



UNIVERSITÀ  
POLITECNICA  
DELLE MARCHE

# Sviluppo di un sistema AI per la detection di imbarcazioni della piccola pesca

- *Davide Ticchiarelli*
- *Giampaolo Marino*
- *Niccolò Ciotti*

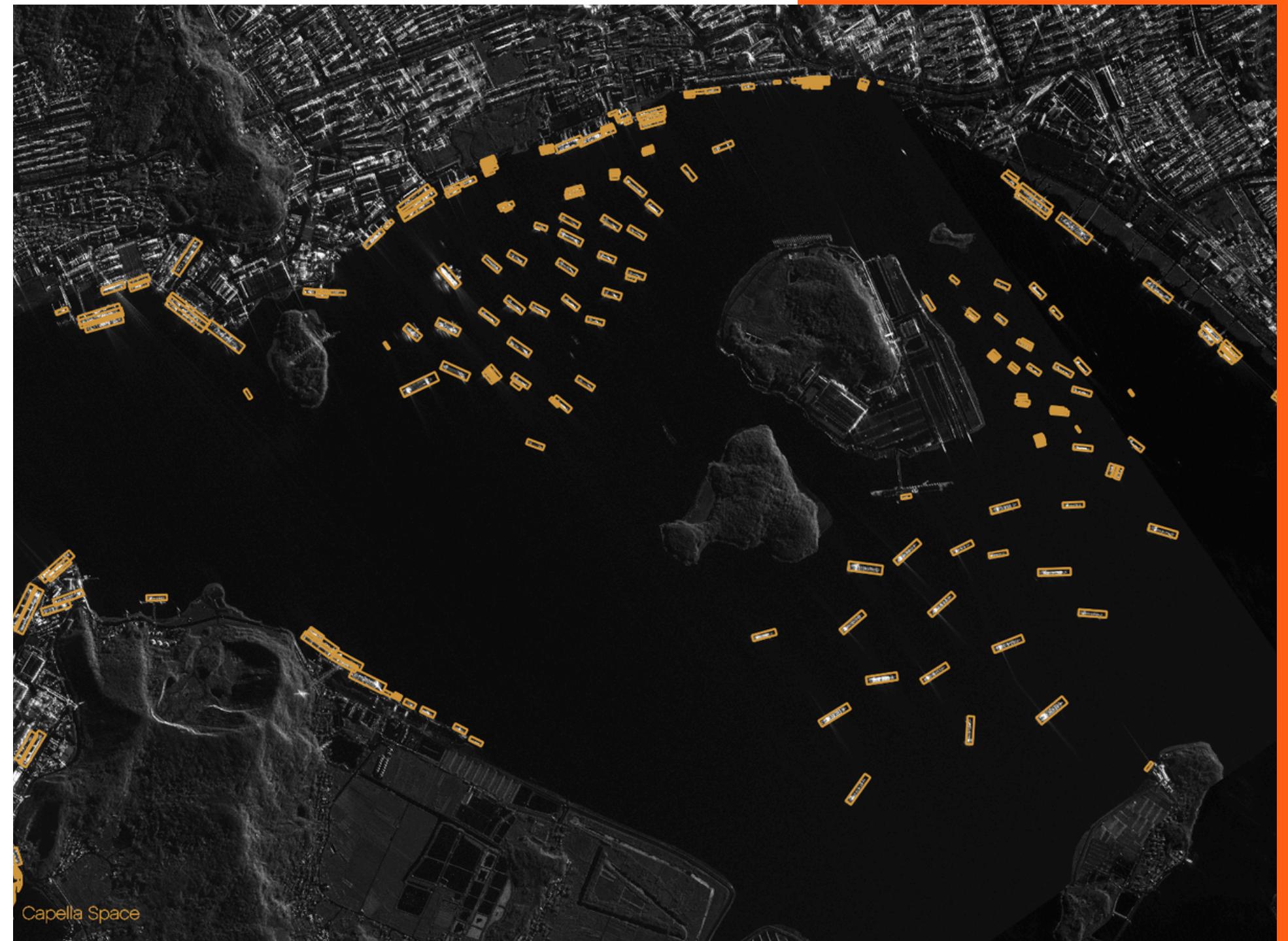


# Indice

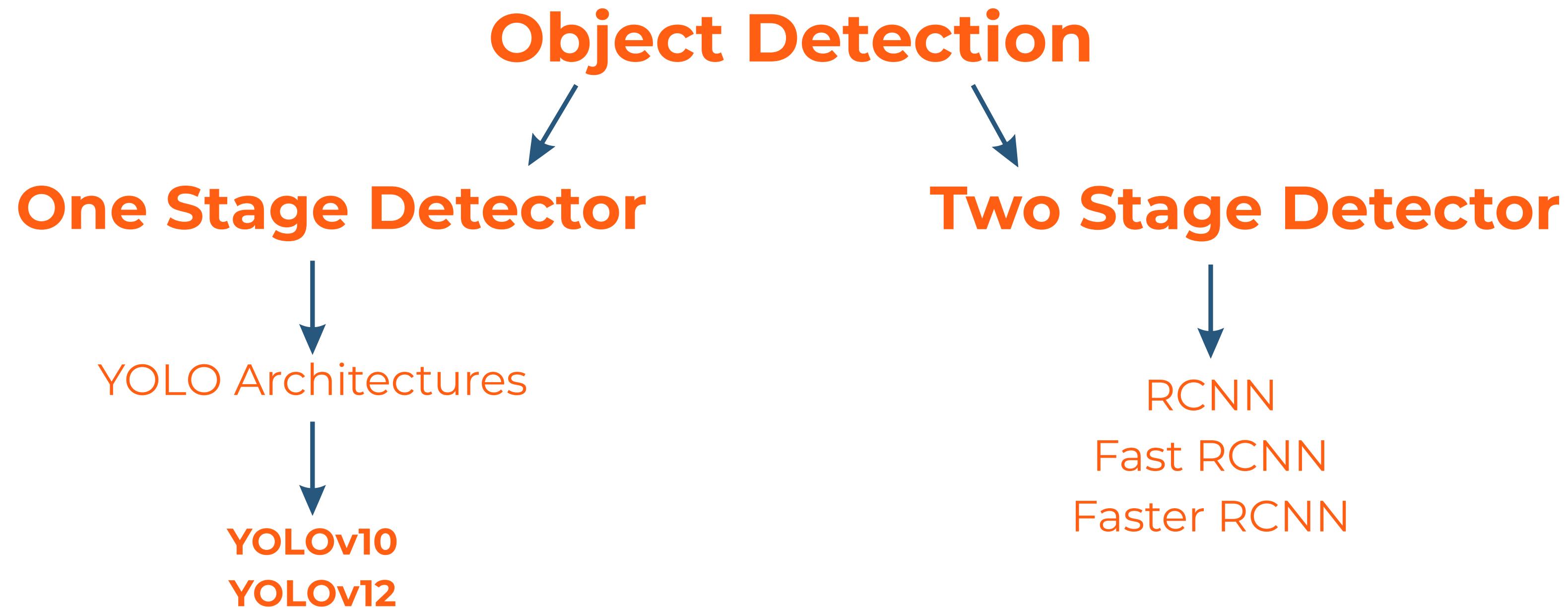
- Introduzione
- Stato dell'Arte
- Metodi e materiali
- Confronto tra baseline
- Modifiche architetturali
- Risultati
- Sviluppi futuri

# Introduzione

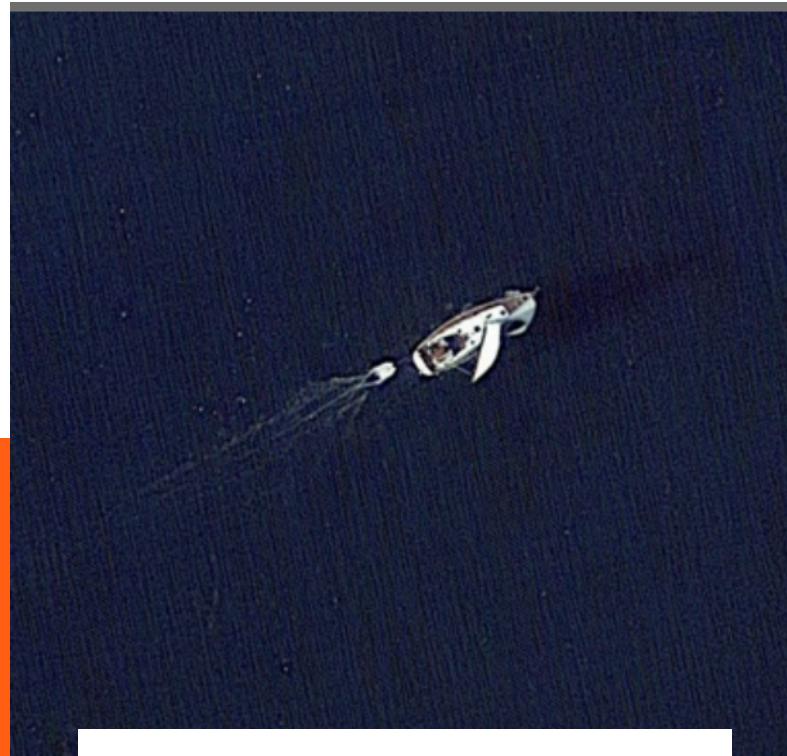
Il progetto si propone di trovare il miglior modello AI capace di individuare imbarcazioni di piccole dimensioni in immagini satellitari. I dati SAR forniscono immagini ad alta risoluzione indipendentemente da luce solare, condizioni meteo o visibilità notturna, consentendo un monitoraggio marittimo continuo e affidabile anche in presenza di nuvole o scarsa illuminazione.



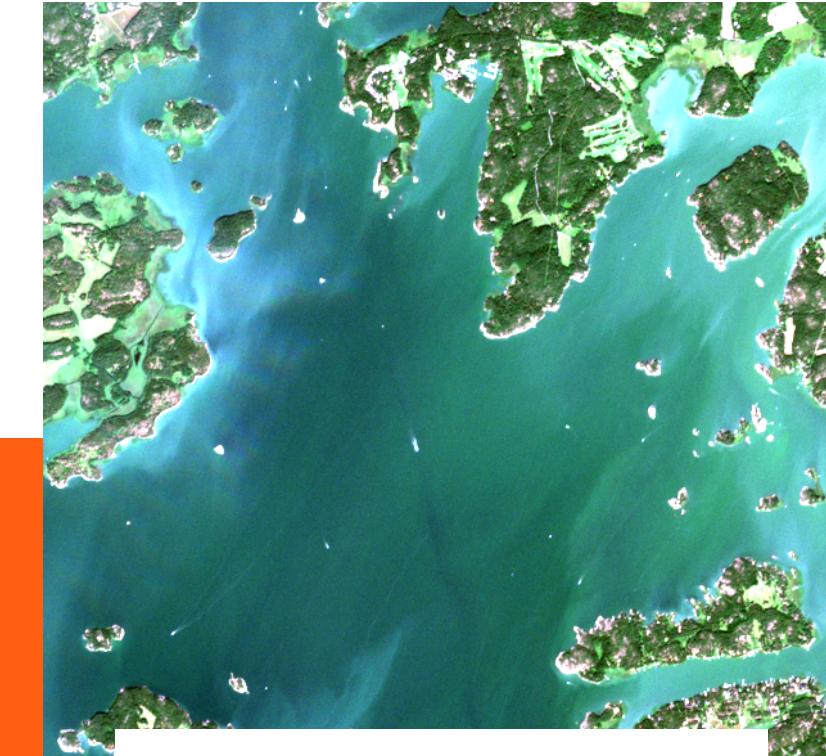
# Stato dell'Arte



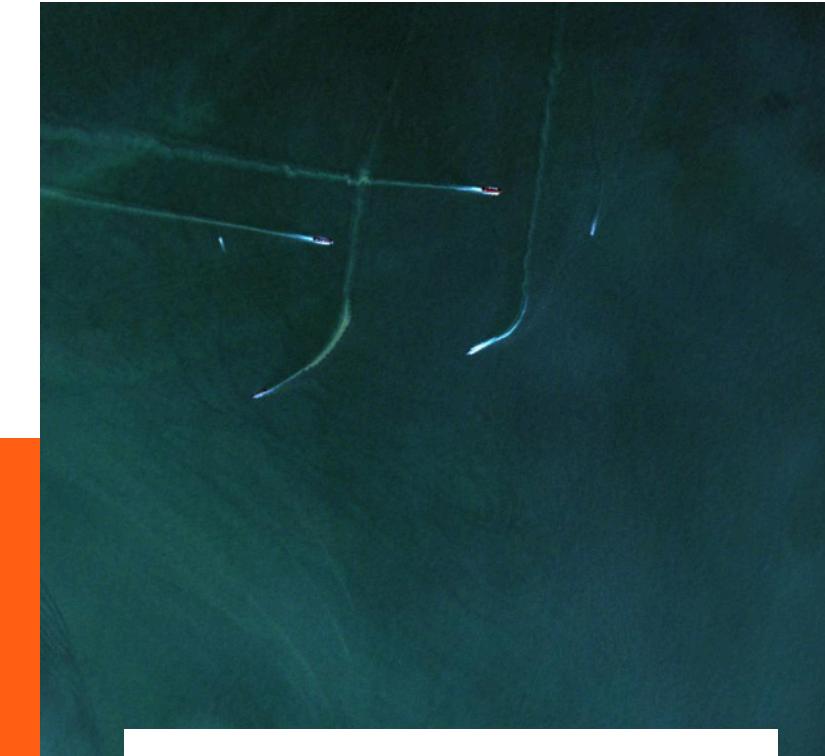
# Metodi e materiali - Dataset



**SDAI**



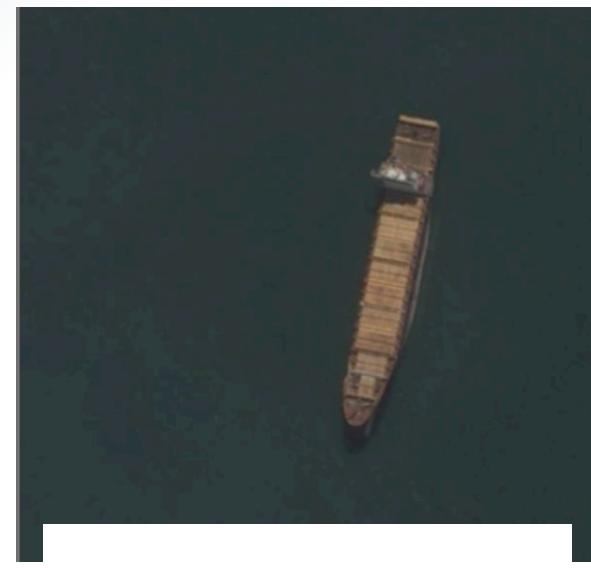
**S2\_FC**



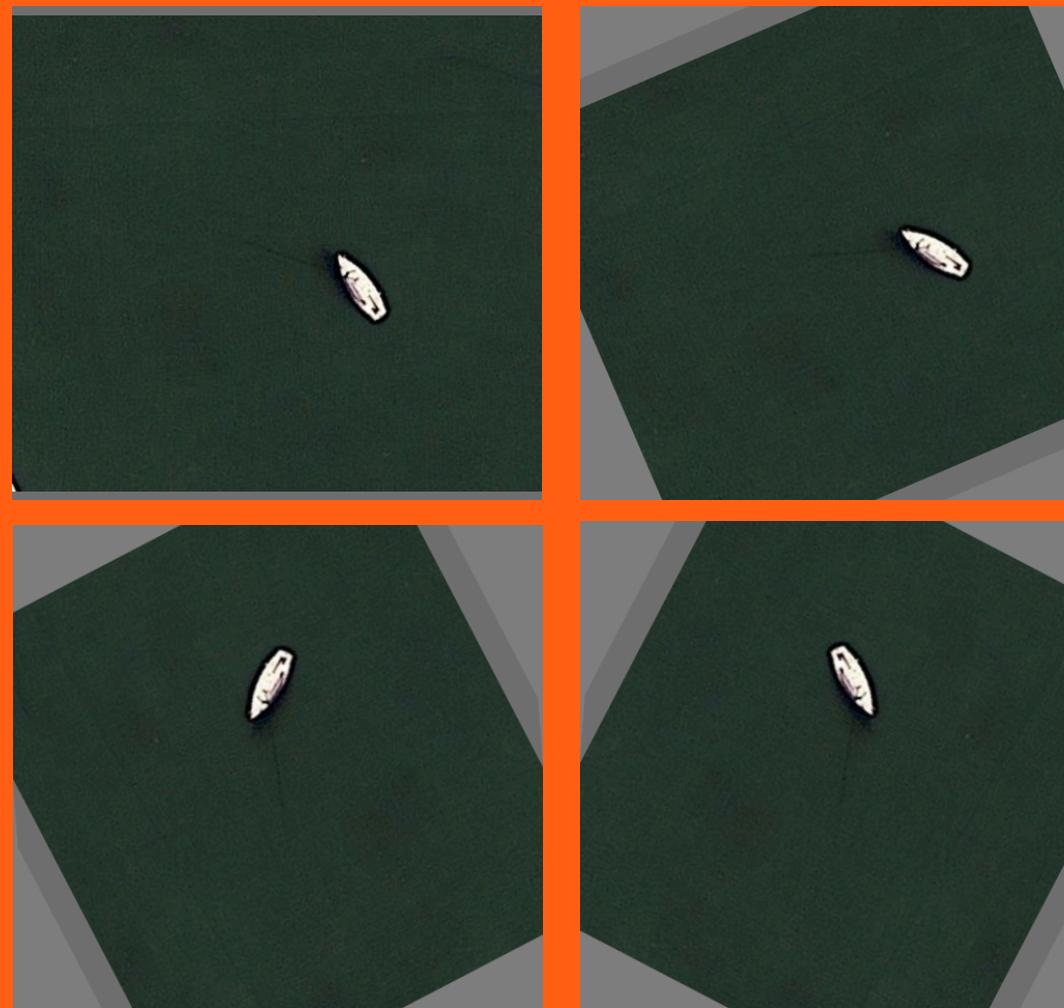
**S2\_Detection**

Il dataset originale SAR comprende diversi sottogruppi di immagini provenienti fonti diverse; tra questi sono stati esclusi quelli basati su immagini Sentinel-1, la cui risoluzione spaziale relativamente grossolana ne compromette l'affidabilità nel rilevamento di imbarcazioni di piccole dimensioni. Di conseguenza, sono stati mantenuti unicamente i sottodataset S2\_Detection, S2\_FC e SDAI, tutti basati su immagini ad alta risoluzione

# Metodi e materiali - Data augmentation



SDAI



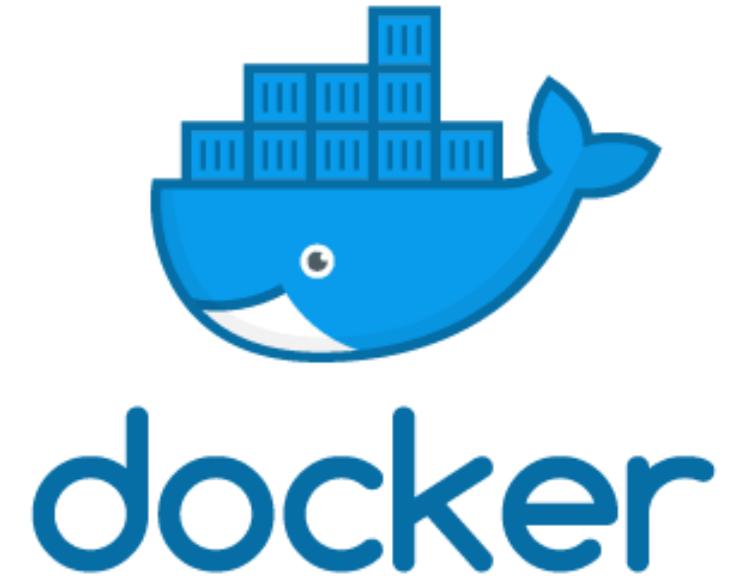
Nel sottodataset SDAI, in cui alcuni dati derivano da riprese effettuate con drone, è stato effettuato un preprocessing aggiuntivo: sono state eliminate le immagini in cui le imbarcazioni risultavano troppo grandi, in quanto avrebbero introdotto un'eccessiva variabilità di scala e avrebbero potuto sbilanciare l'addestramento dei modelli, compromettendone l'efficacia.

Per incrementare ulteriormente la varietà e la dimensione del dataset sono state applicate tre trasformazioni:

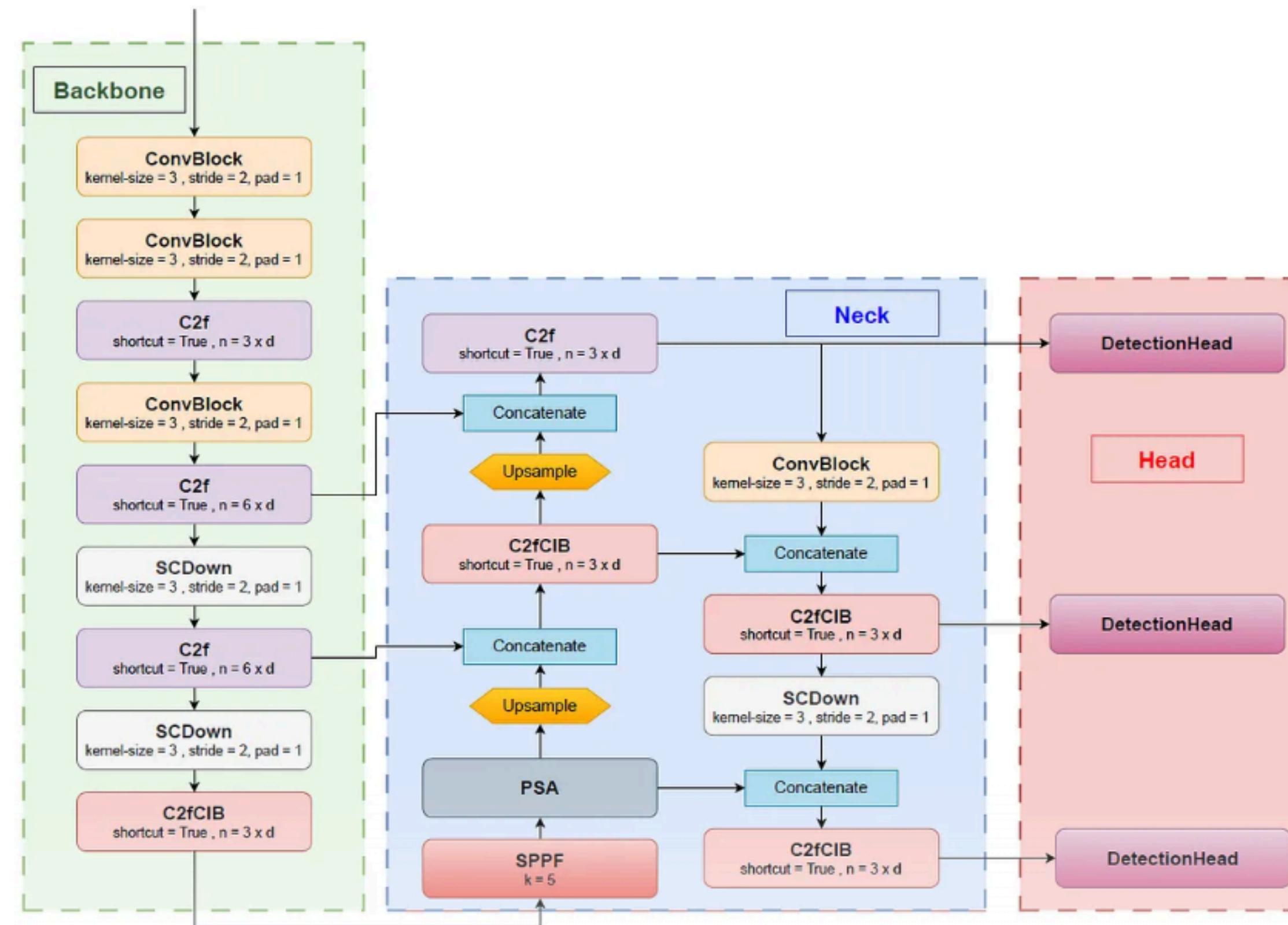
- prima rotazione casuale in senso orario compresa tra  $10^\circ$  e  $35^\circ$ .
- seconda rotazione casuale in senso orario compresa tra  $35^\circ$  e  $95^\circ$ .
- specchiatura/ribaltamento verticale

# Ambiente di sviluppo

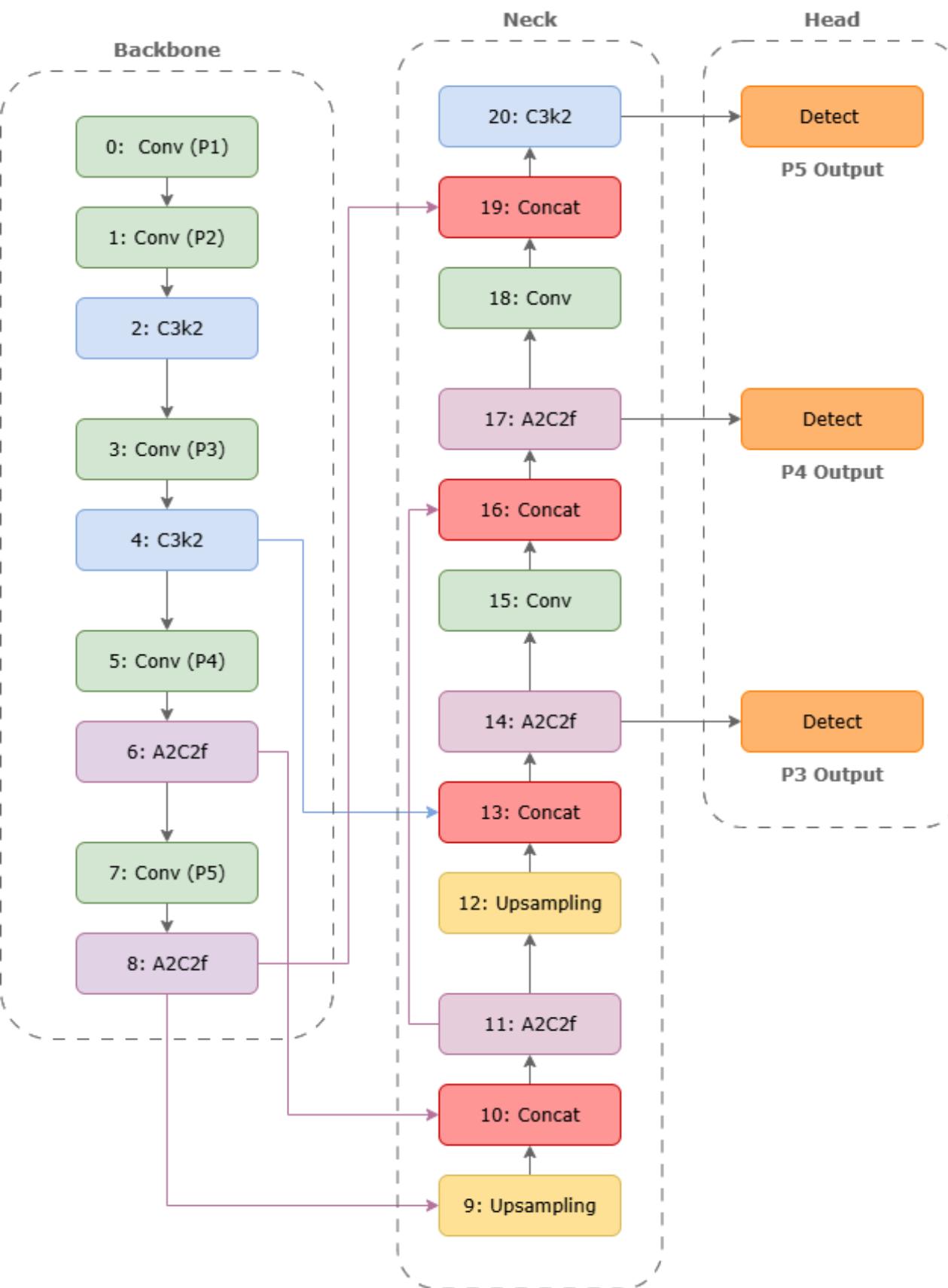
- Server VRAI con CPU Intel(R) Xeon(R) Gold 6254 @ 3.10GHz e GPU NVIDIA RTX A6000.
- Ambiente isolato tramite container Docker.



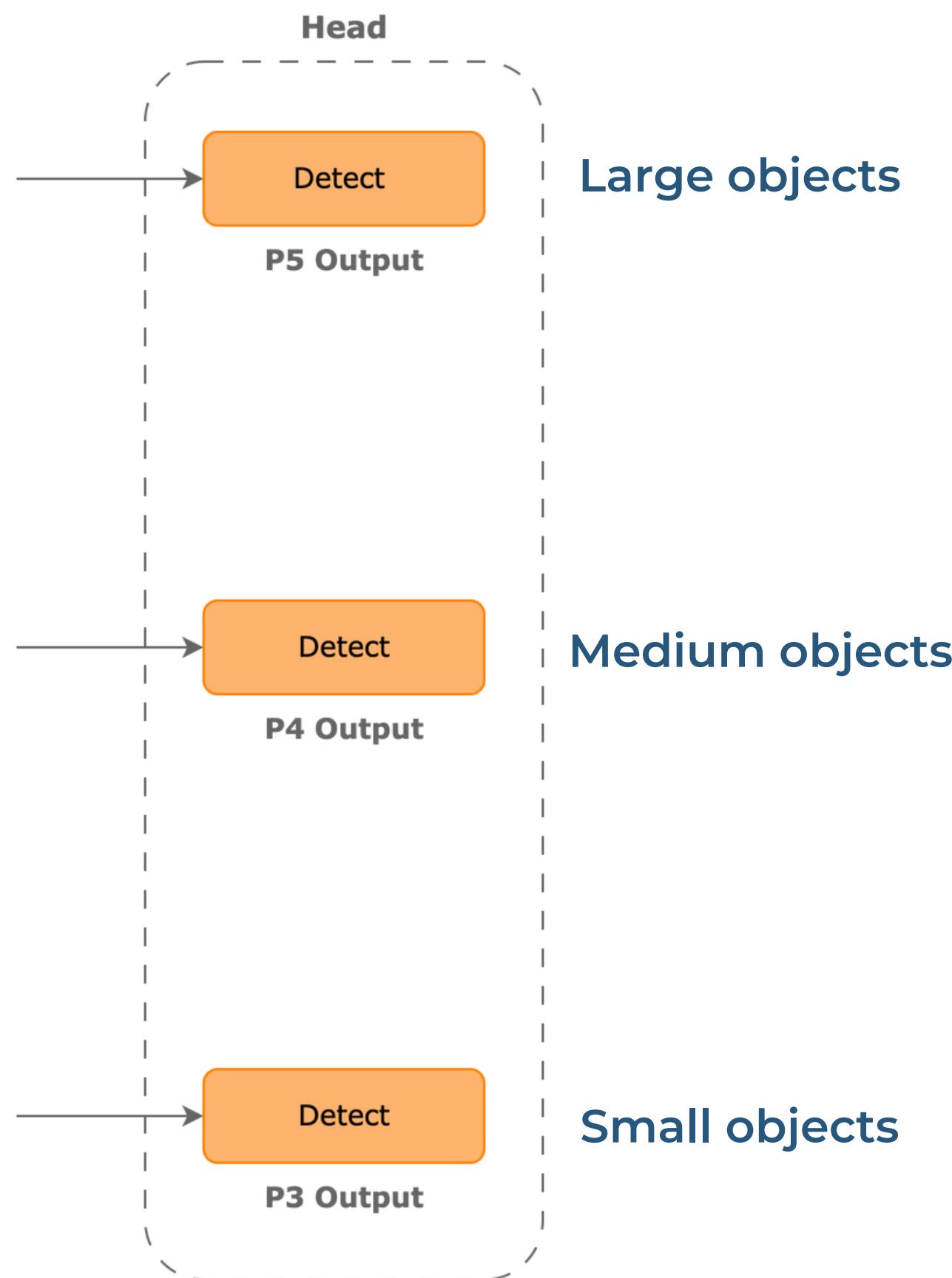
# Architettura YOLOv10



# Architettura YOLOv12



# YOLO V10/V12 - Head



- **3 detection heads (piccoli, medi, grandi oggetti)**
- **Migliora la capacità di rilevare target su tutte le scale**
- **Massimizza precisione e robustezza della rete**

# Confronto baseline

Modello	Precision	Recall	mAP50	mAP50-95	Parameters	GFLOPs
YOLOv10sLight	0,688	0,553	0,62	0,271	4.930.866	19,4
YOLOv10sMedium	0,689	0,563	0,634	0,277	6.057.140	23
YOLOv10mLight	0,704	0,578	0,642	0,281	11.437.218	54,9
YOLOv12sLight	0,691	0,589	0,643	0,273	4.514.321	15,3
YOLOv12sMedium	0,719	0,593	0,652	0,281	5.201.490	17,5
YOLOv12mLight	0,732	0,587	0,649	0,281	13.276.817	56,7

# Scelta iperparametri



- Gli iperparametri derivano da un tuning effettuato in un progetto precedente.
- A causa di risorse limitate e tempi ristretti, non è stato possibile ripetere il processo di ottimizzazione.
- È stata quindi adottata una combinazione già validata come ottimale, garantendo comunque buone performance.

Parametro	Valore
epochs	100
batch	64
optimizer	SGD
$lr_0$	0.01
lrf	0.0001
momentum	0.9
weight_decay	0.0005
warmup_epochs	3

# Modifiche architetturali

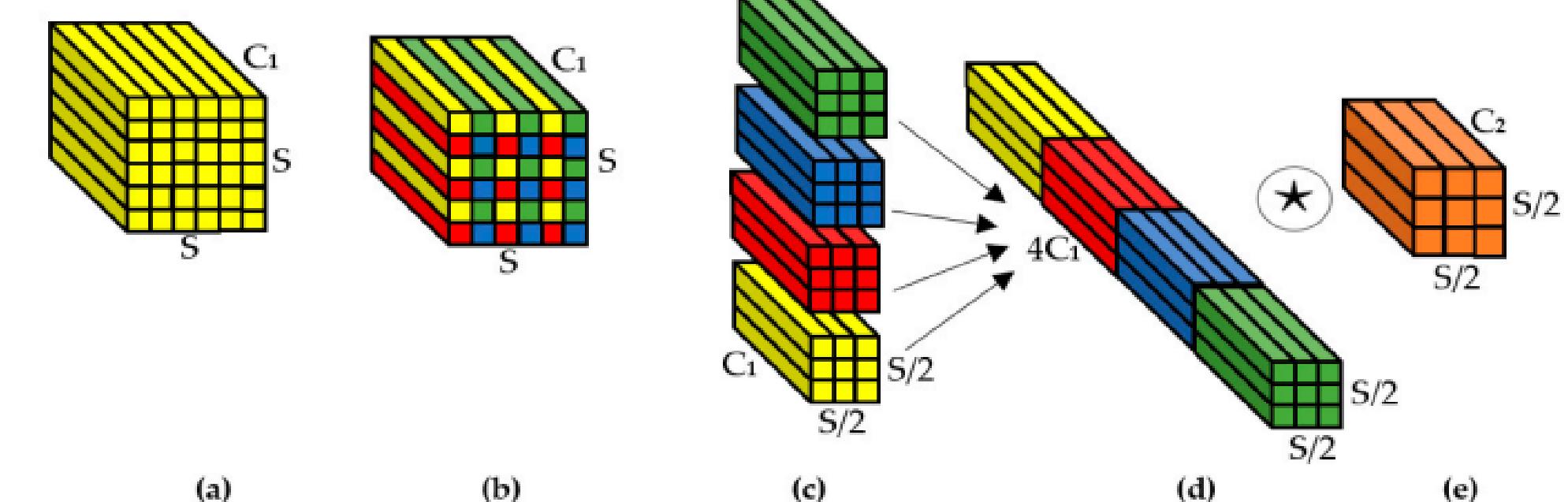
- **SPDConv - Space-to-Depth Convolution**
- **SPANet - Small Path Aggregation Network**
- **Modulo ECSPP**
- **Modulo DLC2f**

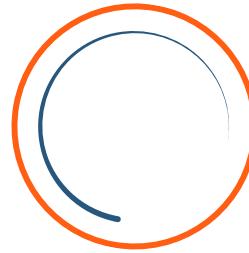


# SPDConv - Space to Depth Convolution

## SPDCONV: POTENZIAMENTO PER PICCOLI OGGETTI

- Sostituisce le convoluzioni di downsampling (eccetto la prima) con una combinazione space-to-depth + convoluzione standard.
- Riduce la risoluzione trasferendo le informazioni sui canali, preservando i dettagli fini.
- Migliora la sensibilità verso i piccoli oggetti, limitando la perdita di informazioni nelle prime fasi del backbone.
- Ideale per scenari come il rilevamento di imbarcazioni in immagini Sentinel-2.

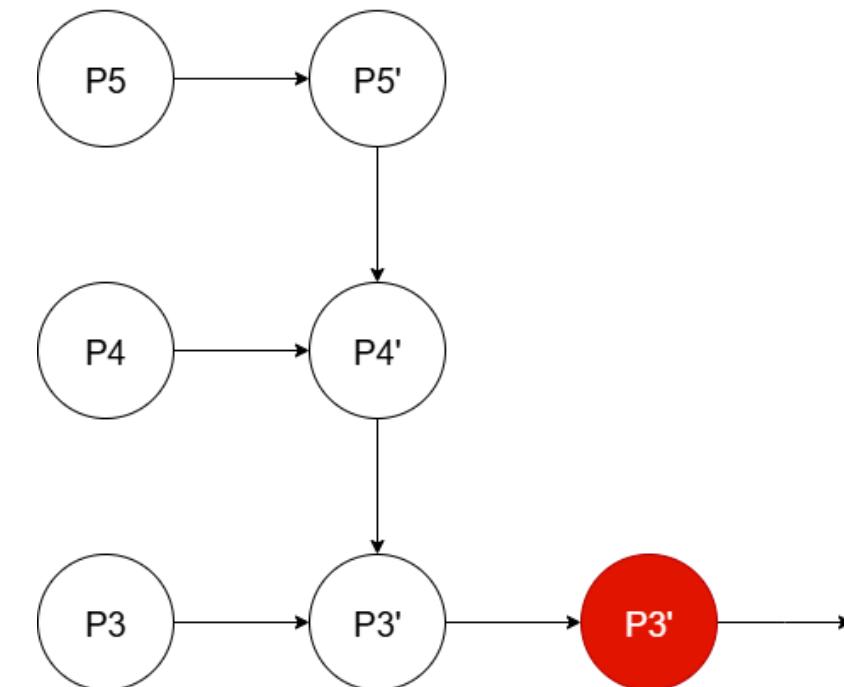




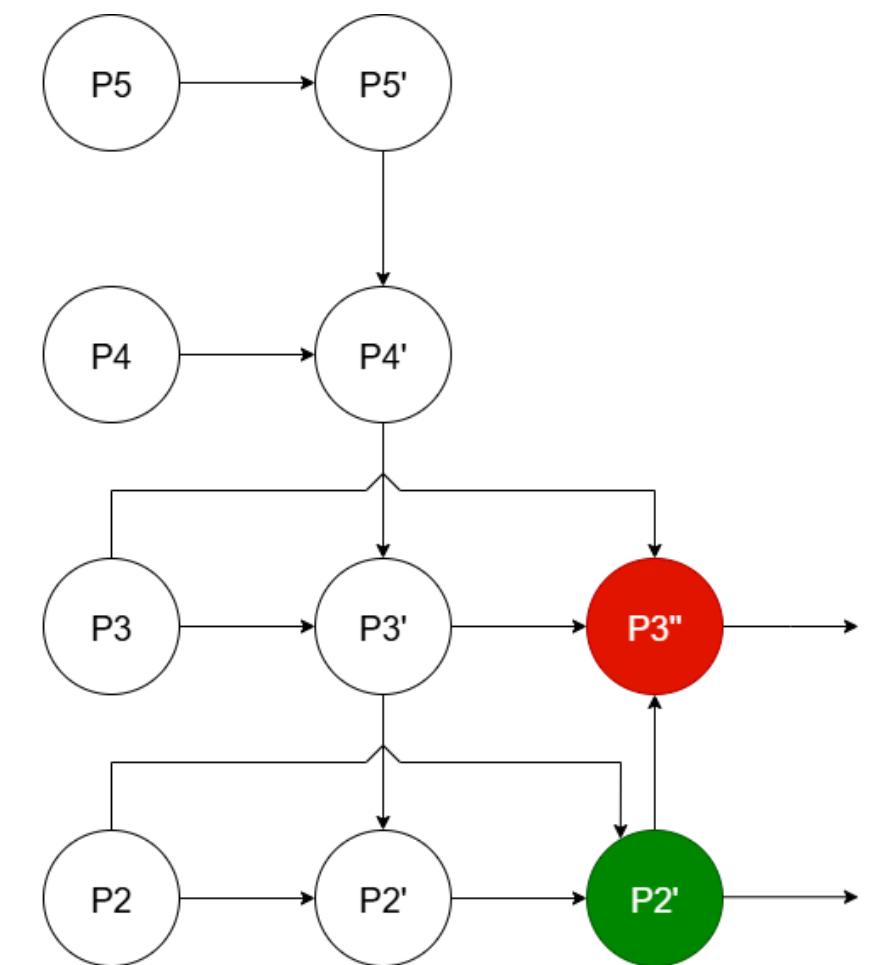
# SPANet - Small Path Aggregation Network

## SPANET: FUSIONE AVANZATA PER PICCOLI OGGETTI

- Estende il classico PANet integrando il livello P2, ricco di dettagli spaziali.
- Combina P2, P3, P4, P5 per una fusione di dettagli fini e informazioni semantiche.
- Introduce SPDConv nella fase bottom-up, migliorando precisione e riducendo i parametri.
- Aumenta la sensibilità e la robustezza nel rilevamento di oggetti piccoli.



a)

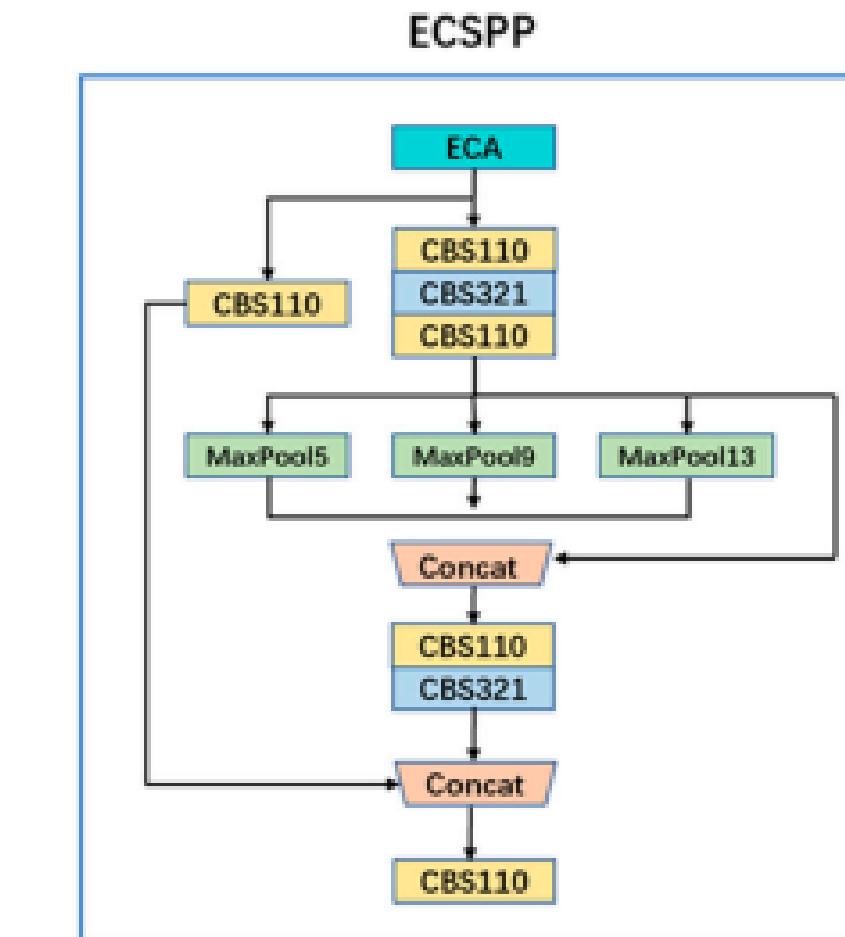


b)

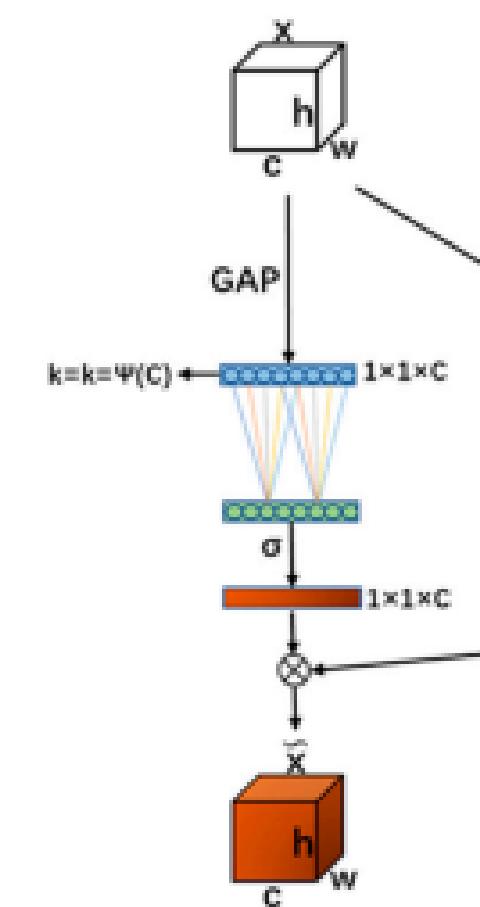
# Modulo - ECSPP

## ECSPP: POTENZIAMENTO MULTI-SCALA DEL BACKBONE

- Combina ECA-Net (attenzione sui canali) e SPPCSPC (estrazione multi-scala).
- Ricalibra le feature map per migliorare la discriminazione e la localizzazione.
- Applica ECA-Net prima della biforcazione tra convoluzioni e SPP, poi unisce i percorsi con ulteriori convoluzioni.
- Testato su YOLOv10 al posto di SPPF, ha migliorato la capacità di rilevamento di oggetti piccoli.



(a)

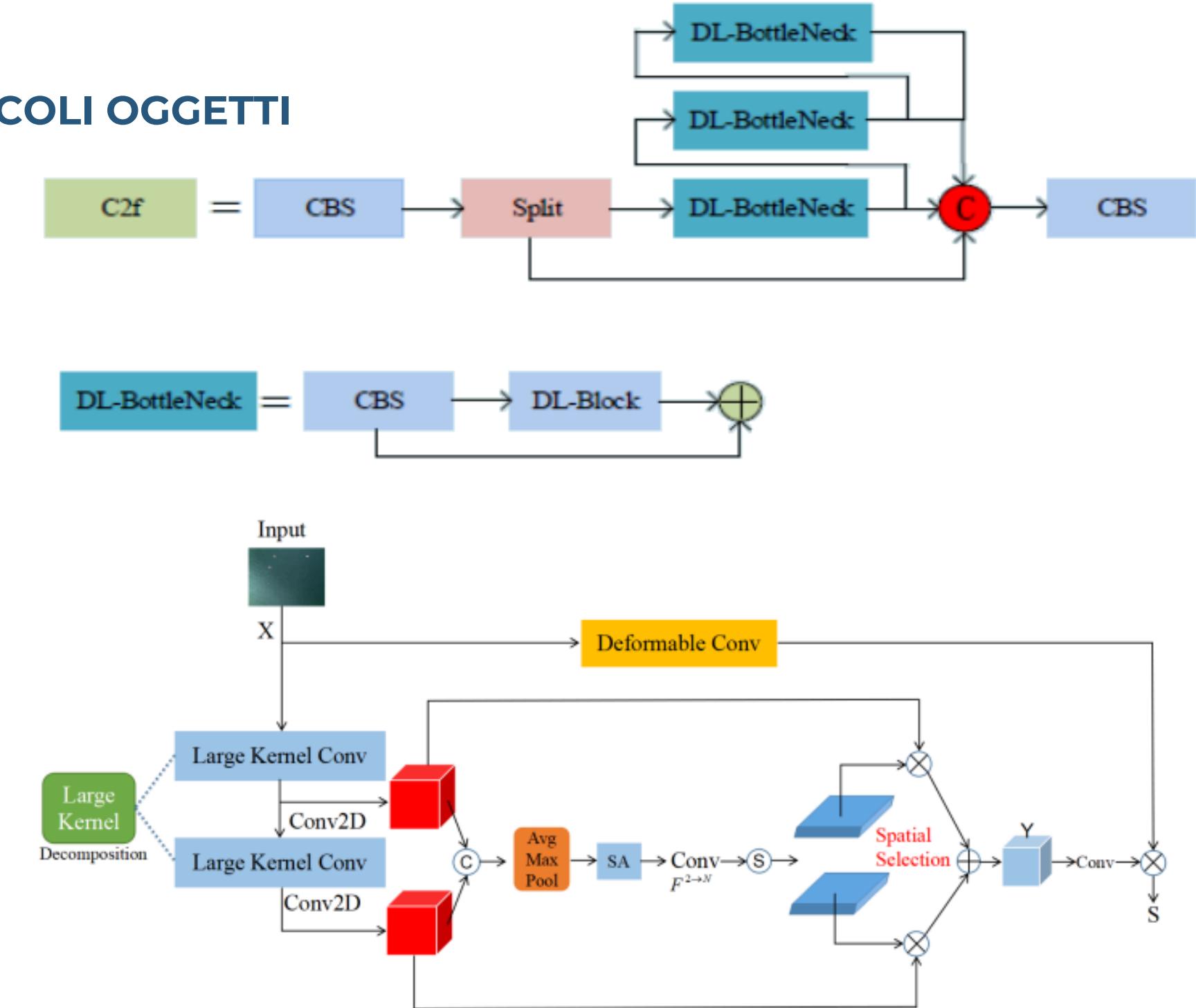


(b)

# Modulo - DLC2f

## DLC2F: ESTRAZIONE AVANZATA DI FEATURE PER PICCOLI OGGETTI

- Evoluzione del blocco C2f con integrazione del DL-Block (Deformable Large Kernel).
- Combina Deformable Convolution (adattiva) e Large Kernel Convolution (ampio campo recettivo).
- Usa attenzione spaziale per raffinare le feature e focalizzarsi sulle regioni rilevanti.
- Integrato sui livelli P2 e P3 del backbone per migliorare la localizzazione di piccoli target.



# Risultati

Modello	Precision	Recall	mAP50	mAP50-95	Parameters	GFLOPs
YOLOv10mLight	0,704	0,578	0,642	0,281	11.437.21	19,4
YOLOv10mLight-SPD-Conv	0,711	0,575	0,642	0,281	22.124.817	77,5
YOLOv10mLight-SPD-SPANet	0,725	0,581	0,652	0,284	22.529.730	93,2
YOLOv10mLight-SPD-SPANet-ECSPP	0,717	0,587	0,650	0,285	30.664.581	99,7
YOLOv10mLight-SPD-SPANet-P2	0,746	0,704	0,728	0,314	22.669.700	104,1
YOLOv10mLight-SPD-SPANet-P2-DLC2f	0,723	0,694	0,721	0,320	23.089.748	105

+5,966% +21,799% +13,396% +11,744%

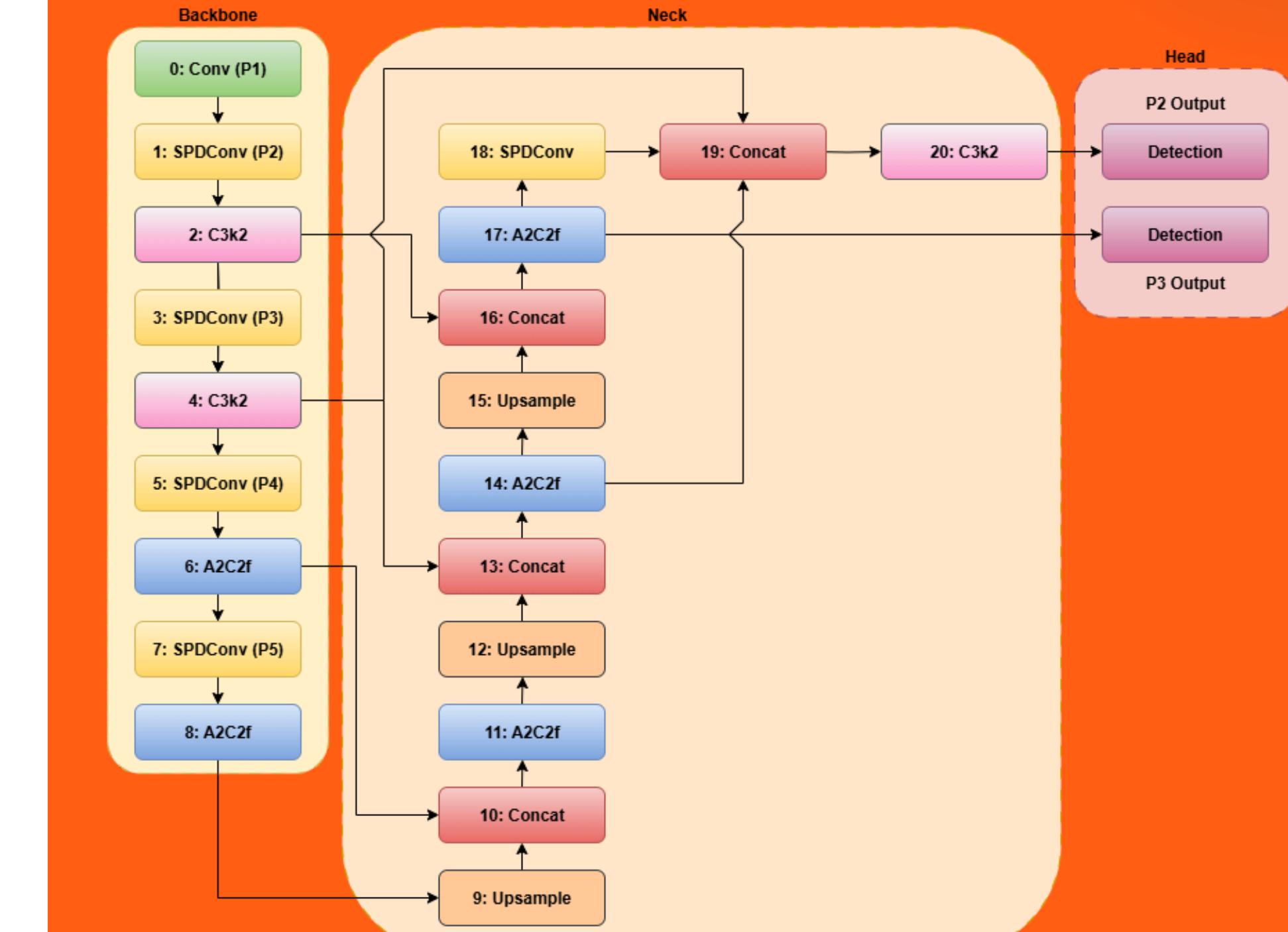
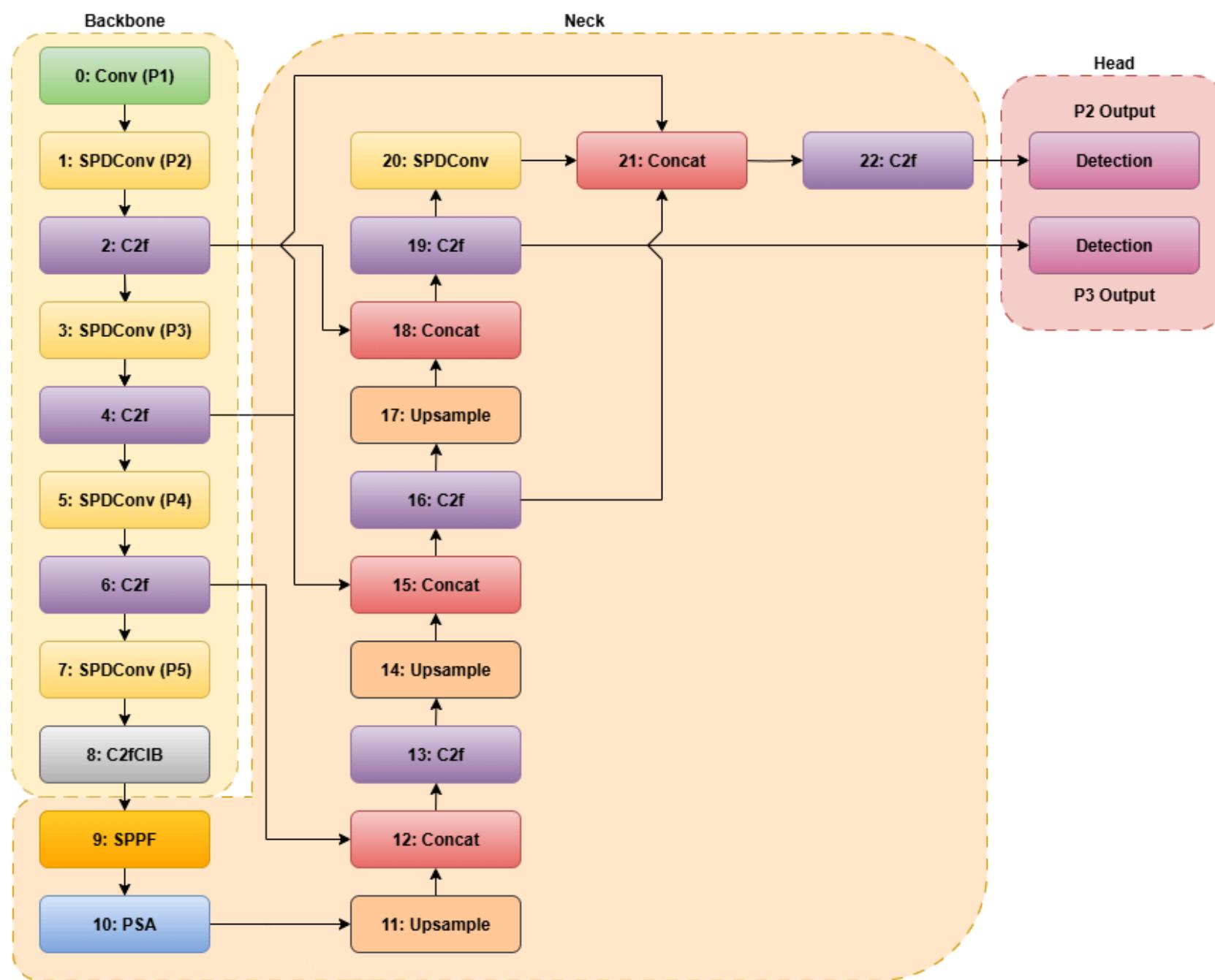
Rispetto a YOLOv10mLight!

Modello	Precision	Recall	mAP50	mAP50-95	Parameters	GFLOPs
YOLOv12mLight	0,732	0,587	0,649	0,281	13.276.817	56,7
YOLOv12mLight-SPD-Conv	0,719	0,6	0,659	0,29	29.424.65	118,9
YOLOv12mLight-SPD-SPANet	0,744	0,606	0,668	0,292	32.145.553	157,5
YOLOv12mLight-SPD-SPANet-P2	0,77	0,736	0,751	0,325	32.212.818	164,3
YOLOv12mLight-SPD-SPANet-P2-DLC2f	0,76	0,72	0,741	0,318	31.000.490	147,5

+5,191% +25,383% +15,658% +11,529%

Rispetto a YOLOv12mLight!

# Architetture finali



YOLOv10mLight  
custom

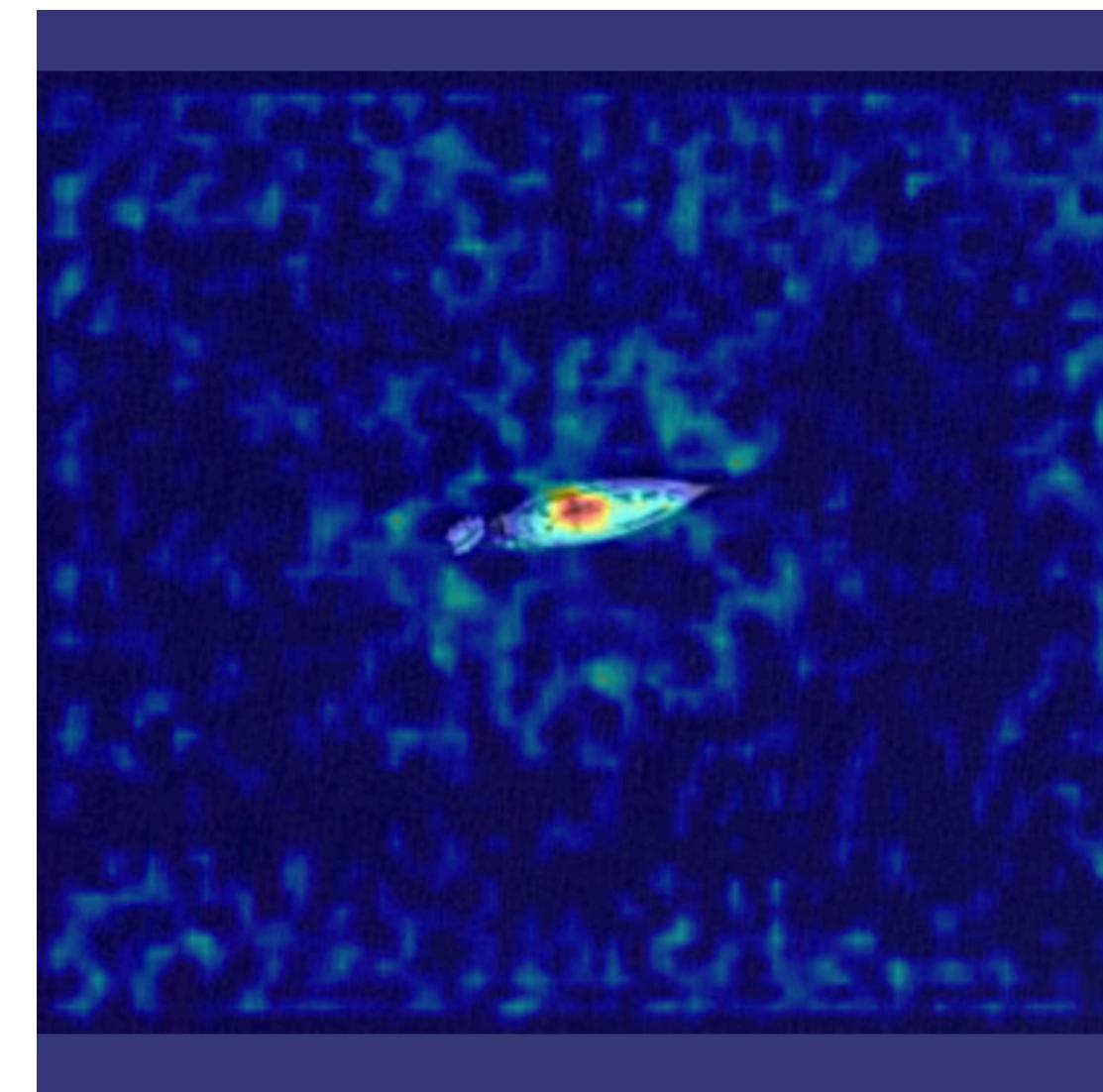
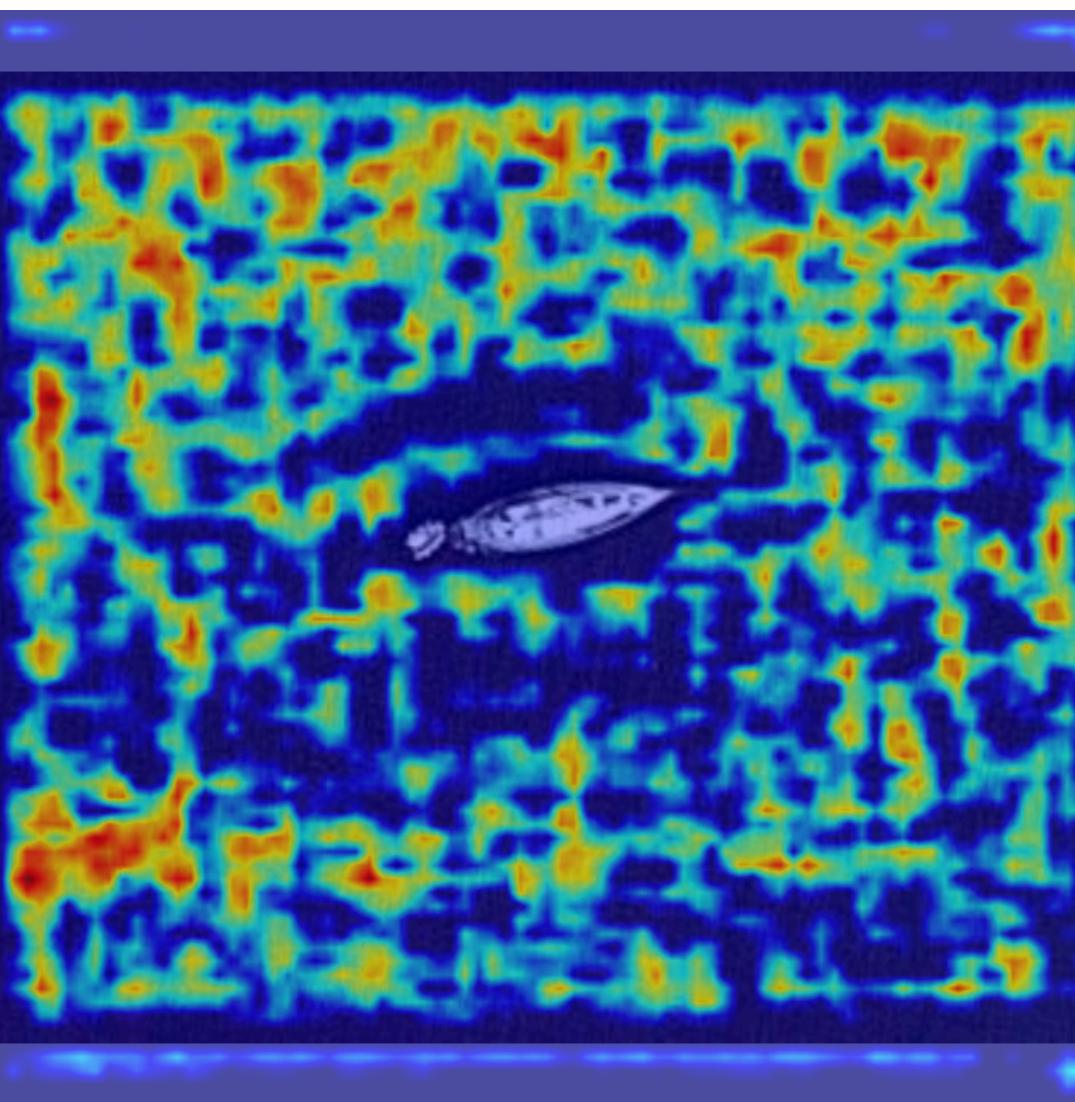
YOLOv12mLight  
custom

# Risultati - Emissioni CO<sub>2</sub>



Modello	CO <sub>2</sub> (kg)
YOLOv10mLight	0.631
YOLOv10mLight-SPD-SPANet-P2	0.729
YOLOv12mLight	0.687
YOLOv12mLight-SPD-SPANet-P2	0.971

# Risultati - Explainability





# Sviluppi futuri

- Espansione del dataset
- Tuning iperparametri
- Branch Super-Resolution
- Studio di loss specializzate su piccoli target
- Estensione al task di classificazione
- Integrazione di moduli di atezione come SCAM

**Grazie per  
l'attenzione**

