MPVision: An End-to-End Computer Vision Pipeline for Rapid Aquatic Microplastic Classification & Quantification

Yuewen Luo[1], Ungku Zoe Anysa Ungku Faiz[2], Chelsea Rochman[2], Florian Shkurti[3,4,5]
[1]Division of Engineering Science, University of Toronto, Toronto, Ontario, Canada
[2]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada
[3]Department of Computer Science, University of Toronto, Toronto, Ontario, Canada
[4]University of Toronto Robotics Institute, Toronto, Ontario, Canada
[5]Vector Institute for Artificial Intelligence, Toronto, Ontario, Canada
{sofia.luo, ungkuzoe.ungkufaiz}@mail.utoronto.ca,
chelsea.rochman@utoronto.ca, florian@cs.toronto.edu

Abstract

Manual analysis of microplastics from images taken under a stereo microscope demands significant time, typically requiring between 30 and 60 minutes on average per sample, creating a bottleneck for large-scale studies. We introduce MPVision, an end-to-end computer vision pipeline automating segmentation, classification, and quantification of microplastics, reducing analysis time by nearly an order of magnitude once images are available. First, the Segment Anything Model (SAM) with a ViT-H backbone generates candidate particle masks and corresponding cutouts from microscope images. Second, an eight-layer CNN distinguishes true plastics from non-plastic debris, achieving 98.11 % accuracy. Third, a seven-layer CNN classifies confirmed particles into yellow polyethylene (PE), pink polystyrene (PS), or blue polyethylene terephthalate (PET), achieving 95.45 % accuracy. Finally, binary masks yield sub-millimeter geometric features. Validated on 253 lake mesocosm images from a large freshwater research facility, MPVision provides a scalable, reproducible workflow bridging remote-sensing and laboratory analyses, enabling rapid, high-throughput microplastic quantification for ecological and toxicological research.

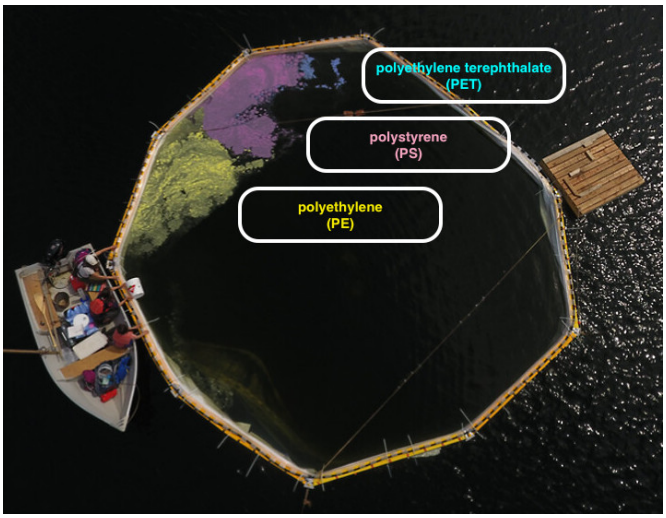
1. Introduction

Due to global transport via atmospheric dust [5], water flow [4], and biogeochemical cycling [7], microplastics—defined as synthetic polymer particles less than 5 mm in size [2, 6]—have emerged as a widespread contaminant raising concerns for aquatic ecosystems and human health [1]. Understanding their transformation and distribution is crucial for predicting accumulation patterns, assessing wildlife exposure, and improving toxicity testing [10].

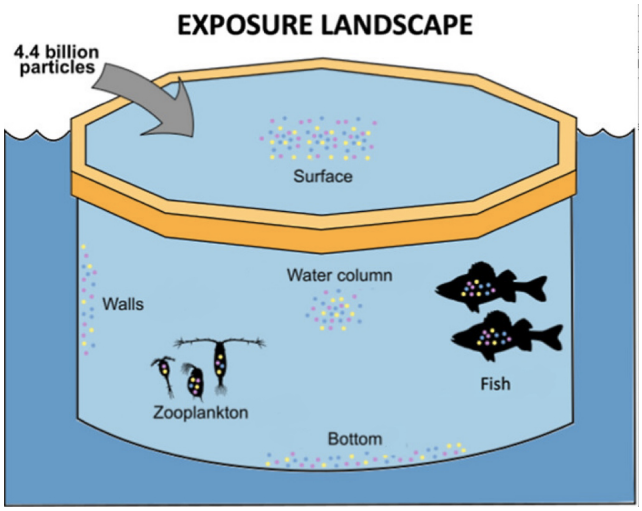
To investigate these dynamics under controlled yet ecologically relevant conditions, we leveraged datasets from Rochman et al. [10], who conducted a large-scale in-lake mesocosm experiment at Lake 378 in the IISD–Experimental Lakes Area, Northwestern Ontario, Canada. As shown in Fig. 1a, they introduced mixtures of three representative plastic polymers—yellow polyethylene (PE; $\sim$0.93 g/cm^3), pink polystyrene (PS; $\sim$1.0 g/cm^3), and blue polyethylene terephthalate (PET; $\sim$1.4 g/cm^3)—chosen for their prevalence in consumer products and environmental relevance [8]. Water samples were collected over a 10-week period to track particle distribution across ecological compartments (see Figs. 1a and 1b). Microplastics were size fractioned (212, 106, and 53 µm).

Traditional microplastic characterization is typically performed either by directly counting and identifying particles under a stereo microscope or by manually analyzing microscope images—both of which are labour- and time-intensive. For one water sample, it requires approximately 30 microscope images for particles >212 µm, 55 for 106–212 µm, and 150 for 53–106 µm. For large-scale ecological studies, manual analysis times create a significant bottleneck.

To overcome these inefficiencies, we introduce *MPVision*, an automated computer vision pipeline that accelerates microplastic analysis by focusing on one of the most challenging tasks: reliable detection and classification of microscope-captured microplastic particles. We benchmarked manual image-based assessment against our automated *MPVision* pipeline, holding image acquisition constant. Across three representative water samples, manual image analysis required on average **32 minutes** for parti-



(a) Close-up drone image of a lake mesocosm showing the three types of microplastics as they are being added. Included from [10] with permission.



(b) Schematic of microplastic distribution across ecological compartments in a lake mesocosm. Included from [10] with permission.

Figure 1. Visual overview of microplastic deployment and distribution.

cles larger than 212 µm and **54 minutes** for particles in the 106–212 µm range. In contrast, *MPVision* processed the same images in approximately **5 minutes** and **6.5 minutes**, respectively. These results demonstrate that, once images are available, *MPVision* reduces analysis time by nearly an order of magnitude.

Beyond efficiency, the pipeline also delivers high classification accuracy across particle types. Our system leverages the Segment Anything Model (SAM) for instance segmentation, followed by a two-stage classification framework and geometric measurement module, all wrapped in an interactive user interface.

The pipeline has three main components. First, SAM generates candidate masks for particles within microscope images. We then introduce a binary classification stage to distinguish microplastics from non-plastic segments, trained on **9744** non-plastic and **1098** plastic segments, achieving **98.11%** accuracy. The binary classifier is implemented as a deep convolutional neural network (CNN) capturing fine-scale spatial features to discriminate plastic-like segments from non-plastic debris and imaging artifacts.

In the second stage, validated plastic particles are classified into yellow (PE), pink (PS), or blue (PET) categories. This colour classification CNN is trained on **512** yellow, **360** pink, and **226** blue samples, achieving **95.45%** accuracy.

Finally, for each particle, we compute maximum length, width, and pixel area based on segmentation mask geometry, which are essential metrics for ecological and toxicological interpretation [10]. The full system is packaged into a user-friendly desktop application that supports batch image processing, interactive result visualization, and tabular data export, making it practical for real-world ecological workflows.

Our current validation uses lake mesocosm imagery with three representative polymers. We focus on the size fractions most visually distinguishable: particles larger than 212 µm and particles between 106–212 µm ranges. Sub-100 µm particles are outside the scope of our current validation and represent an important direction for future work. Although this setup constrains color, shape, and size variability, the modular architecture of MPVision could be extended to other polymers and imaging conditions in future studies.

2. Related Work

2.1. Vision-based monitoring of water quality

Early computer-vision approaches to aquatic contamination focused on biological hazards such as harmful algal blooms (HABs). Samantaray *et al.* [11] demonstrated that deep CNNs can detect algal scums in both ground- and aerial-view images using a fusion of Faster R-CNN, SSD, and R-FCN. Training on 427 hand-curated images, they achieved mean average precision (mAP) up to 0.71 at real-time speeds (25–30 fps), establishing that transfer learning is crucial when annotated data are scarce.

2.2. Deep-learning detection of floating macro-debris

More recently, focus has shifted to anthropogenic debris like floating plastics. Panwar *et al.* introduced **AquaVision**, a RetinaNet detector trained on the AquaTrash dataset (369 images), obtaining an mAP of 0.8148 and per-class

precision above 0.74 [9]. Their benchmarks confirmed that RetinaNet offers a favorable speed-accuracy tradeoff for environmental applications.

Armitage *et al.* mounted cameras on vessels and trained a YOLOv5s model to detect floating plastics in video streams. Evaluating on 9,824 test frames, they reported 95.2% accuracy for presence/absence and class-level mAP of 0.68 while processing at 40 fps on edge hardware [3]. Their analysis showed performance degrades as object coverage falls below 0.5% of the frame.

2.3. Gap analysis and our contribution

Existing systems target macro-scale debris ($\geq$ 1 cm) in complex scenes and lack the geometric detail needed for microplastic ecology. We present the first end-to-end pipeline using SAM for segmentation, CNNs for classification, and extraction of sub-millimeter morphological features. Our method reduces per water sample processing time by nearly an order of magnitude, while achieving high accuracy, specificity, and sensitivity on a new lake-mesocosm dataset—bridging the gap between remote sensing and laboratory workflows.

3. Methodology

As shown in the system pipeline (Fig. 2), the workflow of *MPVision* consists of four core stages: (1) segmentation using the Segment Anything Model (SAM), (2) binary classification of segmented objects as plastic or non-plastic, (3) colour-based classification of plastic particles into yellow PE, pink PS, or blue PET, and (4) geometric analysis to extract morphological features.

Each microscope image is processed using SAM (ViT-H backbone) with custom hyperparameters (e.g., IoU threshold of 0.86, minimum mask area of 100 pixels) to produce binary masks and cut-out images. These segments are passed through an 8-layer CNN trained on 9744 non-plastic and 1098 plastic instances to identify true plastic particles. Positive detections are then classified by a 7-layer CNN trained on 512 yellow, 360 pink, and 226 blue particles to determine plastic type.

For each confirmed plastic particle, the binary mask is analyzed to calculate physical dimensions, including maximum width, length, and area. The conversion ratio utilized are **315 pixels/mm** and **208 pixels/mm** for particles of size $106 - 202\,\mu\text{m}$ and $> 202\,\mu\text{m}$, respectively.

All results are displayed in a custom graphical interface allowing users to browse results and export structured Excel files, enabling rapid analysis in under **10 minutes per water sample**.

3.1. Microplastic Dataset

We focus on microplastics in the $>212\,\mu\text{m}$ and 106–212 µm ranges, which are most visually distinguishable. From three

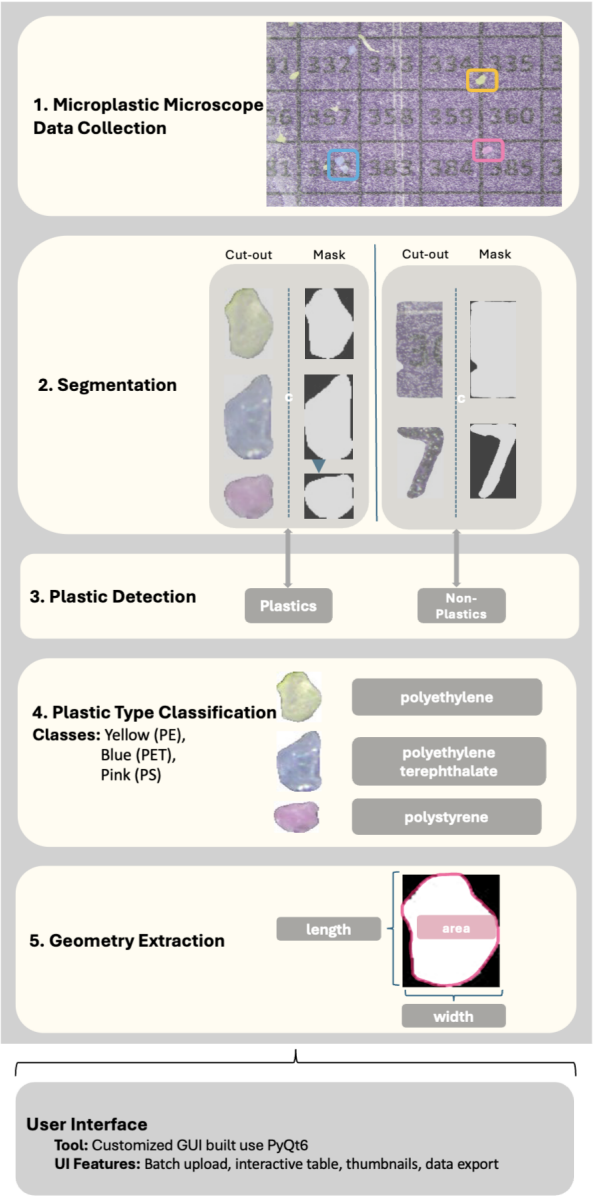


Figure 2. End-to-End pipeline for automated microplastic detection and quantification.

representative samples, there are **88** microscope images for the $>212\,\mu\text{m}$ fraction and **165** images for the 106–212 µm fraction. Using SAM, candidate regions were segmented, resulting in **9744** non-plastic and **1098** plastic particle cut-outs. Among plastic samples, **512** are yellow polyethylene (PE), **360** are pink polystyrene (PS), and **226** are blue polyethylene terephthalate (PET). Both classification models used a 70/20/10 train/validation/test split.

Ground truth annotations were provided by trained staff at the Rochman Lab. Plastic particles were manually identified and extracted from filter papers under a dissecting mi-

croscope. Technicians performing the labelling had formal training and demonstrated proficiency in plastic identification. All plastic labels were visually confirmed by at least one additional reviewer to ensure labelling consistency.

For non-plastic labels, we adopted a conservative exclusion strategy: any segmented region produced by SAM that did not spatially overlap with a manually identified plastic particle was labelled as non-plastic. This method allowed us to efficiently generate a large negative set without introducing false positives from plastic contamination.

This protocol ensures that the dataset used to train and evaluate our pipeline is both representative and rigorously annotated, supporting reliable benchmarking of segmentation and classification performance.

3.2. Segmentation

We employ the Segment Anything Model (SAM) with a ViT-H backbone using `SamAutomaticMaskGenerator` with task-specific hyperparameters: **32 points per side**, IoU prediction threshold of **0.86**, stability score threshold of **0.97**, and minimum mask area of **100 pixels**. SAM produces binary masks and corresponding cut-out images for downstream classification. Examples are illustrated in Fig. 2. This automated segmentation eliminates the need for extensive manual annotation and enables scalable pre-processing of microscopy data, providing high-quality inputs for subsequent classification and quantitative morphometric measurements.

3.3. Plastic Detection

We classify each particle candidate as plastic vs. non-plastic with a CNN architecture. Inputs are resized to 128×128 pixels and normalized. The model stacks six convolutional blocks (batch normalization, ReLU, pooling, dropout) followed by two fully connected layers (total size $\approx$ **139**M learnable parameters). In practice, the network learns polymer-specific morphology and texture while suppressing spurious cues, allowing it to ignore artifacts such as printed grid numbers and organic debris.

To promote robustness and generalization, we train with standard geometric and photometric augmentation and regularization.

3.4. Plastic Type Classification

We subtype plastic particles into yellow PE, pink PS, and blue PET with a compact CNN. Inputs are resized to 150×150 pixels and normalized. The model stacks five convolutional blocks (batch normalization, ReLU, pooling, dropout), followed by two fully connected layers and a final softmax over the three classes (total size $\approx$ **0.805**M parameters). The network emphasizes chromatic and textural cues characteristic of each polymer while suppressing spurious markings.

To promote robustness and generalization, we use standard geometric and photometric augmentation during training, together with conventional regularization.

3.5. Geometry Extraction

Geometric features are computed from each plastic particle's binary mask generated in the segmentation stage. We extract mask dimensions for maximum width and length, and count white pixels to compute area. All measurements are converted from pixels to millimetres using a calibration factor that depends on particle size. The conversion ratio is determined based on the length of the filter paper grid (2 mm): for particles between $106\,\mu\mathrm{m}$ and $202\,\mu\mathrm{m}$, we use **315 pixels/mm**; for particles larger than $202\,\mu\mathrm{m}$, we use **208 pixels/mm**. Unlike traditional manual approximations with ruler, our method provides precise size estimation by directly integrating pixel-level information.

3.6. Graphical User Interface

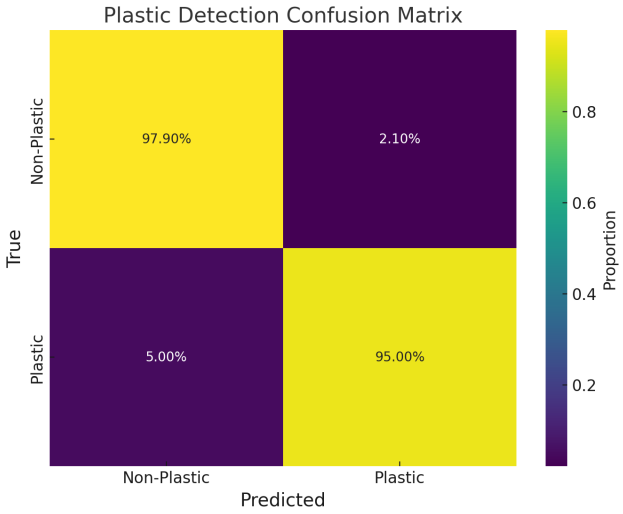
We developed a standalone graphical user interface (GUI) using PyQt6 to make the pipeline accessible to non-technical users. The application allows users to upload microscope images and process them through the complete workflow with a single interface. Output is presented in an interactive results table containing particle-level information, including preview images, colour classification, and size measurements. The GUI is coupled with a backend controller that coordinates model inference and data processing, thereby enabling reproducible and automated execution without reliance on programming expertise or command-line operations.

4. Results and Discussion

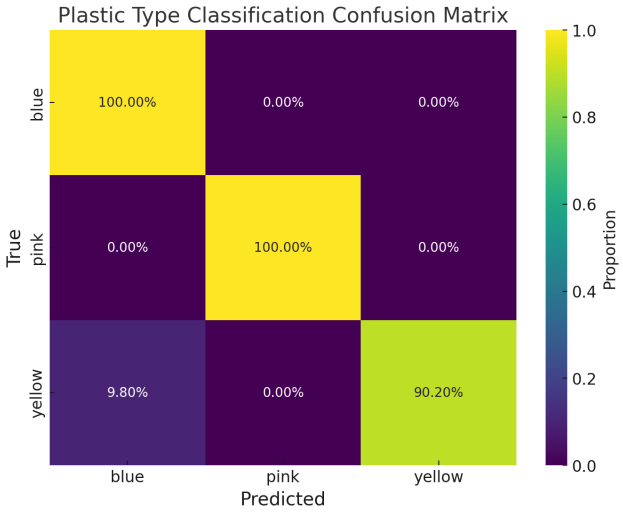
We evaluate the performance of our *MPVision* pipeline in terms of classification accuracy and runtime efficiency, presenting results for both binary plastic detection and multi-class plastic type classification. We also compare our system's execution time against conventional manual analysis.

Metric	Plastic Detection CNN	Plastic Type Classification CNN
Accuracy	98.11%	95.45%
Sensitivity	94.55%	96.73%
Specificity	98.51%	97.73%
Precision	91.51%	94.05%
F1-score	93.80%	95.01%

Table 1. Quantitative evaluation of binary plastic detection and multi-class type classification models.



(a) Confusion Matrix for the plastic detection CNN (row-normalized %).



(b) Confusion Matrix for the plastic type classification CNN (row-normalized %).

Figure 3. Confusion matrixes for the two classification CNNs (row-normalized %).

Test	Size (μm)	Time (Manual → MPVision)	Reduction (%)
Water Sample 1	> 212	70 min → 6 min	91.4
	106–212	90 min → 8 min	91.1
Water Sample 2	> 212	15 min → 5 min	66.7
	106–212	51 min → 6 min	88.2
Water Sample 3	> 212	10 min → 4 min	60.0
	106–212	20 min → 5 min	75.0
Average	> 212	**32 min → 5 min**	**84.4**
	106–212	**54 min → 6.5 min**	**88.0**

Table 2. Comparison of manual image-based analysis and MPVision processing times across three water samples (excluding image acquisition).

4.1. Plastic Detection Performance

As summarized in Tab. 1, the plastic detection model demonstrates strong performance. On a held-out test set of 1,084 instances, it achieves an overall accuracy of **98.11%**, with recall (sensitivity), specificity, macro-averaged precision, and F1-score of **96.73%**, **97.73%**, **94.05%**, and **95.01%**, respectively. According to the confusion matrix in Fig. 3a, class-wise analysis reveals a sensitivity of **95.00%** and specificity of **97.90%** for detecting plastic particles.

These results indicate that the model effectively suppresses false positives from visually similar non-plastic materials while maintaining a high true positive rate for plastics—crucial for accurate downstream particle typing and ecological interpretation.

4.2. Plastic Type Classification Performance

Our second-stage model assigns plastic particles to one of three polymer categories. The model achieves a test accuracy of **95.45%**, with a sensitivity of **95.45%** and specificity of **97.73%** as shown in Tab. 1.

According to the confusion matrix in Fig. 3b, per-class performance shows perfect sensitivity (**100.0%**) for both pink and blue plastics, with corresponding specificities of **100.0%** and **94.25%** calculated using raw counts. Yellow plastics are classified with a sensitivity of **90.20%** and a specificity of **100.0%**. These results confirm the suitability of color-based cues as reliable features for identifying these specific polymers under controlled conditions. However, this performance may not generalize to environmental samples, where yellowish or clear particles are common.

4.3. Efficiency Gains Over Manual Workflow

As shown in Tab. 2, our pipeline offers substantial improvements in processing speed. In our benchmarking across three representative water samples, manual image-based analysis required on average **32 minutes** for particles larger than 212 µm and **54 minutes** for particles in the 106–212 µm range. In contrast, our system processed the same images in approximately **5 minutes** and **6.5 minutes**, respectively.

This corresponds to an average reduction of **84%** in analysis time for the $>$212 µm fraction and **88%** for the 106–212 µm fraction, enabling rapid, scalable processing of large image sets with minimal user intervention. The integration of all components into a single GUI application further enhances usability in research and monitoring contexts.

5. Conclusion

We have presented *MPVision*, an end-to-end computer vision pipeline that automates the detection, classification, and quantification of microplastics in mesocosm stereo microscope imagery. Through the integration of the Segment Anything Model with a two-stage CNN framework and pre-

cise geometry extraction, our system decreases per-sample processing time from roughly 30–60 minutes to 5–7 minutes, representing nearly an order-of-magnitude reduction. Using the lake mesocosm dataset, we demonstrate 98.11% accuracy for plastic detection, 95.45% accuracy for plastic type classification, and more than 80% reduction in overall analysis time.

We deliver a GUI application that enables environmental researchers to batch process image folders, review outputs, and export annotated results without coding. This integration of advanced segmentation, deep learning, and user-friendly software bridges the gap between high-throughput sensing and labor-intensive laboratory workflows.

Our modular design readily accommodates new polymer categories, alternative segmentation approaches, and adaptation to other imaging modalities. Future work will extend the pipeline to sub-100 µm particles via higher-resolution imaging and unsupervised domain adaptation, and integrate results with transport models for predictive risk assessment. This toolset will accelerate reproducible microplastic research and provide a foundation for next-generation environmental monitoring.

References

[1] Zeynep Akdogan and Basak Guven. Microplastics in the environment: A critical review of current understanding and identification of future research needs. *Environmental Pollution*, 254:113011, 2019. 1

[2] Anthony L. Andrady. The plastic in microplastics: A review. *Marine Pollution Bulletin*, 119(1):12–22, 2017. 1

[3] Sophie Armitage, Katie Awty-Carroll, Daniel Clewley, and Victor Martinez-Vicente. Detection and classification of floating plastic litter using a vessel-mounted video camera and deep learning. *Remote Sensing*, 14(14):3425, 2022. 3

[4] Melanie Bergmann, Sophia Mützel, Sebastian Primpke, Mine B. Tekman, Jürg Trachsel, and Gunnar Gerdts. White and wonderful? microplastics prevail in snow from the alps to the arctic. *Science Advances*, 5(8):eaax1157, 2019. 1

[5] Janice Brahney, Natalie Mahowald, Marje Prank, Gavin Cornwell, Zbigniew Klimont, Hitoshi Matsui, and Kimberly Ann Prather. Constraining the atmospheric limb of the plastic cycle. *Proc Natl Acad Sci U S A*, 118(16), 2021. 1

[6] Christopher Blair Crawford and Brian Quinn. Plastic production, waste and legislation. *Microplastic pollutants*, 30: 39–56, 2017. 1

[7] John Paul Dees, Mohamed Ateia, and Daniel L. Sanchez. Microplastics and their degradation products in surface waters: A missing piece of the global carbon cycle puzzle. *ACS ES&T Water*, 1(2):214–216, 2021. 1

[8] Alice A. Horton, Alexander Walton, David J. Spurgeon, Elma Lahive, and Claus Svendsen. Microplastics in freshwater and terrestrial environments: Evaluating the current understanding to identify the knowledge gaps and future research priorities. *Science of The Total Environment*, 586: 127–141, 2017. 1

[9] Harsh Panwar, P. K. Gupta, Mohammad Khubeb Siddiqui, Ruben Morales-Menendez, Prakhar Bhardwaj, Sudhansh Sharma, and Iqbal H. Sarker. Aquavision: Automating the detection of waste in water bodies using deep transfer learning. *Case Studies in Chemical and Environmental Engineering*, 2:100026, 2020. 3

[10] Chelsea M Rochman, Kennedy Bucci, Desiree Langenfeld, Rachel McNamee, Cody Veneruzzo, Garth A Covernton, Gloria H Y Gao, Mira Ghosh, Rachel N Cable, Ludovic Hermabessiere, Raul Lazcano, Michael J Paterson, Michael D Rennie, Rebecca C Rooney, Paul Helm, Melissa B Duhaime, Timothy Hoellein, Kenneth M Jeffries, Matthew J Hoffman, Diane M Orihel, and Jennifer F Provencher. Informing the exposure landscape: The fate of microplastics in a large pelagic In-Lake mesocosm experiment. *Environ. Sci. Technol.*, 58(18):7998–8008, 2024. 1, 2

[11] Arabinda Samantaray, Baijian Yang, J. Eric Dietz, and Byung-Cheol Min. Algae detection using computer vision and deep learning. *arXiv preprint arXiv:1811.10847*, 2018. 2