

Sistema di detection di rifiuti plastici su specchi d'acqua

De Ramundo Marco - Perani Xavier

Agosto 2024

Abstract

Obiettivo: ottenere un modello che possa individuare e classificare oggetti sulla superficie di uno specchio d'acqua come un fiume o un lago per poter riconoscere e individuare rifiuti plastici come bottiglie e buste di plastica.

Introduzione

L'inquinamento ambientale dovuto ai rifiuti plastici è un problema che influisce tutti noi, nella fattispecie per quanto riguarda la salvaguardia degli oceani. Le due principali fonti di inquinamento di plastica negli oceani deriva dalla pesca e dai rifiuti gettati nei fiumi. Per poter contrastare la seconda causa è necessario poter effettuare operazioni di monitoraggio per prevenire che rifiuti dannosi possano raggiungere il mare mentre per la prima occorrono strumenti che rendano il recupero dei rifiuti più semplice e meno dispendioso. Anche solo riconoscere, identificare e individuare i rifiuti darebbe una grossa mano allo scopo. Per questo motivo può aver senso adoperare soluzioni quali modelli convoluzionali per effettuare detection degli oggetti in questione per poi agire di conseguenza con la raccolta e la rimozione dall'acqua della plastica.

Dataset e modello

Il dataset scelto per il nostro scopo è `kili-technology/plastic_in_river` che raccoglie immagini di laghetti e bacini d'acqua con oggetti e/o rifiuti in

superficie. Il dataset è stato creato per la Kili's Community Challenge, sfida nata per invogliare la community di sviluppatori a sfruttare i modelli di deep learning per contrastare il problema dell'inquinamento degli oceani. La sfida si è svolta nel febbraio 2022 e il dataset è rimasto accessibile pubblicamente anche dopo la fine dell'evento¹.

Il dataset è composto da un totale di 4259 immagini, già divisi in tre sottoinsiemi per training, validazione e test. Nella tabella 1 è possibile vedere la ripartizione delle istanze per sottoinsieme. Ogni immagine ha anche un file di testo con la lista degli oggetti presenti all'interno, classe dell'oggetto e posizione spaziale rispettando il data format usato da YOLO.

Subset	Numero istanze
Train	3407
Val	425
Test	427
Totale	4259

Tabella 1: Distribuzione delle immagini negli insiemi di train, val e test

Le categorie a disposizione per poter classificare gli oggetti identificati sono 4 e sono:

- 0 : PLASTIC_BAG
- 1 : PLASTIC_BOTTLE
- 2 : OTHER_PLASTIC_WASTE
- 3 : NOT_PLASTIC_WASTE

La distribuzione delle classi tra le istanze non è uniforme ed è uno dei possibili problemi da affrontare durante l'addestramento del modello. Il rischio è quello di avere difficoltà nel riconoscimento delle categorie che sono meno rappresentate all'interno del dataset. Nella tabella 2 è possibile vedere il numero di istanze per classe all'interno dei vari subsets.

Le dimensioni delle immagini sono varie ma generalmente di qualità alta: la dimensione in larghezza supera i 1000 pixel e gli 800 pixel in altezza.

¹Attualmente (luglio-agosto 2024) si riscontrano problemi nell'acquisizione del dataset tramite API. Probabilmente il proprietario del dataset ha spostato o tolto il dataset dal proprio server. Per il momento non si sa se il dataset tornerà a disposizione senza problemi

Classe	Train	Val	Test
PLASTIC_BAG	1250	167	85
PLASTIC_BOTTLE	6276	785	854
OTHER_PLASTIC_WASTE	3345	296	122
NOT_PLASTIC_WASTE	1414	212	111

Tabella 2: Distribuzione delle istanze all'interno del dataset

Nella immagine 1 è possibile in ordine vedere la distribuzione delle istanze nel dataset, la grandezza dei box per istanza all'interno dell'immagine, la distribuzione delle coordinate x e y del box nell'immagine, infine la grandezza del box per istanza.

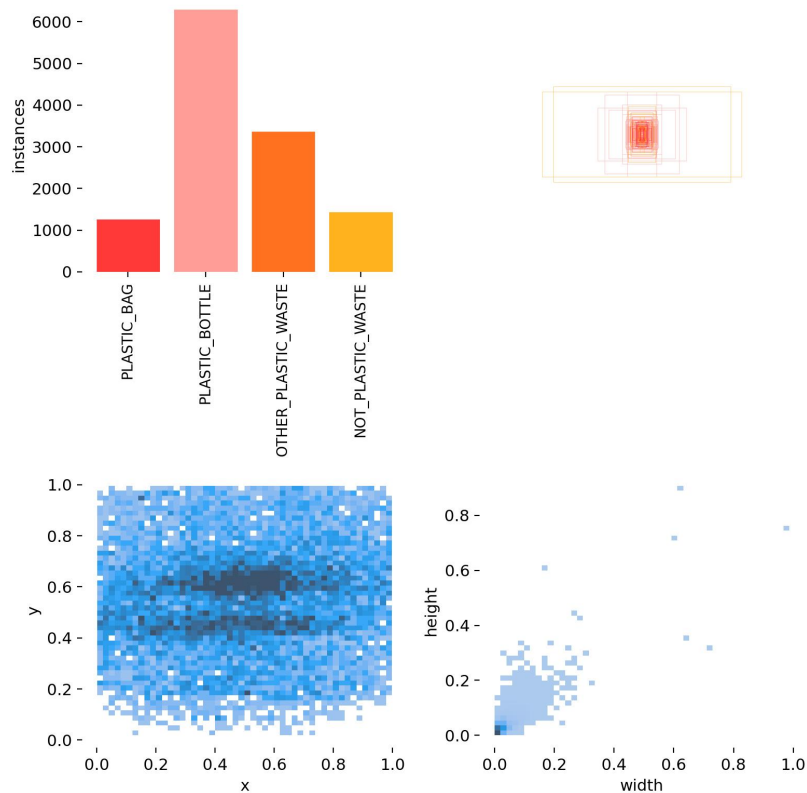


Figura 1: Caratteristiche istanze nel dataset

Da notare come la maggior parte delle istanze hanno una dimensione della box rispetto all'immagine ridotta, in genere inferiore al 20% della larghezza e/o altezza dell'immagine. In genere la maggior parte delle istanze, essendo rifiuti sulla superficie dell'acqua, sono concentrati oriz-

zontalmente nella parte centrale dove è plausibile trovare l'orizzonte dello specchio idrico.

La dimensione ridotta del box potrebbe essere utile da tenere in considerazione per quanto riguarda l'iperparametro `imgsz`, image size, ovvero le dimensioni delle immagini in ingresso passate al modello.

1 Descrizione del Modello YOLOv8 per Object Detection

YOLO (You Only Look Once) è una delle architetture di deep learning più utilizzate per il rilevamento degli oggetti (object detection) in tempo reale. La versione 8 di YOLO, denominata YOLOv8, rappresenta un ulteriore miglioramento rispetto alle versioni precedenti, combinando un'elevata accuratezza con una velocità di esecuzione che la rende ideale per applicazioni in scenari complessi e dinamici, come il riconoscimento dei rifiuti plastici nei bacini idrici.

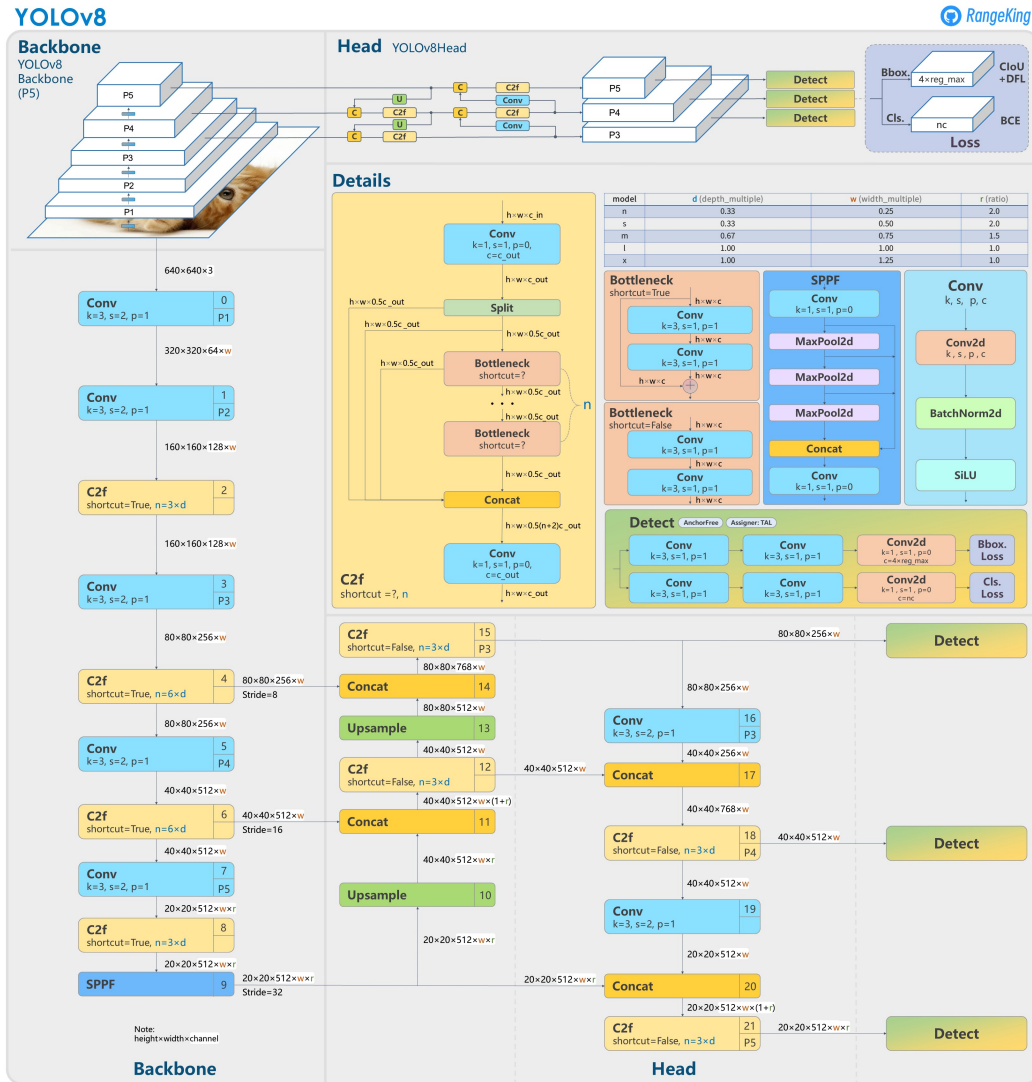
1.1 Panoramica di YOLOv8

YOLOv8 è progettato per fornire un bilanciamento ottimale tra precisione e velocità. La sua architettura è costruita su principi consolidati delle reti convoluzionali, ma con innovazioni che ne migliorano l'efficienza e la capacità di generalizzazione. In particolare, YOLOv8 sfrutta blocchi convoluzionali più leggeri e tecniche di ottimizzazione avanzate che riducono il tempo di inferenza senza sacrificare l'accuratezza.

1.2 Architettura di YOLOv8

L'architettura di YOLOv8 si basa su un approccio end-to-end che suddivide l'immagine in una griglia e applica convoluzioni multiple per predire le bounding box e le classi degli oggetti all'interno di ciascuna cella della griglia. YOLOv8 utilizza tecniche avanzate come la *Path Aggregation Network* (PANet) per migliorare l'integrazione delle informazioni a diversi livelli di profondità della rete, il che è fondamentale per rilevare oggetti di varie dimensioni e scale. Di seguito è riportato un diagramma semplificato dell'architettura:

Il modello YOLOv8 incorpora anche una *Neck* e una *Head* ottimizzati per massimizzare la capacità di rilevamento attraverso una migliore aggregazione delle caratteristiche e una maggiore flessibilità nelle predizioni



cisione, risultando particolarmente adatte per l'esecuzione su hardware con risorse limitate o in contesti dove è necessaria l'elaborazione in tempo reale.

- **YOLOv8 Small:** Questa versione è progettata per situazioni in cui la velocità è critica e le risorse hardware sono limitate. YOLOv8 Small utilizza un numero ridotto di filtri convoluzionali e strati, il che riduce il numero di parametri complessivi e, di conseguenza, il tempo di inferenza. Nonostante la sua leggerezza, mantiene un'accuratezza adeguata per compiti di object detection su scenari meno complessi o con immagini di bassa risoluzione. È ideale per applicazioni che richiedono una rapida risposta e dove la potenza computazionale è limitata.
- **YOLOv8 Medium:** La versione Medium di YOLOv8 rappresenta un'interessante via di mezzo tra precisione e velocità. Questa architettura è più complessa rispetto alla variante Small, con un numero maggiore di strati e parametri, che consente al modello di catturare dettagli più fini e di lavorare meglio su dataset più complessi o con immagini ad alta risoluzione. È particolarmente indicata per scenari dove un compromesso tra velocità e accuratezza è accettabile, garantendo comunque la capacità di rilevare oggetti in tempo reale.

1.4 Vantaggi di YOLOv8 per il Riconoscimento dei Rifiuti Plastici

La scelta di YOLOv8 per il riconoscimento dei rifiuti plastici nei fiumi è giustificata da diversi fattori:

- **Rilevamento in Tempo Reale:** YOLOv8 è noto per la sua capacità di eseguire object detection in tempo reale, il che è cruciale per applicazioni di monitoraggio continuo nei fiumi. La possibilità di rilevare e classificare i rifiuti plastici in tempo reale consente interventi immediati, riducendo il rischio che i rifiuti possano propagarsi ulteriormente.
- **Efficienza Computazionale:** Le versioni Small e Medium di YOLOv8 sono ottimizzate per girare su hardware con risorse limitate, come droni o sistemi embedded. Questo le rende particolarmente adatte per applicazioni sul campo, dove la potenza di calcolo potrebbe essere un vincolo.

- **Accuratezza Elevata:** Nonostante la sua velocità, YOLOv8 mantiene un'accuratezza competitiva grazie all'uso di tecniche avanzate come l'addestramento con *label smoothing* e l'adozione di una griglia più fine per la previsione delle bounding box. Questo è essenziale per distinguere efficacemente i rifiuti plastici da altri detriti naturali presenti nei fiumi.
- **Robustezza in Ambienti Variabili:** YOLOv8 è progettato per funzionare in diverse condizioni di illuminazione e in presenza di rumore visivo, caratteristico di ambienti naturali come i fiumi. La capacità del modello di generalizzare bene a scenari non ideali è un vantaggio chiave per garantire il successo del sistema di rilevamento.

1.5 Considerazioni Finali

In conclusione, YOLOv8, nelle sue varianti Small e Medium, rappresenta una scelta ottimale per il progetto di riconoscimento dei rifiuti plastici nei fiumi. La combinazione di efficienza computazionale, accuratezza e capacità di operare in tempo reale rende YOLOv8 particolarmente adatto per implementazioni sul campo. L'architettura del modello, progettata per massimizzare le prestazioni in contesti reali, assicura che il sistema sia in grado di operare efficacemente in ambienti complessi e variabili come quelli fluviali.

2 Metriche di Valutazione delle Performance del Modello

Nell'ambito del riconoscimento automatico dei rifiuti plastici nei fiumi utilizzando la rete convoluzionale YOLO (You Only Look Once), l'accurata valutazione delle performance del modello è cruciale. Il contesto operativo, caratterizzato da variabilità ambientali e la presenza di oggetti confondenti, rende fondamentale l'impiego di metriche di valutazione che forniscano una visione completa dell'efficacia del modello. In questa sezione, discuteremo in dettaglio le metriche principali utilizzate per valutare il modello, giustificando la loro rilevanza rispetto all'obiettivo specifico di riconoscimento dei rifiuti plastici.

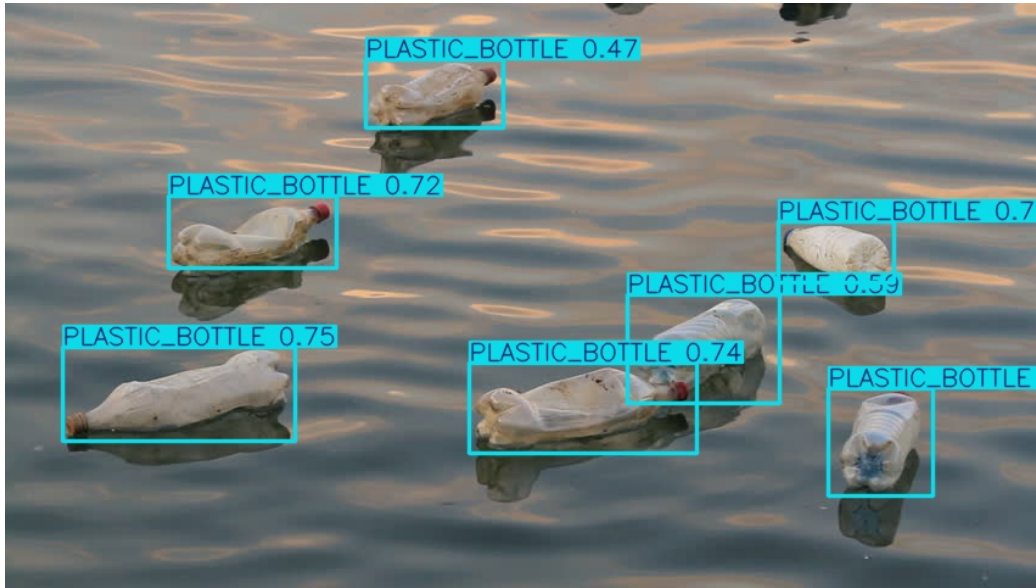


Figura 3: Esempio di output di YOLOv8 per il rilevamento di rifiuti plastici nei fiumi.

2.1 Precision

La *Precision* è una metrica fondamentale nel contesto del riconoscimento dei rifiuti plastici. Questa misura indica la percentuale di rifiuti classificati correttamente tra tutti quelli identificati come rifiuti dal modello. In altre parole, la precision ci dice quanto possiamo fidarci delle predizioni positive effettuate da YOLO. La formula per calcolarla è la seguente:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

dove:

- *TP* (True Positive): il numero di rifiuti plastici correttamente rilevati dal modello;
- *FP* (False Positive): il numero di oggetti non plastici erroneamente identificati come rifiuti plastici.

Nel contesto dei fiumi, dove possono essere presenti detriti naturali o altre forme di inquinamento, una precision elevata è essenziale per ridurre i falsi allarmi. Un'elevata precision significa che il modello è in grado di distinguere efficacemente i rifiuti plastici da altri oggetti, riducendo il rischio di identificare erroneamente materiali innocui come inquinanti, il che è cruciale per evitare sprechi di risorse in operazioni di pulizia.

2.2 Accuracy

L'*Accuracy* rappresenta la proporzione complessiva delle predizioni corrette (sia positive che negative) rispetto a tutte le predizioni effettuate dal modello. Sebbene l'*accuracy* sia una metrica comunemente utilizzata per valutare i modelli di machine learning, essa ha alcune limitazioni, specialmente **in scenari con classi sbilanciate**, come nel nostro caso. La formula per l'*accuracy* è:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

dove:

- *TN* (True Negative): il numero di oggetti non plastici correttamente esclusi dal modello;
- *FN* (False Negative): il numero di rifiuti plastici non rilevati dal modello.

In un ambiente come i fiumi, dove i rifiuti plastici possono rappresentare una piccola frazione rispetto agli altri oggetti, l'*accuracy* potrebbe non riflettere adeguatamente le capacità del modello. Un alto valore di *accuracy* potrebbe essere fuorviante, poiché potrebbe essere raggiunto semplicemente non rilevando quasi nessun oggetto come plastica (minimizzando i falsi positivi a scapito dei veri positivi). Per questo motivo, l'*accuracy* deve essere valutata in combinazione con altre metriche, come la *precision* e la *recall*.

2.3 Recall

La *Recall* misura la capacità del modello di identificare correttamente tutti i rifiuti plastici presenti. La *recall* è particolarmente importante in scenari dove la priorità è minimizzare il numero di rifiuti plastici non rilevati (falsi negativi), che possono avere un impatto ambientale significativo se lasciati nei corsi d'acqua. La formula per calcolare la *recall* è la seguente:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

Nel contesto del riconoscimento dei rifiuti plastici, un alto valore di *recall* significa che il modello è efficace nel rilevare la maggior parte dei rifiuti, riducendo al minimo il rischio che rifiuti plastici possano sfuggire all'attenzione. Tuttavia, un aumento della *recall* potrebbe portare a una

diminuzione della precision, poiché il modello potrebbe iniziare a classificare erroneamente più oggetti come plastica per evitare di perdere i rifiuti effettivi. Pertanto, è fondamentale trovare un equilibrio tra precision e recall.

2.4 F1-Score

L'*F1-Score* è la media armonica tra precision e recall, offrendo una misura bilanciata che tiene conto sia dei falsi positivi che dei falsi negativi. Questa metrica è particolarmente utile in scenari come il riconoscimento dei rifiuti plastici, dove il dataset può presentare classi sbilanciate e dove è importante non sovra-ottimizzare il modello per una sola metrica a scapito dell'altra. La formula per l'*F1-Score* è:

$$F1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

Nel contesto del riconoscimento dei rifiuti plastici, l'*F1-Score* è **particolarmente indicato**, poiché consente di valutare l'efficacia complessiva del modello, bilanciando l'accuratezza con la capacità di rilevamento. Un *F1-Score* elevato indica che il modello YOLO è in grado di mantenere un buon equilibrio tra identificare correttamente i rifiuti plastici e minimizzare gli errori.

2.5 mean Average Precision (mAP)

Il *mean Average Precision* (mAP) è una delle metriche più utilizzate e riconosciute nella valutazione delle performance di modelli di *object detection*. Il mAP rappresenta la media delle *Average Precision* (AP) calcolate su tutte le classi considerate nel modello. La *Average Precision*, a sua volta, è l'area sotto la curva Precision-Recall per ciascuna classe. Il mAP è particolarmente rilevante per il nostro caso di studio, poiché:

- Riflette le prestazioni complessive del modello su più classi di rifiuti plastici (se categorizzati in diverse tipologie);
- Tiene conto della capacità del modello di rilevare i rifiuti plastici in condizioni variabili, come differenti condizioni di illuminazione, presenza di acqua torbida, e variazioni nella forma e dimensione dei rifiuti;

- Considera le variazioni nella soglia di rilevamento del modello, fornendo un'indicazione della sua robustezza e capacità di generalizzazione.

Nel caso del riconoscimento dei rifiuti plastici, un mAP elevato indica che il modello è in grado di rilevare correttamente i rifiuti in diverse situazioni operative, mantenendo un buon equilibrio tra precisione e richiamo per ciascuna classe considerata. Questo è cruciale per garantire che il sistema di monitoraggio automatico sia affidabile in ambienti reali e possa contribuire efficacemente alla riduzione dell'inquinamento nei corsi d'acqua.