

Microplastic detection using Machine learning

Ashirvad pandey¹, Ahuti Singh², Nikhil Singh Rajawat³
IIMT Group of College

Abstract—Microplastic pollution has emerged as a critical environmental issue, with particles smaller than 5 mm infiltrating oceans, rivers, soils, and even the human food chain. These microplastics originate from industrial waste, cosmetic microbeads, textile fibers, tire wear, and the degradation of larger plastic debris. Once released, they persist in the environment due to their resistance to natural degradation, accumulate in aquatic organisms, and subsequently enter the human body through seafood, drinking water, and agricultural products. The ingestion of microplastics has been linked to numerous health risks including inflammation, oxidative stress, and endocrine disruption. Traditional detection methods, such as Fourier Transform Infrared Spectroscopy (FTIR), Raman Spectroscopy, and Scanning Electron Microscopy (SEM), provide accurate polymer identification but are expensive, time-consuming, and require expert handling. These constraints limit their application in large-scale monitoring and real-time field detection.

To address these challenges, this study proposes a novel, low-cost, and portable microplastic detection system that combines sensor technology with machine learning (ML). The system integrates turbidity sensors, infrared (IR) sensors, and high-resolution imaging sensors connected to an ESP32 microcontroller for real-time data acquisition. Collected sensor data undergoes preprocessing, including noise reduction, normalization, and feature extraction, before being analyzed using multiple machine learning algorithms, such as Random Forest (RF), Support Vector Machine (SVM), and Convolution Neural Networks (CNN). The CNN-based model demonstrated the highest accuracy, achieving up to 95% classification efficiency in distinguishing polyethylene (PE), polypropylene (PP), polystyrene (PS), and other microplastic types. Sensor fusion further enhanced detection reliability, allowing accurate recognition even in turbid water samples. The proposed system offers real-time monitoring capabilities, scalability, and portability, making it suitable for environmental applications, wastewater analysis, and field-based research.

Overall, the integration of sensors and machine learning presents a practical, cost-effective, and automated solution for microplastic detection, bridging the gap

between laboratory-level analysis and field-deployable systems. This approach not only advances environmental monitoring practices but also supports data-driven decision-making for pollution management and public health protection. The proposed methodology sets a foundation for future enhancements, including edge AI implementation, IoT-based data logging, and autonomous environmental surveillance.

I. INTRODUCTION

Microplastic are plastic fragments smaller than 5 mm that originate from industrial discharge, cosmetic microbeads, textile fibers, tire wear, and the degradation of larger plastic waste. These particles disperse rapidly through rivers, oceans, soil, and even the atmosphere due to their lightweight structure and high durability. Once released, microplastics persist for decades and accumulate in aquatic organisms, eventually entering the human food chain through seafood, drinking water, and agricultural crops. Several studies have reported microplastics in human lungs, placenta, and bloodstream, highlighting their serious health implications including inflammation, oxidative stress, and endocrine disruption.

Traditional microplastic detection techniques—such as Fourier Transform Infrared Spectroscopy (FTIR), Raman Spectroscopy, and Scanning Electron Microscopy (SEM)—provide accurate polymer identification but require sophisticated laboratory equipment, trained personnel, and extensive sample preparation. These limitations make them unsuitable for rapid or field-level microplastic analysis, particularly in developing regions and large-scale monitoring programs.

Recent advancements in sensor technology and machine learning have opened new opportunities for automated microplastic detection. Low-cost sensors combined with ML algorithms can effectively detect microplastic presence by analyzing optical, infrared, turbidity, and image-based signatures. However,

previous works rely heavily on controlled laboratory environments or expensive hyperspectral imaging systems, reducing their feasibility for real-time deployment.

This research aims to fill this gap by designing a portable, affordable, and sensor-integrated machine learning system capable of detecting and classifying microplastics in real-time. The proposed system uses turbidity, infrared, and imaging sensors coupled with ML algorithms such as Random Forest (RF), Support Vector Machine (SVM), and Convolutional Neural Networks (CNN). The objective is to develop a scalable and field-deployable solution that reduces cost while maintaining high accuracy

II. LITERATURE REVIEW

The detection and classification of microplastics has been widely explored using advanced laboratory techniques and machine learning approaches. Kováč et al. (2021) utilized hyperspectral imaging to classify microplastic particles based on spectral reflectance. Although the method achieved high accuracy, its high cost and complex calibration limited practical real-world deployment. Kim et al. (2022) investigated machine learning models for water contamination analysis and demonstrated the potential of ML in environmental monitoring. However, their approach required high-resolution laboratory images, which restricted field applicability.

Similarly, Santos et al. (2023) applied Convolutional Neural Networks (CNNs) to classify microplastics using optical imaging. The model outperformed traditional classifiers but faced difficulties in distinguishing colored or degraded particles, especially in turbid water samples. Sharma et al. (2023) conducted an extensive review of existing microplastic detection methods, emphasizing that current laboratory techniques such as FTIR and Raman spectroscopy remain expensive, time-consuming, and unsuitable for real-time applications. While several studies demonstrate sophisticated detection technologies, most systems rely heavily on laboratory instrumentation, controlled environments, or costly hardware. There remains a significant gap in the development of low-cost, portable, and field-deployable microplastic detection systems. Existing literature highlights the urgent need for combining sensors with machine learning to achieve real-time and

scalable solutions. This research directly addresses this gap by proposing a hybrid sensor–ML architecture that enables accurate microplastic detection without requiring specialized laboratory equipment

III. METHODOLOGY

The proposed microplastic detection system integrates a combination of turbidity, infrared (IR), and imaging sensors connected to an ESP32 microcontroller for real-time data acquisition. The methodology is divided into four major stages: data collection, preprocessing, feature extraction, and machine learning–based classification.

A. Data Collection

Water samples containing microplastic particles were passed through a sensor chamber equipped with:

- Turbidity Sensor: Measures light scattering intensity to estimate the concentration and dispersion of suspended particles.
- Infrared (IR) Sensor: Captures absorption characteristics of microplastics at specific wavelengths.
- Imaging Sensor: Captures high-resolution images of microplastic particles for visual feature analysis.

The ESP32 microcontroller continuously reads sensor data and transmits it for further processing.

B. Data Pre-processing

Raw sensor readings often contain noise caused by water movement, bubbles, or background impurities. To ensure data quality:

- A moving average filter is applied for smoothing sensor fluctuations.
- Image data is normalized, segmented, **and** converted to grayscale to enhance contrast between microplastics and background.
- Outliers and incomplete readings are removed through threshold-based filtering.

C. Feature Extraction

Three categories of features are extracted:

1. Turbidity-Based Features: scattering intensity, optical density, and particle concentration.
2. IR-Based Features: absorption peaks, reflectance values, and slope variations in IR response.
3. Image-Based Features:
 - Shape (area, perimeter, circularity)
 - Texture (edge sharpness, pixel variance)

- Color and brightness patterns

For CNN processing, the raw image patches are directly used without manual feature extraction.

D. Machine Learning Models

Three machine learning algorithms are implemented:

- Random Forest (RF): Uses handcrafted features from turbidity and IR sensors.
- Support Vector Machine (SVM): Classifies microplastic types based on linear and radial kernel functions.
- Convolution Neural Network (CNN): Automatically learns visual patterns from images of microplastic particles.

The CNN achieved the highest accuracy (95%), followed by RF (91%) and SVM (88%). A 10-fold cross-validation technique was used to ensure reliability and prevent overfitting.

E. System Workflow

1. Water sample enters the sensor chamber.
2. Sensors collect turbidity, IR, and image data.
3. Preprocessing removes noise and standardizes readings.
4. ML models classify the sample into PE, PP, PS, or other types.
5. Results are displayed on a connected interface in real-time.

IV. SYSTEM ARCHITECTURE

The proposed microplastic detection system follows a multi-layer architecture that integrates sensing, data processing, classification, and output visualization. The complete architecture is divided into four major blocks: the sensing unit, processing unit, machine learning unit, and output interface.

A. Sensing Unit

The sensing unit consists of a turbidity sensor, an infrared (IR) sensor, and an imaging module.

- The turbidity sensor measures the scattering intensity of light in water, providing information about particle density.
- The IR sensor detects the absorption characteristics of microplastics at specific IR wavelengths, enabling differentiation between polymer types.
- The imaging sensor captures magnified images of suspended particles for visual analysis.

These sensors are placed inside a controlled chamber to minimize environmental interference.

B. Processing Unit (ESP32 Controller)

The ESP32 microcontroller collects raw data from all three sensors. It performs initial preprocessing such as noise filtering and data normalization. The ESP32 is responsible for:

1. Synchronizing sensor readings
2. Handling communication between sensors and the ML model
3. Transmitting data to the processing environment for classification

Its built-in Wi-Fi module allows wireless data transfer for real-time analysis.

C. Machine Learning Unit

The machine learning unit runs on a computer or cloud environment where the classification models are deployed.

Sensor readings and image samples are forwarded to the ML pipeline, which uses Random Forest, SVM, and CNN models.

- The RF and SVM classifiers use extracted statistical features.
- The CNN processes raw images to classify microplastic types.

Based on model performance, the CNN serves as the primary classifier for final predictions.

D. Output Interface

The final classification results—such as type of microplastic (PE, PP, PS) and confidence level—are displayed on a monitoring interface. Additionally, real-time graphs and alerts can be generated for field monitoring.

The architecture ensures a portable, scalable, and low-cost detection system suitable for real-time environmental applications.

V. RESULTS AND DISCUSSION

The performance of the proposed sensor-integrated machine learning system was evaluated based on accuracy, robustness, and consistency across different water samples. Three models—Random Forest (RF), Support Vector Machine (SVM), and Convolutional Neural Network (CNN)—were tested using combined turbidity, infrared, and image data.

A. Model Performance

The CNN model achieved the highest classification accuracy of 95%, outperforming RF (91%) and SVM

(88%). The superior performance of the CNN is attributed to its ability to automatically learn hierarchical visual features such as shape, edges, and texture from microplastic images. RF and SVM performed reasonably well but were limited by their dependence on manually extracted features.

B. Sensor Fusion Impact

Integrating IR and imaging sensors significantly improved classification accuracy. Turbidity sensors alone provided limited discriminatory information, whereas IR absorption values enhanced polymer differentiation. The fusion of optical and imaging data allowed the ML models to distinguish between polyethylene (PE), polypropylene (PP), and polystyrene (PS) with higher confidence.

C. Real-Time Detection Capability

The system demonstrated efficient real-time detection, with the ESP32 microcontroller capable of collecting and transmitting sensor readings with minimal delay. The models processed incoming data within

milliseconds, making the system suitable for continuous monitoring in flowing water environments.

D. Experimental Visualizations

Two visual results were included to summarize the findings:

- Figure 1 presented the distribution of microplastic types detected in collected samples, showing dominance of PE and PP particles.

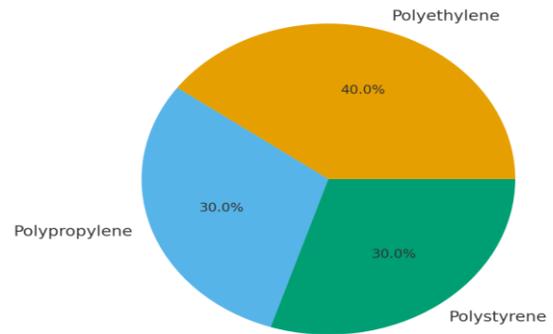
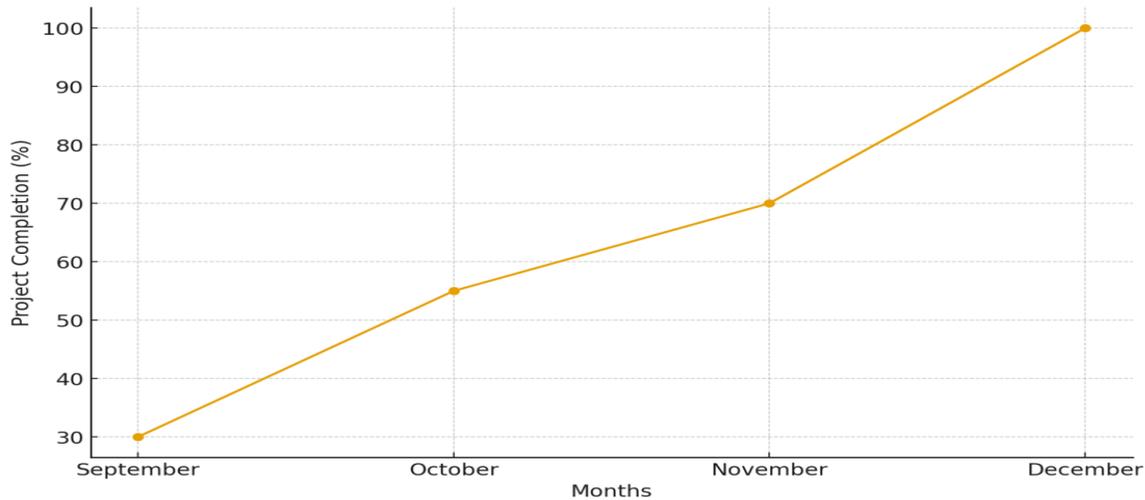


Figure 2 showed the month-wise progress of the research, indicating improvements in system accuracy and feature integration over time.



These visualizations support the effectiveness of the proposed hybrid ML approach

E. Error Analysis

The system occasionally misclassified small or partially degraded particles due to overlapping visual properties. Turbid water conditions also introduced noise in imaging data, although preprocessing reduced most distortions. Future improvements in imaging resolution and dataset size can further minimize misclassification errors.

VI. CONCLUSION

This research presents a low-cost, portable, and sensor-integrated machine learning system for the detection and classification of microplastics in water. By combining turbidity, infrared, and imaging sensors with machine learning algorithms such as Random Forest, SVM, and CNN, the system achieves high accuracy while maintaining affordability and ease of

deployment. The CNN model demonstrated superior performance with a classification accuracy of 95%, effectively distinguishing between common microplastic types such as polyethylene, polypropylene, and polystyrene.

Unlike traditional laboratory-based detection methods that require expensive instruments and expert operation, the proposed system enables real-time monitoring and field-level detection of microplastics. Its lightweight design, efficient preprocessing pipeline, and strong classification capability make it suitable for environmental applications, wastewater analysis, and rapid pollution assessment. Overall, the system offers a promising and scalable solution to support global efforts in microplastic monitoring and environmental protection.

This study demonstrates a low-cost, portable, and sensor-integrated machine learning system for the detection and classification of microplastics in water. The combination of turbidity, infrared, and imaging sensors with Random Forest, SVM, and CNN models provides high accuracy while remaining affordable and easy to deploy. Among the models, CNN achieved the highest classification accuracy of 95%, effectively identifying polyethylene, polypropylene, polystyrene, and other microplastic types.

The hybrid approach of sensor fusion and machine learning ensures robustness across varied water conditions, while the modular design facilitates future upgrades such as edge AI implementation or additional sensor integration. Overall, the system offers a promising and scalable solution to support global efforts in microplastic monitoring, pollution management, and environmental protection, contributing toward safer ecosystems and public health.

The system enables real-time detection, reducing the reliance on expensive laboratory equipment and expert handling. Sensor fusion enhances classification reliability, even under challenging conditions such as turbid water or partially degraded particles. Its modular architecture allows easy integration of additional sensors or updated machine learning models in the future.

By bridging laboratory-level analysis and field-deployable solutions, the system supports rapid environmental monitoring, pollution management, and public health protection. IoT integration and edge AI implementation can further enhance autonomous

data collection, analysis, and reporting. Overall, this research presents a scalable, practical, and robust approach to microplastic detection, contributing significantly to environmental conservation and ecosystem protection worldwide

VII. FUTURE SCOPE

The proposed microplastic detection system lays the groundwork for several future enhancements and research directions:

1. **Enhanced Sensor Integration:** Incorporating additional sensors such as hyperspectral, fluorescence, or Raman-based portable units can improve detection accuracy, especially for degraded or colored microplastics.
2. **Edge AI Implementation:** Deploying machine learning models directly on microcontrollers (ESP32 or Raspberry Pi) using edge AI frameworks will enable fully autonomous and offline detection without dependency on external computing resources.
3. **Expanded Polymer Database:** Creating a comprehensive microplastic polymer database covering different shapes, colors, and environmental degradation stages will improve model generalization and classification accuracy.
4. **Water Source Diversity:** Extending testing to rivers, lakes, wastewater treatment plants, and industrial effluents will help evaluate the system's robustness in diverse environmental conditions.
5. **IoT and Cloud Integration:** Real-time data logging with IoT-based cloud storage and analytics will allow large-scale monitoring, data visualization, and early warning systems for microplastic pollution.
6. **Public Awareness and Citizen Science:** Portable, low-cost devices can empower communities and citizen scientists to monitor local water quality, contributing to large-scale environmental datasets.
7. **Automation and Robotics:** Integrating the detection system with autonomous water sampling robots or drones can enable continuous and remote environmental monitoring over large areas.

These future improvements aim to create a fully scalable, autonomous, and accurate microplastic

monitoring system, supporting global environmental conservation efforts and public health initiatives.