

# 结合E-YOLO和水体指数约束的大面幅影像水利设施检测

许泽宇<sup>1,2</sup>, 沈占锋<sup>1,3</sup>, 李杨<sup>1,4</sup>, 李均力<sup>5,6</sup>, 王浩宇<sup>1,2</sup>, 李硕<sup>1,4</sup>,  
焦淑慧<sup>1,2</sup>, 李苓苓<sup>7</sup>

1. 中国科学院空天信息创新研究院 国家遥感应用工程技术研究中心, 北京 100101;
2. 中国科学院大学 电子电气与通信工程学院, 北京 100049;
3. 中国科学院大学 资源与环境学院, 北京 100049;
4. 中国科学院空天信息创新研究院 遥感科学国家重点实验室, 北京 100101;
5. 中国科学院新疆生态与地理研究所 绿洲与荒漠国家重点实验室, 乌鲁木齐 830011;
6. 新疆遥感与地理信息系统应用重点实验室, 乌鲁木齐 830011;
7. 应急管理部国家减灾中心, 北京 100124

**摘要:** 水利设施对水资源与水量调度、自然湿地生态保护与修复、资源和生态功能的利用及经济效益发展有重要作用。传统方法统计水利设施位置、数量等依赖于汇编资料, 存在耗时长、资料更新不够及时以及具体地理位置不详等缺点, 遥感为大规模监测水利设施提供了新的可能。本文以YOLO v3网络为基础, 结合水利设施的特点, 提出了一种基于大面幅影像快速检测水利设施的算法, 主要分为两个方面: (1) 改进的YOLO算法(E-YOLO)。E-YOLO提出PPA特征融合方法和等比预测框与四特征图交叉预测方法, 对小样本等问题进行优化; 改进损失函数, 突出置信度损失; 同时使用迁移学习的方法, 读取特征提取部分的预训练模型参数。(2) 基于E-YOLO算法和水体指数约束的大面幅水利设施检测算法。通过水体指数约束滑动步幅来解决影像面幅大、目标尺度小的问题, 同时降低漏检率和误检率, 再结合轮廓合并方法, 优化检测结果。本研究中采用高分二号影像数据实现大面幅影像水利设施检测, 实验结果表明: E-YOLO算法可以明显提高水利设施检测效果, 相比平均F2精度相比YOLO v3提高了1.25%。且有更好的稳定性; 水体指数约束的大面幅检测方法可以在保证效率的情况下提高检测精度, 其F2精度相比大步幅和小步幅方法分别提高了3.72%和2.70%, 为遥感水利设施检测提供了良好方案。

**关键词:** 水利设施, 遥感检测, E-YOLO, 大面幅影像, NDWI

**引用格式:** 许泽宇, 沈占锋, 李杨, 李均力, 王浩宇, 李硕, 焦淑慧, 李苓苓. 2022. 结合E-YOLO和水体指数约束的大面幅影像水利设施检测. 遥感学报, 26(10): 2083-2093

Xu Z Y, Shen Z F, Li Y, Li J L, Wang H Y, Li S, Jiao S H and Li L L. 2022. Detection of water conservancy facilities in large-format image combining E-YOLO algorithm and NDWI constraint. National Remote Sensing Bulletin, 26(10): 2083-2093 [DOI: 10.11834/jrs.20200124]

## 1 引言

水利基础空间要素是自然与人工构造物的有机集成, 主要包括流域自然特征要素和流域水利工程分布要素, 其中水利工程分布要素包含水库、堤防、水闸、渠道等(许莹等, 2016), 本文中研

究的水利设施主要指水利工程枢纽、水闸、渠首, 不包括水库、堤防。水利设施在发电、灌溉、洪灾防控等方面发挥着重要的作用(Liu等, 2013), 传统方法统计水利设施位置、数量等数据依赖于汇编资料, 耗时长且存在资料更新不够及时、具体地理位置不详等缺点, 而遥感方法可以对水利

收稿日期: 2020-05-22; 预印本: 2020-08-19

基金项目: 国家重点研发计划(编号: 2021YFC1523503); 国家自然科学基金(编号: 41971375, 41671034); 新疆第三次科学考察专项(编号: 2021xjkk1403); 新疆维吾尔自治区重点研发任务专项(编号: 2022B03001-3)

第一作者简介: 许泽宇, 研究方向为遥感信息计算。E-mail: xuzeyu@aircas.ac.cn

通信作者简介: 沈占锋, 研究方向为遥感信息计算。E-mail: shenzf@aircas.ac.cn

设施进行快速检测, 包括无法获得统计资料的区域, 在水利设施的定期评估等方面有重要应用价值。目前, 利用遥感影像对河湖等自然要素的研究很多(沈占锋等, 2013; 刘康等, 2019; 王宁等, 2020), 但对于水利设施的研究较少(许莹等, 2016)。Li等(2020)在构建大规模遥感影像分类基准RSI-CB时, 分类结果中虽包含了水利设施, 但仅为简单分类, 且未对水利设施的具体提取方法进行分析; 许莹等(2016)通过建立规则集对水利空间要素中的自然要素和人工要素都进行了提取, 提取结果中河流、湖泊等提取精度较高, 而水闸等水利设施提取精度较低。

基于深度学习的目标检测算法, 有很高的检测精度, 经典的目标检测算法有Faster R-CNN (Ren等, 2017)、SSD (Liu等, 2016)、YOLO (Redmon等, 2016; Redmon和Farhadi, 2017和2018)等, 其中YOLO算法检测精度高、时间快, 在遥感影像目标检测中具有很好的效果, 尤其是在舰船、飞机等地物的检测应用最广泛(Hoeser等, 2020), 在水利设施检测中研究还不多。针对舰船目标检测, 马啸等(2019)通过减少模型层数和参数对YOLO网络模型进行了改进, 获得了轻量级的网络结构, 有较好的综合性能; 王玺坤等(2020)基于YOLO v3-Tiny算法, 增加了特征映射模块, 并引入残差网络, 改进后的算法优于SSD和YOLO v3算法; 针对飞机检测, 吴杰等(2021)基于YOLO v3-Tiny算法, 通过分组卷积和增加尺度检测的方法, 提高了低质量影像中飞机检测的效果; 郑志强等(2019)将YOLO v3网络中的两个残差网络模块替换为两个密集网络模块, 提高了遥感影像中飞机检测的准确率。这些优化方案体现了在遥感目标检测中YOLO算法的有效性。

遥感影像面幅大, 待检测地物相对于影像尺度一般很小, 直接进行检测效率较低(陶泽远, 2021)。张作省(2018)通过机场检测算法, 缩小检测视野、锁定检测范围, 以减少计算复杂。目前研究多是以飞机检测为例, 其聚集地一般是机场、跑道、停机坪等, 均为块状分布, 而水利设施位置常与河流、湖泊相关, 为线状或块状分布; 另外, 水利设施检测对精度的需求高于对速度的需求, 由于遥感影像的复杂性, 仅在候选区检测可能存在漏检问题。因此其候选区域的选择方案并不适用于水利设施检测。

本文基于YOLO v3算法提出了遥感影像水利设施检测的方法E-YOLO。E-YOLO使用PPA特征融合方法进行低层和高层特征的融合, 使用等比预测框和四特征图交叉方法进行预测, 并结合损失函数优化和迁移学习, 提高在小样本水利设施数据上的检测精度。在此基础上, 提出基于水体指数约束的大面幅水利设施检测方法, 通过水体指数约束滑动窗口步幅, 同时降低漏检率和误检率, 得到了更优的检测效果。

## 2 方法原理

本文算法主要研究两个方面: 一是针对遥感影像中水利设施检测的特点, 改进深度学习算法; 二是以改进后的深度学习算法为核心, 针对影像和水利设施尺寸差异较大的问题, 提出大面幅影像水利设施提取算法。

### 2.1 水利设施检测算法

#### 2.1.1 YOLO算法

经典的目标检测网络R-CNN分为两步, 首先生成可能存在目标的区域, 再对区域进行识别分类(Girshick等, 2014)。Faster R-CNN通过将两个步骤作为一个网络的两个分支, 缩短运算时间, 而YOLO系列算法则直接使用一个网络, 同时得到目标的位置和类别信息, 很好的提高了算法效率(Redmon等, 2016)。本文研究在YOLO v3算法的基础上进行改进, YOLO v3使用了多尺度预测的方法, 即使用3个不同尺度的特征图, 以不同尺度的特征图检测不同尺度的物体, 从而在多种尺度上都能得到比较精细准确的检测效果(Redmon和Farhadi, 2018)。YOLO v3使用Darknet-53网络实现特征提取, 相比较YOLO v2中的Darknet-19网络添加了残差单元, 达到了优于ResNet的精度, 也提高了目标检测的效果(He等, 2016)。

#### 2.1.2 E-YOLO算法

本文提出的E-YOLO算法是在YOLO v3基础上的增强型YOLO算法(Enhanced YOLO), 记为“E-YOLO”算法。主要改进如下:

(1) 提出PPA (PixelShuffle with Point-wise Addition) 特征融合方法, 由此减少原模型的多层卷积层, 比原方法更精简。

(2) 采用等比预测框与四特征图交叉预测的

方法。由于样本量少，初始预测框大小通过等比推断确定，而不使用k-means方法，提高鲁棒性。

(3) 改进损失函数，突出置信度损失，增大检测准确性。

(4) 采用部分迁移学习的方法，以同时满足小样本训练和网络结构改进需求。由于样本量少，保留原DarkNet-53特征提取部分结构，训练时读取预训练模型进行初始化。其它部分由于结构改进，不读取预训练模型。

本文提出PPA融合方法，其流程如图1所示。首先将高层特征通过1×1的卷积块提升维度，再通过PixelShuffle算法超分辨率上采样，上采样后通道数与待融合的底层特征通道数相同，因此直接使用“point-wise addition”逐点相加的方法进行融合。通过该方法融合后，特征图不需要进行过多的卷积操作，因此原YOLO v3算法中每次融合后的5次卷积均缩减为1次，结构更精简高效。

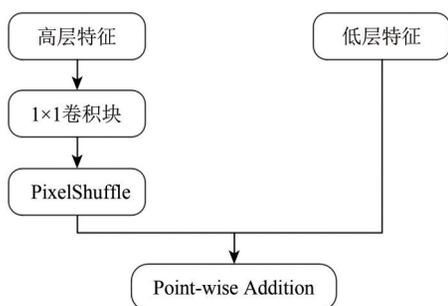


图1 PPA模块结构图

Fig. 1 Structure diagram of PPA

PPA模块中PixelShuffle算法最早用于超分辨率重建(Shi等, 2016)，与双线性插值法不同，该算法通过子像素卷积(sub-pixel convolution)实现特征图的放大(图2)，其上采样过程可训练，以实现更好的效果。在PixelShuffle上采样方法中，图像通道数缩减，本文设置中，通道数缩减为原来的1/4，图像行、列放大2倍。在PixelShuffle上采样前，通过1×1卷积将高层特征图通道数提升为底层特征图通道数的4倍，这样通过PixelShuffle后的高层特征图与底层特征图的通道数相同，且放大后的特征图尺寸也与对应底层特征图相同，因此直接与底层特征图逐点相加融合，在高效的同时保留了更多的信息。

由于样本数据较少，使用K-Means算法获得的预测框大小偶然性大，因此本文设计了等比预测框与四特征图交叉预测的方法。由于输出特征

图大小成等比关系(2倍)，在预测框设定时，也以等比方法为主体，公比设为2，根据等比方法确定主体预测框大小，进而计算其他预测框大小，计算公式为：

$$la_n = \begin{cases} (2^{(n-3)} + 2^{(n-1)/2}) \cdot k, & n \text{ 为奇数} \\ 2^{n/2} \cdot k, & n \text{ 为偶数} \end{cases} \quad (1)$$

式中， $n=1, 2, 3, \dots, 9$ ，表示预测框的序号， $k$ 为给定的初始框大小，经多次实验，本文中取7为最佳值。由式(1)得到9个预测框大小 $la_1$ 到 $la_9$ ，将其用于4个不同尺度的特征图的预测中，如图3所示。其中 $b_1-b_4$ 为输出特征图，不同大小的预测框与不同尺度的输出特征图相对应。特征图对应预测框时存在部分重叠，以满足更复杂情况的检测需求。

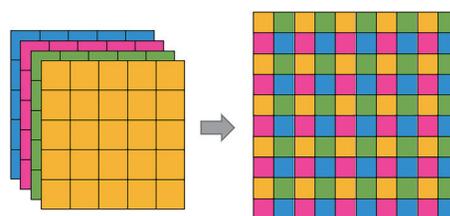


图2 子像素卷积算法示意图

Fig. 2 Schematic diagram of sub-pixel convolution

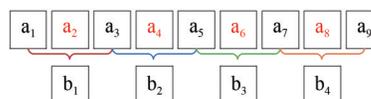


图3 特征图与预测框关系图

Fig. 3 Relationship between feature map and anchors

YOLO v3中损失函数由3部分组成，分别用来评估位置形状、置信度和类别，由于遥感水利设施检测只有一类物体，因此本文取消了类别损失。在水利设施检测实际应用中，由于会对检测结果进行后续整理，因此置信度比位置准确度更为重要。本研究中增加了置信度损失权重。最终改进后的损失函数表示为：

$$loss = loss_{giou} + \alpha loss_{obj} \quad (2)$$

式中， $loss_{giou}$ 表示位置和形状损失， $\alpha loss_{obj}$ 表示置信度，通过参数 $\alpha$ 调节位置和置信度损失权重，经多次实验，本文中取 $\alpha = 3$ 为最佳值。

由于样本数据较少，采用部分迁移学习的方法。特征提取部分DarkNet-53结构基本不变，仅调整了各层输入尺寸，因此在这部分读取预训练模型参数。而在特征融合和YOLO层检测部分，对

网络结构进行了较大改进调整, 不读取预训练模型参数。

整体上 E-YOLO 算法与原 YOLO v3 算法相比, 网络结构由 222 层缩减到 209 层, 更为精简, 其网

络结构见图 4。其中样本子区域大小为  $1000 \times 1000$  像素, 为了减少 GPU 资源消耗, 在输入影像时, 将大小调整为  $576 \times 576$  像素, 比子区域略小。

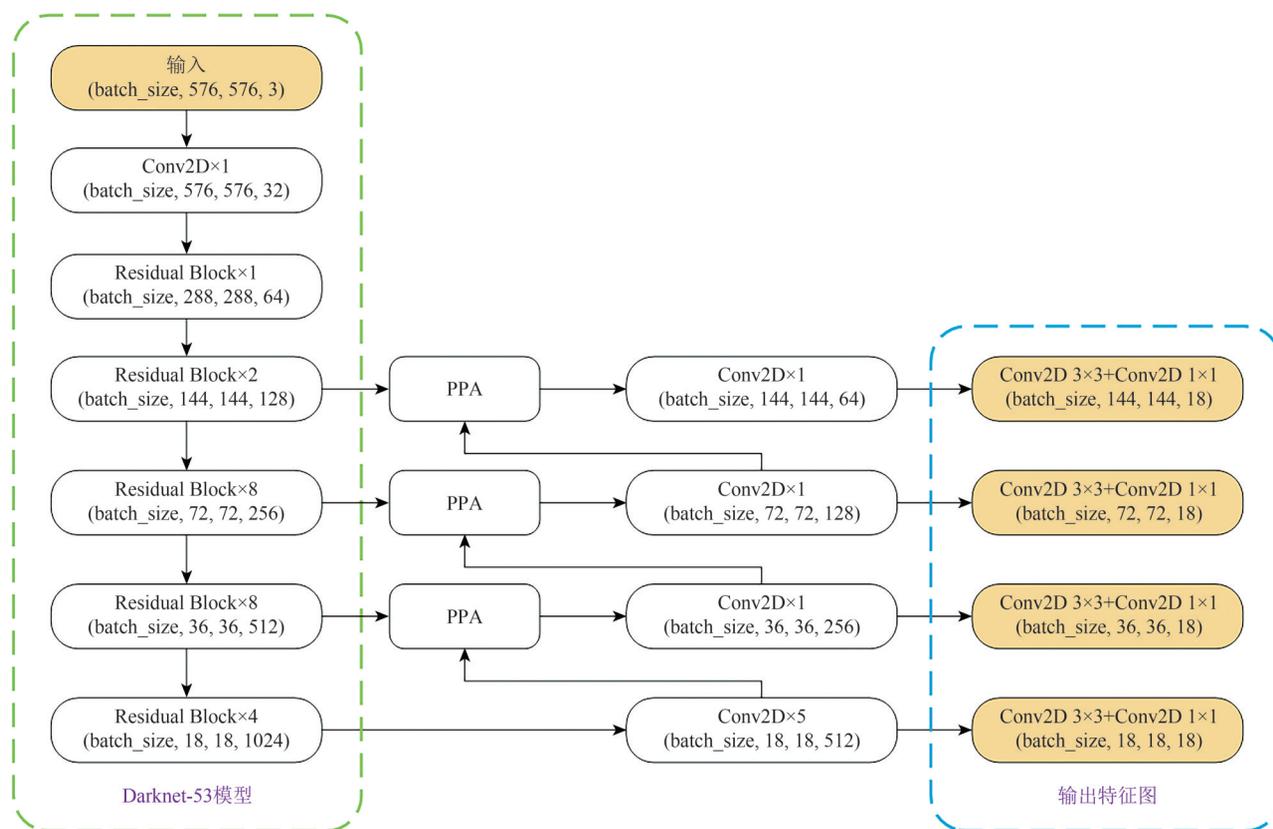


图4 E-YOLO 结构图

Fig. 4 Structure diagram of E-YOLO

## 2.2 大面幅遥感影像检测

针对水利设施的特点, 结合 E-YOLO 算法的优势, 本文提出大面幅遥感影像水利设施检测方案。由于大面幅影像的无法直接进入卷积神经网络 (Audebert 等, 2018), 因此需要每次读取子区域进行检测, 通常通过滑动遍历的方法, 完成对整景影像的检测。常用的大面幅遥感影像检测方案是先确定候选区域, 然后在候选区域中检测, 针对水利设施检测的特殊性, 本文进行以下讨论:

(1) 与飞机所在的机场等不同, 水利设施所在的区域 (如河、湖、人工渠等) 情况复杂: 既有人工区域, 又有自然区域; 既有线状区域, 又有面状区域; 既有连续区域, 又有间断区域。

(2) 研究区存在季节性水域, 且水坝等水利设施位于水域边缘, 先检测水域易出现漏检误差。

因此, 本文研究中均进行整景影像遍历。在

使用深度学习方法进行整体遍历时, 通常会遇到地物被不同子区域切割等边界误差问题。常见的解决方案是通过减小滑动步幅, 使不同子区域间存在重叠, 这类方法在城市密集地物提取中起到了很好的效果 (Audebert 等, 2018), 由于水利设施是稀疏地物, 在影像中范围占比较小, 全影像重叠检测会消耗大量资源, 且在明显没有水利设施的区域过多检测会提高误检概率, 影响检测精度。因此本文提出水体指数约束步幅的检测方法, 实现在水域集中检测, 非水域快速检测, 以同时降低漏检率和误检率, 检测流程如图 5。

滑动遍历时滑动窗口子区域大小为  $1000 \times 1000$ , 每个子区域首先进行筛选判断, 如果该区域值全为 0 或全为同样数值, 则不进行检测。在此基础上, 计算子区域的水体指数 NDWI (Normalized Difference Water Index) (McFeeters, 1996), 公式如下:

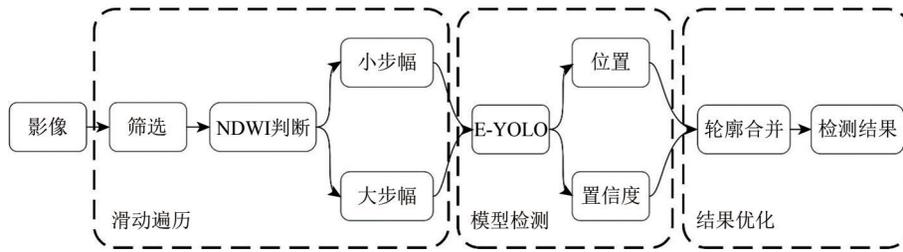


图5 大篇幅水利设施检测流程图

Fig. 5 Flow diagram of large-scale water conservancy facilities

$$NDWI = \frac{GREEN - NIR}{GREEN + NIR} \quad (3)$$

式(3)中, GREEN表示绿波段值, NIR表示近红外波段值, 水体指数值越大则该位置像是水体的概率越大, 该指数在水体提取中有显著效果(骆剑承等, 2009; 洪亮等, 2019)。根据NDWI得到判断函数:

$$stride = \begin{cases} S_{small}, & (num(NDWI > w_0) > n_0) \\ S_{big}, & \text{其他} \end{cases} \quad (4)$$

式中,  $S_{small}$ 取750像素,  $S_{big}$ 取1000像素, NDWI阈值 $w_0$ 取0,  $num(NDWI > w_0)$ 表示当前子区域中水体指数大于阈值的像素个数, 如果个数大于阈值 $n_0$ , 则步幅为 $S_{small}$ , 如果个数小于阈值 $n_0$ , 则步幅为 $S_{big}$ 。因此, 当该子区域存在水体的可能性大时, 该区域附近检测步幅缩短, 相邻子区域部分重叠, 从而在水域附近检测密集, 其他地方检测稀疏, 增大水体附近检测成功率, 且在其他区域保持高速率。本文实验中, 取 $w_0 = 0$ ,  $n_0 = 100$ 。子区域检测完成后, 根据E-YOLO算法得到的预测位置生成矢量文件, 将置信度写入矢量文件的属性中, 方便后续调整阈值以及检查整理。

轮廓合并是指有重叠检测结果的区域进行合并。传统目标检测中也有重叠现象, 通常采用非极大值抑制(NMS)方法处理(Neubeck和van Gool, 2006), 即通过IOU判断重叠度, 当重叠度超过阈值, 去掉置信度低的矩形框。但在此处重叠现象是由于滑动遍历产生, 即同一地物被检测多次的情况, 且存在同一地物被分在不同检测框检测, 因此, 此处不采用NMS方法去掉矩形框, 而是将检测得到的所有轮廓合并, 得到最终的结果, 以便得到更完整的检测结果(图6)。

### 3 实验结果与分析

#### 3.1 样本区与数据集

本文以新疆塔里木河流域为样本区, 塔里木

河流域周围是天山南坡-昆仑山-阿尔金山等高原山区, 中间是塔里木盆地, 河流水系有140多条河流组成, 均发源于盆地周边高原山区, 河流呈向心分布汇入盆地(周莉荔, 2010)。2018年6月, 研究团队在新疆塔河流域进行考察, 结合高分GF-1/GF-2数据, 对塔河流域关键水利枢纽、水库及生态闸坝进行了遥感制图及信息建库, 按照阿克苏河、叶尔羌河、塔河干流、开孔水系4个部分调研, 对塔河关键水利设施进行较为全面的覆盖。图7为研究区及典型水利设施示意图。



图6 轮廓合并示例

Fig. 6 A example of the contour merge

图7中, 主要水域指主要水系、二级线状河流、季节性面状河流、非季节性面状河流、湖泊和水库。其中喀群渠首、和田河交叉工程和开都河第一分水枢纽为融合后高分二号影像。根据实地调查结果及在该区域获取的高分二号影像数据, 构建水利设施数据集, 其空间分辨率为0.8 m。由于水利设施本身较少, 且需要保留部分整体影像测试样本, 对训练集进行旋转、镜像处理, 以扩充数据集, 使用交叉检验方式, 即轮换训练样本与测试样本, 每组540张影像为训练样本, 180张影像为测试样本, 共设4组进行轮换。

#### 3.2 实验环境与评价标准

本文实验环境为Windows 10系统, GTX 1060 6G显卡, 使用Pytorch框架, 编程语言为Python。采用精确率(Precision)、召回率(Recall)和综合精度分数作为结果的评价指标, 以样本标签为实

际值, 网络模型输出结果为预测值; 针对某一地物, 属于该地物为正类, 不属于为负类, 实际和预测均为正类的目标标记为  $TP$ , 实际为负类, 预测为正类的目标标记为  $FP$ , 实际为正类, 预测为负类

的目标标记为  $FN$ , 得到准确率和召回率为:

$$Precision = \frac{TP}{TP + FP} \quad (5)$$

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

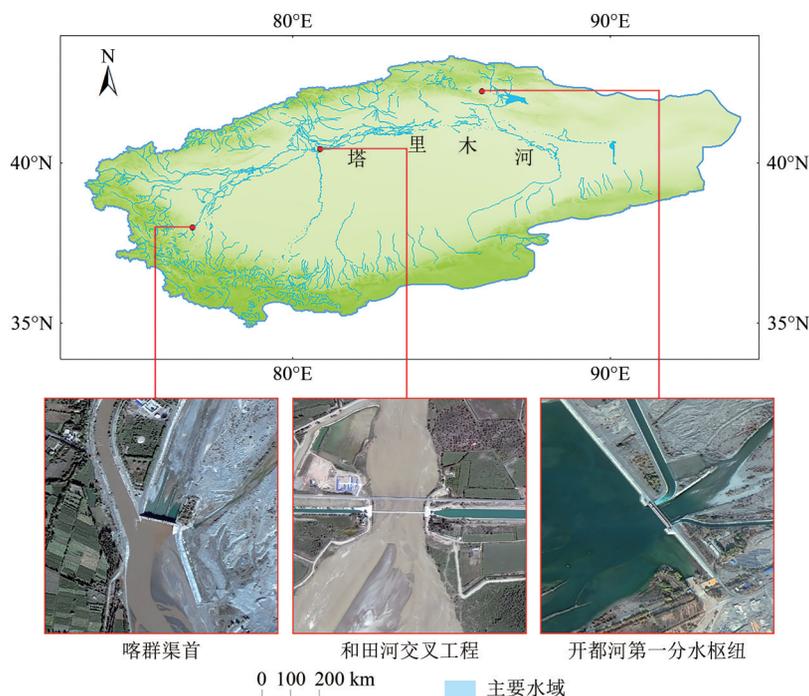


图7 塔里木河流域及典型水利设施示意图

Fig. 7 Tarim River basin and typical water conservancy facilities

在此基础上, 得到  $F$  分数:

$$F = \frac{(1 + \alpha^2)Precision \cdot Recall}{\alpha^2Precision + Recall} \quad (7)$$

由式 (7) 可知, 精确率 Precision 表示检测正确的目标占有所有预测结果的比值, 精确率高则误检率低; 召回率 Recall 表示检测正确的目标占有所有正确水利设施区域的比值, 召回率高则漏检率低。当  $F$  分数中  $\alpha = 1$  时, 召回率和精确度占有同样比重, 此时为  $F1$  分数, 当  $F$  分数中  $\alpha = 2$  时, 更偏重召回率, 此时为  $F2$  分数 (方浩文等, 2019)。由于水利设施检测更注重查全率, 为减少漏分现象, 本文中以  $F2$  分数作为模型效果的最终精度评判。由于使用交叉检验的方法, 在算法对比时, 记录  $F2$  分数的最大值、最小值和平均值, 且计算交叉检验中各算法  $F2$  分数的标准差, 以表示多个  $F2$  精度的分散程度, 用来评估算法的稳定性。

### 3.3 E-YOLO 算法结果对比

研究了原始 YOLO v3 算法、使用 PPA 特征融合方式的 YOLO v3 算法 (记为 YOLO v3 PPA) 和 E-YOLO 算法, 并与目标检测中的经典算法 SSD 进

行了对比, 4 种算法均使用了部分迁移学习的方法, 均训练 50 epoch, 得到实验结果如表 1。

表 1 多种算法水利设施检测结果\*

Table 1 Detection results of various algorithms for hydraulic facilities

算法	Min_F2	Max_F2	mF2	SDF2
SSD	0.726	0.814	0.769	0.034
YOLO v3	0.844	0.899	0.879	0.022
YOLO v3 PPA	0.848	<b>0.922</b>	0.884	0.027
E-YOLO	<b>0.871</b>	0.916	<b>0.890</b>	<b>0.017</b>

注: \*加粗为每列最优值。

表 1 中为交叉检验试验中各类精度值, Min\_F2 表示多次实验中  $F2$  精度最低值, Max\_F2 表示多次实验中  $F2$  精度最高值, mF2 表示多次实验中  $F2$  精度平均值, SDF2 表示多次实验中  $F2$  精度标准差 (Standard Deviation of F2)。由表 1 可知, SSD 算法各项精度弱于 YOLO 系列算法。YOLO v3 PPA 算法, 即在 YOLO v3 基础上使用 PPA 融合方式的算法, 在多次实验中有精度最高的表现, 其  $F2$  最大值达到了 0.922, 但其最低精度为 0.848, 相差较

大。同时根据标准差可看出, 其算法稳定性较差, 最终平均F2得分也不高。而使用了等比预测框与四特征图交叉预测的E-YOLO算法, 则体现了更稳定的精度, SDF2仅为0.017, 在精度稳定的同时, 其F2精度平均值也是各算法中的最优值, 相比YOLO v3提高了1.25%。

图8为各种算法的检测效果示意图, 其中第一行水利设施点仅YOLO v3 PPA和E-YOLO算法正

确检测到; 第二行中仅在影像右上角存在为水利设施, 除E-YOLO算法外, 其他算法均有多余检测, 即误检现象; 第三行YOLO v3和YOLO v3 PPA没有检测出水利设施, 虽然SSD算法检测到了水利设施, 但E-YOLO算法的检测效果更好。综合来看, E-YOLO算法能比其他算法更有效地对水利设施进行检测。

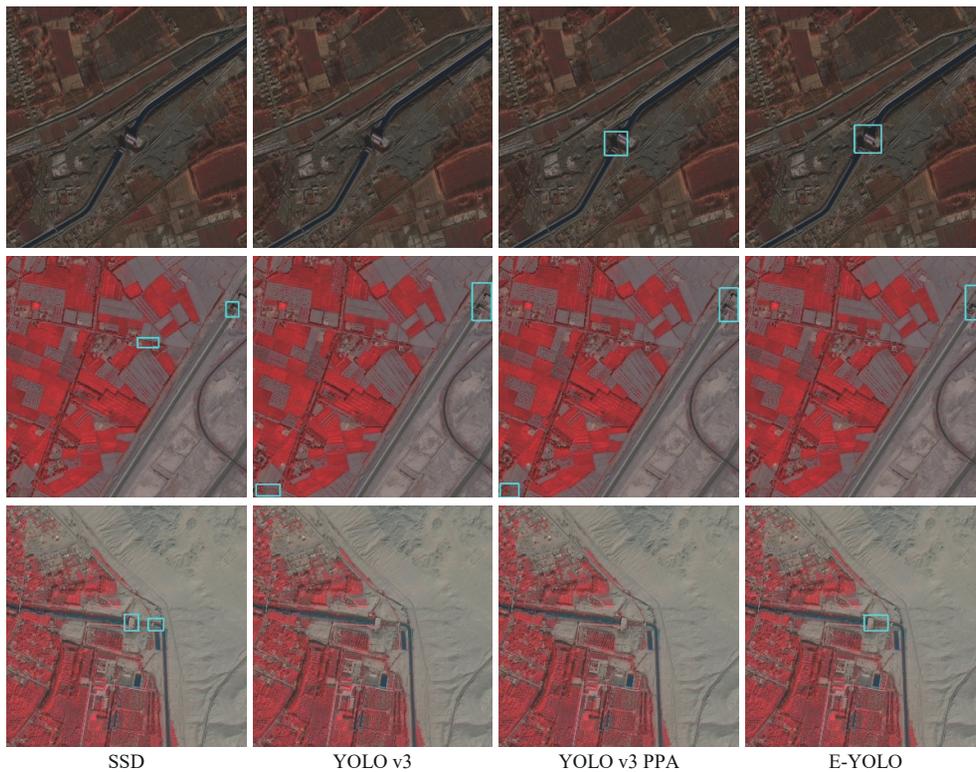


图8 塔里木河流域及典型水利设施示意图

Fig. 8 Tarim River basin and typical water conservancy facilities

### 3.4 大面幅检测方案对比

根据本文提出的水体指数约束的大面幅检测方案进行实验分析, 使用大面幅高分二号影像进行测试, 结果见表2。

表2 大面幅影像检测结果

Table 2 Large format image detection results

步幅	召回率	精确率	F1	F2	时间/s
1000 px	0.837	<b>0.706</b>	0.766	0.807	<b>224.77</b>
750 px	<b>0.884</b>	0.623	0.731	0.815	376.12
NDWI约束	<b>0.884</b>	0.691	<b>0.776</b>	<b>0.837</b>	305.98

注: 加粗为每列最优值。

由表2可知, 步幅为1000像素的方案中, 精度最高, 用时最短, 但其召回率和F2分数较低,

由于检测没有重叠区域, 水利设施处检测次数少、存在边缘漏检情况, 使整体漏检数偏多, 而步幅为750像素的方案中, 召回率明显提升, 但由于在没有水利的区域检测次数多, 其误检数偏多, 精确度较低, 因此最终的F2分数也不高, 且其检测时间很长。根据水体指数调整步幅的方案, 可以实现在水域集中重叠检测和非水域少量快速检测, 同时减少了漏检和误检现象, 其召回率与步幅为750像素的方案相近, 漏检少, 其精确率与步幅为1000像素的方案相近, 误检少, 因此水体指数约束方法的F2综合精度分数是3种方案中的最高值, 相比大步幅和小步幅方法分别提高了3.72%和2.70%。

图9展示了不同大面幅检测方案结果的对比效

果, 由于影像面幅大且水利设施稀疏, 水利设施以点表示, 左下角为细节放大图。在图9中, 第一行放大区域步幅 1000 像素的方案存在漏检, 而步幅 750 像素和本文方案由于在水域检测密集, 避免了漏检现象; 第二行, 步幅 750 像素的方案存在明显的误检现象, 通过放大图可以看出, 将其他建筑错误检测为了水利设施, 而本文提出的方法在

非水体区域步幅较大, 检测不密集, 降低了类似误检情况的发生率; 第三行, 步幅为 1000 像素的算法则提取了一处水利设施的两个小区域, 而步幅为 750 像素和水体指数约束算法, 得到了更完整的检测结果。综合对比可以看出, 使用水体指数约束的方法同时避免了漏检和误检现象, 且检测效果更完整。

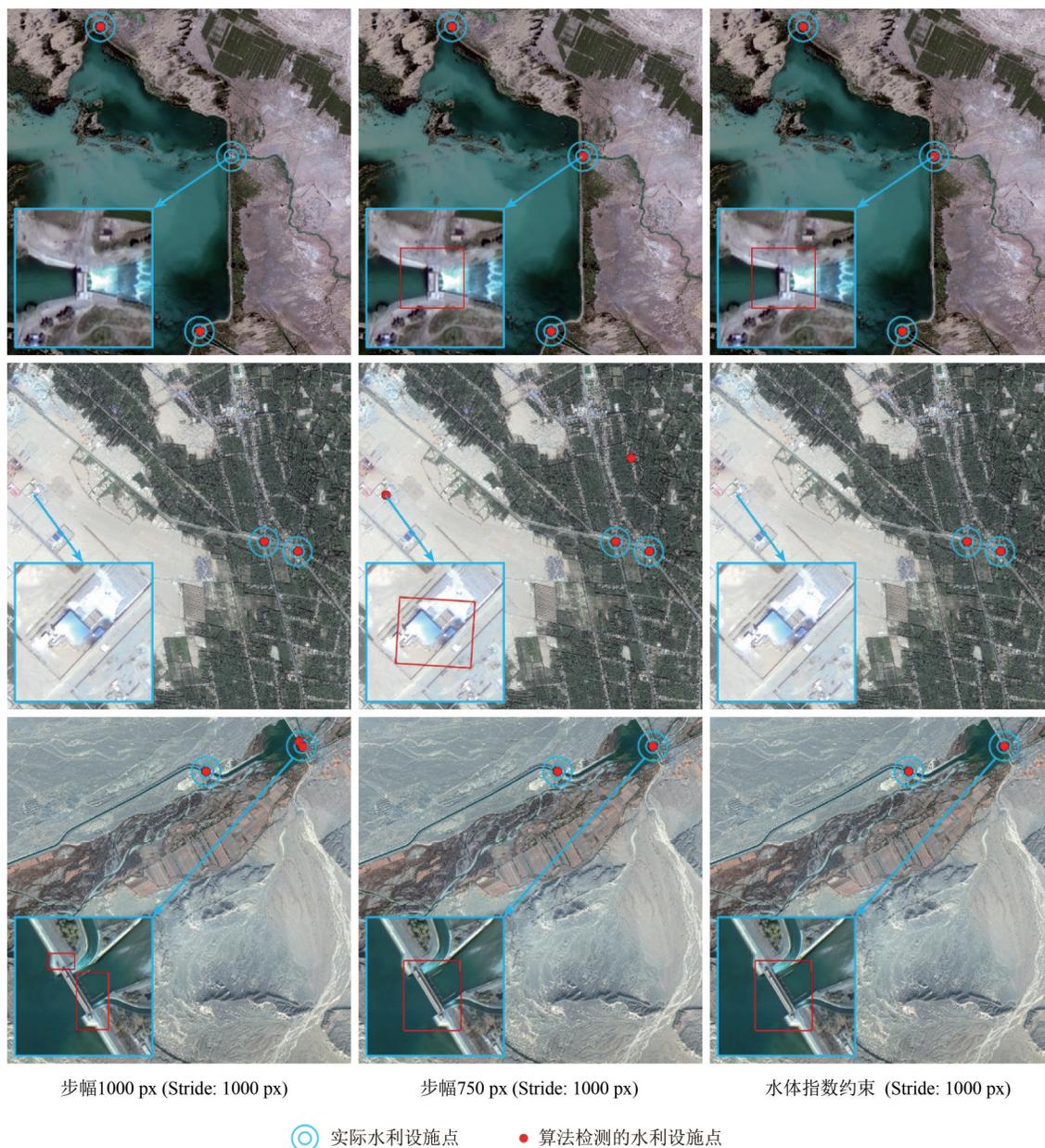


图9 不同大面幅检测方案结果对比图

Fig. 9 Comparison of the results of different large-format image detection methods

## 4 结论

本文针对遥感影像中的水利设施检测, 以新疆塔里木河流域为研究区, 以 YOLO v3 算法为基

础, 提出了更有效的水利设施检测方案。在检测算法方面, 提出改进后的 E-YOLO 算法, 提出 PPA 特征融合方法和等比预测框与四特征图交叉预测方法, 结构更加精简高效, 且优化了小样本

问题;改进损失函数,突出置信度损失,使水利设施检测更准确;同时使用迁移学习的方法,读取特征提取部分的预训练模型参数。改进后的E-YOLO算法平均F2精度相比YOLO v3提高了1.25%,且稳定性更强。在大面幅影像目标检测方面,提出水体指数约束步幅和轮廓合并的方法,同时降低了漏检率和错检率,其F2精度相比大步幅和小步幅方法分别提高了3.72%和2.70%,达到了最优的精度。本文方法在大面幅高分辨率影像中对水利设施有很好的提取效果,但在中低分辨率影像中的水利设施尺寸较小,且形状、纹理等特征信息不明显,其自动提取方法及效果有待进一步研究。

### 参考文献(References)

Audebert N, Le Saux B and Lefèvre S. 2018. Beyond RGB: very high resolution urban remote sensing with multimodal deep networks. *ISPRS Journal of Photogrammetry and Remote Sensing*, 140: 20-32 [DOI: 10.1016/j.isprsjprs.2017.11.011]

Fang H W and Shi H J. 2019. Satellite image recognition and classification method based on deep learning. *Computer Systems & Applications*, 28(10): 27-34 (方浩文,施华君. 2019. 基于深度学习的卫星图像识别分类方法. *计算机系统应用*, 28(10): 27-34) [DOI: 10.15888/j.cnki.csa.007081]

Girshick R, Donahue J, Darrell T and Malik J. 2014. Rich feature hierarchies for accurate object detection and semantic segmentation// *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH: IEEE: 580-587 [DOI: 10.1109/CVPR.2014.81]

He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//*Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]

Hoeser T, Bachofer F and Kuenzer C. 2020. Object detection and image segmentation with deep learning on earth observation data: a review—part ii: applications. *Remote Sensing*, 12(18): 3053 [DOI: 10.3390/rs12183053]

Hong L, Huang Y J, Yang K, Peng S Y and Xu Q L. 2019. Study on urban surface water extraction from heterogeneous environments using GF-2 remotely sensed images. *Journal of Remote Sensing*, 23(5): 871-882 (洪亮,黄雅君,杨昆,彭双云,许泉立. 2019. 复杂环境下高分二号遥感影像的城市地表水体提取. *遥感学报*, 23(5): 871-882) [DOI: 10.11834/jrs.20198064]

Li H F, Dou X, Tao C, Wu Z X, Chen J, Peng J, Deng M and Zhao L. 2020. RSI-CB: a large-scale remote sensing image classification benchmark using crowdsourced data. *Sensors*, 20(6): 1594 [DOI:

10.3390/s20061594]

Liu J G, Zang C F, Tian S Y, Liu J G, Yang H, Jia S F, You L Z, Liu B and Zhang M. 2013. Water conservancy projects in China: achievements, challenges and way forward. *Global Environmental Change*, 23(3): 633-643 [DOI: 10.1016/j.gloenvcha.2013.02.002]

Liu K, Ren H G, Li S Y and Qin B Y. 2019. Automatic extraction of Tibet Plateau frozen lake based on Tiangong-2 multi-spectral data. *Infrared and Laser Engineering*, 48(3): 40-46 (刘康,任海根,李盛阳,覃帮勇. 2019. 基于天宫二号多光谱数据的青藏高原冻湖自动提取. *红外与激光工程*, 48(3): 40-46) [DOI: 10.3788/IR-LA201948.0303004]

Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C Y and Berg A C. 2016. SSD: single shot multibox detector//*Proceedings of the 14th European Conference on Computer Vision*. Amsterdam: Springer: 21-37 [DOI: 10.1007/978-3-319-46448-0\_2]

Luo J C, Sheng Y W, Shen Z F, Li J L and Gao L J. 2009. Automatic and high-precise extraction for water information from multispectral images with the step-by-step iterative transformation mechanism. *Journal of Remote Sensing*, 13(4): 610-615 (骆剑承,盛永伟,沈占锋,李均力,郜丽静. 2009. 分步迭代的多光谱遥感水体信息高精度自动提取. *遥感学报*, 13(4): 610-615) [DOI: 10.11834/jrs.20090405]

Ma X, Shao L M, Jin X and Xu G L. 2019. Improved YOLO model and its application in ship target recognition. *Telecommunication Engineering*, 59(8): 869-874 (马啸,邵利民,金鑫,徐冠雷. 2019. 改进的YOLO模型及其在舰船目标识别中的应用. *电讯技术*, 59(8): 869-874) [DOI: 10.3969/j.issn.1001-893x.2019.08.001]

McFeeters S K. 1996. The use of the normalized difference water index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7): 1425-1432 [DOI: 10.1080/01431169608948714]

Neubeck A and van Gool L. 2006. Efficient non-maximum suppression//*Proceedings of the 18th International Conference on Pattern Recognition*. Hong Kong: IEEE: 850-855 [DOI: 10.1109/ICPR.2006.479]

Redmon J, Divvala S, Girshick R and Farhadi A. 2016. You only look once: unified, real-time object detection//*Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV: IEEE: 779-788 [DOI: 10.1109/CVPR.2016.91]

Redmon J and Farhadi A. 2017. YOLO9000: better, faster, stronger// *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, HI: IEEE: 6517-6525 [DOI: 10.1109/CVPR.2017.690]

Redmon J and Farhadi A. 2018. YOLOv3: an incremental improvement. *arXiv: 1804.02767v1* [DOI: 10.48550/arXiv.1804.02767]

Ren S Q, He K M, Girshick R and Sun J. 2017. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6): 1137-1149 [DOI: 10.1109/TPAMI.2016.2577031]

- Shen Z F, Xia L G, Li J L, Luo J C and Hu X D. 2013. Automatic and high-precision extraction of rivers from remotely sensed images with Gaussian normalized water index. *Journal of Image and Graphics*, 18(4): 421-428 (沈占锋, 夏列钢, 李均力, 骆剑承, 胡晓东. 2013. 采用高斯归一化水体指数实现遥感影像河流的精确提取. *中国图象图形学报*, 18(4): 421-428) [DOI: 10.11834/jig.20130409]
- Shi W Z, Caballero J, Huszár F, Totz J, Aitken A P, Bishop R, Rueckert D and Wang Z H. 2016. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network//*Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, NV: IEEE: 1874-1883 [DOI: 10.1109/CVPR.2016.207]
- Tao Z Y. 2021. Detection and recognition of base in large-format remote sensing images. *Huazhong University of Science & Technology* (陶泽远. 2021. 大幅面遥感影像阵地目标检测与识别方法研究. 华中科技大学) [DOI: 10.27157/d.cnki.ghzku.2021.004412]
- Wang N, Cheng J H, Zhang H Y, Cao H J and Liu J. 2020. Application of U-net model to water extraction with high resolution remote sensing data. *Remote Sensing for Land and Resources*, 32(1): 35-42 (王宁, 程家骅, 张寒野, 曹红杰, 刘军. 2020. U-net模型在高分辨率遥感影像水体提取中的应用. *国土资源遥感*, 32(1): 35-42) [DOI: 10.6046 /gtzyy.2020.01.06]
- Wang X K, Jiang H X and Lin K Y. 2020. Remote sensing image ship detection based on modified YOLO algorithm. *Journal of Beijing University of Aeronautics and Astronautics*, 46(6): 1184-1191 (王玺坤, 姜宏旭, 林珂玉. 2019. 基于改进型YOLO算法的遥感图像舰船检测. *北京航空航天大学学报*, 46(6): 1184-1191) [DOI: 10.13700/j.bh.1001-5965.2019.0394]
- Wu J, Duan J, He L Q, Li Y C and Zhu W T. 2021. Research on aircraft detection algorithm of DS-YOLO network in remote sensing images. *Computer Engineering and Applications*, 57(1): 181-187 (吴杰, 段锦, 赫立群, 李英超, 朱文涛. 2021. DS-YOLO网络在遥感图像中的飞机检测算法研究. *计算机工程与应用*, 57(1): 181-187) [DOI: 10.3778/j.issn.1002-8331.1909-0409]
- Xu Y, Zhang Y J and Zhang Q. 2016. Rule sets building of water conservancy facilities spatial element extraction based on high resolution remote sensing image. *Geospatial Information*, 14(5): 71-74 (许莹, 张友静, 张琴. 2016. 基于高分影像的水利空间要素提取规则集构建. *地理空间信息*, 14(5): 71-74) [DOI: 10.3969/j.issn.1672-4623.2016.05.023]
- Zhang Z X. 2018. Fast Detection of Aircraft Target in Large-Scale Optical Remote Sensing Image. *Changchun: University of Chinese Academy of Sciences (Changchun Institute of Optics, Fine Mechanics, Chinese Academy of Sciences)* (张作省. 2018. 大幅面光学遥感影像飞行器目标快速检测. 长春: 中国科学院大学(中国科学院长春光学精密机械与物理研究所))
- Zheng Z Q, Liu Y Y, Pan C C and Li G N. 2019. Application of improved YOLO v3 in aircraft recognition of remote sensing images. *Electronics Optics and Control*, 26(4): 28-32 (郑志强, 刘妍妍, 潘长城, 李国宁. 2019. 改进YOLO V3遥感图像飞机识别应用. *电光与控制*, 26(4): 28-32) [DOI: 10.3969/j.issn.1671-637X.2019.04.006]
- Zhou L L. 2010. Study on water resource market allocation in Turim river basin. *Xinjiang Agricultural University* (周莉荔. 2010. 塔里木河流域水资源市场配置研究. *新疆农业大学*)

## Detection of water conservancy facilities in large-format image combining E-YOLO algorithm and NDWI constraint

XU Zeyu<sup>1,2</sup>, SHEN Zhanfeng<sup>1,3</sup>, LI Yang<sup>1,4</sup>, LI Junli<sup>5,6</sup>, WANG Haoyu<sup>1,2</sup>, LI Shuo<sup>1,4</sup>, JIAO Shuhui<sup>1,2</sup>, LI Lingling<sup>7</sup>

1. National Engineering Research Center for Geomatics, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China;

2. School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China;

3. College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China;

4. State Key Laboratory of Remote Sensing Science, Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China;

5. State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Ürümqi 830011, China;

6. Key Laboratory of GIS and RS Application, Xinjiang Uygur Autonomous Region, Ürümqi 830011, China;

7. National Disaster Reduction Center of China, Beijing 100124, China

**Abstract:** Water facilities play an important role in water scheduling, ecological protection and restoration of natural wetlands, utilization of resources and functions, and development. The traditional methods of statistics on the location and count of water conservancy facilities

rely on compiled data, which has disadvantages such as time-consuming, untimely data update, and unknown specific geographic locations. Remote sensing provides new possibilities for large-scale detection of water conservancy facilities. Aiming at the problem of detection of water conservancy facilities with remote sensing images, this study proposes a large-scale image detection algorithm. Based on the YOLO v3 network and the characteristics of water conservancy facilities, the study was divided into two main aspects: (1) We improved the YOLO algorithm and obtained the E-YOLO algorithm. We proposed a PPA feature fusion method and a four-feature map cross prediction method with proportional prediction box to optimize the problems of small samples. Besides, we improved the loss function by highlighting the loss of confidence. In addition, we used the transfer learning method to read part of the feature extraction parameters of the pre-trained model. (2) With the improved E-YOLO algorithm as the core, a large-area water conservancy facility detection algorithm combined with the water body index constraint was obtained. Aiming at the problem of large image size with a small target scale, we used the water body index to constrain the sliding step to reduce the missed detection rate and false detection rate at the same time. Then we combined the network output with the contour merging method to optimize the detection results. We used the GF-2 data for this study. The experimental results show that: the E-YOLO algorithm can significantly improve the detection effect of water conservancy facilities. Compared with YOLO v3, the average F2 score of E-YOLO is increased by 1.25% and the E-YOLO algorithm has a better stability. The large-area detection method constrained by the water index can improve the detection accuracy while ensuring efficiency. Compared with the large-step and small-step methods, its F2 accuracy is increased by 3.72% and 2.70%, respectively. Our method provides a good solution for the detection of water conservancy facilities.

**Key words:** water conservancy facilities, remote sensing detection, E-YOLO, large image, NDWI

**Supported by** National Key Research and Development Project (No. 2021YFC1523503); National Natural Science Foundation of China (No. 41971390, 41671034); The Third Comprehensive Scientific Expedition to Xinjiang (No. 2021xjkk1403); The Key Research and Development Program of Xinjiang Uygur Autonomous Region (No. 2022B03001-3)