DOI:10.19652/j.cnki.femt.2305725

轻量化 YOLOv7-tiny 的水下压印字符识别*

李卓润 李 波 邱鹏程 刘 洪
(中国地质大学(武汉)机械与电子信息学院 武汉 430074)

摘 要:自动化水下字符识别技术能通过编号更高效地定位追踪水下设备,是管理和维护水下设备的关键。针对该任务目标区别较小和水下场景中干扰等问题,并考虑其检测速度需求,基于 YOLOv7-tiny 模型,提出一种轻量化的改进模型。首先采用 MobileNetV3 作为新的特征提取网络对整体框架进行轻量化处理;然后引入 PConv 至 ELAN 模块中,减少 Neck 层的计算量;最后将置换注意力机制应用至 Head 层,提升了模型对字符定位的表达能力。实验结果表明,改进后的模型相较于原模型的平均精度均值(mAP)提高了 2.4%,参数量和计算量分别减少 30.0%和 38.5%,检测速度提升 30.8%。改进后的模型在水下字符识别任务中具有更高的效率和精度,为推进并实现水下自动化识别编号设备的部署提供了可行性。

关键词:水下字符识别;YOLOv7-tiny;轻量化;PConv;置换注意力
中图分类号: TP391 **文献标识码**:A **国家标准学科分类代码**: 520.2

Lightweight YOLOv7-tiny underwater embossed character recognition

Li Zhuorun Li Bo Qiu Pengcheng Liu Hong
(School of Mechanical Engineering and Electronic Information, China University of Geosciences, Wuhan 430074, China)

Abstract: Automated underwater character recognition technology can more efficiently locate and track underwater equipment through numbers, which is the key to managing and maintaining underwater equipment. In view of the problems such as the small target difference of the task and interference in the underwater scene, and considering its detection speed requirements, this article proposes a lightweight improved model based on the YOLOv7-tiny model. First, MobileNetV3 is used as a new feature extraction network to lightweight the overall framework. Then PConv is introduced into the ELAN module to reduce the calculation amount of the Neck layer. Finally, the displacement attention mechanism is applied to the Head layer to improve the model′s ability to position characters. expression ability. Experimental results show that compared with the original model, the mAP of the improved model is increased by 2.4%, the amount of parameters and calculations are reduced by 30.0% and 38.5% respectively, and the detection speed is increased by 30.8%. The improved model has higher efficiency and accuracy in underwater character recognition tasks, providing feasibility for promoting and realizing the deployment of underwater automated identification and numbering equipment.

Keywords: underwater character recognition; YOLOv7-tiny; lightweight; PConv; shuffle attention

0 引 言

随着当今对水下资源的不断开发和探索,我国水下设备的种类也逐渐增多,但对于水下工程设备的编号识别、溯源和统计较为困难[1]。考虑到对水下设备的后期管理需求,水下设备编号的在线检测至关重要。设备编号的在线检测本质上是对编号的每个字符进行实时识别。然而由于水下检测环境的特殊性,识别过程中不仅会面临光线较暗、光照不均匀的光学问题[2],还会碰到字符被水下泥沙遮挡和水流波动的环境干扰问题。设备上的压印字符自身还存在与背景对比度不明显的客观问题。且考虑到当前实际应用中水下字符识别技术的自动化程度较低,主要依赖人

收稿日期:2023-11-01
* **基金项目**:国家重点研发计划(2023YFC2813104)、国家重点研发计划(2023YFC3007004)项目资助

眼来进行相关信息的识别，效率低成本高。故面对水下环境下出现的多种类型的字符，快速且精确地识别设备编号是水下设备管理过程中必须攻克的一项难关。因此探索高效水下压印字符识别技术具有一定的实际工程应用。

针对字符识别的工作，研究人员做了大量的研究，汪志成等[3]提出一种基于 LeNet-5 所改进的压印字符识别算法，解决了传统压印识别领域精度低速度慢的问题。黄慧宁[4]对针对钢材压印字符提出一个字符倾斜的预处理方法并利用改进 YOLOv2 算法对矫正后的字符进行准确识别。宫鹏涵[5]就枪械压印字符与其背景色一致的问题，使用 YOLOv5 算法实现了对钢印字符的高效且稳定的识别。Zhang 等[6]就复杂场景中不同分布的铸件浮雕凹凸字符的识别进行了研究，提出了一种基于 YOLOv5[7]的铸造浮雕字符识别方法，该方法准确率高、处理速度快、权值文件较小，但依然达不到可以移动部署的水平。综上，常规环境下的压印字符识别任务的研究较多且效果优良，但在水下环境中的研究较少且未考虑其运用的部署性能。与此同时考虑到水下环境移动识别会面临的各种困难，也给水下压印字符的辨识带来了挑战。因此，研发高效准确的轻便水下字符识别技术具有重要意义。

水下环境的字符识别任务可以被视为对预先单个字符的目标检测与鉴别任务。Wang 等[8]提出的 YOLOv7 作为目标检测任务中的佼佼者，其能在相同体量下比 YOLOv5 精度更高且速度快 120%，在交通监控、工业检测等多个领域中表现突出。本文根据水下部署的应用需求，选择 YOLOv7-tiny 小模型作为研究对象并应用于该识别任务中。借助 YOLOv7-tiny 的卓越检测速率，能够基本完成水下字符的实时识别。为了提高模型的检测性能，本文针对水下环境采集的图像特性以及检测速率的要求，在 YOLOv7-tiny 的基础上提出一种改进模型。为了解决水下识别面临的各种问题，应在数据集中增加问题场景的数据量；为了解决移动端部署的需求，模型应尽量轻量化以减少内存消耗。实验证明改进后的算法相比原始模型有更小的体量、更快的检测速度和稳定的精度，做到了轻量化和精度的平衡。同时在面临干扰情况时也能较好完成识别任务，并满足移动设备部署要求。

1 基准模型

YOLOv7-tiny 是对 YOLOv7 进行模型精简后得来的轻量级架构，具有更好的边缘部署性能。该网络模型整体由输入端(input)、特征提取网络(backbone)、特征融合网络(neck)及输出端(head)4 个部分构成，整体网络结构如图 1 所示。

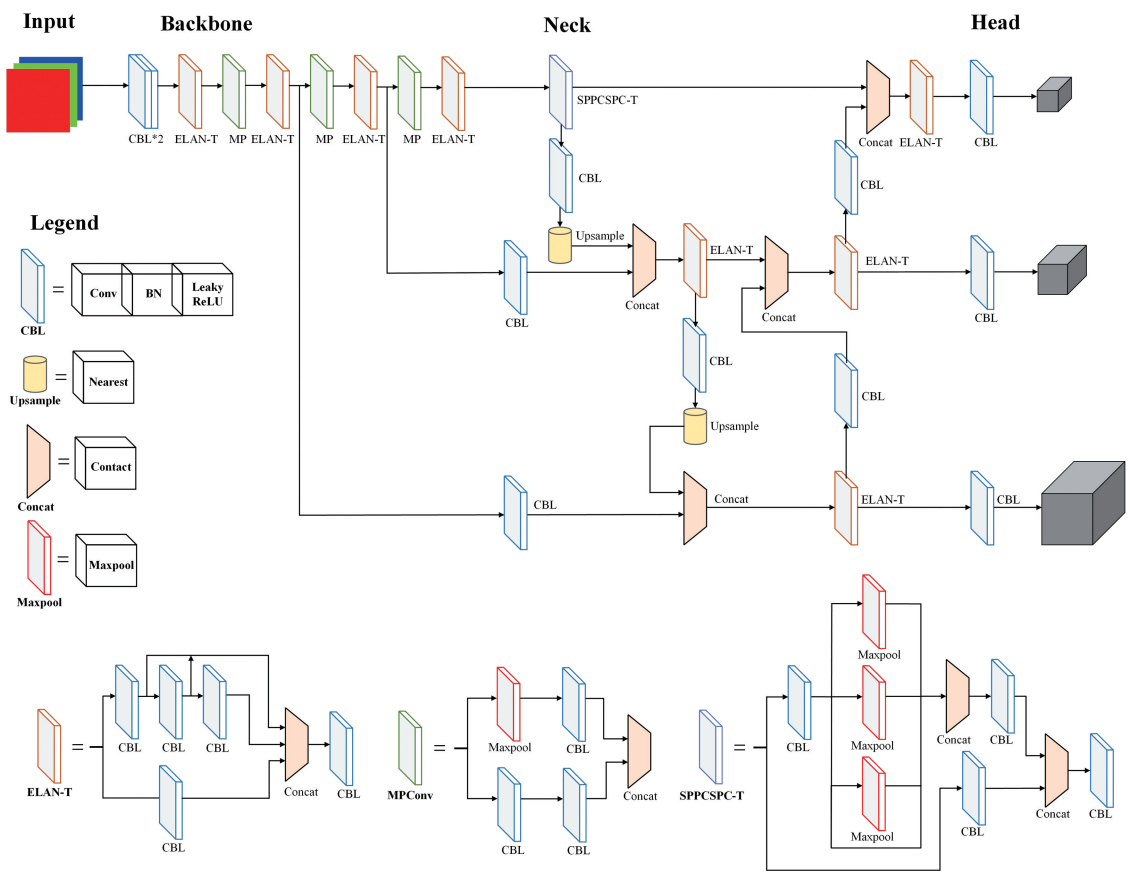

图 1 YOLOv7-tiny 整体网络结构

Fig. 1 Overall network structure of YOLOv7-tiny

与 YOLOv7 相比，tiny 在特征提取网络中，将高效层聚合网络(efficient layer aggregation network，ELAN)消减计算块简化为 ELAN-T；在特征融合网络中，保留 SPPCSPC 并与 PANet 进行多尺度特征聚合；在输出端采用标准卷积代替 REPConv 调整通道数。该基准模型在保证检测精度的同时考虑了参数小速度快的优势，适合应用于水下字符的实时检测需求，但在该任务下模型结构仍存在如下不足之处。

1)在模型的特征提取网络中大量使用了 ELAN 类模块，导致主干网络中的参数数量较多、计算量较大，但考虑到字符识别场景交叉干扰较少，特征提取难度小于多目标重合的复杂场景，故在检测速度上仍有优化空间。

2)在特征融合层同样多次使用 ELAN 类模块造成特征的冗余，考虑到在检测设备上部署的轻便性，消除冗余数据可以减小模型体积便于移动使用。

3)在水下弱光环境里，探照拍摄会造成字符表面反光或光照分布不均匀等情况。水下字符表面附着泥沙遮挡也同时也会对检测效果造成影响。该框架由于其轻量化的设计会造成上述环境下检测精度较差，无法准确识别到相关字符。

本文将以 YOLOv7-tiny 网络架构为基础，探讨如何采用更轻量化的方案来聚合识别任务所需要的特征。在确保特征足够丰富的前提下，将尽可能减少参数量和计算量，以优化整个架构对在水下受环境影响字符的识别能力。最终做出相关改进来优化模型对水下字符识别的效率。

2 本文方法

为增加水下环境下的字符识别任务的效率，在保证精度的同时对模型进行简化，本文对 YOLOv7-tiny 的网络结构进行改进，改进后的网络结构如图 2 所示。1)借助一种轻量化网络 MobileNetV3 替换主干网络来瘦身以提高检测速度；2)将 PConv 融入 ELAN 生成新的 ELAN-P 结构，减少冗余计算和内存访问且保证精度；3)添加 SA 注意力机制提升模型的特征表达能力。

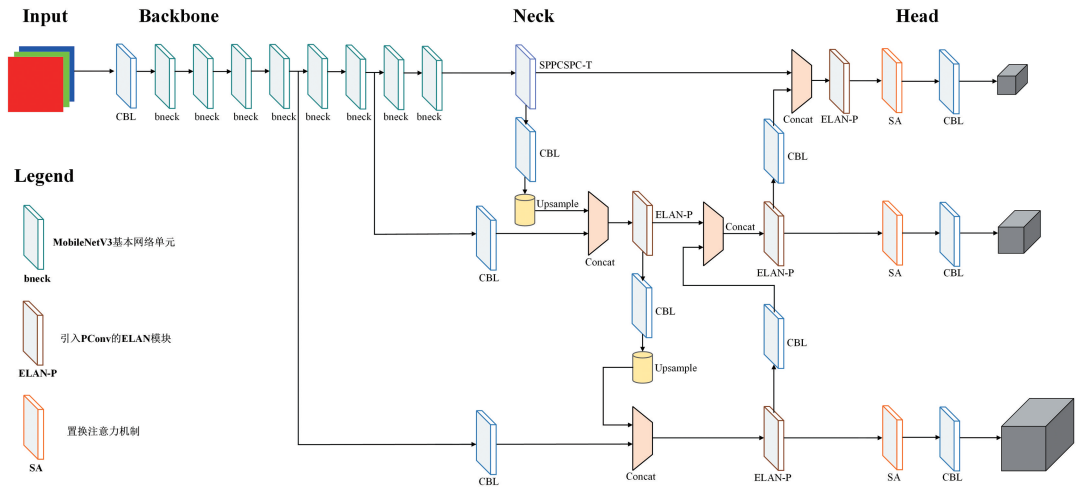

图 2 改进后网络结构

Fig. 2 Improved network structure

2.1 MobileNetV3 网络

MobileNetV3[9]是谷歌的第 3 代轻量级网络架构，相比于其他网络架构，其能在小幅降低精度的代价上大幅减少模型的复杂程度。在其内容上，延续了前两代提出的深度可分离卷积和反向残差结构外，重点加入了压缩激励模块的轻量级注意力机制并更换激活函数为计算量更少的 h-swish(x)。本文将在骨干网络中引入 MobileNetV3 的基本单元 bneck 块来代替网络中的原模块。

MobileNetV3 的基本网络单元 bneck 如图 3 所示，该结构首先让输入特征经过一个 1×1 卷积进行升维操作，增加通道数；扩张后特征层再通过 3×3 深度可分离卷积块分离出各层主要特征；后对特征层施加轻量级注意力机制，其主要任务是对特征层进行平均池化，再经过 ReLU 和 hard-swish 激活函数调整特征图每个通道的权重，后与前特征层相结合完成注意力模块的添加；最后通过逐点卷积进行降维处理掉网络训练时的无效参数将通道数压缩到与输入相同，利用残差边将处理后的特征与输入特征相加。

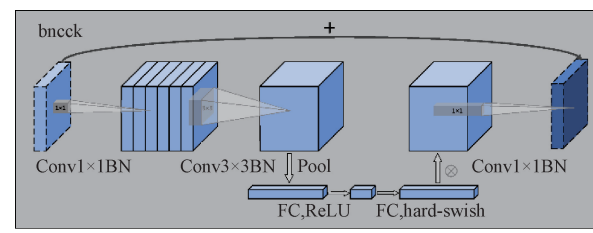

图 3 MobileNetV3 基本网络单元

Fig. 3 Basic network unit of MobileNetV3

由于 MobileNetV3 中所使用的深度可分离卷积相比于普通卷积来说分离了卷积过程的通道和空间的相关性，对每个单通道逐点进行卷积操作，在特征图通道与空间信息完整的情况下能显著降低参数和计算量。假设输入尺寸大小为 $D_i \times D_i \times C_i$（其中 D_i 为特征图的边长，C_i 为输入通道数），卷积核大小为 $F \times F \times C_i$（其中 F 为卷积核尺寸），输出特征图为 $D_i \times D_i \times C_o$（$C_o$ 为输出通道数）。则深度卷积的参数量为 $Q_d = F \times F \times C_i + C_i \times C_o$，计算量为 $S_d = D_i \times D_i \times C_i \times F \times F + D_i \times D_i \times C_i \times C_o$；相同大小的标准卷积参数量为 $Q_c = F \times F \times C_i \times C_o$，计算量为 $S_c = D_i \times D_i \times F \times F \times C_i \times C_o$。故两种卷积参数量之比为：

$$r_q = \frac{Q_d}{Q_c} = \frac{F \times F \times C_i + C_i \times C_o}{F \times F \times C_i \times C_o} = \frac{1}{C_o} + \frac{1}{F \times F} \tag{1}$$

计算量之比为：

$$r_s = \frac{S_d}{S_c} = \frac{1}{C_o} + \frac{1}{F \times F} \tag{2}$$

由比例公式可以验证出深度可分离卷积在参数量和计算量上的优化效果[10]，使用该网络块的 MobileNetV3 可以在完成模型轻量化工作，提升了运行速度[11]。因此，本文选择 MobileNetV3 网络替换原 YOLOv7-tiny 的骨干网络，用于进行特征提取。

2.2 ELAN-P 结构

部分卷积[12]（partial convolution，PConv）是一种更高效的卷积神经网络结构，可以提高网络的性能和效率。其利用了特征映射的冗余性，将常规卷积应用于一些输入通道，而其余部分保持不变。PConv 具有比常规卷积更低的浮点运算，减少了冗余计算和内存访问的数量，并且可以更有效地提取空间特征。普通卷积和部分卷积的原理对比如图 4 所示。PConv 特殊在只对部分输入通道使用普通卷积的方式进行空间特征提取，而其余部分不受影响。其中输入和输出特征映射具有相同的通道，c 表示输入特征映射的通道数，c_p 表示 PConv 中用于空间特征提取的通道数，瘙釆表示卷积。

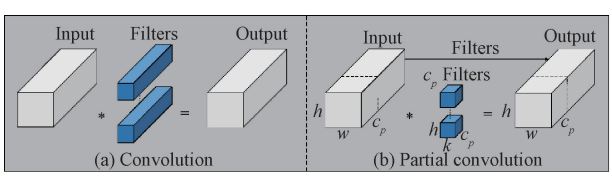

图 4 两种卷积原理对比

Fig. 4 Comparison of two convolution principles

空间特征提取的通道数 c_p 根据输入特征图通道之间的冗余程度决定，如果不同通道之间的冗余较大，则使用部分通道进行计算可以减少计算量和内存访问，但较大的 c_p 值可能会降低精度。根据性能要求，因此需在精度和计算效率之间需要找到平衡。根据实验和验证结果，可以选择最佳的 c_p 值，以保持合理的精度水平同时减少计算冗余和内存访问。在本文中 c_p 被设置为 $\frac{1}{4}c$。在性能方面，PConv 的浮点运算次数仅为 $h \times w \times k^2 \times {c_p}^2$，其中 h、w 分别为卷积核的高度和宽度，由于 $c_p = \frac{1}{4}c$，所以 PConv 的 FLOPs 仅为普通卷积的 1/16。此外，其内存访问量为 $h \times w \times 2c_p + k^2 \times {c_p}^2 \approx h \times w \times 2c_p$，约为普通卷积的 1/4。因为 PConv 只需读取和处理部分通道的数据，而传统的卷积需要读取和处理所有通道的数据。故 PConv 在减少浮点运算次数和内存访问量有很好的效果。

在 YOLOv7-tiny 的 ELAN-T 模块中，将其中的某些卷积替换为 PConv 以生成 ELAN-P 结构，可以在保持良好的特征提取能力的同时，有效地减少冗余计算和内存访问[13]。由于 PConv 的输出特征图与输入特征图具有相同的通道数，使 PConv 能够与 ELAN 模块较好融合。融合后的 ELAN-P 模块结构如图 5 所示。

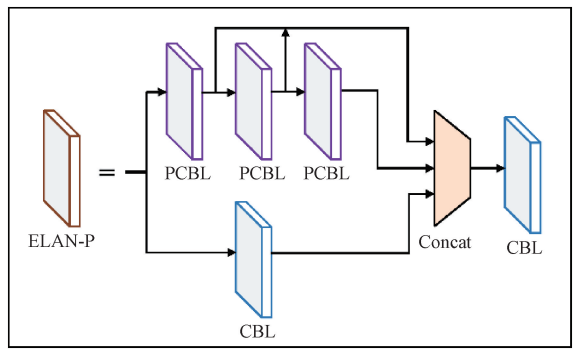

图 5 ELAN-P 结构

Fig. 5 Structure of ELAN-P

2.3 置换注意力机制(Shuffle Attention，SA)

SA 是一种融合通道注意力和空间注意力的轻量化模块，其能关注到输入特征图中的重要信息来提高模型的性能[14]。与其他注意力模块相比，SA 引入了通道洗牌来打乱原特征图的通道顺序，进一步增强了注意特征的多样性。利用洗牌和注意力结合的方式，让 SA 模块在各种复杂场景下具有更好的泛化能力。

SA 模块主要分为分组特征、融合注意力和聚合特征 3 个部分。输入的特征映射 $\boldsymbol{X} \in \boldsymbol{R}^{C \times H \times W}$ 沿通道维度被分为 G 组子特征，表示为 $\boldsymbol{X} = [X_1, \cdots, X_G]$，$X_k \in \boldsymbol{R}^{C/G \times H \times W}$，其中每组子特征层 X_k 将捕捉训练后的准确语义反馈，从而使特征提取更加精准详细。后将 X_k 沿通道维度分为两个子特征图 $X_{k1}, X_{k2} \in \boldsymbol{R}^{C/2G \times H \times W}$，通道注意力分支采用通道之间的关系生成通道注意力图，并使用全局平均池化生成 s 来嵌入全局信息（$\boldsymbol{F}_{gp}$），通过在空间维度 $H \times W$ 上来收缩 X_{k1}，s 的计算公式为：

$$s = \boldsymbol{F}_{gp}(X_{k1}) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} X_{k1}(i, j) \tag{3}$$

然后通过一个门控和 Sigmoid 激活函数来创建一个紧凑的特征，输出的通道注意力 X'_{k1} 的公式为：

$$X'_{k1}=\sigma(\boldsymbol{F}_c(s))\cdot X_{k1}=\sigma(W_1 s+b_1)\cdot X_{k1} \tag{4}$$

空间注意力分支捕捉特征之间的空间依赖关系以生成空间注意力图，其对输入特征图执行分组规范化计算，然后使用 $\boldsymbol{F}_c(\cdot)$ 来增强输入。空间注意力的输出 X'_{k2} 的公式为：

$$X'_{k2}=\sigma(W_2\cdot GN(X_{k2})+b_2)\cdot X_{k2} \tag{5}$$

整体输出由两个分支相连，$X'_k=[X'_{k1},X'_{k2}]\in \boldsymbol{R}^{C/G\times H\times W}$。采用通道洗牌的方式实现了跨组信息交换，且输出端和输入端特征图大小一致。SA 模块结构如图 6 所示。本文将 SA 注意力模块添加至 Neck 层与 Head 层的连接处，提高预测能力以获得更佳的网络性能。

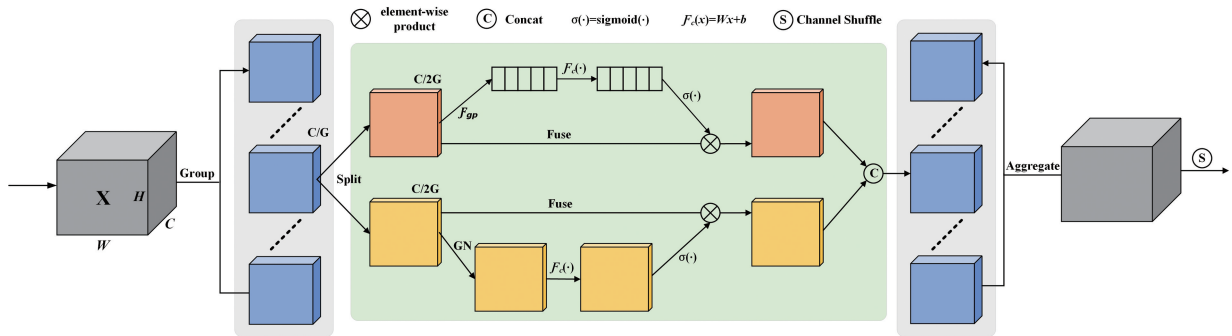

图 6 SA 模块结构

Fig. 6 Structure of SA model

3 实验与分析

3.1 实验数据集

本文数据集来源于相机对水下放置的不同字符平扫拍摄的视频，拍摄设备为工业相机和环形光源，将拍摄的视频进行固定帧数处理得到共 2 340 张图像，其分辨率为 640×640。使用标注工具 LabelImg 对所有数据集进行标注，数据集格式选择为 VOC，保持文件名与图片名一致。数据集总包含了 35 个数据类别，分别是大写字母 A、B、C、D、E、F、G、H、I、J、K、L、M、N、P、Q、R、S、T、U、V、W、X、Y、Z 和阿拉伯数字 0、1、2、3、4、5、6、7、8、9。同时，该数据集涵盖了在水下检测时面临的多种特殊情况，如拍摄造成的过度曝光、水流波动模糊以及字符处有泥沙异物遮挡等问题。本文按照 7∶2∶1 的比例将所有图像等比划分为训练集、验证集和测试集。其中训练集 1 638 张，验证集 468 张，测试集 234 张。为了保证模型拥有更好的训练效果，在训练时对数据集进行 50%概率的 Mosaic 数据扩充。

3.2 实验环境

实验环境为 64 位 Windows11 操作系统，Python3.8.17 版本以及 PyTorch1.9.0 深度学习框架，相关详细硬件信息及训练模型的参数如表 1 所示。

表 1 实验训练环境

Table 1 Experimental training environment

名称	配置
GPU	NVIDIA GeForce RTX4070Ti
CPU	Intel(R) Core™i7－12700F
内存	32 G
CUDA	11.1

3.3 模型评估指标

为了能更好衡量水下场景下的字符识别算法的准确性和实时测量性能，本文对该任务设置的评价指标包括精准率(precision，P)、召回率(recall，R)、平均准确率(average precision，AP)和平均精度均值(mean average precision，mAP)4 个模型精度指标及浮点计算量、参数量(Parameters)和检测速度 3 个衡量模型复杂度的参数。精度指标计算公式如下：

$$P=\frac{TP}{TP+FP} \tag{6}$$

$$R=\frac{TP}{TP+FN} \tag{7}$$

$$AP=\int_0^1 P(R)\mathrm{d}R \tag{8}$$

$$mAP=\frac{1}{N}\sum_{i=1}^{N}AP_i \tag{9}$$

式中：TP 为正样本被成功识别的数量；FP 为负样本被错预为正样本的数量；FN 为正样本被错预为负样本的数量；mAP 则对各类别的平均精度取平均值，并以此作为评价模型性能的重要指标。此外，模型的计算量、参数量和检测速度被用于比较各算法的轻便程度。

3.4 实验结果与分析

为验证该方法在原模型上进步和优化，对本研究中的改进模块设计了一组以 YOLOv7-tiny 为基线的消融实验，实验结果如表 2 所示，其中√表示引入此模块。

由实验结果可知，相比于原始模型，模型 A 是将 MobileNetV3 直接替换原模型的骨干网络，参数量和计算量虽减少了 27.1%和 29.9%，但精度却也出现了小幅下降。为了在追求轻量化的同时保证精度，引入 PConv 以及置换注意力进行对比实验。模型 B 在更换骨干网络的基础

表 2　消融实验结果

Table 2　Results of ablation experiment

模型	改进部分			评价指标			
	MobileNetV3	PConv	SA	mAP@0.5 /%	Parameters /($\times10^6$)	浮点数 /GFLOPs	检测速度 /fps
YOLOv7-tiny	—	—	—	86.6	6.23	13.86	81.86
A	√	—	—	84.3	4.54	9.72	102.42
B	√	√	—	84.1	4.16	8.29	105.14
C	√	—	√	88.2	4.74	9.99	104.34
D	—	√	√	90.2	5.95	12.82	83.39
E	√	√	√	**88.7**	**4.36**	**8.52**	**107.11**

上对 ELAN 结构引入部分卷积结构，减少参数和计算量带来了检测速度的提升，精度保持稳定。模型 C 同样在更换骨干网络的基础上引入置换注意力机制而不引入 PConv，准确率相比未引入时提高了 4.6%，弥补了轻量化网络结构带来的部分精度损失。模型 D 则采用原骨干网络并引入 PConv 和 SA 来与原模型做比较，mAP 值提高 4.1%达到 90.2，但其检测速度的提升较小。考虑到任务的轻量化需求，骨干网络模型的替换不可或缺。模型 E 中引入本文涉及到的所有模块，相比与原模型精度只提高了 2.4%，但在参数量和计算量上都有明显减重效果，加快了模型的检测速度。综上，本文最终形成的算法框架得到了一定优化，在轻量化的同时也较好平衡了精度上的问题，为在实际水下检测场景中的部署提供了可行性。

YOLOv7-tiny 模型与本文改进模型的训练损失值对比如图 7 所示，可以看出，改进后模型收敛速度更快且损失值始终低于原始模型，这一结果表明改进后的模型更具优势。

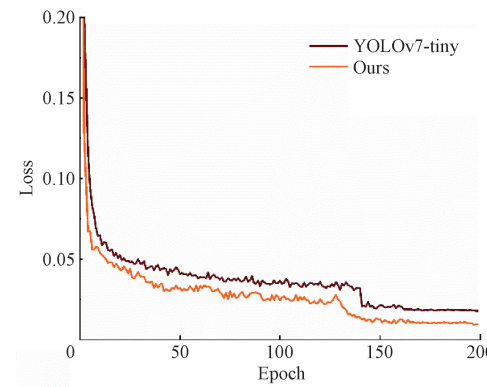

图 7　训练损失对比

Fig. 7　Comparation of train loss

为了验证改进后模型的整体性能及有效性，将其与其他目标检测模型在本文数据集上就准确率、检测速度和模型大小进行对比实验，其中包括主流的轻量化模型、泛化能力较好的复杂模型以及其他研究人员对 v7-tiny 进行改进后的模型，实验结果如表 3 所示。由实验结果可以看出，改进后的模型在准确率上相比 YOLOv4-tiny[15] 提升了 9.1%，且检测速度分别提高了 5.6%，其优势明显。其与 YOLOv7 相比时，在 mAP@0.5 指标虽然下降了 4.2%，但在检测速度上有着大幅提升，更符合移动部署轻量化需求。ECG-YOLO[16] 和 CURI-YOLO[17] 为其他学者基于 YOLOv7-tiny 做出的两种改进模型，与其他改进模型相比具备更好的检测效果。本文的改进方法相比于 ECG-YOLO 在准确率上稍显逊色但模型轻量化效果更好；然而相比于 CURI-YOLO，虽模型体积相比较大，但 CURI-YOLO 损失了太多的精度，实用准确率缺少保证。从实际应用的需求面考虑，实时检测类任务需兼顾准确率和检测速率，而本文提出的改进模型巧妙平衡了精度与模型大小，在自动化检测设备的部署上更具优越性，实用效果更佳。

表 3　对比实验结果

Table 3　Results of comparative experiment

模型	mAP@0.5/%	Parameters/($\times10^6$)	浮点数/GFLOPs	Weight/MB	检测速度/fps
YOLOv4-tiny	81.3	6.057	6.957	23.13	101.44
YOLOv7	92.6	37.09	103.35	143.70	18.71
ECG-YOLO	89.2	5.19	11.32	16.27	96.91
CURI-YOLO	76.7	**2.23**	**2.98**	**7.44**	**118.83**
本文	88.7	4.36	8.52	12.55	107.11

另外，为了探究改进模型在不同水下检测环境下对字符的识别能力，对上述不同的模型同时选取数据集中 7、3 两个字符在水下正常、过曝、波动和遮挡 4 种状态下的情况进行对比实验，实验效果如图 8 所示。实验结果表明，

图 8　不同场景下的实验对比

Fig. 8　Comparation of experiment in different scenarios

在正常和过曝的情况下，实验中的各个模型均具备很好的识别效果。但在有水波干扰和泥沙遮挡的情况下，各模型的识别准确率都出现了下降，但本文提出的改进模型有稳定的识别效果，对水下特殊情况下的字符识别有一定的应对能力，满足识别需求。

4　结　论

本文针对水下字符识别任务，提出了一种基于YOLOv7-tiny的轻量化目标检测模型，该模型同时具备高效和准确的特性。将主干网络更改为MobileNetV3在保证模型准确性的同时有助于减少模型中的参数量和计算量；利用PConv改进ELAN模块提高了模型的性能和效率；在模型中添加SA注意力机制提高了对图像中字符的检测性能。改进后的模型改善了模型的准确性和轻量化程度，提高了其在任务场景中的实际应用价值。

实验结果表明，本文提出的算法在精度方面小幅提升，在参数量和计算量上有大幅减少，且对易混淆的字符也有较好的识别效果，为水下智能化字符识别领域提供了一定参考价值。但在实际水下检测设备部署时情况会更为复杂，拍摄情况会受到水深的影响，随着水深的增加光线会被水逐渐吸收，使得拍摄出的图像颜色倾向蓝绿。除此之外，水下环境中还可能会出现热扰动的现象，对字符的拍摄观测产生一定的影响。未来将会在水下更深的区域进行研究和特征处理，进一步优化提高实际场景下的应用价值。

参 考 文 献

[1] NAIEMI F, GHODS V, KHALESI H. Scene text detection and recognition: A survey[J]. Multimedia Tools and Applications, 2022, 81(14): 20255-20290.

[2] 赵新年. 基于深度学习的水下字符识别算法研究与应用[D]. 成都：电子科技大学，2023.
ZHAO X N. Research and application of underwater character recognition algorithm based on deep learning [D]. Chengdu: University of Electronic Science and Technology of China, 2023.

[3] 汪志成，何坚强，翁嘉鑫，等. 基于改进 LeNet-5 的压印字符识别[J]. 计算机仿真，2022，39(2)：441-446.
WANG ZH CH, HE J Q, WENG J X, et al. Embossed character recognition based on improved LeNet-5[J]. Computer Simulation, 2022, 39(2): 441-446.

[4] 黄慧宁. 基于深度学习 YOLOv2 算法的钢材压印字符识别研究[D]. 南宁：广西大学，2021.
HUANG H N. Research on steel stamping character recognition based on deep learning YOLOv2 algorithm[D]. Nanning: Guangxi University, 2021.

[5] 宫鹏涵. 基于 YOLOv5 算法的钢印字符识别方法[J]. 兵器装备工程学报，2022，43(8)：101-105，124.
GONG P H. Steel seal character recognition method based on YOLOv5 algorithm[J]. Journal of Ordnance

Engineering, 2022, 43(8): 101-105,124.

[6] ZHANG Z, YANG G, WANG C, et al. Recognition of casting embossed convex and concave characters based on YOLOv5 for different distribution conditions[C]. 2021 International Wireless Communications and Mobile Computing (IWCMC). IEEE, 2021: 553-557.

[7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.

[8] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.

[9] QIAN S, NING C, HU Y. MobileNetV3 for image classification[C]. 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE). IEEE, 2021: 490-497.

[10] 梁秀满,安金铭,曹晓华,等. 基于改进 MobileNetV3 烧结断面火焰图像识别[J]. 电子测量技术,2023, 46(14):182-187.
LIANG X M, AN J M, CAO X H, et al. Based on improved MobileNetV3 sintering cross-section flame image recognition[J]. Electronic Measurement Technology, 2023, 46(14): 182-187.

[11] 赵子文,金永,陈友兴,等. 基于改进 YOLVOv5s 的 X 射线图像粘接缺陷实时检测[J]. 国外电子测量技术, 2023,42(4):181-186.
ZHAO Z W, JIN Y, CHEN Y X, et al. Real-time detection of X-ray image bonding defects based on improved YOLVOv5s[J]. Foreign Electronic Measurement Technology, 2023, 42(4): 181-186.

[12] CHEN J, KAO S, HE H, et al. Run, Don't Walk: Chasing higher FLOPS for faster neural networks[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 12021-12031.

[13] 章芮宁,闫坤,叶进. 轻量化 YOLO-v7 的数显仪表检测及读数[J/OL]. 计算机工程与应用,1-11[2024-04-09]. http://kns.cnki.net/kcms/detail/11.2127.TP.20230718.1838.014.html.
ZHANG R N, YAN K, YE J. Lightweight YOLO-v7 digital meter detection and reading[J/OL]. Computer Engineering and Applications, 1-11 [2024-04-09]. http://kns.cnki.net/kcms/detail/11.2127.TP.20230718.1838.014.html.

[14] ZHANG Q L, YANG Y B. SA-Net: Shuffle attention for deep convolutional neural networks[C]. ICASSP 2021—2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2021: 2235-2239.

[15] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection[J]. arXiv preprint arXiv:2004.10934, 2020.

[16] HU W, ZOU J, HUANG Y, et al. ECGYOLO: Mask Detection Algorithm [J]. Applied Sciences, 2023, 13(13): 7501.

[17] ZHANG Y, FANG X, GUO J, et al. CURI-YOLOv7: A lightweight YOLOv7tiny target detector for citrus trees from UAV remote sensing imagery based on embedded device [J]. Remote Sensing, 2023, 15(19): 4647.

作 者 简 介

李卓润,硕士研究生,主要研究方向为图像处理与深度学习。
E-mail:zr_li@cug.edu.cn

李波(通信作者),博士,教授,主要研究方向为机器人技术与智能装备、机器视觉技术应用。
E-mail:lbchql@163.com

邱鹏程,硕士研究生,主要研究方向为机器视觉与深度学习。
E-mail:1031519087@qq.com

刘洪,硕士研究生,主要研究方向为图像处理与目标检测。
E-mail:2417778505@qq.com