引文格式：丁乐乐, 王珍, 刘艳飞, 等. 基于深度学习的土地违法施工场景监测研究[J]. 地理空间信息, 2023, 21(4): 72-74. doi:10. 3969/j. issn. 1672-4623.2023.04.017

基于深度学习的土地违法施工场景监测研究

丁乐乐[1]，王　珍[1]，刘艳飞[1]，孟凡效[1*]
（1. 天津市勘察设计院集团有限公司，天津 300000）

摘　要：土地利用保护是国土资源管理的一项重要职能，高点视频影像为土地利用保护提供了一种实时全天候的监测手段。为实现基于视频影像的自然资源疑似违法建设自动化监测和预警，将违法建设、违法施工等监测转化为建筑施工机械等内容的检测，提出了一种基于深度学习的土地违法施工检测方法。首先利用视频监控影像对施工机械进行样本训练，然后基于深度学习目标检测方法YOLOv3实现建筑施工机械检测，进而实现疑似违法建设预警。实验证明，该方法可从视频影像中自动化检测出施工机械，且可靠性较高，可为全天候耕地疑似违法施工预警提供技术支撑。

关键词：土地利用保护；视频影像；目标检测；施工机械
中图分类号：P258　　**文献标志码：**B　　**文章编号：**1672-4623（2023）04-0072-03

Research on of Land Illegal Construction Scenario Monitoring Based on Deep Learning

DING Lele[1], WANG Zhen[1], LIU Yanfei[1], MENG Fanxiao[1]
(1. Tianjin Survey Design Institute Group Co., Ltd., Tianjin 300000, China)

Abstract: Land-use protection is an important function of land resources management. High point video provides a real-time all-weather monitoring means for land-use protection. In order to realize the automatic monitoring and early warning of suspected illegal construction of natural resources based on video images, we transformed the illegal constructionmonitoring into the construction machinerydetection, and proposed a land illegal construction detectionmethod based on deep learning. Firstly, we usedvideo surveillance images to trainthe construction machinery samples.And then, werealizedthe construction machinery detection based on deep learning object detection method YOLOv3, so as to realize theearly warning of suspected illegal construction. Experimental results show that this method can automatically detect construction machinery from video images and has high reliability, which can provide technical support for the early warning of suspected illegal construction in all-weather farmland.

Key words: land-use protection, video image, object detection, constructionmachinery

视频数据已经逐步应用于自然资源监管，但对于自然资源疑似违法案件的发现还依赖人工巡查、判断，无法实现视频自动监测、报警。针对这一问题，本文利用人工智能深度学习方法，将违法建设、违法施工等占用基本农田的违法行为监测问题转化为建筑施工机械检测问题。通过检测出视频监控区域内是否出现挖掘机、运土车等建设机械，实现疑似违法建设行为自动预警[1-13]。

1　基于深度学习的建筑施工机械检测

本文提出的方法流程主要包括3个部分，即视频影像预处理、视频施工机械检测、根据真实标注值和检测结果计算损失函数3个部分，对于给定视频影像，首先对其进行归一化等预处理，然后输入深度卷积神经网络进行目标检测，得到目标检测结果。在训练阶段，根据真实标准值和检测的结果计算损失函数，计算参数的偏导数，更新网络参数（见图1）。

1.1　数据预处理与视频施工机械检测

利用摄像头对施工现场中常见的施工机械进行了标注，主要包括塔吊、起重机、挖掘机、脚手架、防尘网、建筑材料、推土机、其他机械等。设 $L=\{x_i,y_i|i=1,2\cdots,n\}$ 标注的训练集，其中 x_i 为第 i 张影像；$y_i=\{y_{i,1}\cdots,y_{i,t}\}$ 为第 i 张影像对应的标注数据；$y_{i,t}=R^5$ 为第 i 张影像的第 t 个目标实例，记录目标的矩形框左上角坐标、右下角坐标、以及类别编码。在本文中，由于摄像头拍摄距离、角度等原因，待检测机械目标尺度变化多样，因此本文采用Dark-

收稿日期：2021-12-10；修回日期：2022-03-28。
项目来源：天津市重点研发计划科技支撑重点项目（18YFZCSF00620）；天津市重点研发计划院市合作项目（18YFYSZC00120）。
第一作者简介：丁乐乐（1986—），高级工程师，主要从事卫星精密定位、测绘地理信息相关工作，E-mail：kcydll@126.com。
通信作者：孟凡效（1990—），硕士，主要从事测绘地理信息、精密工程测量等工作，E-mail：739369615@qq.com。

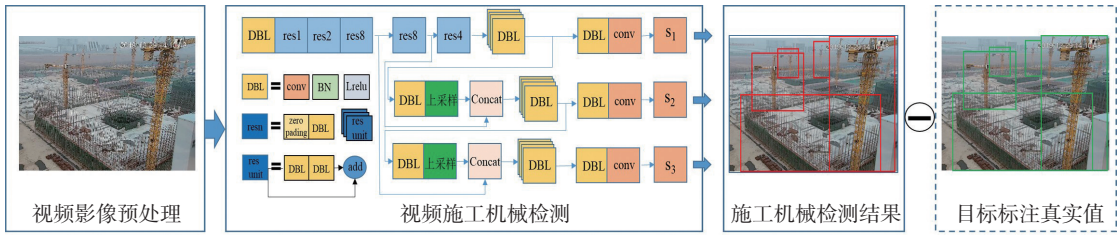


图1　基于深度学习的建筑施工机械检测流程图（CONV表示卷积层，BN表示归一化层，Concat表示将特征叠加）

Net-53[17]进行多尺度特征提取。在DarkNet-53网络中，其主要有卷积层，归一化层和残差网络层构成，通过堆叠多个卷积层和残差网络层达到层次化特征提取的目的。在卷积层，卷积核通过在图像上滑动来提取特征，获得特征图，每一个卷积核对应一幅特征图。由于卷积是一个线性运算，因此往往在每一次的卷积特征图上增加一个非线性映射操作。常采用的非线性激活函数有Sigmoid、Tanh、ReLU、Leaky ReLU等，由于Leaky ReLU的快速收敛性，本文采用Leaky ReLU函数作为非线性激活函数。

$$C^i=\sum_{k=1}^{K}w^{ik}*x^{ik}+b_i \tag{1}$$

$$Y^i=f(C^i) \tag{2}$$

式中，w^{ik}为第i个卷积核的第k个通道；b_i为第i个卷积核对应的偏置；$*$为卷积运算；Y^i为卷积层输出；$f(\bullet)$为响应函数；x^k为输入数据的第k通道。在残差网中，其通过在低层网络和高层网络之间建立跨越连接来保证低层网络向高层网络的信息流通，避免梯度弥散导致的深度网络训练收敛困难问题，其基本构成仍然为卷积层:

$$Y=x+\mathrm{conv}(\mathrm{conv}(x)) \tag{3}$$

式中，$\mathrm{conv}(\bullet)$为卷积操作。对于给定影像x，在DarkNet-53中，通过引出3个分支获得3个不同尺度的特征特征图F_1、F_2和F_3。相较于原始影像尺度，特征图F_1、F_2和F_3分别缩小32倍、16倍、8倍。

对于得到深度基础特征图，将其按照线性缩放关系缩放至输入图大小，则特征图上每一个位置处的像素在原图覆盖一块区域，特征图中的红色像素经过线性映射对应到原图中的红色区域（见图2）。

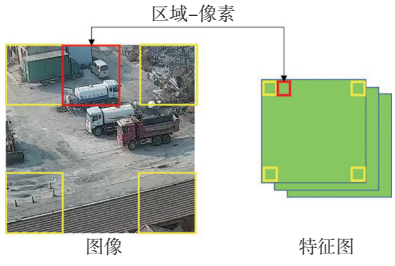


图2　图像局部区域与其特征图像素对应关系

在训练阶段，以其覆盖区域的中心点为核心，枚举3个尺度形状的先验框，由先验框与标注框计算缩放比作为回归目标。首先对标注数据中的标注矩形框进行K-means聚类以获得待检测机械目标的先验框。在本文中，聚类中心点k的值设置为9，经过K-means聚类，将得到9个先验矩形框，然后根据矩形框的面积分别将9个矩形框均匀分配到得到的3个深度基础特征图上。这样对于特征图上的像素来说，其将负责3个形状尺度矩形框回归。假设第一步获得的3个深度基础特征图F_1、F_2和F_3的大小分别为$w_f\times w_f$，$2w_f\times 2w_f$，$4w_f\times 4w_f$，则深度卷积神经网络将产生$63w_f^2$个矩形候选框，每个框含有一个中心点偏移$(x_{\mathrm{off}},y_{\mathrm{off}})$，宽高缩放系数$(w_s,h_s)$，置信度为$c$，以及属于每一类的类别概率分布为$p$。

1.2　损失函数计算

训练时，将计算$63w_f^2$候选框与真实标注数据之间的交并比以筛选出真正用于网络参数训练的预测矩形框。将候选框分为3种情况：正例、负例和忽略样例。对于正例的筛选，首先任取一个真实标注框，与$63w_f^2$个框全部计算交并比IOU，IOU最大的预测框，即为正例。在$63w_f^2$个框中，正例除外，与任意一个真实标注框的IOU大于阈值（本文中使用0.5），则为忽略样例，忽略样例不参与模型损失函数的计算，即不参与模型参数的更新。对于负例，在$63w_f^2$个框中，正例除外，与全部真实标注框的IOU小于阈值的则为负例。

对于得到的正例，首先计算其与对应的真实框之间的中心点偏移$(x_{\mathrm{off}}^{\mathrm{gt}},y_{\mathrm{off}}^{\mathrm{gt}})$，宽高缩放系数$(w_s^{\mathrm{gt}},h_s^{\mathrm{gt}})$：

$$\left(x_{\mathrm{off}}^{\mathrm{gt}},y_{\mathrm{off}}^{\mathrm{gt}}\right)=\left(x_{\mathrm{gt}},y_{\mathrm{gt}}\right)-(x_p,y_p) \tag{4}$$

$$w_s^{\mathrm{gt}}=w_{\mathrm{gt}}/w_p \tag{5}$$

$$h_s^{\mathrm{gt}}=h_{\mathrm{gt}}/h_p \tag{6}$$

式中，$(x_{\mathrm{gt}},y_{\mathrm{gt}})$为真实标注框中心点在特征空间的坐标；$(x_p,y_p)$为先验框中心点在特征空间的坐标；

w_s^{gt}, h_s^{gt} 为真实框的尺度；h_p, w_p 为先验框的尺度。对于某单一特征层来说，其损失函数为：

$$\begin{aligned}loss_{F_t}=&\beta_{box}\sum_{i=0}^{w_t\times w_t}\sum_{j=0}^{3}1_{ij}^{obj}[(x_{off}-x_{off}^{gt})^2+(y_{off}-y_{off}^{gt})^2]+\\&\beta_{box}\sum_{i=0}^{w_t\times w_t}\sum_{j=0}^{3}1_{ij}^{obj}[(w_s-w_s^{gt})^2+(h_s-h_s^{gt})^2]-\\&\beta_{obj}\sum_{i=0}^{w_t\times w_t}\sum_{j=0}^{3}1_{ij}^{obj}\log(c_{ij})-\beta_{noobj}\sum_{i=0}^{w_t\times w_t}\sum_{j=0}^{3}1_{ij}^{noobj}\log(1-c_{ij})-\\&\beta_{class}\sum_{i=0}^{w_t\times w_t}\sum_{j=0}\sum_{c\in classes}[p_{ij}^{'}(c)\log(p_{ij}(c))+(1-p_{ij}^{'}(c))\log(1-p_{ij}(c))]\end{aligned}\tag{7}$$

式中，$\beta_{box}, \beta_{obj}, \beta_{noobj}, \beta_{class}$ 分别为正例、负例以及全部实例的损失权重，classes 为类别个数。本文有个深度基础特征层，因此监督总损失函数为：

$$loss_{supervised_total}=loss_{F1}+loss_{F2}+loss_{F3}\tag{8}$$

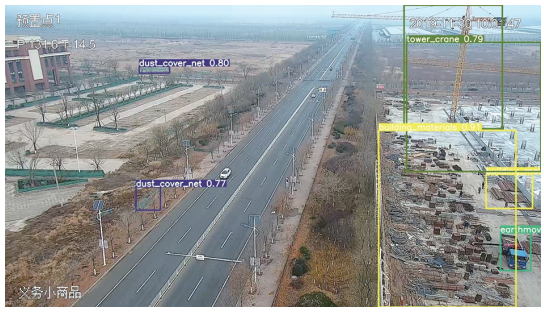
在卷积神经网络的训练中，通过交替前向–后向计算进行参数的训练更新。

2 实验结果分析

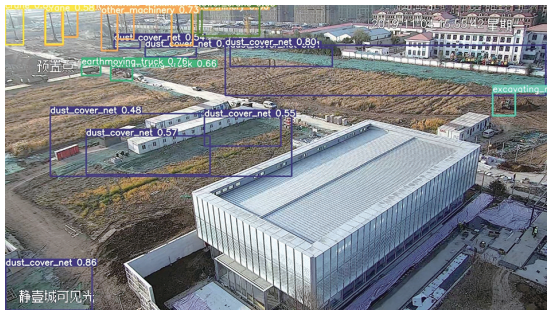
本文共标注了3 951张视频影像用于训练，439张视频影像用于测试。表1给出了采用的方法在测试影像上的检测精度。从表1可以看出本文采用的方法对施工机械的检测精度都在70%以上，平均检测精度达到78.8%，对脚手架、防尘网、建筑材料等目标的精度在80%以上，基本上满足自动化检测需求。图3给出了本文的检测方法在测试数据上检测可视化结果。从图3（a）、（b）中可以看出，本文采用的方法可以将起重机、推土机、防尘网、塔吊等施工相关机械检测出来。

表1 Yolov3对施工机械检测精度mAP

方法	塔吊	起重机	挖掘机	脚手架	防尘网	建筑材料	其他机械	推土机	mAP
Yolov3	71.7	74.1	79.2	80.8	89.1	84.9	78.3	72.6	78.8



a 实验点1视频监测结果



b 实验点2视频监测结果

图3 测试数据施工机械检测可视化结果

3 结 语

本文利用视频数据检测建筑施工机械来监测耕地中的违法建设、违法施工问题，该方法可以实现从视频影像中自动化检测出施工机械，为基于视频影像的全天候耕地疑似违法施工预警提供技术支撑，具有实用价值。

参考文献

[1] 曹凯,江南,吕恒,等. 面向对象的SPOT5影像城区水体信息提取研究[J]. 国土资源遥感,2007(2):27-30

[2] 许高程,毕建涛,王星星,等. 面向对象的高分辨率遥感影像道路自动提取实验[J]. 遥感信息,2012(2):108-111

[3] 徐京萍,赵建华,张丰收,等. 面向对象的池塘养殖用海信息提取[J]. 国土资源遥感,2013(1):82-85

[4] 王理政,林祥国,宁晓刚. 面向对象的高分辨率遥感影像阴影提取方法[J]. 测绘科学,2017,42(3):108-113

[5] 夏乐,李长安,江涛,等. 基于深度卷积神经网络的遥感图像船只识别[J]. 地理空间信息,2021,19(9):7-9

[6] 张应裕,徐敬根. 违法用地动态巡查的路线规划与执法盲区识别[J]. 地理空间信息,2021,19(4):121-125

[7] 吴志锋,郑宇恒,李双琴,等. 基于卫星遥感的输油管廊沿线的地物识别研究[J]. 地理空间信息,2022,20(2):45-50

[8] 史迪超,黎慧斌,李亭谕,等. 自然资源多时相遥感影像智能匹配方法研究[J]. 地理空间信息,2021,19(9):1-6

[9] 耿欣,雷丽珍,花卉,等. 基于深度学习方法的耕地违建自动提取[J]. 地理空间信息,2022,20(3):18-24

[10] 王化娟. 一种基于遥感影像疑似违法用地图斑提取方法的研究[J]. 测绘与空间地理信息,2020,43(3):175-177

[11] 赵逸君,黄牧涛. 国土环境资源遥感影像在线智能优化解译分类研究[J]. 国土资源信息化,2021(4):47-52

[12] 吕良勇. 关于高标准建设“慧眼守土”自然资源动态监管平台建设的初步探索[J]. 浙江国土资源,2019(9):20-21

[13] 薛丰昌,宋肖依,唐步兴,等. 视频监控的城市内涝监测预警[J]. 测绘科学,2018,43(8):50-55