doi:10.6041/j.issn.1000-1298.2026.22.022

基于轻量化分割与热图回归的农业场景路径提取方法

李大寨　刘晨博　章纪龙　王嘉豪
(北京航空航天大学机械工程及自动化学院, 北京 100191)

摘要: 为解决农业非结构化环境下机器人自主导航路径提取过程中路径断裂、中心线偏移、实时性差以及复杂环境适应性弱等问题,本文提出了一种基于轻量化分割与热图回归的农业场景导航路径提取方法。采用改进的 Deeplabv3 + 语义分割模型,结合轻量化 MobileNetV3 作为主干特征提取网络,有效提升了模型在嵌入式设备中推理速度,同时保持了较高的分割精度。基于语义分割获得的可行驶区域掩码图,设计了一种 Mobile - Unet 轻量化热图回归网络,通过加权均方误差与感知损失联合的组合损失函数,增强了模型对热图高响应区域关注能力,并提高了热图预测结果细节还原能力与结构一致性。提出一种基于热值引导贪心搜索与非极大值抑制相结合的导航路径提取算法,能够在复杂农业场景下从热图中高效、稳定提取连续平滑的导航路径。在自建的农业场景数据集上开展试验,验证了语义分割模型、热图回归网络及路径提取算法性能。结果表明,本文方法在保证语义分割精度基础上,提升了模型在 Jetson TX2 嵌入式平台上的推理速度;在热图回归任务中,所设计的 Mobile - Unet 模型结合 Mask 输入方案获得了最低均方误差(MSE)与最佳结构相似性指标(SSIM),且满足实时推理需求。所提方法兼顾了精度、速度与实际部署能力,具备良好的农业机器人导航应用前景。

关键词: 农业场景; 路径提取; 轻量化分割; 热图回归
中图分类号: TP242; S24　　**文献标识码**: A　　**文章编号**: 1000-1298(2026)22-0262-11　　**OSID**:

Agricultural Scene Navigation Path Extraction Method Based on Lightweight Semantic Segmentation and Heatmap Regression

Li Dazhai　Liu Chenbo　Zhang Jilong　Wang Jiahao
(*School of Mechanical Engineering and Automation*, *Beihang University*, *Beijing* 100191, *China*)

Abstract: Aiming to solve the problems of path discontinuity, centerline deviation, poor real-time performance, and weak adaptability to complex environments in unstructured agricultural environments of robot navigation path extraction, a navigation path extraction method for agricultural scenes was proposed based on semantic segmentation and heatmap regression. Firstly, an improved Deeplabv3 + semantic segmentation model was constructed by integrating the lightweight MobileNetV3 as the backbone network, which effectively improved the inference speed of the model in embedded devices while maintaining a high segmentation accuracy. Secondly, based on the drivable area mask obtained by semantic segmentation, a Mobile - Unet lightweight heatmap regression network was designed. Through the combined loss function of weighted mean square error and perceptual loss, the model's ability to focus on high-response areas of the heatmap was enhanced, and the detail restoration and structural consistency of the heatmap prediction results were improved. Finally, a navigation path extraction algorithm combining heat value-guided greedy search and non-maximum suppression (NMS) was proposed to efficiently and stably extract continuous and smooth navigation paths from the heatmaps in complex agricultural environments. Experiment was carried out on a self-built agricultural scene dataset, which verified the performance of the semantic segmentation model, heatmap regression network and path extraction algorithm. Results demonstrated that the proposed method maintained high segmentation accuracy while achieving superior inference speed on the Jetson TX2 embedded platform. In heatmap regression tasks, the proposed Mobile - Unet combined with Mask input achieved the lowest mean square error (MSE) and the highest structural similarity (SSIM), meeting the requirements for real-time navigation. Overall, the

收稿日期: 2025 - 06 - 21　修回日期: 2025 - 11 - 01
基金项目: 国家自然科学基金项目(51975022)
作者简介: 李大寨(1967—),男,副教授,博士,主要从事智能机器人研究,E-mail: lidazhai@buaa.edu.cn

proposed method took into account accuracy, speed and practical deployment capabilities, and had good application prospects in agricultural robot navigation.

Key words: agricultural scene; path extraction; lightweight segmentation; heatmap regression

0　引言

以智能控制和信息技术为核心的农业机器人，已成为全球智慧农业领域的核心组成部分[1-2]。农业机器人核心优势在于其自动化的特点，能够显著提高劳动效率，降低劳动成本。而自主导航技术则是实现这一优势的关键，它能使机器人根据不同地形与作物布局精准执行各类农业作业任务。

早期学者致力于 GNSS 导航系统研究，旨在借助此类系统使机器人在农业场景下实现较为精准的绝对定位[3-4]。但这类方法对作物枝叶及障碍物遮挡敏感，定位精度易受影响[5]。随着激光雷达、RGB-D 相机、IMU 等传感器的发展，自主导航技术研究取得了新的突破，许多学者致力于将其应用于农业场景中[6-8]。然而，农业环境动态特性使得单一传感器很难表现出更高的定位精度，于是研究者开始尝试多种传感器融合方法[9-10]。这些方法提升了系统环境感知与定位精度，但涉及多个传感器之间时间同步与空间标定，系统集成与实际部署困难，并且经济成本较高。

近几年基于机器视觉的自主导航得到了学者青睐[11]，为机器人在作业场景中寻找可行驶区域、避障以及路径规划提供了强力的支撑。Yang 等[12]利用 Segnet 和 Unet 训练语义分割网络，再对果园路况掩码区域进行处理，实现导航路径。Panda 等[13]提出了一种端到端的自主导航框架 Agronav，首先对输入的 RGB 图像进行语义分割，得到道路掩码，再对掩码区域进行语义线检测，输出道路中心线。He 等[14]使用 MVS 算法识别导航路径，利用 Otsu 分割创建了二进制图像，最终利用最小二乘法生成导航线。Liu 等[15]提出了一种单阶段导航路径提取网络 NPENet，直接将道路中心线作为神经网络预测目标，为果园环境中移动机器人提供准确的道路信息。总的来说，上述机器视觉方法都能很好地完成机器人在农业环境下自助导航和路径规划，但同时也存在一些局限：首先大多数方法生成的导航路径均为直线，导致应用场景受限；其次，传统路径提取方法受道路边界形态影响大，易出现中心线偏移、断裂或生成多个分支等现象。

针对上述问题，本文提出一种融合轻量化语义分割与热图回归的农业场景导航路径提取方法。旨在通过视觉感知技术为农业机器人生成安全、连续的导航路径，满足在非结构化场景下自主作业需求。

1　方法和数据

本文所构建的视觉感知与路径生成整体算法框架如图 1 所示。该方法执行流程主要为：对相机采集的 RGB 图像进行轻量化语义分割，提取可行驶区域；引入热图回归模型对分割后的结果进行处理，得到路径位置的二维概率分布；基于预测热图进行后处理，提取精确、连续的导航路径。

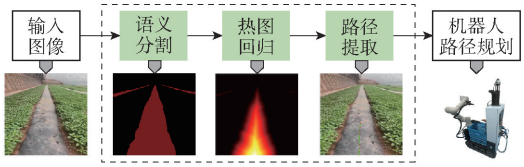


图 1　导航路径提取方法整体流程

Fig. 1　Overall flowchart of navigation path extraction

1.1　试验平台和设备

以自主设计研发的番茄采摘机器人为试验平台，进行图像数据采集，机器人整体结构如图 2 所示。该机器人配备 1 个具有越障能力的履带式底盘，底盘长为 850 mm，宽为 400 mm，高为 270 mm，最快运行速度为 1 m/s，最大负载 100 kg，最大越障高度可达 100 mm。

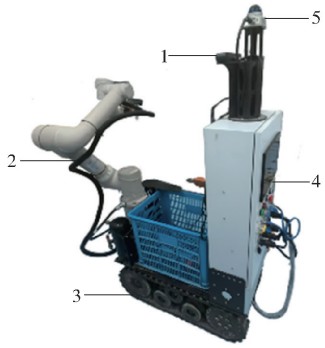


图 2　番茄采摘机器人试验平台

Fig. 2　Tomato picking robot experimental platform

1. 双目相机　2. 机械臂　3. 履带底盘　4. 电控柜　5. 激光雷达

电控柜安装机器人电气系统，包括主控制器、通信系统、供电系统以及逻辑控制电路。主控制器使用 Nvidia 的 Jetson TX2，在此平台上进行机器人软件开发，实现机器人导航与定位算法和底盘运动控制等功能。其配备 256 核 NVIDIA Pascal GPU，核心频率为 1 300 MHz，GPU 算力可达 1.33 TFLOPS，适合深度学习模型的部署和推理。TX2 的 CPU 与内存性能能满足相关需求。

该机器人还配备 1 个多线固态激光雷达以及

1 个深度相机用于感知周围环境。选用 StereoLabs 的 Zed2i 型深度相机,内置 IMU、气压计、磁力计以及偏光片,还可提供图像深度信息和点云信息。相机深度最远可达 20 m,并且在分辨率 1 080 像素下,视频输出帧率达 30 f/s。

1.2 语义分割

1.2.1 图像采集

图像数据采集于内蒙古呼和浩特市果蔬种植大棚,时间为 2024 年 12 月。选择作物有番茄、黄瓜、辣椒等,增强所采集数据通用性。数据采集期间,手动控制机器人保持匀速行驶在田垄之间,模拟真实工作状态。

选取网络上开源数据集,对自有数据集进行扩充,以拓展模型应用场景。Agroscapes 数据集[13]进行了精准像素级标注,包含从不同高度和角度对草莓、亚麻、玉米等 6 种不同作物拍摄图像。此外,BotanicGarden 数据集[16]是 1 个专为机器人在非结构化自然场景中导航而打造的高质量数据集,其场景复杂多样,道路特征明显。以上数据集图像主要取材于室外环境,而自制数据集则多为室内大棚环境,两者结合恰能弥补单一数据集在环境类型上的不足。在数据清洗和筛选后,最终得到由 4 568 幅图像所构成的农业场景道路数据集,数据集组成如表 1 所示。

表 1 数据集组成

Tab. 1 Composition of dataset

名称	国家	图像数量	来源
Agroscapes	美国	987	加州大学洛杉矶分校
BotanicGarden	中国	1 605	上海交通大学
自采集	中国	1 976	内蒙古果蔬大棚

1.2.2 数据集制作

在数据集制作过程中,考虑图像来源不同,需对所有图像进行尺寸归一化。将所有原始 RGB 图像统一缩放至 512 像素 ×512 像素,以保证数据集在空间维度上一致性,消除尺寸差异对模型训练稳定性与预测性能的影响。为了便于模型训练与评估,将数据集划分为训练集、验证集与测试集,划分比例为 70%、20% 和 10%。

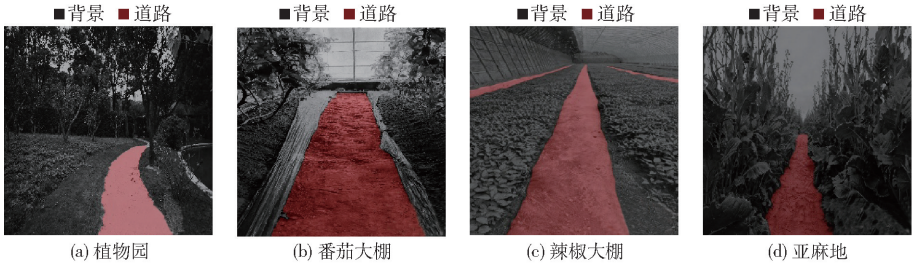
使用开源标注软件 LabelMe 对制作好的数据集进行像素级语义分割标注。由于本文重点在可行驶区域分割和路径提取,故将图像内容主要标注为道路和背景。所有图像数据标注完成后,编写程序脚本,将所有 JSON 格式的标注文件转换为单通道灰度掩码图 Mask。图 3 为数据集中不同场景条件下标签图像。

1.2.3 语义分割网络模型

在语义分割领域,众多网络模型在各种任务场景中都表现出高精度与高鲁棒性,Deeplab 系列是其中 1 种[17]。Deeplabv3 + 作为该系列中的改进版本,在 Deeplabv3 基础上引入解码器结构,进一步提升了对复杂目标轮廓和多尺度信息的分割能力,在非结构化的农业场景中表现不错[18]。

Deeplabv3 + 结构灵活,可搭配多种不同的特征提取网络作为 Backbone。常见的语义分割主干网络有很多,例如 ResNet[19]和 ResNeSt[20],都具有较大的参数量和较高的计算复杂度,对计算平台算力有一定的要求。MobileNetV3 是一种轻量化网络模型,利用深度可分离卷积减少参数量和计算量,有效提升了特征表达能力,并且内存占用小,在移动设备上实时性能更好[21-22]。故以 MobileNetV3 作为 Deeplabv3 + 主干特征提取网络,结合空洞空间金字塔池化模块(ASPP),多尺度地捕捉上下文信息[23],最终通过解码器不断上采样恢复空间分辨率,网络模型结构如图 4 所示。

从图 4 可以看出,输入图像进入编码器后,首先会被主干网络 MobileNetV3 提取特征,其中主要由普通卷积和一系列可分离卷积提取不同层次特征。深层特征语义信息丰富,传入 ASPP 模块进一步处理;中层特征分辨率和通道数适中,用于辅助分割头输入;低层特征保留了很多空间细节信息,直接传入解码器。辅助头利用中层特征信息进行辅助性分割预测,通过 1 个全卷积网络 FCN[24]内部的卷积块提取特征,最终用于计算辅助损失,以促进主网络特征学习。解码器部分利用 ASPP 输出的特征图和由 MobileNetV3 提供的低层特征逐步恢复空间分辨率,



(a) 植物园 (b) 番茄大棚 (c) 辣椒大棚 (d) 亚麻地

图 3 不同场景下标签图像

Fig. 3 Label images in different scenes

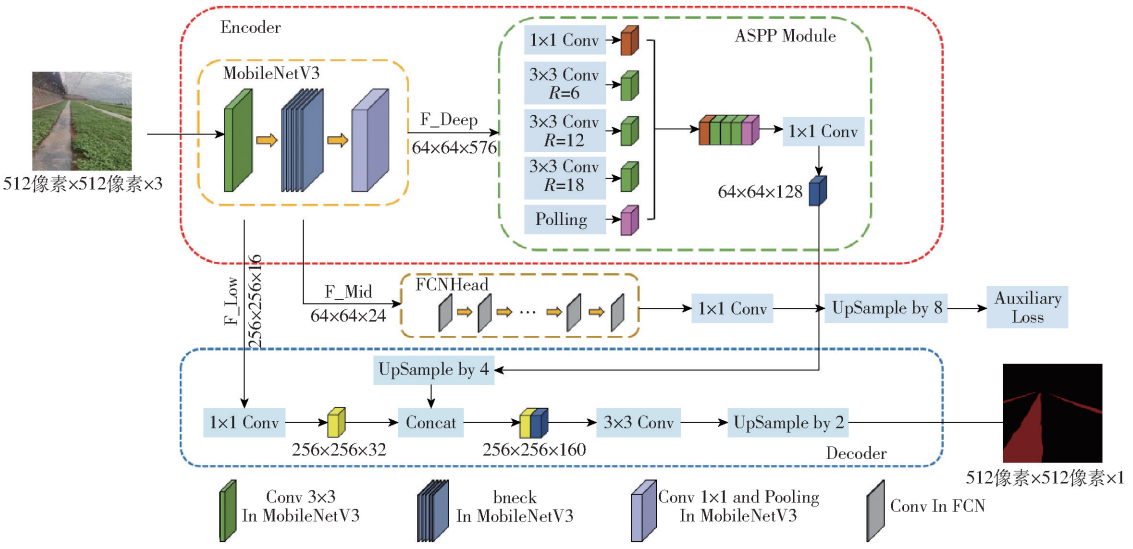


图 4　MobileNetV3 – Deeplabv3 + 模型结构

Fig. 4　MobileNetV3 – Deeplabv3 + model structure

最终得到分割结果。

1.2.4　语义分割损失函数

损失函数选择直接影响模型收敛速度和分割精度。在语义分割任务中，图像中每个像素都要进行分类，损失函数则可表示为模型预测结果和真实标签之间差异，故选用经典的交叉熵损失函数作为优化目标[25]。对于单个像素点的多类别分类问题，交叉熵损失函数可表示为

$$L = -\sum_{c=1}^{M} y_c \ln \hat{p}_c \tag{1}$$

式中　L——交叉熵损失函数

M——类别总数

y_c——像素点在类别 c 上真实标签

$\hat{p}_c$——像素点在类别 c 上预测概率

1.3　热图回归

1.3.1　路径提取原理

在非结构化农业环境中提取机器人路径导航线，通过语义分割将可行驶区域从背景中分离。直接在分割后的掩码图上利用传统几何方法或骨架化方法提取中心线，常常会因为分割边界存在噪声、不连续及形状不规则等问题，导致生成的导航路径出现抖动或不理想偏折。

为了解决上述难题，在分割结果基础上引入基于热图回归导航线提取方法。核心思想是避免对二维坐标的直接回归，而是预测目标在二维图像中出现位置概率分布图。现有研究已证明，在自动驾驶和机器人导航领域，热图回归可有效解决空间感知、定位和车道线检测等问题[26-29]。热图回归方法可将准确的位置坐标建模为不确定性的位置概率，在农业作业场景中极具优势。最终，通过对预测热图进行后处理，生成安全、连续且位于通道中央的导航路径，为机器人自主导航提供参考。

1.3.2　热图数据集构建

为实现热图预测任务，基于可行驶区域掩码图像，通过一系列图像处理步骤，生成连续强度分布的热图标签。实现步骤为：

(1)二值化处理图像，将掩码图中像素为 255 的像素置为 1，背景像素置为 0。

(2)对二值化掩码图进行欧氏距离变换，得到每个掩码区域像素到最近背景像素距离，生成距离分布 $D(x,y)$。

(3)使用高斯滤波器抑制距离变化过程中的噪声波动，进一步增强热值分布连续和平滑。

(4)根据高斯衰减函数计算各个像素处热值，计算式为

$$T(x,y) = \exp\left(-\frac{D^2(x,y)}{2\sigma^2}\right) \tag{2}$$

式中　T——像素热值

D——各像素距离分布

σ——高斯衰减系数

(5)对计算得到的热值进行线性归一化，将像素值重新映射回 0 ~ 255 范围，作为最终热值标签。

经过上述步骤，最终得到由原始 RGB 图像、分割后的二值化掩码图像以及引导性热图标签构成的数据集。

1.3.3　热图回归网络模型

为了实现对热图回归任务高精度预测以及多尺度细节保留，设计一种基于 Mobile – Unet 的轻量化回归网络，在经典 Unet 结构[30]基础上进行改进，采用编码器–解码器结构，模型整体结构如图 5 所示。

编码器使用 MobileNetV3 – Small 特征子网络，

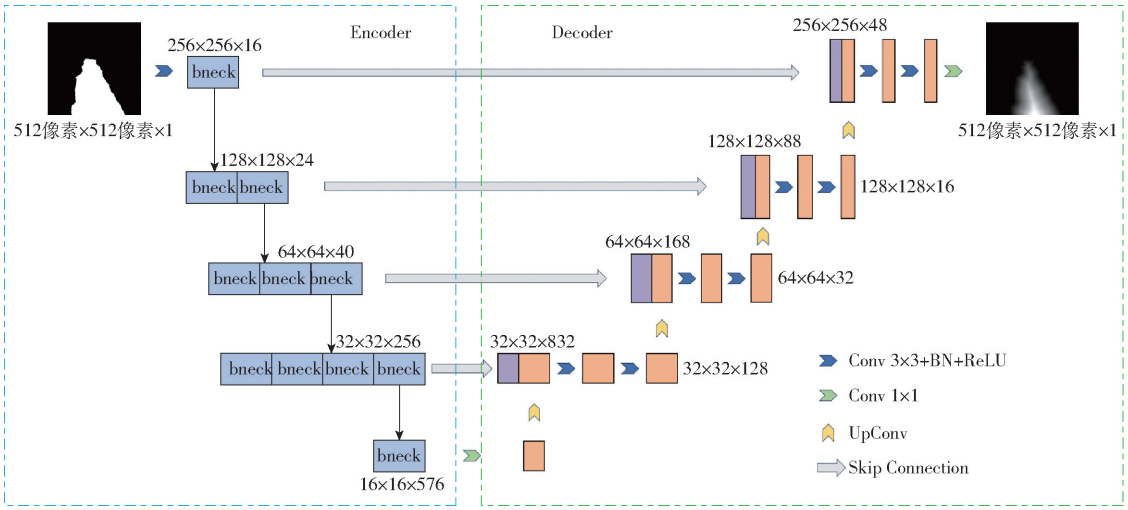


图 5 Mobile－Unet 模型结构

Fig. 5 Mobile－Unet model structure

实现高效特征提取和模型轻量化。为了增强模型对空间细节的保留能力,自定义跳跃连接机制,将编码器阶段不同分辨率下特征图进行保存,在解码器相应阶段与上采样特征图进行拼接。该操作能够有效恢复空间分辨率并且增强特征图细节表达能力[31]。

解码器部分采用 Unet 风格结构,通过上采样卷积模块和由 2 个 3×3 卷积和 ReLU 激活函数组成的卷积块,逐层恢复特征图像的空间分辨率。最终解码器使用 1 个 1×1 卷积将通道数映射为 1,输出尺寸为 512 像素×512 像素×1 热图。

1.3.4 热图回归损失函数

在一般回归任务中,均方误差损失(MSE loss)常作为训练损失函数。在热图回归问题中,则表示每个像素预测值和真实值之间的差距,计算式为

$$L_{\mathrm{MSE}}=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \tag{3}$$

式中 L_{MSE}——均方误差损失

y_i——真实值 $\hat{y}_i$——预测值

均方误差损失平等地衡量模型输出与标签之间每个像素数值差距,但热图回归问题中,受大量背景区域影响,往往无法突出热图中热值区域。其次 MSE 损失仅衡量像素级强度差异,缺乏对高阶空间结构和语义特征的一致性约束,容易导致预测热图热值区域边缘模糊、热点扩散等问题。

为了避免上述问题,本文提出基于加权均方误差(Weighted MSE)与感知损失(Perceptual loss)联合的组合损失函数,兼顾像素强度拟合精度与高层结构一致性。加权均方误差计算式为

$$L_{\mathrm{WMSE}}=\frac{1}{n}w_i\sum_{i=1}^{n}(y_i-\hat{y}_i)^2 \tag{4}$$

式中 L_{WMSE}——加权均方误差

w_i——权重

相比于普通均方误差,加权均方误差多 1 个权重,通过对热核区域赋予更高权重,背景区域赋予较低权重,使模型更加关注热核区域。感知损失基于预训练的 VGG16 网络[32]中间卷积层特征图,衡量预测热图与真实热图在特征空间中的差异,计算式为

$$L_{\mathrm{Perceptual}}=\frac{1}{CHW}\|\phi(\hat{Y})-\phi(Y)\|_2^2 \tag{5}$$

式中 $L_{\mathrm{Perceptual}}$——感知损失

$\phi(\cdot)$——VGG16 网络卷积特征提取函数

C——特征图通道数

H——特征图高度

W——特征图宽度

$\hat{Y}$——预测热图

Y——真实热图

感知损失能够有效增强预测热图结构特征,缓解热核边缘模糊问题。最终组合损失函数为

$$L_{\mathrm{total}}=\alpha L_{\mathrm{WMSE}}+\beta L_{\mathrm{Perceptual}} \tag{6}$$

式中 L_{total}——总损失

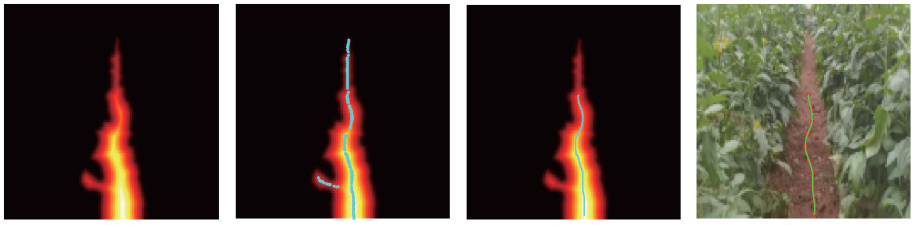
α、β——权重

不同权重表示模型训练时不同倾向。

1.4 导航路径提取算法

导航路径提取直接影响机器人移动轨迹,针对回归模型预测得到的热图,提出一种非极大值抑制结合热值引导贪心搜索的导航路径提取方法。与一般骨架结构化方法不同[33],基于越靠近主干热度扩散概率越大,热值越高的思想,极大利用回归图像本身形态和热值,对于岔路干扰与道路变形有较强的适应性。导航路径生成流程如图 6 所示,实现步骤为:

(1)热图归一化。将输入热图像素值归一化至[0,1]范围内,确保后续处理稳定。



(a) 原始热图　(b) NMS处理　(c) 贪心搜索　(d) 路径映射

图 6　导航路径提取流程

Fig. 6　Navigation path extraction process

(2)局部峰值检测。对热图进行非极大值抑制(NMS),提取所有局部最大值点作为后续路径提取的起始候选点,同时还能辅助观察热图响应分布。

(3)热值引导的路径搜索。将上一步非极大值抑制得到的局部热值最大点作为起点,进行热值引导路径搜索。采用 1 种方向约束下贪心搜索算法,在局部搜索窗口内贪婪选择热值最大、方向合理、热值连续下一路径点,并迭代前进,直到无法推进或到达指定长度。该算法无需全图搜索,能快速获得 1 条热值响应最强且路径方向稳定的主干路线。

(4)路径平滑。将上一步得到的各条路径进行合并得到整体路径,使用滤波器对路径点进行平滑,消除抖动,保持整体趋势。

(5)关键点提取与路径拟合。沿平滑后路径按一定长度均匀采样一组关键点,用于简化表示主路径形态特征。对提取到的关键点进行 B 样条插值拟合,形成高质量导航路径。

2　试验结果与分析

2.1　语义分割试验

2.1.1　试验设置

模型训练统一在 NVIDIA RTX 3090 GPU 上进行,训练批处理大小设为 16,总训练轮次为 50。优化器采用 Adam,初始学习率设置为 0.000 1,权重衰减为 0.000 1。并且所有输入数据都进行数据增强,以增强模型泛化能力。在每一轮训练后进行 1 次模型验证,对当前模型性能进行评估,最终以验证集上平均交并比(mIoU)最佳的模型作为最终模型。

2.1.2　模型结构性能试验

为了检验 MobileNetV3 - Deeplabv3 + 改进模型可行性,选取 ResNeSt50 和 ResNet50 作为 Deeplabv3 + 模型框架主干网络进行训练和预测评估,与本文模型进行效果对比。除此之外,还选取不同分割模型 Unet 和 PSPNet 进行训练,最终在同一测试集上进行性能评估,结果如表 2 所示。

表 2　不同模型性能

Tab. 2　Performance of different models

模型结构	交并比/%	Dice 系数/%	精确率/%	召回率/%	参数量	浮点运算量	内存占用量/MB	帧率/(f·s^{-1})
本文模型	92.00	95.83	96.11	95.55	2.10×10^{6}	9.03×10^{9}	631	6.92
ResNet50 - Deeplabv3 +	92.99	96.37	96.38	96.36	4.12×10^{7}	1.76×10^{10}	972	1.96
ResNeSt50 - Deeplabv3 +	93.59	96.69	96.34	97.04	4.31×10^{7}	1.78×10^{10}	859	1.76
PSPNet	89.17	94.28	96.99	91.71	4.66×10^{7}	1.79×10^{10}	775	2.96
Unet	90.46	94.99	97.01	93.05	2.90×10^{7}	2.03×10^{10}	941	2.80

从表 2 可以看出,搭配不同主干网络的 Deeplabv3 + 模型明显优于 PSPNet 和 Unet 模型。表明 ASPP 模块采样到丰富的特征信息,用于恢复空间结构和轮廓细节,对于分割任务起到了关键作用。在使用 Deeplabv3 + 模型中,以 ResNeSt50 为主干网络的模型取得了最佳 IoU(93.59%),ResNet50 和 MobileNetV3 效果稍逊一点,但 IoU 也达 92.99%。本文模型同样取得了极佳的分割效果,IoU 达 92.00%,Dice 系数为 95.83%,精确率和召回率分别达 96.11%与 95.55%。试验结果表明,该模型虽然比同为 Deeplabv3 + 系列的另外 2 种模型性能指标略低,但优于 Unet 和 PSPNet 分割模型。

尽管本文模型在绝对分割精度上不是最优,稍逊于同系列的其他模型,但 MobileNetV3 优势主要在于实现计算效率与模型性能之间的平衡。从表 2 可以看出,其参数量与浮点运算量均显著低于其他模型,为了进一步测试上述模型在嵌入式设备上的性能,以 NVIDIA Jetson TX2 作为试验平台,测试各个模型图像推理速度与内存消耗。本文模型处理速度最快,帧率达 6.92 f/s,相较于其他模型大幅领先,并且模型推理时内存占用也最低,仅有 631 MB。综合来看,本文模型结构表现最优,在分割效果不逊

于其他模型情况下,图像推理速度远超其他模型,为农业机器人导航实时性要求提供了保障。

2.2 热图回归试验

2.2.1 试验设置

优化器为 Adam,初始学习率为 0.000 3,采用 SetpLR 学习率衰减策略。并且为了提升模型鲁棒性,对输入图像和对应的标签热图进行同步数据增强,主要进行随机水平翻转、随机垂直翻转、随机旋转等。每次训练时批处理大小为 8,训练轮次为 50。

采用均方误差(MSE)、平均绝对误差(MAE)、峰值信噪比(PSNR)以及结构相似性指数(SSIM)作为评估指标。其中 MSE、MAE、PSNR 主要衡量预测结果与真实标签之间的像素差异,SSIM 则用于衡量整体结构分布相似程度。在每一轮训练完成后进行一次模型验证,MSE 作为主要指标。

2.2.2 模型结构与输入类型对比试验

为了验证改进方案可行性,对不同模型结构以及不同模型输入进行了对比。除了本文所提出的以 Mask 作为输入,Mobile - Unet 作为网络模型方案外,分别替换模型输入为 RGB 原图以及模型结构为原版 Unet 进行试验。试验过程中,损失函数统一选用 Weighted MSE 与 Perceptual loss 的组合损失函数,所占权重分别为 1 和 0.05。训练过程中不同方案下加权 MSE 损失和感知损失曲线如图 7、8 所示。

4 种方案均表现出了明显的损失下降趋势,在训练轮次 10 ~ 15 后逐渐收敛,并且最终的损失水平相当。可以看出,以 Unet 作为网络模型方案在训练开始时就表现出较低的损失值,并且收敛速度也更快。

由图 8 可知,以 Mask 作为输入的方案在训练初期时感知损失显著低于 RGB 输入的方案,并且最终收敛后水平也更好,说明使用 Mask 有助于恢复热图结构特征,也证明了语义分割必要性。

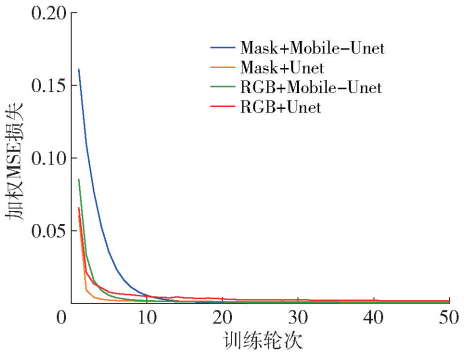


图 7 加权 MSE 损失曲线

Fig. 7 Weighted MSE loss curves

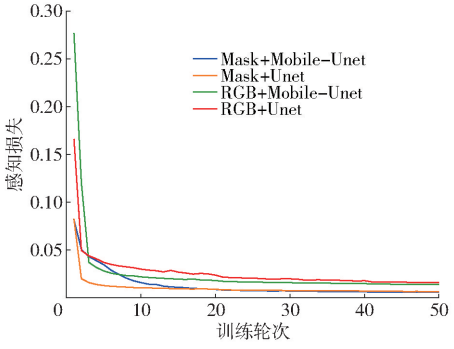


图 8 感知损失曲线

Fig. 8 Perceptual loss curves

为进一步验证各模型在实际热图回归任务中的泛化能力,在统一测试集上对所训练的多个模型进行性能评估,结果如表 3 所示。从表 3 可以看出,不同方案在测试集上的表现有着明显的差异。以 RGB 作为输入的方案在各项指标上均不如对应的 Mask 版本,尤其是峰值信噪比相差较多。Mobile - Unet 模型在 MSE 和 MAE 等误差指标上表现更优,而 Unet 模型表现出了更高的 SSIM。综合来看,本文 Mask + Mobile - Unet 方案效果最佳,MSE 为 8.91×10^{-5},MAE 为 0.004 63,同时 PSNR 指标达 44.00 dB。虽然 SSIM 相对其他模型稍低,但大于 0.9 已表明具有很强的结构一致性,完全满足使用需求。

表 3 不同方案在测试集上性能

Tab. 3 Performance of different solutions

输入类型	模型结构	MSE	MAE	PSNR/dB	SSIM	帧率/(f·s^{-1})
Mask	Mobile - Unet	8.91×10^{-5}	4.63×10^{-3}	44.00	0.921	50.73
RGB	Mobile - Unet	6.83×10^{-4}	7.49×10^{-3}	37.50	0.953	22.36
Mask	Unet	8.11×10^{-4}	8.21×10^{-3}	35.09	0.991	8.63
RGB	Unet	1.47×10^{-3}	1.10×10^{-2}	32.84	0.964	7.30

为验证不同模型在嵌入式平台中实际部署性能,以 NVIDIA Jetson TX2 作为测试环境,进行模型加载与推理测试。从表 3 可以看出,Mobile - Unet 实时性能远强于普通 Unet,而以 Mask 作为输入的模型处理速度最佳,达到 50.73 f/s。综合来看,以 Mask 作为输入,以 Mobile - Unet 作为网络模型方案同时兼顾了预测精度、结构还原度和实时推理速度,是针对该热图回归任务的最优方案。

2.2.3　损失函数性能试验

为了评估感知损失对模型回归性能的影响，设计不同权重性能试验。训练过程中，分别记录加权 MSE 损失和感知损失，观察的变化趋势。训练结束后，将所有模型在同一测试集下进行性能测试，记录测试集上性能指标，结果如表 4 所示。

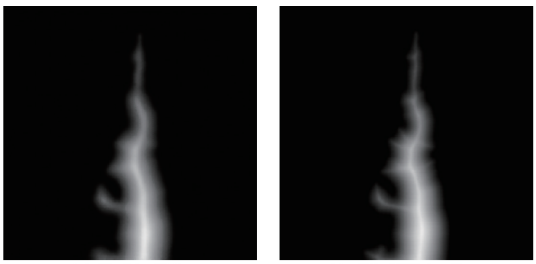
表 4　不同损失函数下性能

Tab. 4　Performance with different loss functions

感知损失权重	权重 MSE 损失	感知损失	MSE	MAE	PSNR/dB	SSIM
0	1.61×10^{-4}	无	5.07×10^{-5}	4.92×10^{-3}	43.88	0.888
0.05	1.73×10^{-4}	5.74×10^{-3}	8.91×10^{-5}	4.63×10^{-3}	44.00	0.921
0.10	1.47×10^{-4}	4.66×10^{-3}	2.92×10^{-5}	3.08×10^{-3}	46.90	0.959
0.20	1.66×10^{-4}	4.23×10^{-3}	3.77×10^{-5}	2.45×10^{-3}	47.48	0.985

由表 4 可知，当感知损失权重逐渐增大时，模型各项性能都得到了一定程度提升。权重 MSE 损失和感知损失都不断降低，PSNR 和 SSIM 同步提升，说明引入感知损失对热图回归模型的图像结构还原能力和像素精度有很大的提升。当感知损失权重提升到 0.1 时，各项指标达到最佳平衡，兼顾了像素误差和结构相似性。进一步提升权重至 0.2 后，虽然感知损失进一步降低，SSIM 也继续提升，但权重 MSE 损失和测试集 MSE 出现反弹。说明过大的感知损失权重使模型过度关注图像结构，但对像素误差优化被限制，最终导致整体回归精度下降。因此，综合训练表现和预测指标，将感知损失权重 0.1 作为最佳权重，应用于最终模型。

无感知损失与感知损失权重 0.1 预测结果如图 9 所示。从图 9 可以看出，在未加入感知损失时，预测热图出现明显的热核扩散和边界模糊现象，细节表达不足。增加感知损失后，热图轮廓边界逐渐清晰，热值分布均匀，热核也更加集中。预测热图表现与测试集定量指标分析结果一致。



(a) 无感知损失　　(b) 感知损失权重0.1

图 9　热图预测结果

Fig. 9　Heatmap prediction results

2.3　导航路径提取试验

为了定量地衡量路径提取效果，选择路径点覆盖热区占比和路径中心性 2 个指标进行评估。路径点覆盖热区占比用于衡量提取路径点与回归热图中高热区域重叠程度，其比例表明路径提取的有效性和精确度。试验时设置适当的热值阈值与路径点热值进行比较，高于阈值的路径点则可判断为在热区内。按照该方法对所有路径点进行计算，最终以平均值作为该条路径的覆盖热区占比。路径中心性用于衡量路径分布在热区内部的均匀性，其数值表示为路径上每一点到热区边界的最小距离与该行热区宽度比例。该值接近 50%，证明路径在热区内分布均匀，距离两侧边界较远，并且导航路径更加平滑和安全。试验时对路径上所有点的中心性取平均，作为整条导航线中心性指标。选取测试集中图像进行路径提取，然后计算各项评估指标，得到分布散点图如图 10 所示。

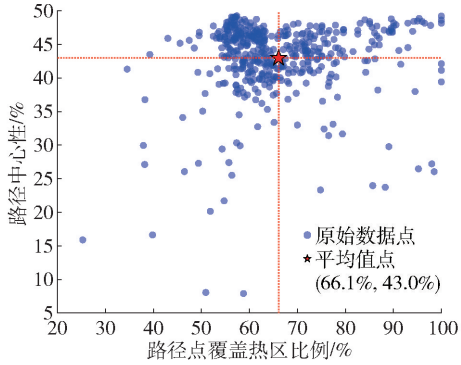


图 10　路径提取指标分布散点图

Fig. 10　Path extraction index distribution scatter plot

从图 10 可以看出，绝大多数图像的覆盖热区占比都超越 50%，并且个别图像达到 100%，表明路径提取效果精确，遵循了热值引导的基本原则，个别图像路径点甚至占据了整个高热值区域。路径中心性则大部分都高于 40%，证明了所提取的导航路径安全可靠，且基本位于道路中心。测试集中所有图像指标平均值分别为 66.1% 和 43.0%，反映所提取路径均有很高的质量。除此之外，选择不同场景下路径提取效果，如图 11 所示。

为了进一步验证该导航路径提取方法在真实机器人导航任务中适用性与安全性，在番茄温室大棚环境中进行了导航试验。番茄采摘机器人在大棚中行驶时，利用本文方法实时生成导航路径，得到路径点在相机像素坐标系下坐标 (u,v)，然后由相机内参反投影到相机坐标系下，得到路径点三维坐标 (X_c,Y_c,Z_c)，转换表达式为



(a) 油菜地 (b) 植物园 (c) 番茄大棚 (d) 亚麻地

图 11 导航路径提取效果

Fig. 11 Navigation path extraction

$$s\begin{bmatrix}u\\v\\1\end{bmatrix}=\boldsymbol{K}\begin{bmatrix}X_c\\Y_c\\Z_c\end{bmatrix} \tag{7}$$

式中 s——深度尺度因子

$\boldsymbol{K}$——相机内参矩阵

基于预先使用 Matlab 工具箱标定的相机与激光雷达之间的外参,可得到路径点在激光雷达坐标系下坐标(X_l,Y_l,Z_l),转换表达式为

$$\begin{bmatrix}X_l\\Y_l\\Z_l\end{bmatrix}=\boldsymbol{R}_{cl}\begin{bmatrix}X_c\\Y_c\\Z_c\end{bmatrix}+\boldsymbol{t}_{cl} \tag{8}$$

式中 $\boldsymbol{R}_{cl}$——外参旋转矩阵

$\boldsymbol{t}_{cl}$——外参平移向量

激光雷达与机器人底盘之间相对位姿固定,通过静态标定关系可将路径点转换到机器人坐标系下,实现路径与导航系统坐标统一。利用 Fast - LIO2 算法[34]实现了激光雷达里程计与地图构建,输出机器人在地图坐标系下位姿序列,得到机器人真实运行轨迹与对应的三维地图,如图 12 所示。

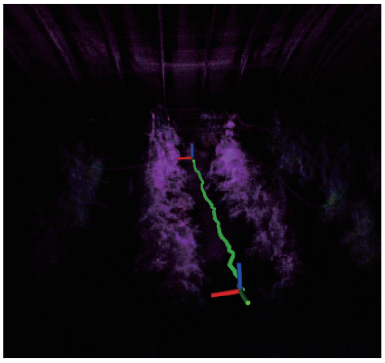


图 12 温室点云地图与导航轨迹

Fig. 12 Greenhouse point cloud map and navigation trajectory

由图 12 可知,两侧高密度的紫色点云对应温室中番茄作物,绿色曲线则是机器人导航过程的真实轨迹。机器人实际行驶轨迹始终保持在两侧番茄作物之间,没有发生偏离或进入作物区域的情况。为进一步量化分析导航效果,在相同温室通道内进行 5 次重复试验,结果如表 5 所示。

表 5 导航试验量化结果

Tab. 5 Navigation experiment results m

组号	最大横向偏差	平均横向偏差
1	0.05	0.03
2	0.08	0.05
3	0.09	0.05
4	0.09	0.06
5	0.06	0.04

从表 5 可看出,最大横向偏差不超过 0.09 m,各组试验平均横向偏差较小,且均不高于 0.06 m。试验结果表明,本文提出的导航路径提取方法能够在真实农业环境中生成安全可靠的导航路径,为机器人在农业场景中自主作业提供了有效的路径参考。

3 结论

(1)为实现机器人在农业场景下的实时导航,构建了结合轻量化语义分割和热图回归的路径提取模型。

(2)设计了以 MobileNetV3 为主干网络的轻量化 Deeplabv3 + 模型,实现农业非结构化场景可行驶区域分割。试验结果表明,该模型 IoU 可达 92%,推理速度达 6.92 f/s,有效平衡了精度与效率,高度适配农业机器人的实时部署需求。

(3)构建了基于 MobileNetV3 与 Unet 的热图回归模型,并设计加权 MSE 与感知损失复合函数以抑制热核模糊。试验结果表明,该模型推理帧率高达 50.73 f/s;复合损失的引入使预测热图 PSNR 达 47.48 dB,SSIM 达 0.985,显著提升了预测热图质量与精度。

(4)通过热图后处理算法实现了导航路径精确提取。试验结果表明,路径平均覆盖热区占比达 66.1%,中心性达 43.0%。真实农业场景实机测试表明,导航平均偏差不高于 0.06 m,全程无作物干涉。该方法满足农业机器人安全、可靠的自主作业需求。

参 考 文 献

[1] 赵春江, 范贝贝, 李瑾, 等. 农业机器人技术进展、挑战与趋势[J]. 智慧农业(中英文), 2023, 5(4): 1 - 15.
Zhao Chunjiang, Fan Beibei, Li Jin, et al. Agricultural robots: technology pragress, chablenges and trends[J]. Smart Agriculture, 2023,5(4):1 - 15. (in Chinese)

[2] 刘成良, 贡亮, 苑进, 等. 农业机器人关键技术研究现状与发展趋势[J]. 农业机械学报, 2022, 53(7): 1 - 22, 55.
Liu Chengliang, Gong Liang, Yuan Jin, et al. Current status and development trends of agricultural robots[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(7): 1 - 22, 55. (in Chinese)

[3] McCool C, Beattie J, Firn J, et al. Efficacy of mechanical weeding tools: a study into alternative weed management strategies enabled by robotics[J]. IEEE Robotics and Automation Letters, 2018, 3(2): 1184 - 1190.

[4] 郑伟员, 梁子安, 周俊, 等. 水田平地机 GNSS 高程数据 EMD 与 S - G 联合降噪研究[J]. 华南农业大学学报, 2024, 45(1): 80 - 87.
Zheng Weiyuan, Liang Zian, Zhou Jun, et al. Research on combined noise reduction of GNSS elevation data of paddy field grader with EMD and SG filter[J]. Journal of South China Agricultural University, 2024, 45(1): 80 - 87. (in Chinese)

[5] Mousazadeh H. A technical review on navigation systems of agricultural autonomous off-road vehicles[J]. Journal of Terramechanics, 2013, 50(3): 211 - 232.

[6] Abanay A, Masmoudi L, El Ansari M, et al. LiDAR-based autonomous navigation method for an agricultural mobile robot in strawberry greenhouse: AgriEco robot[J]. AIMS Electron. Electr. Eng., 2022, 6(3): 317 - 328.

[7] Zhang S, Guo C, Gao Z, et al. Research on 2D laser automatic navigation control for standardized orchard[J]. Applied Sciences, 2020, 10(8): 2763.

[8] 黄炜昭, 谢欢欢, 辛拓, 等. 基于激光雷达与 IMU 融合的变电站智能巡检机器人视觉导航方法[J]. 机械设计与制造工程, 2024, 53(12): 42 - 46.

[9] Choudhary A, Kobayashi Y, Arjonilla F J, et al. Evaluation of mapping and path planning for non-holonomic mobile robot navigation in narrow pathway for agricultural application[C]//2021 IEEE/SICE International Symposium on System Integration (SⅡ). IEEE, 2021: 17 - 22.

[10] Xiong J, Liang J, Zhuang Y, et al. Real-time localization and 3D semantic map reconstruction for unstructured citrus orchards [J]. Computers and Electronics in Agriculture, 2023, 213: 108217.

[11] Li Y, Wang X, Liu D. 3D autonomous navigation line extraction for field roads based on binocular vision[J]. Journal of Sensors, 2019, 2019(1): 6832109.

[12] Yang Z, Ouyang L, Zhang Z, et al. Visual navigation path extraction of orchard hard pavement based on scanning method and neural network[J]. Computers and Electronics in Agriculture, 2022, 197: 106964.

[13] Panda S K, Lee Y, Jawed M K. Agronav: autonomous navigation framework for agricultural robots and vehicles using semantic segmentation and semantic line detection[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 6272 - 6281.

[14] He B, Liu G, Ji Y, et al. Auto recognition of navigation path for harvest robot based on machine vision[C]// International Conference on Computer and Computing Technologies in Agriculture, 2010: 138 - 148.

[15] Liu H, Zeng X, Shen Y, et al. A single-stage navigation path extraction network for agricultural robots in orchards[J]. Computers and Electronics in Agriculture, 2025, 229: 109687.

[16] Liu Y, Fu Y, Qin M, et al. Botanicgarden: a high-quality dataset for robot navigation in unstructured natural environments [J]. IEEE Robotics and Automation Letters, 2024, 9(3): 2798 - 2805.

[17] Chen L C, Papandreou G, Kokkinos I, et al. Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(4): 834 - 848.

[18] Chen L C, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018: 801 - 818.

[19] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770 - 778.

[20] Zhang H, Wu C, Zhang Z, et al. Resnest: split-attention networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 2736 - 2746.

[21] Howard A, Sandler M, Chu G, et al. Searching for Mobilenetv3[C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 1314 - 1324.

[22] 郝夏楠, 庞亚军, 麻月欣, 等. 基于 MobileNetV3 网络的轻量级激光雷达点云图像语义分割[J]. 激光杂志, 2024, 45(12): 174 - 180.
Hao Xia'nan, Pang Yajun, Ma Yuexin, et al. Semantie segmentation of lightweight LiDAR point cloudimages based on MobileNetV3 network[J]. Laser Journal, 2024, 45(12): 174 - 180. (in Chinese)

[23] Zeng W, He M. Rice disease segmentation method based on CBAM - CARAFE - DeepLabv3 + [J]. Crop Protection, 2024, 180: 106665.

[24] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4): 640 - 651.
[25] Cao Z, Liao T, Song W, et al. Detecting the shuttlecock for a badminton robot: a YOLO based approach[J]. Expert Systems with Applications, 2021, 164: 113833.
[26] Liu Y B, Zeng M, Meng Q H. Unstructured road vanishing point detection using convolutional neural networks and heatmap regression[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 70: 1 - 8.
[27] Mendez O, Hadfield S, Bowden R. Markov localisation using heatmap regression and deep convolutional odometry[C]//2021 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2021: 9638 - 9644.
[28] Ates H F, Siddique A, Gunturk B. HMRN: heat map regression network to detect and track small objects in wide-area motion imagery[J]. Signal, Image and Video Processing, 2023, 17(1): 39 - 45.
[29] Wu Z, Zhang J, Zhang L, et al. Bi - HRNet: a road extraction framework from satellite imagery based on node heatmap and bidirectional connectivity[J]. Remote Sensing, 2022, 14(7): 1732.
[30] Ronneberger O, Fischer P, Brox T. U - net: convolutional networks for biomedical image segmentation[C]//International Conference on Medical image Computing and Computer-Assisted Intervention. Cham: Springer International Publishing, 2015: 234 - 241.
[31] 杨断利, 沈洪硕, 陈辉, 等. 基于 YOLO v7 和改进 U - Net 模型的鸡冠肉垂提取与面积计算方法[J]. 农业机械学报, 2025, 56(4): 415 - 426.
Yang Duanli, Shen Hongshuo, Chen Hui, et al. Extraction and area calculation of chicken comb and wattle based on YOLO v7 and optimized U - Net network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(4): 415 - 426. (in Chinese)
[32] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[C]//Proceedings of the 3rd International Conference on Learning Representations, 2015: 1 - 14.
[33] Chang J, Dong N, Li D, et al. Skeleton extraction and greedy-algorithm-based path planning and its application in UAV trajectory tracking[J]. IEEE Transactions on Aerospace and Electronic Systems, 2022, 58(6): 4953 - 4964.
[34] Xu W, Cai Y, He D, et al. Fast - lio2: fast direct LiDAR-inertial odometry[J]. IEEE Transactions on Robotics, 2022, 38(4): 2053 - 2073.

(上接第 246 页)

[21] Li H, Nam S, Lu Z, et al. Biotactip: a soft biomimetic optical tactile sensor for efficient 3D contact localization and 3D force estimation[J]. IEEE Robotics and Automation Letters, 2024, 9(6): 5314 - 5321.
[22] Rayamane P, Herbert P, Munguia-Galeano F, et al. PnuTac: a vision-based pneumatic tactile sensor for slip detection and object classification[C]//2023 29th International Conference on Mechatronics and Machine Vision in Practice (M2VIP). IEEE, 2023: 1 - 6.
[23] Ma C, Ying Y, Xie L. Development of a visuo-tactile sensor for non-destructive peach firmness and contact force measurement suitable for robotic arm applications[J]. Food Chemistry, 2025, 467: 142282.
[24] Jiang R, Fu L, Li Y, et al. High resolution, large area vision-based tactile sensing based on a novel piezoluminescent skin [J]. IEEE Transactions on Robotics, 2025, 41: 2477 - 2494.
[25] Zhang S, Chen Z, Gao Y, et al. Hardware technology of vision-based tactile sensor: a review[J]. IEEE Sensors Journal, 2022, 22(22): 21410 - 21427.
[26] Zhang Y, Kan Z, Yang Y, et al. Effective estimation of contact force and torque for vision-based tactile sensors with helmholtz-hodge decomposition[J]. IEEE Robotics and Automation Letters, 2019, 4(4): 4094 - 4101.
[27] 吴培良,李瑶,牛明月,等. 一种视/触觉融合的柔性物体抓取力估计方法[J]. 机器人, 2024,46(5):544 - 553.
Wu Peiliang, Li Yao, Niu Mingyue, et al. A visual-tactile fusion method for estimating the grasping force on flexible objects [J]. Robot, 2024,46(5):544 - 553. (in Chinese)
[28] Raissi M, Perdikaris P, Karniadakis G E. Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations[J]. Journal of Computational Physics, 2019, 378: 686 - 707.
[29] 张超凡,乔一铭,曹露,等. 基于神经形态的触觉滑动感知方法[J]. 浙江大学学报(工学版), 2023,57(4):683 - 692.
Zhang Chaofan, Qiao Yiming, Cao Lu, et al. Tactile slip detection method based on neuromorphic modeling[J]. Journal of Zhejiang University (Engineering Science), 2023,57(4):683 - 692. (in Chinese)
[30] Sui R, Zhang L, Li T, et al. Incipient slip detection method for soft objects with vision-based tactile sensor [J]. Measurement, 2022, 203: 111906.