Image Processing for Earth Observation

# - Marine Debris Detection -

*Authors:*
Evangelista Santiago
Ethvignot Marie
Jousson Kento

*Professors:*
Prof. Tuia Devis

April 24, 2025

# Contents

# 1 Introduction

## 1.1 Context and Importance

Plastic pollution in marine and coastal ecosystems is a critical environmental and social challenge. It poses severe threats to biodiversity, disrupts ecosystem functions, and impacts human livelihoods dependent on marine resources. Floating marine debris, particularly plastic waste, aggregates into visible structures known as windrows, which can span widths of over 50 meters and lengths exceeding 500 meters. Timely and accurate detection of these debris formations is crucial for coordinating effective cleanup efforts and mitigating their environmental impact. Advances in satellite technology, such as the Sentinel-2 optical sensors with a resolution of 10 to 20 meters, have made it possible to monitor these windrows at scale, opening opportunities for automated detection using machine learning techniques.

## 1.2 Objective of the Project

The primary goal of this project is to develop a robust image classification model capable of identifying the presence of floating marine debris in Sentinel-2 satellite images. The classifier will take advantage of 12-band spectral data to distinguish between debris-laden and debris-free regions. By automating this detection process, the project aims to improve the monitoring capabilities and support targeted intervention strategies in marine environments.

## 1.3 Dataset Description

The project uses a data set consisting of Sentinel-2 image patches, each measuring 32 $\times$ 32 pixels. These patches are labeled with a binary classification: "0" indicating the absence of floating debris and "1" indicating its presence. The dataset comprises:

Training set: 53,535 image patches

Validation set: 7,436 image patches

Test set: 13,386 image patches

The validation and test sets have an equal distribution of positive and negative labels to ensure a balanced performance evaluation. However, the training set is unbalanced, with relatively fewer positive examples. This imbalance presents a key challenge for model training and requires careful consideration of sampling strategies or class weighting techniques.

## 1.4 Challenges

Several challenges must be addressed to achieve the project objectives:

Imbalanced Dataset: The training set has a disproportionately low number of positive examples, which can lead to biased model performance. The effective handling of this imbalance is critical for robust classification.

Spectral Band Sensitivity: The performance of the model will be analyzed in three scenarios: using all 12 spectral bands, using only the RGB+NIR bands, and using only the visible spectrum. This analysis will provide insight into the sensitivity and utility of different spectral inputs.

Labeling Complexity: The data set is derived from a segmentation data set with annotations per pixel. A patch is labeled positive if it contains at least one positive pixel, making classification challenging for images where floating debris occupies only a small fraction of the area.

## 2 Methods

### 2.1 Dataset Preparation

The dataset used in this project comprises Sentinel-2 image patches stored as serialized .pkl files. To facilitate efficient data handling and preprocessing, a custom dataset class, MarineDebrisDataset, was developed using PyTorch's torch.utils.data.Dataset. This class enables seamless integration with PyTorch's DataLoader, which manages batch loading and data shuffling during training and evaluation.

The dataset class reads each .pkl file, extracting the image and corresponding label. Each file contains two elements: a $32 \times 32$ pixel image patch and a binary label indicating the presence or absence of floating debris. This modular approach ensures that the data pipeline remains efficient and scalable across varying dataset sizes.

### 2.2 Data Augmentation and Preprocessing

To improve the model's generalization ability, several data augmentation techniques were applied to the training set. Using the torchvision.transforms library, the following transformations were implemented:

Random Horizontal Flip: Introduces variability by flipping images along the horizontal axis.

Random Vertical Flip: Mirrors images vertically, further diversifying the dataset.

Normalization: Standardizes pixel values to a fixed mean and standard deviation, ensuring consistent input for the model.

These augmentations reduce overfitting by exposing the model to a wider range of scenarios during training. The transformations are applied dynamically during data loading, ensuring that the original dataset remains unaltered.

### 2.3 Handling Class Imbalance

The dataset exhibits a significant imbalance, with relatively few positive examples compared to negative ones. To address this, a WeightedRandomSampler was employed. This sampler assigns higher sampling weights to positive samples, ensuring their adequate representation in each training batch. By mitigating the effects of class imbalance, this approach enhances the model's ability to learn meaningful patterns from minority class examples.

### 2.4 Neural Network Architecture

The core of the classification task relies on a Convolutional Neural Network (CNN), designed using PyTorch's torch.nn module. CNNs are well-suited for image classification tasks due to their ability to capture spatial hierarchies and extract relevant features from image data. While the specific architecture details are modular, the network includes:

Convolutional layers to extract feature maps.

Activation functions (e.g., ReLU) to introduce non-linearity.

Pooling layers to reduce spatial dimensions and focus on salient features.

Fully connected layers that map extracted features to binary classification outputs.

The output layer uses a sigmoid activation function to produce probabilities indicating the presence or absence of floating debris.

### 2.5 Loss Function and Optimization

The binary classification task employs the Binary Cross-Entropy Loss (BCE Loss) as the objective function. This loss measures the discrepancy between predicted probabilities and true binary labels. For optimization, the Adam optimizer is utilized due to its efficiency in handling sparse gradients and adaptability across learning rates. Both the loss function and optimizer are critical components in minimizing classification errors and achieving convergence.

### 2.6 Training and Evaluation

The training loop involves iteratively processing batches of data, with each iteration consisting of:

Forward Pass: Computing predictions from the model.

Loss Computation: Measuring the difference between predictions and ground truth labels using BCE Loss.

Backward Pass: Calculating gradients and updating model parameters using the optimizer.

Evaluation is conducted on the validation and test sets using metrics such as accuracy, precision, recall, and F1-score. These metrics provide a comprehensive view of the model's performance, particularly in handling imbalanced classes.

### 2.7 Spectral Band Analysis

To assess the impact of different spectral bands on classification performance, three experimental setups were implemented:

All 12 Spectral Bands: Utilizing the complete spectral information available in Sentinel-2 images.

RGB + NIR Bands: Focusing on the visible spectrum and near-infrared band, which are commonly used in remote sensing applications.

Visible Spectrum Only: Restricting inputs to the RGB bands to evaluate the model's capability using minimal spectral data.

This analysis provides insights into the sensitivity and utility of various spectral inputs, aiding in the design of efficient and effective remote sensing solutions.

# 3 Results

## 3.1 12 Bands

**Table 3.1:** *Performance Metrics and Confusion Matrix for 12-Band Case*

| Metric | Value |
|---|---|
| Accuracy | 0.9221 |
| Precision | 0.8871 |
| Recall | 0.9673 |
| F1 Score | 0.9255 |

**Confusion Matrix**

| | Predicted Negative | Predicted Positive |
|---|---|---|
| **Actual Negative** | 0.8769 | 0.1231 |
| **Actual Positive** | 0.0327 | 0.9673 |

## 3.2 RGB+NIR Bands

**Table 3.2:** *Performance Metrics and Confusion Matrix for RGB+NIR Bands case*

| Metric | Value |
|---|---|
| Accuracy | 0.9121 |
| Precision | 0.8770 |
| Recall | 0.9586 |
| F1 Score | 0.9160 |

**Confusion Matrix**

| | Predicted Negative | Predicted Positive |
|---|---|---|
| **Actual Negative** | 0.8655 | 0.1345 |
| **Actual Positive** | 0.0414 | 0.9586 |

## 3.3 Visible Bands

Table 3.3: *Performance Metrics and Confusion Matrix for RGB Band Case*

| Metric | Value |
|---|---|
| Accuracy | 0.9009 |
| Precision | 0.8662 |
| Recall | 0.9482 |
| F1 Score | 0.9053 |

**Confusion Matrix**

| | Predicted Negative | Predicted Positive |
|---|---|---|
| **Actual Negative** | 0.8536 | 0.1464 |
| **Actual Positive** | 0.0518 | 0.9482 |

# 4 Conclusion

This project successfully developed a classification model to detect floating marine debris using Sentinel-2 imagery, addressing a significant environmental issue. The Convolutional Neural Network (CNN) achieved strong performance in metrics such as accuracy, precision, recall, and the F1 score, showcasing its effectiveness in distinguishing debris-laden from debris-free regions. Techniques like data augmentation and weighted sampling helped address challenges such as data imbalance and sparse debris in some images.

The results indicate that the 12-band model outperforms the other setups in detecting floating marine debris, achieving the best balance between accuracy, precision, recall, and F1-score. The RGB+NIR setup shows slightly lower performance but still benefits from the additional near-infrared data, while the visible bands setup performs the weakest, suggesting that relying only on RGB limits the model's effectiveness. The confusion matrices reveal that the 12-band model has the lowest false negative rate, meaning it identifies most debris instances, which is critical for environmental monitoring. However, the matrices also show a noticeable false positive rate across all setups, indicating the model occasionally misclassifies debris-free patches, which could impact the efficiency of cleanup efforts. Overall, leveraging multiple spectral bands proves essential for enhancing detection performance and minimizing undetected debris.

By automating marine debris detection, this approach enables scalable and timely monitoring, supporting targeted cleanup efforts. Future work could focus on refining the model for complex scenarios and expanding its application to other forms of environmental pollution. This study contributes to global efforts to conserve marine ecosystems and pave the way for healthier oceans.

# 5 Sources

[1] Image Processing for Earth Observation - Prof. Devis Tuia.
[2] ChatGPT
[3] Github Copilot