

# Supplementary file for the paper “Underwater target detection based on improved YOLOv7”

Kaiyue Liu, Qi Sun, Daming Sun, Lin Peng,<sup>\*</sup>, Mengduo Yang,<sup>\*</sup> and Nizhuan<sup>\*</sup>

**Table S1.** Performance comparison of target detection models based on 5-fold cross-validation of the URPC dataset.

Method	Precision(mean±std)	Recall(mean±std)	mAP@0.5(mean±std)	mAP@0.95(mean±std)
EfficientDet-d0[52]	83.4%±0.04	72.3%±0.05	80.6%±0.05	43.2%±0.02
SSD[21]	77.9%±0.07	73.4%±0.04	76.2%±0.08	37.1%±0.03
RetinaNet-50[53]	73.1%±0.01	66.6%±0.03	73.5%±0.05	34.6%±0.01
Detr[54]	83.9%±0.08	78.8%±0.01	85.4%±0.07	45.6%±0.03
YOLOv5s[27]	87.6%±0.03	79.7%±0.03	84.1%±0.04	48.9%±0.02
YOLOv6n[28]	87.1%±0.02	77.8%±0.03	83.0%±0.08	50.6%±0.02
YOLOv7[29]	86.4%±0.01	81.2%±0.05	86.5%±0.04	51.9%±0.03
YOLOv7-AC	90.3%±0.04	84.0%±0.01	89.8%±0.03	54.0%±0.02

**Table S2.** Ablation comparison of model performance improvement based on 5-fold cross-validation of the URPC dataset.

Model	ResNet-ACmix	AC-E-ELAN	GAM	K-means++	AP(echinus)(mean±std)	AP(starfish)(mean±std)	AP(scallop)(mean±std)	AP(holothuria n)(mean±std)	mAP(mean±std)
YOLOv7	×	×	×	×	78.4%±0.05	89.5%±0.03	90.7%±0.03	87.7%±0.02	86.5%±0.04
	√	×	×	×	87.5%±0.01	90.5%±0.03	87.6%±0.02	84.4%±0.02	87.5%±0.01
	√	√	×	×	91.5%±0.01	89.7%±0.01	91.8%±0.03	83.4%±0.02	89.1%±0.02
	√	√	√	×	91.6%±0.03	91.4%±0.01	90.2%±0.01	84.8%±0.01	89.5%±0.01
	√	√	√	√	92.4%±0.04	91.1%±0.02	92.0%±0.03	83.7%±0.02	89.8%±0.03

**Table S3.** Performance comparison of target detection models based on 5-fold cross-validation of the Brackish dataset.

Method	Precision(mean±std)	Recall(mean±std)	mAP@0.5(mean±std)	mAP@0.95(mean±std)
EfficientDet-d0[52]	94.6%±0.06	87.3%±0.05	93.0%±0.08	62.2%±0.07
SSD[21]	84.3%±0.09	87.4%±0.05	88.2%±0.11	54.8%±0.08
RetinaNet-50[53]	87.1%±0.10	77.9%±0.08	84.7%±0.12	51.3%±0.06
Detr[54]	96.8%±0.04	92.7%±0.06	95.7%±0.07	72.7%±0.05
YOLOv5s[27]	96.1%±0.05	92.8%±0.03	95.4%±0.09	72.1%±0.06
YOLOv6[28]	96.4%±0.06	91.4%±0.12	94.5%±0.14	71.2%±0.08
YOLOv7[29]	95.4%±0.03	93.1%±0.05	95.2%±0.08	74.5%±0.04
YOLOv7-AC	98.8%±0.04	94.5%±0.06	96.4%±0.06	74.2%±0.04

**Table S4.** Ablation comparison of model performance improvement based on 5-fold cross-validation of the Brackish dataset.

Model	ResNet-ACmix	AC-E-ELAN	GAM	K-means++	AP(fish)(mean±std)	AP(small_fish)(mean±std)	AP(crab)(mean±std)	AP(shrimp)(mean±std)	AP(jellyfish)(mean±std)	AP(starfish)(mean±std)	mAP(mean±std)
YOLOv7	×	×	×	×	94.5%±0.04	92.4%±0.04	98.7%±0.02	91.3%±0.04	94.2%±0.03	99.5%±0.03	95.2%±0.08
	√	×	×	×	97.0%±0.03	93.0%±0.04	98.8%±0.02	91.3%±0.03	92.9%±0.02	99.5%±0.01	95.5%±0.05
	√	√	×	×	97.3%±0.02	94.7%±0.03	98.9%±0.04	91.4%±0.01	94.7%±0.03	99.6%±0.03	96.1%±0.04
	√	√	√	×	97.8%±0.02	93.6%±0.04	99.2%±0.03	91.6%±0.03	95.4%±0.03	99.6%±0.02	96.2%±0.05
	√	√	√	√	98.2%±0.03	93.7%±0.04	99.3%±0.02	92.0%±0.0	95.6%±0.03	99.6%±0.02	96.4%±0.06