doi:10.6041/j.issn.1000-1298.2024.02.002

基于轻量化 U－Net 网络的果园垄间路径识别方法

侯文慧　周传起　程　炎　王玉伟　刘　路　秦　宽
（安徽农业大学工学院，合肥 230036）

摘要：针对目前果园垄间导航路径识别方法存在准确性与实时性难以同时兼顾、泛化能力弱等问题，本文在 U－Net 模型的基础上进行优化，采用 MobileNet－v3 Large 作为 U－Net 的主干特征提取网络，并在跳跃连接处引入坐标注意力机制（Coordinate attention，CA），构建轻量化路径识别模型。以该模型分割的垄间可行驶区域为基础，利用最小二乘法重塑可行驶区域边缘点，并进一步提取垄间导航线。首先采用数据增强的草莓垄间数据集进行模型训练，并进一步迁移到葡萄和蓝莓数据集上进行权重微调，以提高模型适应能力。最后在相应的验证集上进行导航路径识别，并通过可视化对比不同模型识别结果，以验证模型准确性。试验结果表明，网络模型在草莓、蓝莓和葡萄果园垄间路径识别的平均交并比分别为 98.06%、97.36% 和 98.50%，平均像素准确度分别达到 99.13%、98.75% 和 99.29%。模型处理 RGB 图像分割可行驶区域的理论推理速度可达 19.23 f/s，满足导航实时性和准确性的要求。

关键词：垄间导航；路径识别；语义分割；图像处理；迁移学习；U－Net
中图分类号：TP242.6　　**文献标识码**：A　　**文章编号**：1000-1298(2024)02-0016-12　　**OSID**：

Path Recognition Method of Orchard Ridges Based on Lightweight U－Net

HOU Wenhui　ZHOU Chuanqi　CHENG Yan　WANG Yuwei　LIU Lu　QIN Kuan
(*School of Engineering*, *Anhui Agricultural University*, *Hefei* 230036, *China*)

Abstract: In response to the issues of accuracy and speed being difficult to balance simultaneously, as well as the weak generalization ability in the current navigation path recognition methods of fruit ridges, an optimization approach was proposed based on the U－Net model. The optimization involved integrating MobileNet－v3 Large as the backbone feature extraction network for U－Net and introducing coordinate attention at the skip connections to construct a lightweight path recognition model. Based on the drivable area segmented by this model in the inter-ridge, the edge points of the area were reshaped by using the least squares method, and further the inter-ridge navigation lines were extracted. Firstly, the model was trained on the augmented strawberry interrow dataset, and further migrated to the grape and blueberry datasets for weight fine-tuning to improve the model's adaptability. Finally, the navigation path was identified on the corresponding verification set, and the recognition results of different models were compared visually to verify the accuracy of the model. Experimental results demonstrated that the model achieved an average intersection over union of 98.06%, 97.36%, and 98.50% for strawberry, blueberry, and grape interridge navigation path segmentation accuracy respectively, and the average pixel accuracy reached 99.13%, 98.75%, and 99.29%. The theoretical reasoning speed of the model for segmenting of RGB images was up to 19.23 f/s, and the average time from image input to completed path extraction was 0.211 s, meeting the requirements of real-time navigation and accuracy. A method of path extraction based on semantic segmentation was proposed, which provided a general method for the navigation of agricultural machinery equipment in interridge operation.

Key words: ridge navigation; path recognition; semantic segmentation; image processing; transfer learning; U－Net

收稿日期：2023－09－21　修回日期：2023－12－20
基金项目：安徽省高校科研计划项目（2022AH050872）和电气传动与控制安徽省重点实验室开放基金项目（DQKJ202203）
作者简介：侯文慧（1985—），女，讲师，博士，主要从事智慧农业研究，E-mail：hwh303@ahau.edu.cn
通信作者：秦宽（1989—），男，副教授，博士，主要从事智能农机装备研究，E-mail：qinkuan@ahau.edu.cn

0　引言

垄栽是一种在种植床或垄上种植果实作物的栽培方法。面对垄间作业对智能农机的需求，解决智能农机在垄间作业时所面临的信息感知和路径导航问题具有重要意义[1]。目前智能农机装备导航主要方法有全球定位系统(GPS)导航、激光雷达导航和机器视觉导航[2-3]。其中，基于 GPS 的自主导航系统在农田环境下提供的定位信息并不完全可靠[4]，并且这种定位系统价格昂贵，制约了其在农业领域中的应用；同时，激光雷达在复杂田间环境中极易受到障碍信息的干扰，如杂草、延伸出来的作物枝叶等[5]；随着传感器技术和机器视觉算法的快速发展，基于视觉传感器的导航系统因其价格低廉、定位可靠，成为路径导航研究的热点[6-8]。但是农业场景复杂多变，垄间路径导航不仅需要算法具有良好的鲁棒性，更需要其根据实时路况调整参数，对算法的泛化能力和实时处理性能提出了较高要求[9]。目前基于传统图像处理的机器视觉技术需要依赖人工进行特征设计和选择，容易受到田间光照、路况等因素干扰，因此此类方法往往只能在特定环境下有效，抗干扰能力差，泛用性差[10-12]。

深度学习可以自动挖掘图像特征，不依赖人工选择，因此避免了人为干扰因素，提升了算法智能性和鲁棒性，广泛应用于农业场景目标识别、分割和分类[13]。近年来 SALEEM 等[14]分析了深度学习在植物病害检测和分类、杂草鉴别分割、果实计数等农业场景应用中，相比机器学习所展现的优势，以及在农业实际应用中存在的问题和发展方向。针对农业场景下非结构化道路识别问题，深度学习为视觉导航提供了新思路。基于卷积神经网络的语义分割技术可实现像素级的分类，以区分植物和可行驶区域[15-17]，因此可以为视觉导航提供可靠指导[18-20]。目前已有众多学者就非结构化道路语义分割技术展开了一系列研究[21-22]。韩振浩等[23]利用 U－Net 网络和多段三次 B 样条曲线拟合的导航路径识别方法，在弱光、普通光以及强光照条件下进行试验，得到分割交并比分别为 89.52%、86.45%、86.16%，路径平均像素误差为 9.5 像素；饶秀勤等[24]保留 U－Net 模型特征跳跃连接的优势，提出了一种 Fast－U－Net 模型，分别对棉花、玉米和甘蔗导航场景进行试验，得到平均交并比分别为 79.10%、77.10% 和 94.00%；李云伍等[25]面向丘陵山区田间道路，使用 FCN 实现了场景语义分割，平均交并比达到 73.20%；BADRINARAYANAN 等[26]提出了一种名为 SegNet 的深度全卷积神经网络架构，用于语义像素分割，进一步提高了自动驾驶的图像语义分割精度；KIM 等[27]提出了一种基于卷积神经网络图像块分类的透视路径检测方法，对可行驶区域进行分割，分割后最大交并比为 81.0%。

虽然基于语义分割的田垄间道路识别和导航方法取得了一些成果，但仍存在以下问题：①深度学习语义分割算法的高精度是通过大量卷积计算得来，占用的算法资源较大，计算时间较长。因此，如何平衡检测精度与速度是一个关键问题。②当前大多研究关注于单一作物垄间场景的路径分割算法，其缺乏通用性，需要考虑一种能够适应于多个场景的泛用算法。③当前的导航线生成方式还是以单一图像检测生成为主，缺乏一种实时的、连贯的导航线检测流程。

针对以上问题，本文提出一种基于轻量化 U－Net 的果园垄间图像实时分割网络模型，在草莓垄间导航路径识别的基础上进行葡萄、蓝莓垄间小样本的迁移学习，以验证模型的分割准确性和泛用性。基于此分割结果提取导航线，并利用垄间道路图像或视频检验该方法的导航精度和检测速度，以期满足田垄间路径导航需求。

1　图像采集与数据集制作

1.1　图像采集

从垄间场景出发，根据作物生长高度、姿态以及果园种植方式的差异性，选择 3 种果类作物代表常见的果实垄间导航场景。植株偏低矮且种植密集的代表果类作物为草莓；植株中等高度且种植密度较稀疏的代表果类作物为蓝莓；植株较高且种植密度稀疏的果类作物代表为葡萄。为了提高方法对光照的鲁棒性，分别在强光、正常光和弱光条件下采集果实垄间数据。数据采集地点位于安徽省合肥市包河区的开放农园，其中草莓数据于 2023 年 4 月采集，葡萄和蓝莓于 2023 年 7 月采集。手动遥控四轮差速智能小车采集平台(图 1)以速度 1 m/s 在垄间行走，模拟机具正常作业状态。行走间，利用其上搭载的 Intel RealSense Depth Camera D455 型摄像头进行数据采集，视频数据存入车载计算机(工控机)内。摄像头通过型材和云台固定架横向安装于底盘上盖，面向车头所在方向，镜头距底盘最前部水平距离 0.03 m 处，距地面高度 0.75 m，向下倾斜 25°左右，如图 1 所示。摄像头采集图像分辨率为 1 280 像素×720 像素，采集帧率为 30 f/s。考虑到网络计算需求，先将采集到的图像进行图像尺寸的归一化裁剪，将像素尺寸规整为 512 像素×512 像素，如图 2 所示。

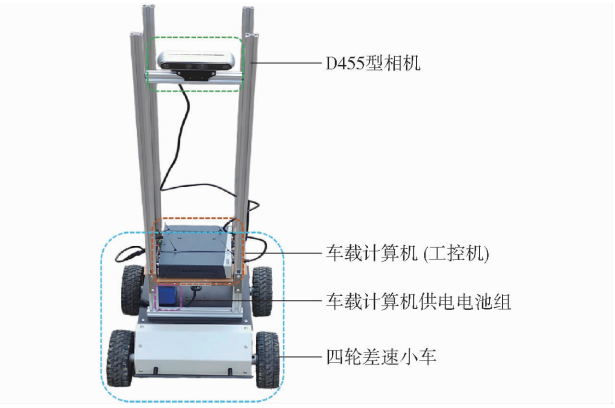

图 1 智能小车采集设备

Fig. 1 Intelligent car acquisition equipment

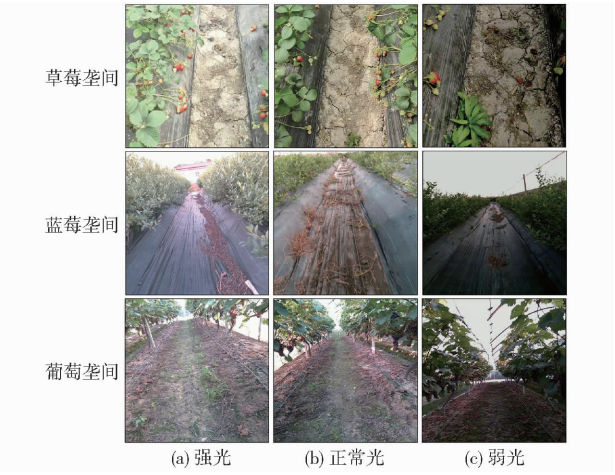

图 2 不同环境下的垄间图像

Fig. 2 Image data between ridges in different environments

1.2 数据增强

为了扩大样本量,并且提高训练样本的复杂程度,以提高模型在复杂环境下的准确性和泛用性,采用旋转 90°、镜像翻转、调整图像亮度等方法进行数据增强。其中,旋转 90°、镜像翻转是为了提高数据集中各种垄间道路导航方向的复杂性,亮度的随机变换是为了模拟实际环境中不均匀、不同天气的自然光照环境,亮度调整范围小于 30%。数据增强效果如图 3 所示。

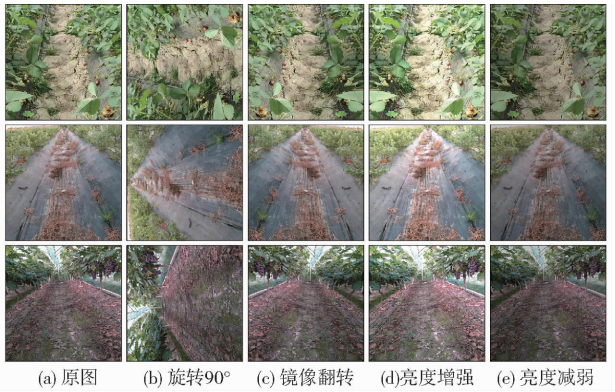

图 3 数据增强

Fig. 3 Data enhancement

1.3 数据集制作

经过归一化裁剪和数据增强操作后,随机选用草莓垄间图像 1 835 幅、蓝莓垄间图像 250 幅和葡萄垄间图像 250 幅,利用 Labelme 软件对垄间可行驶区域像素进行标注,得到可行驶区域和背景二分类标签,标签制作结果如图 4 所示。随后以比例 9∶1将数据集划分为训练集与验证集,以构成草莓、蓝莓和葡萄垄间数据集,如表 1 所示。

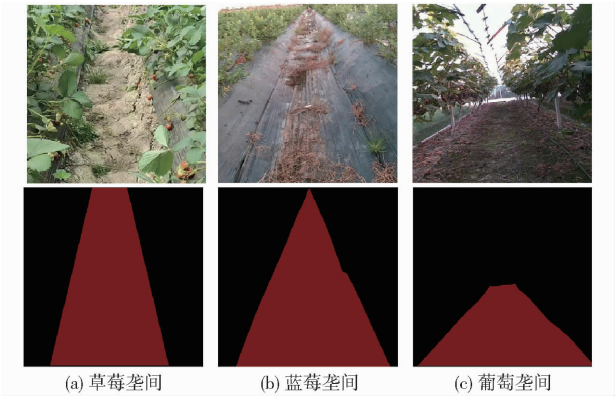

图 4 标签制作结果

Fig. 4 Label creation results

表 1 数据集结构设置

Tab. 1 Data set structure settings

数据集种类	图像数量	原始图尺寸/(像素×像素)	归一化后尺寸/(像素×像素)	训练集数量	验证集数量
草莓	1 835	1 280×720	512×512	1 650	185
蓝莓	250	1 280×720	512×512	225	25
葡萄	250	1 280×720	512×512	225	25

2 基于语义分割的果园垄间路径提取方法

利用智能小车采集平台拍摄垄间视频,获取垄间图像;对采集图像进行裁剪和数据增强处理,提高数据多样性,以进一步优化数据集质量;利用 Labelme 软件进行数据集的标签制作。

为了提高网络实时性,先将 U-Net 的主干网络轻量化,采用轻量级的网络模型 MobileNet-v3 Large 网络作为 U-Net 的主干特征提取网络;同时,为了解决轻量化的特征提取网络检测精度降低的问题,在跳跃连接处引入 Coordinate attention(CA)注意力机制;将改进的 U-Net 分割网络与以 ResNet50 为主干网络的 U-Net 网络(ResNet50-Unet)、以 VGG16 为主干网络的 U-Net 网络(VGG16-Unet)、Deeplabv3+网络和 PSPnet 网络进行分割效果对比,评估改进的 U-Net 分割网络性能;利用路径提取模块对分割出的可行驶区域图像进行边缘检测和路径提取,并以图像和视频验证路径提取方法的准确性。该方法总体流程如图 5 所示。

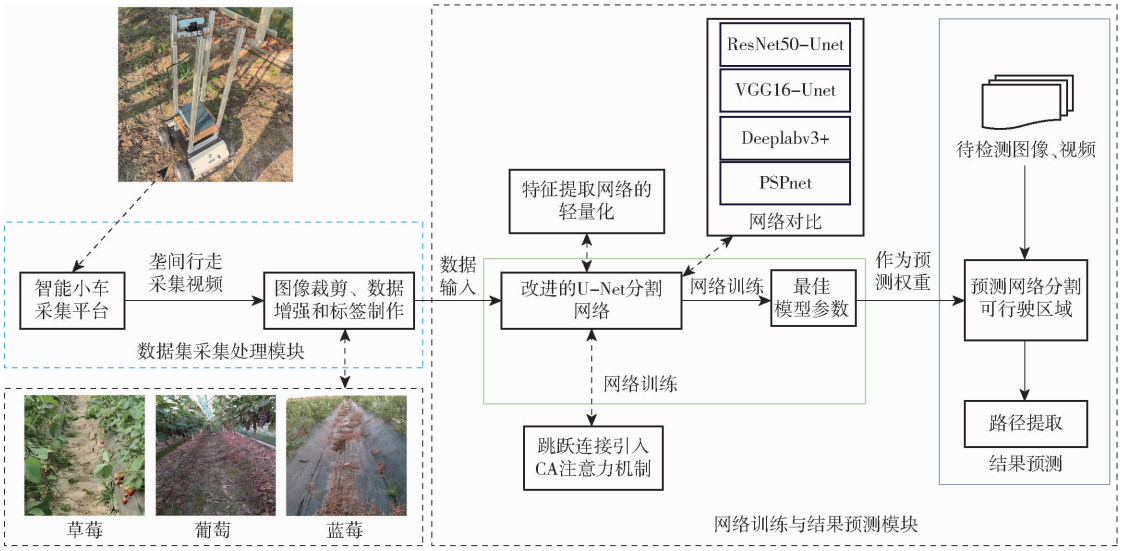

图 5　基于轻量化 U－Net 网络的垄间路径提取总体流程图

Fig. 5　Overall flow chart of crop interrow path extraction based on lightweight U－Net segmentation network

2.1　网络改进

2.1.1　改进的 U－Net 语义分割网络

U－Net[28]是一个使用全卷积的语义分割约束网络，其结构完全对称分布，其中浅高层分辨率特征图用来解决像素的定位问题，深低层分辨率特征图用来解决像素的分类问题。并且网络还采用跳跃连接的方式将特征图进行维度拼接，极大地保留了更多的位置和特征信息，是一种适合小样本任务的语义分割网络。U－Net 主干特征提取网络常用 VGG[29]和 ResNet[30]网络，极大地提高了 U－Net 网络的学习能力。然而，U－Net 较深的卷积层赋予了网络强大学习能力的同时，也降低了网络运行效率。

为了平衡智能农机对路径识别方法精准性与实时性的需求，对 U－Net 网络结构进行优化：

(1)为了满足实时性需求，在 U－Net 网络的编码器部分舍弃卷积层数较深的 VGG 和 ResNet 网络，采用一种轻量化的卷积网络——MobileNet－v3 Large 作为主干特征提取网络，并将 MobileNet－v3 Large 的第 1、3、6、12、15 个 bneck 块的输出作为 5 个初步有效特征层。

(2)为了避免主干特征提取网络轻量化导致模型精度下降，同时为了提高网络对垄间道路可能存在的杂草、枝叶遮挡等因素的抗干扰能力，在 U－Net 跳跃连接处引入 Coordinate attention(CA)注意力机制，加强网络对目标的关注。

改进后的 U－Net 网络结构如图 6 所示。

2.1.2　主干特征提取网络结构

MobileNet－v3[31]是 Google 团队在 2019 年提出的一种轻量级的卷积神经网络架构，其 bneck 模块结构如图 7 所示。

MobileNet－v3 网络保留了 MobileNet－v1 的深度可分离卷积以及 MobileNet－v2 的线性瓶颈的逆残差结构，引入了一种轻量的注意力模块 SE(Squeeze and excitation)，并且用 H－Swish 激活函数代替 Swish 激活函数，在保证准确性的同时，减少了计算时间，提高了网络计算速度，激活函数计算公式为

$$\text{H-Swish}[x]=x\frac{\text{ReLU 6}(x+3)}{6} \tag{1}$$

其中　$\text{ReLU 6}(x+3)=\min(6,\max(0,x+3))$　(2)

根据网络深度、宽度和计算量等方面，MobileNet－v3 被分为 MobileNet－v3 Large 和 MobileNet－v3 Small 两个版本。尽管相对于 Small 版本，Large 版本在检测速度方面稍逊一筹，但在面对杂草、枝叶遮挡等干扰因素的垄间道路检测任务中，Large 版本具有更高的准确性，其检测速度也足以应对垄间道路检测。因此，本文选择 MobileNet－v3 Large 作为 U－Net 结构的主干特征提取网络。表 2 为 MobileNet－v3 Large 网络的配置，其中第 1、3、6、12、15 个 bneck 结构块的输出被用作 5 个初步有效的特征层，传递给后续的加强特征提取网络。

2.1.3　CA 注意力机制

CA 是一种轻量化的全面关注特征层的空间信息和通道信息的新型注意力机制[32]。它的提出者认为之前的通道注意力机制，例如 Squeeze-and-excitation attention(SE)机制等，虽对于提升网络模型的性能有一定的作用，但它们通常会忽略特征层的位置信息，而位置信息对于模型生成空间选择性注意力图很重要。因此，CA 注意力机制将位置信息

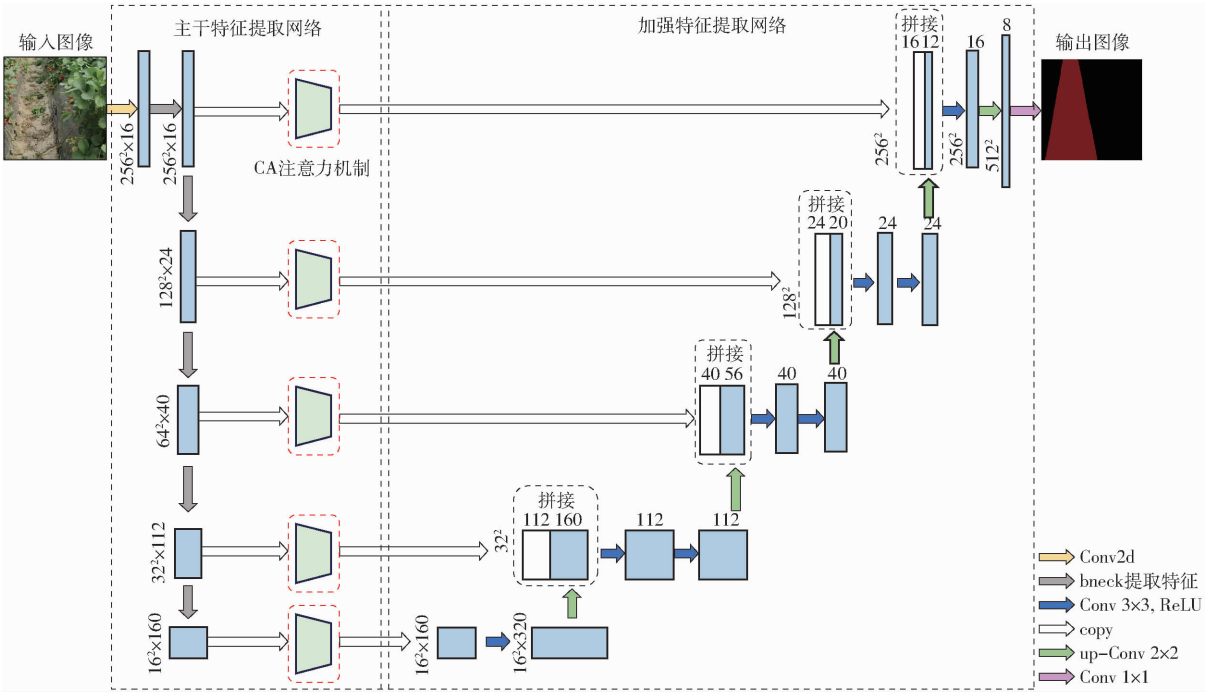

图 6 改进的 U－Net 网络结构简图

Fig. 6 Improved U－Net network structure diagram

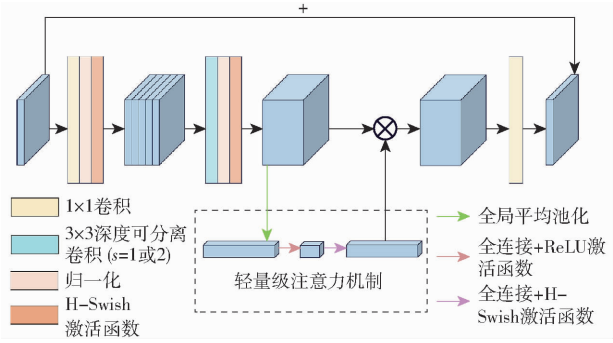

图 7 MobileNet－v3 网络的 bneck 结构简图

Fig. 7 Bneck structure diagram of MobileNet－v3 network

嵌入到通道注意力中，其工作原理如图 8 所示。

(1)将输入进来的特征层($x_c(i,j)$)分别在宽度和高度两个方向上进行全局平均池化，即

$$I_c^h = \frac{1}{W}\sum_{0\leqslant i<W} x_c(h,i) \tag{3}$$

$$I_c^w = \frac{1}{H}\sum_{0\leqslant j<H} x_c(j,w) \tag{4}$$

式中 I_c^h、I_c^w——高度、宽度上的特征图

W——输入特征层在宽度上的维度

H——输入特征层在高度上的维度

表 2 MobileNet－v3 Large 网络配置

Tab. 2 MobileNet－v3 Large network configuration

序号	输入特征尺寸	操作类型	扩张因子	输出通道数	是否引入注意力	激活函数	步幅
0	$512^2\times3$	Conv2d		16		H－Swish	2
1	$256^2\times16$	bneck, 3×3	16	16		ReLU	1
2	$256^2\times16$	bneck, 3×3	64	24		ReLU	2
3	$128^2\times24$	bneck, 3×3	72	24		ReLU	1
4	$128^2\times24$	bneck, 5×5	72	40	√	ReLU	2
5	$64^2\times40$	bneck, 5×5	120	40	√	ReLU	1
6	$64^2\times40$	bneck, 5×5	120	40	√	ReLU	1
7	$64^2\times40$	bneck, 3×3	240	80		H－Swish	2
8	$32^2\times80$	bneck, 3×3	200	80		H－Swish	1
9	$32^2\times80$	bneck, 3×3	184	80		H－Swish	1
10	$32^2\times80$	bneck, 3×3	184	80		H－Swish	1
11	$32^2\times80$	bneck, 3×3	480	112	√	H－Swish	1
12	$32^2\times112$	bneck, 3×3	672	112	√	H－Swish	1
13	$32^2\times112$	bneck, 5×5	672	160	√	H－Swish	2
14	$16^2\times160$	bneck, 5×5	960	160	√	H－Swish	1
15	$16^2\times160$	bneck, 5×5	960	160	√	H－Swish	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

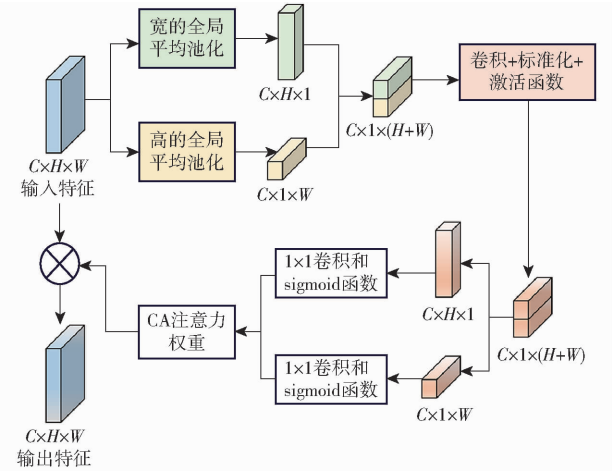

图 8　CA 注意力机制结构简图

Fig. 8　CA attention mechanism structure diagram

（2）将 I_c^h、I_c^w 这两个特征图合并，并经过卷积、标准化和非线性激活函数处理后得到一个特征表示 F，即

$$F=\delta(f([I_c^h,I_c^w])) \tag{5}$$

式中　δ——非线性激活函数

f——卷积、标准化处理函数

（3）将这个特征表示 F 分成宽度和高度方向上的两个特征图，再将宽度和高度转置回原来的维度。对于每个特征图，通过 1×1 卷积调整通道数后，应用 sigmoid 函数获得宽度和高度方向上的注意力情况，即

$$g_c^w=\sigma(f_w(F^w)) \tag{6}$$

$$g_c^h=\sigma(f_h(F^h)) \tag{7}$$

式中　σ——sigmoid 函数

f_w、f_h——宽度、高度上的转置卷积

F^w——特征表示 F 在宽度上的特征图

F^h——特征表示 F 在高度上的特征图

g_c^w、g_c^h——宽度、高度方向上的注意力权重

（4）将这两个注意力权重乘以原有的特征图，得到一个新的特征图

$$y_c(i,j)=x_c(i,j)g_c^h(i)g_c^w(j) \tag{8}$$

2.2　迁移学习

在深度学习任务中，一般会有 2 个问题：①因缺少样本数量导致网络出现过拟合现象。②网络收敛速度慢。在现实应用中，获取足够带有标签的训练样本往往成本较高。考虑到果实垄间导航任务在不同垄间环境中具有一定的相关性，可以通过迁移学习将已学到的模型参数共享给新模型，以减少新任务对样本数量的需求，并加快新网络模型的收敛速度。

迁移学习是机器学习领域中的一种通过将已学的知识迁移到新的学习任务中来加速新知识学习速度的常用方法[33]。在深度学习领域内，迁移学习同样可以用来加快模型学习效率[34]。在迁移学习中，已有知识称为源域，新的学习目标称为目标域，其示意图如图 9 所示。

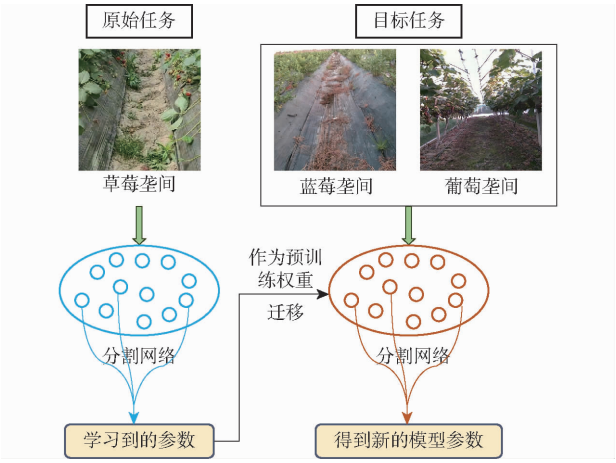

图 9　迁移学习示意图

Fig. 9　Transfer learning diagram

采用模型迁移，假设所提出的分割模型在垄间道路提取上具备一定的泛用性，不同垄间道路可以共享部分模型参数。将源域特征上学习获得的模型用于目标域特征上，并通过小样本学习来得到新模型。源域为草莓垄间特征，目标域为葡萄和蓝莓垄间特征。源域与目标域存在一定的联系，但又有不同之处。在草莓垄间训练得到的网络权重，作为葡萄和蓝莓垄间的预训练权重，通过小样本训练微调，得到新的网络模型。

2.3　路径提取

根据分割模型得到的可行驶区域，可进一步提取导航线。由于垄间作物是根据种植床或垄种植的，为了便于管理以及最大程度地利用土地资源，种植床或垄一般都以一定间距平行分布，即使存在规整误差，垄间道路整体上仍呈直线分布。因此，本文采用最小二乘法进行垄间路径导航线拟合。

2.3.1　最小二乘法

最小二乘法是依据已知点到直线距离平方和最小的思想来拟合一条直线的方法，具备拟合精度高、拟合速度快的优点。设拟合线为 $x=a+by$，其计算公式如下：

（1）已知点集 $\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\}$，进行线性回归，即

$$x_i=a+by_i+\varepsilon_i \quad (i=1,2,\cdots,n) \tag{9}$$

式中　ε_i——y_i 受随机因素影响的系数

（2）计算点到直线的距离平方和，即

$$Q(a,b)=\sum_{i=1}^{n}(x_i-a-by_i)^2 \tag{10}$$

(3)分别对 a、b 求偏导,即

$$\begin{cases}\dfrac{\partial Q}{\partial a}=-2\sum\limits_{i=1}^{n}(x_i-a-by_i)\\ \dfrac{\partial Q}{\partial b}=-2\sum\limits_{i=1}^{n}y_i(x_i-a-by_i)\end{cases}\tag{11}$$

(4)求解 $\hat{a}$、$\hat{b}$,即

$$\begin{cases}\hat{a}=\bar{x}-\hat{b}\bar{y}\\ \hat{b}=\dfrac{\overline{xy}-\bar{x}\bar{y}}{\overline{y^2}-\bar{y}^2}\end{cases}\tag{12}$$

其中

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i\tag{13}$$

$$\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i\tag{14}$$

$$\overline{xy}=\frac{1}{n}\sum_{i=1}^{n}x_iy_i\tag{15}$$

$$\overline{y^2}=\frac{1}{n}\sum_{i=1}^{n}y_i^2\tag{16}$$

(5)拟合直线方程为

$$x=\bar{x}+b(y-\bar{y})\tag{17}$$

2.3.2 导航线拟合

基于最小二乘法拟合导航线,步骤如下:

(1)利用分割网络模型对输入的待检测图像(图 10a)进行分割,得到可行驶区域分割结果(图 10b)。

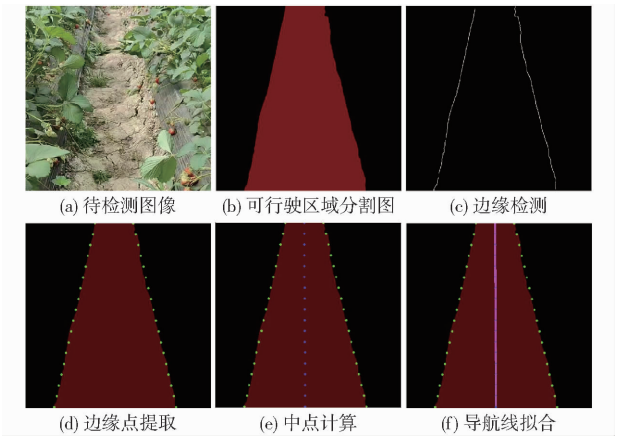

图 10 导航线拟合过程

Fig. 10 Navigation line fitting

(2)利用 Canny 算子对可行驶区域进行边缘检测,获得边缘信息(图 10c)。

(3)得到左右边缘信息后,为了改善规整误差以及枝叶遮挡导致的边缘点参差不齐的情况,利用一元一次方程的最小二乘回归分别将左右边缘点进行线性拟合,使左右边缘点呈线性排列。

(4)在高度方向上分别从左右边缘点中以 30 像素间距取点(图 10d),并计算左右边缘点之间的中点(图 10e),中点(x_{mid},y_{mid})计算公式为

$$\begin{cases}x_{\mathrm{mid}}=\dfrac{x_l+x_r}{2}\\ y_{\mathrm{mid}}=\dfrac{y_l+y_r}{2}\end{cases}\tag{18}$$

式中 (x_l,y_l)——左边界点像素坐标

(x_r,y_r)——右边界点像素坐标

(5)利用中点进行线性最小二乘回归得到导航线(图 10f)。

3 试验结果与分析

3.1 试验平台

试验选用的处理器为 Intel(R) Core(TM) i9-12900K,内存为 32 GB,在装有 NVIDIA GeForce RTX3090 GPU 的 Windows 10 操作系统的计算机上进行。试验中的所有程序基于 Python 3.8 语言,在 PyTorch 深度学习框架下实现。

3.2 超参数设置

草莓垄间数据训练选用的超参数为:迭代次数为 100 次,单次批量训练图像数量为 4,采用 Adam 自适应学习率优化器,初始学习率为 0.001,且最小学习率为 0.000 01,学习率衰减系数为 0.1。采用 BCEDiceLoss 作为训练的损失函数,该损失函数是由 Dice Loss 和二分类交叉熵损失(Binary Cross Entropy Loss,BCE Loss)按权重加权计算。在迁移学习中,训练得到的权重模型作为初始模型,采用蓝莓和葡萄垄间小样本数据进行再次训练,参数选用与草莓垄间数据训练保持一致。

3.3 模型训练与性能对比

为了比较模型性能,基于草莓垄间数据集分别训练 ResNet50-Unet、VGG16-Unet、Deeplabv3+、PSPnet 和改进的 U-Net 共 5 个网络模型,损失值曲线如图 11 所示。利用训练好的各个网络模型分割草莓垄间可行驶区域,并进行可视化检测效果对比,结果如图 12 所示。

本文还采用训练用时、理论推理速度、模型参数量、模型内存占用量以及分割精度等量化指标评估网络模型。其中,理论推理速度指网络对输入图像的像素点做出判断、分割生成掩码图所需时间,不包括图像输入的加载时间、图像检测完成后的结果与原图混合所需时间以及输出图像的保存写入时间;分割精度评估主要采用平均像素准确度(mPA)和平均交并比(MIoU),计算公式为

$$mPA=\frac{1}{N+1}\sum_{i=0}^{n}\frac{P_{ii}}{\sum\limits_{j=0}^{n}P_{ij}}\times100\%\tag{19}$$

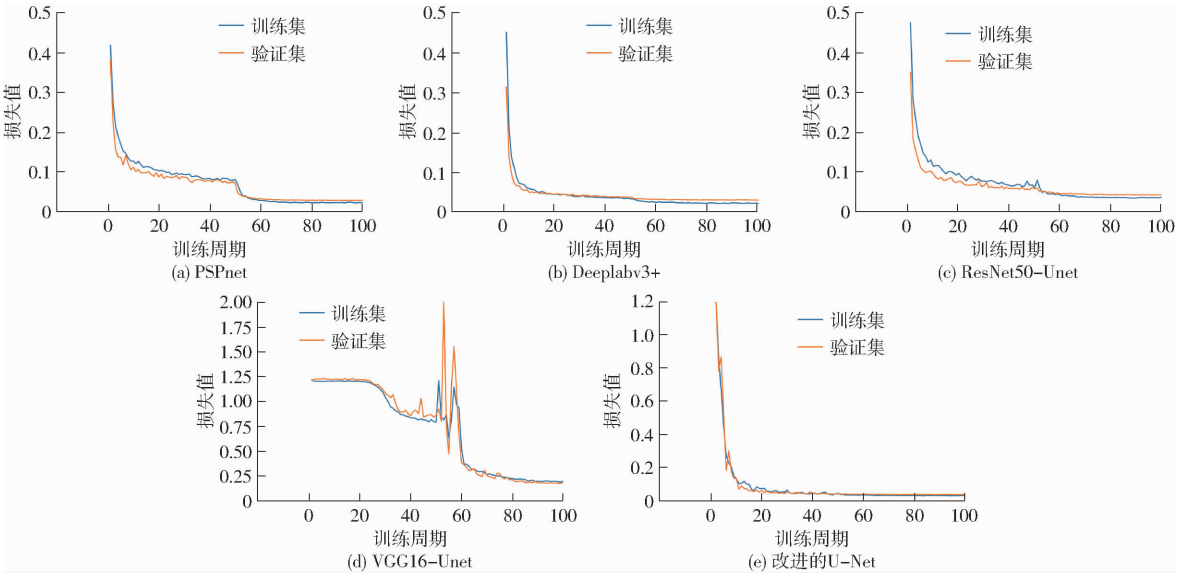

图 11　各模型训练损失函数曲线

Fig. 11　Curves of training loss function for each model

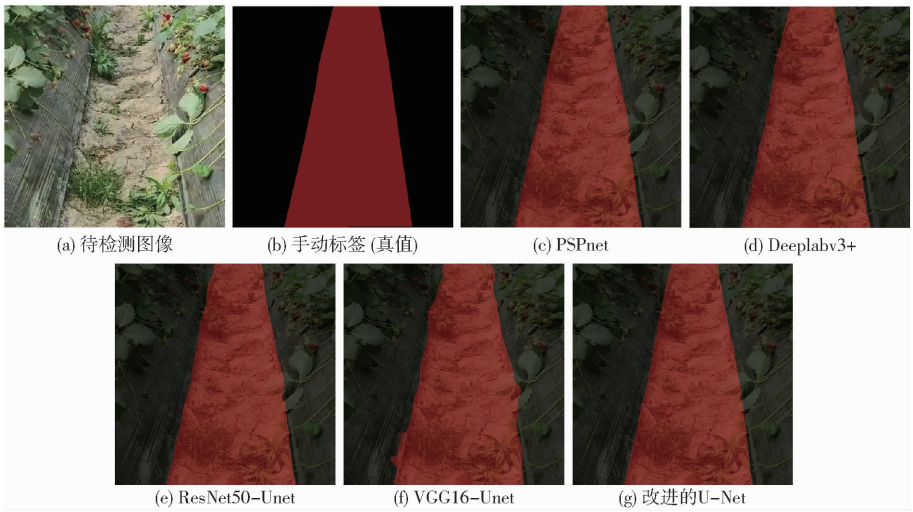

图 12　各模型分割结果可视化

Fig. 12　Visualizations of segmentation results of each model

$$MIoU = \frac{1}{N+1}\sum_{i=0}^{n}\frac{P_{ii}}{\sum_{j=0}^{n}(P_{ij}+P_{ji})-P_{ii}}\times 100\% \tag{20}$$

式中　N——目标类别数量

P_{ii}——真实值为 i 且预测值也为 i 的像素数量

P_{ij}——真实值为 i 但预测值为 j 的像素数量

P_{ji}——真实值为 j 但预测值为 i 的像素数量

本文的分割任务主要是提取垄间道路，因此 $N=1$，总类别数量要加上背景，即 $N+1$。

表 3 为不同网络模型的性能指标。由表 3 可知，本文提出的改进 U－Net 网络在各项指标上均优于其他 4 个网络。该网络平均交并比可达 98.06%，平均像素准确度可达 99.13%，远高于 VGG16－Unet 网络的 91.48% 和 95.87%。尽管相对于 ResNet50－Unet、Deeplabv3＋和 PSPnet 网络，本网络在 MIoU 和 mPA 指标上的提升并不显著，但本网络训练用时、模型参数量以及模型内存占用量都小于这 3 个网络。在网络推理速度上，本网络的理论推理速度可达 19.23 f/s，同样优于其余网络。

综上所述，改进的 U－Net 网络在果园垄间路径提取上具有良好的准确性与实时性，可以作为农业垄间路径导航的依据。

3.4　网络模型迁移学习性能验证

基于草莓垄间数据集训练的网络模型，利用蓝莓和葡萄垄间数据集在有无进行迁移学习的情况下检测网络性能，结果如表 4 所示。

改进的 U－Net 在蓝莓和葡萄垄间数据集上表现出色，性能较好。经过迁移学习后，该网络性能仍然保持优异，仅在蓝莓数据集时，mPA 略低于

表 3 网络性能参数对比

Tab. 3 Comparison of network performance parameters

网络模型	训练用时/s	MIoU/%	mPA/%	模型内存占用量/MB	模型参数量	理论推理速度/(f·s^{-1})
PSPnet	9 873.66	97.76	98.90	178.0	4.671×10^7	6.41
Deeplabv3 +	8 995.37	97.75	98.91	22.9	5.813×10^6	11.76
ResNet50 - Unet	11 942.67	97.74	99.03	167.0	4.393×10^7	4.33
VGG16 - Unet	15 117.54	91.48	95.87	94.9	2.489×10^7	3.86
改进的 U - Net	8 280.15	98.06	99.13	22.4	3.300×10^6	19.23

表 4 引入迁移学习前后网络模型性能

Tab. 4 Network model performance before and after introduction of transfer learning %

数据集	网络模型	MIoU		mPA	
		迁移学习前	迁移学习后	迁移学习前	迁移学习后
蓝莓	PSPnet	29.33	95.84	49.07	97.84
	Deeplabv3 +	30.17	86.64	50.17	99.11
	ResNet50 - Unet	30.17	96.88	49.18	98.32
	VGG16 - Unet	42.87	88.06	59.23	94.16
	改进的 U - Net	32.47	97.36	51.88	98.75
葡萄	PSPnet	71.81	98.12	85.26	99.07
	Deeplabv3 +	51.68	98.15	67.29	99.11
	ResNet50 - Unet	48.06	98.36	64.25	99.25
	VGG16 - Unet	42.13	92.70	59.64	96.52
	改进的 U - Net	61.65	98.50	75.48	99.29

Deeplabv3 + 。从整个数据训练结果来看,改进的 U - Net 的学习效果优于 ResNet50 - Unet、VGG16 - Unet、Deeplabv3 + 和 PSPnet。结果表明改进的 U - Net 网络在果园垄间道路提取任务中具有出色的性能和适应性。在不同果园垄间数据集上的训练和测试表现都优于其余网络,尤其是经过迁移学习后,性能优异。因此,改进的 U - Net 网络可以高效地识别和提取不同果园垄间的路径,并具备在不同环境下的快速适应性。

3.5 导航线提取结果评价

为了评估导航线提取的准确性,采用垄间路径提取模型得到的导航依据点坐标与真实标签的中点坐标之间的平均像素偏差作为评估指标。考虑到真实环境中智能农机对近处导航信息更为关注,在测试阶段,针对每种垄间类型分别随机选取强光、正常光和弱光条件下的 3 组样本,每组样本选择沿相机拍摄视角方向的前 10 个导航点进行像素偏差和平均像素偏差的计算。

采用 VideoCapture 视频类,利用所提出的垄间路径提取模型对强光、正常光和弱光条件下 3 组视频数据进行检测。视频数据导航线提取流程如图 13 所示。首先将输入视频逐帧转换成图像,然后将图像传递到垄间路径提取模型得到导航线信息,最后将得到预测路径图像重新组合为视频帧,从而重新生成视频。

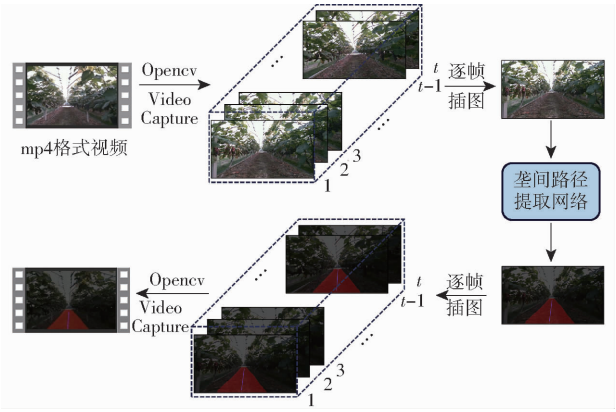

图 13 视频数据导航线提取流程图

Fig. 13 Flowchart for extracting navigation lines from video data

经过检测试验,改进的 U - Net 分割网络从获取图像到完成路径导航提取,单帧图像平均耗时为 0.211 s,导航线提取效果如图 14 所示。表 5 为相应的导航线偏差结果。由表 5 可知,不同垄间导航线精度误差基本处于 1 ~ 4 像素之间,平均偏差分别为 1.47、2.27、2.20 像素,精度较高。视频数据的导航线提取结果如表 6 所示。由表 6 可知,不同光照环境下的视频动态检测平均速率可达 2.27 f/s 以上,满足果园垄间路径导航的需求。

基于轻量化 U - Net 果园垄间路径识别方法可以有效地识别果园垄间路径,在不同光照环境下具备良好的鲁棒性和检测速率,满足果园垄间路径导航任务对于准确性和实时性的需求。

4 结论

(1)提出了一种轻量化的路径识别方法,在 U - Net 模型的基础上,采用 MobileNet - v3 Large 网络替换了 U - Net 的主干特征提取网络,缩减了网络模型体积,从而提升了网络检测速度;为了确保网络在轻量化的情况下仍能保持较好的准确性,在跳跃连接处引入 CA 注意力机制,以提高特征关注度。在草莓、蓝莓、葡萄垄间分别进行试验

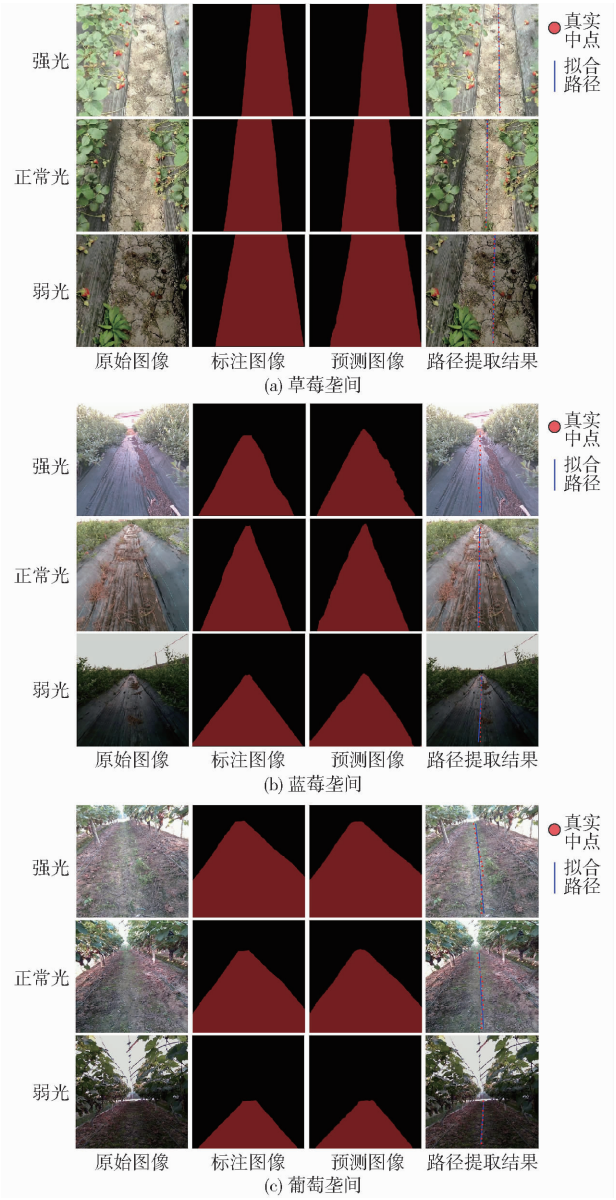

图 14　单帧图像的导航线提取效果

Fig. 14　Effect of extracting navigation lines from single frame images

以测试模型分割性能，得到识别精度 MIoU 分别为 98.06%、97.36%、98.50%，mPA 分别为 99.13%、98.75%、99.29%，模型内存占用量为 22.4 MB，理论推理速度为 19.23 f/s。

(2)引入迁移学习，将大样本草莓垄间数据集训练的网络模型迁移至蓝莓和葡萄垄间路径提取上。试验表明小样本迁移学习可拓展网络在不同垄间路径提取的适用性，为垄间路径导航提供了一种通用方法。

(3)基于网络分割的可行驶区域，通过 Canny 算子检测其边缘信息点，并利用最小二乘法重塑边缘信息，拟合道路中点，进而得到导航线。在草莓、蓝莓、葡萄垄间不同光线环境下分别进行试验，获得导航线平均像素偏差为 1.47、2.27、2.20 像素，单帧图像导航线提取的平均耗时为 0.211 s，视频数据动态检测平均速率为 2.27 f/s。结果表明，本文提出的路径提取方法具有较好的准确性和实时性，可为果园垄间路径导航提供参考。

表 5　导航线偏差结果

Tab. 5　Navigation line deviation result　像素

数据类型	光照类型	点序号										平均像素偏差
		1	2	3	4	5	6	7	8	9	10	
草莓垄间	强光	1	4	3	2	1	1	1	1	0	1	1.47
	正常光	1	1	1	1	1	1	1	3	3	3	
	弱光	4	2	2	2	2	1	0	0	0	0	
蓝莓垄间	强光	5	4	4	4	4	4	3	3	3	2	2.27
	正常光	4	3	2	2	1	0	1	1	2	3	
	弱光	2	1	1	1	1	1	1	1	2	2	
葡萄垄间	强光	3	3	3	3	1	5	7	4	3	4	2.20
	正常光	3	1	1	3	4	4	2	1	2	0	
	弱光	1	2	1	1	0	0	0	1	2	1	

表 6　视频数据导航线提取结果

Tab. 6　Video data navigation line extraction results

视频类型	光照类型	原视频总帧数/帧	原视频帧率/(f·s^{-1})	检测速率/(f·s^{-1})	平均检测速率/(f·s^{-1})
草莓垄间	强光	2 557	30.0	2.24	2.27
	正常光	2 380	30.0	2.31	
	弱光	1 719	30.0	2.27	
蓝莓垄间	强光	492	30.0	2.28	2.28
	正常光	884	30.0	2.29	
	弱光	1169	30.0	2.27	
葡萄垄间	强光	650	30.0	2.31	2.31
	正常光	800	30.0	2.33	
	弱光	829	30.0	2.29	

参　考　文　献

[1] 张漫，季宇寒，李世超，等. 农业机械导航技术研究进展[J]. 农业机械学报，2020，51(4)：1－18.
ZHANG Man，JI Yuhan，LI Shichao，et al. Research progress of agricultural machinery navigation technology[J]. Transactions of the Chinese Society for Agricultural Machinery，2020，51(4)：1－18. (in Chinese)

[2] ROSHANIANFARD A，NOGUCHI N，OKAMOTO H，et al. A review of autonomous agricultural vehicles (the experience of Hokkaido University)[J]. Journal of Terramechanics，2020，91：155－183.

[3] XIE D B，CHEN L，LIU L C，et al. Actuators and sensors for application in agricultural robots：a review[J]. Machines，2022，10(10)：913.

[4] GAO X Y，LI J H，FAN L F，et al. Review of wheeled mobile robots' navigation problems and application prospects in

agriculture[J]. IEEE Access, 2018, 6: 49248 - 49268.

[5] LI Y, IBANEZ-GUZMAN J. LiDAR for autonomous driving: the principles, challenges, and trends for automotive LiDAR and perception systems[J]. IEEE Signal Processing Magazine, 2020, 37(4): 50 - 61.

[6] 李云伍, 徐俊杰, 王铭枫,等. 丘陵山区田间道路自主行驶转运车及其视觉导航系统研制[J]. 农业工程学报, 2019, 35(1): 52 - 61.
LI Yunwu, XU Junjie, WANG Mingfeng, et al. Development of autonomous driving transfer trolley on field roads and its visual navigation system for hilly areas[J]. Transactions of the CSAE, 2019, 35(1): 52 - 61. (in Chinese)

[7] KNEIP J, FLEISCHMANN P, BERNS K. Crop edge detection based on stereo vision[J]. Robotics and Autonomous Systems, 2020, 123: 103323.

[8] OPIYO S, OKINDA C, ZHOU J, et al. Medial axis-based machine-vision system for orchard robot navigation[J]. Computers and Electronics in Agriculture, 2021, 185: 106153.

[9] LIN Y K, CHEN S F. Development of navigation system for tea field machine using semantic segmentation[J]. IFAC-PapersOnline, 2019, 52(30): 108 - 113.

[10] 韩树丰,何勇,方慧.农机自动导航及无人驾驶车辆的发展综述[J].浙江大学学报(农业与生命科学版),2018,44(4): 381 - 391,515.
HAN Shufeng, HE Yong, FANG Hui. Recent development in automatic guidance and autonomous vehicle for agriculture: a review[J]. Journal of Zhejiang University (Agric. & Life Sci.), 2018, 44(4): 381 - 391,515. (in Chinese)

[11] 周学成, 肖明玮, 梁英凯,等. 基于改进 DeepLabv3 + 的火龙果园视觉导航路径识别方法[J].农业机械学报, 2023, 54(9): 35 - 43.
ZHOU Xuecheng, XIAO Mingwei, LIANG Yingkai, et al. Navigation path recognition between dragon orchard using improved DeepLabv3 + network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(9): 35 - 43. (in Chinese)

[12] 宋怀波, 何东健, 辛湘俊. 基于机器视觉的非结构化道路检测与障碍物识别方法[J]. 农业工程学报, 2011, 27(6): 225 - 230.
SONG Huaibo, HE Dongjian, XIN Xiangjun. Unstructured road detection and obstacle recognition algorithm based on machine vision[J]. Transactions of the CSAE, 2011, 27(6): 225 - 230. (in Chinese)

[13] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436 - 444.

[14] SALEEM M H, POTGIETER J, ARIF K M. Automation in agriculture by machine and deep learning techniques: a review of recent developments[J]. Precision Agriculture, 2021, 22(6): 2053 - 2091.

[15] GARCIA-GARCIA A, ORTS-ESCOLANO S, OPREA S, et al. A survey on deep learning techniques for image and video semantic segmentation[J]. Applied Soft Computing, 2018, 70: 41 - 65.

[16] KAMILARIS A, PRENAFETA-BOLDU F X. A review of the use of convolutional neural networks in agriculture[J]. Journal of Agricultural Science, 2018, 156(3): 312 - 322.

[17] 肖珂, 夏伟光, 梁聪哲. 复杂背景下果园视觉导航路径提取算法[J]. 农业机械学报, 2023, 54(6): 197 - 204,252.
XIAO Ke, XIA Weiguang, LIANG Congzhe. Visual navigation path extraction algorithm in orchard under complex background [J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(6): 197 - 204,252. (in Chinese)

[18] GRIGORESCU S, TRASNEA B, COCIAS T, et al. A survey of deep learning techniques for autonomous driving[J]. Journal of Field Robotics, 2020, 37(3): 362 - 386.

[19] 张新钰, 高洪波, 赵建辉,等. 基于深度学习的自动驾驶技术综述[J]. 清华大学学报(自然科学版), 2018, 58(4): 438 - 444.
ZHANG Xinyu, GAO Hongbo, ZHAO Jianhui, et al. Overview of deep learning intelligent driving methods[J]. Journal of Tsinghua Univ. (Sci. & Technol.), 2018, 58(4): 438 - 444. (in Chinese)

[20] MOZAFFARI S, AL-JARRAH O Y, DIANATI M, et al. Deep learning-based vehicle behavior prediction for autonomous driving applications: a review[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(1): 33 - 47.

[21] GUASTELLA D C, MUSCATO G. Learning-based methods of perception and navigation for ground vehicles in unstructured environments: a review[J]. Sensors, 2021, 21(1): 73.

[22] 孟庆宽, 杨晓霞, 张漫,等. 基于语义分割的非结构化田间道路场景识别[J]. 农业工程学报, 2021, 37(22): 152 - 160.
MENG Qingkuan, YANG Xiaoxia, ZHANG Man, et al. Recognition of unstructured field road scene based on semantic segmentation model[J]. Transactions of the CSAE, 2021, 37(22): 152 - 160. (in Chinese)

[23] 韩振浩, 李佳, 苑严伟,等. 基于 U - Net 网络的果园视觉导航路径识别方法[J]. 农业机械学报, 2021, 52(1): 30 - 39.
HAN Zhenhao, LI Jia, YUAN Yanwei, et al. Path recognition of orchard visual navigation based on U - Net[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(1): 30 - 39. (in Chinese)

[24] 饶秀勤, 朱逸航, 张延宁,等. 基于语义分割的作物垄间导航路径识别[J]. 农业工程学报, 2021, 37(20): 179 - 186.
RAO Xiuqin, ZHU Yihang, ZHANG Yanning, et al. Navigation path recognition between crop ridges based on semantic[J]. Transactions of the CSAE, 2021, 37(20): 179 - 186. (in Chinese)

[25] 李云伍, 徐俊杰, 刘得雄,等. 基于改进空洞卷积神经网络的丘陵山区田间道路场景识别[J]. 农业工程学报, 2019,

35(7)：150－159.
LI Yunwu, XU Junjie, LIU Dexiong, et al. Field road scene recognition in hilly regions based on improved dilated convolutional networks[J]. Transactions of the CSAE, 2019, 35(7): 150－159. (in Chinese)
[26] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(12): 2481－2495.
[27] KIM W S, LEE D H, KIM Y J, et al. Path detection for autonomous traveling in orchards using patch-based CNN[J]. Computers and Electronics in Agriculture, 2020, 175: 105620.
[28] RONNEBERGER O, FISCHERF P, BROX T. U－Net: convolutional networks for biomedical image segmentation[C]// Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015: 18th International Conference, Munich, Proceedings, Part Ⅲ 18. Springer International Publishing, 2015: 234－241.
[29] SIMONYAN K, ZISSERMANZ A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
[30] 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.
[31] 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.
[32] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713－13722.
[33] BOSILJ P, APTOULA E, DUCKETT T, et al. Transfer learning between crop types for semantic segmentation of crops versus weeds in precision agriculture[J]. Journal of Field Robotics, 2020, 37(1): 7－19.
[34] WANG D, DENG L, NI J, et al. Recognition pest by image-based transfer learning[J]. Journal of the Science of Food and Agriculture, 2019, 99(10): 4524－4531.

(上接第 15 页)

[98] JIN H, LI A, WANG J, et al. Improvement of spatially and temporally continuous crop leaf area index by integration of CERES－Maize model and MODIS data[J]. European Journal of Agronomy, 2016, 78: 1－12.
[99] RUANE A C, HUDSON N I, ASSENG S, et al. Multi-wheatmodel ensemble responses to interannual climate variability[J]. Environmental Modelling & Software, 2016, 81: 86－101.
[100] TOAN T L, LAUR H, MOUGIN E, et al. Multitemporal and dual-polarization observations of agricultural vegetation covers by X-band SAR images[J]. European Journal of Nutrition, 2017, 56(3): 1339－1346.
[101] 解毅, 王鹏新, 刘峻明, 等. 基于四维变分和集合卡尔曼滤波同化方法的冬小麦单产估测[J]. 农业工程学报, 2015, 31(1): 187－195.
XIE Yi, WANG Pengxin, LIU Junming, et al. Winter wheat yield estimation based on assimilation method combined with 4DVAR and EnKF[J]. Transactions of the CSAE, 2015, 31(1): 187－195. (in Chinese)
[102] LI H, CHEN Z, LIU G, et al. Improving winter wheat yield estimation from the CERES-wheat model to assimilate leaf area index with different assimilation methods and spatio-temporal scales[J]. Remote Sensing, 2017, 9(3): 190.
[103] CHENG Z Q, MENG J H, QIAO Y Y, et al. Preliminary study of soil available nutrient simulation using a modified WOFOST model and time-series remote sensing observations[J]. Remote Sensing, 2018, 10(1): 21.
[104] 蒙继华, 程志强, 王一明. WOFOST 模型与遥感数据同化的土壤速效养分反演[J]. 遥感学报, 2018, 22(4): 546－558.
MENG Jihua, CHENG Zhiqiang, WANG Yiming. Simulating soil available nutrients by a new method based on WOFOST model and remote sensing assimilation[J]. Journal of Remote Sensing, 2018, 22(4): 546－558. (in Chinese)
[105] 陶飞,刘蔚然,刘检华,等.数字孪生及其应用探索[J].计算机集成制造系统,2018,24(1):1－18.
TAO Fei, LIU Weiran, LIU Jianhua, et al. Digital twin and its potential application exploration[J]. Computer Integrated Manufacturing Systems, 2018,24(1): 1－18. (in Chinese)
[106] PENG B, GUAN K, CHEN M, et al. Improving maize growth processes in the community land model: implementation and evaluation[J]. Agricultural and Forest Meteorology, 2018, 250: 64－89.
[107] BOAS T, BOGENA H, GRUNWALD T, et al. Improving the representation of cropland sites in the Community Land Model (CLM) version 5.0[J]. Geoscientific Model Development, 2021, 14(1): 573－601.
[108] PYLIANIDIS C, OSINGA S, ATHANASIADIS I N. Introducing digital twins to agriculture[J]. Computers and Electronics in Agriculture, 2021, 184: 105942.