., Cam, B.C., Akbas, E., Kalkan, S., 2018. Localization recall precision (LRP): a new performance metric for object detection. In: Proceedings of the European conference on computer vision (ECCV), pp. 504–519.

Postcron. (2023). Image Splitter. Ttps://Postcron.Com/Image-Splitter/Editor/Es/Upload-Image.

Python, 2024. Python Software Foundation (3.6.13).

Queiroz, F.C.B.P., Lima, N.C., da Silva, C.L., Queiroz, J.V., de Souza, G.H.S., 2021. Purchase intentions for Brazilian recycled PET products—circular economy opportunities. Recycling 6 (4), 75.

Radhakrishnan, T., Manimekalan, A., Ghosh, D., Prasanna, R., 2024. Identifying high-vulnerable garbage accumulation areas in Coimbatore City, India: an AHP-GIS approach for effective waste management. Environ. Sci. Pollut. Res. 31 (14), 21797–21810.

Ramkumar, Mu., Balasubramani, K., Santosh, M., Nagarajan, R., 2022. The plastisphere: a morphometric genetic classification of plastic pollutants in the natural environment. Gondwana Res. 108, 4–12.

Ranjan, V.P., Joseph, A., Sharma, H.B., Goel, S., 2023. Preliminary investigation on effects of size, polymer type, and surface behaviour on the vertical mobility of microplastics in a porous media. Sci. Total Environ. 864, 161148.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788.

Rodríguez, J., Zapata, C., Torres, M., 2022. Importancia de los parámetros de exposición fotográfica y el geoposicionamiento en la calidad de productos obtenidos en modelaciones fotogramétricas con UAVs. Encuentro Internacional de Educación en Ingeniería, pp. 1–21. https://doi.org/10.26507/paper.2385.

Roebroek, C.T.J., Laufkötter, C., González-Fernández, D., van Emmerik, T., 2022. The quest for the missing plastics: large uncertainties in river plastic export into the sea. Environ. Pollut. 312, 119948.

Roth, B., Luz de Araujo, P.H., Xia, Y., Kaltenbrunner, S., Korab, C., 2024. Specification overfitting in artificial intelligence. Artif. Intell. Rev. 58 (2), 35.

Ryan, P.G., 2015. A brief history of marine litter research. In: Marine Anthropogenic Litter. Springer International Publishing, pp. 1–25.

Sapkota, R., Meng, Z., Karkee, M., 2024a. Synthetic meets authentic: leveraging LLM generated datasets for YOLO11 and YOLOv10-based apple detection through machine vision sensors. Smart Agric. Technol. 9, 100614.

Sapkota, R., Ahmed, D., Karkee, M., 2024b. Comparing YOLOv8 and mask R-CNN for instance segmentation in complex orchard environments. Artif. Intell. Agric 13, 84–99.

Sapkota, R., Flores-Calero, M., Qureshi, R., Badgujar, C., Nepal, U., Poulose, A., et al., 2025. YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series. Artif. Intell. Rev. *58* (9), 274.

Schreyers, L.J., van Emmerik, T.H.M., Bui, T.-K.L., van Thi, K.L., Vermeulen, B., Nguyen, H.-Q., Wallerstein, N., Uijlenhoet, R., van der Ploeg, M., 2024. River plastic transport affected by tidal dynamics. Hydrol. Earth Syst. Sci. 28 (3), 589–610.

Shamskhany, A., Li, Z., Patel, P., Karimpour, S., 2021. Evidence of microplastic size impact on mobility and transport in the marine environment: a review and synthesis of recent research. Front. Mar. Sci. 8, 760649.

Singh, N., Walker, T.R., 2024. Plastic recycling: a panacea or environmental pollution problem. npj Mater. Sustain. 2 (1), 17.

Teng, C., Kylili, K., Hadjistassou, C., 2022. Deploying deep learning to estimate the abundance of marine debris from video footage. Mar. Pollut. Bull. 183, 114049. https://doi.org/10.1016/j.marpolbul.2022.114049.

Tharani, M., Amin, A.W., Rasool, F., Maaz, M., Taj, M., Muhammad, A., 2021. Trash Detection on Water Channels, pp. 379–389.

Tong, Q., Liang, G., Bi, J., 2022. Calibrating the adaptive learning rate to improve convergence of ADAM. Neurocomputing 481, 333–356.

Vázquez-Rowe, I., Kahhat, R., Lorenzo-Toja, Y., 2017. Natural disasters and climate change call for the urgent decentralization of urban water systems. Sci. Total Environ. 605–606, 246–250.

Wagner, S., Hüffer, T., Klöckner, P., Wehrhahn, M., Hofmann, T., Reemtsma, T., 2018. Tire wear particles in the aquatic environment-a review on generation, analysis, occurrence, fate and effects. Water Res. 139, 83–100.

Wang, L.C., Chun-Te Lin, J., Dong, C.D., Chen, C.W., Liu, T.K., 2021. The sorption of persistent organic pollutants in microplastics from the coastal environment. J. Hazard. Mater. 420, 126658.

Weber, R., Watson, A., Forter, M., Oliaei, F., 2011. Persistent organic pollutants and landfills-a review of past experiences and future challenges. Waste Manag. Res. 29 (1), 107–121.

Wolf, M., van den Berg, K., Garaba, S.P., Gnann, N., Sattler, K., Stahl, F., Zielinski, O., 2020. Machine learning for aquatic plastic litter detection, classification and quantification (APLASTIC-Q). Environ. Res. Lett. 15 (11), 114042.

Wu, X., Pan, J., Li, M., Li, Y., Bartlam, M., Wang, Y., 2019. Selective enrichment of bacterial pathogens by microplastic biofilm. Water Res. 165, 114979.

Xu, M., Yoon, S., Fuentes, A., Park, D.S., 2023. A comprehensive survey of image augmentation techniques for deep learning. Pattern Recogn. 137, 109347.

Yan, E., Luo, L., Ceze, L., 2021. Characterizing and taming resolution in convolutional neural networks. In: 2021 IEEE International Symposium on Workload Characterization (IISWC). IEEE, pp. 189–200.

Zailan, N.A., Mohd Khairuddin, A.S., Khairuddin, U., Taguchi, A., 2021. YOLO-based network fusion for riverine floating debris monitoring system. In: 2021 International Conference on Electrical, Communication, and Computer Engineering (ICECCE), pp. 1–5.

Zhang, H., Li, G., Wan, D., Wang, Z., Dong, J., Lin, S., Deng, L., Liu, H., 2024. DS-YOLO: a dense small object detection algorithm based on inverted bottleneck and multi-scale fusion network. Biomim. Intell. Robot. 4 (4), 100190.

Chapter 4

Conclusion and future outreach

“Plastic pollution is not just an ocean problem, it is a people problem.”
—adapted from Sylvia Earle.

CNNS, to the best of our knowledge, have marked a milestone in the development of computer vision since their rise a few decades ago. With the increase in computational power, more robust models based on DL were proposed, e.g., the YOLO family and Very Deep Convolutional Networks. Applications like object detection, segmentation, and classification have been widely adopted in different scientific fields. From our critical analysis, we determined that the exploration of ML techniques for plastic waste detection in aquatic environments (i.e., rivers, oceans, lakes) started around 2018, while the use of more robust DL architectures emerged at the beginning of the current decade. Although transformers represent the state of the art in computer vision today, CNNs are still widely used in this field.

Recent studies show a clear trend towards the application of DL-based approaches, particularly TL techniques, which is also the case in this research through the use of YOLOv11. However, training these architectures requires large datasets, which entails significant fieldwork efforts and associated costs in technical staff, logistics, transportation, and time for acquiring aerial images. On the one hand, the resolution of satellite imagery is not sufficient to detect MPs at small scales, while it has proven useful for identifying larger artificial plastic objects in marine environments, highlighting its potential for environmental monitoring at broader spatial scales. On the other hand, the use of RGB and multispectral cameras mounted on UAVs has become more common for acquiring high-resolution imagery. Evidence shows that model performance tends to improve when multispectral cameras are used, as they allow for the detection of smaller MPs.
In addition, high computational capacity is needed, often requiring powerful GPUs or access to cloud services. These requirements likely limit research in developing regions such as Latin America and Africa, where, until recently, no published scientific studies were found in this specific line. Consequently, most of the advances come from developed countries, with a greater focus on the detection of MPs.

Within this context, our study provides valuable insights into the performance of DL-based detection models, specifically YOLOv11, under unfavorable conditions of solid waste management. We found that objects with well-defined shapes and uniform coloration were more easily detected like tires (mAP50 > 0.80). In a second level, plastic bags, although also monochromatic, posed more difficulties due to their irregular shapes, requiring large numbers of images for effective learning (mAP50 > 0.65). Wooden bars, raffia bags, and PET bottles achieved mAP50 values above 0.55, indicating that further image data are needed for more reliable detection. Based on our field experience, this may be explained by the difficulty of recognizing these objects when mixed with organic debris or with each other. For furniture (mAP50 between 0.49 and 0.60), the challenge lies in the limited number of examples available, despite their defined shapes. The others class also showed poor detection performance (mAP50 < 0.35), likely due to the heterogeneous mix of objects it included and the small sample size. In general terms, the detection capacity of the model decreases in areas of high waste concentration, where MPs often appear mixed with other materials, such as crossed wooden bars, nearly buried tires, or plastic bottles covered with mud and debris. Therefore, YOLOv11 demonstrated higher precision in detecting tires, black and colored plastic bags, followed by raffia bags and wooden debris. While the results for other categories suggest the need for additional training images, it is important to highlight that, under ideal conditions, where objects are clearly visible and not mixed with other waste, we found the model was able to detect them effectively.

From a plastic waste perspective, the presence of MPs in aquatic ecosystems, namely rivers, is a direct consequence of inadequate solid waste management, and has been increasingly reported as a critical environmental problem in Peru and other Latin American countries in the past decade. Understanding how waterways influence the transport of these inorganic elements is crucial for analyzing the interactions between the environment, MPs, and human activities. Based on more than a year of fieldwork along the final stretch of the river Rímac, in Lima (Peru), and after thorough interpretation of the results obtained, the following conclusions can be drawn:

i Solid waste tends to accumulate along riverbanks, particularly in areas with greater human activity, such as pedestrian and vehicle bridges, residential zones, and parks. During flooding periods, the concentration of visible waste decreases due to the increased flow and higher transport capacity. However, continuous human input of waste along the river margins eventually leads to its deposition into the riverbed, creating an ongoing cycle in which plastics remain present throughout the year. Our model was efficient in detecting waste accumulation throughout the hydrological year and visualized the changes across the period analyzed. Further research would be needed to quantify the fraction of plastic waste that becomes submerged in flooding periods, and the fraction that becomes mobile and continues its course to the ocean.

ii Denser or heavier inorganic materials often become buried or partially embedded during medium- or high-flow periods (NovemberDecember and AprilMay). This indicates that transport rates vary by material type: items such as tires and furniture persist longer in the river channel, whereas lighter plastic bags are more easily carried away.

iii Regarding the concentration of inorganic waste, plastic bags are the most common MPs in the urban rivers analyzed in this study in the city of Lima (i.e., Rímac and Chillón rivers), whereas PET bottles are less frequent. Due to their market value, PET bottles are often collected by informal recyclers, who sometimes enter the riverbeds to retrieve them. As a result, higher concentrations of PETs are typically found in areas that are difficult for humans to access.

iv Materials such as iron and copper wires, although present in lower concentrations, are often buried within riverbank debris and are also targeted by informal recyclers. Field observations showed the frequent presence of homeless individuals living along the riverbeds, building makeshift shelters under bridges and using furniture and clothing as improvised beds and coverings.

v Overall, the situation experienced reflects a close interaction between human population and urban river ecosystems, particularly in rivers along the Peruvian coast where population density is highest. In this context, it is likely that much of the solid waste reaching the ocean originates from these human interactions in the final stretches of rivers, where anthropogenic pressure is greatest.

The findings of this research confirm that plastic pollution in urban waterways is primarily associated with the use of plastic bags for storing and disposing of solid waste. Nevertheless, this issue is not limited to a single source but reflects a more complex problem with multiple dimensions that requires comprehensive and coordinated solutions. Although Peru took an important step with the approval of Law No. 30884 in 2018 to regulate the use of single-use plastics, our findings show that stronger efforts are still required to effectively curb plastic pollution. While the intervention of regional and municipal governments in monitoring compliance with this law is necessary, we argue that greater governance efforts are still required to implement policies that enhance and optimize solid waste management. For instance, the expansion of the existing network of sanitary landfills, which still do not cover all the major cities and towns across the nation, while gradually closing open dumps, appears as the most urgent and effective measure, as this would reduce significantly the rampant mismanaged waste throughout many regions in Peru.

However, not all plastic waste can be attributed to the lack of sufficient waste disposal management sites, as littering also plays an important role in the accumulation of MPs in Peruvian watersheds, beaches, or soils. Hence, other types of measures beyond the investment in waste disposition infrastructure must be explored. In fact, in December

2021 the Peruvian government banned the production and sales of expanded polystyrene (EPS). EPS is a low-density plastic polymer that degrades and disperses rapidly in the environment. Its elimination from the technosphere in Peru has been interpreted as a measure in the correct direction, but insufficient to control the enormous flow of plastic littering. Hence, the inclusion of additional policies, such as the implementation of taxes to discourage the consumption of single-use plastics, or the introduction of alternative materials that can substitute single-use plastic while maintaining its functionalities, must be explored.

In parallel, it is essential to intensify efforts to actively involve citizens as a central part of the solution. More insistent recycling campaigns, the establishment of municipal-run collection centers, and the promotion of environmental awareness in schools could prove valuable in the medium and long terms. The plastic pollution challenge in Peru has shown to be complex and requires the involvement of multiple stakeholders, including private companies, whose role is key not only in ensuring transparency in their production processes but also in supporting recycling and awareness campaigns, initiatives already underway to some extent. Furthermore, we believe that adapting successful public policies from developed countries could offer viable alternatives, such as direct incentives for citizens to return plastic containers in exchange for a refund of the deposit paid upon purchase, as well as enacting legislation that mandates stricter separation in waste collection and disposal systems.

In line with these international experiences, the restoration of the river Rímac, one of the main rivers of the Peruvian capital and a project proposed by the Autoridad Nacional del Agua (National Water Authority ANA), can only succeed if existing initiatives, such as recycling campaigns, are reinforced and adapted to current realities. The evidence of informal dumping sites along certain stretches of the river, driven both by local waste disposal needs and by informal businesses seeking economic benefit, reveals a structural challenge. Addressing this issue requires stronger monitoring by municipal governments, the establishment of formal landfills, and the implementation of timely household waste collection systems. These measures are essential for reducing plastic pollution and ensuring the long-term sustainability of river restoration efforts.

Within this broader framework, technological tools can also play a key role in supporting solid waste management. In this regard, while this research employs an existing CNN-based model for the detection of debris, primarily MPs, in river systems, building on this contribution, we propose that an automatic detection and counting system for MPs items along watercourses could significantly enhance solid waste management and support river ecosystem restoration. Rapidly identifying areas of higher density and understanding the spatial distribution of MPs is feasible based on the results obtained in this thesis and can provide valuable insights for decision-makers, particularly municipal governments, to improve the allocation of resources for waste collection and management. Furthermore, the results from adapting the YOLO model for the automatic quantification of MPs within a Peruvian context demonstrate its potential to bridge the existing information gap regarding their concentration dynamics in rivers, thereby deepening our understanding of their role as transport pathways to the ocean. Further calibration using imagery from other coastal rivers could enhance detection accuracy and allow for broader generalization in rivers affected by solid waste management challenges across the region. Nonetheless, the current results have already confirmed its detection capability in a second river (i.e., the river Chillón), with favorable outcomes for the most frequent items, particularly plastic bags. Finally, once greater generalization of the model has been achieved, establishing monitoring systems in major coastal rivers, particularly those flowing through urban areas, should be considered a crucial next step towards improving the estimation and analysis of MP fluxes, their concentrations, interactions within aquatic systems, and their connection to marine environments.

4.1 Future outreach

The effectiveness of plastic item detection is closely linked to the number of representative images, image resolution, and label quality. Future work should aim to expand the diversity of item classes in the training dataset to facilitate the application of the methodology to other rivers along the Peruvian coast. Establishing monitoring systems in major coastal rivers, particularly those that pass through urban areas, should be considered a next step toward improving the estimation and analysis of MP flows, their concentration, interactions within aquatic systems, and their connection to marine environments.

Bibliography

Aliani, S., Casagrande, G., Catapano, P., and Catapano, V. (2020). *Polarquest 2018 Expedition: Plastic Debris at 82ř07 North*, pages 89–116. Springer International Publishing.

Anderson, S. (2024). Health policy watch news: Un plastic pollution treaty derailed as fossil fuel nations block production limits.

Angelini, Z., Kinner, N., Thibault, J., Ramsey, P., and Fuld, K. (2019). Marine debris visual identification assessment. *Marine Pollution Bulletin*, 142:69–75.

Ballengee, B., Cohen, D., Fröschl, M. R., Valerio, M. A. G., Handschuh, S., de. Lutz, C., Maggic, M., Ruiz, R. M., Matysik, R. M., Rapp, R., Reichle, I., Sauer, M., Schwaha, T., Vendl, A., Vesna, V., Wagner-Lawlor, J., Yoldas, P., and ebjani, R. (2021). Plastic ocean : art and science responses to marine pollution. page 270.

Bautista-Zamudio, P. A., Flórez-Restrepo, M. A., López-Legarda, X., Monroy-Giraldo, L. C., and Segura-Sánchez, F. (2023). Biodegradation of plastics by white-rot fungi: A review. *Science of The Total Environment*, 901:165950.

Bell, G. (1998). Planning for change: Housing, land and sustainable development. *Journal of Planning and Environment Law*, pages 8 – 19.

Boldrini, A., Gaggelli, N., Falcai, F., Polvani, A., Talarico, L., Galgani, L., Cirrone, R., Liu, X., and Loiselle, S. (2024). Emerging contaminants from bioplastic pollution in marine waters. *Water*, 16:3676.

Brander, S. and Betjemann, P. (2023). *Marine Pollution*, pages 78–98. Routledge.

Buenos-Aires (2009). Ley 13868 | ministerio de ambiente.

Buenos-Aires (2016). Fin del anexo anexo-resoluciÓn nř 341/apra/16 separata del boletín oficial de la ciudad de buenos aires nř 267.

Campanale, C., Dierkes, G., Massarelli, C., Bagnuolo, G., and Uricchio, V. F. (2020). A relevant screening of organic contaminants present on freshwater and pre-production microplastics. *Toxics*, 8:100.

Choudhury, M., Sharma, A., Pervez, A., Upadhyay, P., and Dutta, J. (2022). *Growing Menace of Microplastics in and Around the Coastal Ecosystem*, pages 117–137.

Circulaelplastico (2018). £cuáles son los principales puntos de la ley chao bolsas plásticas?

Coram, A., Abreo, N. A. S., Ellis, R. P., and Thompson, K. F. (2021). Contribution of social media to cetacean research in southeast asia: illuminating populations vulnerable to litter. *Biodiversity and Conservation*, 30:2341–2359.

Cortesi, I., Masiero, A., Giglio, M. D., Tucci, G., and Dubbini, M. (2021). Random forest-based river plastic detection with a handheld multispectral camera. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIII-B1-2021:9–14.

Cortesi, I., Masiero, A., Tucci, G., and Topouzelis, K. (2022). Uav-based river plastic detection with a multispectral camera. In *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, volume 43, pages 855–861. International Society for Photogrammetry and Remote Sensing.

Dullah, A. A. M., Daud, A., Mallongi, A., Zakir, M., Jafar, N., Indar, I., Wahiduddin, W., Putro, G., and Birawida, A. B. (2025). A systematic review and bibliometric analysis of the process of human health risk assessment to microplastics exposure through seafood consumption. *Ecological Engineering & Environmental Technology*, 26:280–291.

Edo, G. I., Itoje-akpokiniovo, L. O., Obasohan, P., Ikpekoro, V. O., Samuel, P. O., Jikah, A. N., Nosu, L. C., Ekokotu, H. A., Ugbune, U., Oghroro, E. E. A., Emakpor, O. L., Ainyanbhor, I. E., Mohammed, W. A.-S., Akpoghelie, P. O., Owheruo, J. O., and Agbo, J. J. (2024). Impact of environmental pollution from human activities on water, air quality and climate change. *Ecological Frontiers*, 44:874–889.

Flores-Díaz, A., Arango, J. A., Calvo, D. C., Rangel-Mendez, R., and Ontiveros-Valencia, A. (2025). Physicochemical and microbial treatments for plastics, microplastics, and nanoplastics. *Environmental Development*, 54:101181.

Freitas, S., Silva, H., and Silva, E. (2021). Remote hyperspectral imaging acquisition and characterization for marine litter detection. *Remote Sensing*, 13:2536.

Glüge, J., Wang, Z., Bogdal, C., Scheringer, M., and Hungerbühler, K. (2016). Global production, use, and emission volumes of short-chain chlorinated paraffins a minimum scenario. *Science of The Total Environment*, 573:1132–1146.

Gómez, C. A. P., Torres, M., Pozo, K., and Berrío, D. E. (2024). Identificación de microplásticos en heces de lobos marinos (otaria flavescens) en el puerto la poza, región del bio bío, chile. *Revista de Investigaciones Veterinarias del Perú*, 35:e25593.

Hahladakis, J. N., Velis, C. A., Weber, R., Iacovidou, E., and Purnell, P. (2018). An overview of chemical additives present in plastics: Migration, release, fate and environmental impact during their use, disposal and recycling. *Journal of Hazardous Materials*, 344:179–199.

He, S., Wang, H., Li, H., and Zhao, J. (2021). Principle of machine learning and its potential application in cli-mate prediction. *Journal of Autonomous Intelligence*, 4:13.

Hernández-González, J., Inza, I., Granado, I., Basurko, O. C., Fernandes, J. A., and Lozano, J. A. (2019). Aggregated outputs by linear models: An application on marine litter beaching prediction. *Information Sciences*, 481:381–393.

Ho, C. M., Feng, W., Li, X., Ngien, S. K., Yu, X., Song, F., Yang, F., and Liao, H. (2025). Microplastic distribution and its implications for human health through marine environments. *Journal of Environmental Management*, 382:125427.

Hong, S. H., Shim, W. J., and Jang, M. (2024). *Chemicals associated with marine plastic debris and microplastics: Analyses and contaminant levels*, pages 141–179. Elsevier.

Hu, B., Wang, S., Yan, J., Zhang, H., Qiu, L., Liu, W., Guo, Y., Shen, J., Chen, B., Shi, C., and Ge, X. (2024). Review of waste plastics treatment and utilization: Efficient conversion and high value utilization. *Process Safety and Environmental Protection*, 183:378–398.

HUE, H., DONG, L., HIEN, T., NGUYEN, T., and PRADIT, S. (2021). Assessment of microplastics contamination in commericial clams in the coastal zone of vietnam. *Applied Ecology and Environmental Research*, 19:4977–4991.

ItaNagy, D., VázquezRowe, I., and Kahhat, R. (2022). Developing a methodology to quantify mismanaged plastic waste entering the ocean in coastal countries. *Journal of Industrial Ecology*, 26:2108–2122.

Jambeck, J. R., Geyer, R., Wilcox, C., Siegler, T. R., Perryman, M., Andrady, A., Narayan, R., and Law, K. L. (2015). Plastic waste inputs from land into the ocean. *Science*, 347:768–771.

Johnson, H., Keane, K., McGillivray, L., Akhtar-Khavari, A., Chambers, L., Barner-Kowollik, C., Lauchs, M., and Blinco, J. (2025). Reforming plastic packaging regulation: Outcomes from stakeholder interviews and regulatory analysis. *Sustainable Production and Consumption*, 54:52–63.

Kylili, K., Artusi, A., and Hadjistassou, C. (2021). A new paradigm for estimating the prevalence of plastic litter in the marine environment. *Marine Pollution Bulletin*, 173:113127.

Lavender, S. (2022). Detection of waste plastics in the environment: Application of copernicus earth observation data. *Remote Sensing*, 14:4772.

Lebreton, L. C. M., van der Zwet, J., Damsteeg, J.-W., Slat, B., Andrady, A., and Reisser, J. (2017). River plastic emissions to the worlds oceans. *Nature Communications*, 8:15611.

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521:436–444.

Lerpiniere, D. J., Wilson, D. C., and Velis, C. A. (2025). Official development finance in solid waste management reveals insufficient resources for tackling plastic pollution: A global analysis of two decades of data. *Resources, Conservation and Recycling*, 212:107918.

Ljubomir, J., Seba, S., and Ellik, A. (2009). Marine litter: a global challenge | unep - un environment programme.

Lorenzo-Navarro, J., Castrillón-Santana, M., Sánchez-Nielsen, E., Zarco, B., Herrera, A., Martínez, I., and Gómez, M. (2021). Deep learning approach for automatic microplastics counting and classification. *Science of The Total Environment*, 765:142728.

MADS (2016). Resolución 0668.

MADS (2022). Ley 2232 "por la cual se establecen medidas tendientes a la reducciÓn gradual de la producciÓn y consumo de ciertos productos plÁsticos de un solo uso y se dictan otras disposiciones".

Maes, C. (2022). Evaluation of the functionality of compostable bioplastics in food applications. In *TAPPI European PLACE Conference*, volume 2022-Octob.

Martin, C., Parkes, S., Zhang, Q., Zhang, X., McCabe, M. F., and Duarte, C. M. (2018). Use of unmanned aerial vehicles for efficient beach litter monitoring. *Marine Pollution Bulletin*, 131:662–673.

Mattsson, K., Jocic, S., de Lima, J. A., Hansson, L.-A., and Gondikas, A. (2024). *Nanoplastics in aquatic environmentsSources, sampling techniques, and identification methods*, pages 381–397. Elsevier.

Moore, C. and Phillips, C. (2011). Plastic ocean : how a sea captain's chance discovery launched a determined quest to save the oceans. page 358.

Moy, K., Neilson, B., Chung, A., Meadows, A., Castrence, M., Ambagis, S., and Davidson, K. (2018). Mapping coastal marine debris using aerial imagery and spatial analysis. *Marine Pollution Bulletin*, 132:52–59.

N., S. S., Haldar, S., Chanda, V., Muralidaran, Y., and Mishra, P. (2025). *Impact of microplastics on economic condition in underdeveloped nations*, pages 481–508. Elsevier.

Nawaz, F., Islam, Z. U., Ghori, S. A., Bahadur, A., Ullah, H., Ahmad, M., and ullah Khan, G. (2025). Microplastic and nanoplastic pollution: Assessing translocation, impact, and mitigation strategies in marine ecosystems. *Water Environment Research*, 97.

OCEANA (2019). £qué medidas de la ley de plásticos aún están pendientes?

Padermshoke, A., An, Y., Nguyen, T. V., Kobayashi, Y., Ito, H., and Takahara, A. (2025). Prooxidant-based polyolefins exhibiting no evidence of biodegradation under marine environments. *Marine Pollution Bulletin*, 214:117697.

Padha, S., Kumar, R., Dhar, A., and Sharma, P. (2022). Microplastic pollution in mountain terrains and foothills: A review on source, extraction, and distribution of microplastics in remote areas. *Environmental Research*, 207:112232.

Paglia, E. (2021). The swedish initiative and the 1972 stockholm conference: the decisive role of science diplomacy in the emergence of global environmental governance. *Humanities and Social Sciences Communications*, 8:2.

Pagliaro, M. (2019). Biodegradable and compostable bioplastics: Inflection point now in sight. *Chimica Oggi/Chemistry Today*, 37:70–71.

Peruano, E. (2018). Ley - nř 30884; ley que regula el plástico de un solo uso y los recipientes o envases descartables.

Pinochet, J., Thiel, M., and Urbina, M. (2024). How plastic litter sunk by biofouling recovers buoyancy - the role of benthic predation. *Science of The Total Environment*, 952:175910.

PlasticsEurope (2019). Plastics-the facts 2019 an analysis of european plastics production, demand and waste data.

por Naturaleza, C. (2019). 10 puntos para entender la ley de plásticos aprobada en el congreso.

Ramon-Gomez, K., Ron, S. R., Deem, S. L., Pike, K. N., Stevens, C., Izurieta, J. C., and Nieto-Claudin, A. (2024). Plastic ingestion in giant tortoises: An example of a novel anthropogenic impact for galapagos wildlife. *Environmental Pollution*, 340:122780.

RAP-AL (2023). Tercera reunión del comité intergubernamental de negociaciones sobre la contaminación plástica.

Sannigrahi, S., Basu, B., Basu, A. S., and Pilla, F. (2022). Development of automated marine floating plastic detection system using sentinel-2 imagery and machine learning models. *Marine Pollution Bulletin*, 178:113527.

Sapkota, R., Qureshi, R., Calero, M. F., Badjugar, C., Nepal, U., Poulose, A., Zeno, P., Vaddevolu, U. B. P., Khan, S., Shoman, M., Yan, H., and Karkee, M. (2024). Yolo11 to its genesis: A decadal and comprehensive review of the you only look once (yolo) series.

Scott, K. N. (2024). *Law of the sea in the Plasticene*, pages 122–144. Edward Elgar Publishing.

Senathirajah, K., Bonner, M., Schuyler, Q., and Palanisami, T. (2023). A disaster risk reduction framework for the new global instrument to end plastic pollution. *Journal of Hazardous Materials*, 449:131020.

SUNAT (2023). Impuesto al consumo de las bolsas de plástico (icbp).

Sánchez-García, N. and Sanz-Lázaro, C. (2023). Darwin's paradise contaminated by marine debris. understanding their sources and accumulation dynamics.. *Environmental Pollution*, 324:121310.

Taggio, N., Aiello, A., Ceriola, G., Kremezi, M., Kristollari, V., Kolokoussis, P., Karathanassi, V., and Barbone, E. (2022). A combination of machine learning algorithms for marine plastic litter detection exploiting hyperspectral prisma data. *Remote Sensing*, 14:3606.

Tamoria, R. M. and Schofield, J. (2024). *Islands of the Plastic Age*, pages 476–493. Routledge.

Teng, C., Kylili, K., and Hadjistassou, C. (2022). Deploying deep learning to estimate the abundance of marine debris from video footage. *Marine Pollution Bulletin*, 183:114049.

UE (2019). Directiva - 2019/904 - en - eur-lex.

UN (2015). Transforming our world: the 2030 agenda for sustainable development | department of economic and social affairs.

UN (2022). Intergovernmental negotiating committee on plastic pollution | unep - un environment programme.

UNCED (1992). Agenda 21: Sustainable development knowledge platform.

UNEA (2022). Resolution 5/14:end plastic pollution: Towards an international legally binding instrument.

UNEP (2022). Report of the intergovernmental negotiating committee to develop an international legally binding instrument on plastic pollution, including in the marine environment, on the work of its first session.

UNEP (2023). Report of the intergovernmental negotiating committee to develop an international legally binding instrument on plastic pollution, including in the marine environment, on the work of its third session. Technical report.

UNEP (2024a). Draft report of the intergovernmental negotiating committee to develop an international legally binding instrument on plastic pollution, including in the marine environment, on the work of the first part of its fifth session. Technical report.

UNEP (2024b). Report of the intergovernmental negotiating committee to develop an international legally binding instrument on plastic pollution, including in the marine environment, on the work of its fourth session. Technical report.

Wang, Y., Zhang, L., Qi, L., Yi, X., Li, M., Zhou, M., Chen, D., Xiao, Q., Wang, C., Pang, Y., Xu, J., Deng, H., Liu, L., and Guan, X. (2021). Machine learning: Applications and advanced progresses of radiomics in endocrine neoplasms. *Journal of Oncology*, 2021:1–17.

Wolf, M., van den Berg, K., Garaba, S. P., Gnann, N., Sattler, K., Stahl, F., and Zielinski, O. (2020). Machine learning for aquatic plastic litter detection, classification and quantification (aplastic-q). *Environmental Research Letters*, 15:114042.

Zhu, Y., Yeung, C. H., and Lam, E. Y. (2021). Digital holographic imaging and classification of microplastics using deep transfer learning. *Applied Optics*, 60:A38.