ARTIFICIAL INTELLIGENCE

ASSESSING THE USE OF AI
TO QUANTIFY PLASTIC POLLUTION IN RIVERS

with the support of :

ABOUT SURFRIDER FOUNDATION

Surfrider Foundation is a non-profit organization dedicated to the preservation and the promotion of lakes, rivers, oceans, waves, and coastlines. To date, it regroups more than 18,000 members and operates in 12 countries through its volunteer-led local groups. The organization focuses on three specific areas, in which it has accumulated internationally recognized expertise over the past 30 years: aquatic waste, water quality and user health, and coastal management.
To learn more: surfrider.eu

AUTHORS

→ Surfrider Foundation: Sabine Allou

→ Naia Science: Charles Ollion and Clément Leroux

→ Skills-based Volunteers: Antoine Girard and Maxime Lazennec

ACKNOWLEDGEMENTS

This project was made possible thanks to the invaluable contributions of numerous volunteers and dedicated teams. We extend our sincere thanks to the Microsoft AI School for their support through ShareAI, to the organisations l'École de l'IA, Naia Science, Heuritech, Neoxia, Metsys, Data4Good, and Archi-volt for their pro bono contributions, to the freelancers for their expertise and flexibility, and to all the financial partners who supported the Plastic Origins project. We are very grateful to the volunteers who so generously gave us their time and shared their expertise with us. Finally, thank you to project leaders and members of Surfrider Foundation's Environmental Expertise team, for your work carrying out this ambitious project.

PHOTOGRAPHY

Surfrider Foundation

ILLUSTRATIONS AND GRAPHIC DESIGN

→ kanardo.com

CONTACT

Sabine Allou, Surfrider Foundation's Plastic Origins project manager:
plasticorigins@surfrider.eu

LINK TO THE OPEN SOURCE PROJECT

https://github.com/surfriderfoundationeurope/surfnet

HOW TO CITE THIS REPORT

Allou S., Leroux C., and Ollion C. "Assessing the Use of AI to Quantify Plastic Pollution: A Surfrider Foundation Initiative From Technological Innovation to Action in the Field". 2024

***Cover picture** | Reporting waste stranded on riverbanks using the Plastic Origins app, © Surfrider Foundation.*

TABLE OF CONTENTS

FIGURES TABLE

SUMMARY

This technical report presents Surfrider Foundation's Plastic Origins project, which has explored the use of artificial intelligence to map and quantify plastic pollution in European rivers. Its aim is to empower citizens to contribute to data collection efforts by using a smartphone app to film riverbanks during kayaking excursions.

The AI model developed, named Surfnet, analyses these videos to detect macro-waste by combining computer vision technologies with video tracking. The use of AI enables better standardised and larger-scale data collection as compared to manual methods, which are often prone to observational errors. A training dataset of 5,000 labelled images was created with the help of volunteers, and techniques such as data augmentation were used to enhance performance. The project also highlights a frugal approach by using TinyML technologies to minimise both the financial and environmental costs of AI, particularly by reducing dependency on cloud-based operations.

While the project effectively raises awareness about plastic pollution among citizens and decision-makers, challenges remain, both in the field and technologically, to make it a true diagnostic tool.

Figure 1 *|* ***Left page*** *| Shot of plastic waste stranded in a river, © Surfrider Foundation.*

DEFINITIONS OF TERMS AND ACRONYMS

ACCUMULATION ZONES
Area of a watercourse characterised by a high concentration of litter, often due to slowing of the current, the presence of obstacles, or direct litter inputs.

AFNOR
Association Française de Normalisation. Organisation responsible for developing and promulgating standards in France.

AZURE
Cloud platform developed by Microsoft, offering a wide range of services (storage, calculation, databases, etc.).

CNN
Convolutional Neural Network. Type of artificial neural network particularly suited to computer vision tasks, such as object recognition within images.

COMPUTER VISION
Field of artificial intelligence which aims to enable computers to understand and interpret images and videos.

CPU AND GPU SERVERS
A central processing unit (CPU) is a component of a server that helps manage all the computing tasks necessary for the operation of the operating system and applications. A graphics processing unit (GPU) is a similar, but further optimised, hardware component offering high performance for AI applications.

CSV FILES
Comma-Separated Values. Plain text file format for storing tabular data.

DATA.GOUV
French government platform which provides open public data.

DEEP LEARNING
A sub-field of machine learning that uses multi-layer artificial neural networks to learn complex representations of data.

ETL
Extract, Transform, Load. Computer process of extracting, transforming, and loading data.

FINE-TUNING
The fine-tuning of a pre-trained model to adapt it to a specific task, optimising its performance without the need to undertake a new training session.

INFERENCE
Inference is the use of the trained model to make predictions about new data.

IOS AND ANDROID OPERATING SYSTEMS
Operating systems for Apple mobile devices and Android smartphones and tablets respectively.

LLM
Large Language Model. Large language model, a type of artificial intelligence capable of generating text, translating languages, etc.

MACRO-WASTE
Macro-waste are large pieces of litter (compared to micro-waste), generally measuring over 2.5 centimetres.

MSFD
The Marine Strategy Framework Directive is fundamental piece of European legislation designed to protect and conserve marine environments.

OSPAR
International Convention aimed at protecting the Northeast Atlantic's marine environment.

POSTGRESQL
Open Source Relational Data Management System.

SURFNET
Name of the artificial intelligence model developed by Surfrider specifically to count litter on river banks using video input.

TINYML
A sub-field of machine learning that focuses on developing AI models small enough to run on low-power devices.

TRACKING
Tracking is the process of locating and following moving objects in a video sequence.

TRAINING
the training of an AI model involves feeding it with large amounts of data to allow it to learn to recognize patterns.

TRASHROULETTE
Online labelling platform for annotating images of waste and create the Surfnet training dataset.

1 INTRODUCTION

Every year, an estimated 20 million tonnes of waste enter the ocean, including 8 to 12 million tonnes of plastic waste (Galgani 2016). Currently, all marine and coastal ecosystems are threatened by plastic pollution, which affects not only species (through strangulation, ingestion, etc.) and the seabed (deterioration of the ocean floor), but also human beings (health and socio-economic impacts). 80% of this plastic waste is thought to originate from land-based activities, primarily transported by rivers (GESAMP 1990). As they cross numerous agricultural and industrial areas, as well as urban agglomerations, rivers carry a multitude of elements that will subsequently be found in the sea

Once in rivers, the fate of plastic litter is highly variable, depending on the type of plastic, how it degrades, and the hydromorphological characteristics of the waterway. Part of it will transit directly through the estuary, whereas the remainder will accumulate in retention zones. Plastic is thus trapped or temporarily retained, sometimes for extended periods, by riverbanks and vegetation, before being released during heavy rainfall or flooding events. These prolonged residence times of plastic waste in rivers exacerbate the negative effects on aquatic environments, such as the formation of microplastics through fragmentation.

From 2014 to 2018, the Riverine Input project by the NGO Surfrider Foundation (Bruge 2018) reported precise identification of the typology and quantity of waste stranded at various study sites. Several methods were trialled, including collection and characterisation of litter stranded on riverbanks, counting of floating debris from bridges, deployment of nets and floating barriers to capture waste, as well as waste geolocation based on observations and surveys.

Following these experiments, protocols for collecting data on macro-waste trapped on riverbanks appeared to be a promising means of gaining insight into river pollution (Emmerik 2020). Indeed, macro-waste is a visible and tangible form of pollution, unlike microplastics or debris that is floating, suspended, or deposited on the riverbed. Therefore, it affords the possibility of collecting concrete data on the quantities and composition of waste present in these environments. Nevertheless, as of 2019, there was no internationally recognised protocol to guide the implementation of monitoring activities (González 2017).

This led to our decision to develop an innovative protocol to study the accumulation of plastics on riverbanks, incorporating three fundamental principles. Firstly, that it is easiest to detect waste from the water, where visibility is best, as the observer is at the same level as the waste and the view is not obstructed by vegetation. Secondly, as waste does not accumulate uniformly along waterways, the protocol had to cover large distances allows for a better understanding of its distribution. Finally, the protocol had to leverage the ubiquity of smartphones and be designed for simplicity and accessibility. As part of the protocol, the use of smartphone app would streamline data collection and encourages active citizen participation.

__Figure 2__ | __Left page__| Plastic bottles trapped in riverine vegetation, © Surfrider Foundation.

Furthermore, the integration of artificial intelligence (AI) could enhance this protocol by facilitating the acquisition of large quantities of data over extended distances. With the use of AI-based technologies, it would be possible to automate the detection and classification of waste, as well as improve the efficiency and accuracy of data collection and analysis.

Would it be possible to create a smartphone app that uses AI to facilitate and improve the accuracy of collecting data on riverine waste?

Surfrider then initiated development of the Plastic Origins app (Lepâtre 2023). By collecting data along waterways from a moving platform, the protocol contributes to improved understanding of the scale of plastic pollution over long sections of waterways. The objective is to gather evidence of diffuse pollution, which remains challenging to quantify precisely due to methodological limitations.

Practically speaking, users can either manually report observed waste directly on their smartphones or record a video of the riverbank for processing by an AI model. This use of artificial intelligence helps eliminate 'observer bias.' Each user introduces a unique error level based on their individual experience, knowledge, skills, and level of fatigue. The advantage of using detection instead of direct observation is that it replaces this variable error level (from the observers) with a consistent and known error (that of the detection algorithm), improving the comparability of datasets. AI therefore facilitates the generation of robust and reliable data in a citizen science project.

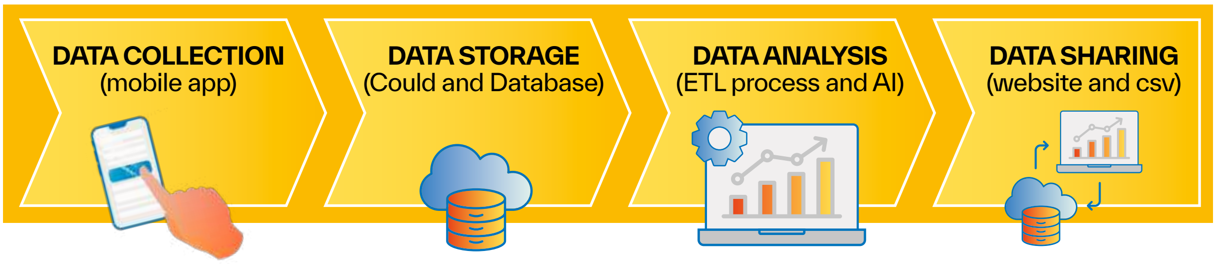
The digital tools (websites, databases, artificial intelligence, etc.) used in the Plastic Origins project were largely developed through the help of volunteers. This significantly reduced the project's overall cost. Without their invaluable support, this project would not have come to fruition. Therefore, sharing our results is not only intended to disseminate our technological expertise but also to acknowledge and celebrate the efforts of the volunteers who made it possible.

The purpose of this white paper is to share the technological advancements in artificial intelligence achieved by Surfrider. In the following, we seek to provide a clear explanation of the technologies used, how they work, the underlying assumptions and strategic choices, as well as their scope and associated limitations. We will also review the results achieved thus far and future perspectives, and in doing so, provide a comprehensive overview of the impact of these developments on our mission and our commitment to preserving aquatic environments.

Figure 3 | ***Above*** | *Sorting table for litter collected by the Riverine Input project, © Lucie Francini.*

2 The Plastic Origins Project

Plastic Origins is a digital tool designed to collect, store, analyse, and share information concerning plastic pollution in rivers. It comprises multiple software components that work in concert.



2.1 DATA COLLECTION

The app enables users to collect of data about plastic waste: the means of transport used for the study, the riverbank being studied, and the type of waste observed. Two modes of data entry are available. In manual entry mode, GPS coordinates from the phone are associated with the observation to generate geolocation data of identified waste. The date and time of data collection are also recorded. In automatic mode, the app allows users to record video with the phone's camera while simultaneously collecting GPS positions.

__Figure 4__ | __Top__ | Waste collection by kayak, © Surfrider Foundation. __Figure 5__ | __Above__ | Simplified diagram of the Plastic Origins digital platform, © Surfrider Foundation.

2.2 DATA STORAGE

he data collected by the Plastic Origins app is stored in two main locations:
— Raw files, including text data, photos, and videos, are stored in the Azure cloud.
— Processed and analysed data is stored in a structured PostgreSQL database.

This approach ensures secure backup and easy access to data from anywhere.

2.3 DATA ANALYSIS

An ETL process (Extract, Transform, Load) is a series of operations that essentially transforms raw data into a usable format for analysis.

Initially, AI object detections are linked to data from automated campaigns and added to manual detections. Next, the GPS data collected by the app is merged with various waste data to create a precise spatial representation of their location. Finally, all this data is cross-referenced with data from a reference 'River' table to identify the rivers where the monitoring took place.[1]

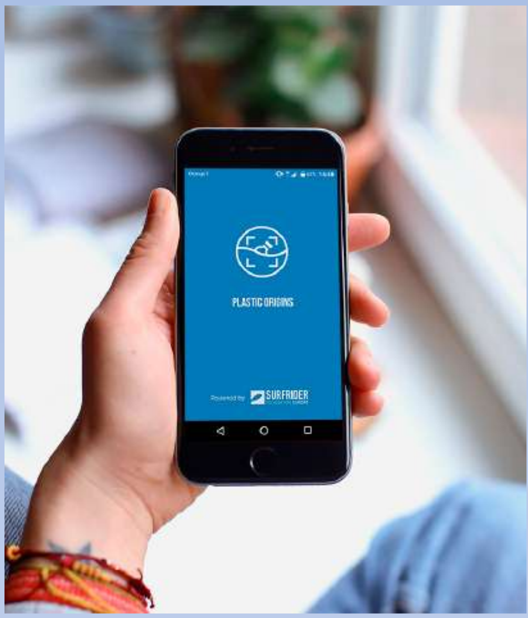
2.4 DATA SHARING

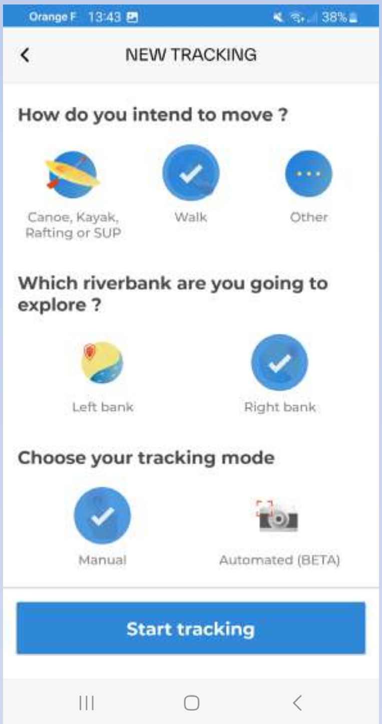
Together, this information allows the construction of a spatiotemporal representation of the presence of litter in rivers. Sharing this data is essential to promote transparency, collaboration, and action against plastic pollution.

An interactive map is available on the Plastic Origins website displaying the location of litter and providing access additional information like the date of the observation and the type of waste it is.[2] Cette carte permet de visualiser l'emplacement des déchets signalés et d'accéder à des informations supplémentaires telles que la date d'observation et le type de déchet.

A data request form is also available through the website, and annual CSV format files are published online via the data.gouv website. Researchers or other interested parties can thus easily access and analyse the data.[3]

Notes | *1.* *[European catchments and Rivers network system]* *(Ecrins)* | *2.* *[www.plasticorigins.eu]* | *3.* *[www.data.gouv.fr]*





The app is free on iOS and Android and is available in 4 languages: *French, English, Spanish, and Italian. Simply create a user profile to access the two acquisition modes available: a manual entry mode in which the user reports the presence of plastic waste, and an automatic mode in which a video recorded by the user is then processed by an Artificial Intelligence.*

3 The AI Model at the Core of the Project

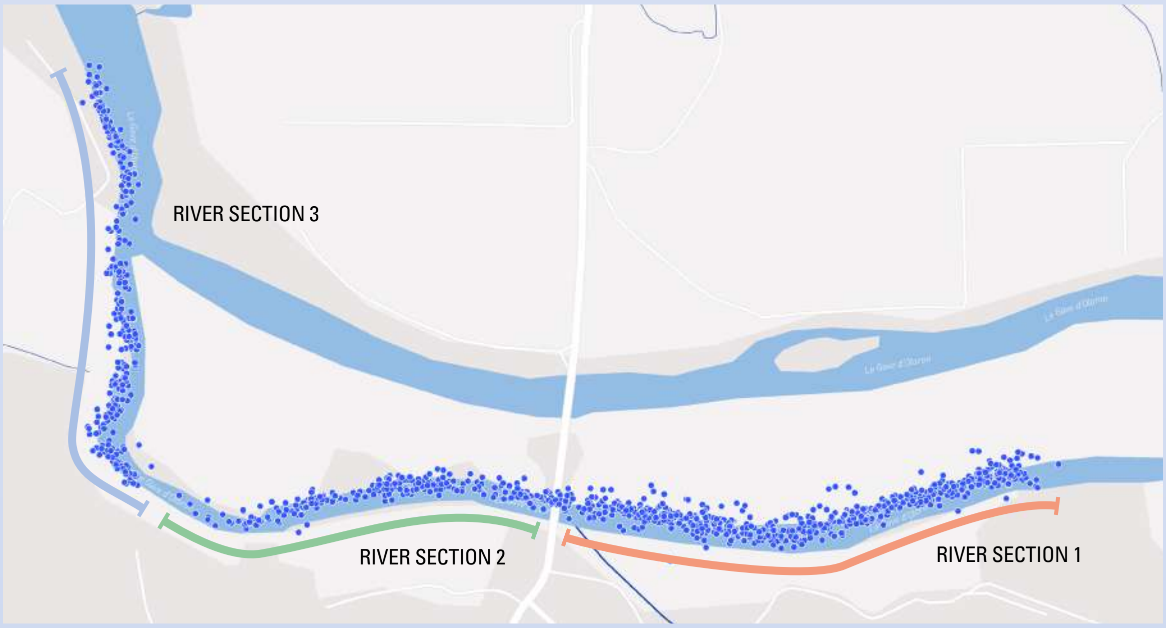
Over the past few years, the progress of Artificial Intelligence has impacted every sector of the economy. Like elsewhere, the fields of ecology and environmental sciences have seen AI-driven initiatives develop.

3.1 AI AS A QUANTITATIVE AND STANDARDISED MEASUREMENT TOOL

We postulate that AI usage can also have a positive impact in these sectors. Specifically, automated image analysis technologies can provide useful indicators regarding natural resources usage or degradation. The analysis of images can, for example, help measure air pollution, quantify the scarcity or misuse of drinking water, or identify changes in coastal landscapes. It thus becomes feasible to rely on these technologies to detect litter and pollutants.[4 5 6] In our case at Surfrider Foundation, we faced the challenge of large-scale data collection in our attempts to study the presence of plastic in rivers. In France alone, rivers cover approximately 623,464 km (Office français pour la biodiversité n.d.),

Figure 6 | Above | Plastic Origins app used on a kayak trip, © Surfrider Foundation.

Notes | 4. Projet Bergson de l'Agence Spatiale Européenne | 5. Projet Foncier innovant de Bercy, Google et Capgemini | 6. Projet DEA Coastlines de Geoscience Australia



The three segments exhibit specific characteristics that were considered in our analysis: *the first segment has a moderate current with dense but non-obstructive vegetation, the second has a strong current with low-lying vegetation that obstructs the view, and the third has a moderate current with tall, sparse vegetation, allowing for good visibility.*

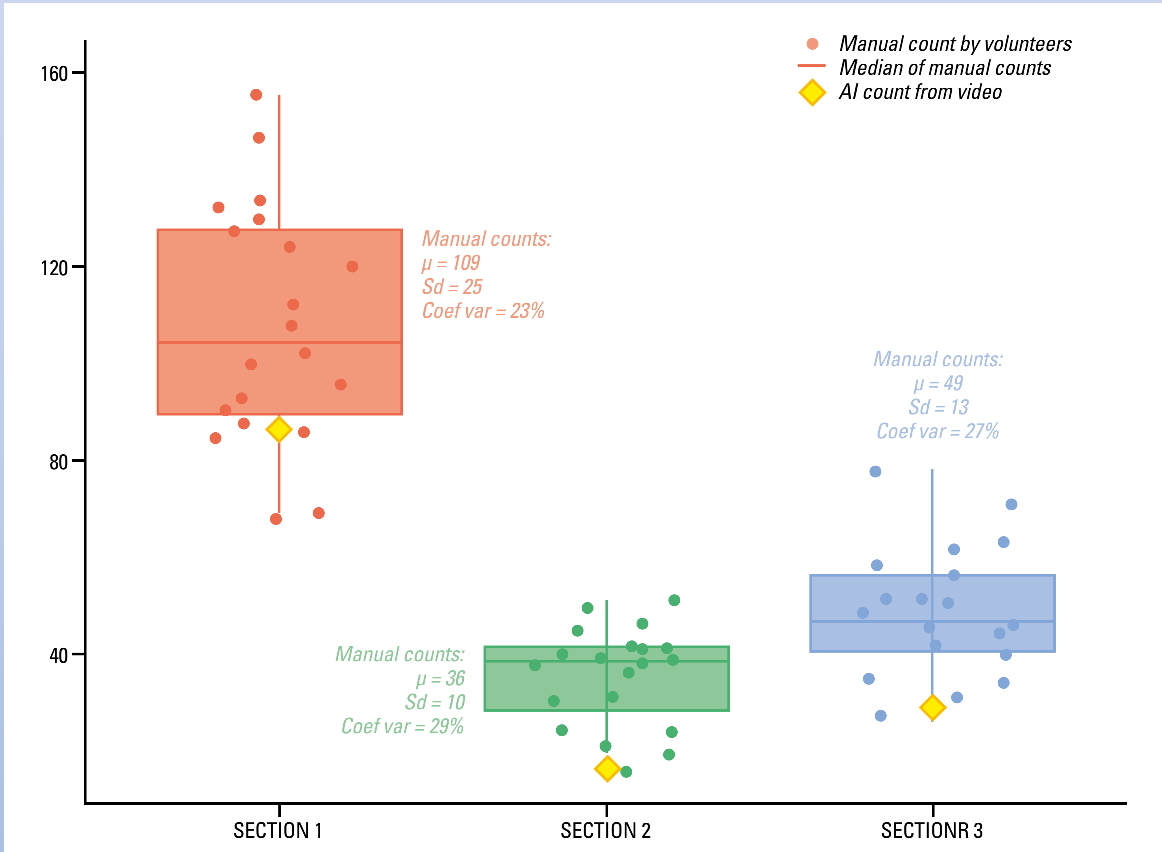


__Figure 7__ | __Top__ | Different sections of the Gave d'Oloron (64 - France) river studied | __Figure 8__ | __Above__| Box plot to compare manual counts by volunteers and AI counts from video (median, standard deviation and coefficient of variation). Adapted from Chagneux 2023.

making relying solely on manual counting methods unfeasible. While manual counting can be useful at small scales, it is prone to fatigue and identification errors, especially for inexperienced observers. Integrating AI into smartphones could serve as a powerful means to significantly enhance data collection and promote the democratisation of plastic waste detection in waterways.

Consequently, one of project's core ideas was to introduce Artificial Intelligence as a quantitative measurement tool usable by thousands of people collecting data under the Plastic Origins initiative.

Moreover, AI enables the standardisation of data acquisition by automating collection, analysis, and large-scale interpretation.

To assess the potential benefits of a standardised approach for our project, we conducted a field experiment on a stretch of river, in which manual counts from 20 pairs of volunteers were compared with the results of AI-based video analysis.

Manual counts conducted by the 20 pairs of volunteers showed significant variability across the three segments studied. In contrast, the AI consistently underestimated the number of items compared to human observers. However, it still followed the same general trends of detected pollution, with a higher number of objects found in T1 than in T2 and T3. This discrepancy may be attributed to challenges in detecting certain types of litter or limitations related to image quality.

The difference observed between the counting methods calls into question the reliability or complementarity of the two approaches. However, since the primary objective is not to obtain an exact count but rather to bring to light the pollution levels in the watercourses studied, we consider these results to be satisfactory. Even if the AI underestimates the total number of waste items, it still enables the standardisation of data acquisition over time. Further work could help refine the automated model to reduce this discrepancy.

> The use of artificial intelligence helps eliminate 'observer bias.' Each user introduces a unique error level based on their individual experience, knowledge, skills, and level of fatigue. The advantage of using detection instead of direct observation is that it replaces this variable error level (from the observers) with a consistent and known error (that of the detection algorithm). It enables citizen science initiatives to generate robust datasets.

3.2 A MINIATURE AI AMONG THE VAST WORLD OF AI

The recent dynamic development of artificial intelligence has been driven by a significant increase in the size of models and their datasets. Progress in the field surpassed a major milestone, particularly since 2023, with the emergence of generative AI. These 'super' models, often referred to as foundation models due to their vast capabilities, have become a dominant force in the AI technology ecosystem. However, they are now under scrutiny regarding the investment required for their deployment and their growing environmental impact.

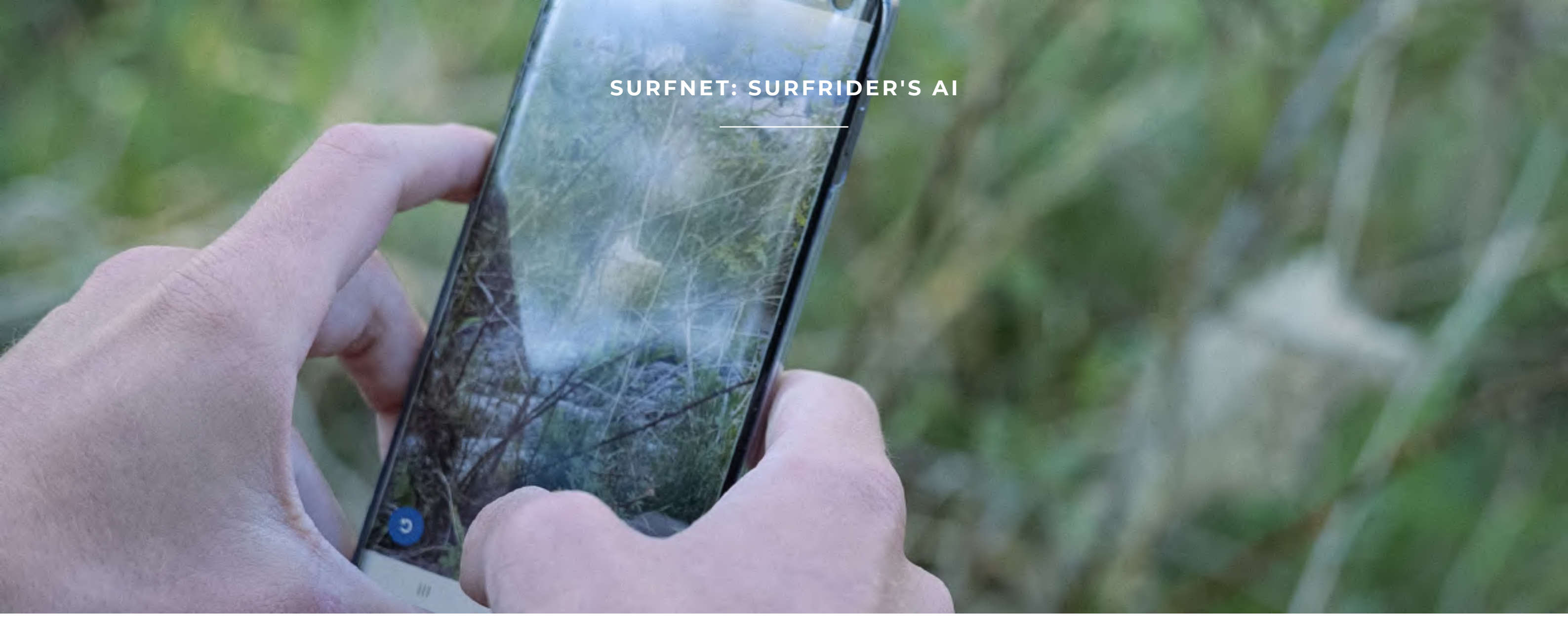
Well-known examples, such as OpenAI's ChatGPT 3.5 (with 175 billion parameters) and, more recently, Meta's Llama 3 (with 70 billion parameters), illustrate the complexity of these artificial intelligence models, which use hundreds of billions of parameters and trillions of data points to achieve their performance levels. These scales are reflected by the substantial infrastructure and investments required — as illustrated by the 24,000 graphical processing units (GPUs) and investments totalling around 100 million dollars needed to train Llama 3 (Meta 2024). These are far beyond the reach of most AI developers and are unsustainable in the long term, given the systems' energy demands. Furthermore, the annual energy consumption of ChatGPT is estimated to be equivalent to that of 30,000 households (Crawford 2024).

Alongside the rise of generative AI and large language models (LLMs), a parallel but less publicised effort is also advancing: the development of TinyML approaches, which aim to minimise model size to enable, among other things, the use of artificial intelligence on mobile phones (Ollion 2023).

To address these various challenges, we quickly turned to TinyML approaches, specifically testing models like MobileNet (with 2 million parameters) and miniaturised versions of Yolov8 (3 million parameters for Yolov8-nano and 11 million parameters for Yolov8-small), with development costs of a few hundred euros and marginally negligible operating costs for mobile devices. The possibility of remote operation is much higher, as there is no need for clusters of thousands of servers to either design or use this artificial intelligence. Finally, although the environmental footprint of these models remains to be calculated precisely, it is certain to be several orders of magnitude lower than current models.

Integrating AI into environmental projects must be accompanied by a comprehensive consideration of the technology's impacts. While the potential benefits are numerous, it is essential to prioritise sustainable and responsible solutions. TinyML, by offering a lighter and more energy-efficient alternative, aligns with this approach.

Figure 9* | *Above | *Volunteer filming waste on riverbanks from her kayak, © Surfrider Foundation.*



4 Surfnet: Surfrider's AI

The first step, essential for what follows, is to define the objective of the Artificial Intelligence model in a precise manner. Indeed, this will determine how data is collected (during both training and operational phases), the R&D effort necessary, and the type of model to be used. Moreover, clearly defining the objectives helps to communicate to all stakeholders involved in the design and use of the AI exactly what is intended to be achieved, thus avoiding misconceptions or unrealistic expectations regarding AI.

The objective of the Surfnet AI is to count visible macro-waste stranded on riverbanks from a kayak using video footage captured by an onboard device.

In the following sections of this document, 'Surfnet' refers to the AI engine developed by Surfrider: it is open-source, available for download, and its use is well-documented to ensure accessibility for any team interested in addressing solid waste-related issues.

4.1 FRAMEWORK FOR USE

To achieve Surfnet's objectives, we designed and tested multiple usage frameworks specifically, how users would interact with the application in practice.

The frameworks needed to address the following requirements:

— Collecting relevant data for Surfrider (enable identification of the relevant types of waste.

— Optimising engine performance (poor image capture, objects that are too small, or items that are too difficult to identify would compromise AI engine effectiveness).

— Provision of an ergonomic user interface to facilitate widespread adoption.

Figure 10 | ***Above*** | *Volunteer taking a picture of stranded waste, © Surfrider Foundation.*

This research represents one of the project's most crucial components, drawing upon expertise in domain knowledge, artificial intelligence, and application design, alongside extensive user testing and feedback.

The framework we ultimately selected comprises the following specifications:
— Users capture footage of the riverbank whilst on the river itself, typically from a kayak.
— The camera is directed perpendicular to the bank at a distance of 2 to 5 metres. Perspective shots with vanishing points are actively discouraged, as objects become unidentifiable.
— Kayak velocity must be maintained at approximately 4 km/h to ensure clear footage without motion blur. Certain sections of waterways must be excluded from analysis due to rapids or obstacles.

The (typically handheld) recording device requires stabilisation. This generally necessitates a two-person operation, with one person steering whilst the other records the riverbank. Although we considered phone stabilisers mounted to either the kayaker or the kayak itself, or the use of GoPro cameras, we determined that to maintain accessibility and simplicity for the highest amount of people, using footage captured via standard smartphone devices was the best option.

A critical consideration is providing tutorials for optimal image capture. To this aim, we trained several volunteers but also uploaded a tutorial directly onto the app, complete with a visual guide for identifying the riverbank.

It is important to note that, despite the stated constraints, we designed Surfnet to be fairly robust in addressing the various challenges of image capture in the field, including:

— Lighting, which varies considerably depending on the orientation of the bank, the position of the sun, and weather conditions
— Object size, which changes based on perspective and distance from the bank
— Object shape, which varies widely due to the nature of plastic waste
— Partial occlusions, as objects may be stuck on the riverbank and in its vegetation
— Numerous other parameters are subject to variation such as blur, humidity, the angle the device is being held, and more.

Figure 11 | Above | Screenshot of the visual guide on the application | Figure 12 | Top | Volunteers filming litter on the riverbanks from their kayak, © Surfrider Foundation.

4.2 TECHNOLOGICAL CHOICES

THE AI ENGINE

With this framework for use and video data capture conditions established, several decisions had to be made regarding actual engine of Surfnet. The first consideration was the method for counting waste. We opted for a standard, state-of-the-art approach, performed in the manner described in detail in subsequent sections:
— Segmenting the video into successive frames
— Automatically detecting each piece of litter within each frame.
— Tracking the litter across successive frames
— Counting the tracked pieces of waste.

THE DETECTION METHOD

The second step is critical as it is where the system identifies and locates litter within an image. We opted to use a type of artificial intelligence called computer vision — a branch of AI aimed at enabling computers to detect shapes, objects, and patterns in images. In our case, these objects are pieces of litter, of which there may be multiple simultaneously, figuring in images of riverbanks. Technically, the approach involves using an artificial neural network — in our case, a Convolutional Neural Network (CNN) — the current state-of-the-art method for object detection.

There are two primary methodological classifications at play in neural networks capable of object detection:
— Object Detection, which identifies each object and provides the coordinates of a bounding box around the object.
— Object Segmentation, which, in addition to detection, precisely outlines the object.

In principle, for counting purposes, only the first method is required. Consequently, we chose to

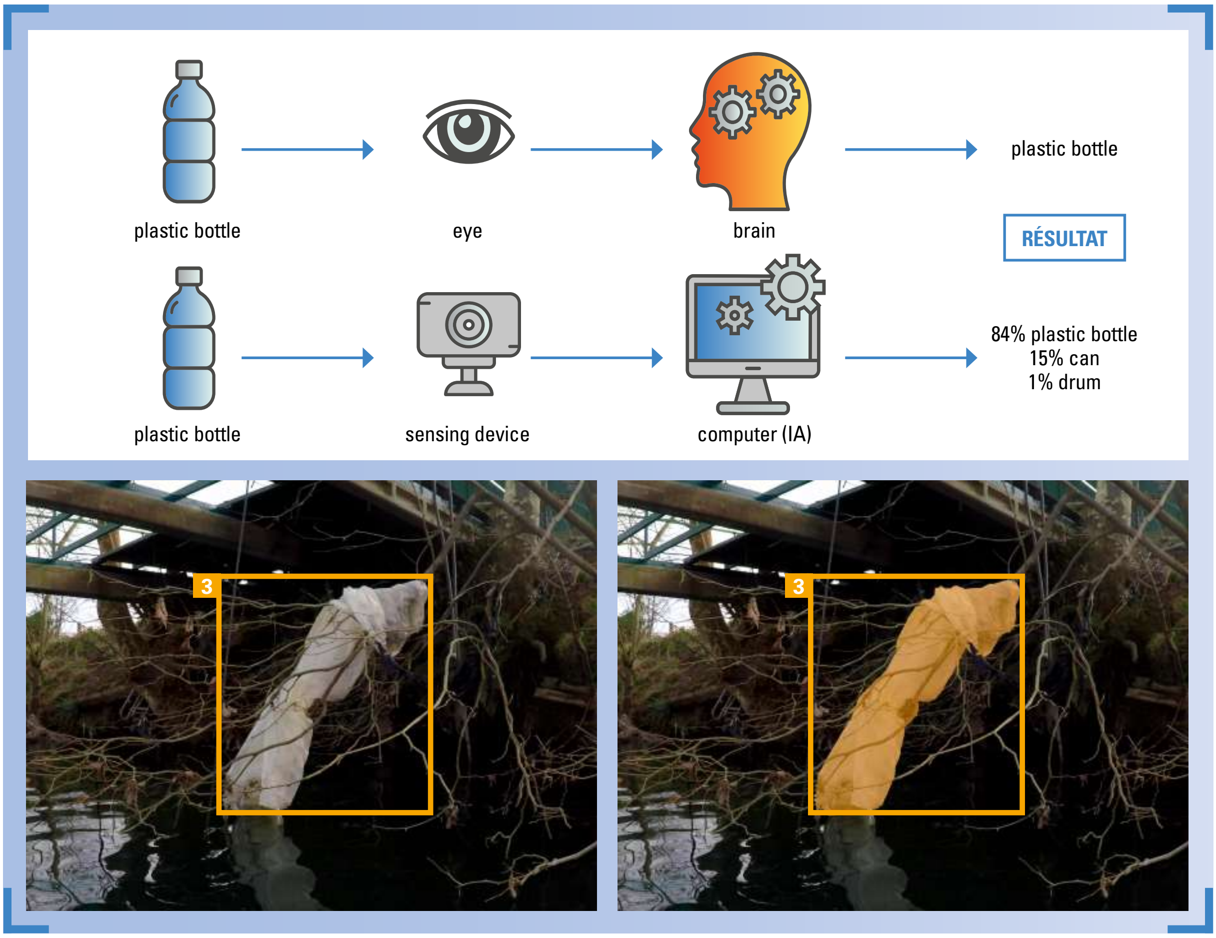


Figure 13 | Top | Diagram comparing computer vision, the branch of AI chosen in the project, to human vision. | Figure 14| Above | Comparison between object detection (left) and segmentation (right) train Surfnet to detect objects using the dataset described in the next section.

MODEL CHOICE
Currently, there are numerous neural network architectures for object detection. Throughout the development of the project, we evaluated various models and model versions (ResNet, centerNet, YOLOv5, YOLOv8) and iterated through multiple data and training strategies to optimise performance. The key considerations we prioritised were:

— Model size (i.e., the quantity of parameters within the model): the larger the model, the higher the training and computational resources.
— Open-source community engagement and model portability: we targeted open-source models with active communities to ensure developmental longevity and avoid obsolescence or maintenance issues.
— Model efficacy in object detection, particularly concerning small objects and performance with limited dataset sizes. It is worth noting that contemporary computer vision neural networks are 'pre-trained', which allowed us to leverage existing generic training before fine-tuning with our specific dataset, thereby significantly enhancing performance metrics.

Over the course of our experimental work, we assembled a comprehensive dataset and subsequently trained the neural networks to recognise litter within this dataset. The following sections elaborate upon the decisions made and the resultant outcomes of our process of development.

4.3. A UNIQUE TRAINING DATASET

The performance of AI models is heavily dependent upon both the quality and quantity of their training data. Indeed, the dataset serves as the raw material that the machine's learning is based on. In our case, the data consists of a collection of annotated photos with information about the litter visible in each image. In this way, the dataset provides the AI with essential examples for learning to recognise and analyse various forms of waste observable on riverbanks.

Once the training phase is complete, the **trained model** is set to analyse new images extracted from video footage. This constitutes the **inference phase**, during which the AI detects objects through predictive analysis.

To optimise our AI model's performance during the inference phase, the training dataset must accurately represent the reality of riverbanks to be recorded by the Plastic Origins app. Generally, models trained under conditions closer to reality demonstrate better capability in waste detection, even when confronted with footage containing high amounts of background variation. In practical terms, the training images needed to depict waste across entire stretches of riverbank, comprising the

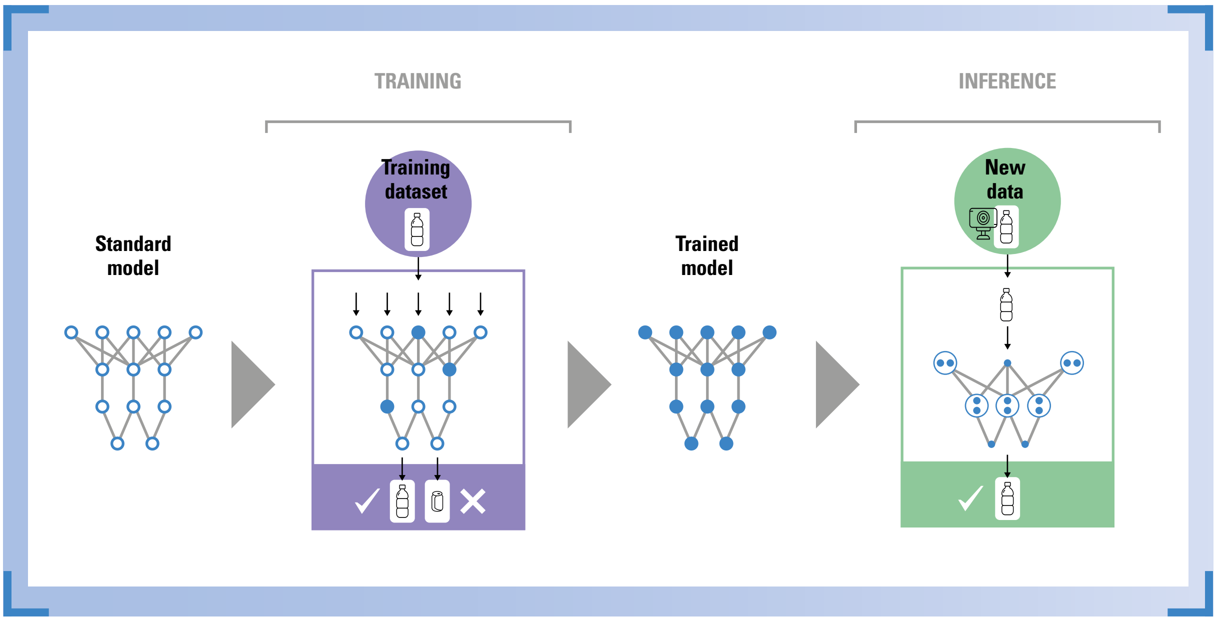


Figure 15 | Above | Diagram of the training and inference phases of an AI model.

During inference, the trained AI effectively says: "Based on my previous observations, there is a high probability that the object at this specific image location is a plastic bottle."

morphological complexities and diversity of the surrounding vegetation.

However, no existing dataset suited our specific context of interest. The most well-known existing waste dataset, TACO[7], contains numerous images of waste, albeit captured at relatively close distances, often in urban or domestic environments.

Therefore, we compiled our own database to train Surfnet AI. With the support of extensive volunteer efforts in the field, we compiled thousands of photographs documenting waste materials left on riverbanks. These images capture waste in real-world conditions, covering a broad range of backgrounds, lighting conditions, viewing angles, and image quality.

Once all the images were collected, one crucial step remained: data labelling. Each image had to be manually annotated, meaning that each piece of waste had to be enclosed within a bounding box to mark its precise location on the photo, and each item of waste needed to be assigned a category *(categories are described in Section 4.4).*

This data labelling process is time-consuming but essential. In 2019, unlike today, no free, open-source annotation tools were available. Therefore, we created an online labelling platform: www.Trashroulette.com. With the support of a detailed tutorial, hundreds of people annotated images and contributed to the project over the internet.

The creation of this labelling platform allowed us to build a dataset of approximately 5,000 images with a total of 8,000 annotations. While this is small compared to the massive datasets used by companies like Google, it grew much larger than TACO in terms of volume. Moreover, our data is highly specific but also representative of a range of natural environments.

The dataset thus provided a solid foundation for Surfnet's initial training. However, it still does not allow us to achieve sufficient performance levels. This suggests that with additional annotated images, our model's performance could be significantly enhanced *(several dataset enhancement techniques are described in Section 4.5).*

The dataset is available as open source through the following link: universe.roboflow.com/surfrider-foundation-europe/plastic-origins

Please note that this version only contains roughly 4,300 images, as some were filtered out before publication for quality reasons.

4.4. THE DIFFERENT CATEGORIES OF WASTE TO DETECT

To accurately assess riverine plastic pollution, the development of a system to categorise the types

***Figure 16** | **Above** | Variety of images of litter in rivers, © Surfrider Foundation.*

***Notes** | **7.** tacodataset.org*

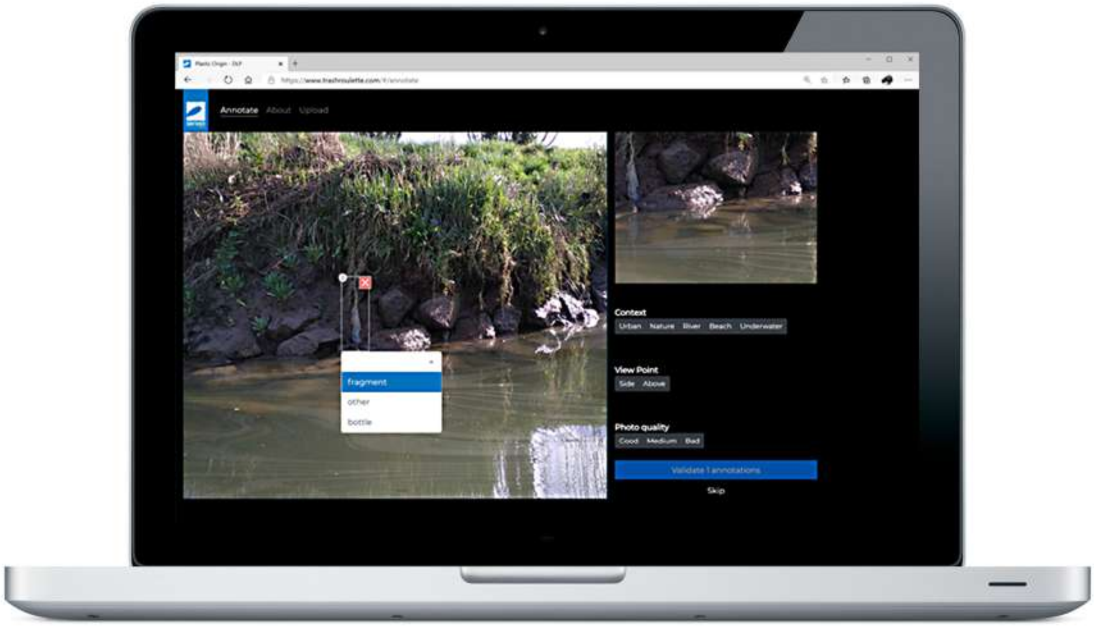


*Figure 17 | **Above** | View of the Trashroulette labelling platform.*

of waste materials observed on riverbanks is essential. There are some existing classification systems developed by experts, such as OSPAR/MSFD systems, which provide the standardised lists used in marine waste monitoring (Hanke 2013). These classification systems encompass diverse categories based on material type (plastic, metal, glass, etc.) and specific subcategories describing objects (plastic bottles, fishing nets, cans, etc.). They are primarily based on semantic characteristics and designed to categorise collected waste, enabling physical manipulation for precise typological identification.

In field conditions, people can readily identify waste objects through their comprehension and familiarity with the materials. For instance, an individual will recognise a crushed plastic bottle as such, despite its altered shape, because they can make abstractions of variations in the size, shape and colour of objects. The AI, however, relies exclusively on visual characteristics for object detection. When an object is deformed, dirty, partially hidden, or in an advanced state of decay, the AI may encounter difficulties classifying it accurately.

Working alongside waste management experts, we developed a custom macro-waste classification system that balances ease of annotation and practicality in a computer vision capacity (Chagneux 2023). This classification comprises ten categories organised into three groups:
es, cans, drums, and tyres, which are frequently found in rivers and easily recognisable by their defined shapes.
— Fragmented items, often found along riverbanks but displaying more varied appearances.
— Two additional categories for waste items that are either clearly identifiable but do not fit the previous categories or are unidentifiable.

These ten categories were used during the image labelling process in the creation of the model's training dataset *(Appendix 1).*

While they accurately represent the diversity of waste typically found on riverbanks, the distribution of labels per category is uneven. Some categories have significantly more annotations than others, such as the "Sheet, tarp, fragment" one, which accounts for half of the annotations.

4.5. MULTIPLE APPROACHES TO IMPROVE THE DATASET

To address both our relatively small dataset and the imbalance in image counts across categories, we explored several techniques to artificially increase the overall number of images and the number of images per waste category:

1. Data Augmentation
2. Synthetic Data Generation
3. Foundation Models for Annotation Assistance

1. Data augmentation is a technique used to artificially expand the size of our small dataset by applying geometric and colour-based transformations to images. Although these transformations alter the image, they must maintain the object's recognisable appearance: objects must not be obs-

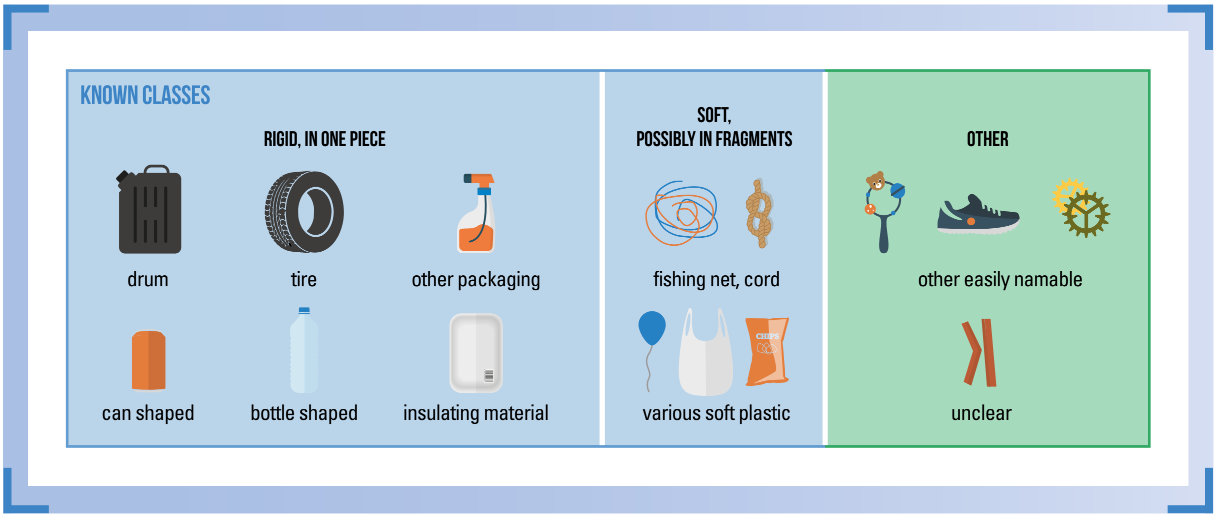
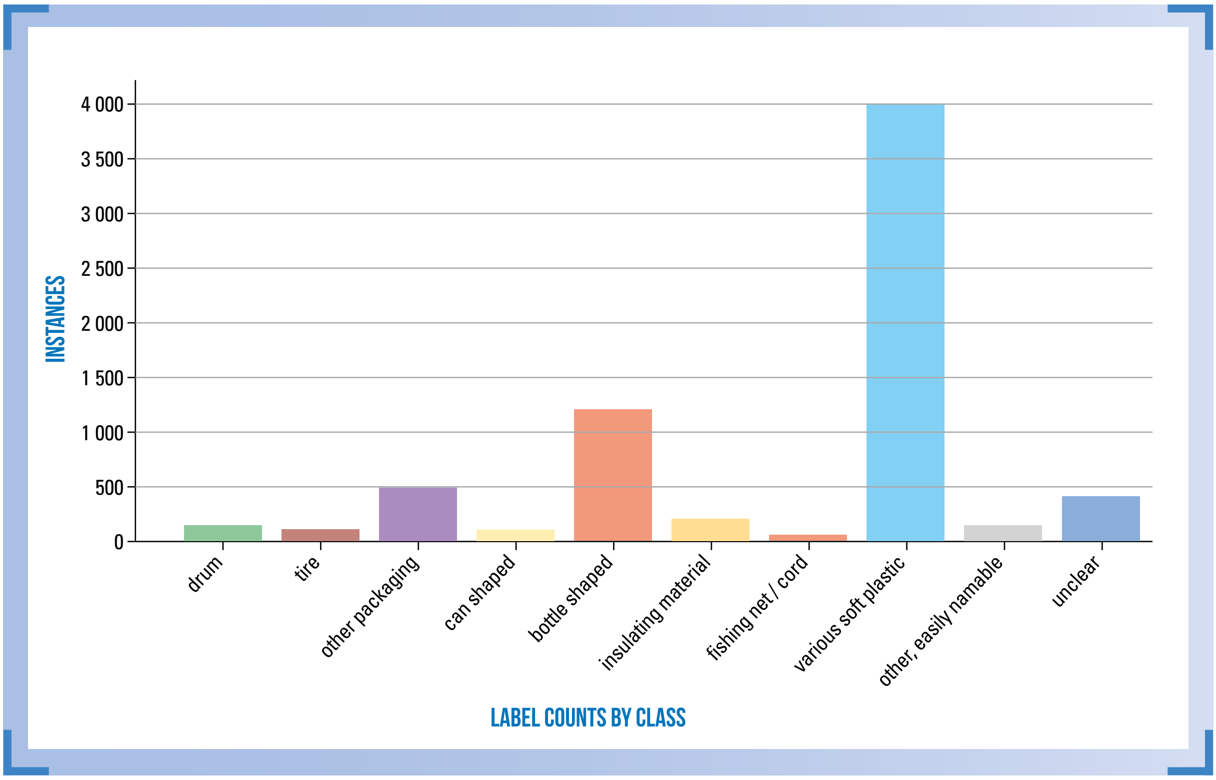


Figure 18 | Above | The ten categories of litter used to train Surfnet. | Figure 19 | Below | Distribution of the number of labels per category. 8,000 labels were annotated on the 5,000 images of the dataset.



cured or overly distorted to remain detectable. Common augmentations include flipping the image (horizontally in our case), adjusting colour or contrast, distorting the image, and applying blur filters. This technique not only introduces artificial diversity into the images, effectively increasing the number of training images, but it also improves the system's robustness against variations caused by common disturbances like blur, contrast shifts, or lighting changes, which frequently occur due to camera movement and positioning relative to the sun. With carefully selected augmentations, we achieved approximately a 20% improvement in final performance.

2. Synthetic data generation is a technique that creates training images artificially. Less commonly used than the previous method, it was tested in the context of Plastic Origins as follows:

— Extract images of riverbanks without waste from footage of the riverbank.
— Using a segmented waste object database from TACO, cut out these objects and paste them on the riverbank image in a random location.http://taco-dataset.org/
— Ensure visual consistency by adding blending effects, positioning the object correctly on the bank, and other visual enhancements.
— Automatically generate the corresponding annotation based on the placement of the object.

This method can artificially produce thousands of images, provided we have a sufficient supply of background images (clean riverbanks) and waste objects to insert (segmented items). It is essential to differentiate these synthetic images from real annotated ones: they are neither authentic nor flawless, and should solely be used as training aids, while the quality of models must be evaluated exclusively on real data. Our 2023 tests showed a modest improvement of around 5% when synthetic data was added. However, the effort required to produce realistic results was significant, sometimes

Figure 20 | Above | Multiple images resulting from the increase in data: rotations, mirroring, changes of contrast, brightness, colorimetry, etc.

Figure 21 | *Above* | *Example of synthetic data generation – waste item cut out and inserted onto a riverbank image, © Surfrider Foundation.*

demanding more time to calibrate the synthetic generation than to simply annotate new images. As a result, this approach has been set aside for now. An example is shown next page.

3. Lastly, foundation models for annotation assistance represent a more recent approach, enabled by the emergence of large, general-purpose models such as SAM[8], OwlVit[9], and DINO[10]. These large models (100 to 1,000 times the size of the pre-trained vision models mentioned earlier) are far more agnostic to data and task, making them useful for improving our training data. In the Plastic Origins project, we principally used SAM ("Segment Anything"), a model capable of outlining any type of object. Since the annotations in our database were primarily bounding boxes around waste items, they lacked the precision of SAM's segmentations. However, by applying SAM to each bounding box, it was possible to obtain a precise outline of each object. This process requires a large model like SAM, but it only needs to be applied once per training image and thus requires little resources. The major caveat is that it would not be feasible to use SAM in production (i.e., for each new video) due to its computational intensity. Using more precise training data (segmentation instead of bounding boxes) improved our lightweight object detection model's performance by about 10%. An example of automatically segmented annotated data is shown below.

4.6. METRICS AND DETECTION RESULTS IN STILL IMAGES

METRICS: INDICATORS OF MODEL PERFORMANCE

By utilising the annotated dataset, it is possible to

In resource-limited situations, with limited computational power and a relatively small dataset like ours, leveraging pre-trained vision models, data augmentation, and foundation models proves simple to implement and yields excellent results. However, synthetic data generation, while potentially beneficial, also requires more consequential and objective-specific investments.

Notes | ***8.*** *segment-anything.com* | ***9.*** *console.cloud.google.com/vertex-ai/publishers/google/model-garden/owlvit-base-patch32?pli=1* | ***10.*** *ai.meta.com/blog/dino-v2-computer-vision-self-supervised-learning* train the computer vision model to detect and localise waste materials in images that are relatively similar to those in the dataset. However, artificial intelligence models are not infallible and exhibit a certain degree of variability in their performance. They are generally judged according to several criteria, including:

— **Recall:** The extent to which the model comprehensively detects existing waste (i.e., minimising missed detections).
— **Precision:** The degree to which the model makes accurate detections versus errors (e.g., mistaking a rock for waste).

Training quality is often assessed by combining these two metrics into an overall measure known as the **F1 score**, which represents a kind of average.
There are numerous other metrics, such as those related to the precision of the predicted bounding box coordinates compared to reality in the field. Rapidly defining the pertinent metrics for a given application is crucial, as each use case has its own unique specifications. In our case, we made a distinction between two types of metrics: those associated with object detection in still images, and those related to enumeration in video footage (the ultimately most important business metric). The rest of this section will focus on detection metrics, while the subsequent section (4.5.) will address counting metrics.

In our context of counting objects, the primary priorities were to avoid missing objects (i.e., achieve high recall) and minimise erroneous detections (i.e., maintain high precision), with less emphasis on precise localisation of objects (tolerating some bounding box coordinate errors). Furthermore, we were interested in obtaining both a general count of waste objects and a per-category breakdown. Therefore, we report two scores: a **multiclass score** considering the detected object's category (counted as correct only if the annotation matches), and a **single-class score** focused uniquely on correctly identifying the presence of an object, irrespective of its type.

THE RESULTS

Given our preference for lightweight models, we opted for a compact YOLOv8-s (small) architecture, comprising 11 million parameters - offering the right balance of performance and speed of execution.

It is important to interpret these results, as scores in isolation provide limited informative value.

In all the images in our dataset, the model detected an average of 60% of the waste materials. Detailed analysis reveals that the smallest objects were often missed, primarily due to limited input image resolution and the distance of the camera from the riverbanks. The model also regularly gave false positives (69% precision, implying 31% of predictions were incorrect). Our analysis shows this often was duplicate detections of amorphous objects (e.g., a plastic tarpaulin detected as two separate pieces with a single annotation). Additionally, high concentrations of white rocks on the riverbanks were sometimes misinterpreted as pieces of plastic.

Finally, multiclass performance was significantly lower than single-class. An object detected in an incorrect category reduces both precision (as it is considered a false positive) and recall (as the original object is now considered undetected). This indicates confusion between certain categories, typically with "Food packaging" logically overlapping with "Sheet, tarp, fragment". Similarly, "Easily namable objects" and "Unclear objects" exhibited comparable issues. These results suggest that the three least well-defined categories from an annotation perspective were also generating the most model confusion, which is an expected outcome. Updating the category definitions or improving labelling could help mitigate this problem.

METRIC	F1 SCORE	RECALL	PRECISION
Single-class	0.63	60%	69%
Multiclass	0.42	39%	46%

Figure 22 | Above | Results obtained on Yolo-v8 after several iterations. F1 score, recall and precision.

Figure 23 | ***Above*** | *Example of image with error: the largest piece of plastic is detected, but the small one to its left is not.*

4.7. TRACKING WASTE IN VIDEO FOOTAGE

After detailing our AI module's capacity for object recognition in still images, we add the complexity of following objects through the successive frames of a video - a process referred to as "Tracking". Surfnet must not only detect objects in each image but also trace their movement over time to maintain continuity in object tracking and avoid repeated counts of the same item. Doing so is no easy task with the simultaneous occurrence of false positives, false negatives, and a moving camera (and therefore moving objects) between each frame.
Object tracking in complex environments with a mobile camera is a challenging problem that requires substantial R&D effort. We partnered with a research laboratory to launch a CIFRE (French Business and Academic Partnership) PhD thesis focused on advancing this subject. This collaboration resulted in a publication in a scientific journal (Chagneux 2023) and the development of an open-source tracking engine.[11]

Here is a simplified explanation of the steps involved in tracking:

1. Initial detection
Wastes are detected in an image by our Surfnet AI.

2. Motion Prediction
The challenge is then to track these waste items over time. In other words, if we detect a piece of waste in one image, it needs to be associated with

These results may appear disappointing at first glance. This perception can complicate communication with involved teams, as they typically expect a 99% precision level to feel confident in the system. However, these performance levels are often sufficient, as computer vision is just one component of the overall system, and object tracking can recover a significant number of errors, ultimately providing much better accuracy in the overall count.

Notes | ***11.*** *github.com/surfriderfoundationeurope/surfnet* a detection in the following frame. For each detection, a statistical algorithm attempts to predict where it will be in the next image. It's as if we were drawing a circle around each item, saying, "It should be somewhere around here".

3. Search and Association
Then, in the next image, the algorithm searches within each prediction zone:

— If an item is found in the zone, it is likely the same as in the previous image
— If nothing is found, the search circle expands
— If a new item appears elsewhere, a new tracking instance begins

4. Uncertainty Management
The algorithm also uses probabilities to handle uncertain cases, such as when an item temporarily disappears, or multiple items are close together.

Below is an example where waste items are detected and then tracked through successive images:

— The first detected item is tracked by the red circle. The red circle moves but maintains the same size, indicating successful tracking from one frame to the next.
— For the second item, represented in brown, the brown circle expands to illustrate the widened search area, as the item is not found in the second and third images.
— A pink circle appears in the third image, indicating the detection of a third waste item.

There are several metrics for evaluating tracking systems, as described in our publication. In our case, the emphasis was on the final tally of waste items, i.e., the number of waste items that were tracked for more than 5 consecutive frames (the best compromise based on various tests).

To illustrate, we will use the example of our study area of 3 sections of the Gave d'Oloron River, where we also compared human counts versus AI counts *(Paragraph 3.1.)*.

In these same sections, the AI's counts were compared to those of a person reviewing the same footage.

The current AI system counts approximately 2/3 of the waste observed by a human viewer. This performance varied significantly across sections, and is particularly low in section T2, where the current is strongest and the vegetation is lowest, hindering the view of the camera.

These results have room for improvement, not only through training a better model, but also by a better acquisition protocol (to ensure volunteers provide more stable data).

TYPE	SECTION 1	SECTION 2	SECTION 3
Human count from video review	133	26	28
AI count from the same video	86	11	25
AI detection	65%	42%	89%

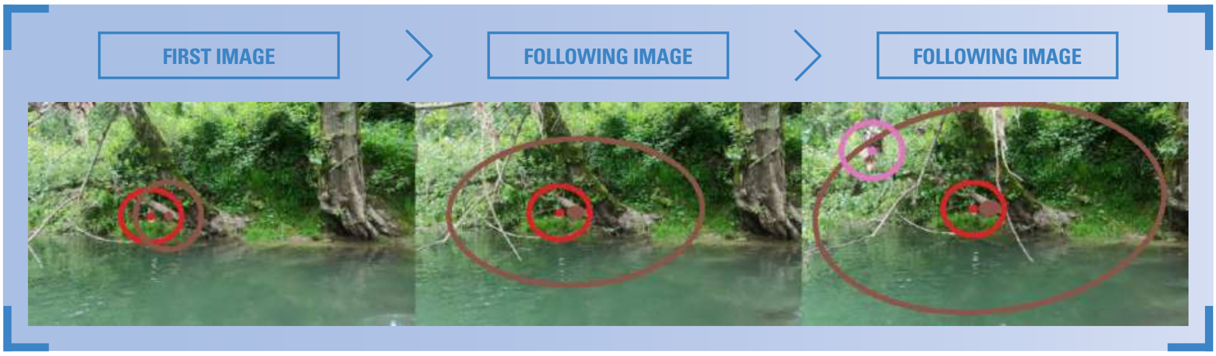


***Figure 24** | **above** | Tracking method through successive images. **Figure 25** | **Top** | Comparison of AI counts with those of a person watching the same video on 3 different river sections.*

Figure 26 *| **Above** | Waste floating just below the water's surface, © Surfrider Foundation.*

5 Discussions

The increasing use of artificial intelligence in citizen science projects is changing the way we think about tackling plastic pollution. Here we look at the opportunities and limitations of this approach in the context of Plastic Origins.

5.1 CITIZEN SCIENCE AND ARTIFICIAL INTELLIGENCE: SYNERGY IN THE FIELD?

Studying plastic pollution in an environment like a river is not an easy feat. The inaccessibility of riverbanks, dense vegetation, currents, and the fragmentation of plastics make the task particularly complex. Recent technological developments like remote sensing are often unavailable due to their cost or lack of applicability in these environments. Here, citizen science offers an innovative alternative. Citizens can contribute to data collection during recreational activities like hiking or canoeing. This collaborative work targets people who are already environmentally conscious, but gets people involved more tangibly in the fight against plastic pollution.

By making AI tools more broadly available, we wanted to simplify and standardise data collection by volunteers. However, the adoption of this automated approach has been slower than anticipated. After two years of routine use, **only 10% of data collection efforts have used the app's automatic mode.**

To recap, in manual mode, the users identify plastic litter themselves, while in automatic mode, they

Since the start of the Plastic Origins project, 1,100 campaigns have been conducted, reporting the presence of nearly 35,000 pieces of litter across over 2,000 km of river in eight European countries (as of September 2024).

Figure 27 | Above | Separating and counting waste after collection, © Surfrider Foundation.

record a video via the app and upload it to a server for later AI analysis *(see paragraph 2)*.
Citizens seem to prefer the playful nature of the manual mode, a preference also observed among the most experienced, trained, and seasoned app users who cover long distances. According to them, this mode allows for visual confirmation of each piece of litter and a sense of direct involvement in environmental protection. Conversely, the automatic mode — with its off-site processing and lack of immediate feedback — does not engage volunteers in quite the same way.
By shifting this off-site system to an on-device setup where AI operates directly on the mobile, real-time visualisation of detection results could be made possible. Making the algorithm's decision-making process more apparent would help users to better understand how the AI assists them, thus fostering greater trust in its capabilities.

5.2. WHAT ARE THE BENEFITS AND ISSUES SURROUNDING THE INTEGRATION OF AI ON OUR MOBILE PHONES?

The various phases of development have culminated in a high-performance, economically viable AI solution that balances detection performance with a cost-effective and sustainable technological approach.
Initially hosted in the Cloud, the impact of the AI developed for Plastic Origins quickly became a point of discussion within the project team. On the one hand, there were concerns about the high costs associated with GPU use for real-time service; on the other, the environmental impact of such a setup seemed unjustifiable.
We therefore opted for a more sustainable solution, implementing innovative engineering (Le Roux 2023). First, AI processing was rescheduled, moving from a 24/7 real-time service to a periodic computation model that ran every few hours or even weekly, depending on the volume of video footage requiring analysis. Simultaneously, we initiated a miniaturisation phase, enabling the processing to be carried out on CPU-based servers, which are less energy-intensive than GPU servers. These initial improvements reduced the operational costs of Plastic Origins' AI by a factor of 10 to 100.

Next, recent technological advancements presented an opportunity for further enhancement. While early AI models were too large for mobile devices, advancements in TinyML now enable us to "migrate" the AI from the Cloud to users' mobile devices, thus cutting costs (Ollion 2023). Running AI models directly on a mobile device eliminates the need to transfer, store, and analyse videos on powerful, costly Cloud servers. The only anticipated energy cost comes from the increased power consumption of the mobile phone.

With this in mind, we embarked on an innovative research and development journey, conducting initial tests to integrate the AI so it could run directly on mobile devices. This shift introduced new engineering challenges, **as we wanted our mobile AI solution to be efficient, ergonomic, accessible, and energy-efficient:**

Efficient: The AI must perform reliably for accurate and useful object counting. To achieve this, we estimate that a minimum of five detections per second are necessary to ensure smooth object tracking. Considerable engineering efforts are required to miniaturise the model while preserving high detection performance. Additionally, since video resolution is capped at 640x640 due to computational speed constraints, detecting smaller pieces of litter becomes more challenging. Currently, tracking methods on mobile devices are still not mature and demand highly specialised technical work.

Ergonomic: Building user trust and ensuring intuitive AI-assisted detection. To make the app user-friendly, we engaged UX designers and conducted user experience workshops to understand how users perceive and interact with the app. We decided to include clear explanations of how the AI functions and developed a tutorial to guide users toward optimal usage and understanding.

Accessible: The app should work on various smartphones, not only the latest models. Despite extensive optimisations, achieving smooth, real-time detection still generally requires a relatively recent phone. Tests conducted on various devices showed good functionality with a Galaxy S10e (2019) or equivalent. However, some newer phones lacking dedicated graphics cards struggle to run the AI at the required speed.

Energy-efficient: The phone should sustain operation long enough to monitor extended stretches of riverbanks. Our tests demonstrated that filming and running the AI for several dozen minutes is feasible. However, energy consumption, particularly by the AI module, varies significantly depending on the device, making this a key consideration moving forward.

__Figure 27__ | __Top__ | French tutorial for onboarding users of the prototype application | __Figure 29__ | __Middle__ | Main screen for real-time detection of litter in rivers via the prototype application | __Figure 30__ | __Bottom__ | End of recording pop-up in the prototype application.

5.3. CAN THE DEVELOPMENT OF SURFNET PROVIDE RELIABLE DIAGNOSTICS OF RIVER POLLUTION?

The overarching aim of the Plastic Origins project is to create a detailed map of plastic waste along European waterways, to provide an overview of the extent and distribution of plastic pollution.
Using this tool, Surfrider's objective is to raise awareness among local, national, and European policymakers regarding the scale of the issue.

In 2019, the Plastic Origins project aimed to create comprehensive diagnostics for river pollution. This process involved classifying rivers according to their contamination levels by using an indicator based on the density of waste along the riverbanks. The goal was to trace the sources of diffuse pollution by identifying the types of waste present.

> The challenge of developing a functional mobile app is well understood; however, the technical feasibility of effectively meeting all outlined constraints remains unproven. Our initial prototypes demonstrated running real-time waste detection and tracking models is possible, yet they highlighted the need for considerable engineering efforts. Current tools are not mature enough to fully support these requirements.

However, inconsistencies between the waste categories used in the AI and manual modes, as well as the challenges people faced to categorise waste accurately, prompted a reassessment of our approach. Consequently, since early 2023, we streamlined the categorisation into three main groups:

— large items that can be easily reported to local authorities;
— Accumulation zones warranting focused efforts (e.g., quantification, prevention, and clean-up);
— and a broad category grouping all other waste types.

Although this simplification has made the data more usable, the original aim of conducting precise diagnostics was shifted towards a more straightforward reporting framework. The newly revised approach limits its potential to produce a valuable indicator for ongoing river monitoring.

Thus, while the citizen science outcomes of this project help raise awareness among both the public and policymakers, they do not meet the requirements for scientifically robust and enduring indicators for river pollution monitoring (Tramoy, 2022).

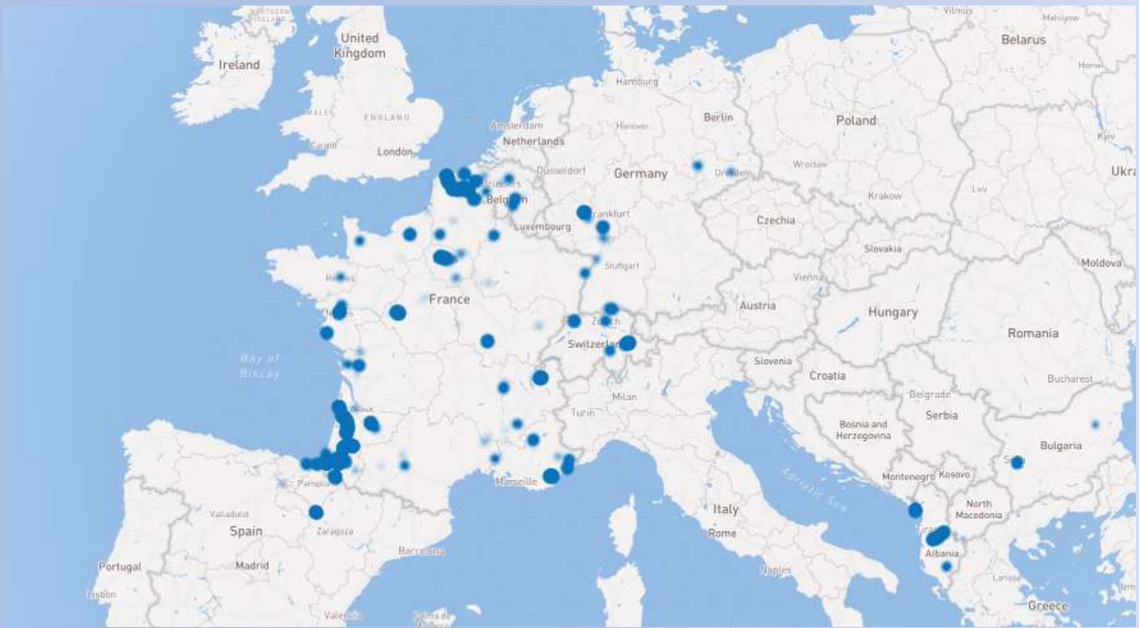
Practically speaking, in the field, the Plastic Origins app is used by partners such as the NGO Découverte et Participation à la Préservation des Milieux (Environmental Discovery and Protection Program - DPPM, 62) and the Parc Naturel Marin du Bassin d'Arcachon (PNMBA - Bassin d'Arcachon Marine Nature Park) to geolocate waste during their river surveillance missions. These organisations also collect and quantify waste based on the OSPAR/DCSMM classification.
Although it has its limitations, the data gathered by Plastic Origins is important for identifying areas of accumulation and guiding more in-depth studies. This may involve applying the OSPAR protocol to river systems, which is being developed as a framework for monitoring European waterways (UNEP 2023).

5.4. HOW CAN WE EVALUATE THE ENVIRONMENTAL IMPACT OF SURFNET AI IN RELATION TO ITS ENVIRONMENTAL OBJECTIVES?

While awareness of the environmental impacts of AI is growing, quantifying and qualifying these impacts remains challenging. However, as AI use rapidly expands and climate issues become ever more pressing, these considerations are increasingly crucial. Historically, attention has been centred on the AI's training phase, yet the growing prevalence of AI calls for parallel scrutiny of the inference phase. Following the EU's work in defining ethical guidelines for AI (European Commission 2024), AI frugality has emerged as a particularly important issue.

Both the development and deployment of AI systems demand considerable quantities of electricity and water while contributing to global CO_2 emissions. Although specific indicators to define AI frugality are not yet fully established, initiatives such as AFNOR's general standard for Frugal AI (AFNOR 2024) already suggest best practices to address these concerns. In the context of Plastic Origins' development, several of these frugality measures were integrated proactively five years ago. Let us cite, for example:
Acculturating Teams to AI Frugality (BP14 in the

AFNOR Standards): To do so, we first engaged in discussions on the environmental impact of our AI solution. We identified significant opportunities to reduce computational power during the AI usage (inference) phase and adapt it more closely to actual operational needs. We then implemented platform updates to optimise performance.
Reusing Pre-trained Algorithms (BP29): we systematically used pre-trained algorithms and models, primarily trained on the well-known ImageNet dataset in computer vision. As a result, our models required only "fine-tuning" rather than extensive initial training, enabling efficient detection performance with significantly reduced training time and computational demand.
Optimising Hardware Usage (BP20): Initially, we



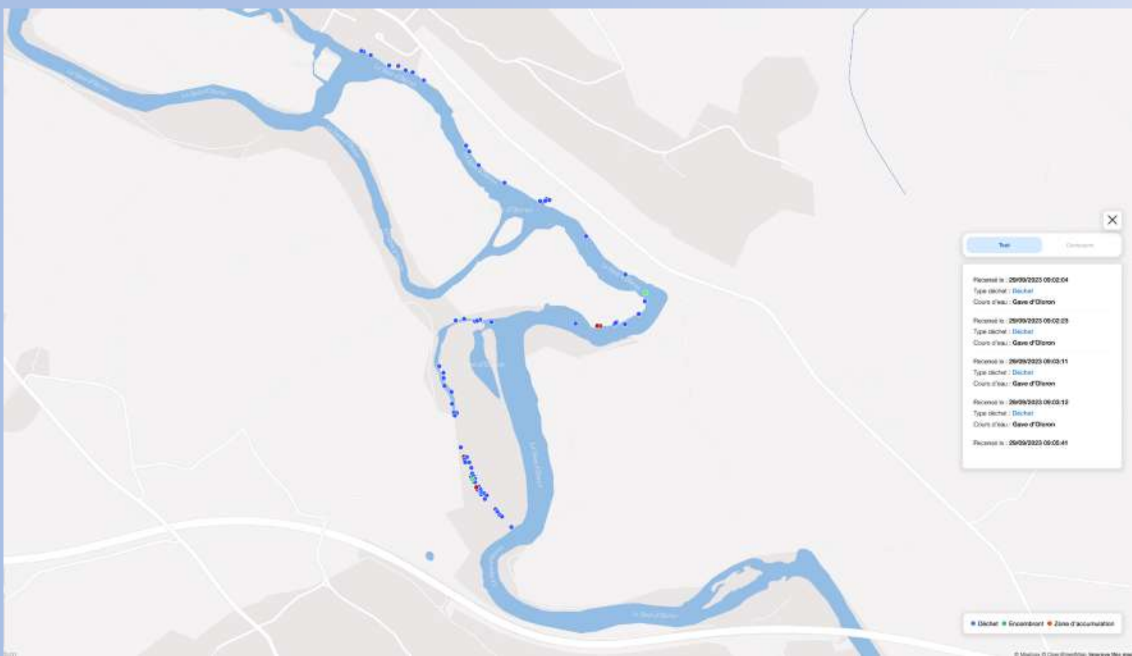


Figure 31 | Top | Heatmap showing pollution reported through the Plastic Origins application | Figure 32 | Bottom | Zoom in on the Sorde-l'Abbaye (40) sector, showing the distribution of waste along the river.

transitioned from a 24/7 operational AI to an asynchronous mode, running only a few hours over the weekend. Subsequently, we initiated our R&D program to deploy AI directly on mobile devices, exploring the potential of eliminating dedicated cloud-based AI servers from the process altogether.

or finally Anticipating AI Project End-of-Life (BP11): the experimental nature of Plastic Origins and its basis in citizen science led us to undertake a pragmatic and budget-conscious approach. To ensure project viability, we proactively addressed sustainability and long-term relevance. Once key results were demonstrated, we established a roadmap to gradually reduce or phase out the digital platform, keeping environmental considerations in mind.

Beyond following best practice guidelines, our standardisation efforts have helped identify tools that provide frugality indicators for AI. For instance, the Green Algorithms calculator (www.green-algorithms.org) and the open-source CodeCarbon software (codecarbon.io) now offer capabilities to estimate electricity consumption and CO_2 emissions both before and during AI development. Such tools pave the way for a new frugal approach to AI development, where designers set predefined limits on energy use, CO_2 emissions, and water consumption relative to the value of the problem the AI seeks to solve. At the very least, these steps open possibilities for more environmentally responsible AI, aligning with a societal shift towards technological sobriety.

Figure 33 | ***Above*** | *Waste stranded on a riverbank, © Surfrider Foundation.*

6 Perspectives and Recommendations

The technological developments of the Plastic Origins project have enabled us to explore the role Artificial Intelligence may play in collecting data on plastic pollution in rivers.

The technological developments of the Plastic Origins project have enabled us to explore the role Artificial Intelligence may play in collecting data on plastic pollution in rivers.

Among the project's key contributions are the compilation of a unique dataset on plastic pollution in waterways and the creation of Surfnet, an AI model specific to the issue. Furthermore, through our work adapting TinyML technologies, we have demonstrated the feasibility of developing lightweight, resource-efficient AI, thereby reducing its infrastructure costs and environmental impact. This marks a significant step forward in optimising AI technologies for increasingly accessible mobile applications.

The mobile application incorporating this AI has shown promising potential in transforming recreational activities like kayaking into opportunities for citizen science. This approach helps people to get involved in combating plastic pollution whilst reinforcing the notion that technology like AI can play a meaningful role in environmentalism.

Nevertheless, the project also encountered both technological and human limits. Although AI has enabled a form of automated waste detection, its performance remains subject to improvement. The quality of videos provided by volunteers is of high importance, and broader adoption of the automatic mode remains a challenge. Indeed, participants expressed a certain reluctance to fully delegate observation to AI, particularly due to the absence of real-time feedback. A native, on-device AI model could potentially overcome this issue. However, that would require very recent smartphones to work correctly, and compromise accessibility concerns crucial for citizen science protocols.

The Plastic Origins project is planned to conclude in December 2025. Despite the levels of performance achieved, the AI developed does not entirely fulfil the initial objectives for waterway diagnostics. The environmental impact of the technology, combined with its mixed results, has led us to question the viability of pursuing this direction, particularly as an alternative protocol is gradually establishing itself as the framework for monitoring hydrographic networks across Europe.

We are confident that our method will contribute to future discussions, and we can draw several key takeaways to inspire other initiatives:

→ Encourage citizen science by simplifying protocols and enhancing real-time result transparency to promote greater engagement and improve data quality.
→ Continue the development of resource-efficient AI models to reduce the environmental footprint of models whilst optimising their operational efficacy on accessible smartphones.
→ Finally, the entirety of the data generated over the course of this project is available as open data to enable further testing and adaptation for diverse objectives. This open and collaborative approach, relatively rare in AI projects, promotes transparency and encourages the appropriation of the technology for more sustainable innovation.

7 Bibliography

AFNOR
«Référentiel général pour l'IA frugale - Mesurer et réduire l'impact environnemental de l'IA.» 2024.

BRUGE, A., BARREAU, C., CARLOT, J., COLLIN, H., MORENO, C., MAISON, P.
«Monitoring Litter Inputs from the Adour River (Southwest France) to the Marine Environment. .» J. Mar. Sci. Eng., 2018.

CHAGNEUX, M., LE CORFF, S., GLOAGUEN, P., OLLION, C., LEPÂTRE, O., AND BRUGE, A.
«Macrolitter Video Counting on Riverbanks Using State Space Models and Moving Cameras.» Computo, 2023.

CRAWFORD, K.
«Generative AI's environmental costs are soaring - and mostly secret.» Nature, 2024.
Emmerik, T. and Schwarz, A. «Plastic Debris in Rivers.» WIREs wATER, 2020.

EUROPEAN COMMISSION
«IA Act : Regulatory framework proposal on artificial intelligence.» 2024.
Galgani, F. «Les déchets marins.» 2016.

GESAMP, JOINT GROUP OF EXPERTS ON THE SCIENTIFIC ASPECTS OF MARINE POLLUTION
«Report of the 20th Session of GESAMP.» 1990.

GONZÁLEZ, D. ET AL.
«Riverine Litter Monitoring – Options and Recommendations.» 2017.

HANKE, G. ET AL.
«Guidance on Monitoring of Marine Litter in European Seas - JRC.» Publications Office of the European Union, 2013.

LE ROUX, C.
«Optimizing and running AI & Data: why Plastic Origins choose Azure Kubernetes Service and Airflow for sustainable operations.» Médium, 2023.

LEPÂTRE, O.
«How can tech help address the issue of plastic pollution in rivers ?» Medium, 2023.

META
«Building Meta's GenAI Infrastructure.» Meta Engineering Blog, 2024.

OFFICE FRANÇAIS POUR LA BIODIVERSITÉ
Les rivières. s.d. www.eaufrance.fr/les-rivieres.

OLLION, C.
«Balancing Performance and Sustainability: How Plastic Origins chooses its AI solution for mobile devices.» Medium, 2023.

OLLION, C.
«Small Scale Computer Vision in 2024.» Medium, 2023.

TRAMOY, R., GASPERI, J., ET TASSIN, B.
«Plastoc : Indicateurs de la pollution en macrodéchets dans l'environnement et estimation des flux issus des eaux urbaines.» 2022.

UNEP, MED WG.550/14.
«Guidelines for Monitoring Riverine inputs of Marine Litter.» 2023.

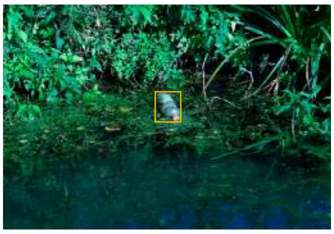
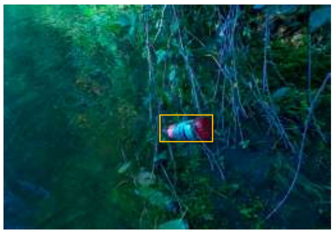
8 Annex

Example of images annotated according to the different categories of the dataset

INSULATION, POLYSTYRENE



BOTTLE-TYPE LIQUID CONTAINER

CAN-TYPE LIQUID CONTAINER



PLASTIC DRUMS



FOOD PACKAGING OTHER THAN BOTTLES OR CANS



TIRES



FISHING LINE AND ROPE



EASILY NAMABLE OBJECTS



UNCLEAR OBJECTS



TARPS, PLASTIC BAG, PLASTIC FRAGMENTS



WWW.SURFRIDER.EU