

DOI:10.16544/j.cnki.cn43-1494/u.2023081923121982

文章编号:1674-599X(2023)06-0001-11

引用格式:李顺龙,郭亚朋.公路建造风险目标的视觉辨识与跟踪研究进展[J].交通科学与工程,2023,39(6):1-11.

Citation: LI Shunlong, GUO Yapeng. Research progress of vision-based identification and tracking of highway construction risk targets[J]. J. Transp. Sci. Eng., 2023, 39(6): 1-11.

# 公路建造风险目标的视觉辨识与跟踪研究进展

李顺龙, 郭亚朋

(哈尔滨工业大学 交通科学与工程学院, 黑龙江 哈尔滨 150090)

**摘 要:**智能辨识与跟踪公路建造过程中的风险目标是安全建造、智能管理的基础,也是公路智能建造的重要内容之一。随着计算机视觉技术的快速发展,国内外许多学者发展了以图像或视频为数据载体,融合计算机视觉智能算法的风险目标辨识与跟踪技术。该研究梳理了基于计算机视觉的公路智能建造风险目标辨识与跟踪领域的研究进展。分别从建造场地重建、风险目标智能辨识与风险目标智能跟踪三方面进行阐述,并对具有代表性研究进行回顾和分析,同时指出现有研究存在的局限和未来研究的方向。

**关键词:**公路智能建造;计算机视觉;风险目标;辨识与跟踪

中图分类号:U415.6

文献标志码:A

## Research progress of vision-based identification and tracking of highway construction risk targets

LI Shunlong, GUO Yapeng

(School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin 150090, China)

**Abstract:** The intelligent identification and tracking of risk targets in highway construction process is the decision-making basis of intelligent management of construction safety and one of the important contents of intelligent construction. With the rapid development of computer vision technology, researchers have developed many risk target identification and tracking technologies with computer vision algorithms that used images or videos as data carriers. This paper systematically reviews the recent research progress in the field of risk target identification and tracking of intelligent highway construction based on computer vision. The construction site reconstruction, risk target intelligent identification and risk target intelligent tracking are classified and elaborated, and the representative work is reviewed and analyzed. Furthermore, the limitations of current research and the future development direction are pointed out.

**Key words:** intelligent highway construction; computervision; risk target; identification and tracking

公路是交通运输物流体系的基础设施,事关一个地区乃至一个国家的社会经济交流和可持续发展。新时代更具战略性、先导性的公路建设方兴未

艾。《“十四五”现代综合交通运输体系发展规划》要求:“十四五”期间,中国公路通车总里程要再增加30.2万km,其中,高速公路里程增加1.2万km。同

收稿日期:2023-08-19

基金项目:国家自然科学基金优秀青年基金(51922034)

作者简介:李顺龙(1982—),男,哈尔滨工业大学教授。E-mail:lishunlong@hit.edu.cn

时,智能建造战略的提出也将有力地促进新一代信息技术与公路工程建造的有机融合,是实现中国公路行业高质量发展的重要依托<sup>[1-3]</sup>。

公路工程的建造载体就是建造场地。在建造过程中,建造场地会在有限的空间和时间内集聚大量的建造人员、建造机械、建造物资、建造材料等建造资源。复杂的人机物料交互和时空约束下建造场地的安全管理极具挑战性<sup>[4]</sup>。根据 HEINRICH 安全理论<sup>[5]</sup>,建造场地事故来自于不安全的参与人员和物体的状态,其中,超过 90% 的安全事故发生的根本原因在于参与人员不安全的行为和工作状态。因此,本研究中风险目标的定义为:与建造人员相关的建造资源,主要包括建造人员、建造运输车辆、建造机械、建造船舶等。只要通过技术手段获取风险目标的时空状态,确保其安全风险降到最低,就能最大限度地保证整个建造场地的安全运营。

风险目标的时空状态获取通常由建造管理人员定期根据风险目标检查单,对各类风险目标进行人工记录和评估,这种方法在中国公路行业高速发展时期,缓解了风险目标状态数据获取困难的局面,但这种基于人工的方法存在数据获取效率低、评估主观性强、获取方式风险高等问题,已不适用于公路智能建造的内在要求<sup>[6-7]</sup>。随着传感技术的发展,实时定位技术(real time location system, RTLS)被广泛应用在风险目标的辨识与跟踪领域。作为非接触式 RTLS 技术的一种,视觉传感技术因其直观性强、成本低、效率高的稠密传感特性被逐渐应用在建造领域。同时,随着计算机视觉技术的快速发展,国内外研究者发展了许多以图像或视频为数据载体,融合计算机视觉智能算法的风险目标辨识与跟踪技术<sup>[8-10]</sup>。

在公路工程建设场地,利用计算机视觉智能算法对视觉传感设备获取的风险目标的图像或者视频进行分析,可得到风险目标的时空状态,包括空间状态(空间位置信息、空间占用信息)、时变的空间信息及目标安全信息等。本研究将对基于计算机视觉的公路智能建造风险目标的时空状态获取技术进行系统性回顾和梳理,对建造场地的视觉重建、风险目标的辨识及风险目标的跟踪的逻辑等相关研究成果进行总结,并对其中代表性工作进行详细介绍,并指出当前基于计算机视觉的公路智能建

造风险目标辨识与跟踪研究存在的局限和未来的发展方向。

## 1 基于计算机视觉的建造场地重建

公路建造场地的重建是实现公路智能建造“数字孪生”的数据空间基础,是统一公路智能建造各子任务成果的框架,其高效实现至关重要。基于计算机视觉的场地重建得益于视觉感知的密集传感特性,可实现大面积、高精度、短时间的场地视觉数据获取,从而为场地重建提供了数据支持。

### 1.1 建造场地二维视觉重建

相比于工厂建筑或房屋建筑建造场地,公路建造场地的占地规模较大。获取大规模建造场地的二维图像的可行途径之一就是遥感,通过卫星或者航空器等平台能够快速对建造场地进行成像,但是这种方式成本较高且分辨率有限。随着无人机技术的发展,以装备有可见光摄像头的低空无人机为平台对建造场地进行多轮次成像后,将多张图像拼成一张高分辨率的建造场地全景图像,是当前最为适合的技术路径之一。BANG 等<sup>[11]</sup>分析了基于图像拼接技术的建造场地二维视觉重建过程中的问题,以图像预处理为重点,提出了场地视频模糊去除和关键帧选取的方法,实现了优于通用图像拼接软件的重建效果。

传统的多图像拼接需要对所有图像进行两两匹配以获取图像的相对空间位置,但无人机在拍摄图像时装备的 GPS 传感器和姿态传感器会记录这些图像的相对空间位置和姿态信息,而且传统的图像特征提取对不同环境的鲁棒性较差且计算量巨大。针对该问题,GUO 等<sup>[12]</sup>采用无人机进行场地图像获取时的坐标信息和姿态信息,提出了多信息融合的场地多图像快速配准方法,同时采用最佳缝合线算法,对场地中速度较快的风险目标因移动产生的鬼影效应进行了有效去除。如图 1 所示,该方法的主要流程:① 根据无人机装备的定位系统和姿态系统获取的图像位置信息,计算出多图像的相对空间位置;② 特征区域提取及区域特征提取与匹配,以计算出相邻图像间的单应性变换矩阵;③ 利用最佳缝合线算法对拼接好的图像进行融合以消除模糊及鬼影效应。

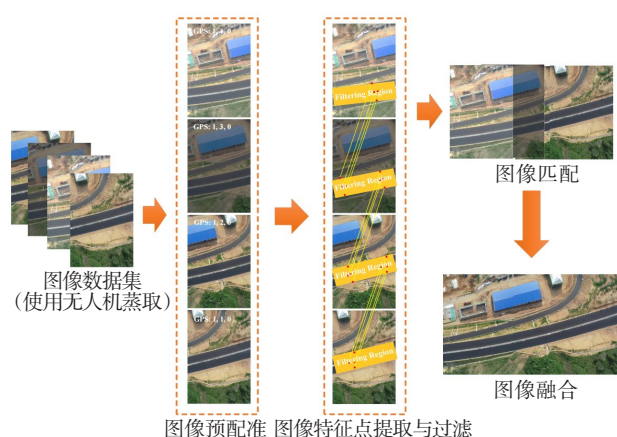


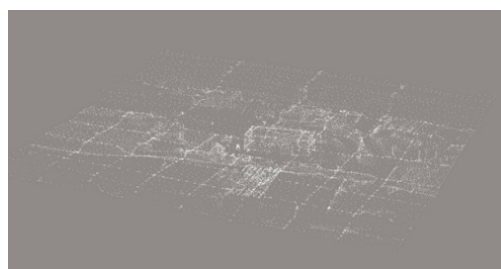
图1 基于无人机和计算机视觉的建造场地二维视觉重构框架

Fig. 1 Two-dimensional visual reconstruction of construction sites based on unmanned aerial vehicle and computer vision

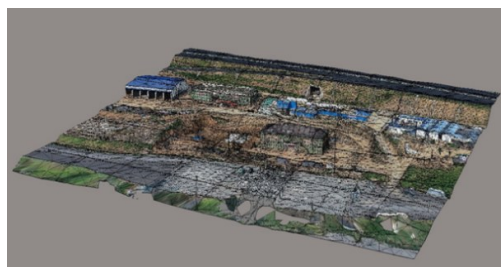
## 1.2 建造场地三维视觉重建

公路建造场地的三维重建目的是恢复场地真实的三维空间相对关系和纹理信息,实现建造场地的高精度数字化。三维重建是连接建造场地 BIM 模型和真实建造状态的重要媒介,可为建造场地的进度监测、安全监管、流程优化等提供重要参考<sup>[13]</sup>。传感技术的进步为建造场地的三维重建提供了激光雷达、卫星遥感、无人机摄像等方式,其中,无人机摄像以其便捷性、低成本、高精度等优点被国内外研究者广泛采用。

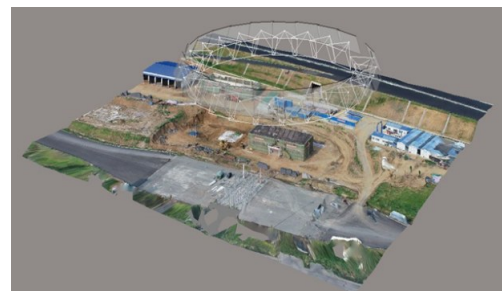
采用无人机摄像进行公路建造场地三维重建的基本流程为:相机标定、多视角图像采集、特征提取与匹配、光束平差、密集点云计算、表面重建、纹理粘贴。进行三维重建的工具除商业软件外,还包括一系列的开源平台算法。这些软件利用传统的三维重建算法,按照基本流程,能够对公路建造场地进行较为完整的三维重建。NIU<sup>[14]</sup>利用装备有单目镜头的无人机倾斜摄像获取建造场地的外观图像数据后,对图像数据进行预处理、空三加密、联合平差、瓦片分割及模型修饰后获取了建造场地的三维模型,如图2所示。将该模型放置在真实地图中加以展示和提供外部访问,初步实现了基于真实地图平台的建造场地二维和三维模型的统一展示。此外,研究者们对三维视觉重建结果的应用主要是利用三维模型与 BIM 模型进行对比、测量基坑或构筑物的几何尺寸或体积、分析风险目标间的相对位置与距离等<sup>[15-20]</sup>。



(a) 三维密集点云



(b) 三角网格重建



(c) 纹理粘贴



(d) 三维实景模型

图2 建造场地三维视觉重建的阶段

Fig. 2 Stages of three-dimensional visual reconstruction of construction sites

## 2 风险目标智能辨识

公路建造现场的风险目标辨识主要是对风险目标空间位置、尺寸、朝向和姿态行为等信息的识别。基于计算机视觉的建造场地风险目标辨识的发展与计算机视觉技术的发展有着相似的路径,都



经历了由传统图像处理方法到机器学习方法、再到深度学习方法的发展过程。风险目标的空间位置、尺寸和朝向辨识通常可同时完成,其主要依托计算机视觉领域的目标检测技术;而风险目标的姿态辨识则是在该辨识结果的基础上进行的,其主要依托计算机视觉领域的姿态识别技术。

## 2.1 风险目标空间位置、尺寸和朝向的辨识

利用计算机视觉领域中的目标检测技术,可以获取图像中风险目标的空间位置、尺寸和朝向等信息。基于传统图像处理的方法主要包括利用色彩空间变换直接分割出风险目标外轮廓、利用背景差分法对运动的风险目标进行提取、利用人工设计特征对风险目标进行匹配和识别等。这些方法针对特定环境的效果较好,但当外部环境发生变化时,则需要重新调整所有参数。基于机器学习的方法主要利用人工设计特征,如:有向梯度直方图(histogram of oriented gradients, HOG)、Haar-HOG、Blob-HOG等。它们提取风险目标的特征后,使用分类器(词袋模型、支持向量机等)对风险目标区域进行识别<sup>[21-24]</sup>。基于机器学习的方法将风险目标的空间位置、尺寸等信息转换成密集滑动窗口的目标分类问题,造成大量区域的无效识别和参数的人工设置。

深度学习为计算机视觉领域的发展带来了新的强力工具。深度学习摒弃了机器学习过度依赖特征提取的特点,最大限度地减少人工设计或人为经验的参与。同时,随着图像或视频数据的获取难度逐渐降低,大规模的数据为基于深度学习的风险目标检测提供了基础。基于深度学习的目标检测的主要技术实现途径是深度卷积神经网络(convolutional neural network, CNN)。基于深度学习的公路智能建造风险目标空间位置和尺寸辨识研究可以分为基于两阶段与单阶段目标检测框架两类。两阶段目标检测框架通常指的是将检测框的候选提取和检测框的分类回归分开来进行训练,先提取候选检测框后,再判断检测框包含目标的类别及精修检测框的位置。该方法的代表性算法则是Faster R-CNN网络。研究者们采用该算法对公路建造过程中的常见风险目标进行检测,如:建造人员的位置及安全装置佩戴情况的辨识、建造车辆的位置和尺寸等<sup>[25-28]</sup>。Retina网络和Mask R-CNN网络也被用来对建造机械的尺寸和位置等进行辨

识<sup>[29-30]</sup>。单阶段目标检测框架则是将检测框的候选提取和检测框的分类回归同时来进行训练。因此,在训练数据和测试速度时,一般采用两阶段方法。该方法的代表性算法则是YOLO系列网络(主要是YOLO v2和v3)和SSD网络<sup>[31-33]</sup>。研究者们直接采用或者改进这些算法对建造人员、建造机械、建造车辆的位置和尺寸进行检测,针对小尺寸的建造人员安全帽检测还开发了特定的注意力模块,并将其插入到SSD网络中<sup>[32]</sup>。为了进一步提升检测速度,研究者们还采用MobileNet等轻量化基网络<sup>[34]</sup>、网络蒸馏技术<sup>[35]</sup>等对单阶段目标检测框架进行精简。

这些研究极大地推动了风险目标空间位置和尺寸辨识的智能化水平,然而在辨识方面涉及甚少。GUO等<sup>[36]</sup>发现过往研究多采用水平检测框来进行风险目标检测的局限性,即水平检测框仅能表征目标位置和尺寸,提出了采用旋转检测框对目标位置、尺寸和朝向进行表征,并发展了基于旋转检测框的单阶段Anchor-based风险目标位置、尺寸及朝向辨识网络,如图3所示。该网络由3个部分组成:①特征提取模块,用来提取风险目标的多层级特征;②特征融合模块,用来对风险目标的多层级特征进行更高效的融合;③旋转检测框候选生成及回归模块,用来生成旋转检测框及对其参数进行精准回归。为进一步提升辨识网络的速度和精度,GUO等<sup>[37]</sup>还提出了Anchor-free的风险目标高效检测模型,该模型抛弃了Anchor-based方法的旋转检测框的候选生成过程,利用点的坐标直接回归旋转检测框的中心点,在此基础上直接回归依托于中心点的其他旋转检测框参数,有效降低了模型的复杂度和运行时间,如图4所示。

## 2.2 风险目标姿态辨识

基于计算机视觉的公路智能建造风险目标姿态辨识指的是以数字图像为输入,通过不同的深度学习网络,最终输出风险目标的姿态或者行为。研究者们主要针对空间自由度高的风险目标(建造人员、建造铰接机械等),通过提取其关键部分或关键节点的位置来判断风险目标的姿态或行为。TANG等<sup>[38]</sup>研发了基于改进的Faster R-CNN网络的建造人员-机械交互状况识别框架,可对建造人员面临的潜在碰撞及安全风险行为进行预警。LUO等<sup>[31,39]</sup>考虑远程监控视频特征融合方法,提出了基于级联

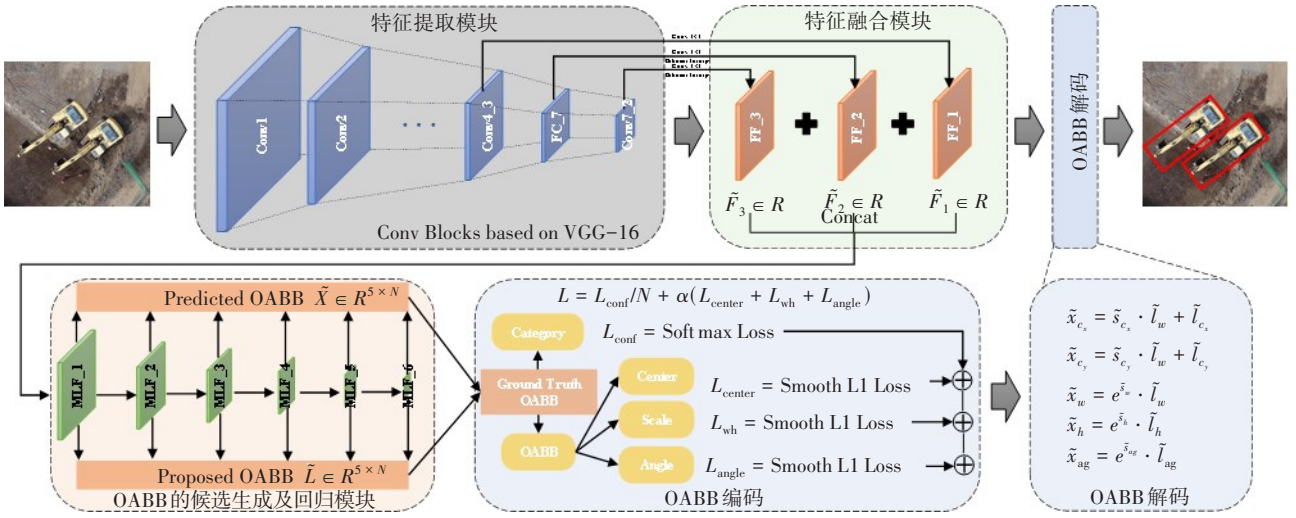


Fig. 3 Anchor-based risk target detection model

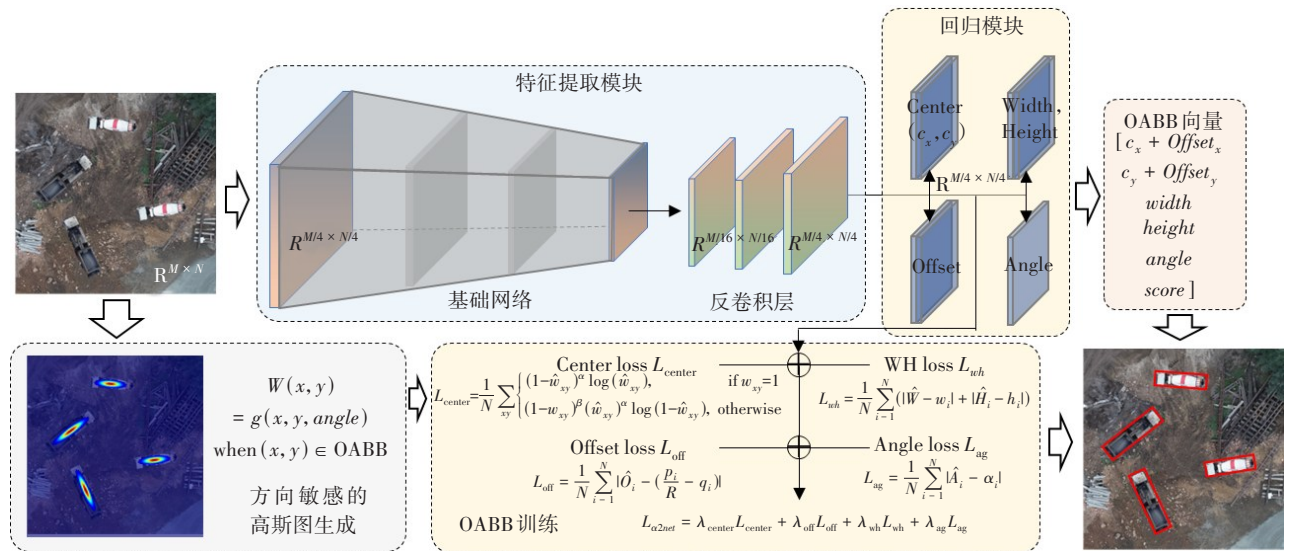


Fig. 4 Anchor-free risk target efficient detection model

统计模型的建造人员动作识别方法,同时还发展了融合深度动作特征和上下文信息的判别式模型,进一步提升建造人员动作的识别准确率,并采用条件随机场来对建造人员关键点空间关系进行建模。另一个主要研究对象则是最常用的建造铰接机械——挖掘机。KIM等<sup>[40]</sup>较早地提出使用两个不同的深度学习模型来构建挖掘机的姿态识别框架,其中一个负责学习和分析挖掘机的视觉多层级特征,由CNN和长短期记忆(long short-term memory, LSTM)网络组成;另一个负责考虑时序关系,输出不同时刻的挖掘机姿态,由LSTM网络构成。LIANG等<sup>[41]</sup>则较早地采用人体关键点检测的思路来对识

别挖掘机的不同姿态,利用人体关键点检测领域中表现优异的Hourglass网络来提取挖掘机的多层级特征,从而实现了不依赖人工标记的挖掘机关键点位置检测。LUO等<sup>[42]</sup>提出融合StackedHourglass网络和Cascaded Pyramid网络的挖掘机关键点检测算法来提取6个点坐标,并利用在线数据进一步提升准确率。CHEN等<sup>[43-44]</sup>提出了采用三维卷积网络,以挖掘机视频为整体输入,对挖掘机的姿态和行为直接进行识别的方法,并利用三维注意力模块和特征金字塔模块对时空特征进行进一步提炼。

深度学习因其强大的建模能力在风险目标姿态或行为辨识领域获得了广泛应用,然而为了



提升辨识精度,这些深度学习模型普遍存在参数量巨大、运算量大的问题,导致对硬件要求高、辨识速度慢。GUO等<sup>[45]</sup>提出基于轻量化技术的风险目标姿态辨识高效网络,从架构、模块和网络层3个级别对深度学习模型进行大幅度精简,实现了在不降低准确率的前提下,参数量的大幅下降和辨识速度的大幅提升。如图5所示,在架构

层次上,采用高效简便的“编码-解码”结构来对风险目标的多层级特征进行提取和恢复;在模块层次上,采用深度可分离卷积DSC来代替传统的卷积操作,在保证特征提取能力的同时,可以大幅降低卷积操作的参数;在网络层次上,人工标记的真实姿态被转换为以关键节点为中心的多个高斯分布。

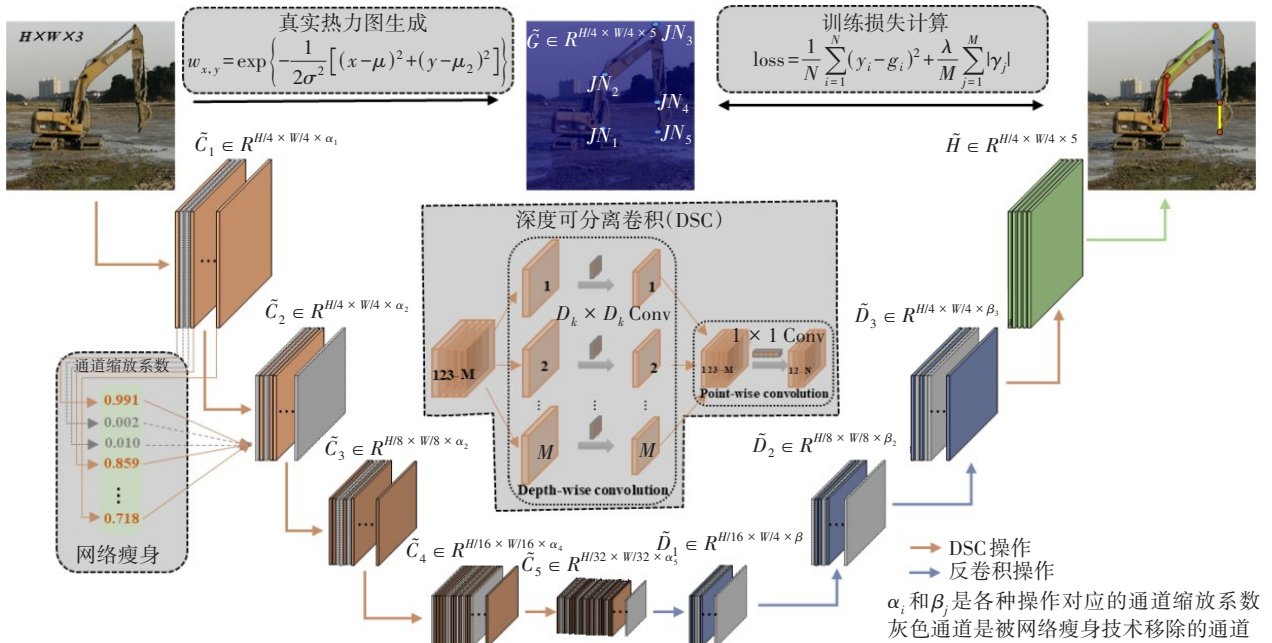


图5 轻量化风险目标姿态辨识模型

Fig. 5 Lightweight risk target pose identification model

### 3 风险目标智能跟踪

基于计算机视觉的公路智能建造风险目标的跟踪指的是在获取到通过风险目标辨识得到的空间位置、尺寸、朝向和姿态行为等信息基础上,考虑时间变化,计算这些信息随时间演化的规律并对规律进行建模,从而得到这些信息的时变模型以进一步预测未来的风险目标信息。

#### 3.1 风险目标二维跟踪

基于计算机视觉的公路智能建造风险目标二维跟踪指的是通过二维数字图像获得时变的以像素为单位的的目标尺寸、轨迹和朝向。较早的相关研究都是基于传统的时序建模方法,以目标像素值或其简易变换、人工设计算子、人工靶标、深度学习多层级特征、检测框结果、分割结果等特征为辨识基础,并采用Mean-Shift算法、KLT光流法<sup>[46]</sup>、TLD算

法<sup>[40]</sup>、卡尔曼滤波<sup>[47]</sup>、核相关滤波<sup>[31]</sup>等技术对目标的时序特征进行建模。

深度学习可增强时序建模的鲁棒性,从而提升目标跟踪的准确性。TANG等<sup>[38,48]</sup>发展了基于上下文信息和深度学习的风险目标时序建模方法,采用时间序列建模的神经网络结构对多目标位置、相对关系及同目标点的距离等特征进行融合并对可观测时间内的信息进行建模和解码预测。BANG等<sup>[49]</sup>基于风险目标的分割结果,采用ConvLSTM,对未来帧的目标所在区域进行预测,相比于之前利用少数几个特征的预测方法,其具有预测结果丰富的优点。LUO等<sup>[50]</sup>采用在单目标跟踪领域表现优异的MDNet来跟踪建造人员,并得到固定时间和空间的感兴趣区域,其首帧初始化采用人为设定的检测框,并使用For循环以应对多目标情况。CHEN等<sup>[51-52]</sup>直接采用三维卷积网络对风险目标的时空特征进行提取,同时对时间和空间的变化进行建模,

有效地考虑时间上下文和空间的交互关系。SON等<sup>[53]</sup>提出 Siamese 网络对建造人员进行跟踪,网络使用深度学习基网络提取图像特征并学习给定图像的相似度,初始帧获得检测结果后,通过网络得到特征图后,再通过互相关层得到下一帧的得分特征图,当下一帧输入到网络中通过得分判断上下帧中风险目标的关系。

风险目标在二维平面内的运动可以通过平移和旋转来定义。当前的研究大多专注于风险目标的平移运动及尺寸的变化,无法对旋转运动进行跟踪。GUO 等<sup>[54]</sup>基于融合旋转检测框的风险目标位

置、尺寸和朝向辨识方法,提出了风险目标平移、尺寸变化和旋转运动的跟踪框架,实现了风险目标二维运动的完整跟踪。如图6所示,首先,利用旋转检测框的5个参数变量对风险目标的外包轮廓进行参数化,建立基于CNN的多层级特征的外包轮廓检测模型,并定义相应的损失函数;其次,将视频帧输入到该检测模型得到检测的外包轮廓,并建立风险目标的运动模型以得到预测的外包轮廓,利用检测外包轮廓对预测外包轮廓进行更新;最后,利用旋转检测框的交占比指标 IOU 来增加、保留或删除多个风险目标的跟踪 ID,最终得到每个风险目标的跟踪状态。

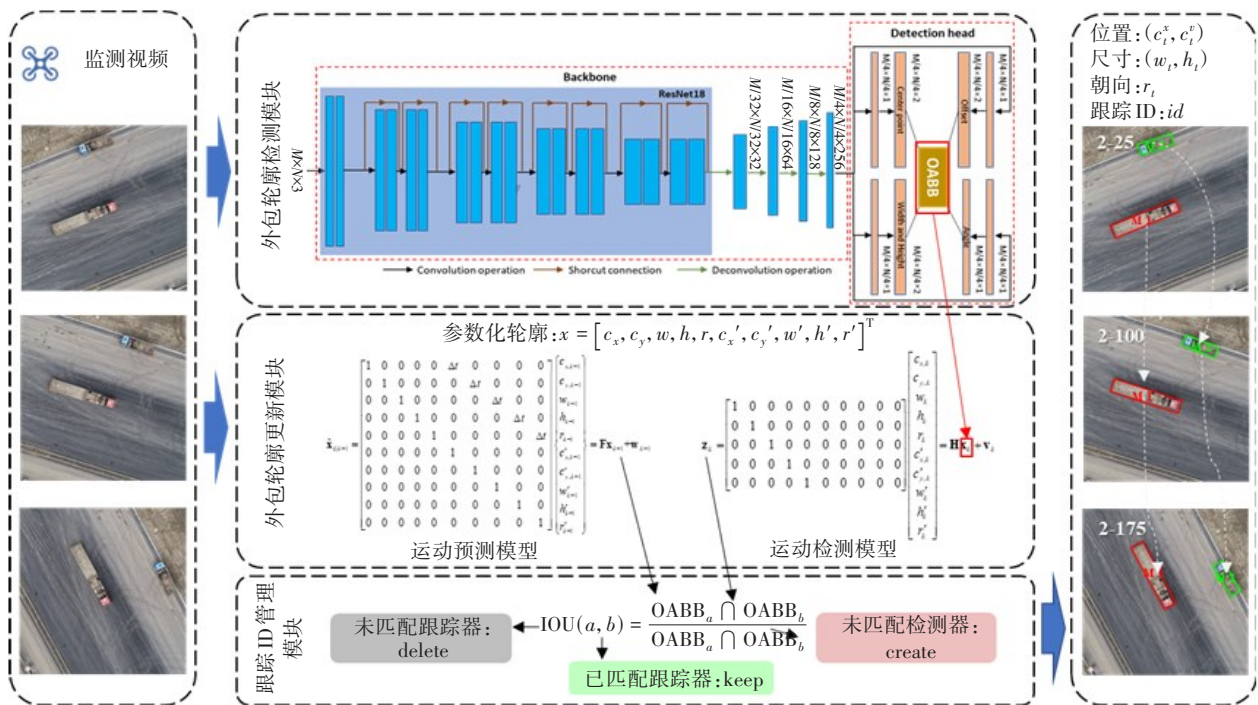


图6 风险目标平移、尺寸变化和旋转运动的跟踪框架

Fig. 6 Tracking framework of risk target translation, size change and rotation motion

### 3.2 风险目标三维跟踪

基于计算机视觉的公路智能建造风险目标三维跟踪是指通过多个二维数字图像获得时变的以实际物理尺寸为单位的的目标尺寸及轨迹。通过单张图像反演真实三维空间是病态问题,无法得到唯一解。因此,需要增加其他的约束条件方能通过图像计算出风险目标真实的三维尺寸或轨迹。研究者们通常采用基于单目视觉和标定的、基于双目视觉的以及基于深度相机的三种方式来增加约束条件。在单目视觉和标定方面,研究者大多通过标定板来获取相机的内部参数,通过在待测区域设置复

杂的已知尺寸的多个对比物来计算相机外部参数,从而实现待测区域风险目标的真实尺寸计算<sup>[55-56]</sup>。在双目视觉方面,研究者利用极线约束的先验条件,使用双目或多目视觉同时采集风险目标的二维图像,通过相机标定、相机姿态估计、二维跟踪及三角测量四步获取各类风险目标的点云数据,计算出三维尺寸和轨迹等信息<sup>[57-59]</sup>。在深度相机方面,不同于一般的RGB(red green blue)相机仅能获取真实三维空间二维投影后的数据,RGB-D(red green blue depth)相机可测量相机到待测目标的距离,从而可以直接获取三维空间点云数据。研究者针对



点云数据,开展了建造人员姿态三维仿真及点云对齐等工作<sup>[60-62]</sup>。

当前,对公路智能建造风险目标三维跟踪的研究大多基于复杂的双目、多目或者深度相机,硬件成本高且感知范围较为有限,而使用单目视觉的三维跟踪又需要复杂的标定过程,这在公路智能建造场地中缺乏大面积感知的可行性。LI等<sup>[63]</sup>抛开单目视觉复杂的标定过程,通过简单的场地设定为射影变换的求解提供了约束条件,建立了较大感知范围的风险目标三维尺寸和轨迹的高效感知框架,并分析了三维跟踪的准确性。如图7所示,该框架可以分为3个部分:①使用架设在合适位置的摄像机获取到的风险目标的二维图像,经过训练好的基于单阶段目标检测框架的风险目标二维定位方法处理后,可以得到风险目标的所在区域;②将建造场地图像进行HSV色彩空间转换后,对风险目标所在区域图像的饱和度通道进行形态学的运算处理,可以得到风险目标的精准轮廓;③基于针孔相机模型和射影变换模型,建立二维图像到三维空间坐标的映射关系,结合时间上下文信息,计算风险目标的精准轮廓对应的风险目标的三维几何参数、速度和朝向等信息。

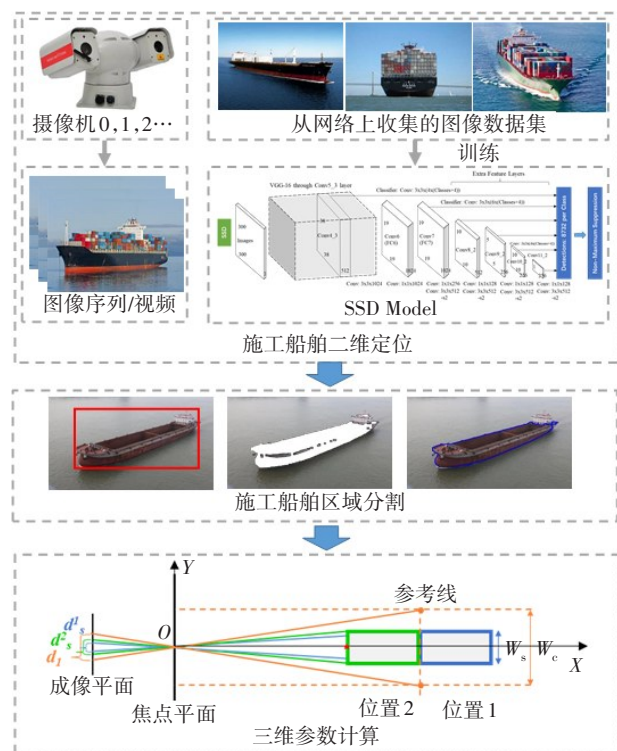


图7 风险目标三维尺寸和轨迹的高效跟踪框架

Fig. 7 Efficient tracking framework of risk target three-dimensional size and trajectory

## 4 结论

通过对近年来公路智能建造领域基于计算机视觉的风险目标辨识与跟踪等的研究进行分析,本研究从建造场地视觉重建、风险目标智能辨识和风险目标智能跟踪三个方面对国内外文献进行了梳理和分析,得出结论:

1) 由于无人机摄影的灵活性和有效性,基于计算机视觉的公路建造场地重建研究大多以无人机拍摄的多视角图像为数据载体,通过先进的图像拼接技术完成建造场地的二维重建,通过三维重建基本完成建造场地三维模型的建立。

2) 基于计算机视觉的风险目标智能辨识研究重点在于风险目标的位置、尺寸、朝向、姿态以及类别等方面,基于深度学习的辨识方法普遍基于传统机器学习的方法,深度辨识模型是该方向研究的主流。

3) 基于计算机视觉的风险目标智能跟踪算法的重点是在构建辨识结果时序上的变化模型,传统时序建模方法和基于深度学习的时序建模方法在二维跟踪上均有良好表现,通过增加约束条件来实现基于单目视觉的目标三维跟踪是可行的途径之一。

4) 基于计算机视觉的公路智能建造风险目标辨识与跟踪是公路建造“数字孪生”的重要内容和实践,对实现公路建造的信息化、智能化、无人化具有重要意义。然而,当前研究还存在一定的局限性:在建造场地的二维和三维视觉重建全流程的自动化程度上有待提高及风险目标的智能辨识和跟踪缺乏具有公信力的、超大规模的公路建造风险目标辨识训练和测试数据集,导致诸多算法缺乏统一的评价标准等问题。未来研究可着重构建建造场地的图像数据获取、预处理及高效重建的全流程框架,建立领域内统一的大规模数据集以及在三维智能跟踪框架中约束条件设置上进一步进行标准化处理。

## 参考文献(References):

- [1] CHEN K, DING L Y. Development of key domain-relevant technologies for smart construction in China[J]. Chinese Journal of Engineering Science, 2021, 23(4):



64. DOI: 10.15302/j-sscae-2021.04.007.
- [2] DING L. The connotation and framework of digital construction [J]. Construction Enterprise Management, 2022(2): 86-89.
- [3] BAO Y, Li H. Artificial intelligence for civil engineering [J]. China Civil Engineering Journal, 2019, 52(5): 1-11.
- [4] FANG W L, DING L Y, LOVE P E D, et al. Computer vision applications in construction safety assurance [J]. Automation in Construction, 2020, 110: 103013. DOI: 10.1016/j.autcon.2019.103013.
- [5] HEINRICH H W, PETERSEN D, ROOS N R, et al. Industrial Accident Prevention: A Safety Management Approach[M]. 5th ed. New York: McGraw-Hill, 1980
- [6] YANG J, PARK M W, VELA P A, et al. Construction performance monitoring via still images, time-lapse photos, and video streams: now, tomorrow, and the future[J]. Advanced Engineering Informatics, 2015, 29(2): 211-224. DOI: 10.1016/j.aei.2015.01.011.
- [7] XU S Y, WANG J, SHOU W C, et al. Computer vision techniques in construction: a critical review[J]. Archives of Computational Methods in Engineering, 2021, 28(5): 3383-3397. DOI: 10.1007/s11831-020-09504-3.
- [8] TEIZER J. Status quo and open challenges in vision-based sensing and tracking of temporary resources on infrastructure construction sites [J]. Advanced Engineering Informatics, 2015, 29(2): 225-238. DOI: 10.1016/j.aei.2015.03.006.
- [9] FANG W L, LOVE P E D, LUO H B, et al. Computer vision for behaviour-based safety in construction: a review and future directions[J]. Advanced Engineering Informatics, 2020, 43: 100980. DOI: 10.1016/j.aei.2019.100980.
- [10] KIM J. Visual analytics for operation-level construction monitoring and documentation: state-of-the-art technologies, research challenges, and future directions [J]. Frontiers in Built Environment, 2020, 6: 575738. DOI: 10.3389/fbuilt.2020.575738.
- [11] BANG S, KIM H, KIM H. UAV-based automatic generation of high-resolution panorama at a construction site with a focus on preprocessing for image stitching[J]. Automation in Construction, 2017, 84: 70-80. DOI: 10.1016/j.autcon.2017.08.031.
- [12] GUO Y P, XU Y, NIU H W, et al. Vision-based full-field panorama generation by UAV using GPS data and feature points filtering [J]. Smart Structures and Systems, 2020, 25: 631-641. DOI: 10.12989/SSS.2020.25.5.631.
- [13] MA Z L, LIU S L. A review of 3D reconstruction techniques in civil engineering and their applications[J]. Advanced Engineering Informatics, 2018, 37: 163-174. DOI: 10.1016/j.aei.2018.05.005.
- [14] NIU H W. Research on safety supervision method of traffic engineering construction site based on aerial image [D]. Harbin: Harbin Institute of Technology, 2019.
- [15] SUNG C, KIM P Y. 3D terrain reconstruction of construction sites using a stereo camera[J]. Automation in Construction, 2016, 64: 65-77. DOI: 10.1016/j.autcon.2015.12.022.
- [16] JIANG W G, ZHOU Y, DING L Y, et al. UAV-based 3D reconstruction for hoist site mapping and layout planning in petrochemical construction [J]. Automation in Construction, 2020, 113: 103137. DOI: 10.1016/j.autcon.2020.103137.
- [17] KIM S, KIM S, LEE D E. 3D point cloud and BIM-based reconstruction for evaluation of project by As-planned and As-built [J]. Remote Sensing, 2020, 12(9): 1457. DOI: 10.3390/rs12091457.
- [18] HU F Q, ZHAO J, HUANG Y, et al. Structure-aware 3D reconstruction for cable-stayed bridges: a learning-based method [J]. Computer-Aided Civil and Infrastructure Engineering, 2021, 36(1): 89-108. DOI: 10.1111/mice.12568.
- [19] JIANG Y H, BAI Y. Low - high orthoimage pairs-based 3D reconstruction for elevation determination using drone [J]. Journal of Construction Engineering and Management, 2021, 147(9): 1-21. DOI: 10.1061/(asce)co.1943-7862.0002067.
- [20] LIU Z G, KIM D, LEE S, et al. Near real-time 3D reconstruction and quality 3D point cloud for time-critical construction monitoring[J]. Buildings, 2023, 13(2): 464. DOI: 10.3390/buildings13020464.
- [21] ZOU J H, KIM H. Using hue, saturation, and value color space for hydraulic excavator idle time analysis[J]. Journal of Computing in Civil Engineering, 2007, 21(4): 238-246. DOI: 10.1061/(asce)0887-3801(2007)21:4(238).

- [22] DU S, SHEHATA M, BADAWY W. Hard hat detection in video sequences based on face features, motion and color information[C]//2011 3rd International Conference on Computer Research and Development. March 11-13, 2011, Shanghai. IEEE, 2011: 25-29. DOI: 10.1109/ICCRD.2011.5763846.
- [23] MEMARZADEH M, GOLPARVAR-FARD M, NIEBLES J C. Automated 2D detection of construction equipment and workers from site video streams using histograms of oriented gradients and colors [J]. Automation in Construction, 2013, 32: 24-37. DOI: 10.1016/j.autcon.2012.12.002.
- [24] XU Y Z, YU G Z, WANG Y P, et al. A hybrid vehicle detection method based on viola-jones and HOG + SVM from UAV images [J]. Sensors, 2016, 16 (8) : 1325. DOI: 10.3390/s16081325.
- [25] FANG W L, DING L Y, LUO H B, et al. Falls from heights: a computer vision-based approach for safety harness detection [J]. Automation in Construction, 2018, 91: 53-61. DOI: 10.1016/j.autcon.2018.02.018.
- [26] FANG W L, DING L Y, ZHONG B T, et al. Automated detection of workers and heavy equipment on construction sites: a convolutional neural network approach[J]. Advanced Engineering Informatics, 2018, 37: 139-149. DOI: 10.1016/j.aei.2018.05.003.
- [27] KIM H, KIM H, HONG Y W, et al. Detecting construction equipment using a region-based fully convolutional network and transfer learning[J]. Journal of Computing in Civil Engineering, 2018, 32 (2) : 1-15. DOI: 10.1061/(asce)cp.1943-5487.0000731.
- [28] SON H, CHOI H, SEONG H, et al. Detection of construction workers under varying poses and changing background in image sequences via very deep residual networks[J]. Automation in Construction, 2019, 99: 27-38. DOI: 10.1016/j.autcon.2018.11.033.
- [29] ROBERTS D, GOLPARVAR-FARD M. End-to-end vision-based detection, tracking and activity analysis of earthmoving equipment filmed at ground level [J]. Automation in Construction, 2019, 105: 102811. DOI: 10.1016/j.autcon.2019.04.006.
- [30] WANG Z C, ZHANG Q L, YANG B, et al. Vision-based framework for automatic progress monitoring of precast walls by using surveillance videos during the construction phase [J]. Journal of Computing in Civil Engineering, 2021, 35 (1) : 1-21. DOI: 10.1061/(asce)cp.1943-5487.0000933.
- [31] LUO X C, LI H, WANG H, et al. Vision-based detection and visualization of dynamic workspaces [J]. Automation in Construction, 2019, 104: 1-13. DOI: 10.1016/j.autcon.2019.04.001.
- [32] WU J X, CAI N, CHEN W J, et al. Automatic detection of hardhats worn by construction personnel: a deep learning approach and benchmark dataset [J]. Automation in Construction, 2019, 106: 102894. DOI: 10.1016/j.autcon.2019.102894.
- [33] NATH N D, BEHZADAN A H. Deep convolutional networks for construction object detection under different visual conditions [J]. Frontiers in Built Environment, 2020, 6: 97. DOI: 10.3389/fbuil.2020.00097.
- [34] WANG L, XIE L B, YANG P Y, et al. Hardhat-wearing detection based on a lightweight convolutional neural network with multi-scale features and a top-down module [J]. Sensors, 2020, 20 (7) : 1868. DOI: 10.3390/s20071868.
- [35] SHEN J, XIONG X, LI Y, et al. Detecting safety helmet wearing on construction sites with bounding-box regression and deep transfer learning [J]. Computer-Aided Civil and Infrastructure Engineering, 2021, 36 (2): 180-196. DOI: 10.1111/mice.12579.
- [36] GUO Y P, XU Y, LI S L. Dense construction vehicle detection based on orientation-aware feature fusion convolutional neural network [J]. Automation in Construction, 2020, 112: 103124. DOI: 10.1016/j.autcon.2020.103124.
- [37] GUO Y P, XU Y, NIU J, et al. Anchor-free arbitrary-oriented construction vehicle detection with orientation-aware Gaussian heatmap[J]. Computer-Aided Civil and Infrastructure Engineering, 2023, 38 (7) : 907-919. DOI: 10.1111/mice.12940.
- [38] TANG S, ROBERTS D, GOLPARVAR-FARD M. Human-object interaction recognition for automatic construction site safety inspection [J]. Automation in Construction, 2020, 120: 103356. DOI: 10.1016/j.autcon.2020.103356.
- [39] LUO X C, LI H, YANG X C, et al. Capturing and understanding workers' activities in far-field surveillance



- videos with deep action recognition and Bayesian nonparametric learning [J]. *Computer-Aided Civil and Infrastructure Engineering*, 2019, 34 (4) : 333-351. DOI: 10.1111/mice.12419.
- [40] KIM J, CHI S, CHOI M. Sequential pattern learning of visual features and operation cycles for vision-based action recognition of earthmoving excavators [C]// *Computing in Civil Engineering 2019*. Atlanta, Georgia. Reston, VA: American Society of Civil Engineers, 2019: 298-304. DOI: 10.1061/9780784482438.038.
- [41] LIANG C J, LUNDEEN K M, MCGEE W, et al. A vision-based marker-less pose estimation system for articulated construction robots [J]. *Automation in Construction*, 2019, 104: 80-94. DOI: 10.1016/j.autcon.2019.04.004.
- [42] LUO H, WANG M Z, WONG P K Y, et al. Full body pose estimation of construction equipment using computer vision and deep learning techniques [J]. *Automation in Construction*, 2020, 110: 103016. DOI: 10.1016/j.autcon.2019.103016.
- [43] CHEN C, ZHU Z H, HAMMAD A. Automated excavators activity recognition and productivity analysis from construction site surveillance videos [J]. *Automation in Construction*, 2020, 110: 103045. DOI: 10.1016/j.autcon.2019.103045.
- [44] JUNG S, JEOUNG J, KANG H, et al. 3D convolutional neural network-based one-stage model for real-time action detection in video of construction equipment [J]. *Computer-Aided Civil and Infrastructure Engineering*, 2022, 37(1): 126-142. DOI: 10.1111/mice.12695.
- [45] GUO Y P, CUI H T, LI S L. Excavator joint node-based pose estimation using lightweight fully convolutional network [J]. *Automation in Construction*, 2022, 141: 104435. DOI: 10.1016/j.autcon.2022.104435.
- [46] REZAZADEH AZAR E, MCCABE B. Part based model and spatial-temporal reasoning to recognize hydraulic excavators in construction images and videos [J]. *Automation in Construction*, 2012, 24: 194-202. DOI: 10.1016/j.autcon.2012.03.003.
- [47] XIAO B, LIN Q, CHEN Y. A vision-based method for automatic tracking of construction machines at nighttime based on deep learning illumination enhancement [J]. *Automation in Construction*, 2021, 127: 103721. DOI: 10.1016/j.autcon.2021.103721.
- [48] CAI J N, ZHANG Y X, YANG L, et al. A context-augmented deep learning approach for worker trajectory prediction on unstructured and dynamic construction sites [J]. *Advanced Engineering Informatics*, 2020, 46: 101173. DOI: 10.1016/j.aei.2020.101173.
- [49] BANG S, HONG Y J, KIM H. Proactive proximity monitoring with instance segmentation and unmanned aerial vehicle-acquired video-frame prediction [J]. *Computer-Aided Civil and Infrastructure Engineering*, 2021, 36(6): 800-816. DOI: 10.1111/mice.12672.
- [50] LUO X C, LI H, CAO D P, et al. Towards efficient and objective work sampling: recognizing workers' activities in site surveillance videos with two-stream convolutional networks [J]. *Automation in Construction*, 2018, 94: 360-370. DOI: 10.1016/j.autcon.2018.07.011.
- [51] CHEN C, ZHU Z H, HAMMAD A, et al. Vision-based excavator activity recognition and productivity analysis in construction [C]// *Computing in Civil Engineering 2019*. Atlanta, Georgia. Reston, VA: American Society of Civil Engineers, 2019: 241-248. DOI: 10.1061/9780784482438.031.
- [52] JUNG S, JEOUNG J, KANG H, et al. 3D convolutional neural network-based one-stage model for real-time action detection in video of construction equipment [J]. *Computer-Aided Civil and Infrastructure Engineering*, 2022, 37(1): 126-142. DOI: 10.1111/mice.12695.
- [53] SON H, KIM C. Integrated worker detection and tracking for the safe operation of construction machinery [J]. *Automation in Construction*, 2021, 126: 103670. DOI: 10.1016/j.autcon.2021.103670.
- [54] GUO Y P, XU Y, LI Z L, et al. Enclosing contour tracking of highway construction equipment based on orientation-aware bounding box using UAV [J]. *Journal of Infrastructure Preservation and Resilience*, 2023, 4: 1-11. DOI: 10.1186/s43065-023-00071-y.
- [55] FENG C, KAMAT V R, CAI H B. Camera marker networks for articulated machine pose estimation [J]. *Automation in Construction*, 2018, 96: 148-160. DOI: 10.1016/j.autcon.2018.09.004.
- [56] ZHAO J Y, HU Y B, TIAN M R. Pose estimation of excavator manipulator based on monocular vision marker system [J]. *Sensors*, 2021, 21 (13) : 4478. DOI: 10.3390/s21134478.

- dictionary and sparse regularization [J]. Structural Control and Health Monitoring, 2019, 26 (6) : 2355. DOI: 10.1002/stc.2355.
- [10] 林肯. 基于动力响应的位移影响线识别及桥梁损伤识别研究[D]. 兰州: 兰州理工大学, 2021. DOI: 10.27206/d.cnki.gsgsu.2021.000805.
- LIN Ken. Displacement influence line identification and bridge damage identification based on dynamic response [D]. Lanzhou: Lanzhou University of Technology, 2021.
- [11] 尚稳齐, 吴德义. 基于动力响应桥梁位移影响线识别方法分析[J]. 南阳理工学院学报, 2022, 14(4): 93-98. DOI: 10.16827/j.cnki.41-1404/z.2022.04.014.
- SHANG Wenqi, WU Deyi. Analysis of bridge displacement influence line identification method based on dynamic response [J]. Journal of Nanyang Institute of Technology, 2022, 14 (4) : 93-98. DOI: 10.16827/j.cnki.41-1404/z.2022.04.014.
- [12] ZHU J S, ZHANG C, LI X T. Structural damage detection of the bridge under moving loads with the quasi-static displacement influence line from one sensor [J]. Measurement, 2023, 211: 112599. DOI: 10.1016/j.measurement.2023.112599.
- [13] HONG Y H, KIM H K, LEE H S. Reconstruction of dynamic displacement and velocity from measured accelerations using the variational statement of an inverse problem [J]. Journal of Sound and Vibration, 2010, 329 (23): 4980-5003. DOI: 10.1016/j.jsv.2010.05.016.
- (责任编辑:欧兆虎;校对:罗容;英文编辑:陈璐)

(上接第11页)

- 10.3390/s21134478.
- [57] CHI S, CALDAS C H. Image-based safety assessment: automated spatial safety risk identification of earthmoving and surface mining activities [J]. Journal of Construction Engineering and Management, 2012, 138 (3) : 341-351. DOI: 10.1061/(asce) co. 1943-7862.0000438.
- [58] PARK M W, KOCH C, BRILAKIS I. Three-dimensional tracking of construction resources using an on-site camera system [J]. Journal of Computing in Civil Engineering, 2012, 26 (4) : 541-549. DOI: 10.1061/(asce)cp.1943-5487.0000168.
- [59] HAN S, LEE S. A vision-based motion capture and recognition framework for behavior-based safety management [J]. Automation in Construction, 2013, 35: 131-141. DOI: 10.1016/j.autcon.2013.05.001.
- [60] KHOSROWPOUR A, NIEBLES J C, GOLPARVAR-FARD M. Vision-based workplace assessment using depth images for activity analysis of interior construction operations [J]. Automation in Construction, 2014, 48: 74-87. DOI: 10.1016/j.autcon.2014.08.003.
- [61] KURIEN M, KIM M K, KOPSIDA M, et al. Real-time simulation of construction workers using combined human body and hand tracking for robotic construction worker system [J]. Automation in Construction, 2018, 86: 125-137. DOI: 10.1016/j.autcon.2017.11.005.
- [62] LEI L, ZHOU Y, LUO H B, et al. A CNN-based 3D patch registration approach for integrating sequential models in support of progress monitoring [J]. Advanced Engineering Informatics, 2019, 41: 100923. DOI: 10.1016/j.aei.2019.100923.
- [63] LI S L, GUO Y P, XU Y, et al. Real-time geometry identification of moving ships by computer vision techniques in bridge area [J]. Smart Structures and Systems, 2019, 23: 359.
- (责任编辑:欧兆虎;校对:李脉;英文编辑:刘至真)