Supporting Information

Lightweight target detection for large-field ddPCR images based on improved YOLOv5

Xingyu Jin [a], Jing Yang [b], Jijun Feng [a,*], Xiaorui Jiang [c], Zhenqing Li [a,*], Jinrong Shen [d], Zhiheng Yu [a], Cunliang Yang [a], Fengli Huang [e], Dunlu Peng [a] and Yoshinori Yamaguchi [f]

[a] Shanghai Key Laboratory of Modern Optical System, Engineering Research Center of Optical Instrument and System (Ministry of Education), School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

[b] Faculty of Engineering, Anhui Sanlian University, Hefei, 230000, China.

[c] National Key Laboratory of Electromagnetic Space Security, Tianjin 300308, China.

[d] State Key Laboratory of ASIC and System, Fudan University, Shanghai 200433, China

[e] The Key Laboratory of Medical Electronics and Digital Health of Zhejiang Province, Jiaxing University, Jiaxing 314001, China.

[f] Comprehensive Research Organization, Waseda University, Tokyo 162-0041, Japan.

Corresponding author: Jijun Feng, fjijun@usst.edu.cn, Zhenqing Li, zhenqingli@163.com

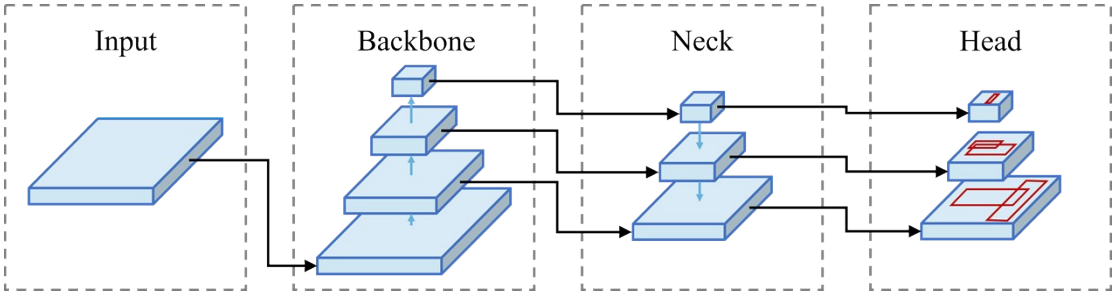

Fig. S1. YOLOv5 network detection workflow. The backbone network captures image features, which are subsequently refined and integrated by the neck network. The head network outputs the final target predictions, detailing both categories and locations.

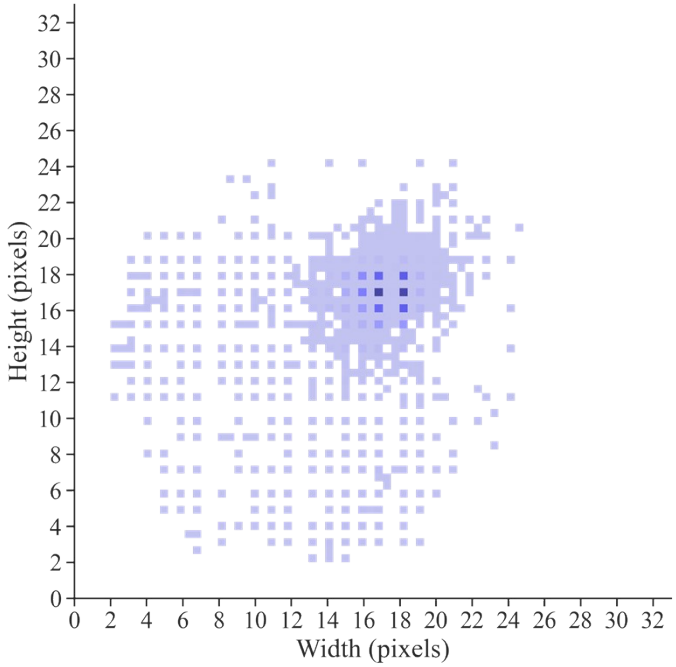

Fig. S2. Size distribution of annotated boxes for targets. The dimensions of all microchambers are confined to a 32×32 pixel, covering the detection ranges of P2, P3, and P4. The majority are concentrated within the region surrounding 14×14 to 20×20 pixels. Since the cropped droplets in the sub-images were considered during annotation, the distribution of annotated detection boxes is wider in size.

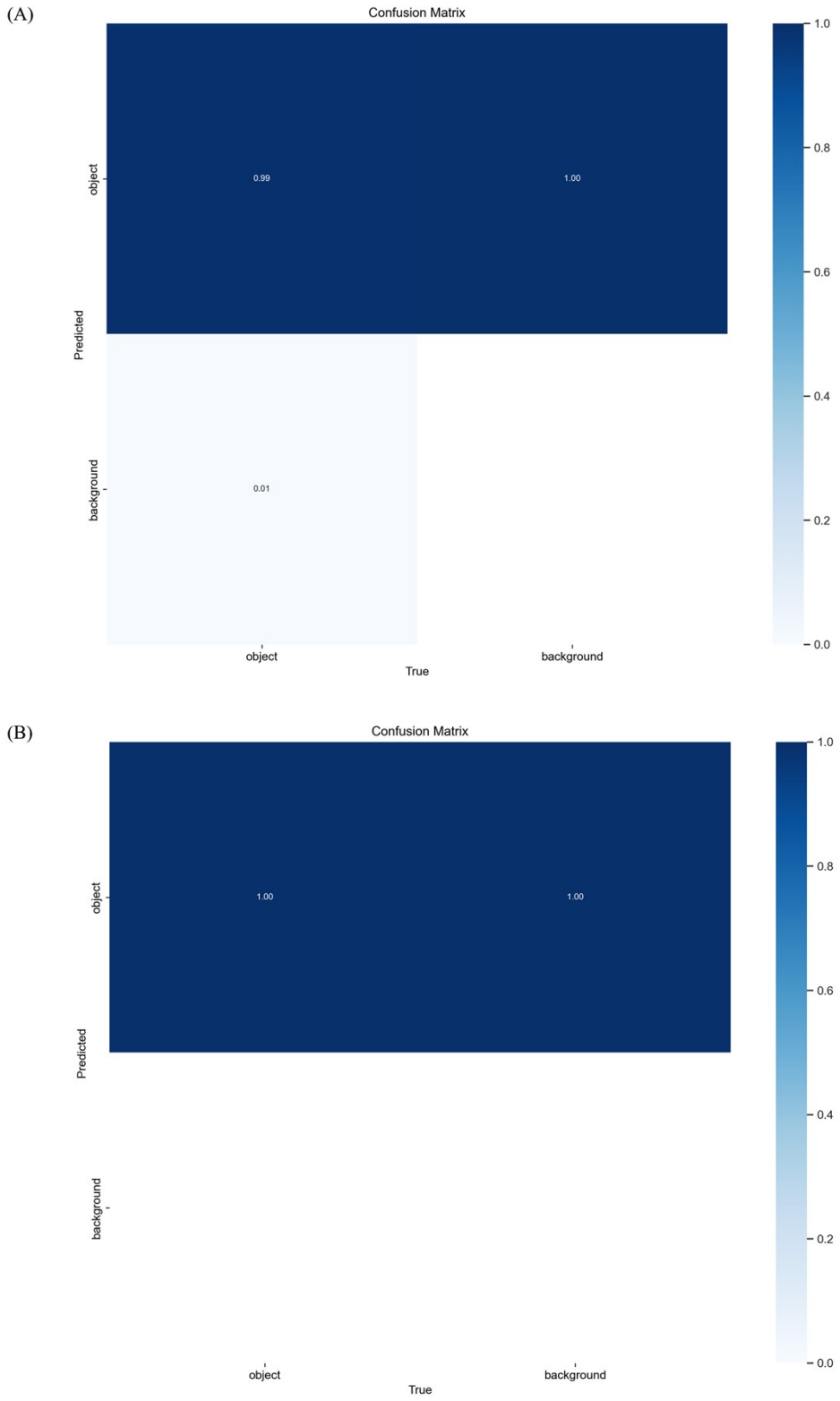

Fig. S3. Confusion Matrices. (A) Results of the original YOLOv5s model. (B) Training results of the final optimized model.

Table S1. Hardware configuration and operating environment.

Device	Details
CPU	Intel Core i7-10750H
GPU	NVIDIA GeForce RTX 3060
Memory	16 G
Graphics Memory	6 G
SSD	512 G
Operating System	Windows 11
Python	3.9.7
PyTorch	1.13.0
CUDA	11.7
CUDNN	8.8.1

Table S2. Comparison of different detection scales.

P2	P3	P4	P5	P	R	mAP(0.5:0.95)	Parameters	Size(MB)	FLOPs(G)
	√	√	√	99.5%	99.3%	77.9%	7012822	14.4	15.8
	√	√		99.5%	99.3%	77.9%	1558340	3.4	10.1
√	√	√		99.2%	99.5%	78.3%	1701526	4.3	12.9

Table S3. Comparison of different feature fusion methods.

Method	P	R	mAP(0.5:0.95)	Parameters	Size(MB)	FLOPs(G)
PANet	99.2%	99.5%	78.3%	1701526	4.3	12.9
FPN	99.5%	99.4%	77.8%	1442902	3.8	11.6
BiFPN	99.3%	99.4%	78.5%	1717910	4.4	13.1

Table S4. Comparison of different module improvement experiments. "B" and "N" represent the backbone and neck networks, respectively, while "S" indicates the insertion of attention mechanism.

B	N	S	P	R	mAP(0.5:0.95)	Parameters	Size(MB)	FLOPs(G)
√			99.5%	99.3%	77.9%	1292030	3.5	9.8
	√		99.2%	99.6%	78.5%	854492	3.7	10.6
√	√		99.5%	99.3%	78.2%	946246	2.9	7.3
		√	99.4%	99.7%	78.7%	1718358	4.4	13.1
√		√	99.4%	99.4%	78%	1292030	3.6	9.8
	√	√	99.5%	99.5%	78.7%	1372574	3.7	10.6
√	√	√	99.3%	99.4%	78.6%	946246	2.9	7.3

Table S5. Results of ablation experiment. Scheme A: Original model without modifications. Scheme B: Improved with scales P2, P3, P4. Scheme C: Improved with BiFPN. Scheme D: Improved with GhostConv, C3Ghost, and SimAM. Scheme E: Improved with network pruning.

Scheme	P	R	mAP(0.5:0.95)	Parameters	Size(MB)	FLOPs(G)
A	99.5%	99.3%	77.9%	7012822	14.4	15.8
B	99.2%	99.5%	78.3%	1701526	4.3	12.9
C	99.3%	99.4%	78.5%	1717910	4.4	13.1
D	99.3%	99.4%	78.6%	946246	2.9	7.3
E	99.5%	99.5%	78.1%	262047	1.5	2.6

Table S6. Comparison of mainstream single-stage object detection algorithms.

Model	P	R	mAP(0.5:0.95)	Parameters	Size(MB)	FLOPs(G)
SSD	80.8%	44.6%	41.9%	26649772	90.6	35
YOLOv3-tiny	99.7%	98.9%	74.1%	8666692	17.4	12.9
YOLOv5s-ShuffleNetv2	99.5%	99.1%	76.5%	842358	2.0	1.8
YOLOv5s-MobileNetv3	99.4%	98.4%	76.3%	1374732	3.1	2.3
YOLOv5n	99.4%	99.3%	77.3%	1760518	3.9	4.1
YOLOv5s	99.5%	99.3%	77.9%	7012822	14.4	15.8
Ours	99.5%	99.5%	78.1%	262047	1.5	2.6