

UNIVERSITÀ POLITECNICA DELLE MARCHE

FACOLTÀ DI INGEGNERIA



*Corso di Laurea Triennale in
Ingegneria Informatica e dell'Automazione*

*Sviluppo di un sistema di intelligenza artificiale a doppio
step per il rilevamento e la classificazione della pesca
ricreativa*

*Development of a Two-Step Artificial Intelligence System
for the Detection and Classification of Recreational Fishing*

Relatore:

DR. ALESSANDRO GALDELLI

Laureando:

MICHELE VIGLIOTTA

ANNO ACCADEMICO 2024-2025

"The best way to predict the future is to invent it."

- Alan Kay, 1971.

Indice

1	Introduzione	5
1.1	Obiettivi	8
1.1.1	Panoramica sull'Ambiente di Sviluppo e sulle Tecnologie Utilizzate	8
1.2	Struttura della Tesi	9
2	IUU Fishing (Illegal, Unreported, Unregulated)	10
2.1	Dimensioni Globali del Fenomeno	11
2.2	La Pesca Ricreativa nel Contesto IUU	14
2.3	Rischi della Pesca Ricreativa	16
2.4	Monitoraggio della Pesca Ricreativa	16
2.5	Limiti del Monitoraggio della Pesca Ricreativa	18
3	Fondamenti Teorici di Computer Vision	19
3.1	Introduzione alla Computer Vision	19
3.1.1	Storia della Computer Vision	19
3.2	Pipeline della Computer Vision	20
3.3	Principi di Elaborazione delle Immagini	21
3.4	Machine Learning vs Deep Learning	22
3.5	Reti Neurali Convoluzionali (CNN)	24
3.5.1	Struttura di una CNN	24
3.6	Object Detection	27
3.7	Classification	28
3.8	Architettura YOLO	28
4	Tecnologie Utilizzate	32
4.1	Google Drive	32
4.2	Google Colab	33
4.3	Visual Studio Code	33
4.4	Python	33
4.5	Ultralytics YOLO	33

5 Progettazione e Sviluppo del Software	34
5.1 Approccio 2-Step	34
5.2 Dataset Utilizzati	35
5.3 Modello di Detection	38
5.3.1 Ottimizzazione dei Parametri - Detection	38
5.4 Modello di Classification	41
5.4.1 Ottimizzazione dei Parametri - Classification	41
6 Risultati	42
6.1 Risultati del Modello di Detection	42
6.1.1 Considerazioni - Detection	43
6.2 Risultati del Modello di Classification	44
6.2.1 Considerazioni - Classification	45
7 Conclusioni	47
7.1 Sviluppi Futuri	48
Bibliografia	50
Elenco delle Figure	56
Ringraziamenti	57

Capitolo 1

Introduzione

In questi ultimi decenni, lo sviluppo tecnologico ha influenzato in modo profondo molte attività connesse alla produzione alimentare e alla gestione delle risorse naturali, coinvolgendo anche il settore della pesca.

Dagli strumenti base per la navigazione e la cattura si è passati a tecnologie sempre più avanzate, quali sistemi satellitari per la localizzazione delle imbarcazioni, ecoscandagli per individuare i banchi di pesce e mappare i fondali, oltre a piattaforme digitali per garantire la tracciabilità delle attività di pesca commerciale. Queste innovazioni hanno migliorato l'efficienza e la produttività, favorendo un monitoraggio ambientale più efficace e una maggiore conformità alle normative vigenti.

In particolare, con l'avvento dell'intelligenza artificiale e grazie alla computer vision, attualmente è possibile sviluppare sistemi automatizzati in grado di rilevare e classificare le attività di pesca, fornendo un valido supporto al contrasto della pesca illegale, non dichiarata e non regolamentata (*IUU – Illegal, Unreported and Unregulated fishing*). La pesca IUU rappresenta infatti una delle principali minacce alla sostenibilità degli ecosistemi marini e alla gestione responsabile delle risorse ittiche. Si stima che la pesca IUU contribuisca a un fatturato globale di circa **10 miliardi di euro** all'anno e che circa il **19%** del pescato annuale provenga da attività riconducibili alla pesca IUU. Tali stime mettono in luce l'ampiezza del fenomeno e l'importanza del monitoraggio [1].

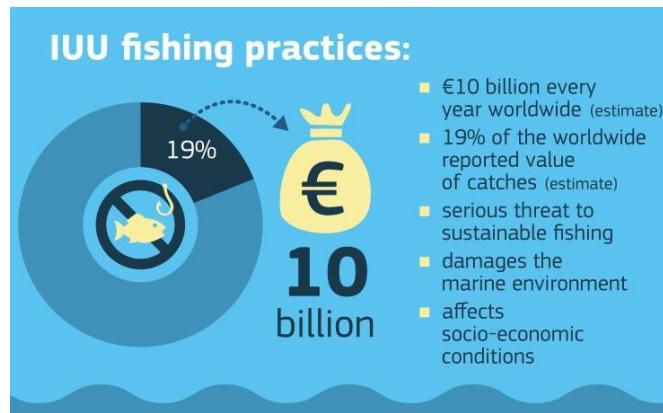


Figura 1.1: Stima dell'impatto della pesca illegale, non dichiarata e non regolamentata. Fonte [1].

In particolare, in Europa, la regolamentazione della pesca vede inizio con l'istituzione della Politica Comune della Pesca (PCP) nel 1982, finalizzata a garantire una gestione sostenibile delle risorse marine e una ripartizione equa della disponibilità di tali risorse tra gli Stati Membri. Nel corso degli anni, la PCP ha subito diverse riforme per rispondere ai cambiamenti economici, sociali e ambientali che hanno reso necessaria un'evoluzione costante della normativa. Per migliorare l'efficacia del monitoraggio, nel 2009 è stato adottato il *Regolamento (CE) n. 1224/2009*, noto come Control Regulation, che ha introdotto un sistema integrato di controllo, ispezione ed esecuzione uniforme in tutti i Paesi UE.

Successivamente, la normativa è stata aggiornata con il *Regolamento (UE) 2023/2841*, approvato nel 2023 ed entrato in vigore nel 2024, che ha modernizzato e rafforzato gli strumenti di monitoraggio [2]. Per avere un quadro completo delle tecnologie adottate, si riportano di seguito gli strumenti di monitoraggio e controllo previsti dalla normativa europea più recente, la cui adozione è obbligatoria o progressiva:

1. *Vessel Monitoring System (VMS)*: principale sistema di monitoraggio installato sulle imbarcazioni da pesca commerciale. Basato prevalentemente su tecnologie satellitari, consente alle autorità di seguire in tempo reale la posizione, la velocità e il percorso delle imbarcazioni, garantendo un controllo efficace delle attività di pesca [3].
2. *Diario di bordo elettronico (ERS – Electronic Reporting System)*: sistema digitale per registrare le catture, le zone di pesca e le specie pescate. Per le imbarcazioni di dimensioni ridotte, è previsto lo sviluppo di applicazioni dedicate per smartphone, al fine di semplificare la registrazione e l'invio dei dati alle autorità competenti [3].

-
3. *Automatic Identification System (AIS)*: sistema di identificazione automatica impiegato per la sicurezza marittima e il monitoraggio delle attività di pesca. Permette alle autorità di identificare le imbarcazioni e monitorarne il traffico in tempo reale, integrando i dati forniti dal VMS [3].
 4. *Remote Electronic Monitoring (REM)*: comprende telecamere, sensori e altri strumenti elettronici installati a bordo, con l'obiettivo di monitorare le operazioni di pesca e verificare il rispetto delle normative. L'adozione è obbligatoria per alcune imbarcazioni di dimensioni maggiori e sarà progressivamente estesa anche a imbarcazioni di dimensioni inferiori [3].
 5. *Strumenti digitali per la pesca ricreativa*: la normativa europea più recente estende l'obbligo di tracciabilità anche alle attività di pesca non commerciale. In particolare, il *Regolamento di esecuzione (UE) 2025/274* del 12 febbraio 2025, che specifica le modalità applicative dell'Articolo 55 della Control Regulation, introduce sistemi elettronici per la registrazione delle catture da parte dei pescatori ricreativi. I dati raccolti, trasmessi attraverso app e piattaforme digitali, verranno aggregati e comunicati alla Commissione dalle autorità nazionali. Inoltre, è prevista la marcatura degli attrezzi da pesca passivi, al fine di garantirne l'identificazione e distinguerli da quelli utilizzati dalla pesca professionale [4].

Dall'analisi delle tecnologie attualmente adottate o in fase di implementazione, emerge che il monitoraggio della pesca ricreativa si basa principalmente sulla raccolta dati manuale, tramite registrazioni volontarie dei pescatori. Questi strumenti, pur essendo funzionali, forniscono solo stime indicative e non garantiscono una visione completa e accurata dell'attività di pesca ricreativa. Il Parlamento Europeo, nel report *"The state of Recreational Fisheries in the EU"* del 2017, sottolinea esplicitamente come la mancanza di dati accurati e completi limiti la possibilità di valutare l'impatto sugli stock ittici, stimare correttamente il valore economico generato e definire politiche di gestione efficaci [5].

Per superare tali limitazioni, il presente lavoro di tesi si concentra sullo sviluppo di un sistema basato su intelligenza artificiale, progettato per rilevare automaticamente le attività di pesca ricreativa sulle spiagge. Questo approccio permette di ottenere dati più precisi, continui e oggettivi rispetto agli strumenti tradizionali. Un ulteriore vantaggio risiede nel fatto che si tratta di un sistema automatico, che non richiede l'intervento diretto dei pescatori o degli operatori, riducendo errori, ritardi e la raccolta di dati incompleti. Complessivamente, il sistema offre potenziali benefici per studi scientifici, gestione sostenibile delle risorse e supporto alle politiche di regolamentazione.

1.1 Obiettivi

L'obiettivo principale di questa tesi è lo **sviluppo di un sistema di intelligenza artificiale per il monitoraggio delle attività di pesca ricreativa attraverso un approccio a due stadi**. Per "due stadi" si intende che l'analisi viene suddivisa in due fasi consecutive: nel primo stadio il sistema individua tutte le persone presenti nella scena (object detection), mentre nel secondo stadio determina se ciascuna persona è un pescatore (classification). Questo approccio concentra l'analisi sulle aree rilevanti, aumentando l'accuratezza del rilevamento e della classificazione rispetto ai metodi tradizionali che eseguono detection e classification simultaneamente. Si basa inoltre sull'osservazione che distinguere una persona con una canna da pesca da una persona comune è relativamente semplice una volta individuate le persone.

L'implementazione utilizza reti neurali convoluzionali e architetture avanzate di deep learning per la computer vision. Il sistema offre vantaggi operativi significativi: riduce i costi di monitoraggio minimizzando la necessità di controlli umani sul campo grazie all'impiego di telecamere fisse e droni, garantisce maggiore oggettività e riproducibilità dei risultati e assicura una scalabilità superiore rispetto agli approcci convenzionali. Inoltre, questo lavoro mira a migliorare la completezza e l'accuratezza dei dati relativi alle attività di pesca ricreativa, affiancandosi alle strategie di raccolta dati già presenti in materia. Tali informazioni potranno essere di supporto per la ricerca scientifica e facilitare lo sviluppo di future politiche di gestione delle risorse ittiche.

1.1.1 Panoramica sull'Ambiente di Sviluppo e sulle Tecnologie Utilizzate

Il linguaggio principale scelto è Python, per la sua flessibilità, la disponibilità di librerie dedicate alla visione artificiale e al machine learning (PyTorch, OpenCV, Ultralytics YOLO), e la facilità di integrazione con altri componenti. Python consente di sviluppare algoritmi di elaborazione dati, pipeline di addestramento dei modelli e strumenti di interazione con l'utente in modo efficiente. L'addestramento dei modelli viene condotto principalmente su Google Colab, sfruttando le risorse GPU per accelerare i tempi di elaborazione, mentre Visual Studio Code (VS Code) è utilizzato prevalentemente per il testing dei modelli.

I dettagli di ciascuna tecnologia, delle librerie utilizzate e delle scelte progettuali specifiche sono riportati nel Capitolo 4. I dati utilizzati per l'addestramento del sistema provengono da due dataset costruiti a partire da immagini annotate preesistenti. Questi dataset sono stati selezionati ed elaborati per le fasi di object detection e classificazione. I dettagli relativi alla preparazione e alla struttura dei dataset sono approfonditi nel Capitolo 5.

1.2 Struttura della Tesi

La tesi è articolata nei seguenti capitoli:

1. *Introduzione*: fornisce il contesto della ricerca, gli obiettivi principali e l'importanza del monitoraggio della pesca ricreativa.
2. *IUU Fishing*: descrive il fenomeno della pesca illegale, ponendo particolare attenzione alle attuali misure di prevenzione. Inoltre, fornisce una definizione della pesca ricreativa, illustrandone la normativa vigente e gli strumenti di monitoraggio.
3. *Fondamenti Teorici*: presenta i concetti chiave alla base del progetto, illustrando i principi della computer vision, le reti neurali convoluzionali (CNN) e l'architettura YOLO (You Only Look Once).
4. *Tecnologie Utilizzate*: illustra gli strumenti software e hardware adottati e l'ambiente di sviluppo.
5. *Progettazione e Sviluppo del Software*: descrive lo sviluppo del software, illustrando l'architettura del sistema, i dataset utilizzati per l'addestramento, la configurazione dei modelli e i test eseguiti.
6. *Risultati*: espone i risultati ottenuti dal modello di detection e dal modello di classification, analizzando le metriche di performance.
7. *Conclusioni*: presenta le conclusioni del progetto, valutando l'efficacia del sistema proposto e identificando le possibili direzioni di sviluppo futuro.

La tesi si conclude con la bibliografia, l'elenco delle figure e i ringraziamenti.

Capitolo 2

IUU Fishing (Illegal, Unreported, Unregulated)

La pesca IUU (Illegal, Unreported and Unregulated) indica un'ampia gamma di attività di pesca svolte violando leggi o al di fuori di un'opportuna regolamentazione. Può essere praticata da imbarcazioni di ogni tipo e dimensione, sia in acque internazionali che in quelle nazionali. Il fenomeno interessa non solo la cattura, ma anche la gestione e la commercializzazione del prodotto pescato.

La distinzione tra pesca illegale, non dichiarata e non regolamentata è stata formalmente definita dalla Food and Agriculture Organization (FAO) nel piano d'azione internazionale *International Plan of Action to Prevent, Deter and Eliminate Illegal, Unreported and Unregulated Fishing (IPOA-IUU)*, nel 2001 [6]. Di seguito sono riportate le rispettive definizioni:

1. *Pesca illegale*: si intendono tutte le attività condotte in violazione delle leggi o regolamenti nazionali o internazionali, ad esempio la pesca senza licenza o in aree vietate.
2. *Pesca non dichiarata*: si intendono le attività non riportate o riportate in modo errato alle autorità competenti, con comunicazioni assenti, falsificate o incomplete.
3. *Pesca non regolamentata*: si intendono attività svolte in aree o con specie coinvolte prive di regole specifiche, oppure condotte da navi senza nazionalità o che operano sotto delle opportune bandiere per evitare controlli.

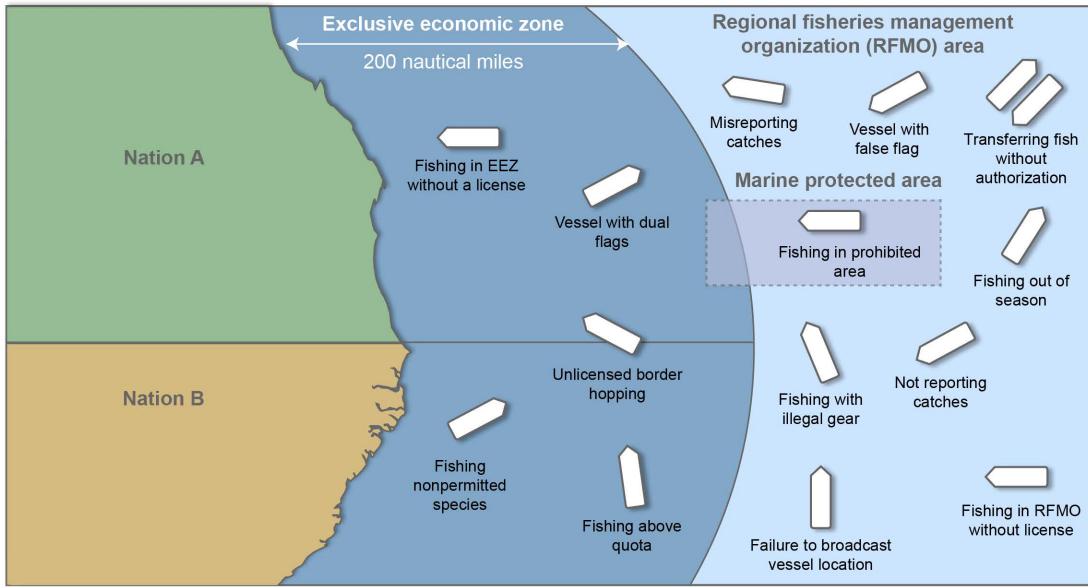


Figura 2.1: Classificazione pesca illegale, non dichiarata e non regolamentata.
Fonte [7].

2.1 Dimensioni Globali del Fenomeno

L'impatto della pesca IUU è complesso da quantificare a causa della natura stessa delle attività che la costituiscono; tuttavia, sono presenti diversi studi che nel corso degli anni hanno fornito delle stime. Uno dei primi è lo studio *"Estimating the Worldwide Extent of Illegal Fishing"* del 2009. Gli autori hanno stimato perdite globali dovute alla pesca illegale e non dichiarata tra 11 e 16 milioni di tonnellate annue, con un valore economico compreso tra 10 e 23,5 miliardi di dollari [8].

Più recentemente, Sumaila et al. [9] hanno condotto uno studio sul commercio illecito di prodotto ittico pescato, conseguenza delle attività di pesca IUU. Si stima una quantità tra gli 8 e i 14 milioni di tonnellate non dichiarate e commercializzate annualmente. A tali catture corrisponderebbero dei ricavati annuali tra i 9 e i 17 miliardi di dollari. Di particolare rilevanza sono le ulteriori stime che riguardano le perdite economiche annue, comprese tra i 25 e 50 miliardi di dollari globalmente, e le perdite per le mancate entrate fiscali per le nazioni, quantificabili tra i 2 e 4 miliardi di dollari.

Sia la FAO che il World Wide Fund for Nature (WWF) confermano tali dati: la FAO stima fino a 26 milioni di tonnellate annue di catture illegali [10], mentre il WWF ritiene che le attività IUU rappresentino l'11-19% della produzione ittica annua, con perdite economiche comprese tra 10 e 23,5 miliardi di dollari [11].

La pesca IUU non si manifesta uniformemente a livello globale, ma è caratterizzata da concentrazioni geografiche peculiari che dipendono da diversi fattori sociopolitici ed economici. Un ottimo strumento per avere una miglior visione sulla distribuzione geografica del fenomeno consiste nell'*IUU Fishing Risk Index*. Sviluppato dalla Poseidon Aquatic Resource Management Ltd. e dalla Global Initiative Against Transnational Organized Crime [12], tale indice fornisce una misura della probabilità che gli stati siano esposti alla pesca IUU e del loro impegno attivo nel contrastarla in modo efficace. L'indice utilizza un punteggio di rischio compreso tra 1 e 5 (da migliore a peggiore) e valuta 152 Stati costieri attraverso 40 indicatori raggruppati in quattro categorie di responsabilità statale:

1. *Coastal*: riguarda la gestione della zona economica esclusiva di ogni stato.
2. *Flag*: riguarda il controllo sulle imbarcazioni registrate sotto la bandiera dei rispettivi stati.
3. *Port*: riguarda la gestione delle attività di pesca nei porti.
4. *General*: riguarda indicatori come vulnerabilità e partecipazione al contrasto di attività illegali.

I risultati vengono rappresentati attraverso un grafico a forma di pesce (Figura 2.2), dove le altezze delle lische corrispondono ai punteggi delle quattro categorie di responsabilità statale.

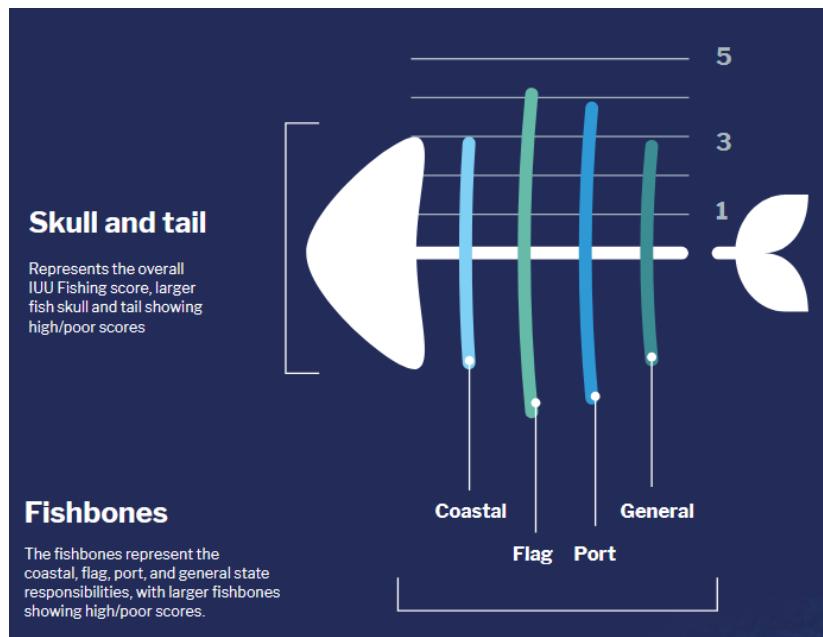


Figura 2.2: Grafico IUU Risk Index. Fonte [12].

Nel report del 2023 [12], il punteggio globale dell'IUU Fishing Risk Index è risultato pari a 2,28 (Figura 2.3), in aumento rispetto a un punteggio del 2,24 del 2021. Ciò evidenzia come una costante lotta a tali attività sia tuttora necessaria e altamente rilevante. L'area geografica più a rischio è l'Asia, seguita dal Medio Oriente e dall'Africa. A livello di bacini oceanici, l'Oceano Indiano occidentale e il Pacifico occidentale presentano i punteggi peggiori. I punteggi individuali delle nazioni mostrano grandi margini di differenza: la Cina presenta il punteggio peggiore con 3,69, mentre la Romania registra il migliore con 1,62.

I paesi in via di sviluppo risultano essere quelli più colpiti dalle attività di tipo illegale, fenomeno coerente con l'assenza di risorse necessarie per l'implementazione di sistemi di monitoraggio e controllo. Questo squilibrio tra vulnerabilità e capacità di risposta sottolinea i vantaggi dell'impiego di sistemi informatici innovativi per la lotta alla pesca IUU, particolarmente in contesti dove le risorse per il controllo tradizionale sono limitate.

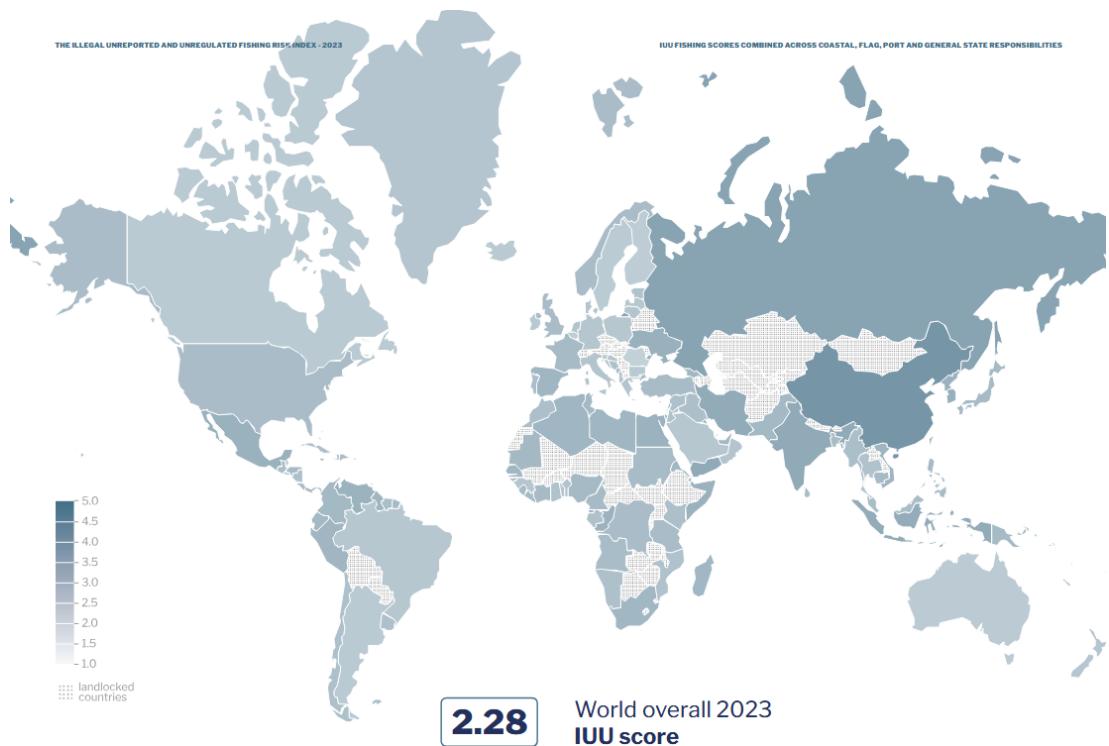


Figura 2.3: Punteggio globale IUU Fishing Risk Index 2023. Fonte [12].

2.2 La Pesca Ricreativa nel Contesto IUU

La pesca ricreativa è la cattura di animali acquatici, soprattutto pesci, non destinati a soddisfare bisogni nutrizionali primari e che di norma non vengono venduti o commercializzati. Nei Paesi industrializzati interessano principalmente le acque interne e le aree costiere [13]. Anche la pesca sportiva rientra nella più ampia categoria della pesca ricreativa, come riportato nel *Regolamento (UE) 2023/2842* [2]. Esistono inoltre diverse categorie principali di pesca ricreativa [14]:

1. *Pesca da barca*: comprende tecniche come, ad esempio, la pesca a lenza, la traina, lo spinning e l'impiego di trappole per pesci.
2. *Pesca da riva*: include attività quali il lancio con rete, la pesca a fondo con canna, lo spinning, la tradizionale pesca di superficie italiana.
3. *Raccolta di molluschi*: può essere svolta sott'acqua o camminando lungo la costa.
4. *Pesca subacquea*: praticata sia da riva che da barca, generalmente con l'impiego di arpioni.

Attività di pesca ricreativa possono essere considerate attività di pesca IUU qualora siano violate regolamentazioni presenti nella zona. Degli esempi possono essere la pesca senza licenza, il superamento delle quote giornaliere, pescare in aree protette, la cattura di specie protette, la mancata registrazione delle catture ove richiesta, il mancato rispetto delle taglie minime di cattura (Figura 2.4) e l'utilizzo di attrezzatura non consentita per la pesca ricreativa.



Figura 2.4: Guida del Parco Nazionale delle Calanques (Francia), che specifica le dimensioni minime per tutte le specie pescate a livello ricreativo 2018. Fonte [15].

L'attività della pesca ricreativa e il suo conseguente impatto possono sembrare marginali, ma in realtà diversi studi dimostrano come sia di assoluta importanza riuscire a monitorare il fenomeno per eliminare le possibili minacce ambientali e problematiche socioeconomiche.

Un report che fornisce un'ottima panoramica su tale attività è quello redatto da Gómez et al. nel 2019 in collaborazione con il WWF, che analizza il fenomeno concentrandosi sull'area costiera mediterranea. Infatti, si stima che in Europa ci siano circa 9 milioni di pescatori ricreativi, che generano annualmente all'incirca 6 miliardi di euro [16]. L'attività risulta molto presente e continua lungo tutte le coste del Mediterraneo, nella mappa sono evidenziate tutte le aree dove hanno luogo potenziali attività di pesca ricreativa (Figura 2.5).

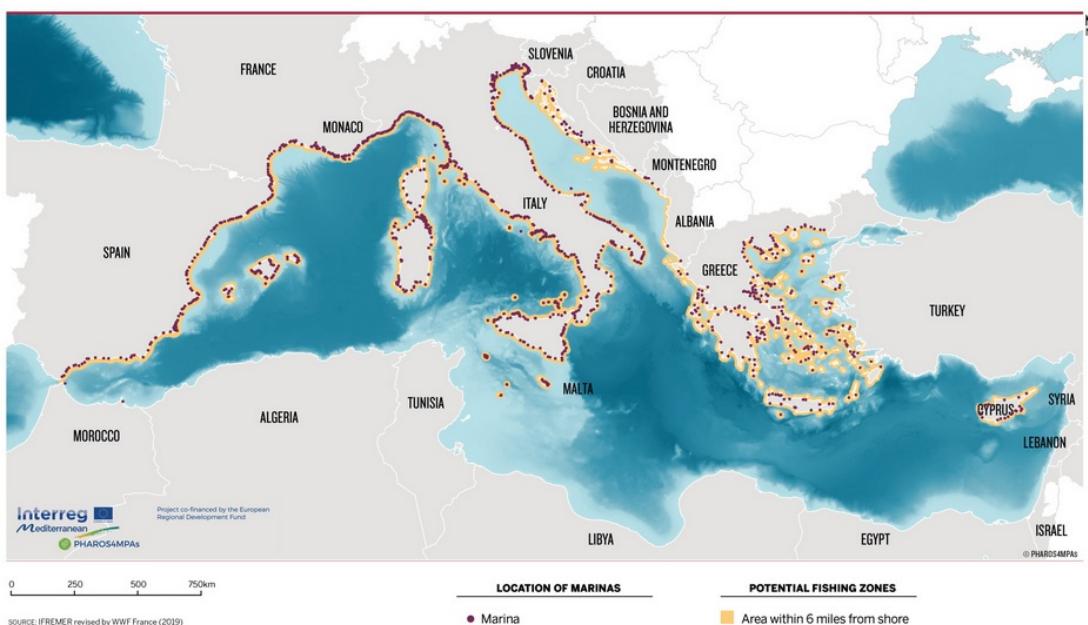


Figura 2.5: Mappa della distribuzione pesca ricreativa. Fonte [16].

Nonostante la scarsità dei dati l'Unione Europea stima che la pratica della pesca possa essere responsabile di più del 10% di tutta la cattura ittica annua. Inoltre, alcuni studi hanno dimostrato che la pressione di pesca esercitata dalla pesca ricreativa in alcune aree risulta comparabile o addirittura superiore a quella della pesca commerciale [16]. Questo scenario evidenzia la necessità di un monitoraggio efficace al fine di sviluppare delle politiche di gestione coerenti ed efficaci.

2.3 Rischi della Pesca Ricreativa

La pesca ricreativa apporta indubbiamente benefici economici rilevanti. Si stima che il valore complessivo dell'attività economica a essa associata sia di circa 10 miliardi di euro, comprendenti 5 miliardi di spese dirette, 2 miliardi di spese indirette e ulteriori 3 miliardi di spese indotte, generando quasi 100.000 posti di lavoro a tempo pieno nell'Unione Europea [16].

Tuttavia, non vanno trascurati i potenziali rischi. Per sua stessa natura, si tratta di un'attività che contribuisce ad aumentare lo sforzo delle risorse ittiche, già interessate dallo sfruttamento ai fini commerciali. I pescatori ricreativi possono rappresentare una minaccia per le specie vulnerabili, le cui popolazioni hanno già subito declini rilevanti.

Oltre alle attività di pesca illegale, purtroppo presenti anche in questo settore, si presentano anche altre criticità, quali: il danneggiamento degli habitat dovuto all'uso di attrezzature non consentite, l'introduzione di specie esotiche utilizzate come esche, l'inquinamento del mare a causa di equipaggiamento perso o abbandonato e i conflitti tra pescatori ricreativi e pescatori professionali.

2.4 Monitoraggio della Pesca Ricreativa

Dopo aver analizzato il quadro generale e l'impatto della pesca ricreativa, risulta evidente che il monitoraggio di questa attività è di fondamentale importanza per garantire la sostenibilità degli stock ittici e valutare il prelievo totale delle risorse marine. Proprio in quest'ottica, negli ultimi anni la direzione comune è stata quella di sviluppare sistemi di controllo sempre più stringenti e accurati. In Europa, come accennato in precedenza, la Control Regulation, il cui ultimo aggiornamento è il *Regolamento (UE) 2023/2842* [3], è il documento ufficiale in cui la Commissione europea stabilisce le norme che regolano le attività di pesca. L'articolo 55 stabilisce la gestione della pesca ricreativa, mentre il *Regolamento di esecuzione (UE) 2025/274* ne definisce le modalità di applicazione. Tale regolamento prevede che ogni Stato membro costiero, a partire dal 15 marzo 2026, dovrà inviare entro il 15 di ogni mese alla Commissione europea i dati relativi alle catture ricreative del mese precedente.

Per facilitare la raccolta dati, è stato sviluppato il sistema RecFishing [17], una piattaforma che consente ai pescatori ricreativi di comunicare le proprie catture in modo anonimo attraverso un European RecFisher ID univoco. Le informazioni richieste includono [4]:

1. *Catture trattenute (non rilasciate)*: quantità (tonnellate) e numero di individui per specie, con suddivisione per area geografica, modalità di pesca, tipologie di attrezzature e lunghezza.

2. *Catture rilasciate*: numero di individui per specie, con la stessa suddivisione per area, modalità, attrezzature e lunghezza.

I pescatori ricreativi utilizzeranno applicazioni mobile su smartphone per comunicare le catture. Il sistema consente di scegliere tra tre modalità di implementazione (Figura 2.6): app commerciali già esistenti, app sviluppate dai singoli Stati membri, app ufficiale sviluppata dalla Commissione europea. Tutte le app, a prescindere dalla modalità selezionata, trasmettono i dati raccolti a un server centrale comune. Il personale autorizzato di ogni Stato membro può accedere ai dati relativi al proprio territorio, per effettuare verifiche ed eventualmente apportare correzioni. In seguito, i dati vengono aggregati e trasmessi alla commissione, rimuovendo i dati personali, garantendo la privacy dei pescatori [18].

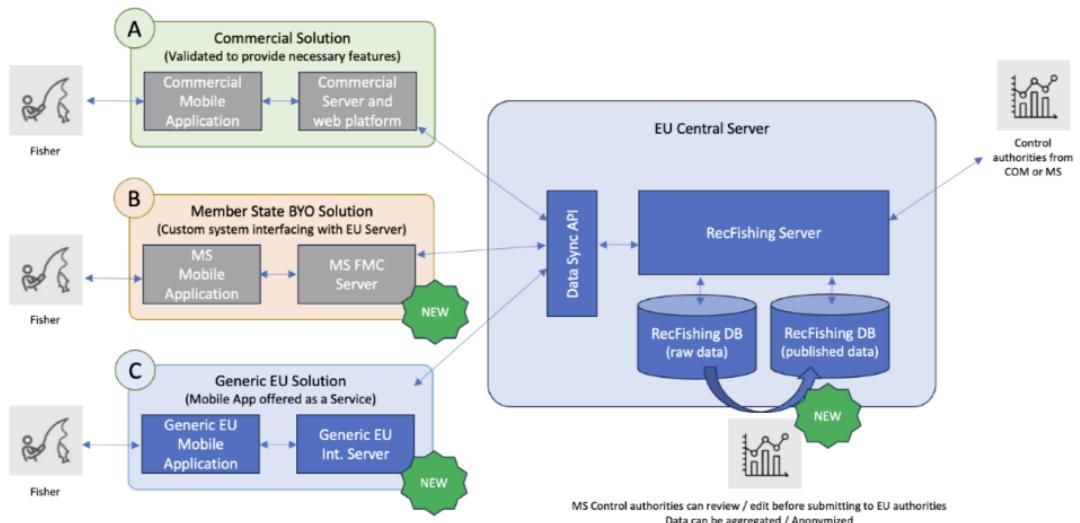


Figura 2.6: Architettura del sistema RecFishing con le tre opzioni disponibili di connessione tra app e server centrale della Commissione. Fonte [18].

È inoltre prevista la marcatura obbligatoria, riconducibile al pescatore, di tutti gli strumenti passivi (come nasse e reti) impiegati nella pesca ricreativa, in modo da promuovere un uso più responsabile delle attrezzature. In particolare, in Italia, per poter praticare la pesca ricreativa in mare è necessario effettuare una comunicazione al Ministero dell'Agricoltura, della Sovranità Alimentare e delle Foreste (MASAF) [19]. Questa comunicazione, che ha valore di autocertificazione, consente di censire i pescatori ricreativi attivi e costituisce la base per il monitoraggio dell'attività ai fini gestionali e statistici.

Per confrontare la situazione con altre aree geografiche, è utile analizzare il monitoraggio negli Stati Uniti. La pesca ricreativa è monitorata dal Marine Recreational Information Program (MRIP), gestito dalla National Oceanic and Atmospheric Administration (NOAA) Fisheries. Il MRIP adotta un approccio

multicanale, combinando sondaggi via e-mail e telefono, interviste dirette presso porti e moli, e in alcune aree l'uso di app mobile per la registrazione in tempo reale delle catture [20].

2.5 Limiti del Monitoraggio della Pesca Ricreativa

Nonostante il crescente impegno nello sviluppo di nuovi sistemi di monitoraggio della pesca sportiva, permangono diverse sfide da superare per raggiungere un'elevata accuratezza dei dati e produrre stime più attendibili sul fenomeno. Ad esempio, i metodi tradizionali di sondaggio presentano limiti significativi e i dati raccolti potrebbero essere sottostimati o distorti.

I sondaggi del MRIP presentano diverse criticità: problemi di campionamento derivanti dalla difficoltà di raggiungere pescatori in aree remote o private e in specifici orari, nonché errori di identificazione delle specie dovuti alla complessità del riconoscimento, tanto per i pescatori quanto per gli intervistatori. Questi fattori compromettono significativamente la validità e l'affidabilità dei dati raccolti [21].

Anche le tecnologie di monitoraggio più recenti presentano problematiche significative. Ad esempio, le app per la segnalazione delle catture da parte dei pescatori possono migliorare la raccolta dati e ridurne i costi, ma presentano importanti limitazioni. Pochi pescatori scaricano o utilizzano regolarmente queste app, causando potenziali bias nei dati. Ciò è dovuto al fatto che si tratta di campioni non probabilistici, bensì volontari. Non possono quindi essere rappresentativi della popolazione e vanno ancora integrati con i sondaggi basati su campioni randomici [22].

Il principale limite dell'attuale monitoraggio della pesca ricreativa risiede nella grande dipendenza dall'intervento umano e dalla partecipazione volontaria. L'obiettivo del lavoro di questa tesi è sviluppare un sistema automatizzato basato su intelligenza artificiale (IA) per migliorare la raccolta dati del fenomeno e aumentare l'accuratezza, integrando tale sistema con i metodi già esistenti. Questa direzione è condivisa da altri ricercatori, occupati nello sviluppo di sistemi per tracciare imbarcazioni destinate alla pesca ricreativa. Poiché queste imbarcazioni sono generalmente sprovviste di sistemi AIS o VMS, a differenza di quelle commerciali, gli studi impiegano immagini satellitari ad alta risoluzione per analizzare gli spostamenti [23].

Capitolo 3

Fondamenti Teorici di Computer Vision

3.1 Introduzione alla Computer Vision

Lo sviluppo del sistema presentato in questa tesi si basa sulle tecniche della computer vision. La computer vision è il campo dell'intelligenza artificiale che permette alle macchine di comprendere e analizzare le informazioni presenti nelle immagini e nei video. Il funzionamento della computer vision si basa sull'utilizzo delle reti neurali convoluzionali (CNN) e del deep learning [24]. Quello della computer vision può essere considerato come un problema inverso, poiché lo scopo è quello di ricostruire informazioni non note o parziali per determinare una descrizione significativa della scena [25].

Pur essendo numerose le applicazioni della computer vision, il presente lavoro si concentra sulle tecniche di object detection e classificazione. Nel contesto della pesca ricreativa e del lavoro svolto in questa tesi, la computer vision rappresenta uno strumento fondamentale per automatizzare il rilevamento delle persone in aree costiere e la successiva classificazione dei pescatori.

3.1.1 Storia della Computer Vision

Le origini della computer vision risalgono agli anni '50 e '60, quando i ricercatori iniziarono a sviluppare algoritmi per elaborare e visualizzare dati. In quegli anni, l'IA si affermò come ambito di ricerca e uno dei suoi obiettivi era quello di replicare la capacità umana di interpretare le immagini. Negli anni '70 si verificarono miglioramenti significativi e fu introdotta la tecnologia di riconoscimento ottico dei caratteri (OCR), che consentì ai computer di leggere ogni tipologia di testo stampato. Tra gli anni 90 e i primi anni 2000, il focus si spostò sullo sviluppo di algoritmi in grado di estrarre caratteristiche chiave dalle immagini e il machine learning iniziò così ad assumere un ruolo sempre più rilevante. In questo periodo furono sviluppate le prime applicazioni di riconoscimento facciale.

Il vero punto di svolta nella computer vision risale al 2012, quando un team dell’Università di Toronto partecipò con una CNN a una competizione di riconoscimento delle immagini. Il modello, noto come AlexNet, superò nettamente le prestazioni dei metodi tradizionali, dimostrando l’efficacia e il potenziale del deep learning [26].

3.2 Pipeline della Computer Vision

Il processo di analisi delle immagini mediante tecniche di computer vision segue una pipeline composta da step fondamentali [27]:

1. *Acquisizione delle immagini*: raccolta di immagini o video attraverso telecamere, sensori o altri dispositivi di acquisizione.
2. *Preprocessing*: miglioramento delle immagini grezze (raw) in modo da facilitare l’elaborazione successiva.
3. *Feature detection*: estrazione delle caratteristiche chiave delle immagini, come i margini, le forme e i pattern ricorrenti. Ciò consente al sistema di concentrarsi sulle parti rilevanti delle immagini per determinare l’output.
4. *Pattern recognition*: confronto tra ciò che viene rilevato nell’immagine e modelli o esempi già noti.
5. *Decision making*: fase finale in cui il sistema utilizza le informazioni raccolte per generare una predizione, come ad esempio l’identificazione di un gatto in una foto.

3.3 Principi di Elaborazione delle Immagini

L’elaborazione delle immagini è un processo fondamentale che consente alla computer vision di eseguire una migliore analisi delle immagini. L’obiettivo principale di questo processo è migliorare la qualità delle immagini grezze (raw) in modo da rendere più semplice e accurata l’estrazione delle caratteristiche rilevanti.

Ad esempio, gli algoritmi di estrazione dei bordi analizzano le variazioni di intensità tra i pixel per determinare i contorni degli oggetti presenti nell’immagine (Figura 3.1).

Alcuni esempi di tecniche di processamento di immagini sono [28]:

- *Blurring*: consiste nella riduzione del rumore presente nell’immagine.
- *Operazioni geometriche*: comprendono modifiche come l’ingrandimento o il rimpicciolimento delle immagini.
- *Modifica delle proprietà visive*: include la regolazione di colore, luminosità, contrasto e altri parametri.

Nel caso di questa tesi è stata utilizzata una tecnica di elaborazione, il ritaglio delle immagini, applicata alle immagini in input al modello di classificazione addestrato. Ciò sarà discusso in maniera più approfondita nel capitolo dedicato allo sviluppo del sistema.

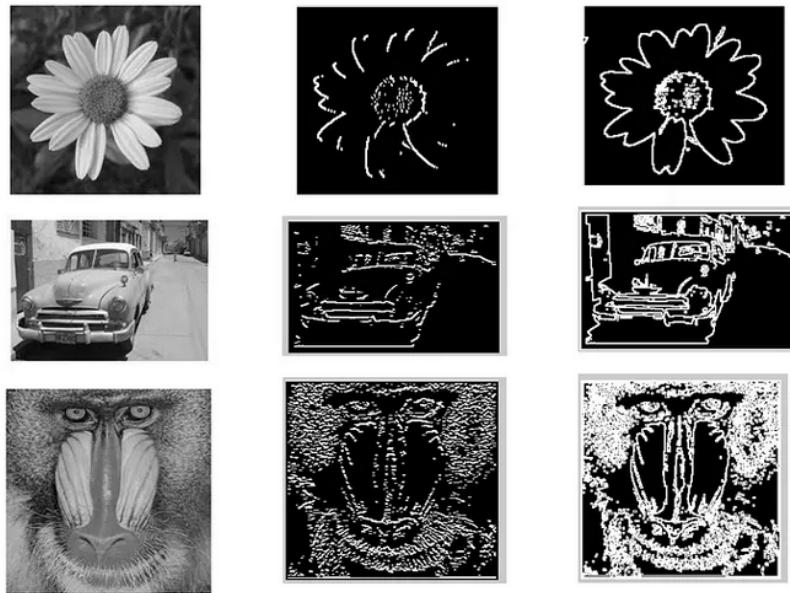


Figura 3.1: Esempio di applicazione dell’algoritmo di edge detection. Fonte [28].

3.4 Machine Learning vs Deep Learning

Il machine learning e il deep learning sono due approcci dell'intelligenza artificiale spesso utilizzati per applicazioni simili, ma presentano importanti differenze. Il machine learning è una branca dell'IA che permette ai computer di apprendere dai dati senza essere esplicitamente programmati per ciascun compito specifico. Gli algoritmi di machine learning richiedono un processo di feature engineering manuale. Ciò significa che è richiesta l'identificazione e la selezione manuale delle caratteristiche rilevanti dei dati da fornire al modello per l'addestramento.

Il deep learning, invece, è una sottocategoria del machine learning basata su reti neurali con molteplici strati (da cui il termine "deep", profondo). La differenza principale risiede nella capacità del deep learning di estrarre e apprendere in modo automatico le caratteristiche significative direttamente dai dati (Figura 3.2). Ciò è possibile grazie all'utilizzo di gerarchie di layer nelle reti neurali profonde: i primi layer apprendono feature di basso livello (come bordi e contorni), mentre i layer successivi ricevono le informazioni dai layer precedenti e si concentrano sul riconoscimento di pattern più complessi. Questo aspetto verrà approfondito nella sezione dedicata alle CNN [29].

Questi modelli presentano vantaggi e svantaggi specifici legati alla loro natura. Il machine learning raggiunge buone prestazioni con dataset di dimensioni ridotte e tempi di addestramento brevi, proprio grazie all'estrazione manuale delle feature. Il deep learning, al contrario, richiede dataset molto grandi ed elevata potenza di calcolo. Tuttavia, l'alta accuratezza raggiungibile, il facile accesso a enormi dataset già etichettati, la disponibilità di GPU sempre più performanti e la possibilità di utilizzare modelli pre-addestrati rendono oggi il deep learning lo strumento preferito in molti ambiti applicativi [29].

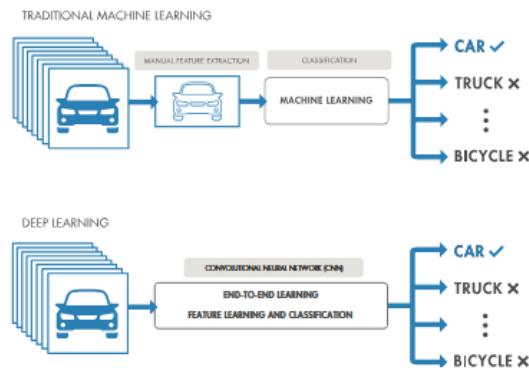


Figura 3.2: Confronto tra machine learning e deep learning. Fonte [29].

Un'altra differenza da sottolineare è quella tra supervised e unsupervised learning (Figura 3.3). È fondamentale chiarire che sia i modelli di machine learning che di deep learning possono essere supervised o unsupervised. Nel caso supervi-

sed, il modello viene addestrato su dati etichettati, che hanno quindi un output corretto corrispondente. In questo modo, il modello fa delle predizioni e le compara con gli output reali, riducendo gli errori e imparando così a fare previsioni accurate su dati nuovi, non visti in precedenza [25]. Al contrario, con l'unsupervised learning, al modello vengono forniti dati privi di etichette. L'obiettivo è quello di scoprire pattern nascosti, strutture o relazioni nel dataset, senza l'intervento umano o la conoscenza preventiva del significato dei dati [25].

Il sistema sviluppato in questa tesi si basa su modelli di deep learning supervisionati che utilizzano architetture CNN.

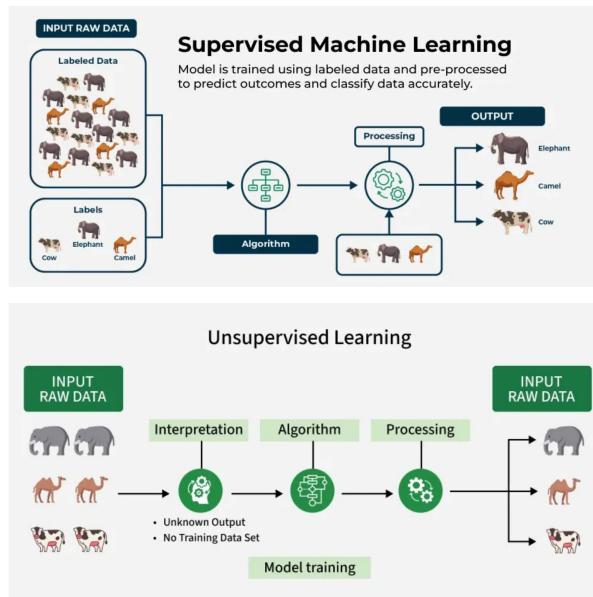


Figura 3.3: Confronto tra supervised learning e unsupervised learning. Fonte [30].

3.5 Reti Neurali Convoluzionali (CNN)

Le reti neurali convoluzionali (Convolutional Neural Networks, CNN) rappresentano l'architettura di riferimento del deep learning per la computer vision [31]. Popolarizzate da Yann LeCun et al. nel 1998 con la rete LeNet-5, progettata per il riconoscimento di cifre scritte a mano, conobbero una rapida diffusione a partire dal 2012 con lo sviluppo di AlexNet. Tale rete dimostrò prestazioni nettamente migliori nella classificazione di immagini rispetto alle tecniche tradizionali [25].

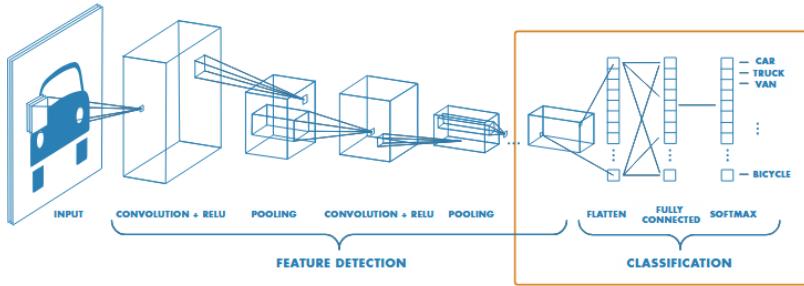


Figura 3.4: Architettura generica di una CNN. Fonte [29].

3.5.1 Struttura di una CNN

Le reti CNN utilizzano un'operazione matematica nota come convoluzione, che consente l'estrazione delle caratteristiche dalle immagini. Una CNN è generalmente costituita da una sequenza di layer principali: layer convoluzionali, layer di pooling, funzioni di attivazione e fully connected layer [32, 33].

Layer Convoluzionali

Questo layer è costituito da un insieme di filtri, o kernel, che scorrono sull'immagine di input generando le feature maps. Queste ultime sono vettori multidimensionali che contengono valori di attivazione, i quali indicano l'intensità della caratteristica rilevata dal filtro.

Ogni filtro ha dimensioni ridotte rispetto all'immagine di input. Durante la convoluzione, ogni filtro si sposta lungo tutta l'immagine e, per ogni posizione, viene eseguito il prodotto scalare tra i valori del filtro e i corrispondenti valori dell'input.

La seguente formula [34] descrive la convoluzione tra un input bidimensionale x e un filtro bidimensionale w . k_h e k_w sono rispettivamente l'altezza e la larghezza del filtro, mentre $y[i, j]$ è il valore risultante nella feature map:

$$y[i, j] = \sum_{m=0}^{k_h-1} \sum_{n=0}^{k_w-1} x[i + m, j + n] \cdot w[m, n]$$

La Figura 3.5 offre una rappresentazione visiva dell'operazione eseguita in questo layer. La matrice gialla rappresenta il filtro (in questo caso 3x3, con valori 1–0–1 / 0–1–0 / 1–0–1), mentre la matrice verde indica l'immagine di input. Nella posizione mostrata, si effettua il prodotto scalare tra ogni elemento del filtro e il corrispondente elemento sottostante dell'input, seguito dalla somma dei risultati. Tale operazione produce il valore che occupa la prima cella in alto a destra della feature map risultante.

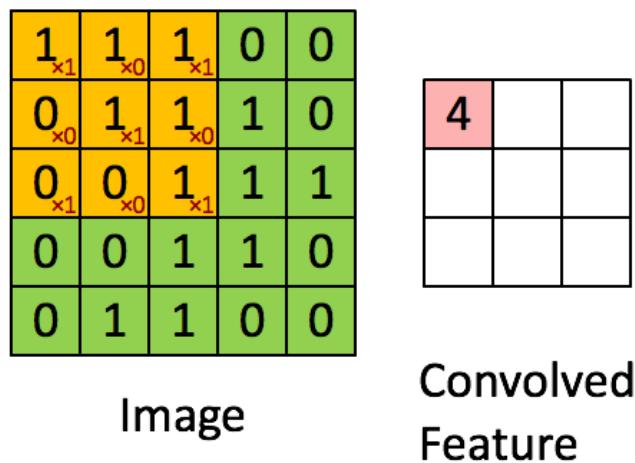


Figura 3.5: Esempio operazione di convoluzione. Fonte [35].

Layer di Pooling

Il layer di pooling è tipicamente collocato tra layer convoluzionali successivi. L'obiettivo è ridurre progressivamente le dimensioni spaziali delle feature map, diminuendo così il numero di parametri e il carico computazionale della rete. L'operazione di pooling più comune è il max pooling, che utilizza un filtro (di solito 2x2) per selezionare il valore massimo all'interno di ciascuna regione della feature map.

La Figura 3.6 mostra come una feature map venga ridotta di un fattore di 2, applicando un filtro 2x2 di max pooling e stride (scorrimento) di 2.

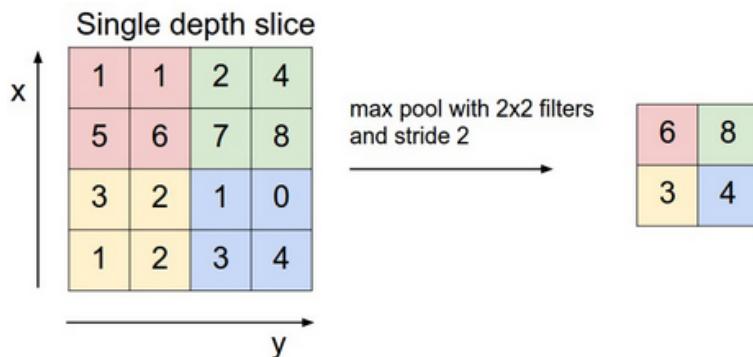


Figura 3.6: Applicazione del maxpooling. Fonte [33].

Funzioni di Attivazione

Dopo ogni operazione di convoluzione, viene applicata una funzione di attivazione non lineare per introdurre capacità di apprendimento di relazioni complesse. La funzione più utilizzata nelle CNN moderne è la ReLU (Rectified Linear Unit), definita come:

$$f(x) = \max(0, x)$$

La ReLU introduce non linearità: azzerà i valori negativi e mantiene quelli positivi, consentendo alla rete di imparare mappature non lineari.

Fully Connected Layer

Si tratta dello step finale (visibile in Figura 3.4). Le feature map vengono appiattite (flatten) in un vettore unidimensionale e passate attraverso uno o più layer completamente connessi. In questi layer, ogni neurone è connesso a tutti i neuroni del layer precedente attraverso pesi appresi durante il training. L'output di ciascun neurone viene calcolato come la somma pesata di tutti gli input più un termine di bias, dove i pesi determinano l'importanza di ciascuna connessione e vengono ottimizzati tramite backpropagation durante l'addestramento [36]. I fully connected layers combinano le feature estratte dai layer convoluzionali

per produrre l'output finale della rete, come le probabilità di appartenenza alle diverse classi in un problema di classificazione.

3.6 Object Detection

L'object detection è una tecnica di computer vision che ha come obiettivo il rilevamento di oggetti in immagini e video [37], associando a ciascuno di essi un'etichetta di classe. Un modello di object detection viene addestrato utilizzando un dataset di immagini annotate, in cui gli oggetti sono assegnati alle loro reali classi di appartenenza. La localizzazione di ciascun oggetto è rappresentata attraverso un bounding box, ovvero un riquadro rettangolare che delimita l'area occupata dall'oggetto nella scena. Per ogni oggetto rilevato, il modello fornisce due informazioni principali: l'etichetta della classe di appartenenza e un punteggio di confidenza che quantifica il grado di certezza della previsione. Nell'esempio (Figura 3.7), il bounding box individua la regione dell'immagine occupata dal gatto, mentre l'etichetta "cat" fornisce la classificazione dell'oggetto (classe cat) e il valore "0.93" indica il livello di affidabilità della previsione (93%).



Figura 3.7: Esempio di output di object detection. Fonte [38].

3.7 Classification

La classificazione delle immagini è una tecnica di computer vision che consente di assegnare un'intera immagine a una tra un insieme di classi predefinite. In questo caso, l'output consiste in un'unica etichetta di classe e un punteggio di confidenza [39]. A differenza dell'object detection, la localizzazione degli oggetti nell'immagine non è rilevante.



Figura 3.8: Esempio di output di classification. Il modello classifica le immagini in due classi: "Frog" e "Toad". Fonte [40].

3.8 Architettura YOLO

Le precedenti reti R-CNN (Region-based Convolutional Neural Network) si basano su una pipeline multi-stage che prevede: la generazione di circa 2000 proposte di regioni, l'esecuzione di una CNN per ogni regione per estrarre le caratteristiche e classificarle, la regolazione delle coordinate dei bounding box e infine l'applicazione di non-maximum suppression per eliminare box duplicate. Questa architettura presenta due principali limitazioni: la complessità della pipeline e i tempi di inferenza molto elevati [41].

L'architettura YOLO è stata sviluppata da Joseph Redmon et al. nel 2016 e si propone di affrontare il problema di object detection come un'unica regressione, che va dai pixel dell'immagine ai bounding box con l'etichetta, offrendo prestazioni superiori. Il paper *You Only Look Once: Unified, Real-Time Object Detection* [42] è quello di riferimento e sarà seguito per illustrare i dettagli di questa architettura.

Pipeline

La pipeline di YOLO prevede tre step: il ridimensionamento dell'immagine di input, l'esecuzione di un'unica rete convoluzionale e la produzione dei bounding box con i valori di confidenza. La caratteristica innovativa sta nel fatto che l'intera immagine viene analizzata in un unico passaggio attraverso la rete, da cui deriva il nome YOLO, che sta per "You Only Look Once".

Funzionamento Generale

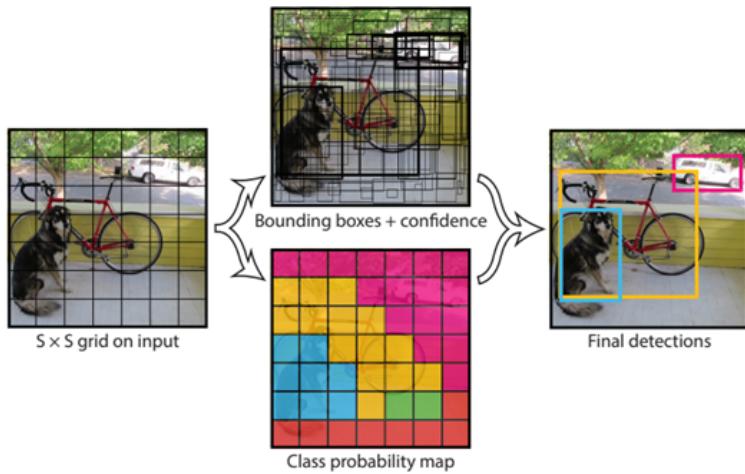


Figura 3.9: Funzionamento generale di YOLO. Fonte [42].

La rete ragiona globalmente sull'intera immagine, facendo predizioni su tutti i bounding box di tutte le classi configurate. L'immagine di input viene suddivisa in una griglia $S \times S$ di celle (Figura 3.9).

Ogni cella della griglia è responsabile degli oggetti il cui centro ricade al suo interno. Ogni cella predice B bounding box e i relativi punteggi di confidenza, che dipendono dal prodotto tra la probabilità che l'oggetto sia nella cella ($\text{Pr}(\text{Object})$) e dall'Intersection over Union (IoU), che misura l'overlap tra il box predetto e la ground truth. Ogni cella predice inoltre C probabilità di classe condizionate, una per ciascuna delle C classi, rappresentate come $\text{Pr}(\text{Class}_i \mid \text{Object})$.

Ogni bounding box è costituito da cinque predizioni: le coordinate x e y del centro della box rispetto alla cella, la larghezza w e l'altezza h della box rispetto all'immagine intera, e il punteggio di confidenza.

In fase di test il punteggio di confidenza viene calcolato moltiplicando le probabilità condizionate di classe per il punteggio di confidenza della box come segue:

$$\text{Pr}(\text{Class}_i \mid \text{Object}) \cdot \text{Pr}(\text{Object}) \cdot \text{IoU} = \text{Pr}(\text{Class}_i) \cdot \text{IoU}$$

Architettura della Rete

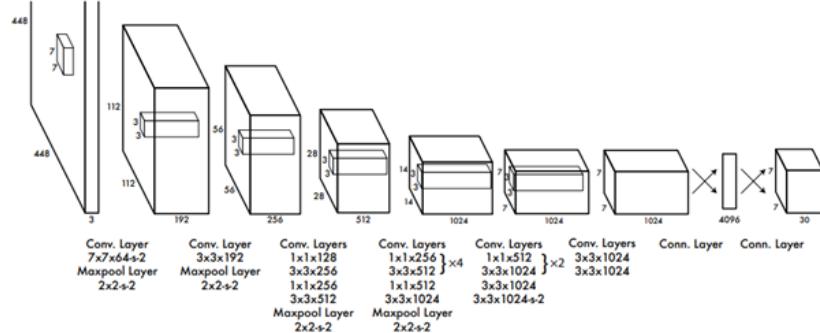


Figura 3.10: Architettura YOLO. Fonte [42].

L’architettura di YOLO è costituita da 24 layer convoluzionali seguiti da 2 layer fully connected. Come visto in precedenza, i layer convoluzionali si occupano dell’estrazione delle feature, mentre i layer completamente connessi predicono le coordinate dei bounding box e le classi di appartenenza.

Vantaggi, Limitazioni e Sviluppi Successivi

Il principale vantaggio di YOLO risiede nella sua elevata velocità: l’elaborazione dell’intera immagine in un’unica passata rende la rete estremamente efficiente, permettendone l’impiego in applicazioni in tempo reale. Inoltre, osservando l’intera immagine durante l’addestramento, YOLO ragiona globalmente sul contesto, producendo meno falsi positivi di background rispetto ad approcci basati su regioni locali.

Tuttavia, YOLO presenta anche alcune limitazioni, tra cui un maggior numero di errori di localizzazione e difficoltà nel rilevamento di oggetti piccoli e raggruppati, a causa della divisione in griglia dell’intera immagine.

Nel corso degli anni, le varie versioni di YOLO hanno introdotto miglioramenti architetturali e funzionali, incrementando le prestazioni in varie metriche e affermando YOLO come una delle architetture di riferimento della computer vision. Di seguito si riporta una sintesi dei principali sviluppi [43]:

- *YOLOv2 (2016)*: miglioramento del modello originale con introduzione di normalizzazione batch, anchor boxes e dimension clusters.
- *YOLOv3 (2018)*: incremento delle prestazioni grazie a una rete backbone più efficiente, ancore multiple e spatial pyramid pooling.
- *YOLOv4 (2020)*: introduzione di Mosaic data augmentation, detection head anchor-free e nuova loss function.

- *YOLOv5*: ulteriori miglioramenti delle prestazioni e aggiunta di ottimizzazione degli iperparametri.
- *YOLOv6 (2022)*: utilizzato in robot di consegna autonomi dell'azienda.
- *YOLOv7*: aggiunta di ulteriori funzionalità come la stima della posa sul dataset COCO keypoints.
- *YOLOv8 (2023)*: nuove funzionalità, supportando una gamma completa di attività di visione artificiale.
- *YOLOv9*: introduzione di metodi innovativi come Programmable Gradient Information (PGI) e Generalized Efficient Layer Aggregation Network (GELAN)
- *YOLOv10*: progressi nell'object detection in tempo reale, con l'introduzione di una head End-to-End che elimina la necessità di Non-Maximum Suppression (NMS)
- *YOLOv11*: ultima versione rilasciata da Ultralytics con ottime prestazioni in compiti di rilevamento di oggetti, segmentazione, stima della posa, tracking e classificazione. È la versione scelta per il progetto presentato in questa tesi.

Capitolo 4

Tecnologie Utilizzate

In questo capitolo vengono descritte sinteticamente le principali tecnologie e gli strumenti impiegati per la realizzazione del progetto oggetto di questa tesi.

4.1 Google Drive

Google Drive è un servizio di archiviazione cloud utilizzato per conservare i dataset e come directory principale del progetto. La cartella **FishersDetector** (Figura 4.1) contiene il file **dataset.zip**, che racchiude i dataset di detection e classification, e il file **training.ipynb**, un notebook Jupyter. Questa soluzione è stata scelta per sfruttare l'integrazione nativa tra Google Drive e Google Colab, facilitando l'accesso ai dati durante l'addestramento dei modelli.

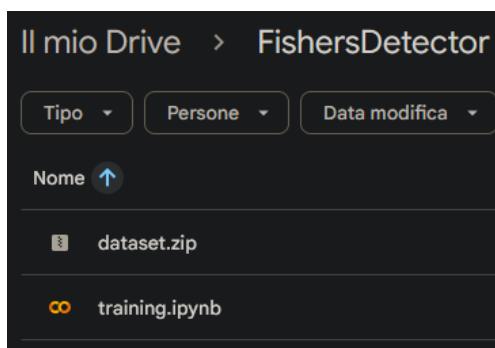


Figura 4.1: Struttura della cartella del progetto in Google Drive.

4.2 Google Colab

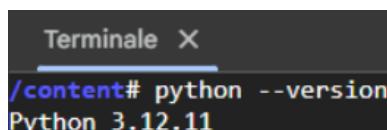
Google Colab è un ambiente di calcolo basato su Jupyter Notebook che fornisce accesso a GPU potenti, essenziali per l'addestramento di modelli di deep learning. Nel progetto, Colab è stato utilizzato per l'ottimizzazione dei parametri dei modelli e per il successivo addestramento. Le GPU impiegate durante l'esecuzione sono state la A100 e la T4, selezionate in base alla loro disponibilità. Il notebook `training.ipynb`, archiviato in Google Drive, viene eseguito in Colab permettendo di montare la cartella di Google Drive nell'ambiente di calcolo e accedere ai dataset. Questa soluzione si è rivelata efficace per l'addestramento dei modelli senza necessità di hardware potente in locale.

4.3 Visual Studio Code

Visual Studio Code (VS Code) è un ambiente di sviluppo integrato (IDE), utilizzato in questo progetto per il testing dei modelli addestrati in Google Colab. Dopo aver completato l'addestramento e il tuning in Colab, i modelli sono stati importati localmente e testati tramite script Python eseguiti in VS Code.

4.4 Python

Python è il linguaggio di programmazione principale utilizzato in questo progetto, nella versione 3.12.11. La sua velocità di sviluppo e l'ampia disponibilità di librerie dedicate al machine learning e alla computer vision lo rendono la scelta preferita per applicazioni di deep learning. Il progetto sfrutta diverse librerie Python, tra cui PyTorch per la gestione delle GPU e Ultralytics YOLO.



```
Terminale X
/content# python --version
Python 3.12.11
```

Figura 4.2: Versione Python.

4.5 Ultralytics YOLO

Ultralytics YOLO è una libreria Python che fornisce implementazioni ottimizzate dell'algoritmo YOLO (You Only Look Once). Nel contesto di questo progetto, Ultralytics YOLO è stata utilizzata per tutte le fasi di sviluppo dei modelli: addestramento su dataset personalizzati, ottimizzazione dei parametri e testing delle performance. È stata scelta la versione YOLOv11 sia per il rilevamento che per la classificazione.

Capitolo 5

Progettazione e Sviluppo del Software

Questo capitolo descrive lo sviluppo del sistema, illustrando l'approccio a due stadi (2-step) che ne costituisce l'architettura fondamentale. Come anticipato nell'Introduzione, il sistema suddivide l'analisi in due fasi consecutive: un primo modello individua tutte le persone presenti nella scena (object detection), mentre un secondo modello verifica se ciascuna persona sia un pescatore (classification).

Questa struttura consente di concentrare l'elaborazione sulle aree rilevanti, aumentando l'accuratezza della classificazione e riducendo interferenze e falsi positivi rispetto agli approcci tradizionali a stadio unico. Nel capitolo vengono inoltre presentati i dataset utilizzati, la ricerca dei parametri ottimali per i modelli addestrati e i test effettuati.

5.1 Approccio 2-Step

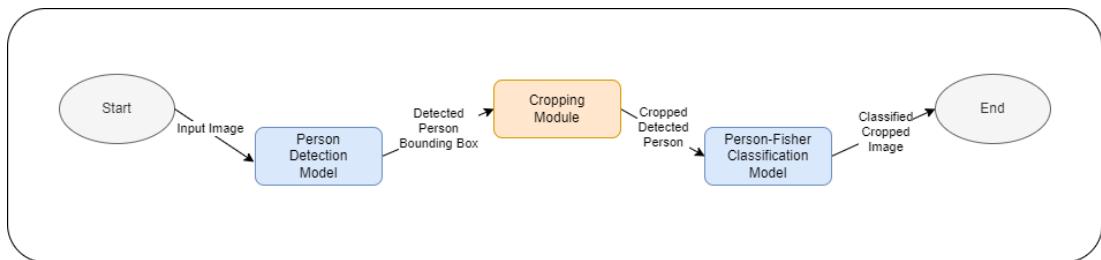


Figura 5.1: Architettura a 2-step. Diagramma realizzato con Draw.io.

Nello sviluppo di questo sistema è stato scelto un approccio a due stadi. Vengono impiegati due modelli: il primo si occupa del rilevamento delle persone all'interno della scena, il secondo svolge la classificazione distinguendo i pescatori dai normali passanti. Si tratta di un approccio consolidato in letteratura,

utilizzato ad esempio nella rilevazione dell'uso dello smartphone alla guida, che dimostra una precisione significativamente più alta rispetto agli approcci a singolo stadio (detection + classification), sebbene compri tempi di elaborazione più lunghi dovuti al coinvolgimento di due modelli [44].

L'architettura proposta (Figura 5.1) è composta da:

1. *Person Detection Model*: è il modello che riceve come input le immagini di spiagge e aree costiere e produce in output le immagini con i bounding box che intercettano le persone presenti nella scena.
2. *Cropping Module*: si tratta di un modulo che riceve in input le immagini contenenti i bounding box delle persone rilevate e produce in output un'immagine ritagliata per ogni bounding box. Il ritaglio viene centrato sulla persona rilevata con altezza pari al doppio di quella della persona stessa.
3. *Person-Fisher Classification Model*: è il modello che riceve come input le immagini ritagliate delle persone e produce in output le stesse immagini con l'annotazione che distingue tra pescatori e non pescatori.

I dataset per l'addestramento dei due modelli sono scelti tenendo presente la struttura di tale approccio.

5.2 Dataset Utilizzati

Sono stati creati due dataset:

1. *Mix Detection*: dataset utilizzato per addestrare il modello di detection per il rilevamento delle persone.
2. *Mix Classification*: dataset utilizzato per addestrare il modello di classification per la distinzione tra pescatori e persone.

Mix Detection

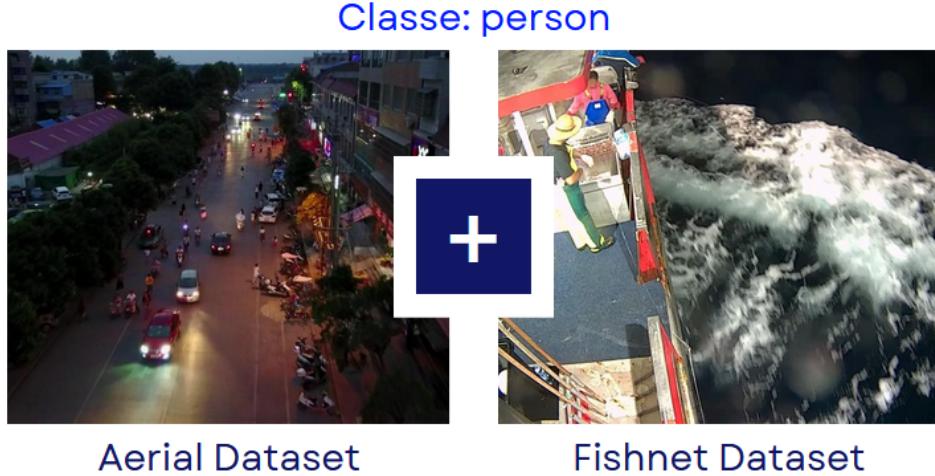


Figura 5.2: Mix Detection dataset. Figura realizzata in Canva con foto prese dal dataset.

Il dataset contiene un'unica classe annotata: "person". È stato costruito unendo e modificando due dataset già esistenti. Il primo di questi è Aerial Person Detection Dataset [45], selezionato poiché contiene immagini di persone da prospettiva aerea. Ciò si abbina bene all'applicazione del nostro sistema ideato per rilevare immagini di aree costiere da telecamere fisse o droni. Da tale dataset sono state rimosse le classi aggiuntive, come i veicoli, e le annotazioni per persone ferme o in movimento sono state unite in un'unica classe "person". Il secondo dataset utilizzato è Fishnet dataset [46]. Quest'ultimo contiene foto di pescatori da telecamere a bordo delle imbarcazioni ed è privo di annotazioni per canne da pesca. Il dataset risultante, Mix Detection, è stato suddiviso nel seguente modo: 1830 immagini di training, 520 immagini di validation e 260 immagini di test.

Mix Classification



Figura 5.3: Mix Classification dataset. Figura realizzata in Canva con foto prese dal dataset.

Il dataset contiene due classi: "person", "fishers". Anche questo è stato costruito unendo due dataset già esistenti. Il primo di questi è il Fishers dataset [47], contenente immagini di persone che pescano con canne da pesca. Il secondo dataset scelto è lo Human Detection Dataset [48]. Quest'ultimo contiene foto di persone ravvicinate, assenti nell'altro e utili in vista di acquisizioni con drone, e immagini di persone che svolgono azioni differenti dalla pesca. Il dataset risultante, Mix Classification, è stato diviso nel seguente modo: 387 immagini di "fishers" e 389 immagini di "person" per il training, 113 immagini di "fishers" e 112 immagini di "person" per la validation, 60 immagini di "fishers" e 59 immagini di "person" per il testing. Le immagini di training sono state incrementate attraverso operazioni di augmentation, tra cui distorsione angolare, rumore sui pixel e conversione in scala di grigi.

5.3 Modello di Detection

Come anticipato in precedenza, si tratta del modello di object detection che ha come obiettivo il rilevamento delle persone nelle immagini. È addestrato e validato sul dataset Mix Detection e l'unica classe da rilevare è "person". L'architettura scelta è YOLO, nella versione **YOLOv11-Large**.

5.3.1 Ottimizzazione dei Parametri - Detection

Per determinare i parametri ottimali sono state scelte manualmente sette configurazioni:

1. *config1_base_auto*: configurazione di base con ottimizzatore automatico (default di YOLO), utilizzata per il confronto con l'ottimizzatore SGD.
2. *config2_base_sgd*: configurazione di base con ottimizzatore SGD (Stochastic Gradient Descent).
3. *config3_lr_basso*: Configurazione con learning rate ridotto, per verificare se una loss più graduale favorisca una convergenza più o meno rapida.
4. *config4_lr_alto*: configurazione con learning rate aumentato, per verificare se la convergenza sia accelerata o la stabilità del modello sia compromessa.
5. *config5_momentum* configurazione con momentum aumentato, per verificare se una maggiore inerzia ai batch precedenti fornisca una loss più regolare.
6. *config6_dflalto*: configurazione con DFL (Distribution Focal Loss) aumentata per migliorare la precisione dei bounding box.
7. *config7_imgsridotto*: configurazione con dimensione delle immagini ridotta e batch size aumentata, in modo da valutare il trade-off tra velocità di elaborazione e accuratezza della detection su immagini a risoluzione inferiore.

I parametri utilizzati per le varie configurazioni sono stati consultati nella documentazione ufficiale di Ultralytics [49]. Nella Tabella 5.1 sono state riportate le configurazioni con i valori dei vari parametri, evidenziando quelli modificati rispetto ai valori di default.

Configurazione	Optimizer	LR	Momentum	DFL	Imgpsz	Batch
config1_base_auto	Auto	0.01	0.937	1.5	640	16
config2_base_sgd	SGD	0.01	0.937	1.5	640	16
config3_lr_basso	SGD	0.001	0.937	1.5	640	16
config4_lr_alto	SGD	0.1	0.937	1.5	640	16
config5_momentum	SGD	0.01	0.9	1.5	640	16
config6_dflalto	SGD	0.01	0.937	2.0	640	16
config7_imgsridotto	SGD	0.01	0.937	1.5	512	32

Tabella 5.1: Configurazioni testate per l’ottimizzazione della detection.

Una volta definite le configurazioni da testare, il modello è stato addestrato con ogni configurazione seguendo un setup comune (Figura 5.4): 50 epoch e patience impostata a 25 per fermare l’addestramento qualora il modello smettesse di migliorare dopo 25 epoch. Le configurazioni da testare sono state definite come lista di dizionari in Python, dove ogni dizionario contiene i parametri caratteristici della configurazione corrispondente.

```
model.train(
    device=0,
    data=config_path,
    project=config["project"],
    name=config["run_name"],
    imgsz=config["imgsz"],
    batch=config["batch"],
    epochs=config["epochs"],
    lr0=config["lr0"],
    momentum=config["momentum"],
    optimizer=config["optimizer"],
    dfl=config.get("dfl", 1.5),
    patience=config["patience"],
    save_period=config["save_period"],
)
```

Figura 5.4: Porzione di Codice Python per l’addestramento del modello con una configurazione generica (tratta dall’ambiente Google Colab).

La Tabella 5.2 mostra le configurazioni e le relative metriche prestazionali. Sono stati evidenziati i valori migliori.

Configurazione	Precision	Recall	mAP50	mAP50-95	Box_Loss
config1_base_auto	0.667	0.429	0.480	0.234	1.582
config2_base_sgd	0.725	0.466	0.532	0.271	1.439
config3_lr_basso	0.737	0.503	0.565	0.299	1.325
config4_lr_alto	0.455	0.270	0.263	0.108	2.148
config5_momentum	0.730	0.487	0.555	0.283	1.406
config6_dflalto	0.708	0.475	0.533	0.269	1.454
config7_imgsridotto	0.648	0.410	0.458	0.227	1.480

Tabella 5.2: Metriche delle prestazioni delle diverse configurazioni.

La metrica principale utilizzata per valutare le prestazioni è la *mean Average Precision (mAP50)*, che misura la precisione media considerando corrette solo le detection con una soglia di IoU maggiore o uguale a 0,5 [50].

La configurazione `config3_lr_basso` risulta la più performante su tutte le metriche mostrate. Al contrario, la configurazione `config4_lr_alto` è la meno efficace. Il learning rate risulta quindi un parametro di importanza critica per le prestazioni del modello.

Per quanto riguarda l'*optimizer*, SGD fornisce risultati migliori rispetto alla configurazione AUTO. Una lieve riduzione del momentum più basso migliora leggermente le prestazioni. L'aumento della *DFL* non ha un impatto significativo. Infine, la riduzione delle *dimensioni delle immagini* e l'aumento del *batch* peggiorano le prestazioni.

I parametri che hanno un impatto maggiore sulle prestazioni sono il *learning rate* e l'*optimizer*. Ciò stabilisce la loro priorità nell'ottimizzazione. Questi risultati suggeriscono inoltre che, per l'applicazione considerata e il dataset Mix Detection, una combinazione di **learning rate basso**, **optimizer SGD** e un **momentum leggermente incrementato** potrebbe rappresentare la configurazione ottimale per massimizzare le prestazioni del modello. Tuttavia, sono necessari ulteriori test per verificare la reale efficacia della configurazione proposta.

5.4 Modello di Classification

Il modello di Classification riceve immagini ritagliate dal Cropping Module contenenti una singola persona e classifica se si tratta di un pescatore (persone con canna da pesca) oppure di un non pescatore (persone senza canne da pesca). È addestrato e validato sul dataset Mix Classification e le classi configurate sono due: "fisher" e "person". L'architettura scelta è YOLO, nella versione **YOLOv11-Large Classification**.

5.4.1 Ottimizzazione dei Parametri - Classification

Per il modello di Classification è stato utilizzato il metodo `tune()` di Ultralytics YOLO [50] al fine di identificare i parametri ottimali.

La ricerca è stata configurata con 30 iterazioni (cioè, combinazioni di parametri), ciascuna addestrata per 30 epoche. La patience è stata impostata a 10.

I risultati mostrano che il modello con il miglior fitness ha raggiunto una *Top-1 accuracy* di **0.99** e una *loss di validazione* di **0.00853**. La Top-1 accuracy è una metrica che indica la percentuale di istanze in cui la classe reale corrisponde alla classe predetta dal modello con la probabilità più elevata. I parametri ottimali trovati sono riportati nella Tabella 5.3.

Parametro	Valore Ottimale
Learning Rate iniziale (lr0)	0.01
Momentum	0.937
Weight Decay	0.0005
Warmup Epochs	3.0

Tabella 5.3: Parametri ottimali individuati dal tuning automatico.

Nonostante l'elevata accuratezza registrata, tali risultati possono suggerire un dataset caratterizzato da dimensioni ridotte, con immagini e classi facilmente distinguibili.

Capitolo 6

Risultati

In questo capitolo vengono presentati i risultati ottenuti dal sistema sviluppato, illustrando le prestazioni raggiunte sui rispettivi dataset.

6.1 Risultati del Modello di Detection

Il modello di detection, addestrato utilizzando la configurazione `config3_lr_basso`, è stato valutato sul set di test del dataset Mix Detection (5.2, composto da 260 immagini. Le metriche di valutazione considerate includono la *mean Average Precision* con soglia IoU pari a 0.5 (*mAP50*), la *mean Average Precision* media calcolata sull'intervallo di soglie IoU da 0.5 a 0.95 (*mAP50-95*), la *Precision* e il *Recall*. I risultati ottenuti sono riportati nella Tabella 6.1.

Metrica	Valore
mAP50	0.561
mAP50-95	0.301
Precision	0.701
Recall	0.498

Tabella 6.1: Metriche del Modello di Detection.

I valori riportati evidenziano una buona capacità del modello nel localizzare e classificare correttamente gli oggetti di interesse. Tuttavia, il valore contenuto del Recall suggerisce che non tutte le persone presenti nelle immagini vengono rilevate.

6.1.1 Considerazioni - Detection

Nel complesso, il modello di detection mostra prestazioni soddisfacenti, soprattutto in relazione al valore di *mAP50* (**0.562**), che indica una buona accuratezza nel posizionamento dei bounding box. Il valore di *Precision* (**0.701**) suggerisce che la maggior parte delle rilevazioni positive è corretta. Tuttavia, il *Recall* (**0.498**) rivela che una parte non trascurabile delle persone presenti non viene rilevata. Queste prestazioni possono essere influenzate da diversi fattori: la variabilità delle condizioni di illuminazione e di posa, la presenza di soggetti parzialmente visibili o distanti e la dimensione limitata del dataset di addestramento. Un miglioramento delle prestazioni potrebbe essere ottenuto attraverso:

1. *Aumento del numero di epoch*: il modello è stato addestrato per sole 50 epoch e la funzione di loss risultava ancora in diminuzione al termine dell'addestramento, suggerendo un possibile margine di miglioramento.
2. *Addestramento con parametri ottimali*: poiché il modello è stato addestrato con la con la configurazione `config3_lr_basso`, è opportuno ripetere l'addestramento utilizzando la configurazione di parametri ottimali suggerita nel Capitolo 5 e valutarne le prestazioni.
3. *Miglioramento del dataset*: l'insieme dei dati potrebbe essere arricchito con immagini provenienti da angolazioni, condizioni di luce e scenari ambientali differenti, oppure ampliato attraverso tecniche di data augmentation per simulare le condizioni reali di ripresa.

Nonostante i limiti riscontrati, i risultati ottenuti restano promettenti e rispondono agli obiettivi prefissati per questa fase del progetto.

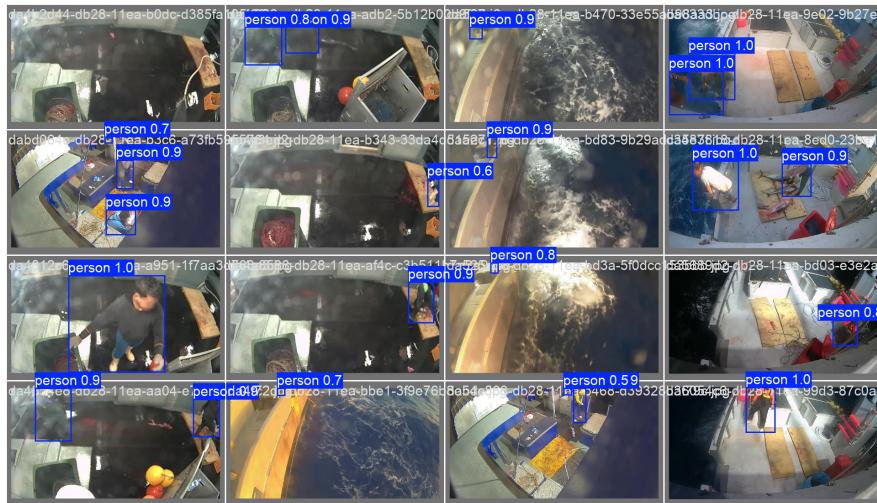


Figura 6.1: Output del modello di detection con bounding box.

6.2 Risultati del Modello di Classification

Il modello ottenuto dal tuning automatico è stato testato sul set di test del dataset Mix Classification (5.3), costituito da un totale di 119 immagini. I risultati mostrano una **Top-1 accuracy** del **99.16%** e un **tempo di inferenza medio** di **0.412** secondi per immagine.

La matrice di confusione (Figura 6.2) mostra come la classe "fisher" è stata predetta correttamente in 60 casi, ma confusa con "person" in un 1 caso. La classe "person" è stata predetta correttamente 58 volte e non è mai stata confusa.

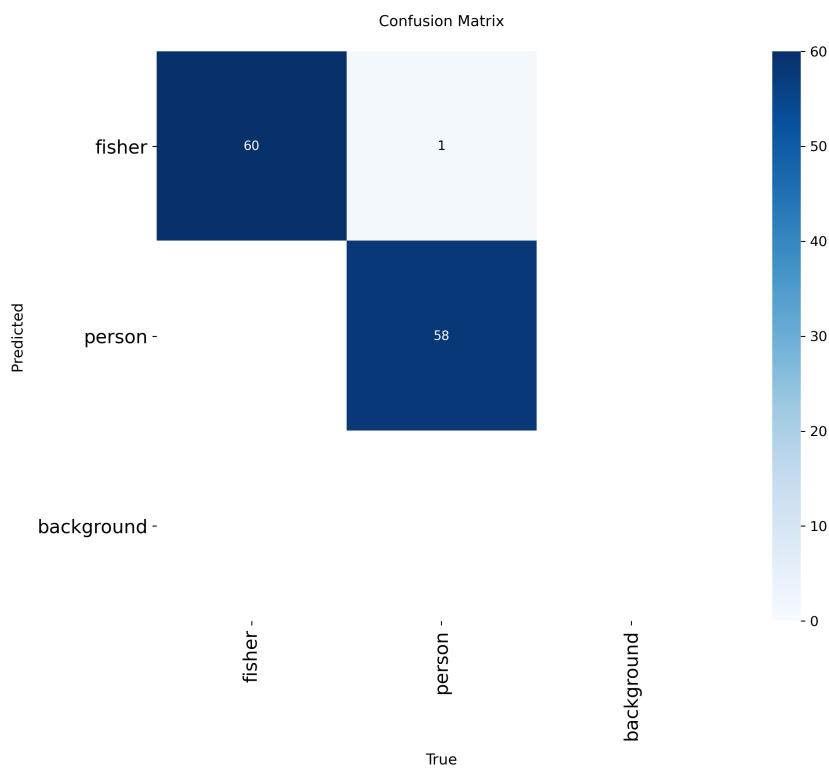


Figura 6.2: Matrice di confusione del test di classification (output del test).

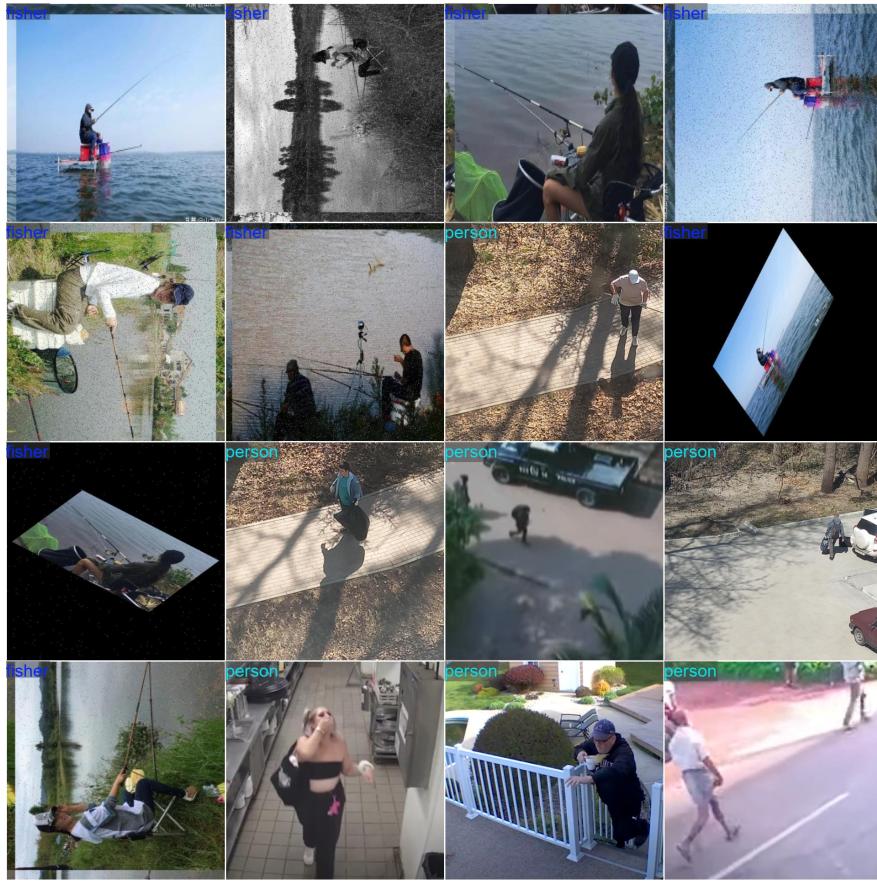


Figura 6.3: Output del modello di classification con le classi fisher e person.

6.2.1 Considerazioni - Classification

Il modello dimostra prestazioni eccellenti in termini di *accuratezza (99.16%)* e *tempo di inferenza (0.412 secondi per immagine)*. Tali risultati potrebbero essere motivati dalla semplicità del task e dal preprocessing del Cropping Module sulle immagini, riducendo la variabilità dei dati.

Tuttavia, restano alcuni aspetti critici da considerare. Il campione di test è ridotto (119 immagini) e potrebbe non rappresentare pienamente la complessità dei casi reali. Inoltre, in alcuni test mirati, è stato osservato che persone presenti su spiagge, che impugnano oggetti con forme riconducibili a canne da pesca, possono essere erroneamente classificate come pescatori. Un esempio significativo è riportato in Figura 6.4, dove un uomo dotato di metal detector viene identificato dal modello come pescatore (fisher).



Figura 6.4: Esempio di errore di classificazione del modello.

Restano quindi necessari ulteriori test su dataset differenti e di dimensioni maggiori per confermare la robustezza e la capacità di generalizzazione del modello.

Capitolo 7

Conclusioni

Il presente lavoro di tesi ha affrontato il problema della rilevazione e classificazione automatica dei pescatori ricreativi in aree costiere mediante un sistema sviluppato in due step consecutivi: detection e classification. L’obiettivo era sviluppare una soluzione efficiente e accurata per il monitoraggio automatizzato della pesca ricreativa in ambiente costiero, con un approccio già consolidato in letteratura che dimostra ottime prestazioni.

Il sistema proposto, basato su deep learning con architettura YOLO, ha raggiunto risultati promettenti per entrambi i task. Nello step di detection, implementato con YOLOv11-Large, il modello nella configurazione migliore (config3_lr_basso) ha registrato una precisione del 73.7%, un Recall del 50.3% e un mAP50 di 0.565, dimostrando buone capacità di rilevamento delle persone. Il modello di classification, basato su YOLOv11-Large Classification, ha raggiunto un’accuratezza del 99.16% con un tempo di inferenza medio di 0.412 secondi per immagine.

Il modello di detection risulta promettente: dopo le prime 50 epoche della configurazione più performante, il modello ha mostrato un trend di miglioramento con una loss che scendeva gradualmente. Sarebbe necessario un training intensivo con un numero di epoche maggiore e con i parametri ottimali individuati. Tuttavia, merita particolare attenzione il Recall piuttosto basso, che rappresenta una limitazione critica per applicazioni di monitoraggio.

Il modello di classification presenta prestazioni elevate, suggerendo che il task potrebbe essere semplificato dal preprocessing eseguito dal Cropping Module e dalla facilità nella distinzione tra persone con canna da pesca e persone comuni. Il modello dovrebbe essere testato su dataset più grandi contenenti immagini con maggiore variabilità e con persone che utilizzano oggetti simili a canne da pesca, in modo da appurare la sua elevata accuratezza.

Questo lavoro è di importanza significativa poiché si basa sull’impiego di tecnologie avanzate per il monitoraggio di attività il cui corretto svolgimento è cruciale per la salvaguardia dell’ambiente e della biodiversità. Come evidenziato in precedenza, le istituzioni stanno progressivamente integrando sistemi informati-

ci per garantire il rispetto delle norme. Tuttavia, è importante che tali sistemi rispettino la privacy.

L'architettura YOLO si è dimostrata sorprendentemente efficace e in continua evoluzione. Nel settore marino, ha già diverse importanti applicazioni: il rilevamento dei metodi che causano maggiore pesca accessoria, la valutazione e stima degli stock ittici, il rilevamento di plastica negli oceani e la conseguente rimozione, nonché il monitoraggio ambientale. Tutto ciò consente di comprendere meglio l'entità dei vari fenomeni e calibrare le nostre politiche di gestione costiera in modo più consapevole e sostenibile.

7.1 Sviluppi Futuri

Il presente lavoro apre diverse prospettive di sviluppo futuro. Il sistema potrebbe essere ottimizzato e ampliato secondo le seguenti direzioni:

1. *Estensione del dataset*: ampliare il dataset con immagini acquisite in diverse condizioni ambientali (variazioni orarie e meteorologiche) e con variazioni geometriche (punti di vista, distanze), per migliorare la generalizzazione del modello in scenari operativi reali.
2. *Rilevazione di ulteriori strumenti da pesca ricreativa*: estendere la detection per riconoscere pescatori che utilizzano strumenti diversi dalle canne da pesca, quali nasse, lenze a mano e altri dispositivi comunemente impiegati. L'inclusione di tali categorie amplierebbe la copertura del sistema e aumenterebbe la capacità di individuare diverse tipologie di attività di pesca.
3. *Estensione del sistema ad ambienti lacustri e fluviali*: estendere l'applicabilità del sistema a laghi, fiumi e bacini artificiali, dove la pesca ricreativa è altrettanto diffusa, attraverso l'acquisizione di dataset per contesti con vegetazione e morfologia territoriale differente.
4. *Ottimizzazione del modello per l'edge computing*: ridurre la complessità computazionale al fine di consentire l'esecuzione del sistema su dispositivi embedded (quali droni o telecamere intelligenti), garantendo capacità di monitoraggio in tempo reale direttamente sul campo.
5. *Integrazione di un sistema di notifica automatica*: sviluppare un modulo per la segnalazione in tempo reale di attività di pesca in aree soggette a fermo pesca. Tale funzionalità aumenterebbe ulteriormente il controllo e la tutela delle risorse ittiche.
6. *Stima del numero di pescatori rilevati*: analizzare i dati raccolti a intervalli regolari (ad esempio annuali) e confrontarli con i dati di comunicazione della

pesca sportiva e con quelli acquisiti dalle app di segnalazione delle catture, al fine di individuare eventuali discrepanze e produrre analisi statistiche più accurate e affidabili sul fenomeno.

L'implementazione di questi sviluppi porterebbe il sistema verso un'applicabilità concreta nel supporto delle politiche di gestione e controllo della pesca ricreativa.

Bibliografia

- [1] European Commission. Tackling illegal, unreported and unregulated (IUU) fishing 2021.
https://oceans-and-fisheries.ec.europa.eu/publications/tackling-illegal-unreported-and-unregulated-iuu-fishing_en.
- [2] European Commission. Control regulation (Regulation (EU) 2023/2842) 2023. https://oceans-and-fisheries.ec.europa.eu/fisheries/rules/enforcing-rules/control-regulation_en.
- [3] European Parliament and Council. Regulation (EU) 2023/2842 of the European Parliament and of the Council 2023.
Official Journal of the European Union, L 2023/2842.
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202302842.
- [4] European Commission. Commission Implementing Regulation (EU) 2025/274 2025.
Official Journal of the European Union, L 2025/274.
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L_202500274.
- [5] European Parliament. Draft Report on the State of Play of Recreational Fisheries in the European Union (2017/2120(INI)) 2017.
Committee on Fisheries, Rapporteur: Norica Nicolai.
- [6] Food and Agricultural Organization for the United Nations. What is IUU fishing? 2025. <https://www.fao.org/iuu-fishing/background/what-is-iuu-fishing/en/>.
- [7] U.S. Government Accountability Office. Combating Illegal Fishing: Clear Authority Could Enhance U.S. Efforts to Partner with Other Nations at Sea.
<https://www.gao.gov/assets/gao-22-104234.pdf>.

- [8] David J. Agnew, John Pearce, Ganapathiraju Pramod, Tom Peatman, Reg Watson, John R. Beddington, Tony J. Pitcher. Estimating the Worldwide Extent of Illegal Fishing 2009.
<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0004570>.
- [9] U. R. Sumaila, D. Zeller, L. Hood, M. L. D. Palomares, Y. Li, D. Pauly. Illicit trade in marine fish catch and its effects on ecosystems and people worldwide 2020.
<https://www.science.org/doi/10.1126/sciadv.aaz3801>.
- [10] Food and Agricultural Organization for the United Nations. Illegal, Unreported and Unregulated (IUU) fishing 2025.
<https://www.fao.org/iuu-fishing/en/>.
- [11] WWF. Illegal, Unreported and Unregulated fishing 2025. <https://www.wwf.org.uk/what-we-do/illegal-unreported-unregulated-fishing>.
- [12] Poseidon Aquatic Resource Management Ltd., Global Initiative Against Transnational Organized Crime. The Illegal, Unreported and Unregulated Fishing Risk Index 2023.
<https://iuufishingindex.net/downloads/IUU-Report-2023.pdf>.
- [13] Food and Agricultural Organization for the United Nations. Technical Guidelines for Responsible Fisheries 2012.
<https://www.fao.org/4/i2708e/i2708e00.pdf>.
- [14] Font, T., Lloret, J., & Piante, C. Recreational Fishing Within Marine Protected Areas In the Mediterranean Marine Protected Areas in the Mediterranean 2012.
https://www.researchgate.net/publication/273604344_Recreational_Fishing_Within_Marine_Protected_Areas_In_the_Mediterranean.
- [15] Parc national des Calanques. Guide à l'usage de la pêche de loisir 2018.
<https://calanques-parcnational.fr/en/node/11749>.
- [16] Gómez, S., Carreño, A., Sánchez, E., Martínez, E., Lloret, J. Safeguarding Marine Protected Areas in the growing Mediterranean Blue Economy. Recommendations for Recreational Fisheries 2019.
https://d2ouvy59p0dg6k.cloudfront.net/downloads/recreational_31july_2_single_page.pdf.
- [17] European Commission. Come funziona il programma RecFishing? 2025.
<https://rcsb-public.herokuapp.com/how>.

- [18] European Commission, Directorate-General for Maritime Affairs and Fisheries. Call for tenders EC-MARE/2024/OP/0003 – RecFishing: an electronic system for catch reporting of recreational fisheries. Tender Specifications. Part 2: Technical Specifications 2024. https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/tender-details/docs/3185ecff-b39e-4928-a893-149b0276a3d-CN/Tender%20Specifications%20-%20Part%20%20-%20Technical%20Specifications_EC-MARE-2024-OP-0003%20-%20F_V1.pdf.
- [19] Capitanerie di porto Guardia Costiera. Comunicazione di pesca sportiva e ricreativa 2025. <https://www.guardiacostiera.gov.it/portale/comunicazione-di-pesca-sportiva-e-ricreativa>.
- [20] NOAA. About the Marine Recreational Information Program 2025. <https://www.fisheries.noaa.gov/recreational-fishing-data/about-marine-recreational-information-program#methods>.
- [21] National Research Council. Review of Recreational Fisheries Survey Methods. Washington, DC: The National Academies Press. 2006. <https://doi.org/10.17226/11616>.
- [22] NOAA. Recreational Electronic Reporting At-a-Glance. <https://www.fisheries.noaa.gov/recreational-fishing-data/recreational-electronic-reporting-glance#what-are-the-benefits-and-limitations-of-angler-reporting-apps>.
- [23] Javier Menéndez-Blázquez, Mara Segovia, David March. Leveraging Earth Observation Data to Monitor Boat-Based Recreational Fishing. <https://doi.org/10.1111/faf.70017>.
- [24] Abirami Vina, Ultralytics. Everything you need to know about Computer Vision in 2025. <https://www.ultralytics.com/blog/everything-you-need-to-know-about-computer-vision-in-2025>.
- [25] Richard Szeliski. Computer Vision: Algorithms and Applications, 2nd ed. 2022. <https://library.huree.edu.mn/data/202295/2024-06-03/Computer%20Vision%20-%20Algorithms%20and%20Applications%202nd%20Edition,%20Richard%20Szeliski.pdf>.
- [26] IBM. Cos'è la computer vision? 2025. <https://www.ibm.com/it-it/think/topics/computer-vision>.
- [27] GeeksforGeeks. Computer Vision Introduction 2025. <https://www.geeksforgeeks.org/computer-vision/computer-vision-introduction/>.

- [28] Abirami Vina, Ultralytics. Computer vision vs image processing: The key differences 2025. <https://www.ultralytics.com/blog/computer-vision-vs-image-processing-the-key-differences>.
- [29] Dimi Fred, MathWorks. Introducing Deep Learning with MATLAB 2018. https://www.academia.edu/37429302/Introducing_Deep_Learning_with_MATLAB.
- [30] GeeksforGeeks. Supervised Machine Learning 2025. <https://www.geeksforgeeks.org/machine-learning/supervised-machine-learning/>.
- [31] Wikipedia. Convolutional Neural Network 2025. https://en.wikipedia.org/wiki/Convolutional_neural_network.
- [32] Ultralytics. Convolutional Neural Network (CNN) 2025. <https://www.ultralytics.com/glossary/convolutional-neural-network-cnn>.
- [33] Stanford University. Convolutional Neural Networks (CNNs / ConvNets) 2025. <https://cs231n.github.io/convolutional-networks/>.
- [34] Wikipedia. Convolutional layer 2025. https://en.wikipedia.org/wiki/Convolutional_layer.
- [35] NVIDIA Developer. Convolution 2025. <https://developer.nvidia.com/discover/convolution>.
- [36] Stanford University. Neural Networks Part 2025. <https://cs231n.github.io/neural-networks-1/>.
- [37] Wikipedia. Object detection 2025. https://en.wikipedia.org/wiki/Object_detection.
- [38] Michele V. CatDetector 2025. <https://github.com/michele-vigliotta/CatDetector>.
- [39] Ultralytics. Image Classification 2025. <https://docs.ultralytics.com/tasks/classify/>.
- [40] User auxeno on kaggle. Frogs Toads and EfficientNet - CV 2023. <https://www.kaggle.com/code/auxeno/frogs-toads-and-efficientnet-cv/notebook>.
- [41] Ultralytics. What is R-CNN? A quick overview 2024. <https://www.ultralytics.com/blog/what-is-r-cnn-a-quick-overview>.

- [42] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection 2016.
<https://arxiv.org/pdf/1506.02640.pdf>.
- [43] Ultralytics. YOLO: A Brief History 2025.
<https://docs.ultralytics.com/#yolo-a-brief-history>.
- [44] Steven Carrell, Amir Atapour Abarghouei. Identification of Driver Phone Usage Violations via State-of-the-Art Object Detection with Tracking 2021. https://www.researchgate.net/publication/354400665_Identification_of_Driver_Phone_Usage_Violations_via_State-of-the-Art_Object_Detection_with_Tracking.
- [45] Aerial Person Detection. Aerial Person Detection Computer Vision Dataset 2022. <https://universe.roboflow.com/aerial-person-detection/aerial-person-detection>.
- [46] The Nature Conservancy. Fishnet.AI: AI Training Data for Fisheries 2022.
<https://www.fishnet.ai/>.
- [47] cio A. 01 Computer Vision Model 2023.
<https://universe.roboflow.com/cio-a-d4yjw/01-3js62>.
- [48] Konstantin Verner. Human Detection Dataset 2021. <https://www.kaggle.com/datasets/constantinwerner/human-detection-dataset/data>.
- [49] Ultralytics. Configuration 2025.
<https://docs.ultralytics.com/usage/cfg/>.
- [50] Ultralytics. Ultralytics YOLO Hyperparameter Tuning Guide 2025.
<https://docs.ultralytics.com/guides/hyperparameter-tuning/>.

Elenco delle figure

1.1	Stima dell'impatto della pesca illegale, non dichiarata e non regolamentata	6
2.1	Classificazione pesca illegale, non dichiarata e non regolamentata	11
2.2	Grafico IUU Risk Index	12
2.3	Punteggio globale IUU Fishing Risk Index 2023	13
2.4	Guida del Parco Nazionale delle Calanques (Francia), dimensioni minime per le specie pescate a livello ricreativo	14
2.5	Mappa della distribuzione pesca ricreativa	15
2.6	Architettura del sistema RecFishing con le tre opzioni disponibili di connessione tra app e server centrale della Commissione	17
3.1	Esempio di applicazione dell'algoritmo di edge detection	21
3.2	Confronto tra machine learning e deep learning	22
3.3	Confronto tra supervised learning e unsupervised learning	23
3.4	Architettura generica di una CNN	24
3.5	Esempio operazione di convoluzione	25
3.6	Applicazione del maxpooling	26
3.7	Esempio di output di object detection	27
3.8	Esempio di output di classification	28
3.9	Funzionamento generale di YOLO	29
3.10	Architettura YOLO	30
4.1	Struttura della cartella del progetto in Google Drive	32
4.2	Versione Python	33
5.1	Architettura a 2-step	34
5.2	Mix Detection	36
5.3	Mix Classification	37
5.4	Porzione di codice Python per l'addestramento del modello con configurazione generica	39
6.1	Output del modello di detection con bounding box	43
6.2	Matrice di confusione del test di classification	44

6.3	Output del modello di classification (classi fisher e person)	45
6.4	Esempio di errore di classificazione del modello	46

Ringraziamenti

In questa sezione vorrei riflettere brevemente sul percorso concluso e ringraziare chi mi è stato accanto. È stato un cammino al contempo stimolante e impegnativo. Le discipline affrontate in questo corso insegnano davvero a essere umili, ricordandoci che c'è sempre qualcosa di nuovo da imparare.

Voglio ringraziare il Dr. Galdelli per aver reso l'apprendimento dell'informatica coinvolgente e appassionante, sia durante il corso di Architetture dei Calcolatori e Cloud Computing, sia durante questo tirocinio. Entrambe le esperienze sono state fondamentali e mi hanno permesso di approfondire tematiche di grande interesse.

Ringrazio la mia famiglia, per avermi sempre supportato e incoraggiato, sprovvandomi e credendo nelle mie capacità. Un ringraziamento sincero anche a tutti i miei amici, che mi hanno aiutato sul piano scolastico, mentale e umano, talvolta anche indirettamente senza saperlo.

A voi vanno i miei ringraziamenti più profondi. Mi reputo fortunato per i legami che ho e per aver avuto l'opportunità di studiare ciò che mi appassiona.