

doi:10.3969/j.issn.1673-9833.2025.01.008

融入注意力机制的水位检测算法

王坤侠^{1,2}, 夏涛¹

(1. 安徽建筑大学 电子与信息工程学院, 安徽 合肥 230601;
2. 安徽省古建筑智能感知与高维建模国际联合研究中心, 安徽 合肥 230601)

摘要: 传统图像处理的水位检测算法易受环境因素的影响, 因此提出一种融入注意力机制的水位检测算法。首先, 利用融入CBAM注意力机制的YOLO v5目标检测模型获取尺身数字类别及其坐标信息; 其次, 通过融入ECA注意力机制的DeepLabv3+语义分割模型实现水尺和背景的分割; 再次, 使用边缘检测算法得出水位线信息; 最后, 根据尺身数字信息和水位线信息将像素值换算为真实水位值。实验结果表明, 优化水位检测算法与人工读数误差在2 cm以内, 满足水位检测标准要求。

关键词: 水位检测; YOLO v5; DeepLabv3+; 注意力机制

中图分类号: TP391

文献标志码: A

文章编号: 1673-9833(2025)01-0057-07

引文格式: 王坤侠, 夏涛. 融入注意力机制的水位检测算法[J]. 湖南工业大学学报, 2025, 39(1): 57-63.

Water Level Detection Algorithm with Attention Mechanism Incorporated

WANG Kunxia^{1,2}, XIA Tao¹

(1. School of Electronic and Information Engineering, Anhui Jianzhu University, Hefei 230601, China; 2. Anhui International Joint Research Center for Ancient Architecture Intellisencing and Multi-Dimensional Modeling, Hefei 230601, China)

Abstract: In view of the susceptibility of traditional image processing water level detection algorithms to environmental factors, a water level detection algorithm has thus been proposed with attention mechanism incorporated. Firstly, the YOLO v5 object detection model, with CBAM attention mechanism incorporated, is utilized to obtain the categorical classification of the water gauge and its corresponding coordinate information. Secondly, the segmentation of water gauge and background can be achieved through the DeepLabv3+ semantic segmentation model with ECA attention mechanism incorporated. Then, an edge detection algorithms is applied to obtain water level information. Finally, the pixel values are converted to the true water level values based on the digital information of the ruler and the water level information. The experimental results show that the error between the optimized water level detection algorithm and manual reading is within 2 cm, which meets the requirements of water level detection standards.

Keywords: water level detection; YOLO v5; DeepLabv3+; attention mechanism

1 研究背景

水位检测是一项被广泛应用于各种工程和水资

源管理领域的重要技术, 对于水电站、水库、河流等水利工程的运行至关重要^[1]。同时, 在城市防洪、河道管理、水资源评估等领域中, 水位检测也扮演着重

收稿日期: 2023-10-20

基金项目: 安徽省住房城乡建设科学技术计划基金资助项目(2023-YF004, 2023-YF113); 安徽建筑大学智能建筑与建筑节能安徽省重点实验室开放课题基金资助项目(1BES2022ZR02)

作者简介: 王坤侠, 女, 安徽建筑大学副教授, 博士, 主要研究方向为人工智能, E-mail: kxwang@ahjzu.edu.cn

要的角色^[2]。传统的水位检测主要通过人工读数记录水位变化和在水位尺上安装电子传感器^[3]记录水位变化,这类方法的安装成本较高、耗费人力且不易维护。随着计算机视觉技术的发展,研究人员开始利用图像处理技术来检测水位。如孙维亚等^[4]利用关键字处理结果和边缘特征计算得到水位信息,张衍等^[5]先使用图像法对标尺进行分割与畸变矫正,再对水位线进行检测和水位值换算,完成水位值检测。张帆等^[6]提出了一种移动激光与视频图像处理相结合的方法,并且根据几何关系计算水位。吴婷等^[7]采用灰度拉伸策略检测水位线,构造了去噪模型,依据灰度直方图找出水面分界线以确定灰度拉伸范围,根据水尺和水平面交界线得出水位值。P. Bharti 等^[8]基于模板匹配算法,利用边缘地图定位 ROI (region of interest),并采取两种并行方法识别水位,一种方法在边缘图中拟合线性回归模型识别水位线,另一种方法是使用滑动窗口计算像素找出水位线。Chen G. J. 等^[9]构建了识别水尺字符多模板匹配算法和检查细化识别字符的序列验证算法,识别和确认尺身字符结果后通过计算得出水位值。

上述文献中提出的水位检测方法对于图像的前期预处理要求均较高,需按照图像个体特征进行二值化、边缘检测和倾斜矫正等前期处理,且传统图像处理技术在阈值参数设置方面受周围环境的影响,复杂背景下的适应能力较差,并不能有效地满足复杂背景下的实时水位检测要求。针对此类问题,本文提出了一种基于融入注意力机制的 YOLO v5^[10]和 DeepLabv3+ 网络^[11]的水位检测算法,使用改进的 YOLO v5 目标检测模型检测复杂背景下水尺尺身数字的类别和坐标信息,以获取水位的粗略信息,然后通过改进的 DeepLabv3+ 语义分割模型分割暴露在水面之上的水尺和背景,并通过边缘检测算法提取水位线坐标信息,综合这两个模型得出的信息实现水位值检测。实验结果表明,本文提出的算法有效提升了水位检测的准确性和鲁棒性,对参数设置要求较低,且具有较强的适应能力,能够适应复杂多变背景下的水位检测需求。

2 水尺刻度识别整体流程

水尺刻度识别即水位检测的算法流程如图 1 所示,具体步骤如下:

1) 利用改进的 YOLO v5 目标检测模型对水尺尺身数字进行识别,得到尺身数字类别和坐标等信息,并筛选识别出的数字,根据坐标信息保留图像最底部

数字的类别和坐标信息。

2) 将图像送入改进 DeepLabv3+ 图像分割模型,抽离出水尺和背景特征部分,得到 mask 图片,再由轮廓检测算法定位水位线坐标信息。

3) 通过线性插值法将图片像素换算为真实物理长度,根据底部数字类别、坐标信息和水位线坐标信息得出水尺水位值。

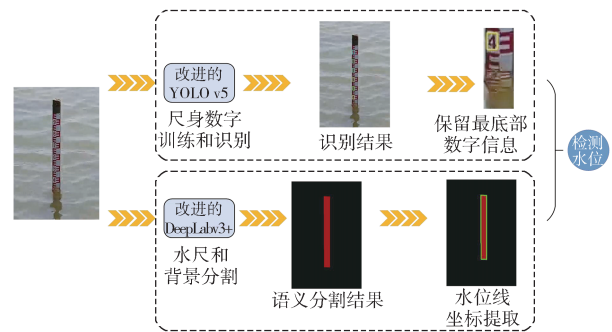


图 1 水位检测流程图

Fig. 1 Flowchart of water level detection

3 改进的 YOLO v5 尺身数字识别算法

3.1 YOLO v5 网络概述

本研究中选择网络深度和宽度最小,但是检测速度最快的 YOLO v5s 作为基础模型。YOLO v5s 的 Input 层采用 Mosaic 数据增强,并且自动计算数据集的最佳锚框值。Backbone 层作为主干网络,由 CSPNet (cross stage partial network)^[12]模块和 SPPF (spatial pyramid pooling-fast)^[13]模块组成,CSPNet 模块的核心在于分割梯度流,将梯度信息传递到不同的网络路径中,能够增强梯度的组合信息,更好地捕捉不同层次的特征。SPPF 模块使用多个小尺寸的池化核级联取代传统 SPP 模块中单个大尺寸池化核的方式,使得网络能够提升不同感受野的特征图表达能力。Neck 层由路径聚合网络 (path aggregation network, PAN)^[14]和特征金字塔网络 (feature pyramid network, FPN)^[15]协同提取图像特征,加强图像浅层信息和深层信息的融合。Output 层采用 GIOU_Loss^[16]作为边界框损失函数,针对多个预测框,YOLO v5s 采用非极大值抑制 (non maximum suppression, NMS)^[17]排除损失函数中贡献值较小的预测结果,从而生成最佳边界框和预测类别。

3.2 CBAM 注意力机制模块

CBAM (convolutional block attention module)^[18]注意力机制分为通道注意力模块 (channel attention module) 和空间注意力模块 (spatial attention module),获取图像的精炼特征,更高效地提升模型

分类结果, CBAM 的结构如图 2 所示。

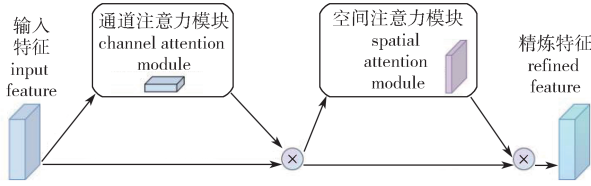


图 2 CBAM 注意力网络结构图

Fig. 2 CBAM attention network structure diagram

通道注意力模块^[19]对输入特征分别进行全局平均池化和全局最大池化, 在通道上实现维度压缩, 再经过共享多层感知机生成通道注意力权重。这使得网络更加关注重要的通道, 从而提取更有意义的特征。其计算式为

$$M_c(F) = \sigma \left(W_1 \left(W_0 \left(F_{avg}^c \right) \right) + W_1 \left(W_0 \left(F_{max}^c \right) \right) \right)。$$

式中: $M_c(F)$ 为通道注意力输出权重值; σ 为 sigmoid 激活函数; W_0 为第一个全连接层权重矩阵; W_1 为第二个全连接层权重矩阵; F_{avg}^c 为平均池化后在空间上的映射; F_{max}^c 为最大池化后在空间上的映射。

空间注意力模块^[20]则将通道注意力模块得到的结果, 通过最大值池化和平均池化, 分成两个大小为 $1 \times H \times W$ 的通道描述, 再通过全连接操作将张量堆叠在一起, 最后通过卷积操作和 sigmoid 激活函数得到权重系数 M_s , 通过学习特征图中不同空间位置的重要性来调整权重, 其计算式为

$$M_s(F) = \sigma \left(f^{7 \times 7} \left(\left[F_{avg}^s; F_{max}^s \right] \right) \right)。$$

式中: $M_s(F)$ 为空间注意力输出权值; $f^{7 \times 7}$ 为 7×7 的卷积运算滤波器; F_{avg}^s 为平均池化后的特征映射, F_{max}^s 为最大池化后的特征映射。

CBAM 注意力机制实现对通道和空间相关性的联合学习, 提高了特征表达能力。首先, 输入特征图 F 与通道注意力模块逐元素相乘, 生成的加强特征图 F' 再和空间注意力模块相乘, 经过 CBAM 注意力机制处理之后的特征图为 F'' , 其计算式如下:

$$F' = M_c(F) \times F, F'' = M_s(F') \times F'。$$

3.3 引入 CBAM 注意力机制的 YOLO v5 检测算法

在 YOLO v5s 网络的 Backbone 层中, C3 模块在特征中蕴含了大量细节信息和位置信息, 但缺乏语义信息。因水尺单个数字在整张图片上的像素占比较小, 常规 YOLO v5s 网络的 Backbone 层会忽略一些小目标信息, 加大了尺身数字识别的误检、漏检率, 故需加强 Backbone 层的目标纹理提取性能, 本文在 YOLO v5s 主干网络中加入 CBAM 注意力机制, 将 Backbone 层 4 个 C3 模块替换为 CBAM 模块, 改进后的网络结构见图 3。如图所示, 融入 CBAM 注意力机制不仅加强了各通道间的信息融合, 还有效提升

了模型对尺身数字点特征和线特征的学习能力。

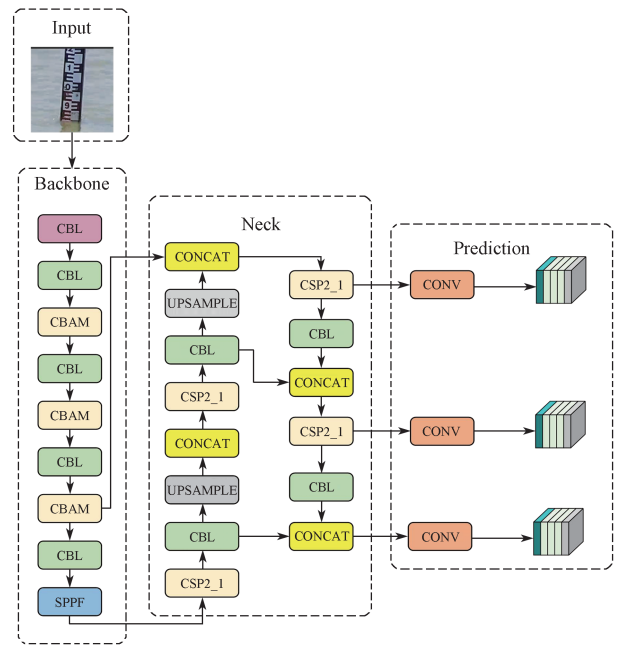


图 3 改进的 YOLO v5s 网络结构图

Fig. 3 Improved YOLO v5s network structure

3.4 尺身数字提取

融入 CBAM 注意力机制的 YOLO v5 模型, 能够有效检测出完整呈现在水面之上的 0~9 共 10 种数字类别, 在数字区域检测结果中, 每个矩形框的左上角和左下角坐标均为已知信息, 依据各输出检测框坐标信息提取图片最底部矩形框所包含的坐标信息, 将提取到的最底部数字类别设为 x , 数字预测框左上角的纵坐标为 h_2 , 左下角的纵坐标为 h_1 。

4 基于 DeepLabv3+ 水位线定位方法

4.1 DeepLabv3+ 模型概述

DeepLabv3+ 模型组合了编码器 (Encoder) 和解码器 (Decoder) 结构^[21], 且解码器使用 Xception 骨干网络和空洞空间卷积池化金字塔 (atrous spatial pyramid pooling, ASPP) 进行多尺度特征提取和通道加权处理。因为本研究只需分割出水尺尺身和背景信息, 语义信息相对较少, 所以本文选取 MobileNetv2^[22] 网络为 DeepLabv3+ 模型主干网络。MobileNetv2 独有的深度可分离卷积^[23]和 Xception 的标准卷积相比, 在运算量上具有明显优势。假设特征图片输入尺寸为 $D \times D$, 输入通道数为 M , 输出通道数为 N , C 为一个卷积核的边长, 则标准卷积的计算量 S_1 和深度可分离卷积计算量 S_2 的比值为

$$\frac{S_1}{S_2} = \frac{D \times D \times C \times C \times M + D \times D \times M \times N}{D \times D \times C \times C \times M \times N} = \frac{1}{N} + \frac{1}{C \times C}。$$

可以看出, 相较于 Xception 的标准卷积, 深度

可分离卷积显著降低了网络的参数量和计算量,大大减少了模型运算时间和复杂度。

在解码阶段,对上一步编码阶段获得的高维信息进行上采样,将其还原为原图像尺寸的像素级别,并对每个像素点进行语义分类,实现语义分割。

4.2 ECA 注意力机制概述

ECA (efficient channel attention) [24] 注意力机制对 SE [25] 注意力机制进行了优化,使用一维卷积实现局部跨通道交互,捕捉各通道间依赖关系,有效解决了 SE 注意力机制在压缩通道所带来学习通道间相互依赖关系紊乱的弊端。ECA 注意力机制结构见图 4。

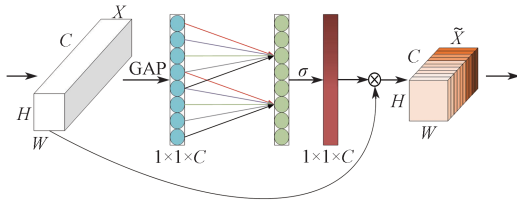


图 4 ECA 注意力网络结构图

Fig. 4 ECA attention network structure

ECA 使用内核大小为 K 的一维卷积 $C_1 D_K$ 实现各个通道权值 ω 的计算,式中 γ 为参数。

$$\omega = \sigma(C_1 D_K(\gamma))$$

通道的维数 K 值越大,局部跨通道交互的范围

越广,这两个参数之间的映射关系如下:

$$C = \phi(K) \approx \exp(\gamma \times K - b)$$

式中: γ 、 b 为调节参数。

在给定通道维数 C 的前提下,卷积核内核大小 K 可自适应确定,计算公式如下:

$$K = \varphi(C) = \left\lfloor \frac{\log_2(C) + b}{\gamma} \right\rfloor_{\text{odd}}$$

式中: $|a|_{\text{odd}}$ 为离 a 最近的一个奇数。

4.3 基于 ECA 注意力机制的 DeepLabv3+ 模型

因只需分割水尺和背景两个类别,语义信息较简单,鉴于常规 DeepLabv3+ 模型采用 Xception 为主干网络,虽然 Xception 有较强的解析力,但其参数量和计算量较大,分割物体种类少的情况下会带来训练困难和收敛速度慢等问题 [26]。为提升运算效率,本文将 DeepLabv3+ 主干网络调整为 MobileNetv2。

由于提取水位线的核心位置在水尺底部和水平面交界处,为提升两者分割精度,在 DeepLabv3+ 模型基础上引入 ECA 注意力机制,其结构见图 5。引入该机制,一方面可有效强调水尺尺身的重要信息,从而提高 DeepLabv3+ 模型的分割精度;另一方面,能够通过抑制冗余特征信息,提升模型运算效率。

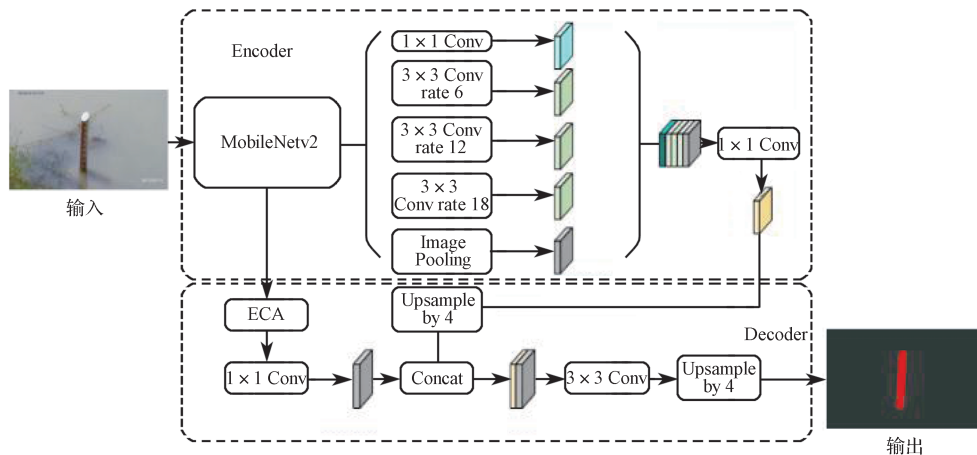


图 5 改进的 DeepLabv3+ 模型结构图

Fig. 5 Improved DeepLabv3+ model structure

4.4 水位线定位方法

根据语义分割模型输出结果,调用 Open CV 视觉库对水尺的分割结果图片进行轮廓检测,并对所有的轮廓信息进行存储。将存储的轮廓信息按照面积大小排序,只保留面积最大的轮廓,并绘制最小包围矩形。矩形的坐标信息是一致的,提取矩形包围框的左下角纵坐标 H 即水平面的纵坐标。

5 水位值运算

根据目标检测得到的底部数字类别 x , 数字预测

框左下角的纵坐标为 h_1 , 左上角的纵坐标为 h_2 , 语义分割获取的水面纵坐标为 H , 通过下式将其换算为真实水位值 h : $h = x - (H - h_1)(h_2 - h_1)/5$ 。

6 实验与结果

6.1 数据集

数据样本通过水文站摄像头拍摄的视频截取、校内拍摄和网络抓取 3 种方式获得,囊括 2 400 张图片,并按 8:2 的比例划分为训练集和验证集。用 Labelimg

标注工具, 按照 VOC 2007 格式对尺身数字 0~9 进行标注, 得到 XML 文件, 并转换为 YOLO v5 需要的 TXT 文件, 完成尺身数字目标检测数据集制作。用 Labelme 标注工具对水尺进行标注, 得到 JSON 文件, 并转换为 DeepLabv3+ 语义分割模型所需 PNG 图像, 以此制作水尺和背景的语义分割数据集。

6.2 实验结果分析

本研究实验环境如下: 操作系统为 Windows 10, CPU 为 Intel E5-2680 v4, GPU 为 1 块 A4000 共 16.9 GB 显存, 内存 32 GB, Pytorch 版本为 1.10, Python 版本为 3.7。

6.2.1 目标检测实验结果

本文采用精度 P 、召回率 R 、平均精度均值 mAP 和平均检测时间评价网络的性能。

P 的计算式为 $P = T_p / (T_p + F_p) \times 100\%$, 式中 T_p 为正确且预测结果是正的图片数量, F_p 为预测错误且预测结果是正的图片数量。

R 的计算式为 $R = T_p / (T_p + F_N) \times 100\%$, 式中 F_N 为预测错误且预测结果是负的图片数量。

mAP 的计算式为 $mAP = \frac{1}{m} \sum \left[\frac{1}{n} \sum P(r) \right]$, 式中 m 为分类的类别数, n 为单个类别中预测的目标数量, $P(r)$ 为召回率为 r 时的精确度值。

$mAP@0.5$ 表示 IoU (intersection over union) 阈值为 0.5 的 mAP , 而 $mAP@0.5:0.95$ 表示 IoU 阈值在 0.5~0.95 的平均 mAP , 该指标可以有效描述模型在不同 IoU 阈值下的性能。

为了定量分析融入 CBAM 注意力机制下的 YOLO v5s 网络的检测性能, 将该算法与 SSD (Single Shot MultiBox Detector)、YOLO v4 和 YOLO v5s 模型检测效果进行对比, 结果如表 1 所示。

表 1 不同算法的检测性能对比结果

Table 1 Comparison results of detection performance of different algorithms

模型	$P\%$	$R\%$	$mAP@0.5/\%$	$mAP@0.5:0.95/\%$	平均检测时间/ms
SSD	90.91	71.43	92.65	55.18	16.9
YOLO v4	97.36	93.39	95.09	61.59	25.0
YOLO v5s	97.51	97.86	98.13	64.18	14.0
CBAM+YOLO v5s	98.48	97.92	98.46	66.33	10.4

由表 1 中的数据可以得出, 改进的 YOLO v5s 的各个评价指标值均高于其他模型指标的对应值, 其中精度分别比 SSD、YOLO v4 和 YOLO v5s 提升了 8.83%, 1.15%, 0.99%, 召回率分别提升了 37.08%,

4.85%, 0.06%, $mAP@0.5$ 分别提升了 6.27%, 3.54%, 0.34%, $mAP@0.5:0.95$ 分别提升了 20.21%, 7.70%, 3.35%。而且在平均检测时间方面, 改进后的算法更有优势。因此改进后的 YOLO v5s 整体性能获得提升, 增加了对水尺尺身数字识别的准确率。

对于水尺尺身 10 种数字单独的准确率, 本文通过建立混淆矩阵图进行可视化, 图 6 展示了本文模型在目标检测自建数据集上的混淆矩阵, 通过分析混淆矩阵, 本文模型在识别 10 种数字类别方面已经达到了很高的准确率。

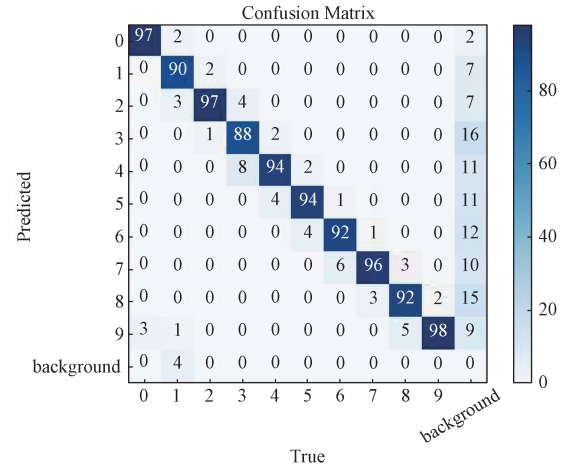


图 6 本文模型在数据集上的混淆矩阵

Fig. 6 Confusion matrix of this proposed model on the dataset

6.2.2 语义分割实验结果

语义分割采用平均像素准确率 MPA 、平均交并比 $mIoU$ 和平均检测时间 3 个评价指标来评估模型对水尺尺身的语义分割精度。其中, MPA 为各类别正确分类的像素数量与总像素数量之比的均值; $mIoU$ 为各个类别的像素预测值与真实值集合的交集与并集之比的平均值。

MPA 计算式为 $MPA = \frac{1}{k+1} \sum_{i=0}^k \left(\frac{p_{ii}}{\sum_{j=0}^k p_{ij}} \right)$, 式中

k 为种类数量, p_{ii} 为预测正确的像素数量, p_{ij} 为 i 类被错误地预测为 j 类的像素数量。

$mIoU$ 计算式为 $mIoU = \frac{1}{k} \sum_{j=1}^k \left(\frac{p_{jj}}{\left(\sum_{j=1}^k p_{ij} + \sum_{j=1}^k p_{ji} - p_{ii} \right)} \right)$,

式中 p_{ji} 为 j 类被错误预测为 i 类的像素数量。

同样的, 为定量分析融入 ECA 注意力机制下的 DeepLabv3+ 网络检测效果, 将该算法与 FCN (fully convolutional networks)、U-net 和 DeepLabv3+ 模型检测效果进行对比, 结果见表 2。由表 2 可知, 改进的 DeepLabv3+ 的 MPA 、 $mIoU$ 和平均检测时间 3 个评价指标值均高于其他模型的对应值, 其中, MPA 分别提升了 8.83%, 7.92%, 1.21%, $mIoU$ 分别提升了

13.13%, 9.76%, 0.21%, 且在检测速度上获得提升。对比可知改进后的 DeepLabv3+ 整体性能获得提升, DeepLabv3+ 在分割水尺尺身和背景的能力有效提升, 能够更加精确地提取水位线数值。

表 2 水尺图像分割量化结果

Table 2 Segmentation and quantization results of water gauge image

模型	MPA/%	mIoU/%	平均检测时间 /s
FCN	90.80	84.71	2.57
U-net	91.57	87.31	2.59
DeepLabv3+	97.64	95.63	0.41
ECA+DeepLabv3+	98.82	95.83	0.23

6.2.3 水位值读数分析

本文选取 4 个水文站水位信息为测试数据, 将本文算法得到的部分水位数据和人工读数进行对比, 结果见表 3。可看出该算法得到的水位值和人工读数绝对误差控制均在 2 cm 以内, 满足水位检测要求。

表 3 4 个站点水位检测结果

Table 3 Water level detection results at four sites

观测站点	本文算法读数 /cm	人工读数 /cm	绝对误差 /cm
站点 A	84.11	84.93	0.82
	84.33	84.50	0.17
	84.63	83.83	0.80
	81.96	82.67	0.71
站点 B	34.21	34.67	0.46
	36.50	35.83	0.67
	35.26	36.00	0.74
	34.60	35.00	0.40
站点 C	12.93	12.16	0.77
	12.50	11.33	1.17
	13.12	13.00	0.12
	12.82	11.77	1.05
站点 D	8.76	9.33	0.57
	8.88	8.33	0.55
	7.56	8.23	0.67
	8.50	8.60	0.10

7 结语

本文提出了一种水位检测算法, 对于目标检测和语义分割进行了改进优化: 首先, 利用 CBAM 注意力机制下的 YOLO v5 检测算法获取尺身数字和坐标信息; 然后, 使用 ECA 注意力机制下的 DeepLabv3+ 语义模型分割出水尺和背景, 在分割图基础上进行边缘检测获取水位线的坐标信息; 最后根据与真实世界物理长度映射得到水尺读数。

实验结果表明, 改进的 YOLO v5 的精度为 98.48%, 高于常规模型的对应值, 而改进的 DeepLabv3+ 网络不仅将准确率提升至 98.82%, 而且显著减少了参数量, 加快了收敛速度, 使得训练更简单、模型运算效率更高。

现实场景测试结果表明, 本文提出的算法可以精确计算出水尺读数, 水位检测精度在 2 cm 以内, 能够满足水尺读数检测要求。

参考文献:

- [1] 齐泓玮, 尚松浩, 李江. 中国水资源空间不均匀性定量评价[J]. 水力发电学报, 2020, 39(6): 28-38. QI Hongwei, SHANG Songhao, LI Jiang. Quantitative Evaluation on Spatial Heterogeneity of Water Resources in China[J]. Journal of Hydroelectric Engineering, 2020, 39(6): 28-38.
- [2] 姚蕊, 杨群涛, 张书亮. 城市暴雨内涝灾害脆弱性研究综述[J]. 水资源保护, 2023, 39(1): 93-100. YAO Rui, YANG Quntao, ZHANG Shuliang. Review on Vulnerability of Urban Rainstorm Waterlogging Disaster[J]. Water Resources Protection, 2023, 39(1): 93-100.
- [3] PAUL J D, BUYTAERT W, SAH N. A Technical Evaluation of Lidar-Based Measurement of River Water Levels[J]. Water Resources Research, 2020, 56(4): e2019WR026810.
- [4] 孙维亚, 王达, 许帅, 等. 基于计算机视觉的水位检测算法[J]. 应用科学学报, 2022, 40(3): 434-447. SUN Weiya, WANG Da, XU Shuai, et al. Water Level Detection Algorithm Based on Computer Vision[J]. Journal of Applied Sciences, 2022, 40(3): 434-447.
- [5] 张衍, 王剑平, 张果, 等. 图像法水位检测研究进展[J]. 电子测量技术, 2021, 44(13): 104-113. ZHANG Kan, WANG Jianping, ZHANG Guo, et al. Review of Image Water Level Detection[J]. Electronic Measurement Technology, 2021, 44(13): 104-113.
- [6] 张帆, 靳晓妍. 基于视频图像的嵌入式水位监测方法[J]. 中国测试, 2022, 48(12): 140-145. ZHANG Fan, JIN Xiaoyan. Embedded Water Level Monitoring Method Based on Video Image[J]. China Measurement & Test, 2022, 48(12): 140-145.
- [7] 吴婷, 褚泽帆, 陈城, 等. 基于灰度拉伸的图像水位识别方法研究[J]. 高技术通讯, 2021, 31(3): 327-332. WU Ting, CHU Zefan, CHEN Cheng, et al. Study on Image Water Level Recognition Based on Grayscale Stretching[J]. Chinese High Technology Letters, 2021, 31(3): 327-332.

- [8] BHARTI P, CHANDRA P, PAPKA M E, et al. An Edge Map Based Ensemble Solution to Detect Water Level in Stream[J/OL]. (2022-01-26) [2023-10-12]. arXiv preprint arXiv:2201.06098. <https://zhuanlan.zhihu.com/p/456380073>.
- [9] CHEN G J, BAI K J, LIN Z C, et al. Method on Water Level Ruler Reading Recognition Based on Image Processing[J]. *Signal, Image and Video Processing*, 2021, 15(1): 33-41.
- [10] MIAO L Z, LIN, ZHOU M L, et al. CBAM-Yolov5: Improved Yolov5 Based on Attention Model for Infrared Ship Detection[C]//International Conference on Computer Graphics, Artificial Intelligence, and Data Processing (ICCAID 2021). Harbin: SPIE, 2022. DOI:10.1117/12.2631130
- [11] REN S S, LIU Q, ZHANG X D. MPSA: A Multi-Level Pixel Spatial Attention Network for Thermal Image Segmentation Based on Deeplabv3+ Architecture[J]. *Infrared Physics & Technology*, 2022, 123: 104193.
- [12] WANG C Y, MARK LIAO H Y, WU Y H, et al. CSPNet: A New Backbone that Can Enhance Learning Capability of CNN[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Seattle: IEEE, 2020. DOI:10.1109/CVPRW50498.2020.00203.
- [13] HE K M, ZHANG X Y, REN S Q, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904-1916.
- [14] LIU S, QI L, QIN H F, et al. Path Aggregation Network for Instance Segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018. DOI:10.1109/CVPR.2018.00913.
- [15] LIN T Y, DOLLAR P, GIRSHICK R, et al. Feature Pyramid Networks for Object Detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu: IEEE, 2017. DOI:10.1109/CVPR.2017.106.
- [16] REZATOFIGHI H, TSOI N, GWAK J, et al. Generalized Intersection over Union: A Metric and a Loss for Bounding Box Regression[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019. DOI:10.1109/CVPR.2019.00075
- [17] ORO D, FERNÁNDEZ C, MARTORELL X, et al. Work-Efficient Parallel Non-Maximum Suppression Kernels[J]. *The Computer Journal*, 2022, 65(4): 773-787.
- [18] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block Attention Module[M]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 3-19.
- [19] 梁 斌, 刘 全, 徐 进, 等. 基于多注意力卷积神经网络的特定目标情感分析[J]. *计算机研究与发展*, 2017, 54(8): 1724-1735.
- LIANG Bin, LIU Quan, XU Jin, et al. Aspect-Based Sentiment Analysis Based on Multi-Attention CNN[J]. *Journal of Computer Research and Development*, 2017, 54(8): 1724-1735.
- [20] ZHU X Z, CHENG D Z, ZHANG Z, et al. An Empirical Study of Spatial Attention Mechanisms in Deep Networks[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul: IEEE, 2019. DOI:10.1109/ICCV.2019.00679.
- [21] CHEN L C, ZHU Y K, PAPANDREOU G, et al. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation[M]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 833-851.
- [22] 傅启凡, 路 茗, 张质懿, 等. 基于语义分割的水位监测方法研究[J]. *激光与光电子学进展*, 2022, 59(4): 0410004.
- FU Qifan, LU Ming, ZHANG Zhiyi, et al. Water Level Monitoring Method Based on Semantic Segmentation[J]. *Laser & Optoelectronics Progress*, 2022, 59(4): 0410004.
- [23] SANDLER M, HOWARD A, ZHU M, et al. MobileNetv2: Inverted Residuals and Linear Bottlenecks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City: IEEE, 2018: 4510-4520.
- [24] WANG Q L, WU B G, ZHU P F, et al. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020. DOI:10.1109/CVPR42600.2020.01155.
- [25] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-Excitation Networks[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(8): 2011-2023.
- [26] 张社荣, 欧阳乐颖, 王 超, 等. 基于DeepLabV3+的骨料图像自动分割算法[J]. *水利水电科技进展*, 2022, 42(6): 28-32, 97.
- ZHANG Sherong, OUYANG Leying, WANG Chao, et al. Automatic Segmentation Algorithm of Aggregate Image Based on DeepLabV3+[J]. *Advances in Science and Technology of Water Resources*, 2022, 42(6): 28-32, 97.

(责任编辑: 廖友媛)