

DOI: 10. 20040/j. cnki. 1000-7709. 2023. 20222376

# 基于 Mask RCNN 的河湖水位识别方法

徐 志<sup>1</sup>, 高 凯<sup>2</sup>, 高希超<sup>2</sup>, 梁犁丽<sup>1</sup>, 易 攀<sup>3</sup>

(1. 中国长江三峡集团有限公司科学技术研究院, 北京 100038; 2. 中国水利水电科学研究院, 北京 100038; 3. 北京市水文总站, 北京 100089)

**摘要:** 随着深度学习和图像识别技术的发展,通过视频实现城市河湖水位监测成为近年来研究热点。为全面监测城市河湖水位,提出了基于 Mask RCNN 的河湖水位识别方法,通过视频画面中的水尺字符及其位置关系得到水尺水位,并使用山东省东营市的现实水位站点监控视频进行了验证。结果表明,在现实水位站点监控视频识别中,水尺水位识别结果与实测数据对比误差小于 2、3、5 cm 的概率分别为 68.5%、76.9%、93.5%,平均误差为 2.1 cm,均方根误差为 3.0 cm,满足《数字孪生流域建设技术大纲(试行)》中对视频识别水尺水位的识别精度要求,故模型识别效果较好,且具有一定应用前景。

**关键词:** 水位识别; 图像识别; 深度学习; Mask RCNN

**中图分类号:** P332.3; [TV11]

**文献标志码:** A

**文章编号:** 1000-7709(2023)07-0023-04

## 1 引言

城市河湖水位是城市洪涝预报和预警的核心指标,也是评价城市洪涝灾害等级、评估洪涝灾害损失的关键因子,快速准确地获取城市河湖水位对城市防洪排涝具有重要意义。目前,多通过水位传感器设备监测城市河湖水位,该方法测量精度高、传输速度快,但传感器设备成本较高,安装数量较少,无法满足城市洪涝灾害的全面监测,同时需投入大量人力物力进行设备定期维护与修复<sup>[1]</sup>。随着数字技术的进步,基于视频图像的水尺水位识别和河道流量监测逐渐成为研究热点<sup>[2-3]</sup>,现有识别水尺水位的方法主要包括:①识别视频画面中的水尺区域,重点在于水面线提取。

②识别水尺字符与水尺刻度得到水尺水位。然而现有水尺水位识别方法的精度受水面线影响较大,对此本文提出一种受水面线影响较小的河湖水位识别方法,即基于 Mask RCNN 的河湖水位识别方法,只需视频画面中可以看清水尺字符,即可识别水尺水位,可为视频识别水尺水位的研究提供参考。

## 2 研究方法

采用深度学习模型 Mask-RCNN 进行水位识别,主要思路见图 1,首先通过 Mask-RCNN 识别视频画面中的水尺区域并进行分割,避免视频画面中背景环境的影响;再利用 Mask-RCNN 模型对分割后的水尺区域进行水尺字符识别,得到水尺区域的“m”字符、“E”字符和数字字符,并记

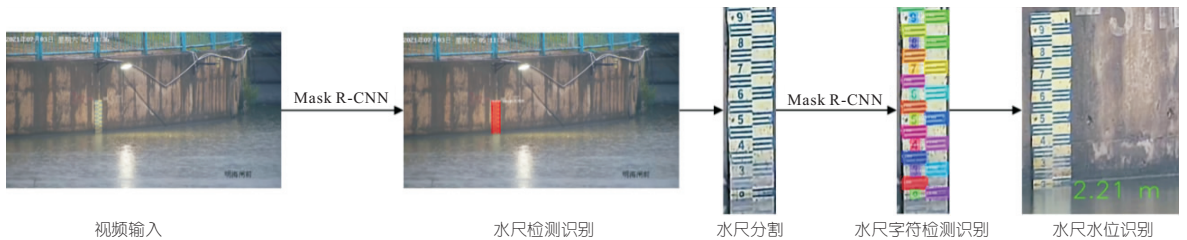


图 1 基于 Mask RCNN 的河湖水位识别思路

Fig. 1 Recognition of river and lake water level based on Mask RCNN

收稿日期: 2022-11-11, 修回日期: 2023-02-07

基金项目: 中国长江三峡集团有限公司科研项目(202203050); 流域水循环模拟与调控国家重点实验室团队项目(SKL2022TS11)

作者简介: 徐志(1990-),男,博士、工程师,研究方向为降水监测、多源降水融合,E-mail: xuzhitc159@163.com

通讯作者: 高希超(1988-),男,博士、工程师,研究方向为水文水资源,E-mail: 999gaoxichao@163.com

录字符在视频画面中的坐标;最后通过水尺字符的代表高度和在视频画面中的坐标位置,计算得到水位数据。

### 2.1 Mask RCNN 模型

#### 2.1.1 模型原理

Mask RCNN 模型网络结构见图 2,将预处理的图像输入模型,通过训练好的主干网络可得到对应的特征图(选用主干网络 ResNet50),对特征图的每个像素点设置预先确定的感兴趣区域(ROD),得到多个 ROI,通过区域候选网络(RPN)对这些 ROI 进行二值分类和边界框回归,过滤掉背景部分的 ROI,再将剩余的 ROI 进行 RoIAlign 操作,得到固定尺寸大小的特征图,最后通过全连接层(FC)预测各标识物的类标签和边界框,通过全卷积网络层(FCN)预测各标识物的对象掩码<sup>[4]</sup>。

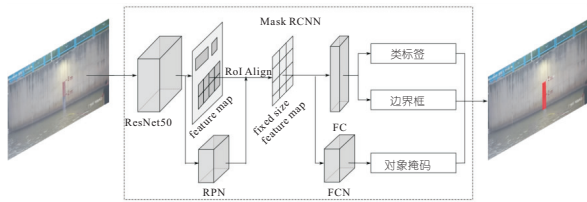


图 2 Mask RCNN 网络结构

Fig. 2 Mask RCNN network structure

#### 2.1.2 模型构建

(1)样本集制作。获取山东省东营市 2021 年汛期 134 个河道水位视频监控站点的监控视频,筛选出约 425 h 时长的包含有水尺的视频影像,并获取与视频同步的频率为 5 min 的实测水位数据,在此基础上共选取 642 张包含有水尺的视频画面制作样本集,其中随机选取 70% 作为训练集,30% 作为验证集,训练 Mask-RCNN 模型。

使用 labelme 标注工具对视频画面中的标识物进行标注,本研究设置 13 类标识物,分别为“water gauge”、“m”、“E”、“0”、“1”、“2”、“3”、“4”、“5”、“6”、“7”、“8”、“9”,其中“water gauge”用于检测与分割视频画面中的水尺区域,“m”用于确定水尺上限高度,“E”、“0”、“1”、“2”、“3”、“4”、“5”、“6”、“7”、“8”、“9”用于确定水尺读数。标注后的数据通过格式转换得到模型训练要求格式的数据。

(2)模型训练与测试。模型基于 TensorFlow 框架构建,模型训练主机操作系统为 Windows10,处理器为 AMD Ryzen 7 5800X 8-Core,显卡为 NVIDIA GeForce RTX 3090,用以实现 GPU 加速训练。模型训练初始学习率为 0.001,训练周期为 40,每个周期内训练 100 步,训练时的损失变化曲线见图 3,包括总损失

(loss)、类标签损失(class\_loss)、边界框损失(bbox\_loss)及对象掩码损失(mask\_loss)的变化。其中图 3(a)为第一个周期训练时 1~100 步的损失变化曲线,图 3(b)为第 1~40 个周期训练的损失变化情况。由图 3 可知,四种损失随着训练的进行整体下降并趋于平稳,经过第一周期训练,模型总损失由 7.9 优化为 2.9 左右,经过 40 个周期训练,总损失值稳定在 1.0 左右。图 4 为该模型在训练和验证过程中的总损失变化趋势,可以看出模型总损失值随着模型训练过程增加而不断下降,最后趋于稳定,在测试过程中,模型总损失值亦很快趋于稳定。

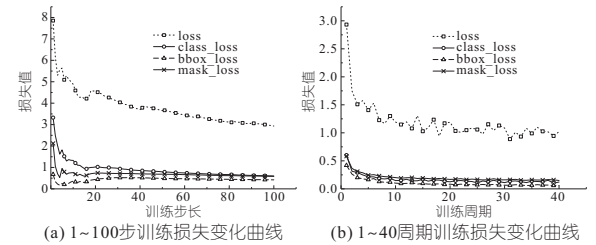


图 3 模型训练损失变化曲线

Fig. 3 Change curve of model training loss

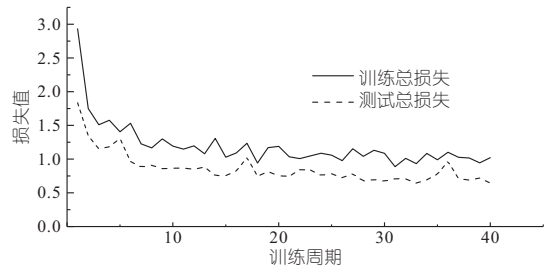


图 4 总损失值变化曲线

Fig. 4 Change curve of total loss value

### 2.2 水位识别

水位识别主要通过分析水尺区域水尺字符的代表高度和相对位置关系实现,计算流程见图 5。

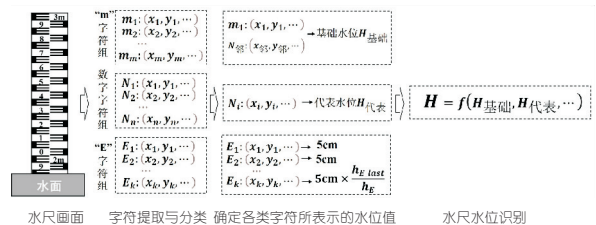


图 5 水尺水位计算流程

Fig. 5 Calculation process of calculating water level of water gauge

根据图 5,主要过程如下。

(1)水尺字符分类。水尺字符一般分为“m”字符、数字字符和“E”字符三类。其中“m”字符为度量单位,与其相邻数字字符组合可以确定水尺上限水位数值;数字字符有两个含义,与“m”字符相邻则代表整米刻度,其余代表整 10 cm 刻度;

“E”字符为更精细的水尺刻度,完整高度为 5 cm,精度为 1 cm。

(2)基础水位值确定。水尺测量范围较大时会包含多组“0~9”的数字字符,每组的基础高度不相同,在水位识别前需确定水尺下边界处于哪一测量区间,并赋予基础水位值,如水尺下边界处于 2~3 m 的测量区间时,基础水位值为 2 m。通过“m”字符的位置、与“m”字符相邻的数字字符的数值和数字字符“0”的数量确定基础水位值,计算公式为:

$$H_{\text{基础}} = N_{\text{邻}} - n_0 - 1 \quad (1)$$

式中, $H_{\text{基础}}$ 为基础水位值,m; $N_{\text{邻}}$ 为与“m”字符相邻的数字字符的数值; $n_0$ 为数字字符“0”的数量。

(3)代表水位值确定。在水尺区域的所有数字字符中选择一个代表数字作为代表水位值,以此为基础进行后续水位识别。基于数字字符在视频画面中的位置对其进行从高到低的排序,位置越低越靠近水面,为避免水面处的数字受水面遮挡识别不准确,选取倒数第二个数字作为代表水位值。

(4)“E”字符代表高度确定。完整的“E”字符代表高度为 5 cm;水面处的“E”字符受水面波动的遮挡,其代表高度可由其像素高度、完整“E”字符的像素高度和完整“E”字符的代表高度计算得到,计算公式为:

$$H_{E_{\text{last}}} = H_E h_{E_{\text{last}}} / h_E \quad (2)$$

式中, $H_{E_{\text{last}}}$ 为水面处“E”字符代表高度; $H_E$ 为完整“E”字符代表高度,取值 5 cm; $h_{E_{\text{last}}}$ 为水面处“E”字符像素高度; $h_E$ 为完整“E”字符像素高度。

(5)水位识别。根据基础水位值、代表水位值和“E”字符代表高度可计算得到水尺水位值,计算公式为:

$$H = N_{\text{邻}} - n_0 - 1 + 10N_{\text{代表}} - 5(n - 1) - 5h_{E_{\text{last}}} / h_E \quad (3)$$

式中, $H$ 为水尺水位值; $N_{\text{代表}}$ 为代表数字字符(水尺露出部分的倒数第二个数字的值); $n$ 为水尺露出部分中位置低于代表数字字符的“E”字符的个数。

### 3 结果

水尺识别包括水尺区域识别和水尺字符识别。其中水尺区域识别的目的是在具有复杂背景的视频画面中检测与分割出水尺区域,避免视频背景对水尺字符识别的影响;水尺字符识别的目的是计算水尺水位。

### 3.1 水尺识别结果

为验证水尺识别模型的准确性,选择山东省东营市明海闸后的河道水位监控视频进行水尺识别验证(图 6),该视频拍摄于 2021 年 7 月 3 日,视频时长 9.8 h,每间隔 5 min 截取一帧视频画面,共 108 帧视频画面用于水尺识别验证。由于水尺区域和水尺字符的特征明显,识别结果的准确性高,除了受水面遮挡导致水尺字符像素尺寸过小造成识别遗漏外,没有识别错误,故本研究所构建的模型可准确识别水尺区域与水尺字符,可用于后续水位识别。

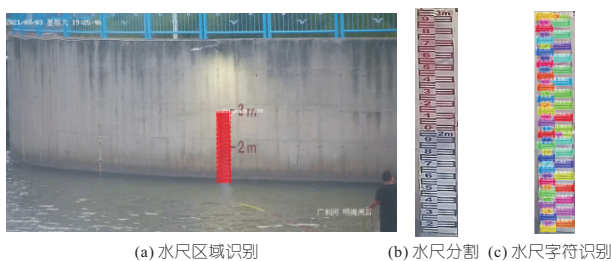


图 6 水尺识别结果

Fig. 6 Identification results of water gauge

### 3.2 水位识别结果

使用山东省东营市明海闸后的河道水位监控视频进行水位识别验证,该站点实测水位数据获取间隔为 5 min,视频共 108 个实测水位数据,水位变化介于 0.00~1.20 m,将对应时刻的 108 帧视频画面输入模型进行水位识别,并将识别结果与实测水位进行对比,结果见表 1 及图 7、8。结果表明,识别结果与实测水位对比误差小于 2 cm 的概率为 68.5%,误差小于 3 cm 的概率为 76.9%,误差小于 5 cm 的概率为 93.5%,平均误差为 2.1 cm,均方根误差为 3.0 cm。满足水尺读数综合误差不超过 0.03 m 的技术要求<sup>[5]</sup>,故模型识别效果较好,具有一定的应用前景。

表 1 明海闸后水位识别结果评价

Tab. 1 Evaluation of water level identification results of Minghai sluice back station

评价对象	误差小于 2 cm 的	误差小于 3 cm 的	误差小于 5 cm 的	平均	均方根
	概率/%	概率/%	概率/%	误差 /cm	误差 /cm
识别水位	68.5	76.9	93.5	2.1	3.0

将本文水位识别方法与水位传感器设备进行对比,以 OTT RLS 雷达水位计为例,该水位计测量精度为±3 mm,测量时间约 20~30 s;本文方法识别结果平均误差为 2.1 cm,均方根误差为 3.0 cm,应用本研究硬件配置的主机识别速度约 2.3 s/帧。对比可知,本文方法识别精度差于水位传感器设备,识别效率优于水位传感器设备,识

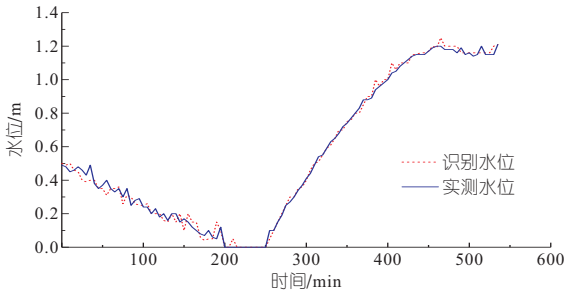


图 7 明海闸后水位数据对比图

Fig. 7 Comparison chart of water level data of Minghai sluice back station

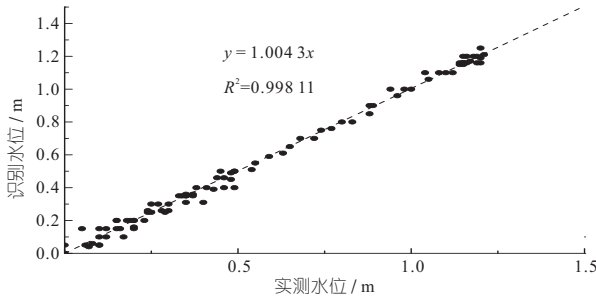


图 8 明海闸后水位识别结果相关性

Fig. 8 Correlation of water level identification results of Minghai sluice back station

别结果满足视频识别水尺水位的精度要求,可以补充无水位传感器设备区域的监测空白,为城市防洪排涝提供数据支撑。

## Water Level Identification Method of Rivers and Lakes Based on Mask RCNN

XU Zhi<sup>1</sup>, GAO Kai<sup>2</sup>, GAO Xi-chao<sup>2</sup>, LIANG Li-li<sup>1</sup>, YI Pan<sup>3</sup>

- (1. Science and Technology Research Inst., China Three Gorges Corporation, Beijing 100038, China;
- 2. China Institute of Water Resources and Hydropower Research, Beijing 100038, China;
- 3. Beijing Hydrology Center, Beijing 100089, China)

**Abstract:** With the development of deep learning and image recognition technology, monitoring the water level of urban rivers and lakes through video has become a hot research topic in recent years. In order to realize the comprehensiveness of urban river and lake water level monitoring, a method of river and lake water level identification based on Mask RCNN was proposed. The water level was obtained by the water level characters and their position relations in the video images, and it was verified by the monitoring video of the real water level station in Dongying City, Shandong Province. The results show that the probability that the comparison error between the water level identification result and the measured data was less than 2 cm was 68.5%, the probability of error less than 3 cm is 76.9%, the probability of error less than 5 cm is 93.5%, the average error is 2.1 cm, and the root mean square error (RMSE) is 3.0 cm, which meets the recognition accuracy requirements of video water gauge level in Technical Outline of Digital Twin Watershed Construction (Trial). Therefore, the model has a good recognition effect and a certain application prospect.

**Key words:** water level identification; image recognition; deep learning; Mask RCNN

\*\*\*\*\*  
(上接第 153 页)

## Influence of Structure and Interval Variation on Hydraulic Characteristics of Vertical Slot Fishway

LV Yang-yang<sup>1</sup>, YANG Lu-hua<sup>1</sup>, WANG Fan<sup>2</sup>, WANG Xu<sup>1</sup>

- (1. College of Hydraulic Engineering, Tianjin Agricultural College, Tianjin 300384, China;
- 2. Zhongshui North Survey, Design and Research Co. LTD., Tianjin 300222, China)

**Abstract:** The structure of fishway is the major factor that fish migrates successfully. The impact of pier head structure and horizontal interval of baffles on the hydraulic characteristics of a vertical slot fishway pool were researched. Twelve different working conditions were composed of two factors, namely, the diversion angle of pier head and the relative position of baffles. The RNG  $\kappa$ - $\epsilon$  turbulence model was used to simulate the hydraulic characteristics of fishway. The results indicate that the influence of pier head structure on hydraulic characteristics is mainly reflected on velocity and turbulent energy, and the smaller the diversion Angle, the smaller the velocity and turbulent energy are. The distance between flow deflector and baffle plate changed the mainstream flow pattern and area size, and then affected the distribution of flow velocity. When the relative position of  $b_0/L=0.15$  or so, the flow pattern in the pool was reasonable, which met the basic requirements of fish migration.

**Key words:** vertical slot fishway; pier head; interval; numerical simulation; hydraulic characteristics

## 4 结论

a. 提出的基于 Mask RCNN 的河湖水位识别方法思路简单、受水面线影响较小、精度较高,具有一定应用前景。

b. 识别水位的关键是水尺字符识别,当在特殊条件下水面产生水尺倒影时,水尺倒影中的水尺字符可能会影响识别结果,后续将进一步研究如何消除水尺倒影的影响。

### 参考文献:

- [1] 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): 89-100.
- [2] 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.
- [3] 陈梦,陈华,黄贵平,等. 河道流量视频测验技术应用研究及误差分析[J]. 中国农村水利水电. <http://kns.cnki.net/kcms/detail/42.1419.TV.20221013.0837.002.html>.
- [4] 任娣声. 基于深度学习的货架状态画像研究与实现[D]. 西安:西安理工大学,2021.
- [5] 中华人民共和国水利部. 数字孪生流域建设技术大纲(试行)[Z]. 2002.