

基于无人机的堤坝裂缝检测研究综述

边凯伦¹

1. 南昌大学 软件学院, 南昌

摘要: 近年来, 随着无人机技术和深度学习的迅猛发展, 基于无人机的堤坝裂缝检测技术得到了广泛关注。本文综述了基于无人机的堤坝裂缝检测方法、难点以及相关研究进展。本文首先梳理了经典目标检测算法的发展和特点, 分析了三类算法的优缺点。其次归纳了无人机堤坝裂缝检测的方法, 整理了公开数据集的信息和自制数据集的做法, 总结了目前无人机堤坝裂缝检测存在的问题。然后综述了近些年关于无人机堤坝裂缝检测的研究进展, 通过对各改进方法分类, 阐述了各方法的特点和贡献。最后对未来该研究的发展方向进行了展望。

关键词: 目标检测; 无人机; 裂缝检测; 堤坝; 深度学习

1 引言

随着社会经济发展, 为满足防洪、发电、航运和灌溉的需要, 我国修建了大量的水利工程, 堤坝是其中重要的建筑物。目前, 我国已建成各类水库9.8万余座、总库容8983亿立方米, 修建的各类河流堤坝超过43万公里[1]。由于结构变形、地震、水流冲击等因素造成的堤坝异常, 如裂缝、锈蚀、剥落等[2], 会严重影响堤坝的安全。国内外已出现多起由于裂缝影响堤坝正常使用的案例。例如位于瑞士西南的泽乌齐尔拱坝, 在极限水位时坝体底部出现约15m的裂缝。位于老挝的桑片-桑南内水电站在溃坝前已发现多处裂缝。在我国, 洞庭湖一线提防由于长期遭受植被和昆虫侵蚀, 出现多处孔隙, 最终导致决堤。就连三峡大坝自本世纪初以来也已发现80余条裂缝[3]。裂缝产生后若不及时干预, 则会破坏堤坝的稳定结构, 造成溃坝, 严重危及周边人民群众生命安全, 给经济发展带来重大损失。

因此, 建立高效精准的堤坝异常检测体系, 定期对堤坝进行裂缝检测, 对堤坝的安全使用具有重要意义。堤坝裂缝检测通常采用传统的人工

目测方法或利用接触式测量仪器进行测量[4]。这些方法存在检测效率低、时效性差、人工成本高、危险系数大[5]等问题。

随着智慧水利、数字孪生流域[6]建设工作的开展以及无人机、深度学习等技术的发展, 研究人员提出利用无人机并结合目标检测技术对堤坝进行裂缝检测, 并取得了一些研究成果[7]。近年来, 基于无人机的堤坝异常检测技术愈加完善, 但针对此技术的综述性文献较少。本文通过阐述现有的基于深度学习的目标检测算法, 分析不同算法的原理和优劣; 通过介绍无人机目标检测的方法和特点, 讨论利用无人机检测堤坝裂缝的难点; 然后, 归纳该研究的最新进展, 阐明各难点的解决方法; 最后, 对全文进行总结, 并对基于无人机的堤坝裂缝检测的发展方向进行了展望。

2 基于深度学习的目标检测算法

在过去的二十年里, 目标检测的发展经历了两个历史时期: 2014年之前的传统目标检测时期和2014年之后的基于深度学习的检测时期[8], 如图1所示。而目前基于深度学习的目标检测算

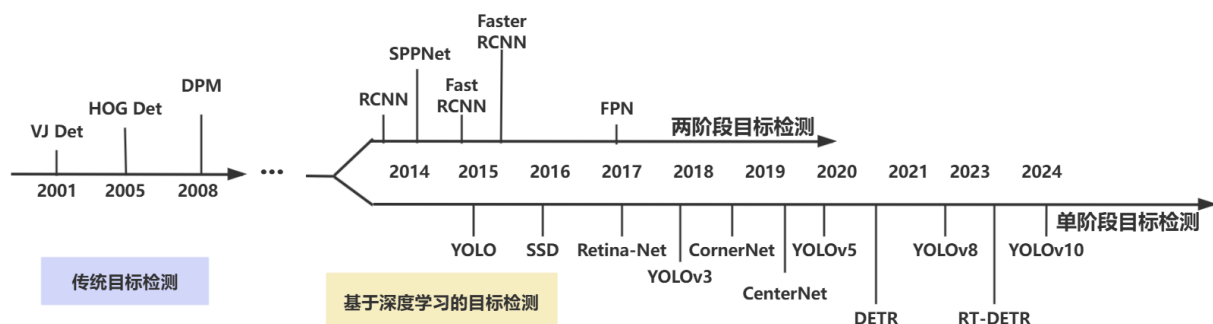


图 1 目标检测发展历史

法根据结构和检测流程可以分为两类：两阶段目标检测和单阶段目标检测。除此之外，随着Vision Transformer的发展，基于Transformer的目标检测算法也得到了许多关注。表1对比了各目标检测算法。

2.1 两阶段目标检测

基于两阶段的目标检测算法又称为基于候选区域的目标检测方法。图2为两阶段目标检测算法的结构。两阶段目标检测算法将检测任务分为两部分，首先利用预选网络检测对象的位置生成候选区域，然后再将所获得的候选区域输入到卷积神经网络中进行特征提取，实现对象分类与定位。

2014年，Girshick等人将候选区域和CNN结合起来，提出了RCNN算法，显著提高了目标检测的准确性[9]。但在大量重叠的候选区域上进行冗余计算导致检测速度极慢。同年He等人引入空间金字塔池化SPP层，提出了SPPNet[10]有效提高了检测速度。2015年，Girshick在R-CNN和SPPNet的基础上使用兴趣区域池化，提出了FastRCNN[11]。Ren等人引入了区域生成网络

RPN，代替了选择性搜索算法，提出了FasterRCNN算法[12]，大幅提高检测效率。后来，研究人员又提出了更多改进方法，包括FPN[13]、RFCN[14]和Light head RCNN[15]等。

两阶段检测算法可以获得较高的准确性和鲁棒性，但检测速度较慢，难以满足无人机实时检测的要求。

2.2 单阶段目标检测

单阶段目标检测算法直接对整张图像进行处理，预测目标的类别和位置，运行速度快，大大增强了检测的实时性。图3为单阶段目标检测算法的结构。

2015年，Joseph等人提出YOLO算法[16]。它将目标检测任务定义为端到端的回归问题，通过单个网络同时预测目标的位置和类别，提升了计算速度。但与两阶段目标检测算法相比，它的检测精度有所下降，特别是对小目标的检测。YOLO的后续版本[17, 18, 19]和SSD算法[20]针对此问题进行改进，提高了小目标的检测精度。2020年，Ultralytics团队开发了YOLOv5算法[21]，引入新的网络架构和新的训练方法，提高了算法的性能。2023年，Ultralytics团队发布了新一代实时目标检测模型YOLOv8[22]，采用先进的骨干网络和颈部架构，设计了无锚点分割头部，保持了精度和速度之间的平衡。最近，YOLO的最新版本YOLOv10被提出[23]，清华大学的研究人员引入了一种新的实时目标检测方法，解决了YOLO以前版本在后处理和模型架构方面的不足。通过消除非极大值抑制（NMS），优化各种模型组件，YOLOv10在显著降低计算开销的同时实现了最优

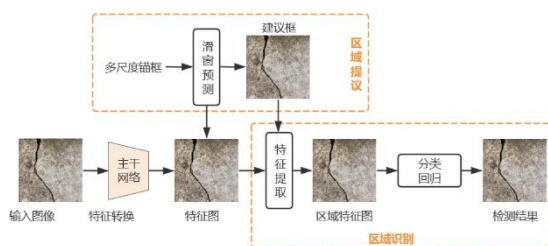


图 2 两阶段目标检测流程

表 1 各目标检测算法性能对比						
	模型	提出年份	主干网络	mAP@0.5:0.95 / %	mAP@0.5 / %	FPS
两阶段 目标检测	RCNN*	2014	AlexNet	---	58.5	~0.02
	SPPNet*	2014	ZFNet	---	59.2	~0.23
	Fast RCNN	2015	AlexNet VGG-16	19.7	35.9	~0.43
	Faster RCNN	2015	ResNet-50	27.2	48.4	5
	FPN	2017	ResNet-101	33.9	56.9	5
单阶段 目标检测	YOLO*	2016	GoogleNet	---	57.9	45
	SSD	2016	VGG-16	28.8	48.5	46
	Retina-Net	2018	ResNet-101+FPN	37.8	57.5	12
	YOLOv3	2018	DarkNet-19	33.0	57.9	20
	CornerNet	2018	Hourglass-104	42.2	57.8	4.1
	CenterNet	2019	Hourglass-104	47.0	64.5	7.8
	YOLOv5	2020	CSPDarknet53+PAN+FPN	37.4	56.8	140
	DETR	2020	ResNet-101	42.0	62.4	28
	YOLOv8	2023	CSPDarknet-53	44.9	---	150
	RT-DETR	2023	ResNet-50	53.1	71.3	108
	YOLOv10	2024	CSPNet	46.3	---	93.5

带*的模型在PASCAL VOC 2007数据集上测试，其余在MS COCO数据集测试。该表来源于相关作者实验数据。

的性能。在速度和精度方面优于大多数现有的目标检测器。从YOLO到YOLOv10，单阶段目标检测方法不断升级，推动着目标检测领域的发展。

单阶段目标检测算法的实时性契合了实际生产需求，因此得到了更加广泛的应用。在基于无人机的堤坝裂缝检测任务中，大多数研究人员多数也采用了此类检测方法。

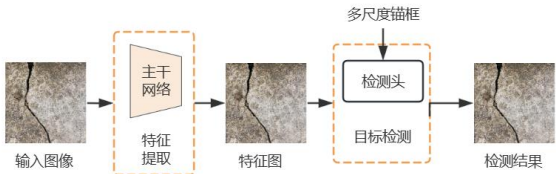


图 3 单阶段目标检测流程

2.3 基于Transformer的目标检测

在目标检测领域，单阶段目标检测算法和两阶段目标检测算法一直是研究的热点。近年来，最初应用于自然语言处理任务中的Transformer方法被成功引入到了计算机视觉领域，为目标检测任务提供了新思路。

Transformer方法由Google团队于2017年提出[24]。为了克服CNN的局限性，Transformer算法抛弃了传统的卷积核，采用注意力机制来动态

地关注图像中的不同区域，从而有效地捕捉目标的全局上下文信息。2020年，Carion等人提出了DETR[25]，他们将目标检测任务视为一个集合预测问题，提出了一种基于transformer的端到端的检测网络，此方法可以不使用锚框或锚点。后来，Zhu等人提出了Deformable DETR[26]。解决了DETR收敛时间长、检测小目标性能有限的问题。2023年，百度团队基于DETR推出了第一个实时端到端目标检测器RT-DETR[27]，通过引入高效混合编码器，解耦尺度内交互和跨尺度融合，高效处理多尺度特征，在保持高精度的同时提供了实时性能。

目前已有许多研究将Transformer引入到无人机堤坝裂缝检测任务中，增强了裂缝位置和形态的特征提取能力，提升了目标检测的精确性。

3 无人机检测堤坝裂缝的方法及难点

随着无人机技术的迅速发展，基于无人机的目标检测技术开始应用于各个领域，逐渐成为复杂环境中检测任务的重要方法，推动了目标检测技术的进一步发展。

3.1 无人机检测方法

目前已存在许多利用无人机进行堤坝裂缝检

测的方法,大致可分为以下几个步骤:

(1) 飞行任务规划。越来越多的无人机已具备自主起降、路径规划的能力。通过无人机控制软件,结合堤坝位置范围等信息进行航线规划。

(2) 堤坝数据采集传输。在任务执行过程中,通过控制摄像头云台转向调整合适的角度和焦距,以获取最佳图像质量。通过传输模块完成图像的高速传输。

(3) 数据处理与裂缝检测。针对无人机获取的图像信息,经过预处理后,通过训练好的目标检测模型进行裂缝检测,获取异常信息。

与传统方法相比,使用无人机进行堤坝检测具有显著优势。首先,无人机覆盖范围广,可以替代检测员进入难以到达和危险的区域,提高了检测的安全性;其次,无人机可以更快地完成堤坝的巡检任务,提高了工作效率;此外,实时数据传输和处理便于检测员及时做出决策,避免出现更大的危害。

3.2 无人机检测数据集

3.2.1 公开数据集

近年来,基于无人机航拍影像的目标检测已成为研究热点,相关的数据集也在快速地更新之中。然而公开的基于无人机拍摄的堤坝裂缝数据集相对较少,但考虑到大部分堤坝为混凝土材质,与混凝土路面、桥梁、房屋等材质相同,由于雨水冲刷、动植物侵蚀所产生的裂缝具有一定参考价值,可使用这类公共数据集进行模型训练[28]。这些公共数据集已经包含了大量的图像和注释信息,可以用于算法的开发和验证。使用公共数据集可以节省数据收集和标注的时间和成本,同时具有较高的数据质量和可靠性。

SDNET2018数据集[29]包含超过56000张混凝土桥面、墙壁和路面的裂缝图像。裂缝窄至0.06毫米,宽至25毫米。该数据集还包含了阴影、表面颗粒、缩放、孔洞等干扰信息。

CRACK500数据集[30]由天普大学研究人员制作,共有500幅裂缝图像及其对应的标签,每幅图像分辨率为2000×1500。每个裂缝图像都有一个像素级的带注释的二值图。

CrackForest数据集[31]主要由城市道路裂

缝图像组成,共329张来自城市不同道路的裂缝图像和相应的裂缝标注信息。另外这些图像中也包含了阴影、油斑、水渍等噪声信息。

CCIC数据集[32]用于混凝土裂缝分类任务。该数据集是从METU大学校园建筑收集的,总共40000张图像,图像分辨率为227×227。

Deepcrack数据集[33]包含多个场景中捕获的537张路面裂缝图像,以及相应的像素级二进制标签。图像的分辨率为544×384。

Crack-detection数据集[34]由6077张桥梁裂缝图像组成,分辨率为224×224。该数据集是由原始数据集通过数据预处理,滤波去除模糊图像,最后随机选择得到的。

Concret crack数据集[35]由SEU-Cracks团队制作,包含了来自不同类型建筑的3107张裂缝图像。该数据集提供了详细的裂缝标注,包括裂缝的形状和位置。所有图像的像素分辨率为512×512。

3.2.2 自制数据集

目前公开的裂缝数据集多为人工手持相机拍摄,与无人机拍摄图像不同;且堤坝裂缝受水流侵蚀、渗透,与路面、房屋等建筑裂缝仍有一定区别。所以公开数据集存在目标尺度不同、背景变化不同、裂缝形态不同等问题。在利用无人机进行堤坝裂缝目标检测的实际研究中,研究人员多数使用自制数据集。自制方法归纳如下:

首先将堤坝区域栅格化,进行区域划分,对每个区域采用全覆盖路径规划算法(弓字形牛耕法或回字路径法)进行路径规划。

在采集过程中,控制无人机匀速定时拍摄。拍摄角度可采取平视、斜视和俯视,拍摄距离为3~20m,拍摄环境应包含强光和暗光。

将获取的原始图像进行裁剪和筛选,在缩放过程中尽量损失更少的信息以适应目标检测器的要求。由于大部分坝面没有裂缝或损伤,因此需由堤坝结构工程师或专家进行识别判断,按照标准水坝裂缝判别依据,将裂缝分为多个类别。数据集应合理安排有裂缝和无裂缝图像的比例。

之后使用LabelImg等标注程序对裂缝损伤进行标注,获得VOC、COCO或YOLO格式的标注数据。在注释、训练和测试之后,应由人工进行结果评

估,防止出现标记错误等情况。

3.3 无人机堤坝裂缝检测中的难点问题

利用无人机进行堤坝裂缝检测虽然有许多优点,但与自然场景下的目标检测不同,依然存在较多需要解决的难点问题。

(1) 小目标问题:无人机影像中裂缝的尺度范围大,裂缝在图像中占据较小的尺寸。裂缝在图片中占据的像素低,导致其自身分辨率低,特征少,检测模型难以提取小目标特征。

(2) 大视场问题:无人机的探测范围较广,且不受地理因素等限制,因此得到的图像视场往往很大。大视场下的目标检测面临着目标分布不均、目标稀疏等问题。无人机采集到的图片中存在大量背景噪声,会造成误检。并且因背景的变化可能会出现目标遮挡或光线过暗、过亮以及曝光异常,对裂缝检测造成干扰,进一步减少裂缝的细节特征,导致检测准确率降低。

(3) 数据样本不足:尽管基于深度学习的方法在目标检测任务中表现出了令人印象深刻的性能,但对带注释的训练图像的大型数据集的需求限制了深度神经网络的可行性。目前基于深度学习的堤坝裂缝检测方法严重依赖于大量的堤坝裂缝图像,公开的裂缝数据图像的缺乏仍然是堤坝表面裂缝检测的挑战。

(4) 实时检测问题:目前大多数检测方法是离线的。通过无人机设备采集数据后,导入计算终端进行分析。这样导致时效性差,资源消耗大。同时对数据传输也提出了很高的要求。目前大多数目标检测模型在计算能力低的嵌入式终端设备上运行速度缓慢,难以部署到边缘设备对堤坝裂缝进行实时检测。

4 基于无人机的堤坝异常检测研究进展

针对上节提到的问题,目前已有许多研究人员提出了解决办法。本节通过分析国内外的研究进展,将这些方法分为小目标检测精度化、检测模型轻量化、少样本利用高效化三类。

4.1 小目标检测精准化

4.1.1 数据增强

在裂缝检测任务中,数据增强是一种增大数据集规模、减轻模型过拟合的有效方法,可以提高模型的检测精度。对于裂缝来说,边缘信息是其重要的特征,因此神经网络应该专注于裂缝边缘的学习。大部分研究仅对数据进行传统的灰度处理和滤波去噪,不能很好地突出裂纹边缘特征。

Wang等人[36]提出一种先验信息强化方法。通过三种裂缝边缘信息图像的叠加来增强先验信息。与传统的滤波去噪、灰度化等强化相比,裂缝的姿态和位置被突出,且在水平方向裂缝趋势稳定,降低了模式识别的难度,有利于训练模型的收敛,可以获得更好的网络训练效果。

在使用数据增强策略时需要根据裂缝特性合理地选择,不当的数据增强策略会引入新的噪声,影响检测模型的性能。

4.1.2 多尺度融合

由于无人机拍摄的特点,裂缝的尺度会根据无人机高度的不同而出现“近大远小”的多尺度特征。为使模型更好地识别不同大小、形状的裂缝,可采用多尺度特征融合的方法。通过提取和整合不同尺度下的图像特征,改善对多尺度裂缝的识别能力。

Zhang等人[37]对YOLOv8模型进行改进,在颈部引入三层的PAFPN结构,并在颈部上下分支的每一层都加入了具有大感受野和低参数量的三明治融合模块。在骨干网中,使用RepVGG模块代替卷积层作为下采样层。这种结构将高层特征和底层特征相结合,增强了模型学习多尺度特征的能力。

李姝等人[38]也对YOLOv8进行改进,在预处理层使用HGStem模块代替Conv模块,通过HGStem的多层卷积和池化操作,提取并融合多尺度特征,可以有效保留并丰富小目标特征信息。

Chen等人[35]改进传统的两级融合,采用多层融合,并使用了三种结构相似的域分解机(FFM)模型。与空间金字塔池化(ASPP)不同,FFM的输入来自不同模块,可以有效提取融合多尺度特征。

虽然多尺度特征融合可以融合无人机拍摄的裂缝的更多特征信息,但这也增加了模型的计算成本,且在特征融合过程中不能剔除噪声干扰。

4.1.3 多模态融合

将不同传感器或模态的图像信息进行融合，可以获得更全面、准确的信息。

红外图像记录了堤坝裂缝的热辐射信息，可以检测可见光中难以察觉的目标。方卫华等人[39]将光学图像与红外图像相结合，提高了模型的检测能力。同时为解决红外图像成像质量低、轮廓模糊等问题，他们采取了基于绝对偏差的中值(MAD)滤波算法和广义形态学算法对红外图像进行预处理，提高了裂缝识别的准确度。

Duan等人[40]利用傅里叶变换，将原始图像和其空域转换的频域，分别输入到两个特征提取模块中，分别提取高级特征并融合，最后根据融合后的特征对裂缝的纹理和轮廓所产生的高低频信号进行区分，从而实现裂缝类别的分类。利用空域和频域的互补关系，使模型具有较高的检测精度。

Wang等人[41]利用北斗卫星导航系统的精确定位功能，结合激光雷达等多种传感器，获得高精度的堤坝裂缝位置和属性数据，通过多种模态数据融合，提升了堤坝裂缝检测算法的检测精度。

4.1.4 引入注意力机制

注意力机制在自然语言处理、计算机视觉和其他领域中得到了广泛的应用。通过模仿人类视觉注意力的方法，增强对数据的关注程度，使模型能够在复杂环境中更有效地处理信息。

吕伏等人[42]基于YOLOv8s模型，在坐标注意力基础上引入三维多分支坐标注意力机制，增加通道维度的信息交互，扩展分支的拆分融合，提高模型全局特征提取能力。在不增加参数量和模型大小的情况下提升检测准确率和速度，使模型的检测性能和资源占用达到平衡。

Zhu等人[43]利用残差网络和信道注意策略，将自适应注意机制模块嵌入到骨干网络中。该模块将单元输入特征图的详细信息补充到输出特征图中，以延缓详细信息丢失的趋势。通过注意机制模块与DeeplabV3+骨干网相结合，可以显著提高模型识别小尺度裂缝的精度。

王伟等人[44]为提高模型对堤坝裂缝的关注度，将GAM注意力模块嵌入到YOLOv5的颈部网络中。使模型能够增强全局跨维度的交互，抵御弥

散信息。

CBAM也是一种注意力机制，与GAM不同的是，它结合了通道注意力和空间注意力两个模块。方卫华等人[39]在YOLOv5的骨干网络中融合CBAM注意力模块，分别进行通道和空间维度上的注意力特征融合，减少参数。使模型在精度提高的同时实现了轻量化。

类似的，黄贲等人[45]改进了YOLOX网络，在主干网络和加强特征提取网络的输出特征图中加入了轻量级CBAM，对主干网络特征图输出和加强网络特征图输出进行注意力权重分配，提高网络对裂缝特征的关注度。

4.1.5 级联检测

级联神经网络通常由多个独立的子网络组成，每个子网络负责完成特定的任务或提取特定的特征。通过将多个子网络串联起来，级联神经网络可以逐步提取和精炼输入数据的特征，从而提高整体模型的性能。

Wang等人[36]提出了一种基于先验信息强化的混凝土裂缝深度学习级联检测算法。该算法分为两个阶段：第一阶段，使用AlexNet网络对背景和裂缝进行分类；第二阶段，使用参数量较多的Inceptionv2网络对裂缝进一步识别。他们通过实验证明了级联神经网络在复杂背景下可以获得理想的检测效果。

4.1.6 图像分割

目标检测和图像分割是计算机视觉中两个重要的基本任务。通过语义分割或实例分割可以将裂缝从复杂背景中分割出来，降低背景对目标检测的影响。目前已有许多研究利用无人机对堤坝裂缝进行分割，达到了不错的精度。因此可以通过学习语义分割损失来改进目标检测。

Feng等人[46]基于SegNet进行改进，提出了一种基于深度卷积网络的像素级裂缝检测方法；Chen等人[47]设计了一个浅编码网络来提取裂缝图像的特征；李怡静等人[48]在U2-Net中融合深度可分离卷积和扩张卷积，扩大了原有模型的感受野，增强了对微小裂缝的检测能力；Zhao等人[49]将线性谱聚类(LSC)超像素分割与增强的GrabCut相结合，融合Sobel边缘算子和GrabCut，

可以有效地对高分辨率无人机图像进行智能分割和堤坝裂缝提取；陈荣敏等人[50]采用ResNet-152作为骨干网络搭建模型提取裂缝特征，并设计新的解码网络层实现了裂缝的像素分割检测。

4.2 检测模型轻量化

4.2.1 知识蒸馏

知识蒸馏是一种模型轻量化技术。通常是将大型教师模型的知识浓缩到小型学生模型以实现模型轻量化。学生模型可以借助教师模型的知识训练，能够在保持相对较高性能的同时减少参数数量并降低计算复杂度。

目前大多数高精度的目标检测算法过于复杂，无法做到实时检测。为了解决这一问题，Zhang等人[51]提出了一种基于改进残差神经网络的动态知识蒸馏混凝土大坝裂缝检测算法。与以往使用固定蒸馏温度的知识蒸馏系统不同，他们在训练过程中通过增加一个对抗性的动态温度模块来动态调节温度，使网络能够在每个训练轮次自动选择合适的蒸馏温度。

动态蒸馏可以帮助模型提高提取裂缝纹理和轮廓等特征的能力，并缓解少样本数据集过拟合的问题。同时实现了模型的轻量化，以便在边缘设备上部署和推理。

4.2.2 模型剪枝

王伟等人[44]将YOLOv5s主干网络替换为改进后的轻量化PMobileNetV2模块，原来的标准卷积被深度可分离卷积替换，减少了模型的计算量和参数量，提高了检测速度。另外他们将卷积核个数减少一倍，同时精简最后一层的结构。通过实验发现检测速度得到提升，并且检测精度几乎不变。

吕伏等人[42]用SPD-Conv替换YOLOv8s的部分标准卷积，在下采样时有效地保留了更多特征信息并加快了推理速度；在C2f模块中采用了更高效的Fast-DBB_Bottleneck模块，结合PConv与DBB结构重参数化叠加，以进一步降低模型计算量；最后通过引入PG-Detect检测头，显著减少计算量，提高模型的计算效率。

李姝等人[38]对YOLOv8骨干网络的通道宽度

进行消减，减少了网络参数；同时删除了对小目标敏感度较低的检测层，进一步减少了模型的参数量和计算量。

Chen等人[35]设计的CrackSeU模型引入了一系列高效和轻量化的设计。例如在特征提取的第一阶段采用最大池化操作，使特征图的空间分辨率减为原来的一半，有效地减少了模型的参数数量和计算消耗。

4.2.3 深度可分离卷积

深度可分卷积是标准卷积的改进，通过拆分空间维度和通道维度的相关性，减少了卷积计算所需要的参数个数，提升了卷积核参数的使用效率，从而达到模型的轻量化。

李姝等人[38]用深度可分离卷积替代YOLOv8中标准卷积，通过逐通道卷积和逐点卷积，实现特征提取和特征降维，降低了模型的复杂度，减少了参数量和计算量，实现了模型的轻量化。

Zhu等人[43]使用exception骨干搭建网络，由于其内部卷积核采用深度可分离卷积，且卷积核的破孔率可自行设定，从而实现了更少的参数，提高了模型裂缝检测的速度。

Chen等人[35]受MobileNetV1启发，在FFM中引入了两个深度可分离卷积来细化特征，降低了模型参数和计算复杂度。

4.3 少样本利用高效化

4.3.1 图像生成

Huang等人[52]提出一种解决堤坝裂缝数据不足的方法。具体做法是先采集少量堤坝裂缝的图像，训练并优化基于无监督的图像转换模型，进而生成新的裂缝样本。通过分类、目标检测和语义分割等任务验证，证明了此方法生成的图像在训练和优化裂缝检测模型方面的潜力。

Xu等人[53]开发了一种基于虚拟场景仿真和图像处理的合成数据生成器，通过少量样本生成大量标记的堤坝表面裂缝数据。这些合成数据与公开的混凝土裂缝图像相结合，用于训练模型，与仅使用混凝土裂缝图像相比，总体裂缝检测平均精度提高了29.2%。

当真实世界的数据集过少或不可用，且收集

和标注真实数据成本过高时，这些生成数据的方法将发挥重要作用，可以解决数据集不足的问题。

4.3.2 迁移学习

迁移学习是解决计算机视觉中少样本训练任务的有效方法。神经网络可以从一个任务中学习知识，并将这些知识应用到另一个独立的任务中。可以提取两个不同但相关的任务之间的共同特征，实现两个任务之间学习能力的相互转移。相比从头开始训练权重，利用迁移学习可以缩短训练时间，提高训练效率。此方法可以使模型在数据量很少的情况下取得可观的性能，达到少样本利用高效化的目的。

Zhang等人[51]采用参数/模型传递法。首先，在mini-ImageNet上训练了一个性能优良的检测模型；然后，冻结其结构和参数，用作裂缝检测任务中特征提取的骨干结构；最后，对一些重要的超参数进行微调，将全连通层的结构修改为双分类器进行裂缝检测，从而将mini-ImageNet的多分类问题转化为两分类问题，识别出背景和裂缝。他们利用迁移学习解决了无人机堤坝裂缝检测任务中样本数量不足的问题。

Li等人[54]提出一种基于两阶段迁移学习的混凝土坝面裂缝智能检测方法。通过跨域和域内两阶段迁移学习将相关领域知识转移到目标领域，学习共享参数和特征。该方法缓解了少样本训练不足和过拟合的问题，减少了裂缝漏检情况，并能在较短的训练周期内达到较高的准确率，同时具有更强的鲁棒性。

李怡静等人[48]通过迁移学习，使目标域共享源域模型参数。利用混凝土裂缝开源数据集和少量无人机堤坝裂缝样本，实现了无人机影像中堤坝裂缝的特征学习和检测任务。既提高了检测精度又大大降低了训练成本。

表2为上述研究进展的归纳汇总。

5 未来研究展望

目前，基于无人机的堤坝裂缝检测任务取得了不错的进展，部分算法虽然具有不错的检测效果，但仍存在较大的改进空间。结合前述问题及研究进展，对该领域未来的发展趋势进行展望，期望可以提供有价值的参考：

（1）实时检测和边缘计算

由于无人机硬件的限制，难以在低功耗的边缘设备部署大量的计算资源。因此需要轻量化的目标检测算法来完成。同时由于实时检测任务的需要，要求轻量化模型具备较快的推理速度，以适用算力受限的无人机平台。但模型复杂度和参数的减少会影响模型的精度，特别是检测裂缝这类小目标时。

因此如何平衡速度与精度，设计一个轻量且高效的模型，实现实时检测和边缘计算，成为无人机目标检测任务的一项挑战。

（2）多模态数据融合

除了常规的视觉单模态图像数据外，可以借助无人机搭载红外摄像头、激光雷达等传感器，采集获取多模态数据。多模态数据融合可以获得更丰富的特征信息，进一步提升堤坝裂缝检测的准确性和鲁棒性。因此，未来基于多模态融合的无人机堤坝裂缝检测技术将会受到更多关注。

（3）高效的注意力模块

随着ViT算法的发展，在堤坝裂缝检测任务中可以通过基于Transformer的编码器和解码器来排除外部复杂环境特征的干扰，让检测过程专注于裂缝等目标。Transformer与多模态数据融合还可以进一步增强细节特征处理能力。目前，注意力机制主要分为通道注意力与空间注意力。如何设计更加有效且轻量化的注意力模块仍值得进一步研究。

（4）数据集获取与标注

目前还没有公开的无人机堤坝裂缝数据集，通过生成对抗网络或大语言模型等方法生成无人机堤坝裂缝样本以扩充数据集，获取成本低、效果好的大量数据具有较好的发展前景。除此之外可以自行创建数据集，但应避免单一类别或单一背景下所获取的图像，应使用多种环境下不同种类的综合样本，以提升无人机在阴雨天气或恶劣条件下的检测能力。

对于数据的标注，可使用大模型等方法进行自动标注，提高标注效率和准确性。另外当标注成本过大时，可采用无监督或半监督训练模式，减少对标注数据的依赖。

数据集的优劣决定了模型训练结果的好坏。为了提高模型的特征提取能力，提升无人机堤坝

表 2 基于无人机的堤坝裂缝检测研究进展				
研究方法	基础检测算法	方法策略	效果	
小目标检测 精准化	数据增强	AlexNet	通过裂缝边缘信息叠加来增强先验信息	优于传统方法，mAP达93.42%
		YOLOv8	引入三层PAFPN结构，加入三明治融合结构	准确率提高13.4%
	多尺度融合	YOLOv8	HGStem代替常规卷积模块	准确率提高5.7%
		UNet	多层融合、使用FFM模型	准确率达到86.39%
	多模态融合	YOLOv5	融合红外图像与可视化图像	提高了检测精度
		CNN	空域和频域信息相结合	分类准确率达99.5%
		PointCloud	结合北斗、激光雷达等信息	裂缝提取精度提升
	引入注意力机制	YOLOv8	引入三维多分支坐标注意力机制	准确率达65.5%，超过DETR
		Xception	注意力机制模块与DeeplabV3+骨干网相结合	准确率达到90.54%
		YOLOv5	颈部嵌入GAM注意力模块	准确率提高1.8%
		YOLOv5	骨干网络融合CBAM注意力模块	小模型检测精度提高
		YOLOX	骨干网络加入轻量化的CBAM模块	准确度提高1.86%
	级联检测	AlexNet	通过AlexNet和Inceptionv2两级神经网络检测	准确率和召回率达到了93.42%和99.60%
	图像分割	---	利用图像分割优化目标检测	降低了复杂背景的影响
检测模型 轻量化	知识蒸馏	ResNet	加入对抗性的动态知识蒸馏模块	模型参数减少，准确率提高3%
	模型剪枝	YOLOv5	主干网改为轻量化的PmobileNetv2	参数减少4.1M，FPS提高17.7
		YOLOv8	SPDCov代替常规卷积，改进C2f模块	参数量减为1/4，FPS达到289
		YOLOv8	对主干网通道宽度削减	参数量降低49.8%
		UNet	特征提取第一阶段采用池化操作	模型参数减少一半
	深度可分离卷积	YOLOv8	逐通道卷积和逐点卷积	体积降低48%，GFLOPs降低17%
		DeeplabV3+	采用空洞自定义的DSC卷积	模型参数减少
		UNet	在FFM中引入两个DSC卷积	降低了模型参数和计算量
少样本利用 高效化	图像生成	Pix2Pix	提出一种无监督图像转换算法	各模型准确率均超过60%
		YOLOX	利用虚拟场景仿真技术合成图像	检测精度提高29.2%
	迁移学习	ResNet	采用参数/模型传递法，多分类改为双分类	准确率达到了98.39%
		UNet	采用两阶段迁移学习	准确率超过85.7%
		U2-Net	在少量无人机堤坝裂缝样本上使用迁移学习	少样本准确率达87.87%

裂缝检测的准确性，建立大规模的公开数据集对研究人员具有重要意义。

（5）大模型的引入

随着大语言模型的发展，视觉大模型得到了广泛关注。目前已出现了Grounding DINO、Co-DETR、GLIP等目标检测大模型。利用大模型可对数据自动化预标注、根据标签提示进行目标检测。未来若能将多模态大语言模型运用到图像增强、特征提取中，同时提高大模型检测速度，达到实时检测的效果，将进一步推动目标检测领域的发展。

6 结 论

堤坝裂缝检测是保障堤坝正常工作、流域安全稳定的重要任务。基于无人机的堤坝裂缝检测精准高效，相关研究已取得一定成果。本文首先介绍了基于深度学习的目标检测算法的发展和原理，然后归纳了基于无人机的堤坝裂缝检测的方法，整理了公开数据集的信息和自制数据集的方法，讨论了目前基于无人机的堤坝裂缝检测存在的问题。紧接着综述了近些年关于无人机堤坝裂缝检测的研究，按照改进方法的类型进行分类，

阐述了各方法的特点和贡献。最后,对未来该领域的发展方向进行了展望。目前,无人机技术和深度学习技术正处于快速发展时期,基于无人机的堤坝裂缝检测技术有望得到进一步发展,并应用于实际工作中。

参 考 文 献

- [1] 苏璇.水利部:全国已建成各类水库9.8万多座修建各类河流堤防43万公里.央视网, 2021.
- [2] 钮新强.大坝安全诊断与加固技术[J]. 水利学报, 2007 (S1): 60-64.
- [3] 魏永晖.三峡枢纽工程卓越的筑坝实践[J]. 中国三峡, 2013, (03): 70-4.
- [4] 黄朝君,杨小云,夏杰.丹江口初期工程大坝上游面水上裂缝检查与处理[J]. 人民长江, 2015, 46(6): 45 48, 74.
- [5] Chen D, Huang B, Kang F. A review of detection technologies for underwater cracks on concrete dam surfaces[J]. *Applied Sciences*, 2023, 13(6): 3564.
- [6] 黄艳. 数字孪生长江建设关键技术与试点初探[J]. 中国防汛抗旱, 2022, 32(2): 16-26.
- [7] Chen D, Huang B, Kang F. A review of detection technologies for underwater cracks on concrete dam surfaces[J]. *Applied Sciences*, 2023, 13(6): 3564.
- [8] Zou Z, Chen K, Shi Z, et al. Object detection in 20 years: A survey[J]. *Proceedings of the IEEE*, 2023, 111(3): 257-276.
- [9] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014: 580-587.
- [10] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2015, 37(9): 1904-1916.
- [11] Girshick R. Fast r-cnn[C]//*Proceedings of the IEEE international conference on computer vision*. 2015: 1440-1448.
- [12] Ren S, He K, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks[J]. *Advances in neural information processing systems*, 2015, 28.
- [13] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017: 2117-2125.
- [14] Dai J, Li Y, He K, et al. R-fcn: Object detection via region-based fully convolutional networks[J]. *Advances in neural information processing systems*, 2016, 29.
- [15] Li Z, Peng C, Yu G, et al. Light-head r-cnn: In defense of two-stage object detector[J]. *arXiv preprint arXiv:1711.07264*, 2017.
- [16] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.
- [17] Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *ArXiv*, abs/1804.02767.
- [18] C. Gao, Q. Cai and S. Ming, "YOLOv4 Object Detection Algorithm with Efficient Channel Attention Mechanism," 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE), Harbin, China, 2020, pp. 1764-1770, doi: 10.1109/ICMCCE51767.2020.00387.
- [19] J. Redmon and A. Farhadi, "YOLO9000: Better, Faster, Stronger," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 6517-6525, doi: 10.1109/CVPR.2017.690.
- [20] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//*Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I* 14. Springer International Publishing, 2016: 21-37.
- [21] Zhang Y, Guo Z, Wu J, et al. Real-time vehicle detection based on improved yolo v5[J]. *Sustainability*, 2022, 14(19): 12274.
- [22] Wang G, Chen Y, An P, et al. UAV-YOLOv8: A small-object-detection model based on improved YOLOv8 for UAV aerial photography scenarios[J]. *Sensors*, 2023, 23(16): 7190.
- [23] Wang A, Chen H, Liu L, et al. Yolov10: Real-time end-to-end object detection[J]. *arXiv preprint arXiv:2405.14458*, 2024.
- [24] Vaswani A, Shazeer N, Parmar N, et al. Attention is all

-
- you need[J]. Advances in neural information processing systems, 2017, 30.
- [25] Carion N, Massa F, Synnaeve G, et al. End-to-end object detection with transformers[C]//European conference on computer vision. Cham: Springer International Publishing, 2020: 213-229.
- [26] Zhu X, Su W, Lu L, et al. Deformable detr: Deformable transformers for end-to-end object detection[J]. arXiv preprint arXiv:2010.04159, 2020.
- [27] Zhao Y, Lv W, Xu S, et al. Detrs beat yolos on real-time object detection[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024: 16965-16974.
- [28] 李怡静, 程浩东, 李火坤, 等. 基于改进 U2-Net 与迁移学习的无人机影像堤防裂缝检测[J]. 水利水电科技进展, 2022, 42(6): 52-59.
- [29] Dorafshan S, Thomas R J, Maguire M. SDNET2018: An annotated image dataset for non-contact concrete crack detection using deep convolutional neural networks[J]. Data in brief, 2018, 21: 1664-1668.
- [30] Yang F, Zhang L, Yu S, et al. Feature pyramid and hierarchical boosting network for pavement crack detection[J]. IEEE Transactions on Intelligent Transportation Systems, 2019, 21(4): 1525-1535.
- [31] Shi Y, Cui L, Qi Z, et al. Automatic road crack detection using random structured forests[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(12): 3434-3445.
- [32] Özgenel Ç F, Sörgüç A G. Performance comparison of pretrained convolutional neural networks on crack detection in buildings[C]//Isarc. proceedings of the international symposium on automation and robotics in construction. IAARC Publications, 2018, 35: 1-8.
- [33] Liu Y, Yao J, Lu X, et al. DeepCrack: A deep hierarchical feature learning architecture for crack segmentation[J]. Neurocomputing, 2019, 338: 139-153.
- [34] Xu H, Su X, Wang Y, et al. Automatic bridge crack detection using a convolutional neural network[J]. Applied Sciences, 2019, 9(14): 2867.
- [35] Chen W, He Z, Zhang J. Online monitoring of crack dynamic development using attention-based deep networks[J]. Automation in Construction, 2023, 154: 105022.
- [36] Wang Q, Li B, Deng Z, et al. Dam crack detection based on deep learning cascade detection algorithm[C]//2023 3rd International Conference on Neural Networks, Information and Communication Engineering (NNICE). IEEE, 2023: 131-134.
- [37] Zhang Z. Drone-YOLO: an efficient neural network method for target detection in drone images[J]. Drones, 2023, 7(8): 526.
- [38] 李姝, 李思远, 刘国庆. 基于YOLOv8无人机航拍图像的小目标检测算法研究[J/OL]. 小型微型计算机系统, 1-13[2024-08-06].
- [39] 方卫华, 杨浩东, 谢双双, 等. 面向堤坝监测的无人机自动巡检系统设计[J]. 江苏水利, 2024, (02): 1-4. DOI:10.16310/j.cnki.jssl.2024.02.002.
- [40] Duan C, Zhang J, Sun W, et al. A Crack Classification Method for Dam Image Based on Double-input Neural Network[C]//2023 IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI). IEEE, 2023: 1742-1748.
- [41] Wang J, Tan C, Hu T, et al. UAV Dam Crack Detection System Based on Beidou and LIDAR[C]//China Satellite Navigation Conference. Singapore: Springer Nature Singapore, 2023: 3-13.
- [42] 吕伏, 傅宇恒, 贺丽娜, 等. 三维多层次特征协同的无人机遥感目标检测算法[J]. 计算机科学与探索, 2024, 18(05): 1301-1317.
- [43] Zhu Y, Tang H. Automatic damage detection and diagnosis for hydraulic structures using drones and artificial intelligence techniques[J]. Remote Sensing, 2023, 15(3): 615.
- [44] 王伟, 雷斌. 基于改进YOLOv5s的无人机路面裂缝检测算法[J/OL]. 农业装备与车辆工程, 1-11[2024-08-06]. <https://kns-cnki-net.wvpu.ncu.edu.cn/kcms/detail/37.1433.TH.20240722.1922.004.html>.
- [45] 黄贲, 康飞, 唐玉. 基于目标检测的混凝土坝裂缝实时检测方法[J]. 清华大学学报(自然科学版), 2023, 63(07): 1078-1086.
- [46] Feng C, Zhang H, Wang H, et al. Automatic pixel-level crack detection on dam surface using deep convolutional network[J]. Sensors, 2020, 20(7): 2069.
- [47] Chen B, Zhang H, Li Y, et al. Quantify pixel-level

-
- detection of dam surface crack using deep learning[J]. Measurement Science and Technology, 2022, 33(6): 065402.
- [48] 李怡静,程浩东,李火坤,等.基于改进U-2-Net与迁移学习的无人机影像堤防裂缝检测[J].水利水电科技进展,2022,42(06):52-59.
- [49] Zhao H, Yin X, Li A, et al. Intelligent Segmentation and Change Detection of Dams Based on UAV Remote Sensing Images[J]. Remote Sensing, 2023, 15(23): 5526.
- [50] 陈荣敏,王皓冉,汪双,等.基于无人机的坝面裂纹缺陷智能检测方法[J].水利水电科技进展,2021,41(06):7-12.
- [51] Zhang J, Bao T. An improved resnet-based algorithm for crack detection of concrete dams using dynamic knowledge distillation[J]. Water, 2023, 15(15): 2839.
- [52] Huang B, Kang F, Li X, et al. Underwater dam crack image generation based on unsupervised image-to-image translation[J]. Automation in Construction, 2024, 163: 105430.
- [53] Xu J, Yuan C, Gu J, et al. Innovative synthetic data augmentation for dam crack detection, segmentation, and quantification[J]. Structural Health Monitoring, 2023, 22(4): 2402-2426.
- [54] Li J, Lu X, Zhang P, et al. Intelligent Detection Method for Concrete Dam Surface Cracks Based on Two-Stage Transfer Learning[J]. Water, 2023, 15(11): 2082.