

# 基于人工智能与随机场的水位图像识别方法研究

武泽恒, 陈 华\*, 黄凯霖, 王 俊, 郭生练, 许崇育

武汉大学水资源与水电工程科学国家重点实验室, 湖北 武汉

收稿日期: 2022年9月14日; 录用日期: 2022年10月21日; 发布日期: 2022年10月28日

## 摘 要

为了提升基于视频图像的自动水位监测的准确度和环境稳定性, 提出一种基于人工智能与随机场的水位图像智能识别方法。在Deeplabv3+语义分割模型的基础上引入注意力机制和条件随机场, 优化水位线识别, 结合摄像头标定结果插值计算水位值。设置3组不同条件的现场水位识别试验, 结果表明所提出的改进算法识别结果更精确, 具备环境适应性, 基本满足水文测验要求。

## 关键词

水位监测, 图像识别, 语义分割, 注意力机制, 条件随机场

## Study on Water Level Image Recognition Method Based on Artificial Intelligence and Random Field

Zeheng Wu, Hua Chen\*, Kailin Huang, Jun Wang, Shenglian Guo, Chongyu Xu

State Key Laboratory of Water Resources and Hydropower Engineering Science, Wuhan University, Wuhan Hubei

Received: Sep. 14<sup>th</sup>, 2022; accepted: Oct. 21<sup>st</sup>, 2022; published: Oct. 28<sup>th</sup>, 2022

## Abstract

In order to improve the accuracy and robustness of automatic water level monitoring based on video im-

作者简介: 武泽恒(2001-), 硕士研究生, 主要从事视频水文监测研究。Email: wu\_zheng@163.com

\*通讯作者 Email: chua@whu.edu.cn

文章引用: 武泽恒, 陈华, 黄凯霖, 王俊, 郭生练, 许崇育. 基于人工智能与随机场的水位图像识别方法研究[J]. 水资源研究, 2022, 11(5): 445-457. DOI: 10.12677/jwrr.2022.115049

ages, a video water level intelligent identification method based on artificial intelligence and random field is proposed. Based on Deeplabv3+ semantic segmentation model, attention mechanism and conditional random field are introduced to optimize water level recognition, and the water level value is calculated by interpolation combining with camera calibration results. Three groups of water level identification experiments with different conditions are set up. The results show that the proposed improved algorithm is more accurate, has environmental adaptability, and basically meets the requirements of hydrological test.

## Keywords

Water Level Monitoring, Image Recognition, Semantic Segmentation, Attention Mechanism, Conditional Random Field

Copyright © 2022 by author(s) and Wuhan University.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

水位作为水文测验的基础要素，是水利工程建设、水资源规划与管理、水生态保护、水灾害防治的重要依据。结合智慧水利总体架构对水文监测自动化、实时化、智能化的要求，目前常用的水位监测技术难以满足智慧水利系统建设的要求。传统的水位监测方法包括设立水尺、采用自记水位计、非接触式水位计等方法。水尺观测依赖人工，效率低下，智能化、自动化水平不足，在恶劣环境中观测风险大。浮子式水位计是目前我国使用最多的自记水位计，具备结构简单、性能稳定的优点，但需要修建测井，土建成本较高，在水位变幅大的情况下易发生机械故障[1]。压力式水位计虽然无需修建水位测井，但长期观测稳定性较差，易受到泥沙淤积影响，设备维护困难。超声波水位计是常用的非接触式水位监测设备，避免了水下环境对测量的影响，但易受到空气温度、湿度、粉尘的影响，测量存在盲区。雷达水位计相比于超声波水位计抗干扰性更强，量程更大，但测量易受到水面漂浮物和雨雪天气的影响[2]。

近年来，随着水利信息化建设以及河长制的推广，视频监控设备成为了河流的标准配置，为水位图像识别方法的研究提供了数据基础。水位图像识别方法基本过程为：首先分割图像中水面与非水面部分，然后识别水位线，并结合水尺或其他参照物的高程信息进行水位值计算。Yu [3]等人利用边缘检测算子定位水尺，使用形态学开运算获取水尺刻度线，建立像素高度查找表推算水位。张振[4]等将水尺图像与标准水尺模板控制点匹配进行水位识别。周衡[5]等采取图像灰度差分法提取水面区域，根据相机标定结果进行曲线拟合识别水位。水位值的计算依赖于水位线的识别，传统的图像处理方法对水尺定位来识别水位线时，通常针对图像中人工选取的单一灰度特征进行处理，在应用时难以适应复杂多变的环境条件。深度学习算法可以在大量水面场景数据驱动下提取图像的泛化特征，相比于传统图像处理算法环境稳定性更强。Bai [6]等人通过训练 SSD (Single Shot MultiBox Detector) 目标检测模型得到水尺区域，根据像素高程关系计算水位值。Jafari [7]等人引入深度语义分割网络进行水位线识别，无需定位识别水尺直接获取水位。

目前运用基于深度学习的语义分割网络在无水尺条件下进行水位识别的研究仍处于起步阶段，将语义分割模型直接用于水位识别任务仍存在问题。第一，已有研究中采取的语义分割模型针对图像中不同位置和通道的特征关注程度相同，未对关键特征加以区分。第二，模型未充分考虑图像中相同类别像素点之间的相关性。同时受制于水面图像数据集的规模与标注质量，现有的语义分割模型对于水面区域的识别效果较差，存在误划分的孤立区域，尤其是边缘处水位线往往存在异常突起和缺失。因此进一步优化用于分割水面区域的深度语义

分割模型，提升水位线的识别效果，以提升水位识别的精度十分必要。

本文在 Deeplabv3+模型的基础上引入 CBAM (Convolutional Block Attention Module)注意力模块提升整体分割精度，利用条件随机场优化水面边缘的分割精度，构建“数字水尺”进行水位读数，实现无水尺条件下水位精确识别，提升了复杂环境下图像水位识别的精度和环境稳定性。

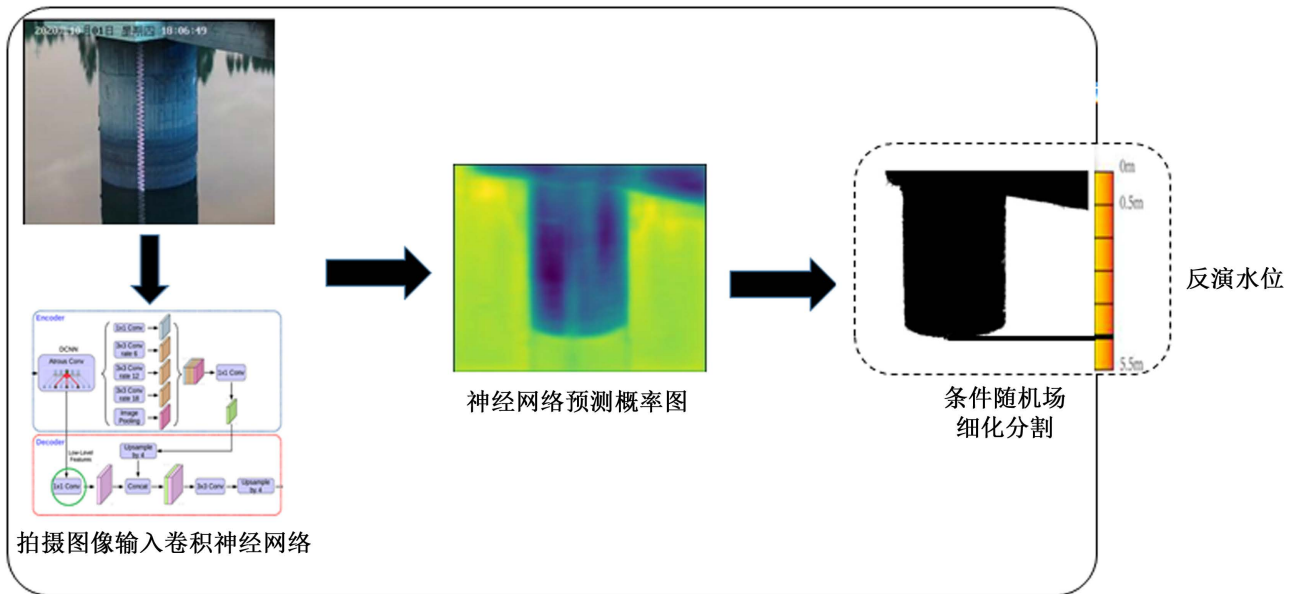


Figure 1. Diagram of the implementation of the water level recognition method  
图 1. 水位识别方法实施流程图

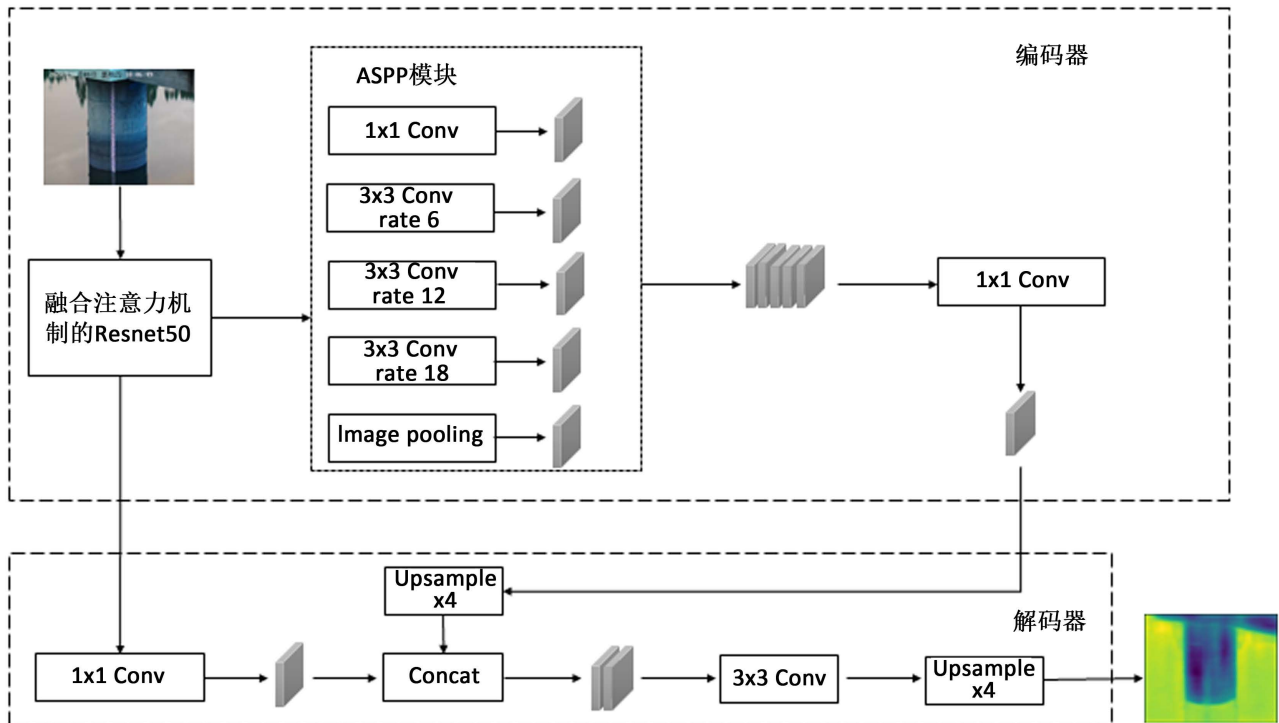


Figure 2. Improved Deeplabv3+ model structure  
图 2. 改进 Deeplabv3+模型结构图

## 2. 水位图像智能识别方法

### 2.1. 计算流程

图 1 为水位识别方法的实施流程，在进行水位识别时，将水面图像输入预训练的语义分割网络，模型会输出预测水面概率分布图，表示每个像素点属于水面区域的概率。结合模型预测结果和原图像的信息进行条件随机场优化，实现水位线的精确定位。由于当摄像机固定后，每个像素点对应的世界坐标也随之固定，采取线性插值即可推算出该点的水位。

### 2.2. 基于人工智能的水位线识别

#### 2.2.1. Deeplabv3+模型

Deeplabv3+ [8]是谷歌公司提出的语义分割模型，采取“编码器-解码器”架构。编码器部分由特征提取网络和空洞空间池化金字塔模块(Atrous Spatial Pyramid Pooling, ASPP)构成。原模型特征提取网络为带空洞卷积的Xception网络或ResNet网络，本文选取融合CBAM注意力模块的ResNet-50网络。多个具有不同空洞率的卷积层并行组成ASPP模块，保证特征图分辨率的同时，整合图像多尺度信息，以提升模型的预测精度。解码器的输入为卷积神经网络提取的浅层特征和ASPP模块输出的深层特征，进行拼接和上采样操作恢复图像信息，输出概率分布图。改进后模型整体结构如图2所示。

#### 2.2.2. CBAM注意力模块

图像处理领域中的注意力机制是对视觉注意力机制的模拟，人眼在接受输入的海量信息时，会选择性聚焦环境中的关键区域，忽视无关区域，从而将视觉信号处理神经元分配给重点区域，为更高级的视觉处理任务提供相关信息[9]。普通卷积神经网络中对于特征空间各通道和区域保持同等程度的关注，但针对不同的任务，特征空间中的不同通道和不同区域的重要性显然不同。本文在ResNet-50网络中引入CBAM模块[10]，训练不同通道和区域相应的权重，提升模型精确识别水面区域的能力。

图3展示了通道注意力模块的结构，给定一个维度为通道数  $C$ ，高度  $H$ ，宽度  $W$  的特征图  $F$  作为输入：

$$F \in R^{C \times H \times W} \quad (1)$$

式(2)展示了通道注意力权重的计算过程，输入特征  $F$  同时进行空间维度上的全局最大池化和平均池化，得到两个池化特征图  $[F_{Max}, F_{Avg}] \in R^{C \times 1 \times 1}$ 。之后再将它们同时输入一个共享的两层感知机，将输出特征进行逐元素对应加和，再输入sigmoid激活函数，生成通道注意力权重矩阵  $Mc \in R^{C \times 1 \times 1}$ 。 $Mc$ 中每一通道的权重，表示该通道对于水面区域信息的重要性和相关性。

$$Mc(F) = \sigma(\text{MLP}(\text{MaxPool}(F)) + \text{MLP}(\text{AvgPool}(F))) \quad (2)$$

式中： $\sigma$ 表示sigmoid激活函数，MLP表示感知机模型运算，MaxPool表示最大池化，AvgPool表示平均池化。

图4展示了空间注意力模块的结构，将输入特征图  $F \in R^{C \times H \times W}$  并行基于通道维度的全局最大池化和平均池化，得到两个池化特征图  $[F_{Max}, F_{Avg}] \in R^{1 \times H \times W}$  后进行通道拼接，得到拼接特征图  $F_{con} \in R^{2 \times H \times W}$ ，之后利用  $3 \times 3$  的卷积层进行卷积操作后进行sigmoid激活，生成空间注意力权值  $MS \in R^{1 \times H \times W}$ ，如公式3所示。

$$Ms(F) = \sigma(f^{3 \times 3}[\text{MaxPool}(F); \text{AvgPool}(F)]) \quad (3)$$

图5展示了嵌入CBAM注意力模块的残差结构。将残差卷积块提取的特征输入CBAM模块中，通道注意力模块和空间注意力模块采取串行连接，将输入特征与提取特征进行残差连接构成残差结构，使神经网络关注到图像中更重要的特征和空间位置。

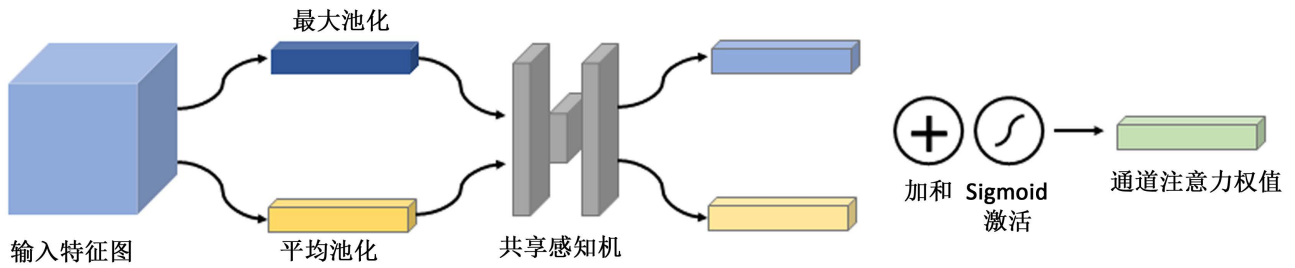


Figure 3. Channel attention module  
图 3. 通道注意力模块

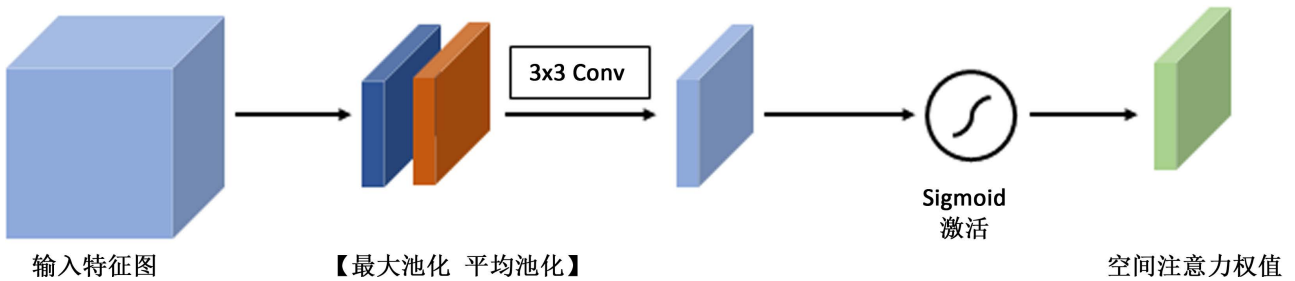


Figure 4. Spatial attention module  
图 4. 空间注意力模块

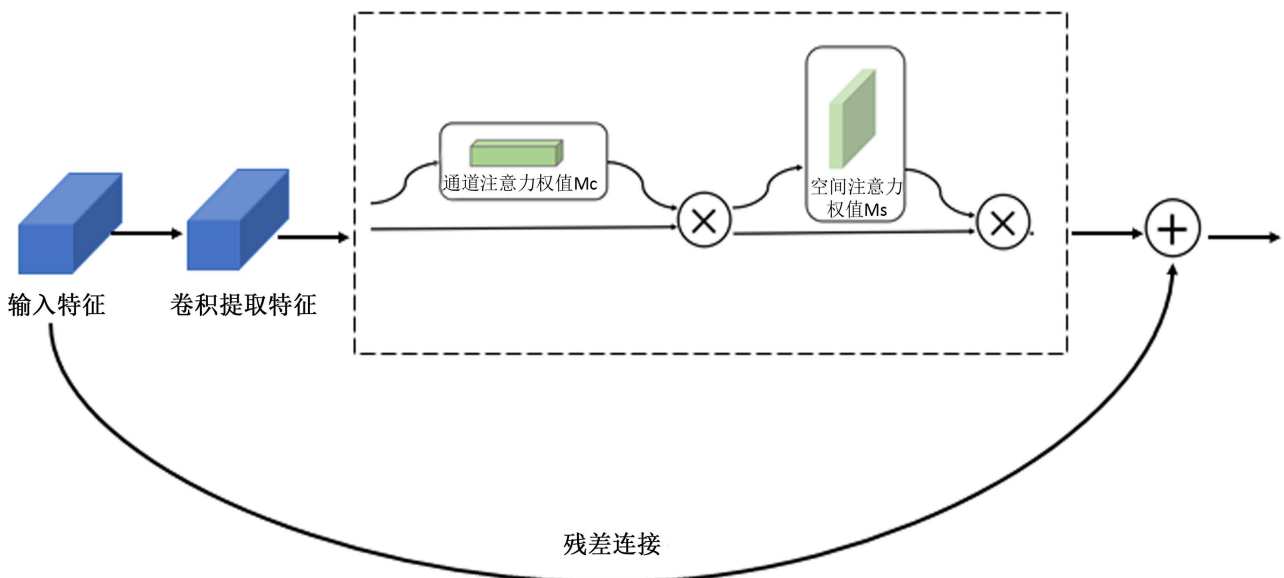


Figure 5. Residual structures embedded in CBAM modules  
图 5. 嵌入 CBAM 模块的残差结构

### 2.3. 基于随机场的水面边缘细化

全连接条件随机场[11]作为边缘细化后处理模块，从原图获取像素点的颜色、位置信息并计算成对势能，从模型输出分割图获取各像素点所属类别的概率并计算单点势能，逐步迭代计算势能最低时各像素的所属类别，综合考虑了图像中的像素强度的相似性和空间位置的连续性，提升了水面边缘分割的精度。

具体地，将像素强度定义为  $I = (I_1, I_2, \dots, I_n)$ ，像素点对应的预测变量定义为  $X = (X_1, X_2, \dots, X_n)$ ，其中预测变量在类别标签集合  $L = \{l_1, l_2, \dots, l_n\}$  中取值，本文中  $n = 2$ ，表示像素所属的类别包括水域和非水域两种类别，



条件随机场定义为像素强度与预测变量集合(I, X)的条件概率分布:

$$P(X|I) = \frac{1}{Z(i)} \exp\left(-\sum_c E(X_c)\right) \tag{4}$$

其能量函数的形式可描述为:

$$E(x) = \sum_i \Psi_u(x_i) + \sum_i \Psi_p(x_i, x_j) \tag{5}$$

单点势能不考虑像素间的联系, 初始化为水面概率分布图的负对数似然函数, 表示根据像素点本身的特征将像素点分类为水面的代价:

$$\Psi_u(x_i) = -\log(P(x_i)) \tag{6}$$

成对势能用于描述像素点之间的联系强度, 通过像素点的颜色信息和位置信息计算。

$$\Psi_p(x_i, x_j) = u(x_i, x_j) \sum K_G^M(f_i, f_j) \tag{7}$$

式中:  $u(x_i, x_j)$  用于判断像素点的连接关系, 约束了像素点之间能量函数传导的条件, 只有两个像素点类别相同时, 能量函数才可以相互传导。

式(7)中  $K_G^m(f_i, f_j)$  为表示像素点之间强度和空间相关性的高斯核函数[12], 其函数形式如公式(8)所示。

$$K_G^M(f_i, f_j) = \omega_1 \exp\left(-\frac{\|p_i - p_j\|^2}{2\sigma_\alpha^2} - \frac{\|I_i - I_j\|^2}{2\sigma_\beta^2}\right) + \omega_2 \exp\left(-\frac{\|p_i - p_j\|^2}{2\sigma_\gamma^2}\right) \tag{8}$$

式中,  $p_i$  代表像素点的坐标,  $I_i$  表示像素点的像素强度。第一项称为表面核, 包括由  $\sigma_\alpha$  控制的像素点的空间距离关系和由  $\sigma_\beta$  控制的像素点的像素强度的关系, 鼓励两个像素强度和空间位置上接近的像素点归为同一类别。第二项为称为平滑核, 目的在于保证图像中同一类别区域的空间连续性, 避免出现孤立区域。

## 2.4. 基于水位图像智能识别的数字水尺

当摄像机被固定后, 图像中每个像素点对应现实空间中的高程也随之确定。因此在进行水位监测前, 应预先进行标定工作。固定摄像机, 结合水尺刻度采集若干特征像素点坐标及其对应的高程, 表示为集合  $\{(u_i, v_i), h_i\}$ , 其中  $(u_i, v_i)$  表示采集的第  $i$  个像素点在图像中的坐标,  $h_i$  为采集的第  $i$  个像素点对应的高程。

在特征点采集后利用线性插值法构建“数字水尺”, 假设“数字水尺”上某点坐标为  $(u, v)$ , 在所述集合中寻找最邻近的像素点, 坐标分别为  $(x_u, y_u)$ 、 $(x_d, y_d)$ , 该点对应的高程  $h$  计算如公式(9)所示:

$$h = h(x_d, y_d) + \left( h(x_u, y_u) - h(x_d, y_d) \right) \frac{\sqrt{(x_u - u)^2 - (y_u - v)^2}}{\sqrt{(x_d - u)^2 - (y_d - v)^2}} \tag{9}$$

当水面区域分割完成后, 利用水面边缘水位线与预先构建的“数字水尺”交点的像素坐标, 即可反演得到该时刻水面边缘相应水位。

## 3. 水位智能识别模型训练

### 3.1. 数据集构建

本文的数据集在清华大学[13]和欧卡智舶[14]公开的水面图像数据集的基础上进行扩充, 除公开数据集中 2500 张左右水面图片和对应的标注结果外, 又自行采集部分水文站点视频监控图像进行标注加入数据集。

数据集包括 3492 张各种天气状况及光照条件下不同场景的水面图片, 范围涵盖白天、夜晚、雨雾、风浪条

件下的野外河道、城市河道、涵洞、池塘、湖泊、海滩等水域场景，部分图片如图 6 所示。

此外，为了扩充水面图像数据集，采取了裁剪、畸变、雾化、高斯模糊等数据增强手段丰富样本。最终用于训练的数据集大小为 10,000 张。其中 80% 的图片划分为训练集，5% 的图片划分为验证集，15% 的图片划分为测试集。



Figure 6. Partial image of the dataset  
图 6. 数据集部分图片

### 3.2. 语义分割模型测试

#### 3.2.1. 测试环境

模型搭建框架为 Pytorch 1.11.0 + CUDA 11.3，操作系统为 Windows10，CPU 为 Intel Core i7-10700，GPU 为 Nvidia RTX3060，显存 12 GB。

模型训练时采用 Adam 优化器，一次训练的样本数目(batchsize)设定为 10，共训练 50 个轮次(epoch)，初始学习率设置为  $10^{-4}$ 。

#### 3.2.2. 评价指标

为了评价模型的对于水面区域的分割效果，选取模型在测试集上的平均像素精度(MPA)和平均交并比(mIoU)作为模型分割效果的评价指标。

MPA 表示每类区域中被正确预测的像素点占该类像素点总数的比例的平均值：

$$MPA = \frac{1}{k+1} \sum_{i=0}^k \frac{P_{ii}}{\sum_{j=0}^k P_{ij}} \quad (10)$$

mIoU 表示每类预测区域和真实区域的像素点的交集和并集之比的平均值：

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k P_{ij} \frac{P_{ii}}{\sum_{j=0}^k P_{ij} + \sum_{j=0}^k P_{ji} - P_{ii}} \quad (11)$$

式中： $k$  代表类别数， $p_{ij}$  表示真实类别为第  $i$  类，识别结果为第  $j$  类的像素点个数。

#### 3.2.3. 分割性能对比

表 1 给出了本文改进算法与 SegNet [15]，UNet [16]，FCN [17]，Deeplabv3+等常用语义分割算法针对水面

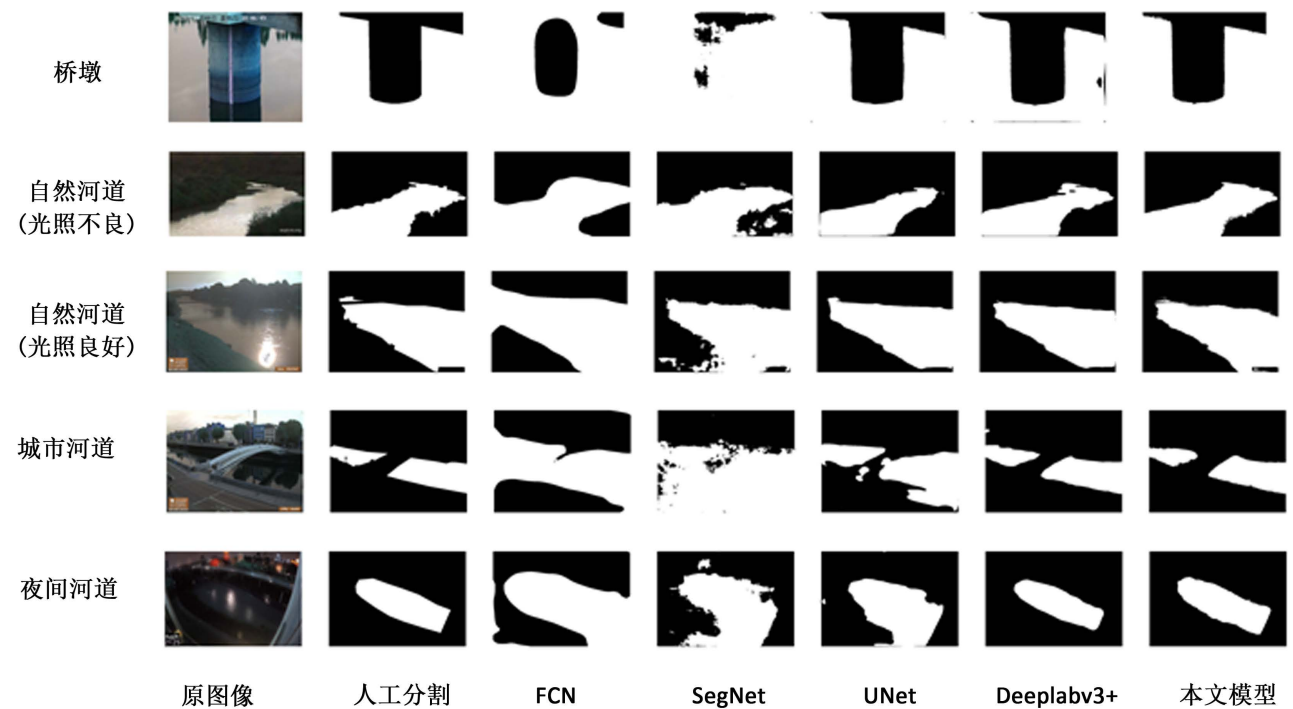
分割任务的平均像素精度和平均交并比指标。

本文改进算法在测试集上平均像素精度达 98.21%，平均交并比达 95.46%，分割效果显著优于其他模型。

图 7 展示了不同模型针对不同场景下的水面区域分割结果，图像中白色区域为水面区域，黑色区域为非水面区域。可以看出 FCN 和 SegNet 在进行水面区域识别时，容易出现将非水面区域误识别为水面区域的情况。使用 UNet 进行分割时，分割结果仍然存在孤立的误分割区域，在光照不良的场景下分割效果不佳。Deeplabv3+ 模型虽然可以较准确地识别出水面区域，但是在水面区域和非水面区域的边缘处分割效果不佳。本文改进算法在各个场景下的分割效果优于其他算法，针对边缘处的分割效果更加精细，能更准确地识别出水面区域。

**Table 1.** Comparison of the performance of different semantic segmentation models for water surface area recognition  
**表 1.** 不同语义分割模型的水面区域识别性能对比

模型名称	MPA (%)	mIoU (%)
FCN	72.81	50.80
SegNet	75.75	52.27
UNet	92.74	84.46
Deeplabv3+	95.29	89.75
本文算法	<b>98.21</b>	<b>95.46</b>



**Figure 7.** Comparison of different semantic segmentation algorithms for recognition of water areas in different scenes  
**图 7.** 不同语义分割算法针对不同场景水面区域识别效果对比

### 3.2.4. 改进效果分析

实验对比了在 Deeplabv3+模型的基础上，分别引入 CBAM 注意力模块，全连接条件随机场(CRF)，以及两个模块组合后模型的分割效果，采用平均像素精度和平均交并比指标，结果如表 2 所示。

从表 2 可以看出，相比于 Deeplabv3+模型，引入 CBAM 注意力模块后 MPA 提升了 2.36%，mIoU 提升了



4.44%；引入 CRF 后 MPA 提升了 2.65%，mIoU 提升了 5.37%；二者组合之后 MPA 提升了 2.94%，mIoU 提升了 5.71%。

图 8 展示了引入不同改进措施的水面区域识别效果。CBAM 模块的引入提升了图像整体的分割精度，CRF 使水面边缘处的分割更加精细。可以看出，CBAM 注意力模块和 CRF 均在不同程度上提高了水面区域识别的准确度。

Table 2. Comparison of the performance of different improvements for water surface area recognition

表 2. 不同改进措施的水面区域识别性能对比

	MPA (%)	mIoU (%)
Deeplabv3+	95.29	89.75
+CBAM	97.65	94.19
+CRF	97.94	95.12
+CBAM+CRF	98.21	95.46

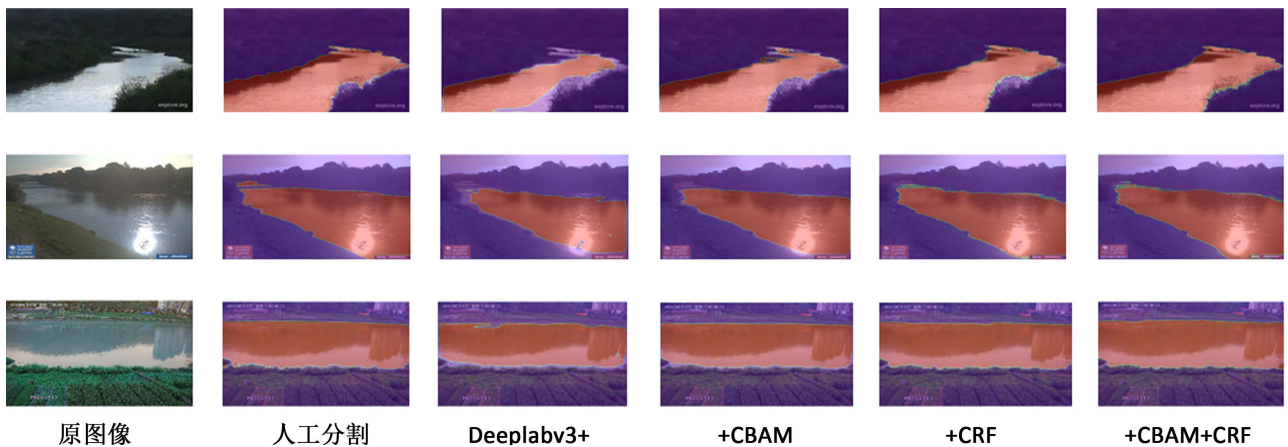


Figure 8. Comparison of different semantic segmentation algorithms for recognition of water areas in different scenes

图 8. 不同改进措施的水面区域识别效果对比

## 4. 水位智能识别方法验证与应用

为了证明本文提出的水位识别方法可以在无需识别水尺的情况下精确测量水位，在武汉大学灌溉排水与水环境综合试验场排水渠、贵阳(三)水文站以及某水库进行水位监测试验进行验证。

### 4.1. 试验方案

#### 4.1.1. 试验步骤

首先固定摄像头，结合水尺或其他能反映现场高程的参照物，采集若干像素点及其对应的实际高程，构建像素坐标与高程的对应关系。之后拍摄水位变化影像，人工观测水尺并记录当前时刻的水位数据。根据采集的水面视频或图像，在不同时刻利用本文算法进行水位识别，并根据《水位观测标准》(GB/T 50138-2010)对水位的自动监测设备的规定进行评价，并分析误差来源。

#### 4.1.2. 结果评定方法

根据《水位观测标准》(GB/T 50138-2010) [18]，自动水位监测设备在置信水平 95% 下的综合不确定度为 3 cm，系统误差为 ±1 cm。

系统误差的计算如公式(12)所示:

$$X'_y = \frac{\sum_{i=1}^N (P_{yi} - P_i)}{N} \tag{12}$$

式中:  $P_{yi}$  为图像识别水位结果,  $P_i$  为人工观测水位结果,  $N$  为观测次数。

随机不确定度的计算如公式(13)所示:

$$X''_y = 2\sqrt{\frac{\sum_{i=1}^N (P_{yi} - P_i - X'_y)^2}{N-1}} \tag{13}$$

综合不确定度的计算如公式(14)所示:

$$X = \sqrt{X'^2_y + X''^2_y} \tag{14}$$

### 4.2. 试验结果

三组试验中, 渠道试验视频时长 5 min, 每间隔 1 s 进行一次水位读数, 水位变化范围为 21~40 cm; 贵阳(三)水文站试验根据 2021 年 8 月 10 日至 2021 年 8 月 14 日连续 5 天的监控图像, 每日 8:00, 14:00, 20:00, 02:00 进行水位读数; 某水库试验选取 2020 年 10 月 1 日至 10 月 7 日的一次落水过程, 2020 年 10 月 16 日至 10 月 18 日的一次涨水过程进行水位识别, 水位变化范围为 1.30~1.84 m。

在运算效率上, 针对每帧图像, 原算法可在 2 s 内完成水位计算, 改进算法可在 5 s 内完成水位计算。在水位识别分辨力上, 当拍摄图像分辨率较高时, 如试验 1 和试验 2, 根据像素 - 高程对应关系构建“数字水尺”

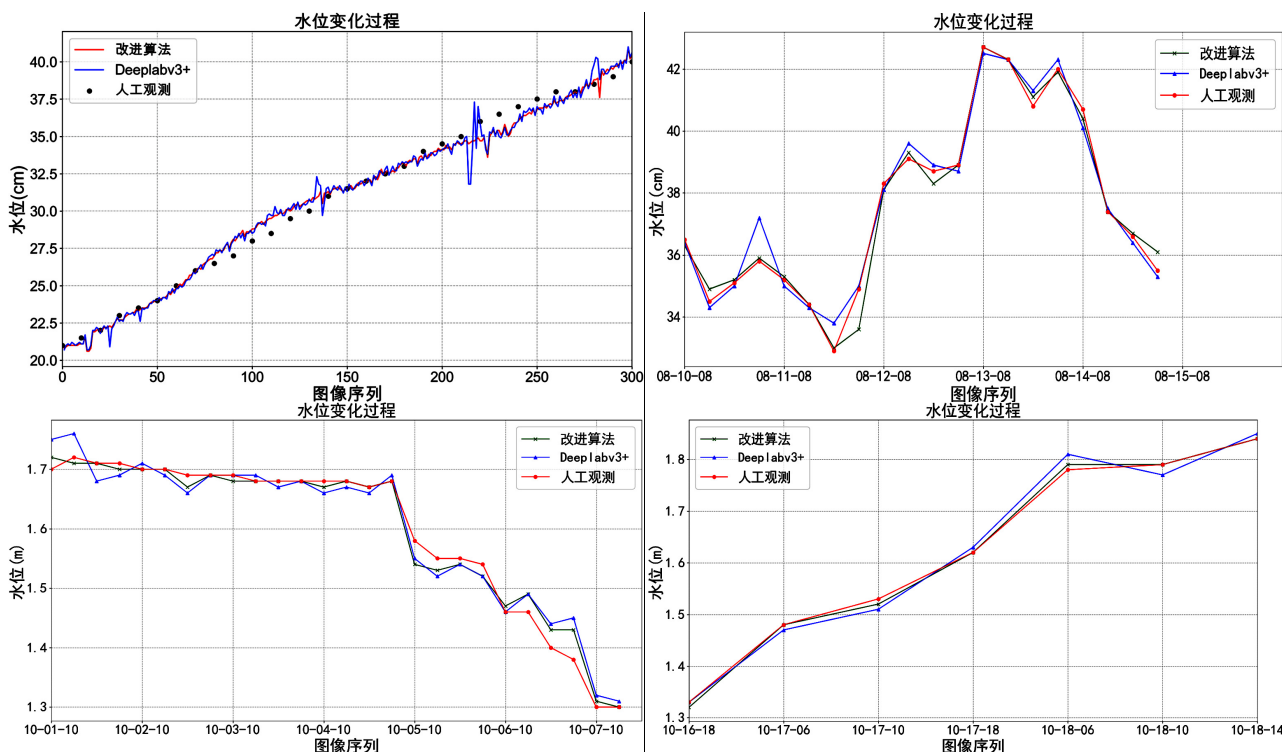


Figure 9. Measurement results of water level recognition test

图 9. 水位识别试验测量结果

的分辨力可达到 0.1 cm；当试验 3 图像分辨率较差的情况下，本方法识别水位的分辨力为 1 cm。

图 9 分别展示了三组试验的水位变化过程，可以直观看出本文算法能够较好地反映水位的变化，并且相比于原 Deeplabv3+算法，改进后的方法观测异常值更少，误差更小。

Table 3. Comparison of the water level recognition test results

表 3. 水位识别试验比测结果

	试验 1		试验 2		试验 3	
	Deeplabv3+	改进算法	Deeplabv3+	改进算法	Deeplabv3+	改进算法
最大误差	-1.500	-1.300	1.400	1.300	3.000	1.000
系统不确定度	0.027	-0.043	0.110	0.025	0.182	-0.030
随机不确定度	1.314	1.201	1.135	0.756	4.936	3.259
综合不确定度	1.314	1.202	1.140	0.757	4.939	3.259



Figure 10. Water level recognition error

图 10. 水位线识别误差



Figure 11. Screen changes caused by camera disturbance

图 11. 摄像机扰动引起的画面变化

表 3 给出了三组试验中原 Deeplabv3+模型和改进算法的比测结果。其中试验 1 和试验 2 符合标准中综合不确定度不超过 3 cm, 系统误差不超过 $\pm 1$  cm 的规定。而试验 3 符合标准中系统误差不超过 $\pm 1$  cm 的规定, 但是综合不确定度均超过 3 cm。经对比, 相比于原 Deeplabv3+算法, 改进算法的误差更小, 识别结果更精确稳定。

### 4.3. 误差分析

当摄像机固定之后, 图像的像素坐标与实际高程的关系也就随之确定, 水位线的识别精度决定了水位计算的精度。当水位线识别出现偏差时, 计算的水位也会出现误差。图 10 展示了试验 1 中某时刻水位线识别误差。

当相机发生扰动时, 预先标定的像素 - 高程对应关系也随之改变, 这时按照预先标定结果进行插值计算同样会造成误差。图 11 展示了试验 3 中由于摄像头扰动造成的图像变化。

## 5. 结论与展望

本文提出一种基于深度语义分割的图像水位监测方法, 在 DeepLabv3+模型嵌入 CBAM 注意力模块提升模型对关键特征的提取能力, 利用全连接条件随机场提升水体边缘的分割精度, 并构建“数字水尺”进行水位值的计算, 极大提升了图像法识别水位的精度和适用性。实验结果表明, 改进的算法在自建数据集上 MPA 达 98.21%, mIoU 达 95.46%, 远优于其他算法。经过三组水位监测试验, 改进算法精度优于原 Deeplabv3+算法, 平均误差不超过 1 cm, 可以适应昼夜环境, 准确监测水位变化情况。

由于训练水面分割网络的数据集规模较小, 且存在系统性标注误差, 使得分割网络识别水面区域边界处仍存在一定误差。后续可以收集更多场景下的水面图像, 扩大数据集规模, 进一步提升模型在不同环境下的泛化能力。针对水面边界处的分割误差, 本文采用了条件随机场作为边缘细化模块, 具有较多超参数, 调参较为复杂, 如何以更高效的手段进行边缘细化仍是值得思考的问题。此外, 本文模型的训练、测试与实地试验基本处于静态条件, 如何在复杂环境引起的水位剧烈波动情况下实现稳定的水位监测还有待针对性研究, 如进一步结合计算机视觉与摄影测量技术监测空间水位曲线, 提升环境稳定性。

## 基金项目

国家重点研发计划(2019YFC1510602)。

## 参考文献

- [1] 张应辉. 浅谈山区型河流水位计的选型[J]. 水利水文自动化, 2008(4): 45-46.  
ZHANG Yinghui. A brief discussion on model selection of water level gauge for mountain river. Water Resources Informatization, 2008(4): 45-46. (in Chinese)
- [2] 安全, 范瑞琪. 常用水位传感器的比较和选择[J]. 水利信息化, 2014(3): 52-54+60.  
AN Quan, FAN Ruiqi. Comparison and selection of common water level sensors. Water Resources Informatization, 2014(3): 52-54+60. (in Chinese)
- [3] YU, J., HAHN, H. Remote detection and monitoring of a water level using narrow band channel. Journal of Information Science and Engineering, 2010, 26(1): 71-82.
- [4] 张振, 周扬, 王慧斌, 高红民, 刘海韵. 标准双色水尺的图像法水位测量[J]. 仪器仪表学报, 2018, 39(9): 236-245.  
ZHANG Zhen, ZHOU Yang, WANG Huibin, GAO Hongmin and LIU Haiyun. Image-based water level measurement with standard bicolor staff gauge. Chinese Journal of Scientific Instrument, 2018, 39(9): 236-245. (in Chinese)
- [5] 周衡, 仲思东. 基于视频图像的水位监测方法研究[J]. 半导体光电, 2019, 40(3): 390-394+400.  
ZHOU Heng, ZHONG Sidong. Research on water level monitoring based on image processing. Semiconductor Optoelectronics, 2019, 40(3): 390-394+400. (in Chinese)
- [6] BAI, G. G., HOU, J. M., ZHANG, Y. W., *et al.* An intelligent water level monitoring method based on SSD algorithm. Measurement, 2021, 185: 110047. <https://doi.org/10.1016/j.measurement.2021.110047>
- [7] JAFARI, N. H., LI, X., CHEN, Q., *et al.* Real-time water level monitoring using live cameras and computer vision techniques. Computers & Geosciences, 2020, 147(5): 104642. <https://doi.org/10.1016/j.cageo.2020.104642>



- [8] CHEN, L. C., ZHU, Y., PAPANDREOU, G., *et al.* Encoder-Decoder with atrous separable convolution for Semantic image segmentation. In Ferrari, V., *et al.* (Eds.), Computer Vision—ECCV 2018. Berlin: Springer, 2018: 833-851. [https://doi.org/10.1007/978-3-030-01234-2\\_49](https://doi.org/10.1007/978-3-030-01234-2_49)
- [9] 王文冠, 沈建冰, 贾云得. 视觉注意力检测综述[J]. 软件学报, 2019, 30(2): 416-439.  
WANG Wenguan, SHEN Jianbing and JIA Yunde. Review of visual attention detection. Journal of Software, 2019, 30(2): 416-439. (in Chinese)
- [10] WOO, S., PARK, J., LEE, J. Y., *et al.* CBAM: Convolutional block attention module. In Ferrari, V., *et al.* (Eds.), Computer Vision—ECCV 2018. Berlin: Springer, 2018: 3-19. [https://doi.org/10.1007/978-3-030-01234-2\\_1](https://doi.org/10.1007/978-3-030-01234-2_1)
- [11] KRÄHENBÜHL, P., KOLTUN, V. Efficient inference in fully connected CRFs with gaussian edge potentials. Advances in Neural Information Processing Systems, 2011, 24: 109-117.
- [12] 甘亦忻. 基于条件随机场的端对端图像语义分割算法研究[D]: [硕士学位论文]. 西安: 西安电子科技大学, 2019.  
GAN Yixin. Research on the end-to-end semantic image segmentation algorithm based on the conditional random field model. Master's Thesis, Xi'an: Xidian University, 2019. (in Chinese)
- [13] LIANG, Y., JAFARI, N., LUO, X., *et al.* WaterNet: An adaptive matching pipeline for segmenting water with volatile appearance. Computational Visual Media, 2020, 6(1): 65-78. <https://doi.org/10.1007/s41095-020-0156-x>
- [14] CHENG, Y., JIANG, M., ZHU, J., *et al.* Are we ready for unmanned surface vehicles in inland waterways? The USVInland-Multisensor dataset and benchmark. IEEE Robotics and Automation Letters, 2021, 6(2): 3964-3970. <https://doi.org/10.1109/LRA.2021.3067271>
- [15] BADRINARAYANAN, V., KENDALL, A. and CIPOLLA, R. SegNet: A deep convolutional encoder-decoder architecture for Image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481-2495. <https://doi.org/10.1109/TPAMI.2016.2644615>
- [16] RONNEBERGER, O., FISCHER, P. and BROX, T. U-Net: Convolutional networks for biomedical image segmentation. In Navab, N., *et al.* (Eds.), Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015. Berlin: Springer, 2015: 234-241. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
- [17] LONG, J., SHELHAMER, E. and DARRELL, T. Fully convolutional networks for semantic segmentation. In 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Manhattan: The Institute of Electrical and Electronics Engineers, 2015: 3431-3440. <https://doi.org/10.1109/CVPR.2015.7298965>
- [18] 水利部长江水利委员会水文局. GB/T 50138-2010. 水位观测标准[S]. 北京: 中国计划出版社, 2010.  
Bureau of Hydrology of Changjiang Water Resources Commission of the Ministry of Water Resources. GB/T 50138-2010. Standard for stage observation. Beijing: China Planning Press, 2010. (in Chinese)