

电力设施多模态精细化机器人巡检关键技术及应用

张辉^{1,2} 杜瑞^{1,2} 钟杭^{1,2} 曹意宏^{1,2} 王耀南^{2,3}

摘要 电力设施巡检对于加快电网基础设施智能化改造和智能微电网建设,提高电力系统互补互济和智能调节能力的需求具有重要作用.近年来,智能巡检机器人开始在电力巡检中广泛应用,在提高电力设施巡检效率和准确性、提升安全性、降低成本和促进电力智能化发展等方面发挥关键作用.本文从电力巡检机器人的智能感知和导航技术出发,重点介绍目标检测、语义分割、自主导航等共性关键技术的国内外发展现状.然后以可见光红外双光融合、可见光图像和点云数据融合、声纹和可见光融合为例,阐述电力场景多模态数据融合方式.并进一步介绍电力部件精准分割和异物检测、线路点云杆塔倾斜检测、输电线路覆冰多模态检测和电力架空线路缺陷分析及台账异常检测等电力设施多模态机器人相关案例.最后探讨电力设施多模态精细化机器人巡检关键技术的发展趋势和所面临的挑战.

关键词 电力设施巡检, 机器人, 智能感知, 多模态, 输电线路

引用格式 张辉, 杜瑞, 钟杭, 曹意宏, 王耀南. 电力设施多模态精细化机器人巡检关键技术及应用. 自动化学报, 2025, 51(1): 20-42

DOI 10.16383/j.aas.c230809

CSTR 32138.14.j.aas.c230809

The Key Technology and Application of Multi-modal Fine Robot Inspection for Power Facilities

ZHANG Hui^{1,2} DU Rui^{1,2} ZHONG Hang^{1,2} CAO Yi-Hong^{1,2} WANG Yao-Nan^{2,3}

Abstract Power facilities inspection plays an important role in accelerating the intelligent transformation of power grid infrastructure and the construction of intelligent microgrid, and improving the complementary and intelligent adjustment ability of power system. In recent years, intelligent inspection robots have been widely used in power inspection. They play a key role in improving the efficiency and accuracy of power facilities inspection, improving safety, reducing costs and promoting the development of power intelligence. This paper starts from the intelligent perception and navigation technology of power inspection robots. This paper focuses on the development status of common key technologies such as target detection, semantic segmentation and autonomous navigation at home and abroad, and then takes visible light infrared dual-light fusion, visible light image and point cloud data fusion, voice-print and visible light fusion as examples. This paper expounds the multi-modal data fusion method of power scene, and further introduces the related cases of multi-modal robots for power facilities, such as accurate segmentation and foreign body detection of power components, tilt detection of line point cloud towers, multi-modal detection of transmission line icing, defect analysis of power overhead lines and abnormal detection of ledgers, etc.. Finally, the development trend and challenges of key technologies for multi-modal fine robot inspection of power facilities are discussed.

Key words Power facilities inspection, robots, intelligent perception, multi-modal, transmission lines

Citation Zhang Hui, Du Rui, Zhong Hang, Cao Yi-Hong, Wang Yao-Nan. The key technology and application of multi-modal fine robot inspection for power facilities. *Acta Automatica Sinica*, 2025, 51(1): 20-42

收稿日期 2023-12-31 录用日期 2024-05-30

Manuscript received December 31, 2023; accepted May 30, 2024

科技创新 2030——“新一代人工智能”重大项目 (2021ZD0114503), 国家自然科学基金重大研究计划 (92148204), 国家自然科学基金 (62027810), 湖南省科技创新领军人才 (2022RC3063), 湖南省十大技术攻关项目 (2024GK1010), 湖南省重点研发计划 (2023GK2068, 2022GK2011), 国网湖南省电力有限公司科技项目 (5216A522001Y, 5216A5240003), 国家电网有限公司科技项目 (5700-202423229A-1-1-ZN) 资助

Supported by National Key Research and Development Program of China (2021ZD0114503), Major Research Plan of National Natural Science Foundation of China (92148204), National Natural Science Foundation of China (62027810), Hunan Leading Talent of Technological Innovation (2022RC3063), Top Ten

Technical Research Projects of Hunan Province (2024GK1010), Key Research and Development Program of Hunan Province (2023GK2068, 2022GK2011), Science and Technology Project of State Grid Hunan Electric Power Company Limited (5216A522001Y, 5216A5240003), and Science and Technology Project of State Grid Corporation of China (5700-202423229A-1-1-ZN)

本文责任编辑 刘华平

Recommended by Associate Editor LIU Hua-Ping

1. 湖南大学机器人学院 长沙 410082 2. 机器人视觉感知与控制技术国家工程研究中心 长沙 410082 3. 湖南大学电气与信息工程学院 长沙 410082

1. School of Robotics, Hunan University, Changsha 410082
2. National Engineering Research Center of Robot Visual Perception and Control Technology, Changsha 410082
3. College of Electrical and Information Engineering, Hunan University, Changsha 410082

电力巡检在保障电力系统稳定运行中具有重要作用. 如图 1 所示, 全国的用电量呈现出稳步上升的态势, 电力巡检对于确保我国面对用电量持续增长挑战时的电力供应安全、稳定和高效至关重要. 随着电力行业加速转型升级, 为了使巡检工作变得更加高效和准确, 应对人工巡检耗时耗力且难以到达高山江河等巡检盲区的挑战^[1-2], 智能巡检机器人在电力巡检中得到广泛应用. 电力巡检机器人具有提高巡检效率、准确性、安全性, 降低成本和促进电力智能化发展的关键作用.

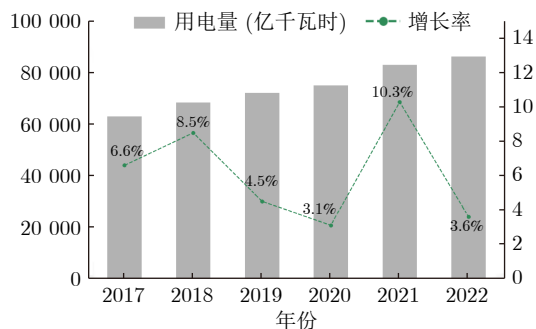


图 1 近年全国总用电量趋势 (单位: 亿千瓦时)

Fig. 1 Trend of national total electricity consumption in recent years (unit: 100 million kW·h)

电力设施精细化多模态机器人巡检是一项集成多种感知技术、机器人技术和人工智能的综合性技术, 旨在提高电力系统的运行安全、稳定性和效率. 该领域涵盖了地面机器人、空中机器人 (无人机)、自主水下机器人等多种类型的机器人系统, 它们通过集成视觉、声音、热成像等多模态感知技术, 对电力设施进行自动化、智能化的检测与评估. 世界各国高度重视电力巡检机器人的发展, 例如美国自上世纪 80 年代起开始研发电力巡检机器人, 并逐步推广应用, 而欧盟和日本也在近年推出了促进电力行业中人工智能应用的具体目标和计划. 我国也发布了《智能能源行动计划》和《“十四五”通用航空发展专项规划》, 提出了加快电力行业智能化改造与推进能源数字化转型的目标和措施, 其中包括智能电网建设、能源大数据平台建设等.

智能感知是实现电力设施巡检的基础关键技术, 可以嵌入到电力设施巡检机器人中, 以实现电力设备的定位和分类, 这可以进一步被认为是机器人传感、自动导航和自主检查的基础^[3-4]. 智能感知通过激光雷达、可见光相机、多光谱相机和红外相机等成像设备进行数据采集, 并对采集到的目标场景数据进行语义理解和状态感知, 记录当前场景的巡检结果. 例如, 变电站红外图像精准检测任务中^[5], 需要对无人机采集到的变电站电流互感器、断路器

和隔离开关等部件进行检测. 在准确得到电力部件在图像中的位置信息后, 利用温度解译对相关电力部件进行分析, 根据是否出现温度异常判断电力设施是否存在潜在过热等危险. 通过这种巡检方式, 可以在出现故障前提前感知潜在风险, 及时采取相应措施, 避免过热产生电力故障甚至严重的火灾, 有效防范了电力故障, 减少了人工巡检的主观干扰, 提升了巡检准确性, 降低了人力成本. 目前, 深度学习和大数据技术的快速发展使得电力场景的智能巡检得到广泛应用. 智能巡检前期通过机器人摄像头、激光雷达等设备采集数据, 并送入目标检测或语义分割模型进行训练, 训练好的模型就可以在很多巡检任务上取得显著性优越结果. 同时, 在成本和安全性能上也实现了对人工巡检的压倒性胜利.

然而, 在实际应用中, 复杂多变的环境给机器人带来了语义理解的难度, 并且由于数据采集本身的局限性, 单一模态的数据难以对巡检场景实现精准数据表征, 巡检结果往往较为粗糙. 例如, 仅使用可见光模态的情况下, 恶劣的天气、较差的照明或低对比度会降低视觉传感器的图像质量, 而杂乱的背景、极端的遮挡或密集的物体可能会降低基于视觉图像的语义分割方法的性能^[6]; 而只使用红外模态完成场景理解后进行温度解译, 会由于红外图像的低分辨率导致电力部件检测的精度不高, 影响后续温度分析的准确性; 对于点云模态, 在杆塔或导线检测任务中, 仅有坐标信息而缺乏颜色和纹理信息会由于难以区分背景与前景而影响检测性能. 因此, 开展电力设施的多模态精细化智能化机器人巡检势在必行. 通过多模态数据融合, 电力巡检机器人可以利用不同数据模态间的互补性, 构建出更精准、更全面、更一致的精细化智能感知模型, 提升特征表达的信息量和判别性, 为电力业务实施提供更加准确、高效、灵活的业务辅助报告.

1 电力设施多模态机器人巡检研究现状

20 世纪 80 年代, 日本学者开始了变电站巡检机器人的研究工作, 并取得了有价值的研究成果. 日本中部电力公司研制尺寸为宽 1.3 m、长 1.4 m、高 1.7 m 的地下管道巡检机器人^[7-8], 对 275 kV 的地下管网内的多种模态数据进行巡检, 包括温度、湿度、水位、超声和彩色视频图像等, 实时反映地下管道的运行情况. 这款机器人通过电池供电, 能够连续工作 2.5 小时, 并且能在变电站设定的充电点进行充电. 20 世纪 90 年代, 日本四国电力公司和东芝公司等研究机构设计了一种适用于 500 kV 变电站的巡检机器人 BIG MOUSE^[9], 这种机器人携带温

度计和摄像机来检测设备是否正常运行,但需要预先在变电站安装轨道。21 世纪初期,加拿大魁北克水电研究院 (IREQ) 开发了一款配备四轮的检测机器人^[10],如图 2(a)所示。该机器人整合了红外热成像和可见光成像系统来巡查变电站的设备,它还配备了一个后台监控系统,允许操作员能够实时从电脑后台监控机器人的工作状态。然而,在那个时期,由于 SLAM (Simultaneous localization and mapping) 技术和多传感器数据融合导航技术还未充分发展,这款机器人主要依赖遥控操作,限制了它的自主操作能力。在此之后,新西兰研制的电力巡检机器人开始采用 GPS 定位,在语音交互和激光避障等功能上实现了多模态信息处理。



(a) 魁北克水电研究院变电站巡检机器人
(a) IREQ substation inspection robot



(b) 国家电网红外全方位视觉导航巡检机器人
(b) State Grid infrared omnidirectional visual navigation inspection robot

图 2 电力设施多模态机器人巡检案例

Fig. 2 Multi-modal robot inspection cases of power facilities

2002 年,国家高新技术研究发展计划(863 计划)开始支持输配电机器人技术方面的研发,包括其原型机的研发和工业应用的开发。国家电网开发了为 500 kV 变电站设计的移动机器人^[11],如图 2(b)所示,该机器人装备可见光和红外摄像头以及方向性麦克风。国家电网电力机器人实验室的官方网站回顾了这些得到国家研发计划支持的机器人项目的

发展历程。自 2005 年起,第一代国家电网智能巡检机器人开始被用于巡检任务。国家电网智能巡检机器人前部配有两个独立的驱动轮,后部装有两个全向轮。自 2008 年以来,国家电网智能巡检机器人开始在 Linux 操作系统下使用软件架构控制,这一架构提升了机器人的性能和稳定性,使得实现复杂的控制算法成为可能。国家电网智能巡检机器人被设计用于在变电站内进行巡检任务^[12-13],负责收集关键的检查数据。这些数据随后被传送至数据中心,由专门的分析团队负责处理和分析。为了有效导航并执行任务,国家电网智能巡检机器人配备了多种导航技术,包括磁性导航、视觉导航以及 GPS 与航位推算组合的方式,从而确保在各种环境下均能高效、准确地完成预定路径的巡检。在到达预设的停止点时,机器人会停下并旋转其搭载的云台,以捕捉并记录图像和温度信息。这些关键信息随即被传回数据中心,用于进一步的分析,以确保变电站的安全和高效运行。

早期的电力设施机器人巡检主要集中在变电站,依赖于固定轨道的导航系统,对变电站设备的多源信息进行监测。随着导航技术、SLAM 技术和机器视觉的发展,电力设施机器人巡检研究重点也趋向于无轨化智能巡检。例如飞行机器人根据杆塔的位置规划飞行路径^[14],需要事先为多个塔杆目标的巡检顺序规划出一条或多条最优路径,在避障的前提下,解决多目标巡检问题。此外,视觉感知作为电力人工智能的主要分支,为机器人多模态巡检奠定深厚的技术基础,感知算法的优劣也成为了巡检机器人的研究重点。为此,本文将综述电力设施多模态精细化机器人巡检中的导航和感知关键技术及相关应用。首先,本文将从电力场景的 2D 视觉感知和 3D 感知入手,分别总结机器人巡检在 2D 和 3D 场景下典型的目标检测和语义分割任务;随后,介绍高精度定位导航和 SLAM 技术对机器人巡检的重要性;之后,阐述多模态机器人巡检中不同数据模态间的融合方法,为多模态机器人巡检做出更完备的数据表示;接下来,分析电力设施多模态巡检的相关实施案例,以及在电力设施精细化机器人巡检中起到的重要作用;最后,列举发展趋势和挑战,为后续多模态精细化机器人巡检提供清晰的研究思路。

2 电力设施巡检感知和导航关键技术

智能机器人巡检感知技术是指机器人搭载摄像头和计算机等硬件设备,通过图像获取和图像处理与分析来感知周围环境。图像获取利用摄像头等设

备获取周围环境的图像信息, 而图像处理与分析则是对这些信息进行处理和分, 提取出有用的特征和信息. 其中, 目标检测和语义分割是图像处理与分析的主要任务, 旨在帮助巡检机器人更加准确地识别目标物体, 例如设施缺陷和损坏等, 从而提高巡检的准确性和效率. 如表 1 和图 3 所示, 现有的巡检场景主要包括 2D 和 3D 场景.

2.1 电力设施机器人巡检 2D 感知技术

2.1.1 2D 目标检测

2D 目标检测作为机器人视觉感知的核心任务, 旨在从二维图像中识别并定位各类物体, 精确地确定其位置和尺寸. 适用于电力设施机器人巡检中精准定位目标设施或缺陷的位置. 早年基于机器学习方法的 2D 目标检测算法, 需要手工提取目标特征, 然后训练分类器对目标进行分类和定位, 如支持向量机^[15] (SVM)、随机森林^[16] (RF) 等. 早在 2005 年, 何洪英等^[17] 就基于支持向量机对高压绝缘子污秽等级进行红外热像检测. 这些算法通常受到手工特征提取的限制, 难以适应复杂的场景和不同尺度的物体. 近年来, 深度学习的发展为 2D 目标检测带来了革命性的进展, 标志性的算法如图 4 所示. 在此基础上, 以 Transformer^[18] 为基本组件的 DETR 系列方法^[19-22] 也成为了现有主流目标检测研究趋势. 在具体技术方面, 基于 Transformer 和卷积神经网络 (CNN) 的方法对图像进行特征提取, 并在特征图上进行检测和定位. 常见的基于深度学习的 2D 目标检测算法大体上分为两种:

1) 两阶段目标检测算法: 两阶段目标检测算法通过分为生成候选框和目标分类与定位的双阶段流程, 实现精确的物体检测. 此类算法首先识别出图像中潜在目标的候选区域, 随后对这些区域进行分类和位置精调. 代表性的两阶段算法包括基于区域的卷积神经网络系列, 如 R-CNN^[23]、Fast R-CNN^[24]、Faster R-CNN^[25] 等, 以及基于金字塔特征的级联卷积神经网络^[26] 算法, 如 Cascade R-CNN^[27] 等.

2) 一阶段目标检测算法: 一阶段目标检测算法直接从输入图像提取特征, 通过单个卷积神经网络

预测目标位置和类别. 这种算法以简洁的网络结构和高速的检测能力为优, 尽管在精度上可能略有不足. 一阶段算法中, 基于锚点的方法尤为流行, 它在各个位置生成多种尺寸和比例的预定义锚点, 通过分类和回归进行目标检测. 常见的一阶段目标检测算法包括 YOLO^[28] (You only look once)、SSD^[29] (Single shot detector)、RetinaNet^[30] 等.

现有的 2D 目标检测技术在电力设施巡检上已经开展了一些研究. 例如, Yi 等^[31] 提出一种新颖的补丁自纹理学习网络 (PSTL-Net), 通过结合自纹理学习模块 (STLM) 和补丁感知空间注意模块 (PSAM), 可以有效地实现输电线路的检测过程, 在振动锤、鸟巢类、缺失绝缘子和悬垂线夹等目标上表现良好; Liu 等^[32] 提出一种用于电力线组件检测的基于图的关系引导网络, 该网络利用区域、图像和类别的相关性, 实现了电力线组件的检测. 虽然 2D 目标检测技术手段也趋近于成熟, 但在电力设施巡检上仍然存在一些限制. 例如, 在输电线路长尾分布下对于小样本类别检测效果不佳, 高分辨率下难以准确感知细微缺陷 (如销钉缺失), DETR 等先进模型受计算资源限制难以部署到巡检机器人或无人机上.

2.1.2 2D 语义分割

相较于目标检测的对象识别和定位, 语义分割侧重于对图像进行像素级别的分类和细致分割, 旨在识别图像中所有区域的类别. 这一技术在自动驾驶、医疗影像处理等领域有着广泛应用. 如表 2 所示, 2D 语义分割主要分为传统方法和深度学习方法. 传统 2D 语义分割包括基于阈值的方法^[33]、基于边缘检测的方法^[34] 和基于图割的方法^[35]. 基于阈值的方法通过设置阈值将图像中像素的亮度或颜色分为不同的类别, 然而, 对于具有复杂背景和变化光照的图像, 这种方法具有局限性. 基于边缘检测的方法通过检测图像中物体的边缘来分割不同的区域, 该方法依赖于边缘的准确性, 当图像中存在较多的干扰线条或背景噪声时, 该方法效果不佳. 基于图割的方法通过将图像分割成图结构, 把分割问题转化为最小割问题, 该方法对于物体形状复杂、

表 1 电力设施机器人巡检智能感知关键技术

Table 1 Key technologies for intelligent perception in robotic inspection of power facilities

场景	技术	电力检测任务	代表方法
2D	目标检测	常用于可见光模态, 输出电力设备缺陷的位置和类型信息	Faster R-CNN ^[25] 、YOLO 系列 ^[28] 、SSD ^[29] 、RetinaNet ^[30] 等
	语义分割	常用于红外模态, 输出电力设备缺陷的热故障区域	U-Net ^[36] 、SegNet ^[37] 、DeepLab 系列 ^[38-39] 等
3D	目标检测	常用于激光点云, 检测电力导线三维位置	PointNet 系列 ^[40-42] 、SESS ^[43] 等
	语义分割	常用于激光点云, 输出杆塔点区域	PointNet 系列 ^[40-42] 、DGCNN ^[44] 等

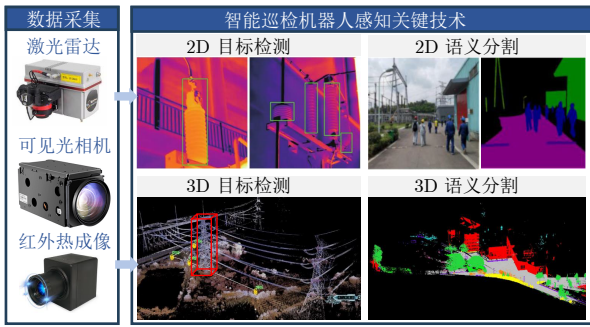


图 3 智能巡检机器人感知关键技术

Fig.3 Key technologies of intelligent inspection robot perception

背景变化的图像分割效果较好,但对于大规模的图像分割,计算复杂度较高。

深度学习快速发展给 2D 图像语义分割技术带来了重大的创新和应用。基于深度学习的语义分割主要分为单网络结构和多网络结构。目前,卷积神经网络是主流的基础网络结构,它通过编码器与解码器的结合来实现语义分割。在单网络结构中, U-Net^[36] 和 SegNet^[37] 是两个比较经典的结构,已经成功应用于医学影像分割和自然场景图像分割等领域。例如,在医学影像领域, U-Net 已经广泛应用于

MRI 和 CT 图像的肿瘤分割等任务中,取得了显著的效果;而在自然场景图像领域, SegNet 则被用于道路、车辆和行人的分割。对于在电力机器人巡检上的应用,主要包括变电站设备漏油检查^[45]、变电站设备准确的红外热图像实例分割^[46]等。

实现技术方面,多尺度信息融合是一种在语义分割中提高精度的常用方法,主要思想是通过在网络中增加多个分支或使用多个卷积核尺度来提取和融合不同尺度的特征。例如, DeepLab^[38] 通过 ASPP^[38] 融合多尺度信息。在此基础上, DeepLab v3+^[39] 采用 ASPP 模块来融合不同尺度的特征,同时使用空洞卷积增加感受野,提高分割结果精度。

2D 语义分割对于像素分类精度有着较高要求,因此带来了大量的人工标注成本。无监督 2D 语义分割方法通过未标注数据进行训练,避免了对大量人工标注数据的需求。常见的无监督分割方法包括基于聚类的方法,例如 K-means 聚类^[47] 和谱聚类^[48] 等。这些方法将图像像素聚类,把相似像素分配至相同类别以实现分割。另外,还包括基于生成模型的方法,如生成对抗网络^[49] (GAN),这些方法使用未标注数据进行训练,生成一组特征向量,然后将这些特征向量用于后续的语义分割任务。弱监督

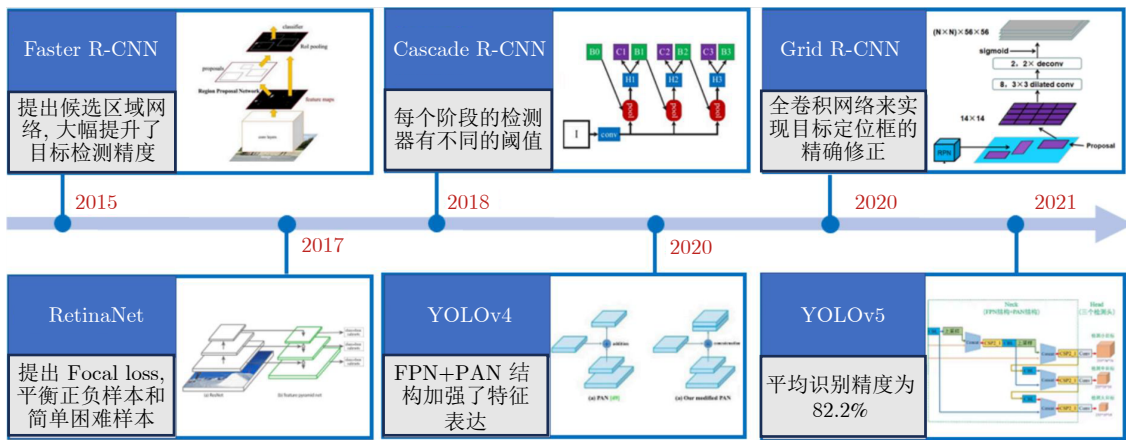


图 4 典型 2D 目标检测网络
Fig.4 Typical 2D object detection networks

表 2 2D 语义分割技术分类
Table 2 Classification of 2D semantic segmentation techniques

方法	分类类型	方法类型	原理或用途
传统方法	原理	基于阈值的方法 ^[33]	根据灰度值大小不同,设定阈值完成分割
		基于边缘检测的方法 ^[34]	对图像边缘线条进行检测
		基于图割的方法 ^[35]	利用图形结构最小割分割图像
深度学习方法	用途	多尺度信息融合 ^[38-39]	增加感受野,提高分割结果精度
		无监督语义分割 ^[40-50]	减少大量人工标注成本
		实时语义分割 ^[51]	节省计算资源,加快推理时间

2D 语义分割结合了少量标注数据和大量未标注数据. 此方法以少量标注数据为基础初始化标签, 再通过标签传播技术将标签扩展至未标注数据, 获得初步分割结果. 随后, 利用这些结果和有限的标注数据来训练模型, 以提高分割精度. 代表性算法如 Semi-supervised Random Walk^[50], 这些方法有效地利用了有限的标注资源, 提高了分割效率和准确性.

实时 2D 语义分割是一项对实时性要求较高的任务, 可应用于电力机器人巡检时的边缘端部署. 为实现实时语义分割, 通常采用轻量级网络结构、硬件加速和分布式计算等技术. ESPNet^[51] 等轻量级实时语义分割网络得到了广泛应用. ENet 以其轻量级设计而突出, 保持高精度的同时实现实时处理速度, 针对实时应用进行优化, 有效平衡了速度与精度. ICNet 基于金字塔结构, 具有高效的多尺度处理能力, 在减少计算量的同时提高分割精度. ESPNet 是一种基于分离卷积的网络, 在保持较高精度的同时大大减少了计算量. 此外, GPU 加速和分布式计算也可以帮助加速语义分割算法, 从而实现实时 2D 语义分割.

尽管语义分割技术已应用到部分电力场景中, 但在某些复杂环境下仍然性能受限. 例如, 广泛分布的配网线路背景跨越山区、湖泊、平原、林地等地貌, 难以实现准确的电力部件分割. 此外, 单一模态的无监督和弱监督方法也难以在电力场景中保持较

高准确率以节省人工标注成本, 在这种情况下, 如何利用不同模态数据的互补性信息实现多模态数据精准场景理解, 是未来电力巡检的主要研究方向.

2.2 电力设施机器人巡检 3D 感知技术

3D 目标检测技术相比 2D 目标检测技术, 可以在更复杂环境下提供更加真实的场景信息, 在三维空间中对检测目标精准识别和定位. 如图 5 所示, 在智能巡检机器人运行过程中, 利用点云和体素等常见 3D 模态, 可以对物体进行完整性检测和形态分析, 提高机器人的巡检效率和精度.

2.2.1 3D 目标检测

3D 目标检测是指在 3D 数据表征下对目标进行识别和定位, 可用于架空线路杆塔和导线检测任务, 为后续智能分析提供精细化的基础. 在数据表征上, 主要分为基于点云的检测方法和基于体素的检测方法. 1) 基于点云的方法: 主要是通过将三维场景转化为点云数据, 使用算法模型处理点云数据, 实现对待测物体的检测和识别. Maturana 和 Scherer^[52] 提出一种基于 Frustum 的 3D 目标检测方法, 通过利用 RGB 图像和点云数据融合, 结合 2D 和 3D 的数据表征, 实现了对目标物体的精细化检测. Charles 等^[40] 提出基于 PointNet 的 3D 目标检测方法, 使用神经网络直接处理点云数据, 提高了检测的精度和速度. 同年, Charles 等^[41] 在 PointNet 的基础上进一步提出基于 PointNet++ 的 3D 目标

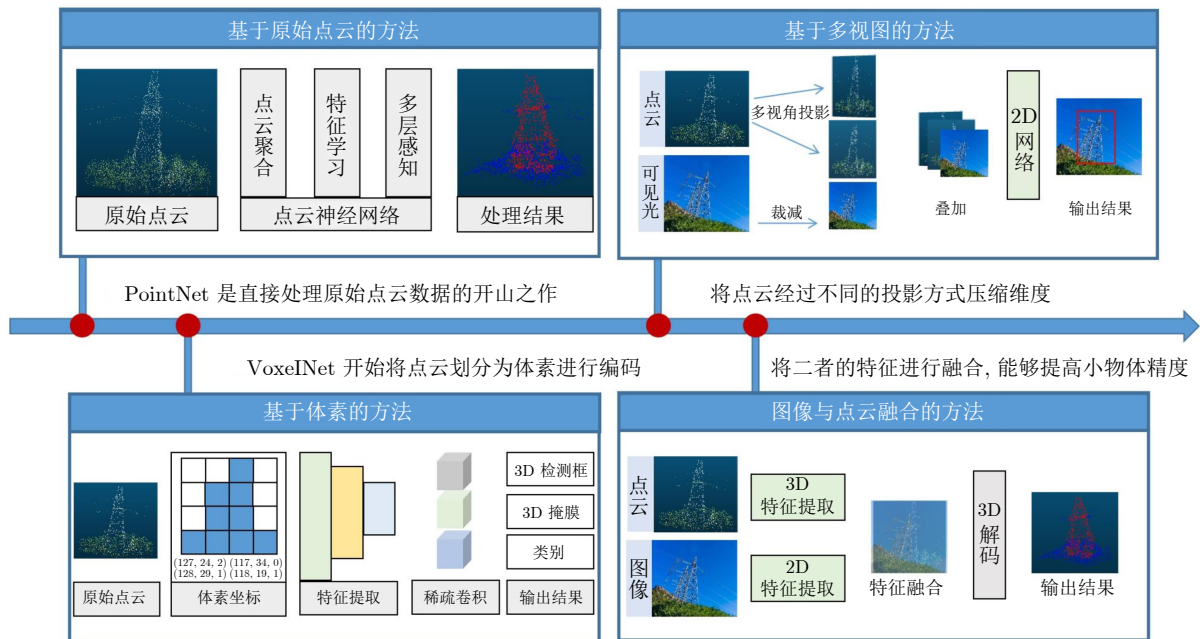


图 5 3D 点云检测和分割方法

Fig. 5 3D point cloud detection and segmentation methods

检测方法,设计多级特征提取和聚合网络,更大程度上提升目标检测的效率和准确性.在数据处理方面,Zhao等^[43]提出一种半监督自我集成的3D目标检测方法(SESS),旨在利用未标记数据和增强学习检测的性能. Shi等^[53]提出在网络中的3D局部感受野池化层和局部信息聚合机制,实现对点云数据精细感知和全局信息的整合.

2) 基于体素的3D目标检测方法:将三维空间划分为多个体素网格,并将每个体素的特征作为输入,通过网络模型对每个体素进行分类和回归来实现目标检测. Zhou和Tuzel^[54]采用体素卷积神经网络的方法,预测每个体素的对象分类和对象边界框,实现目标检测. Ouyang等^[55]把点云数据投影到二维鸟瞰图,将其体素化,快速实现目标检测. Shi等^[42]提出一种联合点云和体素的三维目标检测方法,将点云数据转化为等距的体素,然后采用自适应体素表示来处理同尺度目标. Dong等^[56]利用三维卷积来学习点云特征表示,采用多层3D锚点设计来提高检测准确率. Shi等^[57]提出PV-RCNN++目标检测方法,采用点云中心距离采样和体素空间采样策略提高检测效率,并采用RoI对齐模块和Mask Head模块来进一步提高检测准确率.

2.2.2 3D语义分割

3D语义分割基于三维数据将点云或体素中的每个点或体素分配到相应的语义类别中,需要从点云或体素中提取出特征,如颜色、形状等,然后使用深度学习模型对这些特征进行分类和分割,从而实现点对点或体素中各个部分的语义理解和识别.目前,主流的3D语义分割模型包括PointNet^[40]、PointNet++^[41]、DGCNN^[44]等,PointNet++广泛应用于点云语义分割方法,采用点云金字塔结构,联合局部和全局特征实现高效语义分割.同时,RandLA-Net^[58]等一些研究基于3D注意力机制,增强点云特征的特征能力. PointGNN^[59]和PointASNL^[60]采用图神经网络对点云数据建模.

在电力设施机器人巡检方面,3D目标检测和语义分割主要针对于输电线路杆塔和电力线的提取,以确定电力设施精确的3D坐标位置. Li等^[61]提出一种基于图卷积网络的方法,使用线路结构约束算法来识别各个电力线路,基于多项式的算法重建电力走廊,并从激光雷达(LiDAR)点云中提取电力线和杆塔. 具体来说,Jeong等^[62]提出了不仅可以实时认知输电线路及其环境,还可以通过将激光雷达测量的点云数据与无人机的飞行信息相结合来估计垂度的方法. 实时环境认知解决了体素化映射的坐标变换和概率下采样问题,从而保证有限的

硬件需求,并引入稳健的随机样本共识来有效提取输电线路点云数据,从而准确估计垂度. Munir等^[63]引入了重建高压输电线路导体束中的巨大间隙方法,使用3D体素网格将电源线检索为跨度点,通过将跨度点分成几个段来生成束掩膜,二分图利用来自这些掩膜的束位置信息将具有宽间隙的束段连接起来,最后再次对线束进行分割,形成导体掩膜,以便利用基于图像的算法和概率提取子导体. 针对输电线路附近植被对电力杆塔提取干扰的问题, Huang等^[64]提出一种使用随机森林分类器并结合平滑颜色特征的分类方法,消除了靠近塔腿的植被的错误分类点,在点云中精准提取输电线路塔架.

一些方法通过融合多种传感器数据,探索了多模态3D语义分割,例如SAT3D^[65],这成为了电力设施多模态精细化巡检的基础^[66]. 但目前的多模态感知场景仍然受限,如何在电力设施精细化机器人巡检领域,实现更准确的语义分割和目标检测,还有待进一步的研究和探索.

2.3 智能巡检机器人导航关键技术

2.3.1 高精度自主导航定位技术

智能巡检机器人的关键科学问题之一是实现高精度的自主导航定位技术. 早期的巡检机器人通常依赖于固定轨道的导航系统,这种系统虽然易于控制且准确性高,但因其显著的性能局限和高造价而难以广泛应用或实现高度智能化.

目前,无轨导航控制技术主要包括全球导航卫星系统(GNSS)和惯性导航系统等^[67]. GNSS导航技术包括美国全球定位系统(GPS)和中国北斗导航等,其原理是利用GNSS卫星对巡检机器人进行定位^[68]. 然而,由于电力巡检环境中存在复杂的电磁分布、时钟误差以及通信干扰等因素,对卫星导航的可靠性产生了一定制约,难以在狭小区域实现精准导航.

激光导航技术基于激光进行定位和导航,它主要由激光发射器、反射系统、光电检测组件、数据收集与传输单元构成. 该技术通过反射系统发射激光,当遇到标志物时,标志上的反射装置会将激光信号及其携带的数据经过光电转换装置处理,并将反射信号返回. 随后,数据采集系统会解读这些信息,并将其发送至处理器进行分析. 通过比较这些数据与已知标志的位置,可以精确计算出传感器与标志间的相对位置及方向,进而确定设备的具体位置,实现精准导航. 尽管激光测距在分辨率上有显著优势,但其性能易受环境因素干扰,如何有效去除采集信息中的噪声成为一大挑战. 同时,激光导航的

探测范围受限, 仅依赖激光进行定位导航面临一定的难度。

惯性导航通过电子罗盘和里程计等设备提供相对定位, 但易受到磁场干扰, 存在累积误差问题。因此, 开发智能巡检机器人时, 不仅要选择合适的导航技术, 还需解决各种技术限制, 如电磁干扰、环境干扰、累积误差等问题, 以实现准确、可靠的自主导航定位能力。

2.3.2 定位与地图构建关键技术

SLAM 技术是巡检机器人地图构建的关键技术, 机器人 SLAM 需要标识环境特征在相应坐标系中的位置, 以建立准确的环境模型。另一方面, 需要利用准确的环境地图来对机器人进行准确的自我定位和校正。在电力设施机器人巡检领域, 各种地图模型发挥着不可或缺的作用, 为机器人提供精准的环境信息和导航能力。栅格地图通过简单的二维网格形式直观地反映环境的占用情况, 适用于初步的环境感知和障碍物避让。特征地图利用几何元素精确描述环境特征, 如墙面、树木位置等, 对于需要高度准确性的定位和细节识别尤为重要。拓扑地图通过节点和弧线展现环境的结构关系, 适合于路径规划和导航, 尤其是在结构明确的环境中。混合地图结合尺度和拓扑的优势, 提供一个既包含具体尺度信息又具备清晰拓扑结构的综合地图模型, 增强了机器人对环境的整体理解和灵活应对能力。这些地图模型的综合应用, 为电力设施的自动化巡检提供了强大的支持, 使机器人能够更加高效、安全地完成巡检任务。

SLAM 技术的引入显著提高了巡检机器人在复杂电力设施环境中的自主导航能力。电力设施, 如变电站、输电线路等, 往往分布在地形复杂、环境多变的区域。SLAM 技术通过实时收集周边环境信息, 并利用这些信息进行精确的自我定位和环境地图的构建, 从而确保机器人能够在这些复杂的环境中准确无误地进行巡检, 极大地降低了人工巡检的安全风险和劳动成本。其次, SLAM 技术还能够提高巡检的效率和质量。通过精确的定位与地图构建, 机器人能够有效规划巡检路线, 减少重复或遗漏区域, 确保巡检的全面性。同时, 配合高精度的传感器, 机器人可以在巡检过程中实时检测并记录下电力设施的运行状态和潜在问题, 提供更加准确、详细的数据支持, 便于运维人员及时进行维护与决策。

此外, SLAM 技术还具有良好的可扩展性。随着技术的进步, SLAM 系统可以集成更多高级功能, 如障碍物识别与避障、复杂任务的自主规划与执行等, 进一步扩展机器人巡检的应用场景和能力, 提

升电力设施的运维效率和智能化水平。因此, SLAM 技术不仅能够提升巡检的安全性、效率和质量, 还为电力设施的智能化运维开辟了新的可能性, 是未来电力设施自动化巡检技术发展的关键。

3 电力设施多模态精细化巡检关键技术

3.1 可见光红外双光融合技术

在电力设施巡检中, 单一传感器所获取的图像通常只具备某一方面的信息, 无法满足精细化巡检需求, 因此图像融合技术应运而生。图像融合是从不同模态的图像中提取互补信息, 并对其进行聚合, 以生成更丰富、更有意义的特征表示。如图 6 所示, 红外图像主要依靠物体自身的热辐射进行成像, 可以突出背景中隐藏的热目标, 其不受光照条件、天气的影响, 但对比度较低, 纹理细节不丰富^[69]; 可见光图像通过反射可见光进行成像, 纹理细节和对比度更适合人类的视觉感知, 但在烟雾、夜间等条件下的成像效果较差^[70]。为了更好地获取待巡检电力设施的温度表征, 实时、准确研判电力设施温度异常, 多模态精细化检测往往将两者融合后获得一幅既有可见光图像边缘、又有红外热辐射目标细节信息的互补融合图像。图像融合也可以看作是一种增强技术, 旨在将不同类型传感器获得的图像组合在一起, 生成一幅鲁棒或信息丰富的图像, 以促进后续处理或帮助决策。融合方法的关键是有效的图像信息提取和适当的融合原理, 这允许从源图像中提取有用的信息并将其集成到融合图像中, 而不会在过程中引入任何伪影^[71]。随着目标检测与识别、军事监视等应用需求的不断提高, 红外图像与可见光图像融合技术成为该领域研究的热点。在安防领域, 红外与可见光融合图像可以准确识别黑暗环境、化妆打扮、佩戴眼镜等条件下的人脸, 为商业应用、公安执

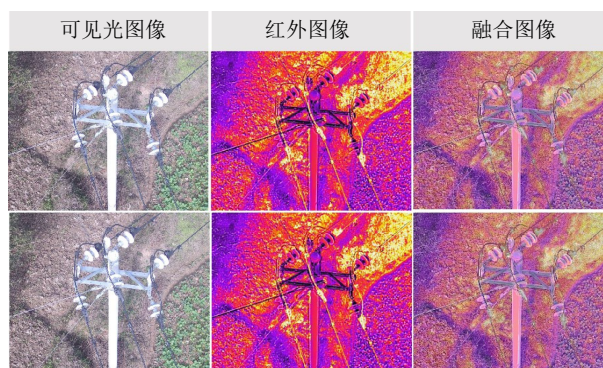


图 6 电力设施红外和可见光融合图像

Fig. 6 Infrared and visible light fusion images of power facilities

法等提供便利需求;在军事领域,红外与可见光融合图像可以实现恶劣环境下隐藏目标的识别与跟踪^[72];在智能交通领域,红外与可见光融合图像应用于行人检测^[73]、车辆识别与车距检测^[74]、道路障碍物分类^[75];在农业生产领域,红外与可见光融合图像可应用于水果的成熟度检测^[76]、植物的病态检测^[77]等.近几十年来,大量红外图像与可见光图像融合方法相继被提出,并在实际应用中得到推广.如表3所示,红外图像与可见光图像双光融合方法大体上可以分为:

1) 基于多尺度变换的方法.通过分辨率分解方法,基于变换域的图像融合方法可获取一系列包含不同层次的子图像,以保留更多的图像细节信息.多尺度变换是应用最为广泛的基于变换域的融合方法,具体可以分为金字塔变换、小波变换和非下采样多尺度多方向几何变换等方法.其中,金字塔变换将图像分解成不同尺度且呈金字塔状的子带图像进行融合.拉普拉斯金字塔(Laplacian pyramid, LP)变换在图像融合上取得了很好的效果^[78],比率低通金字塔^[79]等融合方法同样被应用到可见光和红外图像上.小波变换的概念最早由 Grossmann 和 Morlet^[80]于1984年提出,近年来,小波变换被应用于多频图像分解动态融合策略,实现了高频超声图像与低频超声图像的有效融合,获得细节增强、整体图像质量更好的融合图像^[81].

2) 基于稀疏表示的方法. Yang 和 Li^[82]提出基于稀疏表示(Sparse representation, SR)图像融合方法,其重点在于过完备字典的构造和稀疏系数分解算法^[83].过完备字典的构造方式主要有两种:基于数据模型和基于学习算法的过完备字典.基于数据模型的过完备字典是利用特定的数学模型进行构建的,该方法虽然高效但难以应对复杂数据,可利用基于联合学习策略的平稳小波多尺度字典^[84]和由过完备离散余弦字典与基函数相结合的混合字典^[85]来解决此问题.而基于学习算法的过完备字典是通过训练样本集的方式构造的,常用的有基于最优方向法(Method of optimal directions, MOD)字典^[86]和奇异值分解的稀疏字典训练(Kernel singular value decomposition, K-SVD)字典^[87].

稀疏系数分解算法中,匹配跟踪(Matching pursuit, MP)算法利用原子向量的线性运算,从训练好的过完备字典中选择最佳线性组合的原子以表征图像信息,但其迭代结果只是次优.正交匹配追踪(Orthogonal matching pursuit, OMP)算法^[88]在MP算法基础上进行改进,经OMP算法处理后的原子组合都已处于正交化状态.在精度要求相同的情况下,OMP算法的计算效率高于MP算法.此外,为解决OMP算法和MP算法融合规则设计难的问题,同步正交匹配追踪(Synchronous orthogonal matching pursuit, SOMP)算法^[89]基于OMP算法进行了原子集的改进,其可从不同的源图像中分解出相同子集的字典,从而简化图像融合规则设计,在图像融合领域得到广泛应用.

稀疏表示与传统多尺度变换的图像融合方法相比有两大区别^[90]:一是多尺度融合方法一般都是基于预先设定的基函数进行图像融合,这样容易忽略源图像某些重要特征;而基于稀疏表示的融合方法是通过学习超完备字典来进行图像融合,该字典蕴涵丰富的基原子,有利于图像更好的表达和提取.二是基于多尺度变换的融合方法是利用多尺度的方式将图像分解为多层图像,再进行图像间的融合,因此,分解层数的选择就尤为关键.

3) 基于神经网络的方法.神经网络在图像融合的应用始于脉冲耦合神经网络(Pulse coupled neural network, PCNN)模型.与其他神经网络模型相比,PCNN模型在图像处理 and 融合领域具有独特的优势,它能够模仿生物视觉系统中的脉冲耦合机制,从而在图像信息的提取和融合过程中展现出卓越的性能;与其他传统的神经网络模型相比,PCNN更加注重图像的空间信息与邻域像素之间的相关性,这使得它在处理复杂的图像融合任务时表现得更加优越.通过利用这种独特的脉冲耦合机制,PCNN可以有效地捕捉到图像中的重要特征和细节,从而提高融合图像的质量和精度.

基于深度学习的图像融合方法渐渐成为图像融合领域的主要研究方向,但其在异源图像融合的研究上还处于一个初步发展的水平.深度学习方法是將源图像的深度特征作为一种显著特征用于融合图

表3 可见光红外双光融合方法分类
Table 3 Classification of visible light infrared dual-light fusion methods

方法	优点	缺点
基于多尺度变换的方法	多层次子图像保留了更多图像细节信息	基于预先设定的基函数进行图像融合,易忽略源图像部分重要特征
基于稀疏表示的方法	超完备字典蕴涵丰富的基原子,有利于图像更好的表达和提取	难以应对复杂图像融合
基于神经网络的方法	避免了传统算法手动设计复杂的分解级别和融合规则,并有效保留源图像信息	对计算资源需求较大,暂未大量应用

像的重建, 卷积神经网络是目前最常用的图像融合深度学习方法. Li 等^[91]、Ren 等^[92] 都提出一种基于预训练的 VGG-19 (Visual geometry group-19) 网络, 以提取源图像的深层特征, 获得较好的融合效果. 2019 年, Ma 等^[93] 首次将端到端的生成对抗网络模型用于解决图像融合问题, 避免了传统算法手动设计复杂的分解级别和融合规则, 并有效保留源图像信息.

在技术上, 电力设施多模态精细化机器人巡检主要基于深度学习模型, 利用可见光图像和红外图像的特征级融合提取巡检目标的多模态特征, 再利用解码器用于具体下游任务. 这是因为直接使用图像级的融合难以有一个统一的特征提取器能够同时从不同层面提取可见光和红外的高级互补性特征. 因此在电力场景的多模态双光融合中, 往往设计一个对称双分支网络^[94-96], 分别提取两个模态的判别性特征, 然后在不同层级特征上进行多尺度融合. 简单的拼接和相加通常不能高效地利用多模态特征的独特性, 因此在特征融合上, 大部分技术方法是利用一个模态去引导另一个模态挖掘出与此模态的一致性特征^[97-98], 在获得一致性特征后, 多模态的互补性则得以体现^[99]. 例如在变电站全天候检测中, 当天气恶劣导致能见度下降时, 红外模态可以在一方面反映设备的温度特性, 同时多样的温度分布也使得电力设备呈现清晰的轮廓. 在这种情况下, 低分辨率的红外图像则为能见度下降的可见光图像提供了互补性信息, 保持电力设施多模态精细化机器人巡检的准确性和全面性.

3.2 可见光图像和点云融合关键技术

为了实现更加准确、全面的巡检任务, 可见光图像和点云数据的多模态融合技术被广泛应用于巡检机器人中^[100-102]. 如图 7 所示, 当可见光图像和点云数据同时获取时, 它们提供了不同的信息, 并且

可以互相补充. 可见光图像提供了物体的外观信息, 可以对物体进行分类、定位和识别等; 而点云数据提供了物体的三维形状信息, 可以对物体的几何形状进行分析和建模.

可见光图像和点云数据融合的关键技术包括以下几个方面: 1) 多传感器标定: 巡检机器人通常配备多个传感器, 传感器标定是多模态融合的前提, 需要将多个传感器的坐标系进行统一, 并且确定它们之间的相对位置和姿态; 2) 传感器数据预处理: 将不同传感器采集的数据进行配准、校正和滤波等预处理操作, 使得它们具有一致的参考坐标系和数据精度; 3) 特征提取: 针对可见光图像和点云数据, 分别提取它们的特征信息, 比如颜色、纹理、形状等, 为后续的数据融合提供依据; 4) 数据配准和融合: 由于采集的可见光图像和点云数据来自不同的传感器, 需要对其进行数据配准处理, 将可见光图像和点云数据对应起来, 以便后续进行数据融合. 数据融合算法是实现可见光图像和点云数据融合的核心技术, 将不同传感器采集的数据进行融合, 构建一个综合的三维模型, 可以提高数据的准确性和综合分析能力, 从而实现了对电力设备和线路等的全面检测和监测^[103].

近年来, 由于对可见光图像和点云多模态融合关键技术的研究愈发成熟, 该技术在国内外电力智能巡检机器人上的应用也愈发广泛. 例如, 在电力设备巡检方面, 2020 年, 中国南方电网公司与华为公司合作推出了智能巡检机器人, 如图 8 所示, 配备了可见光相机和激光雷达等多个传感器, 可以使用可见光摄像头和 3D 激光扫描仪等传感器对高压设备进行全面的模态精细化巡检, 包括设备表面的绝缘破损、内部电路板腐蚀等问题; 在输电线路巡检方面, 2021 年, 美国智能制造公司 Locus Robotics 推出一款名为“Hemisphere”的电力输电线路巡检机器人, 配备了可见光相机和激光雷达等多个

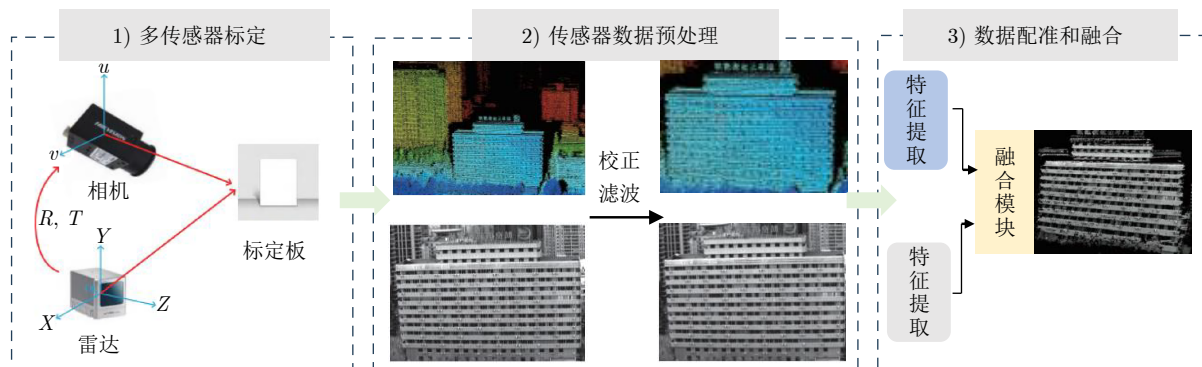


图 7 点云和可见光融合流程

Fig. 7 Point cloud and visible light fusion process



图 8 中国南方电网巡检机器人

Fig.8 China Southern Power Grid inspection robot

传感器,通过多模态融合技术,可实现对输电线路潜在的表面缺陷和杆塔倾斜沉降等问题进行全面巡检和故障检测. 2022年,中国大唐集团公司推出了名为“鸟瞰者”的火电厂巡检机器人,配备了可见光相机、激光雷达和红外相机等多个传感器,可以自主导航对火电厂进行全面检测,包括烟囱、冷却塔等. 从以上国内外的实际应用可知,可见光图像和点云多模态融合技术可以将不同传感器的数据进行融合和处理,提高数据的综合分析和识别能力,从而实现机器人巡检中准确感知多模态的能力.

3.3 声学 and 光学图像融合相关技术

音频数据与图像数据的融合应用在工业领域面向个别较为特殊的零部件异常检测中具有广泛的潜力. 在电力设备巡检任务中,如整流变压器、直流换流阀等电力设备的特殊性^[104],只能通过声纹信号进行运行状态检测. 但现阶段仍存在声纹信号难以与待检设备进行对应等问题,构建面向电力设备可见光与声纹的多模态信息融合算法,完成待检设备与声纹信号匹配是现阶段电力设备巡检研究的重要趋势.

基于声模态与可见光融合的研究最早出现在生

物特征识别领域^[105], Brunelli 和 Falavigna^[106] 最早提出在匹配层上进行人脸与语音的融合识别,其使用概率框架组合线索来确定最可能的身份,取得了较高的识别率. 在生物特征识别的基础上, Noda 等^[107] 提出一个基于深度神经网络 (DNN) 的模型,通过将音频和视频信号融合在一起实现语音识别,该模型采用自编码器和条件随机场来提高识别的准确性和稳定性. 实验结果表明,与仅使用音频或视频信号进行语音识别相比,音视频信号融合的识别效果更好. 基于声纹与图像、视频特征多模态融合的方法在传媒领域同样起到了一定的作用,如张玉珍等^[108] 融合多模态信息来检测足球视频精彩事件,通过从视频中抽取音频、图像、字幕等特征实现足球视频精彩事件检测. 在医疗领域, Xu 等^[109] 提出一种融合语谱图和视频数据用于帕金森病检测的新方法,该方法基于深度卷积对抗生成网络,实现了语谱图和图像数据的有效融合与数据增强,结果表明该方法在检测帕金森病方面具有高准确性.

电力场景下,声纹信号常用来诊断变压器的运行状态^[110-112],而多数多模态方法通常针对语音数据和图像数据的融合,但是由于电力设施产生的音频数据可能因可理解度较低和背景噪声等因素而难以直接分析,导致面向电力工业的音频与可见光数据的有效融合研究相对较少. 此外,电力设施的监测通常要求高度的准确性和效率,这进一步强调了开发高效融合方法的重要性. 随着更多面向电力的多源开放数据库的普及,声学 and 光学图像融合相关技术对于多模态精细化电力巡检有着较强的研究价值. 图 9 为声学 and 光学融合流程,由于数据的异构性,融合主要体现在特征层面. 声学成像和光学成像后的数据通常利用不同层级的特征提取器进行浅层、

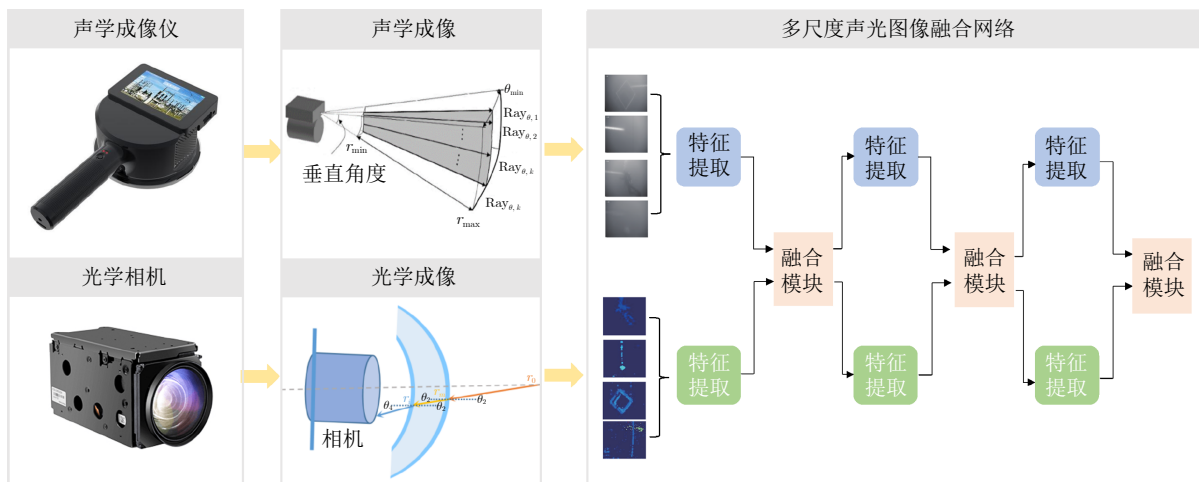


图 9 声学图像和光学图像融合流程

Fig.9 Acoustic image and optical image fusion process

中层和高层特征的融合. 这里的融合是通过融合模块的注意力等方式引导两个模态互相学习, 融合后的特征再输入到下一层级的特征提取器中. 下一层级的特征提取器在更进一步的特征提取后, 再进行本层级的融合. 通过这种在多个尺度下进行模态间信息交互的方式, 可以保留多模态的独立性, 并且互补性更全面, 加强电力设施不同类别缺陷特征表示, 提升巡检效率和准确性.

4 电力设施多模态精细化机器人巡检应用

电力设施多模态精细化机器人巡检技术采用了多种数据来源以实现输电线路、变电站及其他电力设施的全面检测和评估. 主要采用可见光、激光点云和红外热成像等数据形式. 其中, 可见光利用高清晰度的可见光相机捕获电力设施的高分辨率图像. 这些图像能够提供关于设施表面状态的直观信息, 如腐蚀、结构损伤、异物侵入等. 可见光成像是识别视觉特征和表面问题的基础技术, 对于常规检测和日常维护至关重要. 激光点云能够精确地反映被检测对象的几何形态和空间位置. LiDAR 技术在电力巡检中主要用于评估线路走廊的植被覆盖情况、塔架的结构稳定性以及线路间距离的测量, 有助于防止植被侵扰和确保传输线路的安全运营. 红外热成像能够检测设备产生的热量分布, 从而识别过热部件或异常温升情况, 这通常预示着潜在的故障或性能下降. 在电力系统中, 红外热成像技术被广泛用于监测变压器、断路器、接头以及其他关键设备的热态性能, 是预防性维护和故障早期诊断的重要手段. 多模态巡检技术充分利用以上多源数据的互补性, 在输电线路部件和塔架进行精细化巡检. 几种电力设施多模态精细化巡检相关应用典型案例如表 4 所示.

4.1 电力部件精准分割和输电线路异物检测

4.1.1 电力部件精准分割

电力部件分割是电力设施多模态精细化巡检最

常见的应用^[113-115], 其主要目的是实现对电力设备图像中各个部件的自动化检测、分类和识别, 研究对象主要包括架空输电线路、变电站等. 通过电力部件分割, 可以快速精准识别电力设备存在的问题和隐患, 例如设备表面的腐蚀、漏电等. 此外电力设备维护和保养的人力成本较高, 电力部件分割可以实现自动化检测和识别电力设备. 电力部件分割技术还可以用于实现自动化巡检, 例如可以通过对绝缘子进行语义分割来准确定位可能存在的绝缘子破损和断裂风险. 在规模较大的电力场景下, 实现电力设备自动化巡检可以提高巡检效率和准确性. 总之, 通过电力部件分割技术, 可以实现对电力设备的快速、精准检测和识别, 提高电力设备的安全性和维护保养效率, 减少人力和时间成本.

架空输电线路作为电力系统中至关重要的组成部分, 是电力远距离运输的主要方式. 许刚和李果^[116]基于 U-Net 的航空电力线图像语义分割网络进行改进研究, 包括数据集增强方法、色彩空间和损失函数的选择, 并得出两套最佳组合. 然后采用传统通道范数和提出的联合显著度指标对不同色彩空间的网络模型进行稀疏正则化, 并采用网络剪枝大幅降低网络参量和运算量. 左安全等^[117]结合电力线特点, 使用轻量级 PP-LCNet^[118] 替换原始主干网络 Xception^[119], 有效减少了参数量, 并提高了预测速度; 然后对 ASPP 模块增加空洞卷积分支和级联卷积以减少漏分割现象, 同时将空洞卷积分支改为瓶颈结构以减少参数量, 并利用 BAM^[120] 减少了背景分割错误. Mao 等^[121]提出的轻量级双目语义分割模型从双目图像中分割线路相关缺陷. Wang 等^[122]提出一种用于输电线路及其附件语义分割的图像处理框架, 使用一种基于矩阵的线段连接算法来加速检测框架, 并构造背景滤波器和人工轮廓线段特征生成器来提高检测性能. Yang 等^[123]提出一种基于注意力融合的深度分割网络, 实现了端到端的输电线路分割. 唐小煜等^[124]提出优化后的基于 U-Net 的模型来获取绝缘子区域掩膜图像, 实现绝缘子的识别与提取. Yu 等^[125]提出一种基于纹理和形状的有源轮廓模型, 用于在杂乱背景下分割绝缘子, 将

表 4 电力设施多模态精细化巡检应用

Table 4 Application of multi-modal fine inspection in power facilities

典型案例	融合模态	方法原理
电力部件精准分割	可见光+红外	结合可见光的高分辨率和红外图像的温度特性, 实现高精度部件分割
输电线路异物检测	可见光+点云	结合可见光的高分辨率和点云的形态特征, 充分识别线路异物
线路杆塔倾斜检测	可见光+点云	结合可见光的颜色信息和点云的位置信息, 准确分割杆塔, 实现倾斜度检测
线路覆冰多模态检测	可见光+点云	结合可见光的颜色信息和点云的位置信息定位电力线, 通过坐标计算覆冰厚度
台账异常检测	可见光+点云	通过可见光实现台账目标识别, 利用点云构建数字孪生网络以实现台账校准

形状约束合并到基于纹理的活动轮廓模型中,以捕捉绝缘体边界. Tang 等^[126]在原始 U-Net 模型架构中,将多尺度特征与注意力机制和残差学习相结合并应用,并使用注意力模块来实现特征增强,改进绝缘子语义信息和细节纹理的提取. Han 等^[127]提出多目标实例分割的 TAD-D-SOLO 高精度实例分割算法,克服了广泛使用的对象检测和语义分割方法的不足,通过将构成残差结构中的可变形卷积网络与 Transformer 模块合并,解决了边缘拟合差和识别缺失的问题.

为了充分发挥多模态针对电力部件分割中的有效性,消除单一模态带来的数据限制,一些方法把可见光和红外图像融合作为精准分割的基础,通过模态间的互补性突破模态的限制. 以红外图像和可见光图像为例,红外图像虽能提供关键的温度信息,但其相对较低的空间分辨率可能导致设备分割的不准确性,进而影响后续的温度分析和解释. 相反,尽管可见光图像通常拥有较高的分辨率,但会受到光照变化等影响,在光照敏感场景发挥欠佳,难以应用在全天候监测场景中. 并且由于可见光图像缺乏温度信息,无法在热故障检测中发挥作用. 针对输电线路在极端天气下因能见度下降所导致的电力部件分割精度降低的问题, Choi 等^[128]提出一种新的多模态图像特征融合模块,利用可见光和红外图像来提高输电线路检测性能. 具体而言,提出一个多分支特征提取块,其次是一个通道注意力块. 多分支特征提取块使用多个分支提取每个模态输入的代表性特征,引入通道收缩因子,将每个模态的特征图作为输入,并提取通道描述符的底层上下文信息. 多分支特征提取块的输出被联合聚合到通道注意力块中的注意力向量. 最后,注意力向量重新校准所提出模块的每个输入特征. 同时为了减少由于插入模块而导致的额外参数的数量,通道注意力块使用了具有通道收缩因子的 2D 卷积核代替全连接层. 这种方法解决了直接对电力场景多模态数据简单全局平均池化所带来的空间信息被消除的问题,并且减少了大量与每个输入特征向量的通道数成比例的附加权重参数. 在 U-Net 和 UMFNet 上验证以上模块,最终通过多模态方法在提出的输电电力线数据集上将部件分割的平均交并比分别提升至 78.01% 和 86.51%.

上述方法的研究对象仅仅是输电线路中的电力线,尽管取得了较为不错的巡检结果,但是当面对任务更加复杂的其他部件分割时,表现上还存在一定的差距. 这主要是由于巡检机器人在巡检过程中会受到角度变化和与不同电力设备距离的影响,因

此难以从获取的图像中提取足够的上下文信息. 此外,各种电力设备之间巨大的尺度差异进一步增加了高精度对象分割的难度. 针对这些问题, Xu 等^[129]基于双空间图的交互网络,通过可见光图像和红外热图像的协同工作,实现高压输电线路和变电站场景中电力设备的全天时间语义分割. 该网络模型首先提出基于双空间图的跨模态特征融合策略,从全局和语义的角度融合多模态高层特征. 特征图中的像素被聚类,并根据语义信息投影以构建顶点和图模型. 通过内部图与跨图的交互,有效地提取了丰富的上下文信息,并很好地缓解了巨大的模态差异所造成的干扰. 在以上多模态特征表示的基础上,提出多任务解码器协同实现语义分割和边缘检测. 低层特征被集成为权重图,以恢复空间细节和边界信息. 同时,边缘信息显式地转换为语义特征,以提取更微妙的线索和抑制背景噪声,从而实现高精度分割. 最终,该网络模型在提出的变电站和高压输电线路实现了 67.86% 和 78.94% 的平均交并比. 其中变电站包括隔离开关、绝缘子、变压器等部件,高压输电线路包括输电塔、水平绝缘子、柱式绝缘子和挡板等部件. 尽管该模型在高压输电线路和变电站上都取得了良好的机器人多模态巡检结果,但是图卷积的方式仍然造成大量的计算,并且特征融合主要集中在高层,忽略了低层特征融合,可能会错过一些重要信息和提示.

为了解决电力部件分割中大量人工精细化标注所造成的人工成本高的问题, Ma 等^[130]提出一种基于热图像引导的样本掩膜自动标注方法. 通过可见光和红外图像对之间的异质像素配准算法,利用红外图像前景分割结果指导可见光图像中电气设备样本掩膜的提取. 该方法仅采用了图像级标签弱监督,实现了对样本掩膜的自动标注,大大提高了效率,降低了成本. 具体而言,该方法首先利用具有全局搜索能力的进化算法获得初始参数,搜索局部最优解来实现红外图像和可见光图像的配准. 然后利用显著性检测、距离变换和 GrabCut 算法,实现了红外图像由粗到精的分割. 最后通过一个渐进优化模型的分割实例,利用全连接的条件随机场 (CRF) 约束到边界的损失,以指定每个对象的精细详细边界,并解决了分割的电气设备具有复杂结构的困难. 此外,该方法还探讨了自适应学习技术,以解决可见光和红外图像对之间的分辨率差异的问题,提高泛化能力. 最终在变电箱、变电柜、绝缘子和变压器上分别取得了 56.2%、51.5%、75.0% 和 89.0% 的平均准确率. 该方法的提出是弱监督在多模态电力设施巡检的一次深入探索,可有效解决电气设备像素级

标注成本高的问题, 并且具有提供相对准确的掩膜注释和成本较少的优点. 但是在针对输电和配电复杂多变的背景及尺度较小的目标上, 还有一定的局限性.

4.1.2 输电线路异物检测

输电线路异物检测是在多模态机器人拍摄图像基础上, 利用深度学习目标检测等技术检测输电线路上的悬挂异物. 相比于传统人工检测而言, 其利于节省人工成本, 并提高异物实时检测率, 从而大大减小由异物附着而导致的输电线路事故发生概率^[131-132].

由于人工成本、维护难度和环境复杂等原因, 人工巡检难以满足如今输电线路安全运行状态监测的需求, 因此国内输电线路巡检大多采用无人机、机器人等设备代替人工工作, 并结合图像处理技术实现异物检测. 基于输电线路异物检测的图像识别技术又可分为传统特征提取方法和基于深度学习的方法两个类别. 传统特征提取方法, 例如 SIFT 特征提取方法, 需要经过构建 DOG 尺度空间、关键点定位与搜索、方向赋值等步骤, 方法提取特征点少, 实时性不高, 检测精度低^[133]. 随着深度学习的不断发展, 其在输电线路异物检测中的应用不断增加. 图 10 为电力设施精准目标检测框架, 在该框架下, 网络提取异物、变电站设施等感兴趣区域目标, 通过多尺度特征图融合增强对应特征, 在检测头中共享权重层, 并输出目标的类别、位置以及其他限制条件. 例如异物的类型, 诸如风筝、气球、鸟巢等, 以及异物所在输电线路的位置. 主流检测方法如 YOLO、Faster R-CNN、meta R-CNN 等通过卷积神经网络提取图像特征, 并生成候选框进行分类与回归, 实现输电线路异物检测. 基于深度学习的输

电线路异物检测方法检测精度高、实时性强, 能够很好满足工业异物检测要求. 目前, 采用智能目标检测技术对输电线路图像中的异物进行检测, 已成为我国智能电网领域的研究热点^[134].

基于机器视觉的输电线路异物检测是通过由多模态机器人拍摄的输电线路异物图像或视频进行分析和处理, 从而实现输电线路异物的检测. 例如, 对于输电线路中鸟巢的检测, Jiang 等^[135]提出一种能够利用 SSD 对输电线路鸟巢进行识别的方法, 采用 HSV 颜色空间滤波器对检测区域进行滤波, 并验证了该方法在输电线路中鸟巢检测的可行性.

4.2 电力线路点云多模态精细化检测

电力设施多模态精细化巡检在点云上的相关应用主要体现在电力线路中的电力线和杆塔上, 相比于 2D 图像上的目标检测任务和语义分割任务, 更加关注目标的类别和状态信息 (有无缺陷). 点云数据通常侧重于体现目标的精准位置关系, 因此点云主要用于测量电力线和杆塔上的 3D 位置变化, 从而完成语义理解, 得出巡检结果. 一般点云是由坐标组成的点集, 根据点的位置可以识别出位置的形状, 但当场景复杂时, 往往单一模态的点云无法根据结构区分出前景和背景, 特别是当电力线路背景复杂且被植被所包围时. 因此多模态精细化巡检至关重要, 这里的多模态通常体现在点云数据和可见光数据的融合. 在采集阶段, 激光雷达融合了点云的坐标信息和可见光信息, 为每一个点分配颜色属性, 为后续点云精细化检测奠定了良好的数据基础.

4.2.1 线路杆塔倾斜检测

电力线路主要是由电力线、杆塔以及杆塔附件

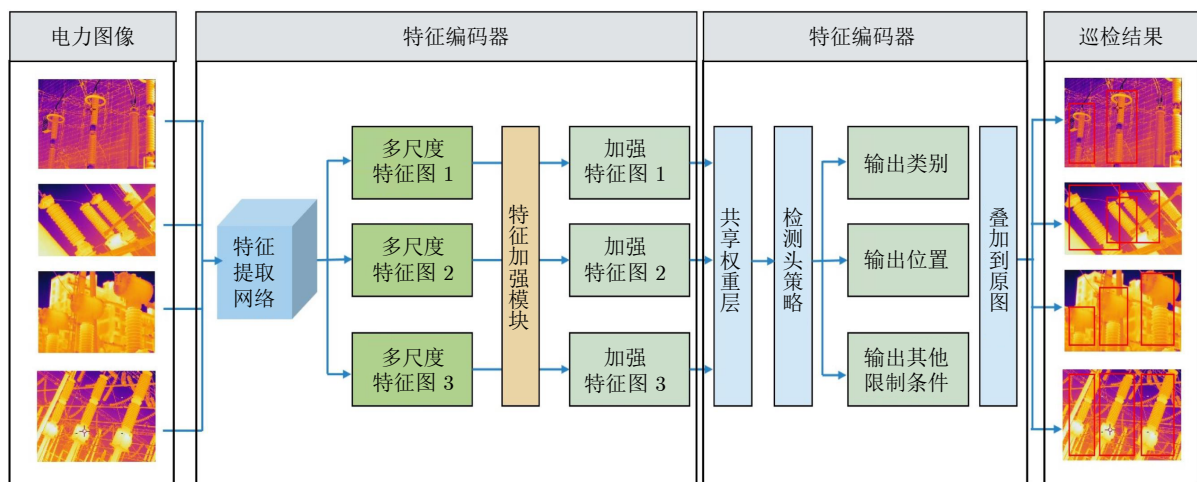


图 10 电力设施精准目标检测框架

Fig.10 Accurate object detection framework for power facilities

组成。其中杆塔是电力线路中最基础的元件，在支撑电力线运行过程中不仅要承受自身和导线的垂直载荷，还需要承受电力线架空后的张力载荷，长期处于自然力、重力等外力作用下，杆塔地基容易发生形变，导致杆塔倾斜，甚至倒塔断线，使供电线路陷于瘫痪，造成重大经济损失。传统的电力线路杆塔倾斜度测量方法主要有铅垂法、平面镜法、经纬仪法^[136]。这些方法不仅实现难度大、施工强度大、测量效率低，而且不能适应现代电网的安全运行和可持续发展。随着中国智能电网的高速发展，智能数字化处理技术广泛应用于电力行业，因此基于深度学习开展线路杆塔倾斜检测成为了研究主流。

数据采集利用 LiDAR 的高频率脉冲能部分地穿透植被，获取高压线下的地形信息，能解决人工和传统航测中存在植被遮挡的问题，是目前杆塔精确定位和重建的有效工具。通过激光雷达实时扫描输电线路现实场景建立实时输电线路三维点云。

具体实施方式如图 11 所示，针对杆塔倾斜的检测与防护，主要涉及杆塔识别和倾斜检测两个环节。首先杆塔识别是利用机载 LiDAR 点云自动确定杆塔的中心位置，以便进行后续的激光点云数据分析和缺陷诊断。这个过程通常用深度学习模型通过 3D 语义分割或目标检测对多模态点云完成。在提取出杆塔区域、确定杆塔每个点的三维坐标后，根据坐标执行相应算法（如最小二乘法）实现杆塔倾斜度的计算。一些相关的案例在杆塔识别和杆塔倾斜上有着较好应用。在杆塔识别方面，Li 等^[137]提出通过垂直投影的点云密度和高度变化来探测杆塔点云，识别杆塔位置。沈小军等^[138]提出将长距离电力线路点云划分为多个小距离尺度空间组，利用子空间高程差异和点云密度特征实现复杂地形条件下的杆塔定位与电力线提取。此外，为应对杆塔点云

提取算法对地面起伏较为敏感以及地面点难以剔除的问题，直接处理原始点云数据实现杆塔提取的轻量级网络被提出，通过神经网络对体素空间特征训练，实现精确提取^[139]。在倾斜检测方面，芦竹茂等^[140]通过拟合塔身结构来计算杆塔中心线，进而计算杆塔倾斜率。徐梁刚等^[141]提出基于杆塔塔身分层投影提取中轴点的激光点云杆塔倾斜检测算法，结果与实测值对比偏差小于 0.90%。应斌等^[142]开展了基于北斗高精度定位技术的输电铁塔滑坡监测试点应用，实现了杆塔形变量监测和地质灾害预警，提升了巡检效率和精益化管理水平。

4.2.2 线路覆冰多模态检测

输电线路覆冰多模态检测是指基于可见光与激光点云融合的多模态数据，通过多模态 3D 目标检测技术实现输电线路覆冰检测。具体而言，针对我国输电线路覆冰检测人工巡检方式存在效率低、人工安全无法保障、持续周期时间长等问题，近年来，国内外利用巡检机器人与视觉结合的方式来实时检测输电线路覆冰情况，当线路覆冰超出阈值时进行预警。这种方法大大减少了人工使用量，提高了检测效率，并大大缩短了线路覆冰处理周期，加强了输电安全^[143]。

与输电线路杆塔倾斜检测方法类似，具体实施方案依然需要通过对线路多模态点云进行 3D 目标检测，识别出电力线路的位置，再通过计算电力线的厚度来实现覆冰前后的探测，比对前后电力线厚度的变化，来精准识别线路是否覆冰以及覆冰厚度。实施方案的核心在于 3D 目标检测方法。多模态 3D 目标检测方法是采集到的线路实时图像提供的颜色、语义特征信息与输电线路点云提供的 3D 结构信息进行融合，即利用主干网络中的点云算子提取点云特征，通过基于采样的点和特征预测 3D 框来

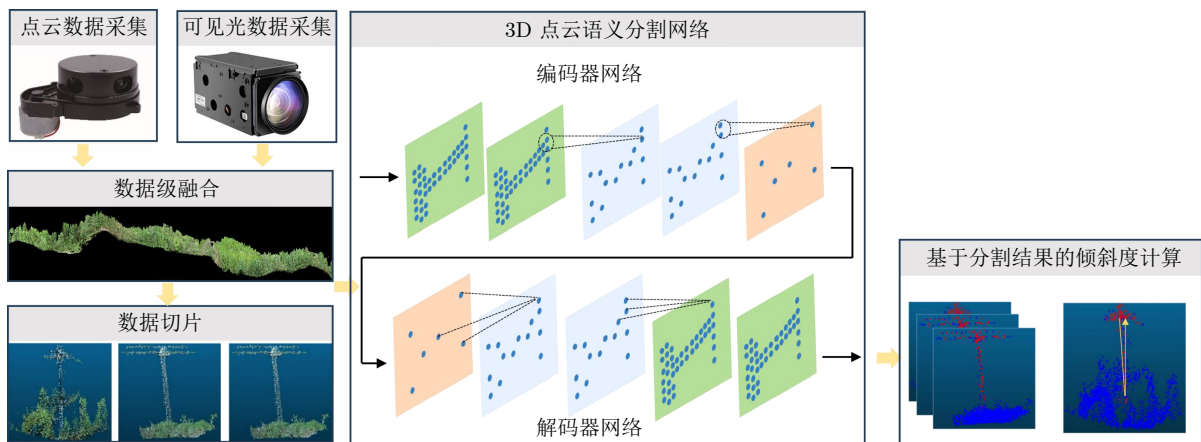


图 11 基于可见光和点云的杆塔倾斜检测流程图

Fig. 11 Flow chart of tower tilt detection based on visible light and point cloud

准确标注出导线的位置关系, 从而实现输电线路覆冰检测. 以 Point-GNN 为例, 文献 [59] 提出一种基于图神经网络的激光雷达点云目标检测方法, 通过固定半径的近邻图对点云进行有效编码来预测图中每个顶点所属的对象类别和形状, 从而实现三维目标检测. 在算法层面, 主流的多模态 3D 目标检测算法包括 PointFusion 和 EPNet 等. 以 EPNet 为例, 使用融合模块在不同尺度特征图上用语义特征来增强对应的点特征, 以获取更多的可辨别特征的表示, 再利用特征细化网络提取紧凑的全局特征, 送入到两个级联的子网络进行分类和回归, 从而实现多模态 3D 目标检测^[144]. 如图 12 所示, 在完成覆冰检测后, 除冰机器人还能采用无线控制、融合敲击振动与碎冰等技术, 实现机器人自主上下线、线路覆冰的旋转击碎和碾压脱落等工作.



图 12 带电除冰机器人
Fig. 12 Electric de-icing robot

4.3 电力架空线路缺陷分析及台账异常检测

输变电台账管理作为实行设备状态检修的基础, 是电力技术管理的一项基础内容, 其管理水平的高低直接关系到变电运行管理、设备管理、安全管理及检修校验管理等一系列输变电管理工作是否可以顺利开展^[145]. 现有大部分配电网在台账治理和校核上多采用人工进行, 时间周期长且效果不好. 而电力设施多模态机器人配合数字孪生技术、三维点云重建技术等新兴技术, 则可以有效代替人工进行配电网台账治理和校核过程, 同时相较于传统的人工台账治理及校核而言, 它具有稳定性好、持续工作时间长、精度和效率高等优点.

目前我国电力架空线路仍面临维护难、管理难, 运行维护电力架空线路的风险大、成本高等一系列问题, 需研制面向电力架空线路巡检的智能电力巡检机器人, 并配合新兴技术建成新一代的电力线路台账拓扑, 实现电力架空线路数字化管理、智能化检测与监测. 针对这一重大需求, 国家电网湖南省电力有限公司电力科学研究院联合湖南大学团队,

研制了数字孪生化的三维电力架空线路缺陷分析及台账异常检测系统, 如图 13 所示. 该系统从电力架空线路相关数据的三维点云重建、电力架空线路主要部件和元素语义分析深度神经网络研发、三维台账拓扑的异常检测技术开发入手, 对包括电力架空线路的三维模型数据、地形地貌数据、材料属性数据等电力架空线路立体环境进行三维点云结构重建, 并对数据进行预处理和清洗, 去除无用信息和异常数据, 创建高精度的电力架空线路数字孪生体. 并设计一种基于迁移学习配合大规模预训练模型的深度神经网络模型, 利用深度学习对电力架空线路三维点云拓扑进行语义分析, 识别线路中的主要部件和元素, 建立语义关系. 设计电力架空线路缺陷分析算法, 利用语义分析结果进行缺陷检测和定位, 分析缺陷类型、程度和可能的影响. 解决了长期以来电力架空线路人工检测难度大、缺陷识别率低等问题, 提高了电力架空线路的维护效率和准确性, 减少运维风险和成本. 该系统依托湖南省大数据平台汇集样本 50 000 余张, 累计支撑线路 2 000 余公里, 发现严重及以上缺陷 600 多个, 实现了工单驱动、自主巡航、缺陷在线识别和录入的全流程线上化闭环.

5 现有技术挑战与未来发展趋势

5.1 现有技术挑战

电力设施多模态精细化机器人巡检的应用旨在整合来自不同传感器的数据 (如可见光、激光点云、红外热成像等), 以提高故障检测的准确性和效率. 尽管这一技术具有显著的优势, 但在实际应用过程中也面临着多种挑战和局限性: 1) 尽管目前红外和可见光图像能够实现不同模态间的配准, 但是在更多源的情况下, 例如在融合非视觉数据 (如源荷数据) 与视觉数据的任务中, 还存在着较大的空间和时间分辨率差距, 这就要求在数据融合过程中进行精确的校准和同步, 以确保数据的一致性和准确性, 这是后续电力多模态数据发展的重大挑战. 2) 多模态巡检产生的数据量巨大, 对巡检机器人数据存储、传输和处理能力提出了较高要求, 尤其是在实时或近实时处理需求的场景中. 轻量化多模态巡检策略消除了因处理多模态数据造成的模型参数量和计算量爆炸问题, 完成了机器人边缘端部署, 对未来多模态巡检机器人的应用至关重要. 3) 多模态数据缺陷稀少且表征不统一, 例如多模态机器人巡检中, 不同的缺陷在不同数据下的表征形式不同, 并且样本稀少, 如何构建统一的数据增强策略、消除风格

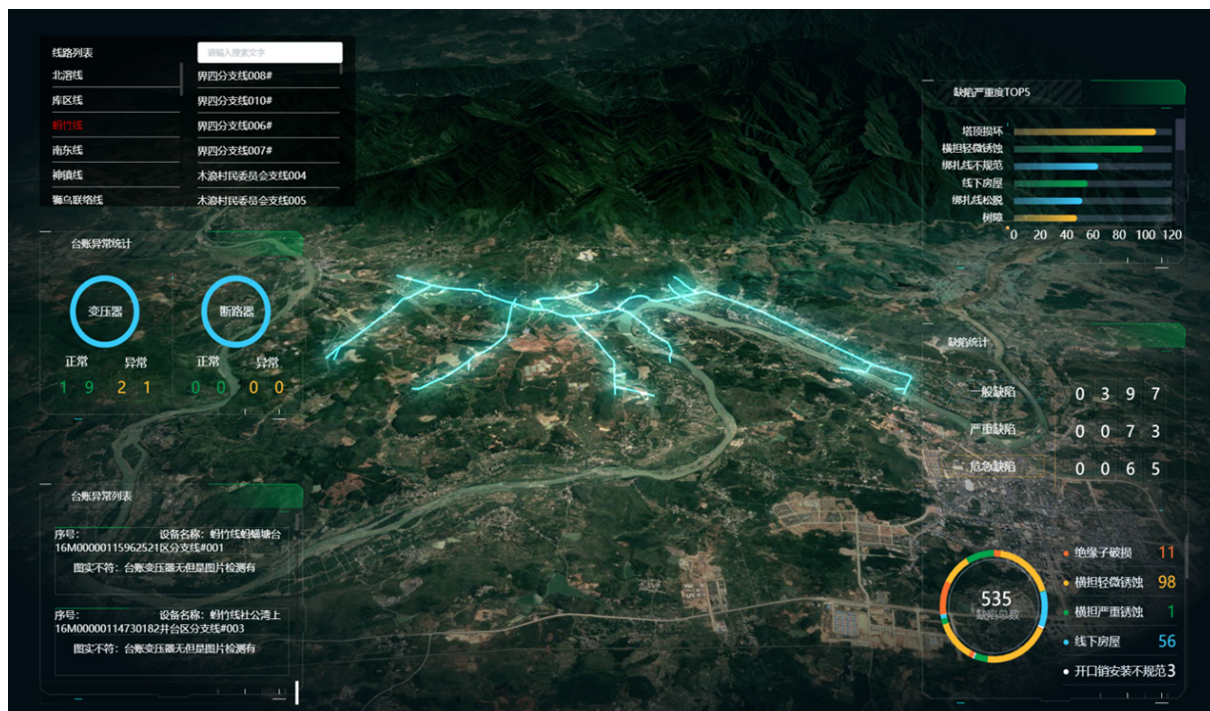


图 13 电力架空线路缺陷检测及三维台账拓扑数字孪生体

Fig. 13 Defect detection of power overhead lines and three-dimensional ledger topology digital twins

迁移和扩散模型等数据增强方式带来的分布不一致问题,是全面应用多模态精细化机器人巡检亟待解决的问题. 4) 多模态机器人巡检通常需要昂贵的设备和技术支持,这可能限制其在资源有限场景下的应用. 例如,激光点云的采集需要较高的成本,这对多模态巡检机器人的推广带来了重大挑战.

5.2 未来发展趋势

5.2.1 多模态精细化机器人巡检技术集成

多模态感知和多源融合技术作为机器视觉领域中重要且具有挑战性的问题,随着研究的不断深入,图像融合领域已经发生了巨大的变化,针对异源、多尺度、不同平台的图像在算法处理和实际应用上都有了一定发展. 为满足多模态精细化机器人巡检需求的多样性和精度,采用多模态巡检技术集成不同传感器(如光学、红外、声学等)实现对电力设施的全方位、多尺度检测是电力设施巡检技术的重要发展方向之一. 而其中的挑战主要集中在传感器数据融合策略研究、环境因素对传感器性能影响的补偿方法以及实时性能优化等方面.

5.2.2 巡检机器人自主学习及智能决策

电力线路巡检过程中存在大量不同种类的故障类型及复杂情况,例如线路老化、设备损坏、短路、过载等. 面对此类复杂情况,以往人工作业环境中

需要由组成的团队来完成. 而随着人工智能领域的不断发展与增强,新一代人工智能技术由以往的简单、单一任务处理工具逐渐发展成为具备“感知-决策-控制”闭环特征的复杂场景多任务处理系统,而在动态环境中构建机器人的感知-决策-控制协同作业机制来替代以往需要群体性团队人工作业完成复杂任务具有重要意义. 基于该方向的研究将赋予电力巡检机器人强大的环境适应能力,使得电力设施多模态精细化机器人具备在各种复杂环境下稳定工作的能力,从而应对复杂电力输变电环境中(如高压线路、变电站等高风险场景)待检目标异质化和多样化干扰因素的挑战.

5.2.3 电力设施多模态大规模预训练模型

随着电力设施多模态精细化巡检的发展,电力场景巡检业务精度和效率得到了极大提升,然而仍存在着巡检数据量大、标注难度高等问题. 因此,面向海量电力场景数据开展电力设施多模态精细化巡检的大规模预训练模型成了未来发展趋势. 尽管现有多模态大模型在文本和视觉等模态上有着优越表现,但是针对电力设施精细化巡检的大规模预训练还面临着数据格式差异大、难以获取统一标准电力巡检数据等挑战.

6 结束语

随着国家提出加快电网基础设施智能化改造和

智能微电网建设, 提高电力系统互补互济和智能调节能力的战略需求, 电力设施巡检有着重要的战略意义. 本文首先对智能巡检机器人的共性关键技术(如目标检测、语义分割、高精度自主导航、定位与地图构建等)进行了综述, 然后介绍了电力设施多模态精细化机器人巡检中多模态数据融合方式, 最后介绍了电力设施多模态巡检机器人应用的相关案例, 并探讨了电力设施多模态精细化机器人巡检面临的挑战和未来发展趋势.

现有多模态技术可以在部分任务上充分利用多种模态的互补性和一致性. 例如, 在可见光和红外图像融合的电力部件分割任务中, 结合可见光图像的详细纹理信息和红外图像的温度分布信息, 可以更准确地识别和定位故障点, 如过热组件等. 并且这种融合有助于减少误报率, 通过同时考虑温度异常和可见损伤, 提供更全面的故障诊断. 此外, 对于室外电力设施的巡检, 特别是在光照条件变化大的环境下, 这种融合方法可以提供更稳定和可靠的检测结果. 但是这种方法在融合过程中需要精确对齐图像, 并处理不同分辨率和尺度的问题, 增加了处理的复杂度. 可见光与点云数据融合可以应用在输电线路杆塔倾斜和覆冰检测中, 点云数据能够提供电力设备和基础设施的精确三维空间信息, 与可见光图像融合后, 可以实现对设备形态和表面状况的全面评估. 并且这种方法特别适用于评估设备的空间位置、结构稳定性和植被侵扰等问题, 对于规划维护和防灾减灾具有重要意义. 但是由于点云数据量通常很大, 处理和融合这些数据需要较高的计算资源和专业的数据处理技术. 将点云数据与可见光图像精确配准也是一大挑战, 需要精确的校准和高效的算法来保证数据的一致性.

未来的电力设施多模态精细化机器人巡检研究方向主要体现在技术集成、自主学习与智能决策能力、大规模预训练模型的开发等方面. 首先, 多模态感知和多源数据融合技术的进步, 为电力设施提供了全方位和多尺度的监测手段, 满足检测需求的多样性和精度要求. 其次, 随着人工智能技术的发展, 新一代人工智能技术的出现使得能够处理复杂场景和多任务的系统, 特别是在动态环境下构建的感知-决策-控制协同作业机制, 为替代传统的人工团队作业提供了可能. 这种进步不仅提升了电力巡检机器人的环境适应能力, 也使得电力设施巡检在高风险场景下面对异质化和多样化干扰因素的挑战时能够稳定工作. 最后, 面对电力设施多模态精细化巡检数据量大和标注难度高的问题, 开发大规模预训练模型成为了未来发展的趋势.

综上, 电力设施多模态精细化机器人巡检通过综合不同类型的数据来提供更全面和准确的电力设备状态评估, 但同时也面临数据处理复杂性、成本和技术挑战. 未来的研究和技术开发需要着重解决这些问题, 以进一步提升多模态融合技术的应用效率和可靠性.

References

- 1 Wu Qing, Zhao Tao, Dian Song-Yi, Guo Rui, Li Sheng-Chuan, Fang Hong-Wei, et al. General type-2 fuzzy logic control for a power-line inspection robot based on FPSO. *Acta Automatica Sinica*, 2022, **48**(6): 1482-1492
(吴庆, 赵涛, 佃松宜, 郭锐, 李胜川, 方红伟, 等. 基于 FPSO 的电力巡检机器人的广义二型模糊逻辑控制. *自动化学报*, 2022, **48**(6): 1482-1492)
- 2 Dian S Y, Chen L, Hoang S, Pu M, Liu J Y. Dynamic balance control based on an adaptive gain-scheduled backstepping scheme for power-line inspection robots. *IEEE/CAA Journal of Automatica Sinica*, 2019, **6**(1): 198-208
- 3 Wang Yao-Nan, Jiang Yi-Ming, Jiang Jiao, Zhang Hui, Tan Hao-Ran, Peng Wei-Xing, et al. Key technologies of robot perception and control and its intelligent manufacturing applications. *Acta Automatica Sinica*, 2023, **49**(3): 494-513
(王耀南, 江一鸣, 姜娇, 张辉, 谭浩然, 彭伟星, 等. 机器人感知与控制关键技术及其智能制造应用. *自动化学报*, 2023, **49**(3): 494-513)
- 4 Zhang Zhen-Guo, Mao Jian-Xu, Tan Hao-Ran, Wang Yao-Nan, Zhang Xue-Bo, Jiang Yi-Ming. A review of task allocation and motion planning for multi-robot in major equipment manufacturing. *Acta Automatica Sinica*, 2024, **50**(1): 21-41
(张振国, 毛建旭, 谭浩然, 王耀南, 张雪波, 江一鸣. 重大装备制造多机器人任务分配与运动规划技术研究综述. *自动化学报*, 2024, **50**(1): 21-41)
- 5 Li J Q, Xu Y Q, Nie K H, Cao B F, Zuo S N, Zhu J. PEDNet: A lightweight detection network of power equipment in infrared image based on YOLOv4-tiny. *IEEE Transactions on Instrumentation and Measurement*, 2023, **72**: Article No. 5004312
- 6 Chen G, Shao F, Chai X L, Chen H W, Jiang Q P, Meng X C, et al. CGMDRNet: Cross-guided modality difference reduction network for RGB-T salient object detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 2022, **32**(9): 6308-6323
- 7 Koji K, Takehisa Y, Tomohiro M, Katsuyoshi S. Development of substation patrol robot, overhead distribution line works manipulator and underground cable conduit monitor robot. *Robot*, 1988, **64**: 65-73
- 8 Masafumi N. Development of a patrolling and inspection robot at 500 kV Transformer substation. (Results of field test). *R&D News Kansai*, 1999, **386**: 13-15
- 9 Allan J F, Beaudry J. Robotic systems applied to power substations—A state-of-the-art survey. In: Proceedings of the 3rd International Conference on Applied Robotics for the Power Industry (CARPI). Foz do Iguaçu, Brazil: IEEE, 2014. 1-6
- 10 Beaudry J, Poirier S. Véhicule Téléopéré Pour Inspection Visuelle Et Thermographique Dans Les Postes De Transformation, Technical Report IREQ-2012-0121, IREQ (Hydro-Québec Research Institute), Canada, 2012.
- 11 Lu Shou-Yin, Qian Qing-Lin, Zhang Bin, Wang Ming-Rui, Li Xiang-Dong, Wang Hong. Development of a mobile robot for substation equipment inspection. *Automation of Electric Power Systems*, 2006, **30**(13): 94-98
(鲁守银, 钱庆林, 张斌, 王明瑞, 李向东, 王宏. 变电站设备巡检机器人的研制. *电力系统自动化*, 2006, **30**(13): 94-98)
- 12 Guo R, Han L, Cheng X Q. Omni-directional vision for robot navigation in substation environments. In: Proceedings of the

- IEEE International Conference on Robotics and Biomimetics (ROBIO). Guilin, China: IEEE, 2009. 1272–1275
- 13 Guo R, Xiao P, Han L, Cheng X Q. GPS and DR integration for robot navigation in substation environments. In: Proceedings of the IEEE International Conference on Information and Automation. Harbin, China: IEEE, 2010. 2009–2012
- 14 Zhou Z Y, Zhang C T, Xu C, Xiong F, Zhang Y, Umer T. Energy-efficient industrial internet of UAVs for power line inspection in smart grid. *IEEE Transactions on Industrial Informatics*, 2018, **14**(6): 2705–2714
- 15 Vapnik V N. *The Nature of Statistical Learning Theory*. New York: Springer, 2000.
- 16 Ho T K. Random decision forests. In: Proceedings of the 3rd International Conference on Document Analysis and Recognition. Montreal, Canada: IEEE, 1995. 278–282
- 17 He Hong-Ying, Yao Jian-Gang, Jiang Zheng-Long, Wang Xin-Xiu, Li Wei-Wei. Infrared thermal image detecting of high voltage insulator contamination grades based on support vector machine. *Automation of Electric Power Systems*, 2005, **29**(24): 70–74, 82
(何洪英, 姚建刚, 蒋正龙, 汪新秀, 李伟伟. 基于支持向量机的高压绝缘子污秽等级红外热像检测. 电力系统自动化, 2005, **29**(24): 70–74, 82)
- 18 Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, et al. Attention is all you need. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: ACM, 2017. 6000–6010
- 19 Carion N, Massa F, Synnaeve G, Usunier N, Kirillov A, Zagoruyko S. End-to-end object detection with transformers. In: Proceedings of the 16th European Conference on Computer Vision. Glasgow, UK: Springer, 2020. 213–229
- 20 Zhu X Z, Su W J, Lu L W, Li B, Wang X G, Dai J F. Deformable DETR: Deformable Transformers for end-to-end object detection. In: Proceedings of the 9th International Conference on Learning Representations. Vienna, Austria: OpenReview, 2021.
- 21 Zong Z F, Song G L, Liu Y. DETRs with collaborative hybrid assignments training. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). Paris, France: IEEE, 2023. 6725–6735
- 22 Zhang S L, Wang X J, Wang J Q, Pang J M, Lyu C Q, Zhang W W, et al. Dense distinct query for end-to-end object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, Canada: IEEE, 2023. 7329–7338
- 23 Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Columbus, USA: IEEE, 2014. 580–587
- 24 Girshick R. Fast R-CNN. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015. 1440–1448
- 25 Ren S Q, He K M, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(6): 1137–1149
- 26 Tang X, Du D K, He Z Q, Liu J T. PyramidBox: A context-assisted single shot face detector. In: Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany: Springer, 2018. 812–828
- 27 Cai Z W, Vasconcelos N. Cascade R-CNN: Delving into high quality object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA: IEEE, 2018. 6154–6162
- 28 Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE, 2016. 779–788
- 29 Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C Y, et al. SSD: Single shot multibox detector. In: Proceedings of the 14th European Conference on Computer Vision (ECCV). Amsterdam, The Netherlands: Springer, 2016. 21–37
- 30 Lin T Y, Goyal P, Girshick R, He K M, Dollár P. Focal loss for dense object detection. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017. 2999–3007
- 31 Yi J F, Mao J X, Zhang H, Zeng K, Tao Z M, Zhong H, et al. PSTL-Net: A patchwise self-texture-learning network for transmission line inspection. *IEEE Transactions on Instrumentation and Measurement*, 2024, **73**: Article No. 5005714
- 32 Liu X Y, Miao X R, Jiang H, Chen J, Wu M, Chen Z H. Component detection for power line inspection using a graph-based relation guiding network. *IEEE Transactions on Industrial Informatics*, 2023, **19**(9): 9280–9290
- 33 Barrow H G, Tenenbaum J M. Recovering intrinsic scene characteristics from images. *Computer Vision Systems*. London: Academic Press, 1978.
- 34 Canny J. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1986, **PAMI-8**(6): 679–698
- 35 Ke L, Tai Y W, Tang C K. Occlusion-aware instance segmentation via bilayer network architectures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, **45**(8): 10197–10211
- 36 Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Proceedings of the 18th International Conference on Medical Image Computing and Computer-assisted Intervention (MICCAI). Munich, Germany: Springer, 2015. 234–241
- 37 Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, **39**(12): 2481–2495
- 38 Chen L C, Papandreou G, Kokkinos I, Murphy K, Yuille A L. Semantic image segmentation with deep convolutional nets and fully connected CRFs. arXiv: 1412.7062, 2015.
- 39 Chen L C, Zhu Y K, Papandreou G, Schroff F, Adam H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany: Springer, 2018. 833–851
- 40 Charles R Q, Hao S, Mo K C, Guibas L J. PointNet: Deep learning on point sets for 3D classification and segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017. 77–85
- 41 Charles R Q, Yi L, Su H, Guibas L J. PointNet++: Deep hierarchical feature learning on point sets in a metric space. In: Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: ACM, 2017. 5105–5114
- 42 Shi S S, Guo C X, Jiang L, Wang Z, Shi J P, Wang X G. PV-RCNN: Point-voxel feature set abstraction for 3D object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE, 2020. 10526–10535
- 43 Zhao N, Chua T S, Lee G H. SESS: Self-ensembling semi-supervised 3D object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE, 2020. 11076–11084
- 44 Wang Y, Sun Y B, Liu Z W, Sarma S E, Bronstein M M, Solomon J M. Dynamic graph CNN for learning on point clouds. *ACM Transactions on Graphics*, 2019, **38**(5): Article No. 146
- 45 Zhao Z B, Liu B, Zhai Y J, Zhao W Q, Su P. Dual graph reasoning network for oil leakage segmentation in substation equipment. *IEEE Transactions on Instrumentation and Measure-*

- ment, 2024, **73**: Article No. 3502415
- 46 Zhao Z B, Feng S, Zhai Y J, Zhao W Q, Li G. Infrared thermal image instance segmentation method for power substation equipment based on visual feature reasoning. *IEEE Transactions on Instrumentation and Measurement*, 2023, **72**: Article No. 5029613
- 47 MacQueen J. Some methods for classification and analysis of multivariate observations. In: Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability. Berkeley, USA: University of California Press, 1967. 281–297
- 48 Shi J B, Malik J. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(8): 888–905
- 49 Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems (NeurIPS). Montreal, Canada: ACM, 2014. 2672–2680
- 50 Jiang B, Zhang Z Y, Lin D D, Tang J, Luo B. Semi-supervised learning with graph learning-convolutional networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE, 2019. 11305–11312
- 51 Mehta S, Rastegari M, Caspi A, Shapiro L, Hajishirzi H. ESPNet: Efficient spatial pyramid of dilated convolutions for semantic segmentation. In: Proceedings of the 15th European Conference on Computer Vision (ECCV). Munich, Germany: Springer, 2018. 561–580
- 52 Maturana D, Scherer S. VoxNet: A 3D convolutional neural network for real-time object recognition. In: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Hamburg, Germany: IEEE, 2015. 922–928
- 53 Shi S, Wang Z, Wang X, Li H. Part- A^2 net: 3D part-aware and aggregation neural network for object detection from point cloud. *Pattern Recognition*, 2022, **122**: Article No. 108242
- 54 Zhou Y, Tuzel O. VoxelNet: End-to-end learning for point cloud based 3D object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, USA: IEEE, 2018. 4490–4499
- 55 Ouyang Z C, Dong X Y, Cui J H, Niu J W, Guizani M. PV-EncoNet: Fast object detection based on colored point cloud. *IEEE Transactions on Intelligent Transportation Systems*, 2022, **23**(8): 12439–12450
- 56 Dong Z, Li J, Li Y. 3DV-Net: Learning multi-level multi-scale 3D features from point cloud for object detection. *Sensors*, 2021, **21**(11): Article No. 3875
- 57 Shi S S, Jiang L, Deng J J, Wang Z, Guo C X, Shi J P, et al. PV-RCNN++: Point-voxel feature set abstraction with local vector representation for 3D object detection. arXiv: 2102.00463, 2022.
- 58 Hu Q Y, Yang B, Xie L H, Rosa S, Guo Y L, Wang Z H. RandLA-Net: Efficient semantic segmentation of large-scale point clouds. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE, 2020. 11105–11114
- 59 Shi W J, Rajkumar R. Point-GNN: Graph neural network for 3D object detection in a point cloud. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE, 2020. 1708–1716
- 60 Yan X, Zheng C D, Li Z, Wang S, Cui S G. PointASNL: Robust point clouds processing using nonlocal neural networks with adaptive sampling. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, USA: IEEE, 2020. 5588–5597
- 61 Li W, Luo Z P, Xiao Z L, Chen Y P, Wang C, Li J. A GCN-based method for extracting power lines and pylons from airborne LiDAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, **60**: Article No. 5700614
- 62 Jeong S, Kim D, Kim S, Ham J W, Lee J K, Oh K Y. Real-time environmental cognition and sag estimation of transmission lines using UAV equipped with 3-D Lidar system. *IEEE Transactions on Power Delivery*, 2021, **36**(5): 2658–2667
- 63 Mumir N, Awrangjeb M, Stantic B. An iterative graph-based method for constructing gaps in high-voltage bundle conductors using airborne LiDAR point cloud data. *IEEE Transactions on Geoscience and Remote Sensing*, 2024, **62**: Article No. 5700316
- 64 Huang J J, Shen Y Q, Wang J G, Ferreira V. Automatic pylon extraction using color-aided classification from UAV LiDAR point cloud data. *IEEE Transactions on Instrumentation and Measurement*, 2023, **72**: Article No. 2520611
- 65 Ibrahim M, Akhtar N, Anwar S, Mian A. SAT3D: Slot attention Transformer for 3D point cloud semantic segmentation. *IEEE Transactions on Intelligent Transportation Systems*, 2023, **24**(5): 5456–5466
- 66 Zhang Gui-Feng, Zhang Zhi-Qiang, Shen Feng. Review of the current situation and development of substation inspection robots. *Yunnan Electric Power*, 2022, **50**(6): 2–8 (张贵峰, 张志强, 沈锋. 变电站巡检机器人现状与发展综述. 云南电力技术, 2022, **50**(6): 2–8)
- 67 Yuan Li, Jiang Tian-Tian, Wei Chun-Ling, Yang Meng-Fei. Advances and perspectives of space control technology. *Acta Automatica Sinica*, 2023, **49**(3): 476–493 (袁利, 姜甜甜, 魏春岭, 杨孟飞. 空间控制技术发展展望. 自动化学报, 2023, **49**(3): 476–493)
- 68 Huang Guan-Wen, Wang Yuan-Yuan, Long Zheng-Xin, Qin Zhi-Wei, Zhang Qin. Technique advancements in GNSS satellite orbit maneuver detection. *Journal of Navigation and Positioning*, 2024, **12**(2): 1–12 (黄观文, 王媛媛, 龙正鑫, 秦志伟, 张勤. GNSS 卫星轨道机动探测技术进展. 导航定位学报, 2024, **12**(2): 1–12)
- 69 Zhang X C, Demiris Y. Visible and infrared image fusion using deep learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, **45**(8): 10535–10554
- 70 Li T, Yu H Y. Visual-inertial fusion-based human pose estimation: A review. *IEEE Transactions on Instrumentation and Measurement*, 2023, **72**: Article No. 4007816
- 71 Sun Z H, Ke Q H, Rahmani H, Bennamoun M, Wang G, Liu J. Human action recognition from various data modalities: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, **45**(3): 3200–3225
- 72 Yuan D, Zhang H P, Shu X, Liu Q, Chang X J, He Z Y, et al. Thermal infrared target tracking: A comprehensive review. *IEEE Transactions on Instrumentation and Measurement*, 2024, **73**: Article No. 5000419
- 73 Tian C, Zhou Z K, Huang Y Q, Li G J, He Z Y. Cross-modality proposal-guided feature mining for unregistered RGB-thermal pedestrian detection. *IEEE Transactions on Multimedia*, 2024, **26**: 6449–6461
- 74 Li X Y, Chen S G, Tian C N, Zhou H, Zhang Z X. M2FNet: Mask-guided multi-level fusion for RGB-T pedestrian detection. *IEEE Transactions on Multimedia*, 2024, **26**: 8678–8690
- 75 Xu Y W, Fan K G, Hu Q, Zhang X T. Positioning of suspended permanent magnet maglev trains using satellite-ground multisensor fusion. *IEEE Sensors Journal*, 2024, **24**(10): 16816–16825
- 76 Cossio-Montefinale L, Ruiz-Del-Solar J, Verschae R. Cherry CO dataset: A dataset for cherry detection, segmentation and maturity recognition. *IEEE Robotics and Automation Letters*, 2024, **9**(6): 5552–5558
- 77 Rayhana R, Ma Z Y, Liu Z, Xiao G Z, Ruan Y F, Sangha J S. A review on plant disease detection using hyperspectral imaging. *IEEE Transactions on AgriFood Electronics*, 2023, **1**(2): 108–134
- 78 Song M, Um G M, Lee H K, Seo J, Kim W. Dynamic residual filtering with Laplacian pyramid for instance segmentation.

- IEEE Transactions on Multimedia*, 2023, **25**: 6892–6903
- 79 Wu S J, Wang Y B. Seismic image dip estimation by multiscale principal component analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 2023, **61**: Article No. 5900410
- 80 Grossmann A, Morlet J. Decomposition of hardy functions into square integrable wavelets of constant shape. *SIAM Journal on Mathematical Analysis*, 1984, **15**(4): 723–736
- 81 Li X X, Cheng X Z, Xu Y W, Jiao Y, Huang W C, Cui Y Y, et al. Multi-frequency ultrasound imaging fusion method based on wavelet transform for guided screw insertion. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 2024, **71**(3): 395–407
- 82 Yang B, Li S T. Multifocus image fusion and restoration with sparse representation. *IEEE Transactions on Instrumentation and Measurement*, 2010, **59**(4): 884–892
- 83 Liu Y, Liu S P, Wang Z F. A general framework for image fusion based on multi-scale transform and sparse representation. *Information Fusion*, 2015, **24**: 147–164
- 84 Yin H T. Sparse representation with learned multiscale dictionary for image fusion. *Neurocomputing*, 2015, **148**: 600–610
- 85 Yang B, Li S T. Pixel-level image fusion with simultaneous orthogonal matching pursuit. *Information Fusion*, 2012, **13**(1): 10–19
- 86 Engan K, Aase S O, Husoy J H. Method of optimal directions for frame design. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing. Phoenix, USA: IEEE, 1999. 2443–2446
- 87 Aharon M, Elad M, Bruckstein A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 2006, **54**(11): 4311–4322
- 88 Zong Z Y, Fu T, Yin X Y. High-dimensional generalized orthogonal matching pursuit with singular value decomposition. *IEEE Geoscience and Remote Sensing Letters*, 2023, **20**: Article No. 7502205
- 89 Liu S J, Ma J J, Cui C K. FPGA implementation of threshold projection orthogonal matching pursuit algorithm for compressed sensing reconstruction. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2024, **71**(3): 1184–1197
- 90 Bai L, Yao S L, Gao K, Huang Y J, Tang R J, Yan H, et al. Joint sparse representations and coupled dictionary learning in multisource heterogeneous image pseudo-color fusion. *IEEE Sensors Journal*, 2023, **23**(24): 30620–30632
- 91 Li H, Wu X J, Kittler J. Infrared and visible image fusion using a deep learning framework. In: Proceedings of the 24th International Conference on Pattern Recognition (ICPR). Beijing, China: IEEE, 2018. 2705–2710
- 92 Ren X Y, Meng F Y, Hu T, Liu Z J, Wang C. Infrared-visible image fusion based on convolutional neural networks (CNN). In: Proceedings of the 8th International Conference on Intelligent Science and Big Data Engineering. Lanzhou, China: Springer, 2018. 301–307
- 93 Ma J Y, Yu W, Liang P W, Li C, Jiang J J. FusionGAN: A generative adversarial network for infrared and visible image fusion. *Information Fusion*, 2019, **48**: 11–26
- 94 Sun F M, Zhang K, Yuan X, Zhao C X. Feature enhancement and fusion for RGB-T salient object detection. In: Proceedings of the IEEE International Conference on Image Processing (ICIP). Kuala Lumpur, Malaysia: IEEE, 2023. 1300–1304
- 95 Zhu X, Liu J, Xiong X Z, Luo Z Q. Maximize peak-to-sidelobe ratio for real-time RGB-T tracking. *IEEE Transactions on Instrumentation and Measurement*, 2024, **73**: Article No. 4502104
- 96 Liu J, Luo Z Q, Xiong X Z. Online learning samples and adaptive recovery for robust RGB-T tracking. *IEEE Transactions on Circuits and Systems for Video Technology*, 2024, **34**(2): 724–737
- 97 Lv Y, Liu Z, Li G Y. Context-aware interaction network for RGB-T semantic segmentation. *IEEE Transactions on Multimedia*, 2024, **26**: 6348–6360
- 98 Liang W L, Shan C F, Yang Y J, Han J G. Multi-branch differential bidirectional fusion network for RGB-T semantic segmentation. *IEEE Transactions on Intelligent Vehicles*, DOI: 10.1109/TIV.2024.3374793
- 99 Chen G, Shao F, Chai X L, Chen H W, Jiang Q P, Meng X C, et al. Modality-induced transfer-fusion network for RGB-D and RGB-T salient object detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023, **33**(4): 1787–1801
- 100 Chen Jian, Liao Yan-Jun, Wang Shi, Zheng Ming-Kui, Su Li-Chao. An adaptive segmentation based multi-mode inter-frame coding method for video point cloud. *Acta Automatica Sinica*, 2023, **49**(8): 1707–1722
(陈建, 廖燕俊, 王适, 郑明魁, 苏立超. 自适应分割的视频点云多模式帧间编码方法. 自动化学报, 2023, **49**(8): 1707–1722)
- 101 Chen L, Luo X. Tensor distribution regression based on the 3D conventional neural networks. *IEEE/CAA Journal of Automatica Sinica*, 2023, **10**(7): 1628–1630
- 102 Ye S Q, Chen D D, Han S F, Liao J. Robust point cloud segmentation with noisy annotations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023, **45**(6): 7696–7710
- 103 Zhang Kai, Yang Peng-Cheng, Peng Kai-Xiang, Chen Zhi-Wen. A deep belief network-based fault evaluation method for multimode processes and its applications. *Acta Automatica Sinica*, 2024, **50**(1): 89–102
(张凯, 杨朋澄, 彭开香, 陈志文. 基于深度置信网络的多模态过程故障评估方法及应用. 自动化学报, 2024, **50**(1): 89–102)
- 104 Liu Yun-Peng, Lai Ting-Yu, Liu Jia-Shuo, Wei Xiao-Guang, Pei Shao-Tong. Vibration voiceprint characteristics and looseness detection method of UHVDC converter valve saturable reactor. *Transactions of China Electrotechnical Society*, 2023, **38**(5): 1375–1389
(刘云鹏, 来庭煜, 刘嘉硕, 魏晓光, 裴少通. 特高压直流换流阀饱和电抗器振动声纹特性与松动程度声纹检测方法. 电工技术学报, 2023, **38**(5): 1375–1389)
- 105 Khotimah K, Santoso A B, Ma'arif M, Azhiimah A N, Suprianto B, Sumbawati M S, et al. Validation of voice recognition in various Google voice languages using voice recognition module V3 based on microcontroller. In: Proceedings of the 3rd International Conference on Vocational Education and Electrical Engineering (ICVEE). Surabaya, Indonesia: IEEE, 2020. 1–6
- 106 Brunelli R, Falavigna D. Person identification using multiple cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1995, **17**(10): 955–966
- 107 Noda K, Yamaguchi Y, Nakadai K, Okuno H G, Ogata T. Audio-visual speech recognition using deep learning. *Applied Intelligence*, 2015, **42**(4): 722–737
- 108 Zhang Yu-Zhen, Wei Dai-Di, Wang Jian-Yu, Dai Yue-Wei. Semantic analysis for soccer video based on fusion of multimodal features. *Computer Science*, 2010, **37**(7): 273–276
(张玉珍, 魏带娣, 王建宇, 戴跃伟. 基于多模态融合的足球视频语义分析. 计算机科学, 2010, **37**(7): 273–276)
- 109 Xu Z J, Wang R F, Wang J, Yu D H. Parkinson's disease detection based on spectrogram-deep convolutional generative adversarial network sample augmentation. *IEEE Access*, 2020, **8**: 206888–206900
- 110 Shi Sheng-Dan, Huang Jin-Jun, Zhu Xiao-Xun, Wang Yu, Qian Bai-Yun. Partial discharge pattern recognition on transformer based on voiceprint SDP-CNN. *Electric Power Information and Communication Technology*, 2022, **20**(10): 105–112
(施胜丹, 黄金军, 朱霄珣, 王瑜, 钱白云. 基于声纹 SDP-CNN 的变压器局部放电模式识别. 电力信息与通信技术, 2022, **20**(10): 105–112)
- 111 Lu Yun-Cai, Liao Cai-Bo, Li Qun, Wang Tong-Lei, Shao Jian, Zhang Yi. Transformer fault diagnosis method based on voiceprint feature and ensemble learning. *Electric Power Engineer-*

- ing Technology, 2023, 42(5): 46–55
(陆云才, 廖才波, 李群, 王同磊, 邵剑, 张一. 基于声纹特征和集成学习的变压器缺陷诊断方法. 电力工程技术, 2023, 42(5): 46–55)
- 112 Wang Huan, Wang Xin, Zhang Feng, Qi Xiao, Chai Fang-Sen, Li Wen-Peng. Transformer voiceprint fault diagnosis based on improved generative adversarial network. *Smart Power*, 2024, 52(4): 24–31
(王欢, 王昕, 张峰, 齐笑, 柴方森, 李文鹏. 基于改进生成对抗网络的变压器声纹故障诊断. 智慧电力, 2024, 52(4): 24–31)
- 113 Lin Ying, Zhang Feng-Da, Li Zhuang-Zhuang, Zheng Wen-Jie, Ge Ning. Large model based interactive segmentation of infrared image for power equipment. *Network New Media Technology*, 2024, 13(2): 53–60, 67
(林颖, 张峰达, 李壮壮, 郑文杰, 戈宁. 基于大模型的红外图像电力设备交互式分割. 网络新媒体技术, 2024, 13(2): 53–60, 67)
- 114 Yang Quan, Fan Shao-Sheng. Visual navigation method for electric power inspection robot based on image preprocessing and semantic segmentation. *Journal of Electric Power Science and Technology*, 2023, 38(6): 248–258
(杨权, 樊绍胜. 基于图像预处理和语义分割的电力巡检机器人视觉导航方法. 电力科学与技术学报, 2023, 38(6): 248–258)
- 115 Ma Qing-Shan, Zhu Jian-Bao, Yu Xin-Chun, Zhang Bin. Power safety belt segmentation algorithm based on improved DSD-LinkNet. *Electrical Automation*, 2023, 45(3): 106–108
(马青山, 朱建宝, 俞鑫春, 张斌. 基于改进 DSD-LinkNet 的电力安全带分割算法. 电气自动化, 2023, 45(3): 106–108)
- 116 Xu Gang, Li Guo. Research on lightweight neural network of aerial powerline image segmentation. *Journal of Image and Graphics*, 2021, 26(11): 2605–2618
(许刚, 李果. 轻量化航拍图像电力线语义分割. 中国图象图形学报, 2021, 26(11): 2605–2618)
- 117 Zuo An-Quan, Qin Lun-Ming, Wang Xi, Bian Hou-Qin, Chen Si-Lin. Powerline semantic segmentation method based on improved DeepLabv3+ model. *Radio Engineering*, 2023, 53(1): 96–104
(左安全, 秦伦明, 王悉, 边后琴, 陈思林. 基于改进 DeepLabv3+ 模型的电力线语义分割方法. 无线电工程, 2023, 53(1): 96–104)
- 118 Cui C, Gao T Q, Wei S Y, Du Y N, Guo R Y, Dong S L, et al. PP-LCNet: A lightweight CPU convolutional neural network. arXiv: 2109.15099, 2021.
- 119 Chollet F. Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, USA: IEEE, 2017. 1800–1807
- 120 Park J, Woo S, Lee J Y, Kweon I S. BAM: Bottleneck attention module. arXiv: 1807.06514, 2018.
- 121 Mao M F, Chen Y, Chen W X, Du W, Zhang M, Mao T Q. Power transmission line image segmentation method based on binocular vision and feature pyramid network. In: Proceedings of the International Conference on Sensing, Measurement & Data Analytics in the Era of Artificial Intelligence (ICSMD). Nanjing, China: IEEE, 2021. 1–4
- 122 Wang L J, Chen Z L, Hua D, Zheng Z X. Semantic segmentation of transmission lines and their accessories based on UAV-taken images. *IEEE Access*, 2019, 7: 80829–80839
- 123 Yang L, Kong S Y, Cui S L, Huang H Y, Liu Y H. An efficient end-to-end CNN network for high-voltage transmission line segmentation. In: Proceedings of the 8th International Conference on Cloud Computing and Intelligent Systems (CCIS). Chengdu, China: IEEE, 2022. 565–570
- 124 Tang Xiao-Yu, Huang Jin-Bo, Feng Jie-Wen, Chen Xi-He. Image segmentation and defect detection of insulators based on U-net and YOLOv4. *Journal of South China Normal University (Natural Science Edition)*, 2020, 52(6): 15–21
(唐小煜, 黄进波, 冯洁文, 陈锡和. 基于 U-net 和 YOLOv4 的绝缘子图像分割与缺陷检测. 华南师范大学学报(自然科学版), 2020, 52(6): 15–21)
- 125 Yu Y J, Cao H, Wang Z Z, Li Y Q, Li K, Xie S Q. Texture-and-shape based active contour model for insulator segmentation. *IEEE Access*, 2019, 7: 78706–78714
- 126 Tang X, Mao J X, Yi J F, Tao Z M, He Z Y. An insulator feature-enhanced segmentation method based on the residual-type attention mechanism. In: Proceedings of the 5th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC). Chongqing, China: IEEE, 2022. 1507–1511
- 127 Han Y F, Han J, Ni Z J, Wang W S, Jiang H Y. Instance segmentation of transmission line images based on an improved D-SOLO network. In: Proceedings of the 3rd International Conference on Power Data Science (ICPDS). Harbin, China: IEEE, 2021. 40–46
- 128 Choi H, Yun J P, Kim B J, Jang H, Kim S W. Attention-based multimodal image feature fusion module for transmission line detection. *IEEE Transactions on Industrial Informatics*, 2022, 18(11): 7686–7695
- 129 Xu C, Li Q W, Jiang X B, Yu D B, Zhou Y Q. Dual-space graph-based interaction network for RGB-thermal semantic segmentation in electric power scene. *IEEE Transactions on Circuits and Systems for Video Technology*, 2023, 33(4): 1577–1592
- 130 Ma J L, Qian K, Zhang X B, Ma X D. Weakly supervised instance segmentation of electrical equipment based on RGB-T automatic annotation. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(12): 9720–9731
- 131 Sun Yang, Li Jia. YOLOv7-tiny transmission line foreign object detection algorithm based on channel pruning. *Computer Engineering and Applications*, 2024, 60(14): 319–328
(孙阳, 李佳. 基于通道剪枝的 YOLOv7-tiny 输电线路异物检测算法. 计算机工程与应用, 2024, 60(14): 319–328)
- 132 Shen Xiao-Feng, Xie Wei, Sun Lu, Li Yi, He Run-Ping. Research on foreign object target detection on transmission lines based on adjacent frame difference method. *Process Automation Instrumentation*, 2023, 44(10): 20–24
(沈晓峰, 谢伟, 孙路, 李轶, 贺润平. 基于相邻帧差法的输电线路异物目标检测研究. 自动化仪表, 2023, 44(10): 20–24)
- 133 Lowe D G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004, 60(2): 91–110
- 134 Yu Yan-Zhen, Qiu Zhi-Bin, Zhou Yin-Biao, Zhu Xuan, Wang Qing. Foreign body detection for transmission lines based on convolutional neural network and ECOC-SVM. *Smart Power*, 2022, 50(3): 87–92, 107
(余沿臻, 邱志斌, 周银彪, 朱轩, 王青. 基于卷积神经网络与 ECOC-SVM 的输电线路异物检测. 智慧电力, 2022, 50(3): 87–92, 107)
- 135 Jiang H, Huang W L, Chen J, Liu X Y, Miao X R, Zhuang S B. Detection of bird nests on power line patrol using single shot detector. In: Proceedings of the Chinese Automation Congress (CAC). Hangzhou, China: IEEE, 2019. 3409–3414
- 136 Yang Zheng. Tilt Monitoring System of Transmission Line Towers Based on the ZigBee and GSM Technology [Master thesis], North China Electric Power University, China, 2012.
(杨征. 基于 ZigBee 和 GSM 技术输电线路杆塔倾斜监测系统研究 [硕士学位论文], 华北电力大学, 中国, 2012.)
- 137 Li Q Q, Chen Z P, Hu Q W. A model-driven approach for 3D modeling of pylon from airborne LiDAR data. *Remote Sensing*, 2015, 7(9): 11501–11524
- 138 Shen Xiao-Jun, Qin Chuan, Du Yong, Yu Xin-Le. An automatic power line extraction method from airborne light detection and ranging point cloud in complex terrain. *Journal of Tongji University (Natural Science)*, 2018, 46(7): 982–987
(沈小军, 秦川, 杜勇, 于忻乐. 复杂地形电力线机载激光雷达点云自动提取方法. 同济大学学报(自然科学版), 2018, 46(7): 982–987)
- 139 Liu Chang-An, Sun Shu-Ming, Zhao Li-Juan. A lightweight network for power tower extraction from laser point cloud. *Laser Technology*, 2021, 45(3): 367–372
(柳长安, 孙书明, 赵丽娟. 基于激光点云实现杆塔提取的轻量级

网络. 激光技术, 2021, 45(3): 367-372)

- 140 Lu Zhu-Mao, Gong Hao, Jin Qiu-Heng, Hu Qing-Wu, Li Jia-Yuan. Tilt rate measurement of power tower based on UAV LiDAR point cloud. *Journal of Applied Sciences*, 2022, 40(3): 389-399
(芦竹茂, 龚浩, 金秋衡, 胡庆武, 李加元. 无人机激光雷达点云电力杆塔倾斜状态测量. 应用科学学报, 2022, 40(3): 389-399)
- 141 Xu Liang-Gang, Shi Lei, Chen Feng-Xiang, Wang Shi-Chun, Long Xin, Wang Di. Transmission line tower tilt detection algorithm based on laser point cloud. *Laser Technology*, 2022, 46(3): 390-396
(徐梁刚, 时磊, 陈凤翔, 王时春, 龙新, 王迪. 基于激光点云的输电线路杆塔倾斜检测算法. 激光技术, 2022, 46(3): 390-396)
- 142 Ying Bin, Tang Bin, Pan Jun-Jie, Guo Zhen. Research on a detection algorithm of power tower based on YOLOv3. *Zhejiang Electric Power*, 2021, 40(5): 53-59
(应斌, 唐斌, 潘俊杰, 郭震. 基于 YOLOv3 的电力杆塔检测算法研究. 浙江电力, 2021, 40(5): 53-59)
- 143 Hao Mei. Key of icing online monitoring for transmission line. *Electronic Technology*, 2020, 49(3): 90-91
(郝美. 输电线路覆冰在线监测的关键技术. 电子技术, 2020, 49(3): 90-91)
- 144 Huang T T, Liu Z, Chen X W, Bai X. EPNet: Enhancing point features with image semantics for 3D object detection. In: Proceedings of the 16th European Conference on Computer Vision (ECCV). Glasgow, UK: Springer, 2020. 35-52
- 145 An Qiu-E. Equipment ledger and spare parts management. *Chemical Equipment and Anticorrosion*, 1999(4): 27-28, 62
(安秋娥. 设备台帐与备件管理. 化工设备与防腐蚀, 1999(4): 27-28, 62)



张 辉 湖南大学机器人学院教授. 主要研究方向为机器视觉, 图像处理和机器人控制.

E-mail: zhanghuihy@126.com

(ZHANG Hui Professor at the School of Robotics, Hunan University. His research interest covers

machine vision, image processing, and robot control.)



杜 瑞 湖南大学机器人学院博士研究生. 2020 年获得湘潭大学硕士学位. 主要研究方向为机器视觉, 图像处理. 本文通信作者.

E-mail: durui@hnu.edu.cn

(DU Rui Ph.D. candidate at the School of Robotics, Hunan University. He received his master degree from Xiangtan

University in 2020. His research interest covers ma-

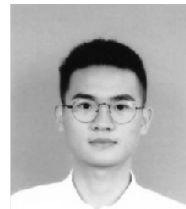
chine vision and image processing. Corresponding author of this paper.)



钟 杭 湖南大学机器人学院副教授. 分别于 2013 年、2016 年和 2020 年获得湖南大学电气与信息工程学院自动化科学专业学士、硕士和博士学位. 主要研究方向为航空机器人, 多机器人系统, 视觉伺服, 视觉导航和非线性控制.

E-mail: zhonghang@hnu.edu.cn

(ZHONG Hang Associate professor at the School of Robotics, Hunan University. He received his bachelor degree, master degree and Ph.D. degree in automation science from the College of Electrical and Information Engineering, Hunan University in 2013, 2016 and 2020, respectively. His research interest covers aerial robotics, multi-robot systems, visual servoing, visual navigation, and nonlinear control.)



曹意宏 湖南大学机器人学院博士研究生. 2021 年获得湖南师范大学硕士学位. 主要研究方向为计算机视觉.

E-mail: caoyihong@hnu.edu.cn

(CAO Yi-Hong Ph.D. candidate at the School of Robotics, Hunan University. He received his master degree from Hunan Normal University in 2021. His main

research interest is computer vision.)



王耀南 中国工程院院士, 湖南大学电气与信息工程学院教授. 1995 年获得湖南大学博士学位. 主要研究方向为机器人学, 智能控制和图像处理.

E-mail: yaonan@hnu.edu.cn

(WANG Yao-Nan Academician at Chinese Academy of Engineering,

professor at the College of Electrical and Information Engineering, Hunan University. He received his Ph.D. degree from Hunan University in 1995. His research interest covers robotics, intelligent control, and image processing.)