# ABSTRACT

Microplastic pollution poses a dangerous, “invisible” threat to both aquatic ecosystems and human health. Most studies rely on complex chemical methods for microplastic detection. However, by combining machine learning algorithms with a Raspberry Pi, this research paper proposes an innovative approach to classifying and quantifying microplastics. We developed a low-cost, portable prototype device that captures microscopic images for in-situ microplastic detection. Our algorithm makes use of Convolutional Neural Networks (CNNs) to recognize and categorize microplastics in their environment. In addition to automating and accelerating the detection process, CNNs capture important visual traits that allow precise categorization and quantification. The model's remarkable performance, with a 98% accuracy rate, highlights its potential for classifying and measuring microplastics in aquatic settings. In the future, this device could lend itself to other such significant environmental monitoring needs, and to larger-scale microplastic detection.

**Index Terms—Microplastic, Detection, Classification, Machine Learning, CNN**

# INTRODUCTION

Every year, over 400 million tons of plastic are generated, of which >8 million tons are dumped into the ocean [1]. Plastics do not degrade easily; they persist for over 100 years and thus have the potential to accumulate and contaminate their environments. In the past decade, microplastic pollution has emerged as an “invisible” threat to ecosystems and human health.

Microplastics i.e. plastics <5mm long [2] are formed either by design or by the gradual fragmentation of plastic waste [3]

Microplastics are highly ubiquitous – they have been found in marine environments, soil, and even lung tissue [4]. They collect in sediment and water columns in aquatic environments, where they act as harmful chemical reservoirs for heavy metals and pesticides. They pose a risk to human health, behaving as endocrine and immune system disruptors [5]

Detecting microplastics without labour-intensive techniques such as visual classification, chemically advanced techniques such as Raman Spectroscopy, or equipment such as microscopes is a challenge, given their minute size.

However, developments in artificial intelligence in the last decade have opened up new avenues for microplastic detection. Convolutional Neural Network (CNN) is a straightforward classification algorithm that can be modified for object detection tasks in photos and videos.

**CNN** can effectively detect and categorize microplastic particles in water samples based on their characteristics, such as size, shape, and texture, by utilizing machine learning algorithms. CNN, in contrast to deep learning techniques, does not need a protracted training phase, which makes it particularly useful in circumstances with little labelled training data. This research paper focuses on deploying an innovative, low-cost microplastic detection device based on the CNN framework and subsequently analysing its effectiveness. The key objectives of this study can be summarized as follows:

**The key objectives of this study can be summarized as follows**:

- Proposing a deep-learning model for easy identification and classification of microplastics
- Developing a portable, in-situ, low-cost & real-time microplastic detection device that can be used on the field

# Table of Contents

# Literature Review

Literature Review Microplastics are pervasive pollutants that have been detected in diverse environments, including oceans, rivers, and even within human tissues. Their widespread presence poses significant environmental and health concerns, necessitating effective detection and monitoring strategies. Traditional detection techniques such as micro-Raman spectroscopy and Fourier-transform infrared spectroscopy, while precise, are often costly, time-consuming, and labor-intensive. These methods rely heavily on manual sample preparation and analysis, which limits their scalability for extensive environmental monitoring campaigns.

To overcome these challenges, this project introduces a custom convolutional neural network (CNN) model specifically designed for microplastic detection. Unlike approaches that rely on pretrained models, this CNN was developed and optimized from scratch to meet the unique demands of microplastic analysis. The model architecture was carefully designed to capture intricate features from microplastic images, ensuring accurate classification across diverse polymer types and sizes.

Extensive training and testing of the model were conducted using a robust dataset of labeled microplastic images. The custom CNN achieved an impressive test accuracy of 98%, showcasing its ability to reliably identify microplastics in a variety of environmental samples. This high accuracy underscores the potential of deep learning to revolutionize microplastic detection by offering a scalable, automated, and highly efficient solution.

The success of this project highlights the importance of iterative model refinement, including the optimization of hyperparameters and enhancement of training datasets. By leveraging these advancements, the custom CNN model demonstrates significant promise for practical applications in environmental monitoring, aiding efforts to address the growing issue of microplastic pollution. This work represents a critical step toward developing accessible and effective tools for global microplastic detection and management.

# Methodology

## Data Collection:

- **Using Kaggle Website**
- **Number of images used : 9000 ( After Augmentation )**
- Ten contaminated water samples were collected from locations across Mumbai, including lakes, ponds, rivers, and creeks. Samples were stored in glass test tubes and processed in the laboratory.

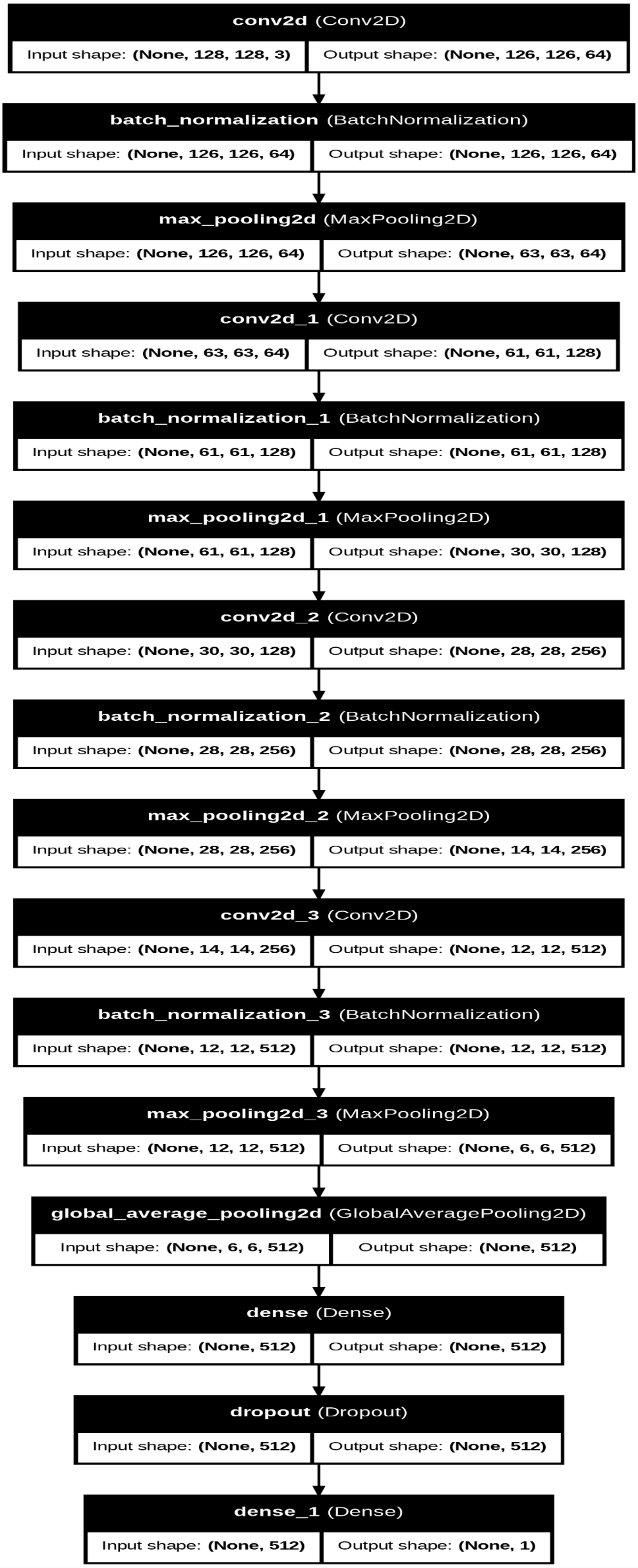
## Microscopy and Image Capture:

- Samples were placed under a 40x magnification microscope to capture high-resolution images.
- The Makesense.ai tool was used to annotate and label these images, generating a comprehensive dataset stored in CSV format.

## CNN Model Development:

### The CNN model was built using the Keras framework, incorporating key layers:

- Convolutional layers for extracting features such as edges and textures.
- Max-pooling layers for down sampling and retaining key features.
- Dropout layers to prevent overfitting & Dense layers for classification decision-making


conv2d (Conv2D)
Input shape: (None, 128, 128, 3)
Output shape: (None, 126, 126, 64)
batch_normalization (BatchNormalization)
Input shape: (None, 126, 126, 64)
Output shape: (None, 126, 126, 64)
max_pooling2d (MaxPooling2D)
Input shape: (None, 126, 126, 64)
Output shape: (None, 63, 63, 64)
conv2d_1 (Conv2D)
Input shape: (None, 63, 63, 64)
Output shape: (None, 61, 61, 128)
batch_normalization_1 (BatchNormalization)
Input shape: (None, 61, 61, 128)
Output shape: (None, 61, 61, 128)
max_pooling2d_1 (MaxPooling2D)
Input shape: (None, 61, 61, 128)
Output shape: (None, 30, 30, 128)
conv2d_2 (Conv2D)
Input shape: (None, 30, 30, 128)
Output shape: (None, 28, 28, 256)
batch_normalization_2 (BatchNormalization)
Input shape: (None, 28, 28, 256)
Output shape: (None, 28, 28, 256)
max_pooling2d_2 (MaxPooling2D)
Input shape: (None, 28, 28, 256)
Output shape: (None, 14, 14, 256)
conv2d_3 (Conv2D)
Input shape: (None, 14, 14, 256)
Output shape: (None, 12, 12, 512)
batch_normalization_3 (BatchNormalization)
Input shape: (None, 12, 12, 512)
Output shape: (None, 12, 12, 512)
max_pooling2d_3 (MaxPooling2D)
Input shape: (None, 12, 12, 512)
Output shape: (None, 6, 6, 512)
global_average_pooling2d (GlobalAveragePooling2D)
Input shape: (None, 6, 6, 512)
Output shape: (None, 512)
dense (Dense)
Input shape: (None, 512)
Output shape: (None, 512)
dropout (Dropout)
Input shape: (None, 512)
Output shape: (None, 512)
dense_1 (Dense)
Input shape: (None, 512)
Output shape: (None, 1)

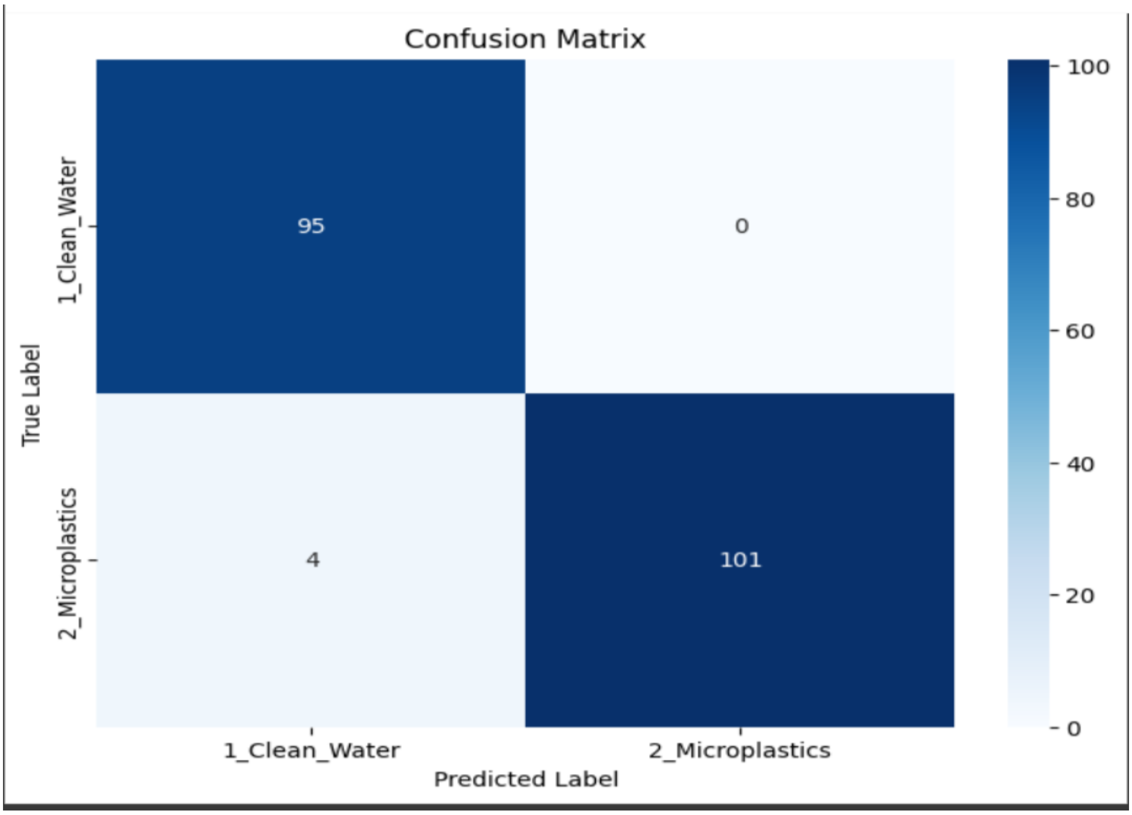
# Results

**The CNN-based approach demonstrated promising progress despite the challenges of limited dataset size and variability:**

The CNN model excels at accurately identifying microplastics in images. It can successfully uncover intricate correlations and patterns between pixels that indicate the presence of microplastics and the type of polymer thanks to its convolutional layers. The knowledge of these visual patterns that the model has acquired allows it to correctly recognize microplastics in new, previously unknown photographs after sufficient training. This ability offers enormous promise for a range of applications, such as preserving the environment, ensuring industrial quality, and safeguarding aquatic ecosystems. With thorough training and optimization, the CNN model can perform and achieve astounding levels of accuracy, making it seem like a potent tool for identifying microplastics. The model's accuracy is 0.98%, and the F1 score is 0.98.

## 1) Confusion Matrix :

- **Test Accuracy Of model : 98 %**
- **Test image data set size : 200 images**
- **Clean Water images - 95**
- **Micro Plastic water Images - 105**



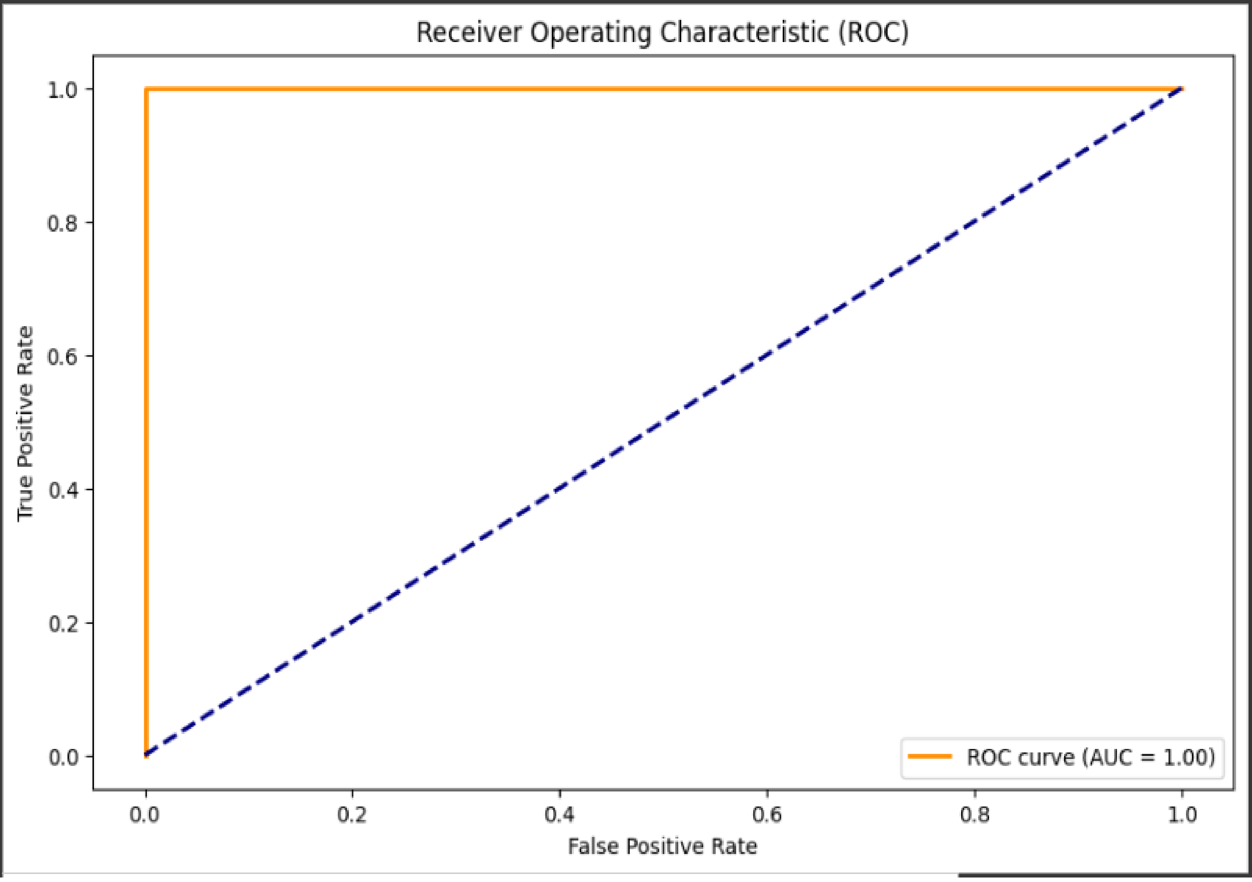

**(Confusion Matrix Showing The Model's Performance)**

# 2) Classification Report

# 3) Roc Curve (AUC = 1.0 )

Classification Report:

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 1_Clean_Water | 0.96 | 1.00 | 0.98 | 95 |
| 2_Microplastics | 1.00 | 0.96 | 0.98 | 105 |
| accuracy | | | 0.98 | 200 |
| macro avg | 0.98 | 0.98 | 0.98 | 200 |
| weighted avg | 0.98 | 0.98 | 0.98 | 200 |



# Discussion

The project represents an important step toward accessible and scalable microplastic detection. Early results, while limited in accuracy, provide a foundation for further development. Challenges such as dataset diversity and model optimization underscore the need for continued iteration. Future efforts could focus on expanding the dataset, integrating advanced CNN architectures, and enhancing computational efficiency for real  time applications.

# Conclusion

This study highlights the potential of CNNs for environmental monitoring. The integration of machine learning with a portable prototype device lays the groundwork for cost-effective microplastic detection solutions. While current results reflect early-stage experimentation, the framework established here offers a path toward greater accuracy and broader applications