引文格式：王腾,孟小亮,蔡冠中,等. 自然资源视频监测目标检测关键技术研究与应用[J]. 地理空间信息, 2024, 22(3): 47-50. doi:10. 3969/j. issn. 1672-4623.2024.03.0011

自然资源视频监测目标检测关键技术研究与应用

王　腾[1,2]，孟小亮[1,3*]，蔡冠中[3]，陈　影[3]，孙洁衍[3]，姚　鹏[4]
（1. 自然资源部华南热带亚热带自然资源监测重点实验室，广东 广州 510500；2. 广东省国土资源测绘院，广东 广州 510500；
3. 武汉大学 遥感信息工程学院，湖北 武汉 430079；4. 广西泰绘信息科技有限公司，广西 桂林 541002）

摘　要：随着机器视觉技术的突飞猛进，智能视频监测为自然资源监管提供了可实时响应的新执法手段。运用基于深度学习的域自适应目标识别技术和基于DEM辅助的高精度单目定位技术，对自然资源违法行为与线索进行快速目标检测；并在此基础上设计了自然资源视频执法监测系统技术架构，在广东开展了实地应用测试。结果表明，该方法对重型机械、土堆、建筑板材、在建砖房、新建板房、红砖堆等类别的识别准确率均高于80%，在1 200 m监测距离内平均定位误差小于3%。

关键词：自然资源；视频监测；目标检测；目标识别；目标定位
中图分类号：P208　　**文献标志码：**B　　**文章编号：**1672-4623（2024）03-0047-04

Research and Application of Target Detection Key Technologies for Video Monitoring on Natural Resources

WANG Teng[1,2], MENG Xiaoliang[1,3], CAI Guanzhong[3], CHEN Ying[3], SUN Jieyan[3], YAO Peng[4]
(1. Key Laboratory of Natural Resources Monitoring in Tropical and Subtropical Area of South China, Ministry of Natural Resources, Guangzhou 510500, China;
2. Surveying and Mapping Institute Lands and Resource Department of Guangdong Province, Guangzhou 510500, China;
3. School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China;
4. Guangxi Taihui Information Technology Co., Ltd., Guilin 541002, China)

Abstract: With the rapid development of machine vision technology, intelligent video monitoring provides a new law enforcement means that can respond in real time for natural resource supervision. In this paper, we used domain adaptive target recognition technology based on depth learning and high-precision monocular positioning technology based on DEM assistance to rapid detect the illegal acts and clues of natural resources. On the basis of this, we designed the technical framework of natural resources video law enforcement monitoring system, and carried out field application tests in Guangdong. The results show that the accuracy of this method in identifying heavy machinery, mounds, building boards, brick houses under construction, new plank houses, red brick piles and other categories is higher than 80%, and the average positioning error is less than 3% within the monitoring distance of 1 200 m.

Key words: natural resource, video monitoring, target detection, target recognition, target location

近年来生态文明建设上升到新的高度，“山、水、林、田、湖、草、矿、海、城”统一监测监管体系是国家生态文明建设的重要保障。《自然资源部信息化建设总体方案》指出，要立足已有基础，统筹整合信息化资源，运用传感器等现代物联网技术，构建全天候监测与预警的感知体系，形成实时快速的自然资源动态监测和态势感知能力，实现对国土空间的全时全域立体监控[1]。结合高精尖新技术研发，综合利用天空地网多途径技术，升级移动巡查、监管、调查等信息系统，以促进有限资源的高效配置与利用，已成为当前国土资源管理的主要手段[2-3]。

土地卫片执法、耕地卫片监督、耕地保护监察等自然资源监管工作中迫切需要全天候、准确、实时、快速的决策机制。目前日常巡查、媒体披露、群众举报、卫片执法是自然资源在线巡查平台的主要发现手段，其中卫片执法主要是通过比对卫星遥感影像时相变化提取疑似违法图斑，在地物目标与变化检测方面已发展得较成熟，但受卫星重访周期影响，时效性较

收稿日期：2023-01-09。
项目来源：自然资源部华南热带亚热带自然资源监测重点实验室开放基金课题资助项目（2022NRMZ02）。
第一作者简介：王腾（1984—），硕士，高级工程师，研究方向为地理信息系统技术及应用，E-mail：wangteng43@hotmail.com。
通信作者：孟小亮（1981—），博士，教授，研究方向为智能空间感知网，E-mail：xmeng@whu.edu.cn。

差[4-5]。随着机器视觉技术的突飞猛进，智能视频监测为自然资源监管提供了可实时响应的新执法手段，可根据塔基或固定杆上高清摄像头获取的视频图像数据进行违法线索智能分析，并在国内部分省尝试应用示范[6]。以卷积神经网络为代表的视觉自动识别检测技术成为目前的研究热点[7-8]，但该技术在复杂变化环境下的识别准确率、视觉定位精度以及多源感知融合监测等方面还面临挑战性问题[9]。目前，计算机视觉领域中域自适应方法被证明能应用于光照条件变化、天气条件变化（有雨、雾、霾等）、季节变化等挑战性场景中；无监督域自适应算法使深度学习模型在不同环境中都能取得良好的效果[10]；利用大量标注样本进行模型训练可提高应用场景中目标识别的准确率[11]。自然资源监测场景中多采用单目非专业量测摄像头定位检测目标对象，价格低廉、测距模型简单、运算速度快[12-13]。本文针对上述需求与难点，设计了立体式一体化自然资源视频执法监测系统的技术架构，利用基于深度学习的域自适应目标识别技术和基于DEM辅助的高精度单目定位技术对自然资源违法行为与线索进行快速目标检测，并通过系统实现与应用验证了关键技术的有效性。

1 总体技术架构

总体技术体系架构包括基础设施、平台服务与业务应用3层。基础设施中的感知层为由摄像机、无人机、卫星遥感、激光雷达等感知手段获取的多模态监测数据；网络层包括4G、5G等网络传输形式；支撑层通过平台应用与存储服务器实现数据的存储、组织与管理，利用样本库服务器支撑平台服务层的模型训练，还可利用边缘计算设备进行前置业务处理与分析，降低中心服务器负载。平台服务层提供目标识别、目标定位、变化检测、多模态数据融合与交互验证等模型与方法的关键技术，以及多源数据标准化与共享、地理空间信息分析与决策等服务。业务应用层可调用平台服务实现视频执法、卫片执法、日常巡查、智能预警、统计分析等自然资源执法监管业务功能。技术体系同时需要安全保障与运维管理两子系统作为支撑。

2 关键技术

2.1 域自适应的违法线索视频目标识别检测技术

借助计算机视觉方法识别真实场景下的违法违章线索行为，依赖实际违法违章数据样本提高识别模型的准确率，同时要考虑天气变化等复杂条件下的监测可靠性，因此本文从构建数据样本库和域自适应方法两个方面开展关键技术研究。

1）根据自然资源智能执法的需求和场景目标的特点，自然资源违法行为往往伴随着一些显著特征的目标类别。根据对违法行为的整理，本文将违法线索分为乱石、土堆、建筑板材、地面硬化、在建砖房、新建板房、红砖堆和重型机械等8类（表1），收集目标数据样本并进行标注，存储于样本库服务器，以支撑目标识别检测模型的研究。

2）利用机器视觉中跨域自适应目标检测方法解决自然资源视频监测中目标识别准确性受恶劣天气、拍摄角度、光照条件等因素影响的难题。其关键技术主要包括图像级适应、实例级适应和域损失的条件对抗。基于迁移学习的无监督跨域目标检测，通过学习到域无关的特征，改善目标域由于恶劣天气、拍摄角度、光照条件等原因产生的域偏差，进而增加模型的鲁棒性、泛化性。

表1 目标识别类别

类别	分类依据	样本数	图例	类别	分类依据	样本数	图例
乱石	水泥砖块（堆）、开采石块堆、大块的石块、砾石堆（呈银白色与灰黑色）	2 217		在建砖房	自带手脚架的砖房，包括带防护网的砖房	2 291	
土堆	防护土堤、泥土堆、沙土土堆，包括带防护网的土堆、堆放在大棚内的土堆	2 750		新建板房	临时板房、竹编板房、木板房、彩钢板房等，不包括四面通风的棚	1 995	
建筑板材	木材、钢材、砌块（堆）、管材，不包括砖房内放置的板材	2 587		红砖堆	大量红砖堆放在一起，不包括红砖墙	2 249	
地面硬化	路面上的凹陷、凸起、裂缝	1 750		重型机械	运输机械、土方机械、起重机械、路面机械	2 345	

基于域自适应的跨域目标检测网络模型见图1，可有效地将在有标签的数据集上训练的目标检测模型迁移到具有明显域差异的无标签注释的数据集上，以减少重新标注数据集投入的人力物力。对于Faster-RCNN目标检测网络，首先将来自于源域和目标域的两张图像分别送入深度卷积网络中提取图像特征，在提取的特征图上添加域分类器以达到图像级域适应的效果，

通过梯度反转层，原来的目标检测分支仍最小化目标检测损失，而域分类分支则最大化域分类损失用于混淆目标域数据与源域数据；然后经过RoI Pooling和全连接层获得实例级的特征，在该特征图上添加一个域分类器达到实例级域适应的效果。总的损失函数为目标检测的分类损失、定位损失和两个域分类器的损失。该模型在自然资源违法监测等目标检测数据集上取得了较好的效果，使复杂变化环境下目标检测的准确率提高了2%~12%[14]。

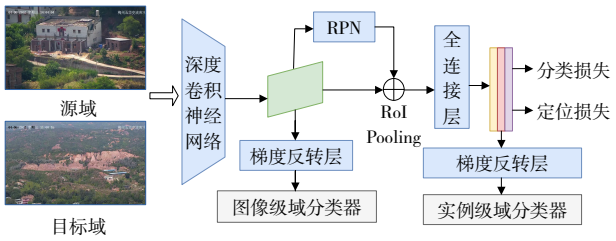


图1　基于域自适应的跨域目标检测框架

2.2　基于DEM辅助的高精度单目定位技术

自然资源视频监测场景中大多数摄像头是单目摄像头，因此单目定位技术是实现视频监测的重要环节。众所周知，一条射线在三维空间中无法形成交点，仅依靠单视影像无法实现目标的精准定位，只能根据空间相似性原理进行目标的概略定位。因此，本文基于DEM辅助的高精度单目定位技术，通过从影像中心到特征点的射线与DEM的交会实现目标的精准定位，算法流程见图2。摄像头内参标定时，采用张正友棋盘标定法对摄像头进行参数标定，在不同焦距下计算光学中心的像素坐标和畸变参数等相机内参。在实际监测中，由于摄像头位置已经固定，因此采用移动棋盘格的方式进行影像数据采集。自然资源监测的摄像机多采用变焦镜头，在固定焦距下，应保证影像中棋盘格网点具有极高的辨识度。数据采集完成后，首先检测每张影像上的棋盘格网点，利用格网点的像素坐标与世界坐标计算影像的单应矩阵；然后利用这些单应矩阵，分步计算得到影像的内外参数和畸变的估计参数；最后采用极大似然法优化估计参数。除标定的相机内参和测量的位置参数外，实现单目定位还需摄像头云平台提供的角度参数；但摄像头云平台安装时并未严格对齐世界坐标系，存在一定的角度差异，因此需要提前校正以获取云台局部坐标系与世界坐标系之间的相对旋转矩阵。首先利用多组地面控制点及其对应的像素坐标计算投影矩阵，并利用已知的内外参数获取像空间坐标系与世界坐标系之间的绝对旋转矩阵；然后将从云台获取的云台局部坐标系与像空间坐标系之间的旋转矩阵的逆矩阵左乘绝对旋转矩阵，即可得到相对旋转矩阵。

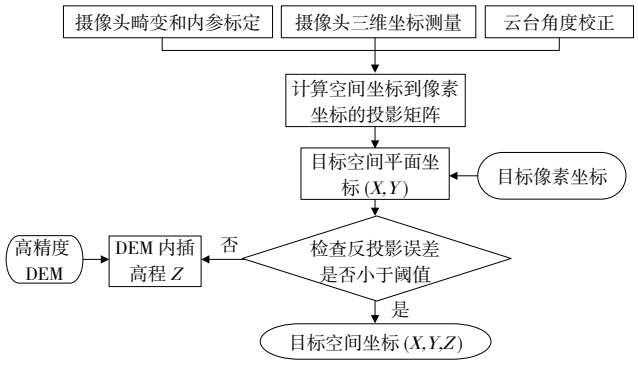


图2　基于DEM辅助的高精度单目定位流程图

基于DEM辅助的高精度单目定位技术的关键在于如何实现射线与不规则DEM面的求交。本文通过高程迭代逼近的方式实现精准定位，具体实现见图3。利用投影矩阵 $\boldsymbol{P}$ 建立从像素坐标 $(x, y, 1)$ 到世界坐标 $(X, Y, Z, 1)$ 的投影变换，即 $(x, y, 1) = \boldsymbol{P}(X, Y, Z, 1)$；但对于一个特征点而言，只能建立两个方程，而未知数有3个，因此预设一个初始高程平面 Z_0 代入该投影变换，可以计算得到初始平面坐标 (X_0, Y_0)，再利用该平面坐标内插DEM，即可得到更接近真实高程的 Z_1 高程面。迭代上述过程，不断更新高程面，直到反投影误差小于某极小阈值为止，即可得到精度极高的目标三维坐标。

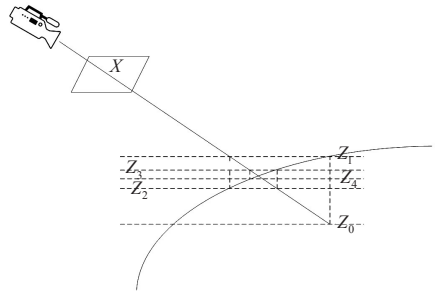


图3　高程迭代逼近过程

3　系统实现与应用实验

3.1　系统功能实现

根据上述总体技术架构和关键技术，本文设计开发了立体化自然资源执法监管系统。系统的主要功能模块包括监管视频管理、违法行为智能分析和智能监管与执法3大类，分为12个子功能模块（图4）。本文以视频执法智能分析中的违法线索目标识别为例，介绍系统功能的设计与实现。系统基于B/S架构和SpringBoot的MVC模式，Web服务部署采用Tomcat，功能模块开发采用Eclipse开发工具和Java编程语言，人机交互可视化界面基于HTML和Javascript编程。违

法线索目标识别的功能界面见图5，用户通过列表选择查看监控摄像头，系统可以抓取监控摄像头的监控图像，利用违法线索目标识别与定位算法分析处理监控图像，通过在监控视频画面中标注出违法线索目标的位置与类别的方式展示识别与定位结果。

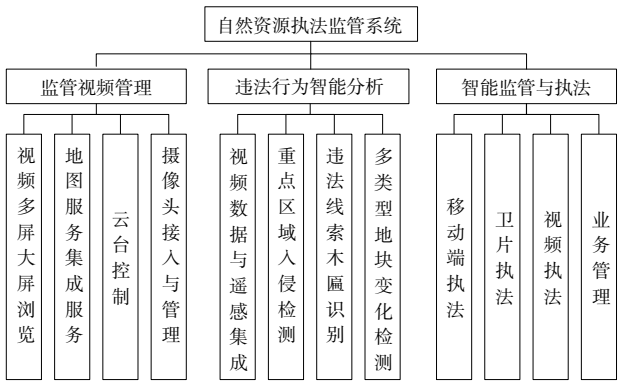


图4 系统功能模块树

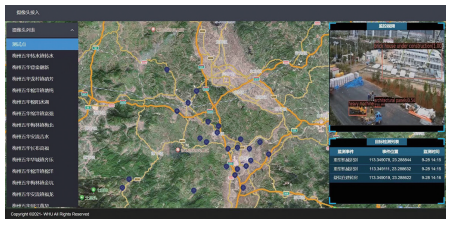


图5 系统功能界面

3.2 应用验证分析

系统在广东省某区县开展了试点应用，本文以2022年测试期间应用效果为例进行分析，采用Faster-RCNN网络，将数据以8∶2的比例划分训练集和验证集，以召回率和准确率为评价指标，评估8类目标的测试结果（表2），可以看出，该方法对重型机械、土堆、建筑板材、在建砖房、新建板房、红砖堆等类别的识别准确率均高于80%，对乱石、地面硬化的识别准确率较低，主要是受其外观特征难区分等因素影响。对于目标定位的应用，本文选取不同距离共6组每组10次的结果进行实验分析（表3），可以看出，在1 200 m监测距离内单目定位平均相对误差在3%以内。

表2 目标识别应用测试结果

类别	识别图像	召回率	准确率
重型机械	664	0.759	0.84
乱石	720	0.697	0.712
土堆	946	0.727	0.871
建筑板材	816	0.725	0.863
地面硬化	737	0.531	0.626
在建砖房	640	0.845	0.912
新建板房	812	0.76	0.804
红砖堆	600	0.768	0.876

表3 单目定位应用测试结果

组号	监测距离/m	检测数	平均测量距离/m	平均计算距离/m	平均测距误差/m	平均测距相对误差/%
1	0~200	10	140.23	141.85	−1.62	−1.16
2	200~400	10	272.34	267.64	4.70	1.73
3	400~600	10	458.53	450.73	7.80	1.70
4	600~800	10	713.68	729.07	−15.59	−2.18
5	800~1 000	10	898.52	877.58	20.94	2.33
6	1 000~1 200	10	1 099.25	1 076.03	23.22	2.11

4 结 语

本文设计了自然资源视频执法监测系统技术架构，利用域自适应目标识别和基于DEM辅助的高精度单目定位技术对自然资源违法行为与线索进行了快速目标检测；并在广东省某区县开展了实地应用测试。结果表明，该方法对重型机械、土堆、建筑板材、在建砖房、新建板房、红砖堆等类别的识别准确率均高于80%，在1 200 m监测距离内的平均定位误差小于3%。

从应用价值角度看，我国在自然资源、生态环境、安全应急等多个行业领域监测中都部署了海量监控摄像头，用于动态实时监控土地、矿产、森林、草原、湿地、海洋等各类资源，但视频监测资源缺乏与地理信息系统的动态、高效、智能的耦合机制，因此搭建融合视频与地理空间信息的自然资源执法监测系统，形成实时、快速的自然资源感知能力，能显著降低建设执法智能视频监控系统的成本，提升自然资源风险预测、预警、预防能力以及应急指挥调度能力，形成智能感知、快速反应、精准指挥、科学决策的现代化自然资源治理体系。

参考文献

[1] 中华人民共和国自然资源部. 自然资源部信息化建设总体方案[R/OL]. (2019-11-22)[2023-01-12]. http://gi.mnr.gov.cn/201911/t20191122_2483072.html

[2] 谢鸣,任建福,李新东,等. 天空地网一体化技术服务自然资源调查监测[J]. 地理空间信息,2022,20(9):32-35

（下转第54页）