

# 基于视觉与深度学习的无人机自主着陆场景感知方法研究进展

王中天, 吴一全\*

(南京航空航天大学电子信息工程学院, 江苏南京 211106)

**摘要:** 随着无人机(Unmanned Aerial Vehicle, UAV)技术的蓬勃发展,其在军事国防、智能交通、设施巡检、灾害救援、农业管理等众多领域的应用日益广泛,成为低空经济发展的核心驱动力。自主着陆作为无人机关键核心技术之一,直接决定了无人机作业的安全性与可靠性,尤其在电池电量不足、气象条件恶化或通信中断等紧急场景下,能有效避免设备损坏与事故发生,是实现无人机完全自动化的关键环节。基于视觉与深度学习的场景感知技术,凭借强大的特征学习和模式识别能力,突破了传统依赖GPS(Global Positioning System)、激光雷达(Light Detection And Ranging, LiDAR)等技术在复杂环境中的局限性,为无人机自主着陆领域带来了全新的解决方案。本文系统综述了基于视觉与深度学习的无人机自主着陆场景感知方法。首先阐述了深度学习在无人机自主着陆中的应用背景和重要性,梳理了从传统传感器驱动到智能感知的技术演进历程。随后详细剖析了不同场景的特征与技术挑战:静态平台着陆聚焦降落标识、跑道检测、地基引导三类场景,核心需求是提升着陆精度与准确率;动态平台着陆涵盖车载陆地、舰艇海上及其他移动平台,需重点解决运动跟踪与干扰抑制问题;特殊场景着陆则面临山区、森林、城市峡谷等复杂环境中的障碍物遮挡、信号干扰、极端气象等多重挑战。本文深入探讨了核心技术体系,包括目标检测、语义分割、姿态估计、光流预测、三维重建等关键技术的原理与应用。同时分析了特征提取优化、语义理解增强及场景适配策略的应用效果与性能表现。最后总结了该领域面临的复杂环境适应性不足、计算资源约束、数据依赖与标注难题等挑战,并对未来研究方向进行了展望。指出通过多源传感器数据融合可提升复杂环境感知能力,开发轻量化模型能适配无人机资源限制,加强仿真与真实场景结合可提高模型泛化能力。本文通过系统地总结与分析,全面呈现了该领域的技术现状与发展脉络,为无人机自主着陆技术的进一步研究与工程应用提供了宝贵的参考和指导。

**关键词:** 自主着陆;无人机;深度学习;计算机视觉;目标检测;语义分割

**基金项目:** 国家自然科学基金(No.61573183)

**中图分类号:** TP391.41;V279

**文献标识码:** A

**文章编号:** 0372-2112(2025)11-4171-28

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20250775

## Research Progress of UAV Autonomous Landing Scene Perception Methods Based on Vision and Deep Learning

WANG Zhong-tian, WU Yi-quan\*

(College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 211106, China)

**Abstract:** With the vigorous development of unmanned aerial vehicle (UAV) technology, its applications in various fields such as military defense, intelligent transportation, facility inspection, disaster relief, and agricultural management have become increasingly widespread, becoming the core driving force for the development of the low-altitude economy. Autonomous landing, as one of the core and key technologies of UAVs, directly determines the safety and reliability of UAV operations. Especially in emergency scenarios such as low battery power, deteriorating weather conditions, or communication disruptions, it can effectively prevent equipment damage and accidents, and is a crucial step towards achieving full automation of UAVs. Scene perception technology based on vision and deep learning, with its powerful feature learning and pattern recognition capabilities, has broken through the limitations of traditional technologies such as GPS (Global Positioning System) and LiDAR (Light Detection And Ranging) in complex environments, bringing a brand-new solution to the

field of UAV autonomous landing. This paper systematically reviews the scene perception methods for UAV autonomous landing based on vision and deep learning. Firstly, it elaborates on the application background and significance of deep learning in UAV autonomous landing, and sorts out the technological evolution from traditional sensor-driven to intelligent perception. Then, it analyzes in detail the features and technical challenges of different scenarios: static platform landing focuses on three types of scenarios - landing marks, runway detection, and ground guidance, with the core demand being to improve landing accuracy and reliability; dynamic platform landing covers land-based vehicles, ships at sea, and other mobile platforms, and needs to focus on solving problems of motion tracking and interference suppression; special scenario landing faces multiple challenges such as obstacle occlusion, signal interference, and extreme weather in complex environments like mountains, forests, and urban canyons. This paper deeply explores the core technical system, including the principles and applications of key technologies such as object detection, semantic segmentation, pose estimation, optical flow prediction, and 3D reconstruction. At the same time, it analyzes the application effects and performance of feature extraction optimization, semantic understanding enhancement, and scene adaptation strategies. Finally, it summarizes the challenges faced in this field, such as insufficient adaptability to complex environments, computational resource constraints, data dependence and annotation difficulties, and looks forward to future research directions. It points out that multi-source sensor data fusion can enhance the perception ability in complex environments, developing lightweight models can adapt to the resource limitations of UAVs, and strengthening the combination of simulation and real scenarios can improve the generalization ability of models. Through systematic summary and analysis, this paper comprehensively presents the current technical status and development trends in this field, providing valuable reference and guidance for further research and engineering applications of UAV autonomous landing technology.

**Key words:** autonomous landing; UAV; deep learning; computer vision; object detection; semantic segmentation

**Foundation Item(s):** National Natural Science Foundation of China (No.61573183)

## 1 引言

无人机是利用无线电遥控设备和自备的程序控制装置的不载人飞机,它可以按照预设程序自主飞行,也可以通过地面控制站进行远程操作,具有灵活性和可控性. 无人机具有多样化的结构设计,根据任务需求,可采用多旋翼、固定翼、混合翼等不同设计,具有不同的飞行特性和环境适应性. 近年来,在军事国防、智能交通、设施巡检、地形测绘、灾害救援、农业管理、环境监测、影视新闻等领域<sup>[1,2]</sup>,无人机不仅提供了新的解决方案,也推动了相关技术的进步. 以无人机为主体的低空经济的快速发展表明它在社会和经济发展中具有巨大的潜力和重要性,随着5G、人工智能和大数据技术的结合,无人机在精细化管理、自动化作业、未来智慧城市、无人驾驶等领域的作用将更加显著.

无人机自主着陆的研究极大提升了无人机的智能化水平,使其能够在无人干预的情况下,安全、精确地降落,这项研究推动了无人机自主飞行技术的发展,是实现无人机完全自动化的重要步骤之一<sup>[3]</sup>. 尤其是在环境复杂的场景下. 例如,当无人机飞行过程中出现电池电量不足、气象条件恶化或通信中断等紧急情况时,自主着陆技术可以确保无人机找到安全着陆点,避免因电力耗尽或失控造成损坏或事故,提高了无人机在多样化任务中的安全性和可靠性. 随着这一技术的不断成熟,无人机的应用场景将进一步扩展.

传统的无人机自主着陆技术主要依赖于精确的传

感器数据与导航系统,基于不同感知原理形成多种技术路径. (1)卫星与惯性导航,采用全球定位系统和惯性导航系统(Inertial Navigation System, INS)实现精确定位,在开阔区域表现良好,但在城市高楼、森林等信号遮挡环境中性能受限<sup>[4]</sup>. (2)计算机视觉感知,利用摄像头实时识别地面场景与着陆点,适用于无GPS(Global Positioning System)环境,但对光照条件与图像噪声敏感. (3)激光雷达探测,通过激光束获取三维环境信息规划着陆区域,可应对复杂地形,然而设备成本高、功耗大<sup>[5]</sup>. (4)地标匹配导航,依赖预先部署的地面标志或路径,适用于固定环境,但牺牲了系统灵活性. (5)多传感器融合,整合视觉与GPS数据,补偿定位偏差并校正累积误差,提升复杂环境适应性,却面临数据同步、传感器标定及计算复杂度高等挑战<sup>[6]</sup>. 这些传统方法虽各有优势,但在动态场景适应性与复杂环境鲁棒性上存在局限.

随着深度学习与人工智能的发展,无人机自主着陆技术迎来革新. 2012年 AlexNet 的提出推动卷积神经网络(Convolutional Neural Network, CNN)在视觉领域的应用,其端到端学习模式可直接从传感器数据提取着陆特征<sup>[7]</sup>. 针对时间序列信息,循环神经网络(Recurrent Neural Network, RNN)及长短期记忆网络(Long Short-Term Memory, LSTM)通过历史状态建模增强动态环境感知能力<sup>[8,9]</sup>. 2017年 Transformer 模型凭借自注意力机制,在多传感器数据融合中高效捕获长距离依赖关系,显著提升模型训练效率与复杂场景适应性<sup>[10]</sup>. 相较于

传统依赖预设规则的感知方式,深度学习驱动的技术通过数据驱动的端到端学习与多模态融合,大幅提升了无人机在复杂环境下的感知精度与自主着陆能力,标志着无人机自主着陆从“传感器驱动”向“智能感知”的演进升级.未来研究将持续聚焦算法鲁棒性与计算效率优化,推动技术在更多场景的实际应用.

谷美颖等人<sup>[11]</sup>全面总结了基于视觉的无人机定位与导航技术,涵盖了图像检索和图像匹配两大类方法,并对其技术特点、适用场景进行了详细分析,但对算法在实际复杂环境下的泛化能力和鲁棒性讨论不足,限制了对算法实用性的深入理解.钟春来等人<sup>[12]</sup>系统总结了基于视觉的无人机自主着陆研究,围绕相机安装方式、着陆平台类型及关键技术对比展开,但分析较为局限,对各类算法的对比研究不够具体.Katkuri等人<sup>[13]</sup>回顾深度学习驱动的计算机视觉在无人机导航中的应用,梳理技术现状与研究轨迹,但未深入探讨技术挑战.马宁等人<sup>[14]</sup>归纳近十年视觉感知与位姿估计算法,重点分析SIFT(Scale-Invariant Feature Transform)与CSRT(Channel and Spatial Reliability Tracker)结合的跑道检测算法,但未探讨自主着陆问题中除位姿估计外其他任务.Hasanalian等人<sup>[15]</sup>全面综述无人机分类、应用及设计挑战,覆盖微型到混合无人机的多场景潜力,但其广泛性导致特定技术细节分析不足,且未纳入最新研究成果.赵良玉等人<sup>[16]</sup>聚焦无人机硬件微型化与低成本化,但对深度学习目标检测算法的研究仅限于2020年及之前.Xin等

人<sup>[17]</sup>将视觉自主着陆分为静态、动态、复杂场景并细分,涉及合作目标、自然场景及多传感器融合,然而缺乏对各分类下具体算法实现的深入解析.Chen等人<sup>[18]</sup>基于YOLO(You Only Look Once)技术筛选分析无人机自主着陆等相关研究,但其对技术落地面临的挑战与限制讨论不够全面,影响读者对其实用性的完整认知.综上,目前发表的综述性文献对于基于视觉与深度学习的无人机自主着陆场景感知的各类关键任务和技术尚需系统全面地阐述.

本文聚焦视觉与深度学习在无人机自主着陆场景感知中的应用,剖析技术原理、典型场景及发展趋势,通过文献分析总结优势与不足,为后续研究提供参考.

## 2 无人机自主着陆的场景研究

本节主要呈现无人机自主着陆场景研究体系,如表1所示.首先围绕降落标识、跑道检测、地基引导三类静平台介绍静态平台着陆场景,接着介绍包含车载、舰艇及其他动平台的动态平台着陆场景,最后介绍特殊着陆场景,剖析各场景关键挑战,如静态聚焦着陆精度与准确率提升,动态侧重着陆时间与干扰率减少,并对应阐释无人机自主着陆关键技术.感知算法涵盖目标检测、语义分割等技术,模型改进依托以改变模型结构为主的特征提取、加入注意力和Transformer结构为主的语义理解,通过网络结构优化等实现;场景适配通过鲁棒化、轻量化,以数据增广、模型蒸馏等达成,构建“场景-挑战-技术”关联框架,为无人机自主着陆研究提供清晰技术路径与逻辑脉络.

表1 无人机自主着陆场景、挑战与对应技术总览表

研究场景	场景分类	主要内容	关键挑战	主要内容	技术分类	研究技术
无人机自主着陆场景研究	静态平台	基于降落标识的静平台	提高着陆精度和准确率	通过优化网络结构设计、引入注意力或Transformer结构等方法	特征提取	模型改进
		基于跑道检测的静平台			语义理解	
		基于地基引导的静平台				
	动态平台	车载陆地动平台	减少着陆时间和干扰	通过增广和自建数据集、剪枝蒸馏量化等方法	鲁棒化	场景适配
		舰艇海上动平台			轻量化	
		其他动平台				
特殊场景	山区、森林、城市峡谷、其他场景	复杂环境、遮挡信号、极端气象、感知干扰、其他挑战	目标检测、语义分割、姿态估计、其他算法	感知算法		

### 2.1 静态平台着陆

静态平台着陆是无人机自主着陆中较为基础的场景,着陆平台位置固定且姿态不变,如空旷场地的固定停机坪、建筑物顶部预设区域等.基于深度学习的方法在此场景下主要通过CNN提取图像特征,快速准确检测识别静态平台上的标识物(如特定颜色、形状图案),经大量不同背景和光照条件的标识物图像训练,模型

能适应复杂环境,定位标识物位置并计算无人机与平台的相对位置和姿态关系,即便在不同光照强度和角度下,也能通过学习标识物的鲁棒特征实现精确识别,为无人机着陆提供引导<sup>[19]</sup>.

#### 2.1.1 基于降落标识的静态着陆场景

在基于降落标识的静态着陆场景中,着陆平台常设置“T”形、“H”形、AprilTag码、QR码、菱形标记、圆形

图案等特定标识物<sup>[20]</sup>,如图1所示.其中,“T”形标识通过边缘检测与矩特征提取,利用独特结构稳定实时指引着陆;“H”形标识借助改进的YOLO、SSD(Single Shot multibox Detector)等模型提升小尺寸和远距标识的检测精度,保障跟踪安全;菱形标记以四条边和四个角的丰富几何特征与背景区分,精准指引着陆;圆形图案如圆环形或复合圆形标识,利用各向同性特点在不同视角下易识别,可抗噪声干扰,提高复杂环境检测精度.



(a) “T”形标识 (b) “H”形标识 (c) 圆形标识 (d) 菱形标识

图1 常见的单一形状降落标识

常见组合降落标识通过结合多种形状与特征,比单一标识提供更丰富信息,可提升无人机自主着陆精度与可靠性.如复合图形标识将“T”与圆形、“H”与矩形等不同形状组合,提供更多可识别特征,像结合“T”形快速定位方向和圆形便于复杂背景识别的优势,适用于多种着陆场景,如图2所示.

含颜色、温度辅助信息的标识,通过加入高对比度颜色组合或温度特征,可在不同光照或热成像下增强辨识度,辅助无人机在强光、阴影、烟雾等复杂环境中精准识别着陆点.

基于视觉基准标识系统的组合标识通常以AprilTag码、QR码、ArUco码等为基础设计而成<sup>[21]</sup>,如



(a) 组合标识1 (b) 组合标识2 (c) 组合标识3 (d) 组合标识4

图2 常见的组合降落标识

基于AprilTag设计的复合着陆图标,通过两级级联检测网络,可在无人机着陆不同高度阶段发挥作用,提高视觉引导信息采集的有效性和准确性.如图3是AprilTag码和QR码.



(a) QR码 (b) AprilTag码

图3 常见的视觉基准标识

它们均用于提供定位和引导信息,但在具体应用方式和特点上有所不同. AprilTag码是专门为机器视觉设计的视觉基准标识系统,帮助无人机在不同高度阶段采集有效视觉引导信息,实现精准着陆. QR码是常见的二维码,信息存储量大、编码范围广、识别速度快且有纠错能力<sup>[22]</sup>. 无人机通过摄像头等视觉传感器获取含标识物的图像,经深度学习算法处理分析,识别标识物的位置、姿态和形状等特征,据此计算自身相对着陆点的位置和姿态偏差,调整飞行轨迹实现精确着陆,相关研究场景见表2.

表2 基于降落标识的静态着陆场景

研究方向	文献	文献类型	着陆场景	模型	实验结果
目标检测	Mu(2023) <sup>[23]</sup>	期刊文章	小黑白方格间隔	YOLOv5	精度为6.6%,着陆速度为0.1 m/s
语义分割	Wu(2024) <sup>[24]</sup>	会议论文	各类静态目标	Deeplabv3+	精度达到93.8%
光流预测	Wang(2020) <sup>[25]</sup>	会议论文	AprilTag标识	FlowNet2	平均误差为21.4%,CPU耗时462 ms
图像分割	乐辉(2022) <sup>[26]</sup>	学位论文	H形标识	CNN	着陆标识图像的分类准确率达到了98.18%

以二维码标识为例,研究人员利用基于CNN的YOLOv5算法对无人机拍摄的图像进行处理.经灰度化、降噪预处理后,图像输入模型,通过卷积层、池化层提取特征,全连接层与分类器完成二维码检测与感知.经过实验,模型的IoU值、精度、召回率分别达到62%、73%、35%,该方法在不同光照条件和背景干扰下都能准确识别,为无人机的精确着陆提供了可靠依据<sup>[23]</sup>.

### 2.1.2 基于跑道检测的静态着陆场景

对于一些需要较长滑行距离的固定翼无人机,通常采用基于跑道检测的静态着陆方式,如图4所示<sup>[27]</sup>.

在这种场景下,无人机需要识别跑道的位置、方向和边界等信息.深度学习在跑道检测中发挥重要作用,



图4 固定翼无人机着陆跑道检测<sup>[27]</sup>

表3为基于跑道检测的静态着陆场景.

以基于目标检测的跑道检测为例,研究人员提出基于改进Faster RCNN的跑道检测方法,通过选择性搜索算法生成图像候选区域,输入模型完成分类回归.经大量数据训练后,模型可精准识别跑道区域.实际飞行

表 3 基于跑道检测的静态着陆场景

研究方向	文献	文献类型	着陆场景	模型	实验结果
目标检测	方琪鸿(2024) <sup>[27]</sup>	学位论文	机场	YOLOv5	距机场端头中点距离 1 000 m 内,误差控制在 5 m 以内,满足无人机自主降落需求
目标检测	Ma(2023) <sup>[28]</sup>	期刊文章	操场	YOLOv7-tiny	精度为 90.5%,mAP0.5 为 93.3%,参数量为 6.2 M
目标检测	Ma(2022) <sup>[29]</sup>	期刊文章	跑道	YOLOv5	mAP 达到 99.5%,mIoU 达到 95.4% 且推理速度达到 64.5 FPS
目标检测	Liu(2024) <sup>[30]</sup>	期刊文章	机场	YOLOX	x 方向的定位误差达到跑道附近 2 m 左右, y 方向上,定位误差约为 2.7 m,在 z 方向上,定位误差约为 1.5 m
目标检测	Li(2024) <sup>[31]</sup>	期刊文章	跑道	YOLO-RWY	AP0.5-0.95 为 76%,参数量为 3.106 M,推理速度为 125.3 FPS
目标检测	Wang(2023) <sup>[32]</sup>	期刊文章	跑道	Faster RCNN	AP 值为 84.37%,提高了 1.85 个百分点
语义分割	Wang (2024) <sup>[33]</sup>	期刊文章	跑道	VALNet	RLD 上 mAP 达到 69.4%,推理速度达到 75.9 FPS,参数量为 12.4 M,GFLOPs 为 43.1
姿态估计	马宁(2022) <sup>[34]</sup>	期刊论文	雨天、雪天、沙漠、城市	Q-YOLOv5	mAP0.5 为 99.5%,mAP0.5:0.95 为 95.4%,速度为 125 FPS,姿态角平均误差为 2.01°,位置平均误差为 0.32 m
高度估计	Zhang(2021) <sup>[35]</sup>	期刊文章	跑道	VIAE-Net	飞行高度为 15~60 m 范围内的测试数据上的 MAE、RMSE 和决定系数 $R^2$ 结果为 0.136 8、0.171 5、0.999 3

试验 AP 值为 84.37%,该方法助力无人机有效检测跑道并使之成功着陆<sup>[32]</sup>。

### 2.1.3 基于地基引导的静态着陆场景

基于地基引导的静态着陆场景中,如地面基站、导航灯塔等地面设备会向无人机发送引导信号,帮助无人机确定着陆位置和轨迹。无人机通过接收这些信号,结合自身的传感器数据实现自主着陆。深度学习技术

可以用于优化地基引导系统的性能,提高着陆的准确性和可靠性<sup>[36]</sup>,如表 4 所示。

利用深度学习算法对地基引导信号进行处理和分析,预测信号的传输延迟和干扰情况,从而调整无人机的接收策略,提高信号的稳定性和准确性。通过对历史信号数据的学习,模型能够准确预测信号的变化趋势,为无人机提供更可靠的引导信息。

表 4 基于地基引导的静态着陆场景

研究方向	文献	文献类型	研究场景	模型	实验结果
目标检测	Khan(2023) <sup>[37]</sup>	会议论文	静、动态着陆	TransLearn-YOLOv7	平均分类精度为 94%,平均召回率为 93.1%,mAP0.5 均为 95.3%,平均 F1 得分为 0.923,推理速度为 24 FPS
目标检测	Xie(2023) <sup>[38]</sup>	会议论文	静态着陆	Faster RCNN	mAP 为 66.58%,提高了 4.62 个百分点
姿态估计	Tang(2023) <sup>[39]</sup>	期刊文章	静态着陆	Block-CNN	锚点检测精度提升了 3 个百分点,位置和姿态估计精度分别提高了 19.5 和 12.7 个百分点

### 2.1.4 小结

静态平台三种着陆方式的难易与方法选择差异明确,排序如下:降落标识着陆(易)<跑道检测着陆(中)<地基引导着陆(难)。降落标识着陆依赖预设视觉标识,特征鲜明、部署灵活,适配低成本场景;跑道检测着陆需识别大面积低纹理区域,应对光照与背景干扰,满足长距离着陆需求;地基引导着陆涉及多系统协同与干扰抑制,技术链路长,实现厘米级高精度定位。静态平台着陆的关键难点是提升精度与准确率,需通过优化网络结构、引入注意力或 Transformer 结构,在特征提取与语义理解层面改进模型,应对技术挑战。

### 2.2 动态平台着陆

动态平台着陆场景难度显著增加,因着陆平台处于运动状态,需无人机实时跟踪平台动态并调整飞行轨迹与姿态以安全着陆。常见动态平台有移动车辆、

舰船等,如无人机降落移动车辆时,先通过深度学习目标检测模型识别车辆,再利用跟踪算法结合自身传感器数据预测车辆位置,据此调整轨迹实现精准着陆<sup>[40,41]</sup>。

#### 2.2.1 车载陆地动平台

车载陆地动平台着陆时,无人机需要考虑车辆的行驶速度、方向和颠簸等因素。如图 5 所示是一种基于视觉和惯性融合的无人机在车载平台上自主着陆方法<sup>[42]</sup>。

无人机通过搭载的视觉传感器实时监测车辆上的标识物,利用深度学习算法识别标识物的位置和姿态变化,同时结合视觉和深度学习对无人机和车辆的相对姿态进行估计。根据估计结果,无人机调整飞行轨迹,实现对车辆的跟踪和着陆<sup>[43-46]</sup>,如表 5 所示。

在实际实验中,将该方法应用于四旋翼无人机在行驶车辆上的自主着陆。实验结果表明,该方法能够有

图5 无人机车载动平台着陆<sup>[42]</sup>

效跟踪车辆的运动,在车辆行驶速度为40 km/h的情况下,无人机多次成功实现稳定着陆,着陆误差率仅为

3%,在较小范围内<sup>[47]</sup>.

### 2.2.2 舰艇海上动平台

舰艇海上动平台着陆面临着更为复杂的环境,如海浪、海风、舰艇的摇摆和起伏等.为了实现无人机在舰艇上的自主着陆,研究人员提出了多种基于深度学习的方法,如表6所示.

利用深度学习算法对舰艇的运动姿态进行预测,通过对舰艇的历史运动数据和传感器信息进行学习,建立舰艇运动模型.在无人机着陆过程中,模型根据实时传感器数据预测舰艇的未来运动姿态,为无人机提供提前运动补偿信息,使其能够更好地适应舰艇的动态变化.

表5 车载陆地动平台

研究方向	文献	文献类型	着陆场景	模型	实验结果
目标检测	王忠言(2022) <sup>[47]</sup>	期刊论文	木板车	YOLOv4	mAP为75%
目标检测	Idrovo(2024) <sup>[48]</sup>	预印本	汽车	YOLOv8x	mAP0.5达到97.7%,mAP0.5-0.95为78.2%,F1得分为0.95
目标检测	Shen(2023) <sup>[49]</sup>	期刊文章	车载标识	AeroNet	检测的准确率为89%,在Intel NUC上和Nvidia TX2上的计算时间分别仅为76 ms和28 ms

表6 舰艇海上动平台

研究方向	文献	文献类型	着陆识别场景	模型	实验结果
目标检测	Ma(2023) <sup>[28]</sup>	期刊文章	船舶平台	YOLOv7-tiny	精度达到90.5%,mAP0.5为93.3%,参数量为6.2 M
目标检测	Ma(2022) <sup>[29]</sup>	期刊文章	舰艇跑道	YOLOv5	mAP达到99.5%,mIoU达到95.4%且推理速度为64.5 FPS
目标检测	Li(2020) <sup>[50]</sup>	期刊文章	舰艇静态标识	SSD	成功率为91.1%,平均处理速度为9 ms
语义分割	刘健(2020) <sup>[51]</sup>	期刊文章	舰艇着舰标注	ERFNet	mIoU为76.35%,计算时间达到22.37 ms
姿态估计	Qiu(2024) <sup>[52]</sup>	会议论文	船舶静态标识	YOLOv7-tiny	mAP为94.0%,推理速度为42.92 FPS
姿态估计	Zhou(2022) <sup>[53]</sup>	会议论文	船舶静态标识	YOLO	YOLO的平均准确率提高,着陆成功平均时间约为0.16 s

结合视觉引导技术,无人机通过识别舰艇上的着陆标识,实现精确着陆.实验验证了该方法在模拟海上环境中的有效性,提高了无人机在舰艇上自主着陆的成功率<sup>[54,55]</sup>.

### 2.2.3 其他动平台

除了车载和舰艇动平台外,还有一些特殊的动平台,如移动的机器人平台、飞行中的其他飞行器等.对于这些动平台着陆场景,同样需要利用深度学习技术解决无人机与动平台之间的相对运动跟踪和精确着陆

问题,如表7所示.

以无人机在移动机器人平台上的自主着陆为例,任毅利用深度学习算法对机器人平台的运动模式进行分类和识别.通过对机器人不同运动状态下的视觉特征和运动数据进行学习,模型能够准确判断机器人的运动方向和速度变化,为无人机提供准确的着陆依据.在实际应用中,该方法能够使无人机在不同运动状态的机器人平台上实现稳定着陆<sup>[56]</sup>.

表7 其他动平台场景

研究方向	文献	文献类型	着陆场景	模型	实验结果
目标检测	任毅(2019) <sup>[56]</sup>	学位论文	“H”标识移动平台	Tiny-YOLOv3	移动平台0.75 m/s,平均着陆误差0.25 m时达到95%成功率
目标检测	Wang(2022) <sup>[57]</sup>	期刊文章	空中无人机动平台	YOLOv3	白天,晚上分别达到8.83 FPS和8.24 FPS,任务无人机在平台无人机上着陆点分布平均达到15 cm

### 2.2.4 小结

车载陆地、舰艇海上及其他动平台的着陆场景差异显著,算法选择需适配各自特性.陆地动平台速度 $\leq 40$  km/h、路面颠簸小,优先轻量化目标检测算法,兼顾实时性与抗干扰性;海上动平台受海风海浪影响,存在

不规则运动及反光、雾天干扰,需结合目标检测与姿态估计算法,补偿摇摆误差;其他动平台运动模式多样,着陆区域与姿态多变,倾向兼具跟踪精度与实时性的算法,适配不同场景需求.动态平台着陆的关键难点是适配多场景差异,需通过增广自建数据集、剪枝蒸馏量

化等方法,实现算法鲁棒化与轻量化,减少着陆时间与干扰,解决动态环境下的技术难题。

### 2.3 特殊场景着陆

特殊场景着陆涵盖了一系列复杂且具有挑战性的环境,如山区、森林、城市峡谷等。这些场景的特点是环境复杂多变,存在大量的障碍物、不规则地形以及信号干扰等问题,对无人机的自主着陆技术提出了极高的要求<sup>[58]</sup>。在山区,地形起伏大,可能存在强风、气流紊乱等情况,无人机需要具备精确的地形感知和避障能力。利用深度学习算法对激光雷达点云数据或视觉图像进行处理,构建三维地形模型,实时检测障碍物,并规划安全的着陆路径,使无人机能够在复杂的气流环境中自主学习和调整飞行策略,确保安全着陆。在森林环境中,树木密集,GPS信号容易受到遮挡而失效,无人机主要依赖视觉和其他传感器进行导航。通过深度学习模型对森林场景图像进行分析,识别出相对开阔、平坦的潜在着陆区域,并利用视觉里程计等技术实现精确的定位和导航。城市峡谷环境则面临着高楼林立、电磁干扰严重等问题。无人机需要通过深度学习算法对城市环境中的建筑物轮廓、街道特征等进行识别和分析,结合地图匹配技术实现定位,同时采用抗干扰通信技术和传感器融合技术,提高系统的可靠性和稳定性,确保在复杂的城市环境中成功着陆。

本节详细探讨了无人机自主着陆的各类场景。静态平台着陆相对基础,包括基于降落标识、跑道检测和地基引导3种方式,主要依靠深度学习精确识别着陆平台特征与位置;动态平台着陆难度较大,常见的车载陆地动平台和舰艇海上动平台,需结合目标检测与跟踪算法实时调整无人机飞行轨迹和姿态;特殊场景着陆涵盖山区、森林、城市峡谷等复杂环境,对无人机的环境感知和避障能力提出了更高要求。通过对不同场景的分析,明确了各场景的特点和挑战,为后续研究奠定了基础。

## 3 无人机自主着陆场景感知关键技术

### 3.1 无人机自主着陆场景感知方法的发展历程

#### 3.1.1 传统的无人机自主着陆

传统无人机自主着陆依赖多技术协同,但局限性显著。导航上,GPS借卫星信号全球定位,为着陆路径规划提供坐标,常规定位精度数米;INS以惯性测量单元为核心,依靠加速度计、陀螺仪数据,在GPS信号不佳时短时间维持精准监测。传感器应用中,气压计基于大气压力测高,因环境干扰难稳定实现分米级高精度测量;超声波传感器在着陆末段数米,可实现厘米级测距;磁罗盘结合地磁场与姿态角,保障飞行方向准确。然而传统技术短板突出:GPS易被遮挡、电磁干

扰,影响路径规划;复杂环境里,传感器易受冲击,如超声波遇障碍物、磁罗盘遇强磁会失效;气压高度计遇反光区域,局部气压场异常会导致其测高逻辑崩坏,无法稳定输出数据,使复杂场景下自主着陆的适应性、可靠性受限<sup>[59,60]</sup>。

#### 3.1.2 基于视觉与深度学习的无人机自主着陆

随着计算机视觉与深度学习发展,无人机自主着陆技术迎来变革,基于视觉的目标检测作为关键环节,在CNN驱动下取得显著进步。早期传统方法依赖SIFT、SURF(Speeded-Up Robust Features)等手工特征,效率与精度受限。深度学习时代,RCNN开创性将CNN引入目标检测,通过选择性搜索生成候选区域并分类,虽提升精度但计算量庞大、速度缓慢。Fast R-CNN通过ROI池化层实现整图一次特征提取,联合训练多任务,大幅提升检测速度与精度。Faster R-CNN进一步引入区域提议网络(Region Proposal Network, RPN),与检测网络共享卷积层,端到端生成高质量候选区域,在无人机着陆中能快速准确识别跑道、平台等目标<sup>[61]</sup>。YOLO系列另辟蹊径,以单阶段回归思路直接预测目标类别与位置,无需候选区域生成,检测速度飙升,满足无人机实时性需求。从YOLOv1到v12,模型持续优化网络结构,增强小目标与复杂场景检测能力<sup>[62-65]</sup>。在无人机自主着陆中,利用大量标注的跑道、平台图像数据训练RCNN与YOLO系列模型,使其能从视觉传感器实时图像中快速检测目标。例如,Faster R-CNN可精准识别野外简易平台,YOLO能快速定位跑道。相较传统特征匹配方法,深度学习方法凭借强大特征学习能力,显著提升检测精度与速度,适应多样化复杂环境,有力推动无人机自主着陆技术发展。

语义分割技术的发展为无人机环境理解带来关键突破。早期传统方法如阈值分割、边缘检测和区域生长,受限于简单灰度或边缘特征,在复杂场景中分割效果差。阈值分割仅适用于对比度明显的图像,边缘检测易受噪声干扰,区域生长对种子点选择敏感,均难以满足无人机复杂着陆场景的需求。深度学习推动语义分割进入全卷积网络(Fully Convolutional Networks, FCN)时代。FCN通过将CNN全连接层替换为卷积层,支持任意尺寸图像输入并输出像素级分类结果,能提取高级语义特征并通过上采样恢复分辨率,使无人机在城市低空等场景中初步区分着陆区域与障碍物。U-Net进一步优化,其对称U型结构通过跳跃连接融合高低层特征,显著提升小目标和精细结构的分割能力,例如在林地环境中精准识别小块空地,为复杂地形下的安全着陆点检测提供支持。近年来,基于Transformer的模型引入自注意力机制,突破传统卷积的局部感受野限制,高效捕捉长距离语义关联,在城市建筑群等大规模场景

中,能准确区分屋顶平台、道路等区域,为无人机提供更全面的着陆依据。

早期无人机姿态估计依赖 IMU (Inertial Measurement Unit),通过加速度计和陀螺仪积分运算获取姿态角,虽响应速度快但存在误差累积和环境敏感问题,如气流震动会导致测量偏差。基于视觉的方法弥补了这一缺陷:单目摄像头凭借成本低、部署便捷的优势广泛应用,但在无人机姿态估计中易受距离、光照变化影响,存在显著尺度不确定性;双目及多摄像头通过视差计算或多视角信息融合,有效提升了姿态估计精度,却同步增加了硬件部署成本、系统复杂度,且需精准校准各摄像头参数,适配标定难度较高的复杂飞行环境<sup>[66]</sup>。深度学习的引入带来革新,CNN 通过训练数据学习图像特征与姿态的映射关系,结合目标检测和语义分割结果,可精确识别着陆目标并重建三维环境。进一步融合 IMU 高频数据与视觉感知优势,显著提升复杂环境下姿态估计的稳定性,使无人机能实时调整飞行姿态,平稳接近着陆点。

在无人机自主着陆技术体系中,深度学习驱动的多模态技术协同发挥作用。光流预测通过分析图像序列像素运动,获取无人机相对环境的速度和方向信息,助力应对气流干扰等动态场景;三维重建利用多视角图像构建环境模型,提供地形起伏、障碍物分布等三维信息,支持提前规划安全着陆轨迹;语义分割则通过像素级分类标注,精准区分安全着陆区域与障碍物,尤其在城市复杂环境中为着陆提供关键依据。这些技术共同构建了从环境感知、姿态调整到路径规划的完整闭环,克服了传统方法对单一传感器的依赖和复杂场景适应性不足的问题,成为提升无人机自主着陆可靠性与精确性的核心驱动力。

## 3.2 无人机自主着陆感知算法

### 3.2.1 基于目标检测的无人机自主着陆

目标检测是无人机自主着陆的重要环节,其目的是在复杂的着陆场景中准确识别出着陆点、障碍物等关键目标<sup>[67]</sup>。随着深度学习技术的发展,其为目标检测提供了强大的工具,使得无人机能够更快速、更准确地感知周围环境。在基于深度学习的无人机自主着陆目标检测中,CNN 是最常用的模型之一,它通过卷积层、池化层和全连接层等组件,能够自动提取图像的特征,从而实现对目标的检测和分类。其中,Faster R-CNN、YOLO 系列、SSD 等算法在无人机目标检测中均得到了广泛应用。为了进一步提高目标检测的准确性和鲁棒性,研究人员提出了许多改进算法。一些算法引入注意力机制,使模型能够更加关注图像中的关键区域,从而提高目标检测的精度;一些算法通过多尺度特征融合,充分利用图像的不同尺度信息,增强对不同大小目标

的检测能力<sup>[68]</sup>;还有一些算法通过修改数据增强方式来增强算法的鲁棒性和泛用性<sup>[69]</sup>。

具体来看,YOLO 系列算法作为单阶段目标检测算法,将目标检测任务转化为一个回归问题,直接在图像上预测目标的类别和位置,其检测速度快、实时性强的优点,非常适合无人机自主着陆这种对实时性要求较高的应用场景,例如 YOLOv5 在无人机目标检测中表现出色,能够快速准确地检测出着陆区域和障碍物,为无人机的安全着陆提供保障<sup>[70,71]</sup>。Faster R-CNN 算法则是基于区域提议的目标检测算法,将目标检测任务分为区域提议生成和目标分类与回归两个阶段。在区域提议生成阶段,使用 RPN 在图像中生成一系列可能包含目标的候选区域;在目标分类与回归阶段,对这些候选区域进行分类和边界框回归,以确定目标的类别和位置,在无人机自主着陆场景中,该算法可以有效地检测出着陆点和障碍物<sup>[61]</sup>。此外,SSD 等目标检测算法在无人机自主着陆过程中,也为无人机准确识别着陆场和相关目标提供了重要支持<sup>[72]</sup>。

如表 8 所示,蔡炳锋<sup>[42]</sup>研究无人机自适应移动着陆,基于深度学习视觉识别与深度估计策略,但移动目标着陆功能处于初步阶段,未解决复杂场景无规律运动着陆问题。王忠言等人<sup>[47]</sup>研究低速平台无人机导航降落技术,提出总体方案但缺乏充足数据集支撑。Park 等人<sup>[73]</sup>开发了 GPS 拒止环境下的深度学习着陆方法,增强无人机自主性,却面临控制精度、环境适应性与成本性能的权衡挑战。Jiang 等人<sup>[74]</sup>利用 YOLO 处理热红外图像,验证昼夜检测有效性,但未探讨更广环境的适应性。Nepal 等人<sup>[70]</sup>对比 YOLOv3/v4/v5 在故障无人机中的应用,强调安全着陆对算法速度与精度的双重要求。Serrano 等人<sup>[75]</sup>对比 YOLOv5~v8 确定基准模型,但是缺乏推理速度实验数据。Rolland 等人<sup>[76]</sup>提出基于机器视觉的火山灰羽流采样技术,验证了模拟实验有效性。Idrovo 等人<sup>[48]</sup>评估了高精度 YOLO 模型在交通领域的应用潜力,认为其适用于自动驾驶集成,但未讨论模型在复杂环境下的泛化能力及交通场景动态适应性问题。Yuan 等人<sup>[77]</sup>提出了一种基于 Vision Transformers 的模型 VitP-RCNN,它通过粒子滤波方法大幅减少候选区域,提高跟踪速度和准确性,但依赖特定的训练数据集和参数调整。Li 等人<sup>[50]</sup>通过重新设计特征提取结构优化 SSD 模型,结合 KCF 检测进行权重校正以增强小目标和尺度变化场景的检测效果,然而数据集规模较小,算法泛化能力较弱。伍瀚等人<sup>[78]</sup>提出了一种目标跟踪模型 TCFNet,能够有效利用时序信息和目标轨迹先验对齐跨视角图像特征,提升跟踪性能,但空间维度上微小的误差可能导致目标关键信息被映射到非目标区域。

从批判性的综合视角来看,YOLO 系列算法的演变

表 8 无人机自主着陆目标检测算法

算法	文献	文献类型	着陆场景	自建数据集	创新点	实验结果	局限
YOLOv3-Tiny	蔡炳锋(2021) <sup>[42]</sup>	学位论文	动态小车	√	分析室内和室外两种情况,自建数据集	mAP达到了98.5%,推理速度达到53.0 FPS	未考虑无规则运动
YOLOv4	王忠言(2022) <sup>[47]</sup>	期刊文章	动态平台	√	通过目标检测与跟踪算法,验证了目标跟踪及自主降落	mAP为75%	数据集较小
YOLOv4-tiny	Park(2023) <sup>[73]</sup>	期刊文章	静态标识	√	通过使用YOLOv4-tiny算法实现了无人机自主着陆	算法平均精度达到90%	未介绍数据集
YOLOv5-s	Jiang(2022) <sup>[74]</sup>	期刊文章	静态平台	√	利用YOLO模型对无人机热红外图像和视频目标检测,通过实地测试	mAP为88.69%,最快的推理速度达到50 FPS	缺乏算法改进方案
YOLOv5l	Nepal(2022) <sup>[70]</sup>	期刊文章	静态标识	√	比较了YOLOv3/v4/v5三种算法在故障无人机自主着陆中的应用效果	精度达到70.7%,mAP达到63.3%,推理速度达到58.8 FPS	精度较低
YOLOv8	Serrano(2023) <sup>[75]</sup>	会议论文	静态目标		综合比较了YOLOv5~v8,确定了基准模型	mAP为89.1%,F1分数为90.8	缺乏推理速度
YOLOv8	Rolland(2024) <sup>[76]</sup>	会议论文	静态平台	√	构建了有质量的数据集,实现了先进的检测模型	算法平均精度值为90.7%	环境灵敏度高
YOLOv8x	Idrovo(2024) <sup>[48]</sup>	预印本	动态平台	√	分析了一系列YOLO模型,模型在车载检测方面表现出色	mAP0.5达到97.7%,mAP0.5-0.95达到78.2%,F1得分为0.95	缺乏泛化能力分析
VitP-RCNN	Yuan(2022) <sup>[77]</sup>	期刊文章	静态标识		基于MobileViT和RCNN构建模型,保证了精度和推理速度	平均精度为87%,参数量为5.7 M,推理速度达到51 FPS	未考虑复杂环境
SSD	Li(2020) <sup>[50]</sup>	期刊文章	动态舰艇	√	使用深度学习技术的目标检测方法,提出位置记忆系统	实现了89%的平均精度,94%的精度和99%的召回率,在Jetson TX2平台达到13FPS	推理速度较慢
TCFNet	伍瀚(2025) <sup>[78]</sup>	期刊文章	动态目标		引入OAN、TAN和CFFN模块协同跨视角特征融合	在MOTA和IDF1指标上分别提升了2.23和1.67个百分点	小目标跟踪有误差

始终围绕检测效率与精度的平衡,其中YOLOv1首次将目标检测转化为端到端的回归问题,极大提升了检测速度但小目标检测效果欠佳,YOLOv3通过多尺度预测头和特征金字塔结构改善了小目标检测能力,YOLOv5则通过轻量化网络设计在保持高精度的同时降低计算量,YOLOv8进一步升级架构实现精度和速度的双重提升.不过YOLO系列存在依赖大量标注数据、泛化性能受场景差异影响大,以及因回归本质导致重叠目标定位精度不足等局限,而Faster R-CNN虽定位准但计算复杂,SSD速度占优但小目标检测精度不及YOLO后续版本,当前引入注意力机制的改进算法存在计算开销大、多尺度特征融合可能导致信息损耗、数据增强或致分布偏移等问题,因此未来无人机自主着陆目标检测

算法的研究需在兼顾效率与精度的同时,强化对复杂真实场景的适应性和鲁棒性.

### 3.2.2 基于语义分割的无人机自主着陆

语义分割通过像素级分类助力无人机理解环境,为自主着陆提供关键信息.基于深度学习的方法采用编码器-解码器结构,实现像素级语义分类.FCN作为首个全卷积网络,将传统CNN全连接层替换为卷积层,直接输出像素级分类结果,可区分无人机着陆场景中的地面、建筑、树木等类别,辅助判断安全区域<sup>[79]</sup>.U-Net通过对称编码器-解码器结构及跳跃连接,融合不同层次特征信息,显著提升复杂场景下的分割精度,能更准确识别无人机拍摄图像中的着陆区域与障碍物,为安全着陆提供精细环境感知支持<sup>[80]</sup>.

如表9所示,刘健等人<sup>[51]</sup>引入非对称残差与弱瓶颈模块,优化下/上采样处理,改进后模型性能提升但推理速度仅22 ms,存在效率瓶颈. Yang等人<sup>[81]</sup>基于语义SLAM (Simultaneous Localization And Mapping) 实现室内无人机导航着陆,通过3D地图构建与语义场景理解提升定位精度. Claudet等人<sup>[82]</sup>采用HPSC替代微控制器进行语义分割实时性能预测,但是算法仅在特定高度训练,且分辨率提升会导致推理时间二次增长. Pu-

tranto等人<sup>[83]</sup>基于U-Net实现无标记着陆区域识别,简化结构并提高可靠性,通过离线处理验证可行性. Loera-Ponce等人<sup>[84]</sup>结合语义分割与风险评估,提升无人机紧急着陆的系统鲁棒性及城市环境适应性. Benjwal等人<sup>[85]</sup>基于UNet++架构优化分割算法,在准确性、IoU等指标上表现优异,模型轻量化优势显著. Ryu等人<sup>[86]</sup>使用DDRNet提高了在各种环境下的跑道识别性能,并在真实飞行测试中验证了该方法的有效性.

表9 无人机自主着陆语义分割算法

算法	文献	文献类型	自建数据集	改进	实验结果	局限
ERFNet	刘健 (2020) <sup>[51]</sup>	期刊文章	√	引入了非对称残差模块和弱瓶颈模块,将特征图尺寸缩小为最初的1/8,使用Adam随机梯度下降法	mIoU为76.35%,计算时间达到22.37 ms	分割推理速度较低
ORB-SLAM2	Yang (2024) <sup>[81]</sup>	期刊文章		提出了一种基于语义SLAM的室内环境下无人机导航和着陆方法	计算时间为25.5 ms	缺乏对比实验定量比较
ConvDeconv	Thomas (2022) <sup>[82]</sup>	期刊文章	√	对多种算法和架构在准确性、内存消耗和推理时间等多个度量空间进行了测试	准确率为95%,平均交并比(mIoU)为89%	提高分辨率将付出高昂代价
U-Net	Putranto (2022) <sup>[83]</sup>	会议论文	√	选择“着陆”像素类别作为着陆区域,并将这些区域转换为二值图像	mIoU达到47.3%	构建的数据集较小,分割准确率偏低
SegFormer	Loera-Ponce (2025) <sup>[84]</sup>	会议论文	√	通过SegFormer架构提供重要的上下文信息,帮助无人机组了解周围地面的风险	Acc为88%,mIoU为41%,F1得分为0.49	未研究泛化能力和鲁棒性且分割精度较低
UNet++	Benjwal (2023) <sup>[85]</sup>	会议论文		架构在准确性和IoU上取得了最高的成绩,同时在F1和AUC得分上也有不错的表现,且模型大小适中	达到91.75%的准确率,准确性显著提高了6.3个百分点	视频中可能出现运动模糊或遮挡等遮挡因素
DDRNet	Ryu(2025) <sup>[86]</sup>	期刊论文	√	结合高低分辨率网络,使用深聚合金字塔池化模块(DAPPM),简化尺寸	实现97.9,mIoU以及0.21 FPS, MAE从3.143下降到1.779	对复杂环境不敏感,鲁棒性不足

FCN以全卷积结构实现像素级预测,突破图像分类局限,但池化操作导致空间信息丢失,分割细节与边界定位较差,易造成无人机着陆场景中目标边界误判. U-Net借助对称结构与跳跃连接,融合多尺度特征提升精度,却存在信息冗余、计算复杂的问题,难以满足无人机实时性需求. 二者均依赖大量精细标注数据,在复杂光照、天气等未覆盖场景下,泛化能力受限.

### 3.2.3 基于姿态估计的无人机自主着陆

姿态估计是确定无人机在空间中的位置和姿态的过程,它是无人机自主着陆的关键技术之一. 准确的姿态估计能够为无人机提供精确的位置和方向信息,使其能够按照预定的轨迹安全着陆. 基于深度学习的姿态估计方法主要利用无人机搭载的摄像头获取的图像信息,通过对图像特征的分析和处理,来估计无人机的姿态. 其中,基于单目视觉的姿态估计方法由于其成本低、结构简单等优点,在无人机自主着陆中得到了广泛

研究. 基于深度学习的单目视觉姿态估计方法通常采用卷积神经网络来提取图像特征,并结合几何约束和优化算法来求解无人机的姿态. 一些方法通过训练深度神经网络,直接从图像中回归出无人机的姿态参数;另一些方法则采用两步法,首先利用神经网络检测图像中的特征点,然后根据这些特征点的几何关系来计算无人机的姿态<sup>[87]</sup>.

如表10所示, Baldini等人<sup>[88]</sup>提出端到端学习系统,实现原始图像与惯性数据到6-DOF全局位姿的映射,为无人机自主导航着陆提供新路径. Tang等人<sup>[39]</sup>设计基于地面视觉的N摄像头联合位姿估计方法,提升固定翼无人机着陆精度与鲁棒性,但尚未验证实际计算资源消耗情况. Zhou等人<sup>[53]</sup>针对单一目标检测简化YOLO网络结构,减少卷积层以降低计算量,虽准确性优于PNP方法,但实时性能存在差距.

直接回归姿态参数的方法凭借端到端训练实现高效输出,契合无人机实时需求,但对训练数据依赖强,

表 10 无人机自主着陆姿态估计算法

算法	文献	文献类型	着陆场景	自建数据集	改进	实验结果	局限
ViNet	Baldini (2020) <sup>[88]</sup>	会议论文	静态平台		接收图像输入流的卷积神经网络、一个集成连续图像之间 IMU 测量的小型 LSTM,以及一个核心 LSTM	姿态估计的准确性明显提高了 25%	模型泛化能力较差
Block-CNN	Tang (2023) <sup>[39]</sup>	期刊文章	地基引导	√	设计了 Block-CNN 架构,构建了自己的数据集	实现了 3 个百分点的锚点检测精度提升,位置和姿态估计精度分别提高了 19.5 和 12.7 个百分点	缺乏计算资源的数据测试
YOLO	Zhou (2022) <sup>[53]</sup>	会议论文	动态船舶	√	通过减少网络层数和增加全连接层来输出 6 维相对位置和姿态	YOLO 的平均准确率提高,着陆成功平均时间约为 0.16 s	推理速度较慢

面对环境变化时泛化性差. 两步法借助特征点与几何约束提升稳定性,却易受特征点检测精度影响,在复杂场景下易误判,且计算耗时久. 两种方法都受限于单目视觉的尺度歧义问题,难以精确定无人机与着陆点距离,影响姿态估计的准确性与着陆安全性.

### 3.2.4 基于其他深度学习方法的无人机自主着陆

深度学习方法在无人机自主着陆领域发挥着关键作用,除了常见的深度学习算法,光流预测、三维重建、

图像分割等技术也为无人机自主着陆提供了重要支持<sup>[89,90]</sup>.

如表 11 所示, Wang 等人<sup>[25]</sup>引入角点插值优化的 LK 稠密光流算法与 FlowNet2 神经网络,提升光流矢量计算的速度与准确性,为无人机动态环境感知提供支持. Dutrannois 等人<sup>[91]</sup>在真实系统中集成 ORB-SLAM3,融合点云地图、相机姿态与多传感器数据,增强运动规划框架的环境建模能力.

表 11 其他无人机自主着陆深度学习算法

算法	文献	文献类型	数据集类型	研究类型	改进	实验结果	局限
FlowNet2	Wang (2020) <sup>[25]</sup>	会议论文	自建	光流预测	采用堆叠的网络结构,将多个网络子模块进行组合,并提出了一个子模块 FlowNet2-SD	平均误差 21.4%,CPU 耗时 462 ms	由于特征点的个数较少且没有迭代优化的过程,所以解算的准确性不高
ORB-SLAM3	Dutrannois (2022) <sup>[91]</sup>	会议论文	公开	三维重建	详细讨论了 SLAM 算法的选择,包括现有的 3D 视觉 SLAM 算法和这些算法的基准测试,实现 ORB-SLAM3	在 ATERMSE 中 720 p 和 376 p 分辨率的精度为 5 cm 和 11.61 cm,同时保持 8 FPS 和 19 FPS 的计算速度	单目相机设置的精度性能较低,不足以满足应用需求,尽管其速度可达 19 FPS

光流预测基于像素运动连续性捕捉环境动态,为着陆提供运动参数,但易受光照与纹理影响,计算复杂,实时性欠佳;三维重建依靠多视角图像构建场景模型,直观呈现地形,不过数据处理慢,纹理匮乏时精度低;图像分割通过语义特征划分场景类别辅助路径规划,却在复杂环境下泛化不足,计算资源消耗大. 三者均存在环境适应难、效率与精度难平衡的问题,需进一步改进.

## 3.3 无人机自主着陆模型改进

### 3.3.1 无人机自主着陆特征提取方向模型改进

在无人机自主着陆模型中,特征提取的准确性与高效性对整体性能起着关键作用. 为优化这一环节,研究人员对模型结构与网络设计进行了多方面改进. 部分学者对经典的 CNN 结构进行创新,例如在传统的网络层级中引入了深度可分离卷积. 这种卷积方式将常规卷积操作拆分为深度卷积和逐点卷积,在大幅减少计算量的同时,能够保留对关键特征的提取能力. 以

MobileNet 系列为代表,其通过这种方式构建轻量级网络,在无人机计算资源受限的情况下,仍可快速提取图像中的着陆点、障碍物等目标的特征,显著提升了模型推理速度,为实时性要求极高的自主着陆任务提供了有力支持. 此外,多尺度特征融合策略也被广泛应用于改进特征提取过程. 传统的 CNN 在不同层对图像特征的提取侧重不同,浅层网络关注细节信息,深层网络则更擅长捕捉全局语义. 通过设计如 FPN (Feature Pyramid Network) 及其变体等结构,将不同尺度、不同层级的特征进行融合,使模型能够同时利用图像的多尺度信息. 在无人机自主着陆场景中,这一策略尤为重要,无论是远处的小型着陆标识,还是近处的大面积障碍物,模型都能通过多尺度特征融合更精准地识别与定位,从而有效提高了着陆精度和准确率<sup>[92]</sup>.

如表 12 所示,吴鹏飞等人<sup>[93]</sup>针对 SSD 网络小目标识别不足的问题,采用 ResNet101 替换 VGG-16 骨干网

络,并引入特征金字塔改进上采样结构,有效提升了着舰标志的识别精度与鲁棒性,但其研究仅停留在定性分析层面,缺乏量化数据支撑,且未在真实场景中验证算法实用性.

Shen 等人<sup>[49]</sup>开发的 AeroNet 系统基于深度学习实现鲁棒目标检测与定位,为无人机与地面车辆协作提供支持,但数据集场景单一,未涵盖复杂环境. Liu 等人<sup>[94]</sup>构建多传感器融合系统,整合 LiDAR、IMU 与视觉数据获取三维位置信息,结合点云语义与几何特征提升着陆区域识别的准确性和鲁棒性. Yi 等人<sup>[95]</sup>设计对

称菱形解码器,利用多阶段特征保留分割对象边界,通过消融实验证明其性能优于主流方法. Dhami 等人<sup>[96]</sup>开发语义分割制图系统识别安全着陆区域,但存在预测噪声. 乐辉<sup>[26]</sup>构建三种 CNN 分类模型,结合 RANSAC 算法提纯 SURF 特征,提高着陆标识特征提取效率与精度. 通过深度可分离卷积和多尺度特征融合等技术对特征提取进行改进,在保障计算效率的同时,显著提升了模型对各类目标特征的捕捉能力,为无人机精准识别着陆环境奠定了坚实基础,有效推动了自主着陆精度与准确率的提升.

表 12 特征提取方向模型改进

方法类别	算法	文献	文献类型	场景	自建数据集	改进	实验结果	局限
目标检测	SSD	吴鹏飞 (2022) <sup>[93]</sup>	期刊文字	静态平台	√	使用 ResNet101 代替 VGG-16 网络,并利用特征金字塔网络结构改进传统上采样结构	具有较强的鲁棒性,能在对实时性影响不大的前提下,有效提升对着舰标志的识别精度	仅给出定性分析,缺乏定量分析
	AeroNet	Shen (2023) <sup>[49]</sup>	期刊文章	车载标识	√	基于深度可分离卷积构建轻量级骨干,检测头采用了宽广的学习系统.	准确率为 89%,在 Intel NUC 上和 Nvidia TX2 上计算时间为 76 ms 和 28 ms	未考虑复杂环境情况
语义分割	深度分割网络	Liu (2022) <sup>[94]</sup>	期刊文章	静态标识		使用标准卷积、深度可分离卷积、残差瓶颈块、金字塔池化模块和特征融合模块	可以实现 68.88% 的 mIoU	精度较低,环境较简单
	CCTseg	Yi (2023) <sup>[95]</sup>	期刊文章	静态标识		包含三个子编码器的级联复合编码器提取多尺度特征,并将其输出到对称菱形解码器,该解码器逐步恢复多尺度特征输入的分辨率	SDD 测试集和 AERO 测试集上的 mAcc 分别为 68.6% 和 71.2%,mIoU 分别为 66.5% 和 70.1%	处理特定类型图像出现细节丢失
	深度分割网络	Dhami (2022) <sup>[96]</sup>	期刊文章	静态物体		使用了 Resnet50 前端的 DenseASPP 模型	Acc 达到 73%,精度达到 78%,F1 指数为 0.73,mIoU 达到 41.5%	模型预测中存在噪声
图像分割	CNN	乐辉 (2022) <sup>[26]</sup>	学位论文	静态标识	√	构建了三种不同结构的 CNN 分类模型,同时提出了使用 RANSAC 算法对 SURF 特征进行提纯	着陆标识图像的分类准确率达到了 98.18%	未详细说明细节和指标

### 3.3.2 无人机自主着陆语义理解方向模型改进

语义理解层面的改进旨在让无人机对复杂的着陆场景有更深入、更准确的认知,以做出更合理的着陆决策. 注意力机制的引入为这一目标提供了有效途径. 在基于深度学习的无人机自主着陆模型中,注意力机制能够引导模型将更多的计算资源聚焦于与着陆相关的关键区域和重要特征. 例如,在无人机拍摄的图像中,通过通道注意力模块,模型可以自动调整对不同通道特征的关注程度,突出显示如着陆区域的颜色、纹理等具有区分性的特征通道,抑制无关背景信息的干扰. 而空间注意力模块则能在空间维度上对图像进行加权,使模型更关注可能存在着陆点或障碍物的特定空

间位置,从而提升了对复杂场景语义理解的准确性,为精确着陆奠定基础. Transformer 架构的融入进一步革新了无人机自主着陆模型的语义理解能力. Transformer 最初在自然语言处理领域取得巨大成功,其基于自注意力机制的设计,能够有效捕捉序列中各元素之间的长距离依赖关系. 在无人机着陆场景中,将图像视为一个序列化的像素集合,Transformer 可以对场景中的各种物体及其相互关系进行全局建模. 与传统的 CNN 相比,Transformer 在处理具有复杂语义结构的场景时表现出更强的优势,例如在面对包含多个不同类型障碍物和不规则着陆区域的复杂环境时,能够更好地理解场景中各元素的空间布局和语义关联,进而帮

助无人机更准确地判断安全着陆路径,提高着陆精度和成功率<sup>[97]</sup>.

如表 13 所示, Xu 等人<sup>[98]</sup>提出基于自注意力机制的 YOLOX 模型,提升无人机复杂场景定位精度,却存在悬停时间过长的缺陷,可能导致低电量安全风险. Chen 等人<sup>[99]</sup>融合 YOLOX 与 Swin-Transformer 构建 Swin-YOLOX,提升特征提取效率,但其高海拔检测性能弱于 A-YOLOX,细粒度任务效果有待优化. Zhang 等人<sup>[100]</sup>提出了一种综合闭环解决方案,通过策略、算法和硬

件组件的协同设计,增强了复杂海洋环境中船载无人机自主着陆的能力,实现了技术上的鲁棒性和操作上的可行性. Li 等人<sup>[31]</sup>提出 YOLO-RWY 跑道检测框架,通过数据增强、可分离核注意力机制与双向特征金字塔网络,实现高效多尺度特征融合与快速收敛,提升检测精度. Dai 等人<sup>[101]</sup>设计轻量级注意力模块改进 MobileNetV3,集成 PANet 和 SPPnet 聚合特征,消除 NMS 需求以减少延迟,但在多样化跑道场景中难以实现高精度检测.

表 13 语义理解方向模型改进

方法类别	算法	文献	文献类型	场景	自建数据集	改进	实验结果	局限
目标检测	A-YOLOX	Xu (2022) <sup>[98]</sup>	期刊	静态标识	√	在 YOLOX 中加入了注意力机制,设计了全新的无人机算法	AP0.5 为 95.5%,检测速度 53.7 FPS,着陆误差在 5 cm 以内	稳定性较差
	Swin-YOLOX	Chen (2024) <sup>[99]</sup>	期刊	静态标识		提出基于 SwinTransformer 的 YOLOX 无人机着陆视觉算法	AP0.5 可以达到 98.7%,检测速度为 38.4 FPS	推理速度较慢
	YOLOv8	Zhang (2025) <sup>[100]</sup>	期刊	静态标识	√	引入 C2f-RVB 模块,集成 ABF 特征融合模块,结合 KCF 算法优化目标跟踪	mAP0.5 为 93.5%,测置信度提高了 10~15 个百分点	极端海况尚未验证
	YOLO-RWY	Li (2024) <sup>[31]</sup>	期刊	静态跑道		整合 EDA 模块,集成大型可分离核注意力 LSKA 模块,提供轻量级注意力机制将 neck 重组 BiFPN 模块	AP0.5-0.95 为 76%,参数量为 3.106 M, FLOPs 为 8.308 G,推理速度为 125.3 FPS	模型精度较低
	YOMO-Runwaynet	Dai (2024) <sup>[101]</sup>	期刊	静态跑道		在 MnasNet 中引入轻量级注意力,改进的 MobileNetV3 为骨干网络,结合 PANet 和 SPPnet	检测准确率 89.5%,关键点检测误差率 0.003,推理速度 90.9 FPS	缺乏复杂场景实验
语义分割	STDC-CT	Jiang (2023) <sup>[102]</sup>	期刊文章	静态标识	√	细节引导、小目标注意力提取器和多尺度上下文信息三个分支网络	在 UAV-City 数据集以 196.8 FPS 速度实现 67.3% 的 mIoU	小目标丢失细节
	Deeplabv3+	Wu (2024) <sup>[24]</sup>	会议论文	静态标识	√	增加了双分支融合模块,结合了 ShuffleNetv2 作为骨干网络,并引入了 CBAM 注意力机制	精度达到 93.8%	未验证泛化能力
	U-Net	Zhang (2025) <sup>[103]</sup>	会议论文	静态标识		增加了特征金字塔模块,并融合编码器和解码器	Acc 达到 96%,精度 80%,召回率达到 81%,F1 指数为 0.82	受限有限的数据集
姿态估计	YOLOv7-tiny	Qiu (2024) <sup>[52]</sup>	会议论文	动态船舶	√	用 CARAFE 算子代替传统的上采样部分,增加 CA 注意力机制	mAP 为 94.0%,比 YOLOv7-tiny 提高了 9.64 个百分点检测速度为 42.92 FPS	检测角度少

Wu 等人<sup>[24]</sup>改进 Deeplabv3+模型,引入 ShuffleNetv2 与 CBAM 注意力机制,提升语义分割精度. Zhang 等人<sup>[103]</sup>提出一种基于改进 U-Net 架构的无人机视觉自主降落区域检测方法,能从不同下降高度评估降落区域,实验结果表现良好,但数据集有限,可能影响算法泛化能力. Qiu 等人<sup>[52]</sup>采用 CARAFE 算子替换传统上采样模块,结合 CA 注意力机制增强复杂海洋环境特征提取

能力,然而该模型仅能检测无人机两个姿态角度,缺乏对俯仰、滚转、偏航三角度的全面识别,且需构建大规模多样化数据集以提升泛化性能. 注意力机制与 Transformer 架构的应用,使无人机模型能够从多维度深入理解着陆场景语义,无论是关键区域聚焦还是全局关系建模,都显著增强了模型对复杂场景的解析能力,为无人机安全、精准着陆提供了可靠的语义理解保障.

### 3.4 无人机自主着陆场景适配

#### 3.4.1 无人机自主着陆场景的模型鲁棒化

为使无人机自主着陆模型在复杂多变的实际场景中稳定运行,模型鲁棒化的场景适配技术至关重要.数据增广是一种常用且有效的手段.通过对原始训练数据进行多样化的变换,如旋转、缩放、裁剪、添加噪声等操作,可以生成大量新的训练样本,丰富数据的多样性.在无人机自主着陆场景中,不同的光照条件、拍摄角度以及天气状况等因素会导致采集到的图像数据存在较大差异.数据增广能够模拟这些实际场景中的变化,使模型在训练过程中接触到更多不同形态的数据,从而增强模型对各种环境变化的适应能力,减少因环境干扰导致的检测错误和着陆失误,提高模型的鲁棒性.自建数据集也是提升模型场景适配能力的关键技术.针对无人机实际飞行的特定场景,如城市建筑密集

区域、山区复杂地形或海上作业环境等,收集并标注专门的图像和传感器数据,构建具有针对性的数据集.这些数据能够更真实地反映无人机在实际应用中可能遇到的场景特征和挑战,相比使用通用数据集,基于自建数据集训练的模型能够更好地学习到特定场景下的着陆相关信息,对场景中的目标和干扰因素有更准确的识别与判断,有效降低外界环境干扰对无人机自主着陆的影响,提升模型在实际场景中的可靠性和适应性<sup>[104]</sup>.

如表 14 所示,Truong 等人<sup>[105]</sup>结合 YOLOv2 与去模糊算法改善运动模糊图像检测,但结果精度较低.Naufal 等人<sup>[106]</sup>基于 YOLOv8 实现无人机测绘中的自动化 GCPs 检测,展现训练速度与模型扩展性优势.Lee 等人<sup>[107]</sup>提出深度学习优化视觉受损环境下的目标检测与标记识别,增强了多旋翼无人机在复杂场景中的着陆能力.

表 14 基于模型鲁棒化的场景适配技术

方法类别	算法	文献	文献类型	场景	自建数据集	改进	实验结果	局限
目标检测	YOLOv2	Truong (2020) <sup>[105]</sup>	期刊	静态平台	√	将 YOLOv2 与 DeepDeblur、DeblurGAN 结合	平均精度为 64.5%,平均召回率为 64.1%,F1 指数为 0.643	实验精度较低
	YOLOv8s	Naufal (2023) <sup>[106]</sup>	会议	静态标识	√	对比了使用 YOLOv8 模型与早期 YOLO 版本的改进,构建了数据集	在验证集上实现了 99.7% 的高精度和 99.3% 的召回率,mAP50-95 为 88.7%	数据集较简单
	RetinaNet	Lee (2022) <sup>[107]</sup>	期刊	静态平台	√	使用深度学习技术的目标检测方法,提出位置记忆系统	实现 89% 的平均精度,94% 的精度和 99% 的召回率,在 Jetson TX2 平台达到 13 FPS	推理速度较慢
语义分割	深度分割网络	Lim (2024) <sup>[108]</sup>	期刊文章	静态标识	√	通过集成的硬件和软件系统,实现了在不同条件下的自主感知、决策和着陆	计算时间为 250 ms 时,语义图像分割的通用评估指标 mIoU 为 67.0%	推理速度较慢
	KDP-Net	Zhang (2024) <sup>[109]</sup>	期刊文章	静态标识	√	KDP 模块采用了核共享机制,核共享机制与预处理阶段的数据增强相结合	在 UDD6 和 SDD 达到 85.25 FPS 和 108.11 FPS,mIoU 达到 76.9% 和 67.14%,在 Jetson Orin 上速度达到 53.7 FPS 和 38.79 FPS	构建的数据集较小

Lim 等人<sup>[108]</sup>开发复杂环境自主着陆系统,展示多地面类型的适用性,但未涉及极端气象条件下的性能验证.Zhang 等人<sup>[109]</sup>提出 KDP-net 网络,利用扩张核捕获层级交互信息,缓解多分辨率图像的特征丢失问题,有效处理小目标与边缘细节缺失.数据增广与自建数据集技术从数据层面出发,通过拓展数据多样性和增强数据针对性,大幅提升了模型对复杂场景的适应能力,有效降低了环境因素对无人机自主着陆的干扰,为模型在实际应用中的稳定表现提供了有力支撑.

#### 3.4.2 无人机自主着陆场景的模型轻量化

在无人机资源受限的情况下,模型轻量化对于实现快速、稳定的自主着陆至关重要.模型量化是一种有

效的轻量化手段,其核心思想是将模型中原本高精度的浮点型参数和计算转换为低精度的整型表示.在无人机自主着陆模型中,大量的卷积和全连接层运算占据了主要的计算资源.通过量化技术,例如将 32 位浮点型参数转换为 8 位整型参数,虽然在一定程度上会牺牲少量精度,但能够显著减少模型的存储需求和计算量,加快模型推理速度.这使得无人机在有限的硬件资源下,仍能快速处理传感器数据,实时做出着陆决策,减少着陆时间延迟,同时降低因计算资源紧张导致的系统不稳定因素,提高了模型在实际场景中的实用性.模型蒸馏和剪枝技术也在无人机自主着陆模型轻量化中发挥着重要作用.模型蒸馏通过将复杂的大型教师

模型的知识迁移到小型学生模型中,使学生模型在保持较高性能的同时实现轻量化.在自主着陆场景中,教师模型可以是经过大量数据训练、性能优良但计算复杂的模型,而学生模型则是针对无人机硬件资源优化后的轻量级模型.通过蒸馏,学生模型能够学习到教师模型对场景特征的判别能力,从而在不损失过多精度的前提下减少模型参数量和计算复杂度.模型剪枝则是通过去除模型中冗余的连接、神经元或层,进一步简化模型结构.在无人机自主着陆模型训练过程中,分析各参数和连接对模型性能的贡献程度,将那些对性能影响较小的部分剪掉,从而降低模型复杂度,减少计算量和存储需求,提升模型在实际场景中的运行效率,减

少着陆过程中的干扰,保障无人机安全、快速地完成着陆任务<sup>[110]</sup>.

如表 15 所示,张琴<sup>[111]</sup>通过改进 YOLOv3 和 Kalman 滤波提高降落精度,但地标数据集仅单目标识别,场景单一.Yu 等人<sup>[112]</sup>利用 CNN 实现着陆目标实时鲁棒检测,通过仿真与真实实验验证有效性,但缺乏复杂环境性能分析.

模型量化、蒸馏和剪枝等轻量化技术,从参数优化、结构精简和知识迁移等多个角度出发,在有限硬件资源条件下,有效提升了模型运行效率,缩短了着陆决策时间,降低了系统运行干扰,为无人机自主着陆的高效、稳定实现提供了技术保障.

表 15 基于模型轻量化的场景适配技术

方法类别	算法	文献	文献类型	场景	自建数据集	改进	实验结果	局限
目标检测	PC-YOLOv3	张琴 (2021) <sup>[111]</sup>	学位	静态标识	√	通过通道剪枝,层剪枝及压缩剪枝策略调整模型	mAP 达到 91%,推理速度达到 43.1 FPS,大小为 110 MB	数据集种类较单一
	YOLO-Tiny	Yu (2020) <sup>[112]</sup>	期刊	静态标识	√	基于 YOLO 设计了算法,在更加轻量化的情况下损失较少的精度	IoU 精度为 83.7%,推理速度为 21 FPS	算法精度较低

### 3.5 关键算法的对比与总结

本节全面阐述了无人机自主着陆的关键技术.如表 16 所示,该表系统对比了基于视觉的无人机自主着陆深度学习算法,涵盖了各系列算法的原理、优势、局限性以及它们适用的场景.

在目标检测领域,YOLO 系列算法以单阶段端到端回归实现快速检测,适合动态场景,但小目标检测能力弱;RCNN 系列算法凭借两阶段检测获得高精度,适用于静态精密任务,却存在推理慢的问题;SSD 算法则平衡了速度与精度,适用于中等复杂度场景.语义分割方面,FCN 分割速度快,U-Net 分割精度高,ERFNet 在速度和远距离检测上有优势.姿态估计中,单目视觉姿态估计成本低但存在尺度问题,多目视觉姿态估计则深度测量精准但硬件要求高.此外,光流预测、三维重建等其他方法也各有优劣,分别适用于不同的无人机自主着陆应用场景.

## 4 数据集与评估

### 4.1 评估指标

为了全面评估基于深度学习的无人机自主着陆系统的性能,需要选择合适的评估指标<sup>[113,114]</sup>.

#### 4.1.1 定位误差

定位误差是无人机自主着陆系统定位准确性的核心指标,反映实际位置与目标定位点的偏差,下面将对常见指标进行介绍.

欧氏距离(Euclidean Distance, ED).计算无人机实际着陆点与目标着陆点在二维平面或三维空间中的直

线距离.

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (1)$$

其中,  $(x_i, y_i, z_i)$  表示第  $i$  个点在三维空间中的坐标.

均方根误差(Root Mean Square Error, RMSE)和平均绝对误差(Mean Absolute Error, MAE).RMSE 对多次着陆偏差平方和平均后取平方根,综合多次试验结果,评估精度稳定性.MAE 计算多次定位偏差的绝对值平均值.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (d_i)^2} \quad (2)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |d_i| \quad (3)$$

其中,  $n$  为试验次数;  $d_i$  为第  $i$  次试验的着陆偏差.

坐标轴分解误差.将三维误差拆分为水平和垂直分量,分别影响水平对齐与高度控制精度,对动态平台或复杂地形着陆至关重要.

最大定位误差.记录多次试验的最大偏差,评估极端情况下的定位可靠性,适用于高安全性场景.

静态平台着陆.优先用欧氏距离和 RMSE,评估定点定位的准确性与稳定性.

动态平台着陆.结合水平/垂直分解误差与 MAE,重点关注实时定位能力.

#### 4.1.2 成功率与运行时间

成功率和运行时间是评估无人机自主着陆系统性能的关键指标.成功率定义为多次着陆试验中,未碰撞且着陆位置在规定误差范围内的试验比例,其直接反映系统可靠性与稳定性,是安全回收和任务完成的关

表 16 基于视觉的无人机自主着陆深度学习算法对比

方法类别	具体方法	原理	优势	局限	适用场景
目标检测	YOLO 系列算法	单阶段端到端回归,网格划分直接预测目标坐标与类别,多尺度特征融合提速	检测速度快,如 YOLOv8 达 125 FPS、模型轻量,适合实时响应	小目标检测弱、复杂环境泛化差	动态平台、实时避障
	RCNN 系列算法	两阶段检测 RPN 生成候选区域后分类回归,如 Faster R-CNN 结合特征金字塔	检测精度高、细节识别准	推理慢、计算资源消耗大	静态精密着陆、高精度障碍识别
	SSD 算法	单阶段多框检测,不同尺度卷积层直接预测目标,无需候选区域生成	速度与精度平衡、中尺度目标检测效果好	小目标精度低于 RCNN、泛化数据不足	城市峡谷等中等复杂度场景的着陆点检测
语义分割	FCN	全卷积网络,全连接层替换为卷积层,上采样实现像素级分类	分割速度快、支持任意尺寸输入,适合场景粗分类	边界模糊、小目标分割弱	开阔场地着陆区、粗筛选
	U-Net	对称 U 型结构,跳跃连接融合高低层特征,精准分割小目标与边界	小目标分割优、边界定位准	计算复杂、推理慢、硬件要求高	森林、山区等复杂地形安全区域识别
	ERFNet	ERFNet 用非对称残差模块提速	ERFNet 速度快对舰艇检测较准	ERFNet 精度低,极端天气适应性差	适用于动态平台跟踪
姿态估计	单目视觉姿态估计	基于 CNN 回归图像特征与姿态参数映射,如 YOLOv5 检测标识点后解算位姿	成本低、结构简单、模型轻量	尺度不确定、复杂光照姿态漂移	消费级无人机、简易着陆
	多目视觉姿态估计	双目/三目通过视差计算深度,结合 PnP 解算姿态	深度测量准、姿态稳定	标定复杂、硬件成本高、计算量大	军事着舰、工业精细作业
其他方法	光流预测(FlowNet2)	CNN 分析图像序列像素运动,估计无人机相对环境速度	实时运动感知、辅助轨迹规划	特征点稀疏、复杂纹理场景失效	动态环境轨迹调整
	图像分割(CNN 分类)	CNN 对着陆标识分类,如 LeNet 识别 H 形标识	标识识别快、模型极小、硬件适配性强	仅适用于简单标识、复杂背景鲁棒性差	简易标识着陆、初级分类

键. 运行时间核心衡量系统实时性,深度学习模型推理时间是重要组成部分,其单位为 s 或 ms,统计平均运行时间评估性能,时间越短实时性越强.

#### 4.1.3 检测准确率

在无人机自主着陆过程中,目标检测和语义分割等任务的准确性对系统性能有着重要影响. 检测准确率用于衡量深度学习模型对着陆点、障碍物等目标的识别和分类能力. 下面将对常见指标进行介绍. 常用的检测准确率指标包括精确率(Precision)、召回率(Recall)和 F1 值.

精确率(Precision). 表示被模型正确识别为目标的样本数占模型识别为目标的总数的比例.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

其中,TP 为真正例;FP 为假正例. 精确率反映了模型识别结果的准确性.

召回率(Recall). 表示被模型正确识别为目标的样本数占实际目标样本总数的比例.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

其中,FN 为假反例. 召回率反映了模型对目标的覆盖

程度.

F1 值(F1 Score). 是精确率和召回率的调和平均值,综合考虑了精确率和召回率两个指标.

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (6)$$

F1 值越高,说明模型在目标检测任务中的性能越好.

## 4.2 数据集

数据集是基于视觉与深度学习的无人机自主着陆场景感知研究的重要基础,其质量和规模直接影响深度学习模型的性能和泛化能力. 在无人机自主着陆领域,常用的数据集主要包括以下几类.

### 4.2.1 公开数据集

公开数据集为研究人员提供了统一的测试和比较平台,有助于推动无人机自主着陆技术的发展. 目前,针对无人机自主着陆的公开数据集相对较少,但一些相关领域的公开数据集也可用于无人机自主着陆研究,表 17 介绍了一些常用的公开数据集.

VisDrone 是面向无人机视觉任务的大规模基准数据集,由天津大学 AISKEYE 团队构建,包含 288 个视

表 17 公开数据集

名称	来源	数据集类型	样本类型	样本数	数据集内容
VisDrone	Du(2019) <sup>[115]</sup>	目标检测	视频、图像	288 视频, 10 209 图像	基准数据集由 288 个视频片段组成, 由 261 908 帧和 10 209 幅静态图像组成, 包括位置、环境、对象和密度。
UAVid	Lyu(2020) <sup>[116]</sup>	语义分割	视频、图像	30 视频, 300 图像	由 30 个视频序列组成, 在倾斜视图中捕获 4K 高分辨率图像。总共对 300 张图像进行了 8 类的密集标记
EVD4UAV	Sun(2024) <sup>[117]</sup>	目标检测、语义分割	图像	3 000	数据集包括大约 3 000 张图像, 这些图像描绘了车辆可能部分或全部被雪覆盖的冬季场景
UAV-Human	Li(2021) <sup>[118]</sup>	姿态估计等	视频	67 428	67 428 个多模态视频序列和 119 个对象, 22 476 帧姿态估计, 41 290 帧人物再识别, 22 263 帧属性识别
AID	Xia(2017) <sup>[119]</sup>	场景识别	图片	10 000	数据集由 30 个航拍场景类型组成, 有 10 000 张图像
UAV123	Mueller(2016) <sup>[120]</sup>	目标跟踪	视频	143	包含 123 个短视频和 20 个长视频, 总大小约为 13.5 GB

视频片段、26 万余帧及万余幅静态图像, 覆盖中国 14 个城市的城乡场景, 涵盖多无人机平台、天气及光照条件, 人工标注超 260 万目标边界框并附带场景可见性、遮挡等属性。该数据集虽为无人机目标检测提供了丰富数据, 但场景覆盖仍存在局限性, 仅聚焦中国城市, 难以全面反映全球不同地理环境、文化背景下的场景特征。标注一致性方面, 大规模人工标注可能因标注人员理解差异, 导致边界框精度与属性定义存在偏差。在使用限制上, 由于不完全针对自主着陆场景, 在训练着陆场景关键目标识别模型时, 需结合其他数据集弥补特定任务的不足。

UAVid 作为高分辨率语义分割数据集, 包含 30 个 4K 视频序列, 300 张 8 类密集标注图像, 聚焦大尺度变化、运动目标识别与时间一致性挑战, 支持无人机着陆场景的语义分割任务, 助力识别着陆区域与障碍物。然而, 其标注可能存在误差, 这不仅影响模型训练的准确性, 不同标注人员对语义类别边界的划分也可能存在分歧, 导致标注一致性难以保证。在使用限制上, 由于数据集规模有限, 在复杂多变的实际着陆场景中, 模型泛化能力可能不足, 需通过数据增强或与其他数据集融合进行预处理验证。

VD4UAV 是针对无人机图像车辆检测的对抗性基准数据集, 包含约 3 000 张不同高度拍摄的冬季场景图像, 提供细粒度注释, 用于评估对抗条件下物体探测器性能, 为复杂环境检测算法提供独特测试场景。但该数据集场景覆盖单一, 仅涵盖冬季场景, 无法模拟其他季节的环境变化对检测算法的影响。标注一致性方面, 对抗条件下的注释可能因对干扰因素的定义和标注标准不统一, 导致数据质量参差不齐。使用限制上, 其仅适用于车辆检测, 难以满足无人机着陆场景中对多种目标检测的需求。

UAV-Human 是大规模多模态数据集, 包含 6.7 万视频序列及超 2.2 万帧姿态估计、再识别数据, 覆盖多城市昼夜场景, 支持无人机视角下的人类行为理解, 为着

陆场景中的动态目标感知提供数据支撑。但场景覆盖存在偏向性, 虽涵盖多城市, 却未充分考虑不同地域人群行为习惯差异。标注一致性方面, 多模态数据标注涉及多个维度, 容易出现模态间数据对齐偏差。使用限制上, 专注于人类行为理解, 在着陆场景的其他关键目标检测与场景理解任务上难以直接应用。

AID 是基于谷歌地球图像的大规模航空数据集, 包含 30 类共 1 万张多源遥感图像, 由专家标注, 覆盖全球不同地区及成像条件, 适用于场景分类算法评估, 其多源特性为无人机着陆环境的场景理解带来多样性挑战。然而, 航空图像与无人机实际着陆场景存在视角和尺度差异, 导致数据适用性受限。标注一致性方面, 专家标注虽有较高权威性, 但不同专家对场景类别的划分标准可能存在细微差别。使用限制上, 主要适用于场景分类, 难以满足着陆场景中对具体目标检测与定位的需求。

UAV123 由商业无人机拍摄, 包含 123 个短视频与 20 个长视频, 经标准化处理为 720 p/30 FPS, 注释以直立边界框形式提供, 背景干净、视角多样, 适用于目标跟踪算法开发, 为无人机着陆过程的动态目标定位提供标注数据支持。但背景干净的设定与实际复杂着陆环境脱节, 标注仅采用直立边界框, 无法准确描述倾斜或不规则目标形状。使用限制上, 数据规模和场景丰富度不足, 难以支撑复杂环境下的高精度目标跟踪任务。

#### 4.2.2 自建数据集

为了满足特定的研究需求, 许多研究人员会根据实际情况构建自己的数据集。自建数据集能够更准确地反映无人机自主着陆的实际场景, 提高模型的针对性和适应性。基于实际飞行采集的数据集, 研究人员通过实际操作无人机进行飞行试验, 在不同的着陆场景下采集图像和传感器数据。具体如图 6 所示。

在静态平台着陆场景中, 采集不同光照条件、天气状况下的着陆点图像; 在动态平台着陆场景中, 记录无人机与动态平台相对运动过程中的数据。通过这种方式获取的数据具有较高的真实性和可靠性, 但采集过



图6 UAV-RGBT数据集及其预处理流程

程较为复杂,需要耗费大量的时间和精力,且受到飞行条件和安全因素的限制,如表18所示。

基于仿真生成的数据集可以利用 AirSim、Gazebo 等仿真软件灵活设置地形地貌、气象条件等参数,快速生成大量多样化的无人机自主着陆虚拟数据,并获取高质量图像和传感器数据,但该类数据集在实际应用中存在明显局限。在场景覆盖上,难以复现真实世界的偶然性与复杂性,与真实场景存在“认知鸿沟”;

标注一致性方面,自动生成的标注无法模拟真实标注因目标模糊等产生的差异,易导致模型误判;使用限制上,仿真与真实场景在物理规律、渲染效果等方面的偏差,使得模型易过拟合,泛化能力差,尽管需将仿真与真实数据结合训练来提升泛化能力,但如何有效融合及平衡权重仍是难题。为了提高模型的泛化能力,通常需要将仿真数据与真实数据相结合进行训练。

表18 自建数据集

来源	数据集类型	数据集基准模型名称	样本类型	样本数	数据集内容
Idrovo(2024) <sup>[48]</sup>	目标检测	YOLOv8x	图片	1 862	包括训练集 1 734 张,验证集 81 张以及测试集 47 张
Xie(2023) <sup>[38]</sup>	目标检测	Faster RCNN	图片	3 828	建立了共计 3 828 张图像的无人机数据集.训练样本与测试样本的比率为 9:1
汪进中(2025) <sup>[121]</sup>	目标检测	UAV-DMDet	图片	5 117	跨季节、跨昼夜、多类别、多尺度的大规模 UAV 视角 RGB-T 多源图像数据集 UAV-RGBT,包含 20 个类别、5 117 对 RGB-T 图像和超 11 万个标注
Putranto (2022) <sup>[83]</sup>	语义分割	U-Net	蒙版	405	将图像和蒙版拆分为相同的大小,图像总数为 405 张图像和蒙版.此数据集有六个类
Loera-Ponce (2024) <sup>[84]</sup>	语义分割	SegFormer	图片	400	包含 400 张鸟瞰图 5 000 × 4 000 像素的高分辨率航拍图像,在离地面 5~30 m 的高度采集
Wu (2024) <sup>[24]</sup>	语义分割	Deeplabv3+	图片	600	数据集用于语义分割实验,它包括 600 张无人机捕获的低空航拍图像,并包括 20 个语义分割类别,该数据集适用于旋翼无人机的自主着陆
Qiu(2024) <sup>[52]</sup>	姿态估计	YOLOv7-tiny	图片	5 145	图像按 7:1:2 的比例分为训练 3 601 张图像、验证 515 张图像和测试集 1 029 张图像
Tang(2023) <sup>[39]</sup>	姿态估计	Block-CNN	仿真	12	包括 8 个模拟和 4 个真实着陆视频,并配有边界框、锚点和地面无人机姿态信息等标签
Zhou(2022) <sup>[53]</sup>	姿态估计	YOLO	仿真	5 400	用 solidworks 构建了船舶模型,并添加了相机传感器来收集图像数据作为输入
Wang(2020) <sup>[25]</sup>	光流预测	FlowNet2	图片	10 000	使用了 1 000 张原始运动场景的图片,每张原图基于修改 10 次不同的运动参数,生成 10 张虚拟第二帧图片和光流图,总共获得 10 000 套训练图片
乐辉(2022) <sup>[26]</sup>	图像分割	CNN	图片	7 050	共得到 7 050 张图像,正样本 4 750 张,负样本 2 300 张
Zhang(2021) <sup>[35]</sup>	高度估计	VIAE-Net	测量值	2 400	数据集包括 2 400 张图像和相应的 IMU 测量值.数据集分为训练集的 80% 和验证集 20%

### 4.3 竞赛

#### 4.3.1 世界大学生全自主无人机飞行挑战赛

世界大学生全自主无人机飞行挑战赛(Unmanned Aerial System)由英国机械工程师学会举办,发起于2014年,是具有全球影响力的大学生航空模型竞赛之一。挑战赛分为设计、开发和演示三个阶段,从前一年的10月份一直持续到竞赛当年的6、7月份,最终通过现场飞行演示和商业案例展示决出最终冠军。要求参赛队伍从头开始设计并制造一架最大起飞重量为10 kg的无人驾驶飞行器,自动完成区域搜索、精准投放、指定路线返航等特定任务。2024年的比赛在英国林肯郡举行,来自全球各地的38支高校队伍参赛。北航航模队首次参赛并荣获冠军,这也是该项赛事发起以来首次有来自中国的高校参赛。比赛要求无人机在 $10\text{ m} \times 10\text{ m}$ 的区域自主起飞,完成区域搜索和定点投弹任务后,在 $10\text{ m} \times 20\text{ m}$ 的指定区域内自主着陆。

#### 4.3.2 中国国际飞行器设计挑战赛

中国国际飞行器设计挑战赛由中国国家体育总局航空无线电模型运动管理中心等主办。设置垂直起降载运、限时载运空投、模拟搜救、太阳能飞机、对地侦察与打击等11个比赛项目,分赛需设至少4项,比赛设大学生组和高中生组。是目前我国唯一集飞行器设计、制造、飞行与控制于一体的赛事。

#### 4.3.3 VisDrone

VisDrone是一个专注于无人机视觉目标检测和追踪的开源项目,由清华大学计算机科学与技术系发起并举办相关的挑战赛,例如VisDrone2023挑战赛分为目标检测和零样本目标检测两个赛道,其关于自主着陆的技术要求有很大的竞争力与挑战性。

本节介绍了评估基于深度学习的无人机自主着陆系统性能的指标和相关数据集。着陆精度、成功率、检测准确率和运行时间等指标可全面评估系统性能;公开数据集如VisDrone、UAVid等为研究提供了统一测试平台,自建数据集则能更贴合实际研究需求。此外,还介绍了世界大学生全自主无人机飞行挑战赛、中国国际飞行器设计挑战赛等相关竞赛,这些竞赛对推动无人机自主着陆技术发展具有积极作用。

## 5 面临的挑战与限制

### 5.1 复杂环境适应性

尽管深度学习推动了无人机自主着陆技术发展,但在复杂环境中仍面临多重挑战。对抗性攻击鲁棒性方面,暴雨导致图像模糊、大雾致视觉传感器失效,传统模型难处理异常输入;小样本学习上,山区、森林等复杂地形样本稀缺,模型泛化能力不足;域泛化层面,城市电磁干扰导致数据传输异常,高楼遮挡反射GPS

信号,均影响模型精度;光照变化也会改变图像特征分布,降低模型鲁棒性。视觉感知还面临更多特殊干扰。反光与光晕场景中,水面、金属平台等光滑表面的强反光会掩盖着陆特征,低光照时段的光晕会模糊关键信息;雨水、灰尘污染时,镜头水膜、沙尘附着会导致图像雾化、遮挡,降低对比度;透明障碍物场景下,玻璃幕墙等透光物体难以与背景区分,还可能产生虚假特征误导姿态估计。动态干扰同样加剧检测难度,城市中行人车辆遮挡着陆点、森林枝叶摆动产生动态纹理干扰、水域波浪致反光区域变化,均影响检测稳定性。极端气象与特殊光照叠加场景更严峻,暴雪覆盖标识且压缩图像动态范围、沙尘暴致图像失真、隧道出入口光线突变致曝光异常,都会造成着陆目标特征失效。这些干扰常叠加出现,进一步放大误差,需在算法抗干扰、多模态融合及环境适应性上深化研究,以提升无人机自主着陆的环境感知能力<sup>[122]</sup>。

### 5.2 计算资源约束

深度学习模型的复杂结构与庞大参数对计算资源需求较高,而无人机作为受限平台,在算力、存储和功耗上面临多重制约。计算能力方面,受限于体积、重量与功耗,无人机通常仅能搭载嵌入式设备(如NVIDIA Jetson模块),其算力远低于高端GPU,处理复杂模型时易出现实时推理瓶颈,导致着陆感知延迟,影响安全性与精度。存储容量方面,模型参数、训练数据及传感器实时数据均需大量空间,而无人机有限的存储空间可能导致模型简化或数据丢失,进一步影响系统性能。功耗层面,无人机依赖电池供电,深度学习计算设备的高功耗会显著缩短续航时间,尽管需在算法优化与设备选型上平衡性能与能耗,但技术实现存在难度。这些资源限制迫使深度学习在无人机上的应用需聚焦轻量级模型设计、高效存储策略及低功耗硬件适配,以突破实时性、泛化能力与续航时间的多重瓶颈<sup>[123]</sup>。

### 5.3 数据依赖与标注难题

深度学习模型的性能高度依赖于大量高质量的训练数据。在无人机自主着陆领域,为了训练出能够准确识别各种着陆场景和应对复杂环境的模型,需要收集丰富多样的着陆数据,包括不同场景、不同气象条件、不同光照条件下的图像和传感器数据等。然而,数据的采集过程面临诸多困难。实际飞行试验存在安全风险,且受到天气、场地等条件的限制,难以大规模进行。通过仿真生成数据虽然可以在一定程度上解决数据量不足的问题,但仿真数据与真实数据之间存在一定的差异,即所谓的“仿真-现实差距”,这可能导致模型在真实场景中的泛化能力下降<sup>[124]</sup>。

标注是另一个难题。为了训练深度学习模型,需要对采集到的数据进行准确标注,标注出着陆点、障碍

物、场景类别等信息。对于无人机自主着陆相关的数据,标注工作通常需要专业知识和经验,且过程繁琐、耗时费力。在标注图像数据时,需要精确地绘制出着陆点的位置、障碍物的轮廓等,对于复杂的场景,标注的准确性和一致性难以保证。标注过程中可能存在人为误差,不同标注人员对同一数据的标注可能存在差异,这会影响到模型的训练效果。此外,随着深度学习模型对数据需求的不断增加,数据标注的工作量也呈指数级增长,进一步加剧了数据标注的难题。

综上,本节分析了基于深度学习的无人机自主着陆技术面临的挑战。复杂环境适应性方面,自然环境中的气象条件和地形地貌,以及人为环境中的电磁干扰和光照变化,都会影响无人机的感知能力;计算资源约束方面,无人机计算能力、存储容量和功耗的限制,制约了深度学习模型的应用;数据依赖与标注难题方面,深度学习对大量高质量数据的需求难以满足,且数据标注存在准确性和一致性问题。

## 6 发展趋势与展望

### 6.1 多源数据融合

面对复杂环境适应性不足的挑战,对抗训练、元学习、域适应等技术已用于增强模型鲁棒性与泛化能力,而多源传感器协同是提升环境感知的核心方向。如麻省理工学院研究所,融合多源数据可显著提升城市场景下无人机自主着陆的安全性与准确性,除传统视觉、激光雷达外,红外传感器、毫米波雷达等设备的协同策略需重点优化。从传感器特性适配来看,红外传感器(含热成像)可捕捉物体热辐射,不受光照限制,夜间或浓烟场景中,能通过着陆平台(如车辆、舰艇)与环境的温度差标注区域边界,规避RGB摄像头因光晕、阴影导致的特征丢失;毫米波雷达可穿透雨、雾、雪,精准测量距离与相对速度,在恶劣气象下补充视觉缺失的深度信息,且对金属目标(如舰艇甲板)的强反射特性,能辅助动态平台轨迹定位,减少反光或植被遮挡的跟踪偏差。协同策略需围绕“时空对齐-特征互补-动态权重分配”展开。时间同步上,利用LSTM对红外(低帧率)与RGB(高帧率)数据的时序偏差建模,如动态舰艇场景中校准两类数据的时间关联,确保目标特征匹配;空间对齐上,通过3D卷积神经网络处理红外热轮廓与RGB纹理的异构性,如透明玻璃场景中,融合红外捕捉的后方热辐射与RGB环境纹理,明确玻璃边界以避免碰撞;动态权重分配则按场景调整传感器占比。同时,通过异构数据编码网络统一多模态数据格式,借助Transformer自注意力机制捕捉长距离依赖,如城市峡谷场景中关联雷达建筑轮廓、红外热源与RGB街道纹理,精准判断安全着陆区<sup>[125-127]</sup>。

实验验证显示,卡内基梅隆大学通过AirSim仿真,毫米波雷达与红外在恶劣天气中可提供可靠信息,融合视觉后能精准识别着陆点;加州大学伯克利分校融合视觉、激光雷达与红外数据,使目标检测准确率提升15%,语义分割mIoU提高10%。因此在技术方面应该针对性构建基于Transformer的跨模态融合框架,借助自注意力机制对齐视觉、激光雷达等多源数据时空特征;引入自适应权重分配算法,根据光照、气象等环境复杂度调整传感器权重;开发轻量化异构数据编码网络统一特征空间,结合LSTM修正时序偏差,提升复杂环境数据融合的准确性与实时性。未来,传感器微型化与低功耗技术将推动多设备集成,基于Transformer的自适应融合机制将进一步提升复杂场景感知精度,为自主着陆提供保障<sup>[128]</sup>。

### 6.2 模型轻量化

为应对无人机计算资源有限的挑战,研究人员将聚焦轻量化深度学习模型开发。通过设计高效网络结构削减参数数量,如运用深度可分离卷积、稀疏卷积优化传统卷积神经网络,同时采用模型剪枝与量化技术剔除冗余,在保障性能前提下降低计算复杂度与存储需求。在国际上,MIT开发的MobileNet系列模型为例,其采用深度可分离卷积技术,将传统卷积操作分解为深度卷积和逐点卷积,在图像分类任务中,参数数量相较于传统卷积神经网络减少约80%,计算量降低75%,却仍能保持相近的识别准确率,此技术已成功应用于小型无人机的目标检测场景,展现了技术可行性。硬件层面,伴随芯片技术革新,适配无人机的高性能低功耗计算芯片将成为发展方向。此外,边缘计算技术的广泛应用可有效分担无人机计算压力,可采用“结构优化+参数压缩”双路径策略,以深度可分离/稀疏卷积重构网络骨干,结合通道剪枝与8位量化减冗余参数;设计着陆专用轻量模块,聚焦着陆点、跑道等核心目标特征提取;搭建“云端训练-边缘部署”架构,借边缘节点分担推理压力,优化推理引擎适配无人机硬件算力与功耗约束。华为推出的Atlas500智能小站,能将无人机图像识别任务的处理时延降低40%,充分展现了边缘计算在无人机自主着陆领域的应用潜力,为无人机自主着陆技术突破提供有力支撑<sup>[129]</sup>。

### 6.3 仿真与真实场景结合

为了解决数据依赖和标注难题,以及提高深度学习模型在真实场景中的泛化能力,未来的研究将更加注重仿真与真实场景的结合。如在韩国团队开展的研究中,基于目标识别和跟踪的无人机自主着陆系统,通过构建逼真的仿真环境模拟复杂着陆场景,生成多模态仿真数据,并与真实飞行数据进行时空对齐融合,大幅提升模型训练效率。同时,利用增强现实技术构建沉

浸式训练场景,使无人机在虚拟环境中模拟真实任务执行,积累多场景应对策略,最终在实际场景验证中取得良好效果,形成“仿真数据生成-虚实协同训练-真实场景验证-仿真环境迭代”的闭环体系,通过 diffusion 或 GAN 算法生成高保真仿真数据,缩小“仿真-现实差距”;用域自适应训练对齐仿真与真实数据分布;开发增强现实混合训练系统,叠加虚拟障碍与真实场景丰富样本;建标注自动化工具,借半监督学习降标注成本,结合跨场景数据迁移提高模型泛化能力.例如在新加坡国立大学在 ROSUnitySim 的研究中,采用轻量化模型架构设计,将仿真训练的轻量化网络部署至无人机边缘设备,在算力能耗降低的同时,有效提升模型在真实场景的实时感知能力,为攻克相关难题提供了实践参考<sup>[130-132]</sup>.

#### 6.4 飞行安全

在无人机自主着陆的安全性方面,发展趋势围绕技术提升与法规完善展开.技术上,更先进的传感器和智能算法将被广泛应用,以此提升环境感知与应对突发情况的能力;多无人机协作、抗干扰和环境适应技术也会不断发展,增强整体安全性.同时,法规与标准会持续完善,促使行业更加重视安全.展望未来,无人机在自主着陆时将具备更强的可靠性和适应性.通过加强安全冗余设计,可降低设备故障风险;凭借技术的进步,无人机能在复杂环境下安全着陆的能力也会进一步提高,从而推动无人机在更多领域的广泛应用<sup>[133,134]</sup>.

本节展望了无人机自主着陆技术的未来发展方向.多源数据融合将综合利用多种传感器数据,提升环境感知能力;模型轻量化旨在降低模型计算复杂度和存储需求,开发更适合无人机的计算芯片;仿真与真实场景结合有助于解决数据依赖和标注难题,提高模型泛化能力;同时,更加注重飞行安全,完善法规标准,推动无人机在更多领域广泛应用.

## 7 结论

本研究系统梳理了基于视觉与深度学习的无人机自主着陆场景感知技术体系,通过对场景特征、核心技术、数据集应用及现实挑战的深度剖析,形成以下具有实践指导意义的专业判断与价值总结.从技术演进逻辑来看,无人机自主着陆已完成从“传感器驱动”向“智能感知”的核心转型.深度学习的介入彻底突破了传统技术对预设规则和单一传感器的依赖,其中目标检测领域的 YOLO 系列算法凭借实时性优势成为动态场景首选,RCNN 系列以高精度适配静态精密着陆需求,语义分割的 FCN 与 U-Net 等算法则分别在快速场景粗分类和复杂地形精细识别中形成功能互补,姿态估计技术通过单目视觉与多目视觉的协同发展,平衡了成本

与精度的矛盾,这些技术的差异化应用构建了覆盖不同场景需求的完整技术矩阵.场景适配能力的强弱已成为技术落地的核心评判标准.静态平台着陆通过降落标识、跑道检测、地基引导的梯度化技术路径,实现了从低成本部署到厘米级定位的需求覆盖,其技术关键在于通过注意力机制与 Transformer 架构优化特征提取与语义理解精度;动态平台着陆则依托轻量化算法设计与多模态数据融合,有效缓解了运动跟踪延迟与环境干扰问题,车载与舰艇平台的算法选型差异验证了“场景特性决定技术路线”的核心原则;特殊场景着陆的技术突破则集中体现为鲁棒化设计与环境建模能力的提升,为无人机在复杂地形与极端条件下的安全着陆提供了关键支撑.

当前技术发展仍面临“三重矛盾”的核心制约:复杂环境下的感知鲁棒性与实际应用需求的矛盾,深度学习模型的计算复杂度与无人机硬件资源约束的矛盾,高质量数据需求与数据采集标注成本的矛盾.这三重矛盾也明确了技术迭代的核心方向——多源数据融合通过整合视觉、激光雷达、红外等异构数据,可突破单一传感器的环境适应性局限;模型轻量化借助深度可分离卷积、剪枝量化等技术,实现性能与资源消耗的最优平衡;仿真与真实场景的协同训练则为解决数据依赖难题提供了高效路径,三者共同构成了突破技术瓶颈的关键抓手.从产业应用价值来看,本研究明确了技术落地的优先级路径:民用领域应优先推进静态平台着陆技术的规模化应用,通过低成本方案覆盖巡检、测绘等常规场景;特种领域需重点突破动态平台与特殊场景的核心技术,满足军事、救援等高端需求;跨场景通用技术则应聚焦多源融合与轻量化模型的研发,构建具备广泛适配性的技术底座.未来,随着这些技术方向的持续突破,无人机自主着陆将实现从“可控安全”向“智能高效”的进阶,实现复杂环境下高可靠、高安全、低成本落地,其应用边界也将进一步拓展至低空物流、城市交通、极端环境作业等更广阔的领域,为低空经济的发展提供核心技术支撑.

#### 参考文献

- [1] SHAH ALAM M, OLUOCH J. A survey of safe landing zone detection techniques for autonomous unmanned aerial vehicles (UAVs) [J]. Expert Systems with Applications, 2021, 179: 115091.
- [2] MANGINA E, O'KEEFFE E, EYERMAN J, et al. Drones for live streaming of visuals for people with limited mobility[C]//2016 22nd International Conference on Virtual System & Multimedia. Piscataway: IEEE, 2017: 1-6.

- [3] GAUTAM A, SUJIT P B, SARIPALLI S. A survey of autonomous landing techniques for UAVs[C]//2014 International Conference on Unmanned Aircraft Systems. Piscataway: IEEE, 2014: 1210-1218.
- [4] 钟映春, 张文祥, 王波, 等. 电力巡检无人机自主降落的引导系统与策略[J]. 光学精密工程, 2022, 30(11): 1362-1373.  
ZHONG Y C, ZHANG W X, WANG B, et al. Navigation system and strategies for electric inspecting UAV autonomously landing[J]. Optics and Precision Engineering, 2022, 30(11): 1362-1373. (in Chinese)
- [5] RAEI H, CHO Y, PARK K. Autonomous landing on moving targets using LiDAR, Camera and IMU sensor Fusion[C]//2022 13th Asian Control Conference. Piscataway: IEEE, 2022: 419-423.
- [6] CHO G, CHOI J, BAE G, et al. Autonomous ship deck landing of a quadrotor UAV using feed-forward image-based visual servoing[J]. Aerospace Science and Technology, 2022, 130: 107869.
- [7] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [8] ZAREMBA W, SUTSKEVER I, VINYALS O. Recurrent neural network regularization[EB/OL]. (2015-02-19)[2025-10-10]. <https://arXiv.org/abs/1409.2329>.
- [9] WANG Y X, ZHU J G, CAO L, et al. Long short-term memory network with transfer learning for lithium-ion battery capacity fade and cycle life prediction[J]. Applied Energy, 2023, 350: 121660.
- [10] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[EB/OL]. (2023-08-02)[2025-10-10]. <https://arxiv.org/abs/1706.03762>.
- [11] 谷美颖, 李航, 张家伟, 等. 基于视觉的无人机定位与导航方法研究综述[J]. 电子学报, 2025, 53(3): 651-685.  
GU M Y, LI H, ZHANG J W, et al. A review of vision-based UAV localization and navigation methods[J]. Acta Electronica Sinica, 2025, 53(3): 651-685. (in Chinese)
- [12] 钟春来, 杨洋, 曹立佳, 等. 基于视觉的无人机自主着陆研究综述[J]. 航空兵器, 2023, 30(5): 104-114.  
ZHONG C L, YANG Y, CAO L J, et al. A review of vision-based autonomous UAV landing research[J]. Aero Weaponry, 2023, 30(5): 104-114. (in Chinese)
- [13] KATKURI A V R, MADAN H, KHATRI N, et al. Autonomous UAV navigation using deep learning-based computer vision frameworks: A systematic literature review[J]. Array, 2024, 23: 100361.
- [14] 马宁, 曹云峰. 面向无人机自主着陆的视觉感知与姿态估计方法综述[J]. 自动化学报, 2024, 50(7): 1284-1304.  
MA N, CAO Y F. A survey on vision-based sensing and pose estimation methods for UAV autonomous landing[J]. Acta Automatica Sinica, 2024, 50(7): 1284-1304. (in Chinese)
- [15] HASSANALIAN M, ABDELKEFI A. Classifications, applications, and design challenges of drones: A review[J]. Progress in Aerospace Sciences, 2017, 91: 99-131.
- [16] 赵良玉, 李丹, 赵辰悦, 等. 无人机自主降落标识检测方法若干研究进展[J]. 航空学报, 2022, 43(9): 025882.  
ZHAO L Y, LI D, ZHAO C Y, et al. Some achievements on detection methods of UAV autonomous landing markers[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(9): 025882. (in Chinese)
- [17] XIN L, TANG Z M, GAI W Q, et al. Vision-based autonomous landing for the UAV: A review[J]. Aerospace, 2022, 9(11): 634.
- [18] CHEN C L, ZHENG Z Y, XU T Y, et al. YOLO-based UAV technology: A review of the research and its applications[J]. Drones, 2023, 7(3): 190.
- [19] GARCÍA-PULIDO J A, PAJARES G, DORMIDO S, et al. Recognition of a landing platform for unmanned aerial vehicles by using computer vision-based techniques[J]. Expert Systems with Applications, 2017, 76: 152-165.
- [20] MORE D S, SURESH S, D'SOUZA J M, et al. Landmark detection for auto landing of quadcopter using YOLOv5[C]//Intelligent Control, Robotics, and Industrial Automation. Singapore: Springer, 2023: 3-12.
- [21] OLSON E. AprilTag: A robust and flexible visual fiducial system[C]//2011 IEEE International Conference on Robotics and Automation. Piscataway: IEEE, 2011: 3400-3407.
- [22] WANG J, OLSON E. AprilTag 2: Efficient and robust fiducial detection[C]//2016 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway: IEEE, 2016: 4193-4198.
- [23] MU L X, LI Q L, WANG B, et al. A vision-based autonomous landing guidance strategy for a micro-UAV by the modified camera view[J]. Drones, 2023, 7(6): 400.
- [24] WU J G, ZHANG Z, HUANG W H. Semantic map construction of UAV autonomous landing in unknown environment[C]//2024 36th Chinese Control and Decision Conference. Piscataway: IEEE, 2024: 5018-5025.

- [25] WANG Z, WANG B, TANG C Y, et al. Pose and velocity estimation algorithm for UAV in visual landing[C]//2020 39th Chinese Control Conference. Piscataway: IEEE, 2020: 3713-3718.
- [26] 乐辉. 基于视觉的四旋翼无人机着陆标识分类与识别研究[D]. 镇江: 江苏大学, 2021.  
LE H. Research on Classification and Recognition of Quad-rotor UAV Landing Sign based on Vision[D]. Zhenjiang: Jiangsu University, 2021. (in Chinese)
- [27] 方琪鸿. 基于单目视觉与惯性融合的无人机自主降落导航技术研究[D]. 成都: 电子科技大学, 2024.  
FANG Q H. Research on Autonomous Landing Navigation Technology for Unmanned Aerial Vehicles Based on Monocular Vision and Inertial Fusion[D]. Chengdu: University of Electronic Science and Technology of China, 2024. (in Chinese)
- [28] MA M Y, SHEN S G, HUANG Y C. Enhancing UAV visual landing recognition with YOLO's object detection by onboard edge computing[J]. *Sensors*, 2023, 23(21): 8999.
- [29] MA N, WENG X R, CAO Y F, et al. Monocular-vision-based precise runway detection applied to state estimation for carrier-based UAV landing[J]. *Sensors*, 2022, 22(21): 8385.
- [30] LIU X X, XUE W H, XU X L, et al. Research on unmanned aerial vehicle (UAV) visual landing guidance and positioning algorithms[J]. *Drones*, 2024, 8(6): 257.
- [31] LI Y, XIA Y, ZHENG G J, et al. YOLO-RWY: A novel runway detection model for vision-based autonomous landing of fixed-wing unmanned aerial vehicles[J]. *Drones*, 2024, 8(10): 571.
- [32] WANG Z Y, ZHAO D, CAO Y F. Visual navigation algorithm for night landing of fixed-wing unmanned aerial vehicle[J]. *Aerospace*, 2022, 9(10): 615.
- [33] WANG Q, FENG W Q, ZHAO H B, et al. VALNet: Vision-based autonomous landing with airport runway instance segmentation[J]. *Remote Sensing*, 2024, 16(12): 2161.
- [34] 马宁, 曹云峰, 王指辉, 等. 基于YOLOv5网络架构的着陆跑道检测算法研究[J]. *激光与光电子学进展*, 2022, 59(14): 199-205.  
MA N, CAO Y F, WANG Z H, et al. Landing runway detection algorithm based on YOLOv5 network architecture[J]. *Laser & Optoelectronics Progress*, 2022, 59(14): 199-205. (in Chinese)
- [35] ZHANG X P, HE Z Z, MA Z, et al. VIAE-net: An end-to-end altitude estimation through monocular vision and inertial feature fusion neural networks for UAV autonomous landing[J]. *Sensors*, 2021, 21(18): 6302.
- [36] WUBBEN J, FABRA F, CALAFATE C T, et al. Accurate landing of unmanned aerial vehicles using ground pattern recognition[J]. *Electronics*, 2019, 8(12): 1532.
- [37] KHAN M U, DIL M, MISBAH M, et al. TransLearn-YOLOx: Improved-YOLO with transfer learning for fast and accurate multiclass UAV detection[C]//2023 International Conference on Communication, Computing and Digital Systems (C-CODE). Piscataway: IEEE, 2023: 1-7.
- [38] XIE M J, CAO Y R, JIANG C H, et al. Object detection in UAV ground-based visual landing process based on improved faster R-CNN[M]//Advances in Guidance, Navigation and Control. Singapore: Springer Nature Singapore, 2023: 5012-5021.
- [39] TANG D Q, SHEN L C, XIANG X J, et al. N-cameras-enabled joint pose estimation for auto-landing fixed-wing UAVs[J]. *Drones*, 2023, 7(12): 693.
- [40] ARAAR O, AOUF N, VITANOV I. Vision based autonomous landing of multicopter UAV on moving platform[J]. *Journal of Intelligent & Robotic Systems*, 2017, 85(2): 369-384.
- [41] TZOUMANIKAS D, LI W B, GRIMM M, et al. Fully autonomous micro air vehicle flight and landing on a moving target using visual-inertial estimation and model-predictive control[J]. *Journal of Field Robotics*, 2019, 36(1): 49-77.
- [42] 蔡炳锋. 基于机器视觉的无人机自适应移动着陆研究[D]. 杭州: 浙江大学, 2020.  
CAI B F. Research on Vision Based Adaptive Landing of UAV on a Moving Platform[D]. Hangzhou: Zhejiang University, 2020. (in Chinese)
- [43] WANG Z L, SHE H P, SI W Y. Autonomous landing of multi-rotors UAV with monocular gimbaled camera on moving vehicle[C]//2017 13th IEEE International Conference on Control & Automation. Piscataway: IEEE, 2017: 408-412.
- [44] BACA T, STEPAN P, SPURNY V, et al. Autonomous landing on a moving vehicle with an unmanned aerial vehicle[J]. *Journal of Field Robotics*, 2019, 36(5): 874-891.
- [45] BOROWCZYK A, NGUYEN D T, NGUYEN A P, et al. Autonomous landing of a quadcopter on a high-speed ground vehicle[J]. *Journal of Guidance, Control, and Dynamics*, 2017, 40(9): 2378-2385.
- [46] JIANG S J, LUO B, LIU J, et al. UAV-based vehicle de-

- tection by multi-source images[C]//the 2nd CCF Chinese Conference. Singapore: Springer, 2017: 38-49.
- [47] 王忠言, 李波, 刘茹艳, 等. 基于 yolov4 算法的无人机单目测距算法[J]. 机械设计与制造工程, 2022, 51(3): 58-62.
- WANG Z Y, LI B, LIU R Y, et al. UAV monocular ranging algorithm based on yolov4 algorithm[J]. Machine Design and Manufacturing Engineering, 2022, 51(3): 58-62. (in Chinese)
- [48] IDROVO P, VALLADOLID J D, DUTAN D, et al. A novel proposal for traffic officer detection in autonomous vehicles using convolutional networks YOLO v3, v5, and v8[EB/OL]. (2024-05-16)[2025-10-10]. <https://www.preprints.org/manuscript/202405.1078>.
- [49] SHEN K, ZHUANG Y, CHEN Y X, et al. AeroNet: An efficient relative localization and object detection network for cooperative aerial-ground unmanned vehicles[J]. Pattern Recognition Letters, 2023, 171: 28-37.
- [50] LI J H, WANG X H, CUI H R, et al. Research on detection technology of autonomous landing based on airborne vision[J]. IOP Conference Series: Earth and Environmental Science, 2020, 440(4): 042093.
- [51] 刘健, 张祥甫, 于志军, 等. 基于改进 ERFNet 的无人直升机着舰环境语义分割[J]. 电讯技术, 2020, 60(1): 40-46.
- LIU J, ZHANG X F, YU Z J, et al. Semantic segmentation of landing environment for unmanned helicopter based on improved ERFNet[J]. Telecommunication Engineering, 2020, 60(1): 40-46. (in Chinese)
- [52] QIU J T, YU F F, XU F R, et al. Improved you only look once model for UAVs/ships relative attitude detection[M]//Proceedings of 2024 Chinese Intelligent Systems Conference. Singapore: Springer Nature Singapore, 2024: 218-226.
- [53] ZHOU R, SHE J Y, QI N M, et al. Pose estimation algorithm for helicopter landing based on YOLO and PNP[C]//Advances in Guidance, Navigation and Control. Singapore: Springer, 2022: 3019-3028.
- [54] XU G L, ZHANG Y, JI S Y, et al. Research on computer vision-based for UAV autonomous landing on a ship[J]. Pattern Recognition Letters, 2009, 30(6): 600-605.
- [55] POLVARA R, SHARMA S, WAN J, et al. Towards autonomous landing on a moving vessel through fiducial markers[C]//2017 European Conference on Mobile Robots. Piscataway: IEEE, 2017: 1-6.
- [56] 任毅. 基于视觉引导的旋翼无人机移动平台自主着陆技术研究[D]. 绵阳: 西南科技大学, 2019.
- REN Y. Research on Vision Based Autonomous Landing of Rotor UAV on a Moving Platform[D]. Mianyang: Southwest University of Science and Technology, 2019. (in Chinese)
- [57] WANG L, JIANG X Q, WANG D, et al. Research on aerial autonomous docking and landing technology of dual multi-rotor UAV[J]. Sensors, 2022, 22(23): 9066.
- [58] ARAFAT M Y, ALAM M M, MOH S. Vision-based navigation techniques for unmanned aerial vehicles: Review and challenges[J]. Drones, 2023, 7(2): 89.
- [59] SARIPALLI S, MONTGOMERY J F, SUKHATME G S. Vision-based autonomous landing of an unmanned aerial vehicle[C]//Proceedings 2002 IEEE International Conference on Robotics and Automation. Piscataway: IEEE, 2002: 2799-2804.
- [60] KONG W W, ZHOU D L, ZHANG D B, et al. Vision-based autonomous landing system for unmanned aerial vehicle: A survey[C]//2014 International Conference on Multisensor Fusion and Information Integration for Intelligent Systems. Piscataway: IEEE, 2014: 1-8.
- [61] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [62] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 779-788.
- [63] KHANAM R, HUSSAIN M. What is YOLOv5: A deep look into the internal features of the popular object detector[EB/OL]. (2024-07-30)[2025-10-10]. <https://arXiv.org/abs/2407.20892>.
- [64] VARGHESE R, M S. YOLOv8: A novel object detection algorithm with enhanced performance and robustness[C]//2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems. Piscataway: IEEE, 2024: 1-6.
- [65] TIAN Y J, YE Q X, DOERMANN D. YOLOv12: Attention-centric real-time object detectors[EB/OL]. (2025-02-18)[2025-10-10]. <https://arXiv.org/abs/2502.12524>.
- [66] ZHANG Z, CHEN L, WANG Q F, et al. Monocular visual pose measurement for autonomous landing in unknown environments[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2025, 35(8): 7592-7604.
- [67] DING J, XUE N, XIA G S, et al. Object detection in aeri-

- al images: A large-scale benchmark and challenges[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022, 44(11): 7778-7796.
- [68] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 936-944.
- [69] WANG H, SONG Z L. Improved mosaic: Algorithms for more complex images[J]. *Journal of Physics: Conference Series*, 2020, 1684(1): 012094.
- [70] NEPAL U, ESLAMIAT H. Comparing YOLOv3, YOLOv4 and YOLOv5 for autonomous landing spot detection in faulty UAVs[J]. *Sensors*, 2022, 22(2): 464.
- [71] JIANG P Y, ERGU D J, LIU F Y, et al. A review of yolo algorithm developments[J]. *Procedia Computer Science*, 2022, 199: 1066-1073.
- [72] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot MultiBox detector[C]//Computer Vision - ECCV 2016. Cham: Springer, 2016: 21-37.
- [73] PARK C H, AHN S M. Autonomous landing of drones using deep learning GPS-denied environments[C]//Proceedings of the Korean Society of Computer Information Conference. Daejeon: Korean Society of Computer Information, 2023: 15-18.
- [74] JIANG C C, REN H Z, YE X, et al. Object detection from UAV thermal infrared images and videos using YOLO models[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2022, 112: 102912.
- [75] SERRANO K K D, BANDALA A A. YOLO-based terrain classification for UAV safe landing zone detection[C]//2023 IEEE Region 10 Symposium. Piscataway: IEEE, 2023: 1-5.
- [76] ROLLAND E G A, GRØNTVED K A R, CHRISTENSEN A L, et al. Autonomous UAV volcanic plume sampling based on machine vision and path planning[C]//2024 International Conference on Unmanned Aircraft Systems. Piscataway: IEEE, 2024: 1064-1071.
- [77] YUAN B X, MA W Y, WANG F. High speed safe autonomous landing marker tracking of fixed wing drone based on deep learning[J]. *IEEE Access*, 2022, 10: 80415-80436.
- [78] 伍瀚, 孙浩, 计科峰, 等. 时序信息引导跨视角特征融合的多无人机多目标跟踪方法[J]. *电子学报*, 2025, 53(3): 728-743.
- WU H, SUN H, JI K F, et al. Temporal-guided cross-view feature fusion network for multi-drone multi-object tracking[J]. *Acta Electronica Sinica*, 2025, 53(3): 728-743. (in Chinese)
- [79] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 3431-3440.
- [80] RONNEBERGER O, FISCHER P, BROX T. U-Net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Cham: Springer, 2015: 234-241.
- [81] YANG L J, YE J, ZHANG Y, et al. A semantic SLAM-based method for navigation and landing of UAVs in indoor environments[J]. *Knowledge-Based Systems*, 2024, 293: 111693.
- [82] CLAUDET T, TOMITA K, HO K. Benchmark analysis of semantic segmentation algorithms for safe planetary landing site selection[J]. *IEEE Access*, 2022, 10: 41766-41775.
- [83] PUTRANTO H Y, NUR IRFANSYAH A, ATTAMIMI M. Identification of safe landing areas with semantic segmentation and contour detection for delivery UAV[C]//2022 9th International Conference on Information Technology, Computer, and Electrical Engineering. Piscataway: IEEE, 2022: 254-257.
- [84] LOERA-PONCE J A, MERCADO-RAVELL D A, BECERRA I, et al. Risk assessment for UAV autonomous landing in urban environments using semantic segmentation[M]//Advances in Artificial Intelligence - IBERAMIA 2024. Cham: Springer Nature Switzerland, 2025: 197-208.
- [85] BENJWAL A, UDAY P, VADDURI A, et al. Safe landing zone detection for UAVs using image segmentation and super resolution[C]//2023 18th International Conference on Machine Vision and Applications. Piscataway: IEEE, 2023: 1-6.
- [86] RYU H, LIM J, LEE H J, et al. An estimation method for vision-based autonomous landing system for fixed wing aircraft[J]. *The International Journal of Robotics Research*, 2025, 44(6): 952-971.
- [87] SEFIDGAR M, LANDRY R. Unstable landing platform pose estimation based on camera and range sensor homogeneous fusion (CRHF)[J]. *Drones*, 2022, 6(3): 60.
- [88] BALDINI F, ANANDKUMAR A, MURRAY R M. Learning pose estimation for UAV autonomous navigation and landing using visual-inertial sensor data[C]//

- 2020 American Control Conference. Piscataway: IEEE, 2020: 2961-2966.
- [89] DE CROON G C H E, HO H W, DE WAGTER C, et al. Optic-flow based slope estimation for autonomous landing[J]. *International Journal of Micro Air Vehicles*, 2013, 5(4): 287-297.
- [90] WANG P K, WU L, QI J X, et al. Unmanned aerial vehicles object detection based on image haze removal under sea fog conditions[J]. *IET Image Processing*, 2022, 16(10): 2709-2721.
- [91] DUTRANNOIS T, NGUYEN T T, HAMESSE C, et al. Visual SLAM for autonomous drone landing on a maritime platform[C]//2022 International Symposium on Measurement and Control in Robotics. Piscataway: IEEE, 2022: 1-7.
- [92] LIN Y X, LAI Y C. Deep learning-based navigation system for automatic landing approach of fixed-wing UAVs in GNSS-denied environments[J]. *Aerospace*, 2025, 12(4): 324.
- [93] 吴鹏飞, 石章松, 黄隽, 等. 基于改进 SSD 网络的着舰标志识别方法[J]. *电光与控制*, 2022, 29(1): 88-92.
- WU P F, SHI Z S, HUANG J, et al. Landing mark identification method based on improved SSD network[J]. *Electronics Optics & Control*, 2022, 29(1): 88-92. (in Chinese)
- [94] LIU F, SHAN J Y, XIONG B Y, et al. A real-time and multi-sensor-based landing area recognition system for UAVs[J]. *Drones*, 2022, 6(5): 118.
- [95] YI S, LI J J, JIANG G, et al. CCTseg: A cascade composite transformer semantic segmentation network for UAV visual perception[J]. *Measurement*, 2023, 211: 112612.
- [96] DHAMI H S, IGNATYEV D, TSOUSDOS A. Semantic segmentation based mapping systems for the safe and precise landing of flying vehicles[J]. *IFAC-PapersOnLine*, 2022, 55(22): 310-315.
- [97] ZHUANG J D, CHEN X, DAI M, et al. A semantic guidance and transformer-based matching method for UAVs and satellite images for UAV geo-localization[J]. *IEEE Access*, 2022, 10: 34277-34287.
- [98] XU Y, ZHONG D S, ZHOU J H, et al. A novel UAV visual positioning algorithm based on A-YOLOX[J]. *Drones*, 2022, 6(11): 362.
- [99] CHEN R B, XU Y, BIN SINAL M S, et al. Swin-YOLOX for autonomous and accurate drone visual landing[J]. *IET Image Processing*, 2024, 18(14): 4731-4744.
- [100] ZHANG Y, LIU X, XIAO C S, et al. Research on autonomous landing method of shipborne uncrewed aerial vehicle based on visual recognition[J]. *IEEE Access*, 2025, 13: 179994-180005.
- [101] DAI W, ZHAI Z J, WANG D Z, et al. YOMO-runwaynet: A lightweight fixed-wing aircraft runway detection algorithm combining YOLO and MobileRunwaynet[J]. *Drones*, 2024, 8(7): 330.
- [102] JIANG B, CHEN Z H, TAN J T, et al. A real-time semantic segmentation method based on STDC-CT for recognizing UAV emergency landing zones[J]. *Sensors*, 2023, 23(14): 6514.
- [103] ZHANG Q T, XIA Q Y, WEI L S, et al. A vision-based method for UAV autonomous landing area detection[C]// *Intelligence Science V*. Cham: Springer, 2025: 204-213.
- [104] MUMUNI A, MUMUNI F. Data augmentation: A comprehensive survey of modern approaches[J]. *Array*, 2022, 16: 100258.
- [105] TRUONG N Q, LEE Y W, OWAIS M, et al. SlimDeblurGAN-based motion deblurring and marker detection for autonomous drone landing[J]. *Sensors*, 2020, 20(14): 3918.
- [106] NAUFAL C, SOLANO-CORREA Y T, MARRUGO A G. YOLO-based multi-scale ground control point detection in UAV surveying[C]//2023 IEEE Colombian Caribbean Conference (C3). Piscataway: IEEE, 2024: 1-5.
- [107] LEE M, SHIN S G, JANG S, et al. Visual-based landing guidance system of UAV with deep learning technique for environments of visual-detection impairment[J]. *International Journal of Control, Automation and Systems*, 2022, 20(5): 1735-1744.
- [108] LIM J, KIM M, YOO H, et al. Autonomous multirotor UAV search and landing on safe spots based on combined semantic and depth information from an onboard camera and LiDAR[J]. *IEEE/ASME Transactions on Mechatronics*, 2024, 29(5): 3960-3970.
- [109] ZHANG Z Q, ZHANG Y F, XIANG S, et al. KDP-net: An efficient semantic segmentation network for emergency landing of unmanned aerial vehicles[J]. *Drones*, 2024, 8(2): 46.
- [110] XUE C, XIA Y L, WU M J, et al. EL-YOLO: An efficient and lightweight low-altitude aerial objects detector for onboard applications[J]. *Expert Systems with Applications*, 2024, 256: 124848.
- [111] 张琴. 基于视觉辅助定位的无人机智能降落系统研究与实现[D]. 南京: 南京邮电大学, 2020.
- ZHANG Q. Research and Implementation of UAV Intelligent Landing System Based on Vision-Assisted Posi-

- tioning[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2020. (in Chinese)
- [112] YU L J, LUO C, YU X R, et al. Deep learning for vision-based micro aerial vehicle autonomous landing[J]. *International Journal of Micro Air Vehicles*, 2018, 10(2): 171-185.
- [113] DAVIS J, GOADRICH M. The relationship between precision-recall and ROC curves[C]//*Proceedings of the 23rd International Conference on Machine Learning*. New York: ACM, 2006: 233-240.
- [114] ZHENG Z H, WANG P, LIU W, et al. Distance-IoU loss: Faster and better learning for bounding box regression[J]. *Proceedings of the AAAI Conference on Artificial Intelligence*, 2020, 34(7): 12993-13000.
- [115] DU D W, WEN L Y, ZHU P F, et al. VisDrone-DET2020: The vision meets drone object detection in image challenge results[C]//*Computer Vision - ECCV 2020 Workshops*. Cham: Springer, 2020: 692-712.
- [116] LYU Y, VOSSELMAN G, XIA G S, et al. UAVid: A semantic segmentation dataset for UAV imagery[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 165: 108-119.
- [117] SUN H M, GUO J C, MENG Z B, et al. EVD4UAV: An altitude-sensitive benchmark to evade vehicle detection in UAV[C]//*2024 IEEE Intelligent Vehicles Symposium*. Piscataway: IEEE, 2024: 545-552.
- [118] LI T J, LIU J, ZHANG W, et al. UAV-human: A large benchmark for human behavior understanding with unmanned aerial vehicles[C]//*2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE, 2021: 16261-16270.
- [119] XIA G S, HU J W, HU F, et al. AID: A benchmark data set for performance evaluation of aerial scene classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(7): 3965-3981.
- [120] MUELLER M, SMITH N, GHANEM B. A benchmark and simulator for UAV tracking[C]//*Computer Vision - ECCV 2016*. Cham: Springer, 2016: 445-461.
- [121] 汪进中, 戴顺, 张秀伟, 等. 无人机视角多源目标检测数据集 UAV-RGBT 及算法基准[J]. *电子学报*, 2025, 53(3): 686-704.
- WANG J Z, DAI S, ZHANG X W, et al. UAV-RGBT multispectral object detection dataset and algorithm benchmark[J]. *Acta Electronica Sinica*, 2025, 53(3): 686-704. (in Chinese)
- [122] HAUSMANN P, MEESS H, ELGER G. Image segmentation based emergency landing for autonomous and automated unmanned aerial vehicles[EB/OL]. (2022)[2025-10-10]. [https://www.icas.org/icas\\_archive/ICAS2022/data/papers/ICAS2022\\_0725\\_paper.pdf](https://www.icas.org/icas_archive/ICAS2022/data/papers/ICAS2022_0725_paper.pdf).
- [123] NGUYEN P H, ARSALAN M, KOO J H, et al. Light-DenseYOLO: A fast and accurate marker tracker for autonomous UAV landing by visible light camera sensor on drone[J]. *Sensors*, 2018, 18(6): 1703.
- [124] HINNIGER C, RÜTER J. Synthetic training data for semantic segmentation of the environment from UAV perspective[J]. *Aerospace*, 2023, 10(7): 604.
- [125] GRLJ C G, KRZYNAR N, PRANJIĆ M. A decade of UAV docking stations: A brief overview of mobile and fixed landing platforms[J]. *Drones*, 2022, 6(1): 17.
- [126] SHAH S, DEY D, LOVETT C, et al. AirSim: High-fidelity visual and physical simulation for autonomous vehicles[C]//*Field and Service Robotics*. Cham: Springer, 2018: 621-635.
- [127] SHAHAM T R, SCHWETTMANN S, WANG F, et al. A multimodal automated interpretability agent[EB/OL]. (2025-02-11)[2025-10-20]. <https://arxiv.org/abs/2404.14394>.
- [128] HE H Q, LI C C, YANG R H, et al. Multisource data fusion and adversarial nets for landslide extraction from UAV-photogrammetry-derived data[J]. *Remote Sensing*, 2022, 14(13): 3059.
- [129] HOWARD A, SANDLER M, CHEN B, et al. Searching for MobileNetV3[C]//*2019 IEEE/CVF International Conference on Computer Vision*. Piscataway: IEEE, 2019: 1314-1324.
- [130] FAN B K, LI Y, ZHANG R Y, et al. Review on the technological development and application of UAV systems[J]. *Chinese Journal of Electronics*, 2020, 29(2): 199-207.
- [131] HU Y C, MENG W. ROSUnitySim: Development and experimentation of a real-time simulator for multi-unmanned aerial vehicle local planning[J]. *Simulation*, 2016, 92(10): 931-944.
- [132] BAIDYA R, JEONG H. Simulation and real-life implementation of UAV autonomous landing system based on object recognition and tracking for safe landing in uncertain environments[J]. *Frontiers in Robotics and AI*, 2024, 11: 1450266.
- [133] KAKALETSIS E, SYMEONIDIS C, TZELEPI M, et al. Computer vision for autonomous UAV flight safety: An overview and a vision-based safe landing pipeline example[J]. *ACM Computing Surveys*, 2021, 54(9): 1-37.
- [134] 王巍, 解慧, 魏忠诚, 等. 不确定需求下无人机任务分

配的两阶段鲁棒优化方法[J]. 电子学报, 2024, 52(10): 3552-3561.

WANG W, XIE H, WEI Z C, et al. Two-stage robust

optimization method for UAV task assignment under uncertain demand[J]. Acta Electronica Sinica, 2024, 52(10): 3552-3561. (in Chinese)

#### 作者简介



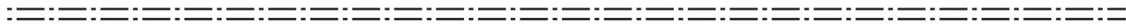
王中天 男, 2002年生, 安徽蚌埠人. 现为南京航空航天大学电子信息学院博士研究生. 主要研究方向为模式识别、计算机视觉、图像处理等.

E-mail: zt\_wang01@163.com



吴一全 男, 1963年生, 江苏启东人. 现为南京航空航天大学教授、博士生导师. 主要研究方向为视觉检测与图像测量、遥感图像处理与理解、红外目标检测与识别、视频处理与智能分析等.

E-mail: nuaaimage@163.com



## 勘 误

本刊2022年第50卷第8期《深度学习框架下群组行为识别算法综述》(作者:邓海刚,王传旭,李成伟,林晓萌)一文中,基金项目应为“国家自然科学基金(No. 61672305)”. 特此更正.

《电子学报》编辑部