doi:10.6041/j.issn.1000-1298.2025.05.003

基于车载成像与深度卷积神经网络的地表残膜识别方法

吕继东[1,2]　翟志强[1,3]　孟庆建[1,2]　苗璐鹏[1,2]　陈　悦[1,2]　张若宇[1,2]

(1.石河子大学机械电气工程学院,石河子 832003;2.农业农村部西北农业装备重点实验室,石河子 832003;
3.兵团智慧农场数字化装备技术创新中心,石河子 832003)

摘要:针对残膜回收机实际作业过程中存在多种相似非目标场景干扰,目标场景图像背景复杂且地表残膜尺寸小、破碎度大、无固定轮廓导致残膜覆盖率难以准确评估的问题,提出基于车载成像和深度卷积神经网络的地表残膜识别方法。构建了一种基于多重特征增强的 SE-DenseNet-DC 分类模型,在 DenseNet121 模型每个稠密块的非线性组合函数前后引入通道注意力机制增强有效特征信息通道的权重,然后引入多尺度串联空洞卷积替换原始模型第 1 层卷积提升感受野并保持细节敏感度,实现目标场景图像的有效提取;构建了一种基于细节信息增强和多尺度特征融合的 CDC-TransUnet 分割模型,在 TransUnet 模型的编码器部分引入 CBAM 模块提取更加细微和精确的全局特征,在跳跃连接部分引入 DAB 模块融合多尺度语义信息并弥补编码和解码阶段特征之间的语义差距,然后在解码器部分引入 CCAF 模块减少上采样丢失的细节信息,实现目标场景图像复杂背景中地表残膜的精准分割。试验结果表明,SE-DenseNet-DC 分类模型对目标场景图像的分类准确率、查准率、查全率和 F1 值分别达到 96.26%、91.54%、94.49% 和 92.83%,CDC-TransUnet 分割模型对目标场景图像中地表残膜分割平均交并比(MIOU)达到 77.17%,模型预测残膜覆盖率与人工标注残膜覆盖率决定系数(R^2)为 0.92,均方根误差(RMSE)为 0.23%,平均相对误差为 2.95%,单幅图像评估时间平均为 0.54 s。本文方法在残膜回收机回收后地表残膜覆盖率监测评估中具有较高的准确率和较快的推理速度,为残膜回收机回收质量实时准确评估提供技术支撑。

关键词:棉田;残膜回收;车载成像;深度卷积神经网络;识别

中图分类号:S24　　**文献标识码**:A　　**文章编号**:1000-1298(2025)05-0026-12　　**OSID**:

Surface Residual Film Recognition Method Based on Vehicle-mounted Imaging and Deep Convolutional Neural Networks

LÜ Jidong[1,2]　ZHAI Zhiqiang[1,3]　MENG Qingjian[1,2]　MIAO Lupeng[1,2]　CHEN Yue[1,2]　ZHANG Ruoyu[1,2]

(1. *College of Mechanical and Electrical Engineering*, *Shihezi University*, *Shihezi* 832003, *China*
2. *Key Laboratory of Northwest Agricultural Equipment*, *Ministry of Agriculture and Rural Affairs*, *Shihezi* 832003, *China*
3. *Technology Innovation Center of Smart Farm Digital Equipment*, *Xinjiang Production and Construction Corps*, *Shihezi* 832003, *China*)

Abstract: Aiming to address the challenges in accurately assessing residual film coverage due to interference from multiple similar non-target scenarios, complex background textures in target scene images, and the small size, high fragmentation, and irregular contours of residual films during the operational process of residual film recovery machinery, a residual film recognition method was proposed based on vehicle-mounted imaging and deep convolutional neural networks. A multi-feature-enhanced SE-DenseNet-DC classification model was developed by integrating channel attention mechanisms before and after the nonlinear combination functions in each dense block of the DenseNet121 architecture, the model enhanced the weighting of effective feature channels. Additionally, the first-layer convolution of the original model was replaced with multi-scale cascaded dilated convolutions to expand the receptive field while preserving sensitivity to fine details, enabling effective extraction of target scene images. Furthermore, a CDC-TransUnet segmentation model was constructed with enhanced detail

收稿日期:2024-11-04　修回日期:2024-12-13

基金项目:国家重点研发计划项目(2022YFD2002400)、国家棉花产业技术体系岗位科学家项目(CARS-15-17)和石河子大学高层次人才科研启动项目(RCZK202442)

作者简介:吕继东(2000—),男,硕士生,主要从事图像处理和智慧农业研究,E-mail:2798778283@qq.com

通信作者:张若宇(1980—),男,教授,博士生导师,主要从事数字棉花技术与装备研究,E-mail:ry248@163.com

information and multi-scale feature fusion. In the encoder of the TransUnet framework, CBAM modules were introduced to capture finer and more precise global features. DAB modules were embedded in the skip connections to fuse multi-scale semantic information and bridge the semantic gap between encoder and decoder features. CCAF modules were then incorporated into the decoder to mitigate detail loss during upsampling, achieving precise segmentation of residual films against complex backgrounds in target scenes. Experimental results demonstrated that the SE - DenseNet - DC classification model achieved classification accuracy, precision, recall, and F1 score of 96.26%, 91.54%, 94.49%, and 92.83%, respectively, for target scene image classification. The CDC - TransUnet segmentation model achieved an average intersection over union (MIOU) of 77.17% for surface residual film segmentation. The coefficient of determination (R^2) between the predicted and manually annotated film coverage was 0.92, with root mean square error (RMSE) of 0.23%, and average relative error of 2.95%. The average evaluation time was 0.54 s per image. This method demonstrated high accuracy and rapid processing capabilities for real-time monitoring and evaluation of residual film coverage post-recovery, providing robust technical support for quality assessment in residual film recovery operations.

Key words: cotton fields; residual film recycling; vehicle-mounted imaging; deep convolutional neural network; identification

0　引言

棉花是我国重要的战略物资,在国民经济和社会发展中占据重要地位[1]。覆膜植棉具有增加土壤贮水量、增温保墒、缩短生长周期等优点,可以显著提高棉花产量[2]。但农用地膜材料主要为聚乙烯,在自然条件下降解时间超过 200 a,如果不进行及时有效回收,将造成严重的土壤面源污染[3]。残膜污染治理是一项系统工程[4],除开发有效的残膜回收机具外,开展高效精准的残膜回收质量评估,对指导残膜回收作业补贴,引导残膜污染高效治理具有重要意义。

近年来相关学者已实现利用无人机低空成像进行农田地表残膜识别的初步探索[5-8]。尽管这些研究在残膜识别方面取得了一定进展,但多集中于春播前或秋收后空旷地表上的残膜。针对残膜回收机作业过程中的残膜覆盖率进行快速、准确的监测评估,作业环境复杂多变,无人机难以实时跟踪每台回收机具的具体位置和状态,使其难以满足残膜回收机回收质量监测评估需求。

传统的残膜回收机回收质量监测评估方式是通过农业环境执法人员采用手机拍照的方法,进行残膜回收机回收质量简单评断。该方法繁琐、低效,而且受主观经验的影响,观测结果因人而异,无法使用统一标准进行衡量。为实现农田残膜精准回收,文献[9-10]采用近地面成像设备对残膜回收机收后地表残膜进行准确识别和定位,但缺乏智能残膜识别系统,难以满足残膜回收机回收质量实时准确的监测需求。因此,亟需一种智能残膜回收质量监测评估方法,实现地表残膜覆盖率快速准确评估的同时减少人工干预,从而提高监测评估的时效性和准确性。

车载视觉成像具有空间分辨率高、成本低和实时性强等特点[11],在作物信息实时获取方面得到广泛应用[12-16]。因此采用车载视觉成像系统聚焦残膜回收机作业过程中所采集图像的地表残膜覆盖率的快速准确评估。然而在残膜回收机实际作业的复杂环境下,实时采集到的图像数据中存在多种相似非目标场景干扰,常常存在非目标场景被误分类的问题,严重影响残膜覆盖率的准确评估。在图像数据集中精准提取出目标场景图像,成为残膜覆盖率准确评估的关键前提。因此,本文针对识别目标图中存在多种相似非目标场景干扰,目标场景图像背景复杂且地表残膜尺寸小、破碎度大、无固定轮廓导致残膜覆盖率难以准确评估的问题,提出一种基于车载成像和深度卷积神经网络的地表残膜识别方法,并通过田间试验验证本文方法的可靠性和有效性。

1　材料与方法

对作业场景图像数据进行数据标注,并进行数据增强和划分数据集等数据预处理;接着,构建 SE-DenseNet-DC 分类模型对地膜回收后目标场景图像进行有效提取;将提取出的目标场景图像进行数据增强等数据预处理;构建 CDC-TransUnet 语义分割模型对图像残膜覆盖率进行评估。主要技术路线如图 1 所示。

1.1　图像数据采集与处理

1.1.1　试验区域与图像数据采集

试验区域位于新疆维吾尔自治区石河子市北泉镇和昌吉回族自治州玛纳斯县包家店镇棉花种植区域(43°37′N,85°97′E;44°23′N,86°19′E)。共随机

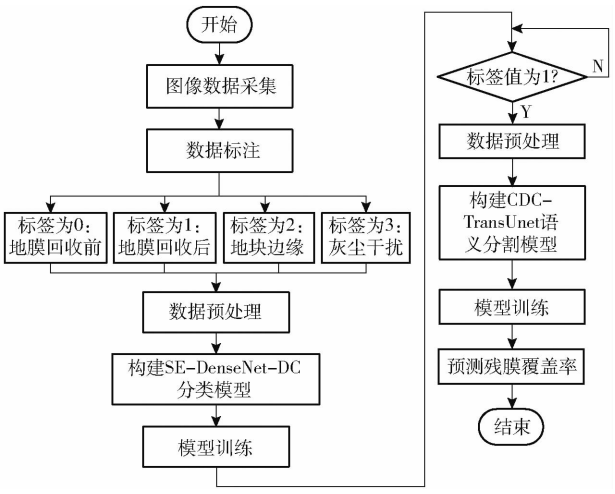

图 1 技术路线图

Fig. 1 Technology roadmap

选取 10 块常年覆膜种植棉花的地块，图像数据采集时间为 2023 年 10 月 20 日至 11 月 1 日，采集时间为 10:00—19:00。此时试验田中棉花已经收获，但地膜和秸秆尚未回收，地表存在大量待收地膜和秸秆。每块试验田尺寸为 500 m × 200 m，采样过程中，残膜回收机沿地块边界绕圈采样，相机安装于驾驶室右侧，以最大限度地采集地膜回收后目标场景图像，部分试验区域及采样过程如图 2 所示。

采用弹齿式残膜回收机（光大 11CMJD - 2.0 型，光大农机有限公司）搭载鹰眼运动相机（FIREFLY6S 型，深圳鹰眼航拍科技有限公司）拍摄采样区域内图像，相机镜头角度为 30°，离地高度 1.8 m。拍摄地膜回收后目标场景图像和地膜回收前、地块边缘和灰尘干扰 3 种常见的非目标场景图

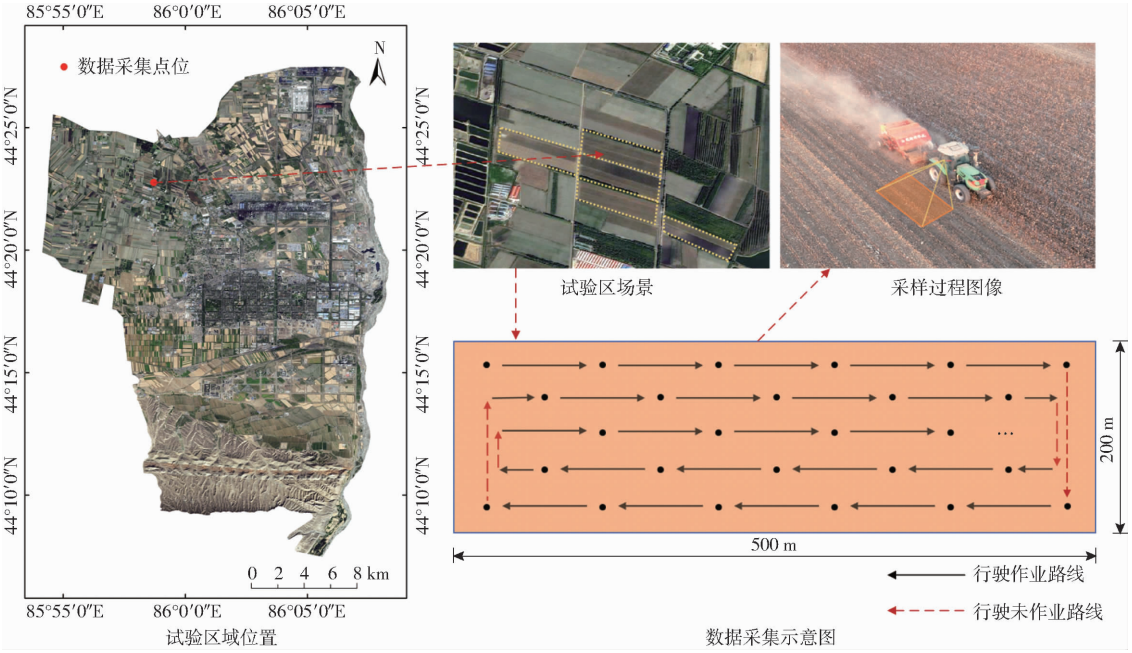

图 2 试验区域示意图

Fig. 2 Experimental area schematic

像，如图 3 所示，图像分辨率为 4 608 像素 × 3 456 像素，格式为 JPG，相机无畸变。为提高图像样本多样性和模型在复杂作业环境下的实用性，对不同天气条件下图像数据进行采集。每块棉田沿直线方向采用 Python 编写的采集算法进行拍摄，设定速率为 10 f/min，环境风速小于 3 级，此时地表残膜无明显摆动。通过照度计（TA8126B 型，苏州特安斯电子实业有限公司）记录不同天气条件下的光照强度：晴天光照强度为 70 ~ 110 klx，此时地表存在土块局部反光现象，多云天光照强度为 30 ~ 70 klx，此时地表土块无明显反光现象。

1.1.2 作业场景分类图像数据预处理

在残膜回收机复杂作业环境下，光照强度、光照角度和相机抖动等环境因素都会对图像数据的质量

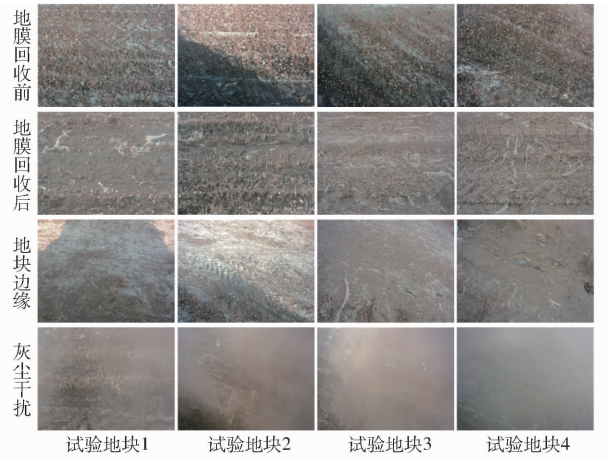

图 3 不同地块目标场景图像和常见非目标场景图像

Fig. 3 Target scene images of different plots and common non-target scene images

造成影响，使得原始数据并不能总是适合分类模型，将大大降低实际应用中模型分类准确性。因此，需要对原始数据进行预处理以提高样本多样性，进而提高分类模型在各种作业环境下的适用性和鲁棒性。

首先对采集的 2 009 幅作业场景图像数据进行数据标注：地膜回收前场景图像标注为 0、地膜回收后场景图像标注为 1、地块边缘场景图像标注为 2、灰尘干扰场景图像标注为 3。随后采用亮度变化、旋转、裁剪、高斯噪声和椒盐噪声等数据增强手段对数据集进行扩充。通过调整图像亮度和对比度，降低光照因素的影响；通过调整旋转角度，丰富不同角度图像数据，降低模型对某些图像属性的依赖；通过随机添加高斯噪声和椒盐噪声来干扰图像的 RGB 值，以模拟不同拍摄质量，降低相机抖动的影响和模型过拟合的风险；通过裁剪对数据集扩充，进一步提高模型的准确性和鲁棒性[17-18]。

由于图像采集在残膜回收机复杂作业环境下存在不确定性，导致所采集的图像数据分布不均衡，尤其是灰尘干扰场景的图像样本数量较少，因此本研究按照比例 6∶2∶2划分训练集、验证集和测试集，保留更多的数据作为验证集和测试集，以减少模型在训练集上过拟合的风险，同时提供更充分和准确的模型评估。共采集 2 009 幅原始图像，数据增强后生成 16 072 幅图像样本，数据集各类样本统计信息如表 1 所示。

表 1　作业场景图像数据集样本数

Tab. 1　Composition of operation scene image dataset

图像类型	原始图像	数据增强			
		训练集	验证集	测试集	总计
地膜回收前	565	2 712	904	904	4 520
地膜回收后	821	3 941	1 314	1 313	6 568
地块边缘	502	2 410	803	803	4 016
灰尘干扰	121	581	194	193	968
总计	2 009	9 644	3 215	3 213	16 072

1.1.3　回收后地表残膜分割图像数据预处理

从作业场景图像数据集中随机选取 450 幅地膜回收后图像样本，并进行了人工标注。首先在 Photoshop CS5 中对图像中的残膜区域进行手动标注，并填充为蓝色。随后，采用阈值分割方法进行二值化处理，其中残膜区域标注为 1，背景区域标注为 0，并将标注结果保存为 PNG 格式文件。为了确保标注质量，标注完成后进行了样本复检。为提高语义分割模型泛化能力和数据多样性，将地膜回收后场景图像与对应的标签图按照作业场景图像数据相同的数据增强方法进行了预处理，并按照比例 8∶1∶1对训练、验证和测试数据集进行划分。经过数据增强处理后，原始 450 幅图像扩充为 3 600 幅。

1.2　作业场景分类模型构建

DenseNet[19]通过层与层之间的密集连接，最大化了网络中的信息流动。通过相互连接所有层，DenseNet 允许每层获取前一层的额外输入，同时将本层提取的特征传递给后续层，从而有效保留了前向特征信息。DenseNet121 通过密集连接促进特征重用，可以全面捕捉图像的细节和结构，对于需要注重图像深层次细节信息的作业场景分类任务，深层网络 DenseNet121 对其具有更强的适用性。

在残膜回收机实际作业的复杂环境下，采集到的图像数据中存在多种相似非目标场景干扰，非目标场景与目标场景图像特征相近。利用原始 DenseNet121 模型对目标场景图像进行提取时往往会出现非目标场景被误提取的现象，这是由于 DenseNet121 模型未考虑不同通道间的相关性[20]，在特征提取的过程中对有效特征信息的筛选不足。因此，为提高模型对目标场景图像的提取能力，需要对各特征通道的权值进行矫正。而在特征提取的过程中，SE 通道注意力机制模块[21-22]能够增强有效特征信息通道的权重，降低并压缩无用特征信息通道的权重，从而最大化有用信息的利用效率。因此本研究通过 SE-Net 的基本模块 SE-Block 引入通道注意力机制，将 SE-Block 嵌入到 DenseNet 中每个稠密块(Dense Block)中的非线性组合函数前后构建 SE-Dense Block，从而构建特征通道加权的 DenseNet(SE-Dense Block)，以提高模型对相似特征的识别能力。SE-Dense Block 结构如图 4 所示。

空洞卷积(Dilated convolution)可以有效增大感受野，以较少的参数纳入更大的上下文信息[23]。在 DenseNet 中，图像输入首先经过一个 7×7 的卷积层，参数量大且无法提升网络的性能。因此为了缓解加入通道注意力机制模型计算量增加的问题，同时保留更多的细节信息。引入空洞卷积替换原始模型的第一层卷积层。太大的空洞卷积无法提取细粒度信息，太小的空洞卷积无法提取大尺度信息，并且感受野以倍数增加能够更有效地捕捉到不同尺度的信息，形成更明显的层次化特征提取结构[24]。因此，引入空洞率为 1、2、4 不同参数组合的多尺度串联空洞卷积，在增大感受野的同时保持对细节的敏感度，以进一步强化模型的特征提取能力。将引入通道注意力机制和空洞卷积的网络模型称为 SE-DenseNet-DC，模型整体结构如图 5 所示。图中，Dilated Conv 为空洞卷积模块，Conv 为卷积模块，Pooling 为最大池化下采样，SE-Dense Block 为嵌入

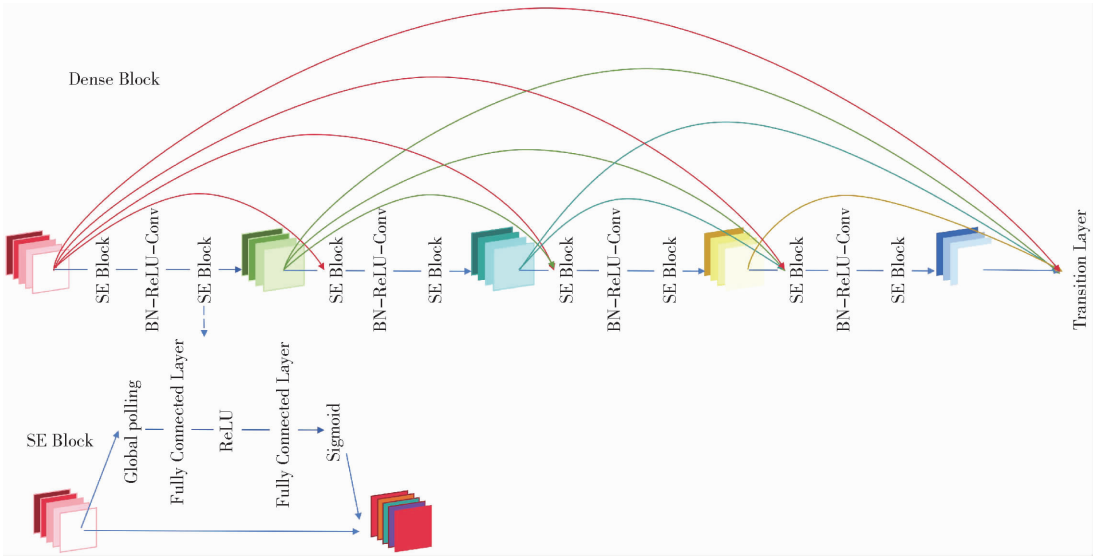

图 4 SE－Dense Block 网络结构图

Fig. 4 SE－Dense Block network structure diagram

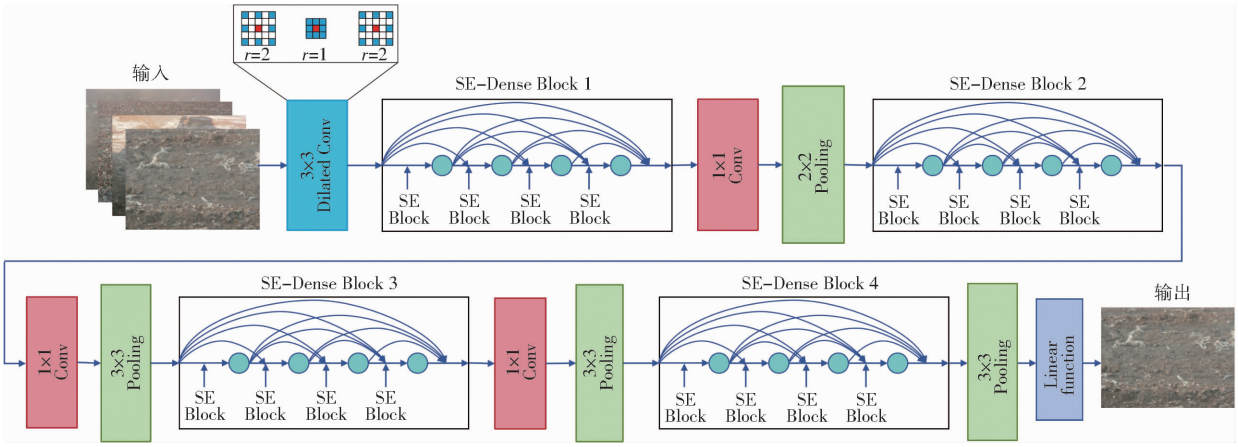

图 5 SE－DenseNet－DC 分类模型结构图

Fig. 5 Structure diagram of SE－DenseNet－DC classification model

SE－Block 的稠密块，Linear function 为线性函数，r 为空洞卷积空洞率。

1.3 收后地表残膜分割模型构建

TransUNet 网络采用混合 CNN－Transformer 架构，用重视全局信息的 Transformer 结构和重视局部图像特征的 CNN 进行混合编码来提高分割效果，其前向传播过程为输入图像先通过 CNN 编码器和 Transformer 编码器提取深层次特征图，然后通过解码器逐步恢复分辨率进行最终的分割预测[25]。在 TransUNet 基础上提出改进的 CDC－TransUNet 网络，其架构主要由 6 个关键组件组成：CNN 部分、CBAM 模块、DAB 模块、Transformer 层、跳跃连接和 CCAF 模块。具体实现如下：

（1）车载成像系统采集的图像中地表残膜尺寸各异、形状复杂多变，且存在大量性状相似的反光土块、棉铃壳和棉花等因素干扰，需要模型在特征提取时更好地平衡细节和全局信息，Transformer 通过其自注意力机制能够有效捕获全局特征，但其注意力机制仅限于单向聚焦位置属性，而在局部特征的多角度、多维度分析上存在局限，且传统的 U－Net 架构擅长通过卷积操作提取细致的局部特征，但缺乏全面整合上下文信息和获取全局语义细节的能力。

（2）卷积块注意力模块（Convolutional block attention module，CBAM）结合通道注意力（Channel attention）和空间注意力（Spatial attention）2 种注意力机制，能够有效地提升图像特征表达能力[26]。通过先对通道维度进行加权，再对空间维度进行加权，CBAM 模块能够有效地捕获图像中的关键特征区域，并抑制冗余的背景信息。在 Transformer 层之前嵌入 CBAM 模块，以提高模型的特征表达能力，将特征图细化到 Transformer 中。首先编码器 CNN 部分对输入的地表残膜图像进行处理，通过 3 个连续的卷积块逐步提取关键特征，系统地获取感受野集中的关键特征。随后，CBAM 模块从空间和通道 2 个维度对特征进行双重注意力机制处理，在增强特征表示深度的同时又有效保留了输入图像的原始特

征。此后，调整后特征被传至 Transformer 框架中以提取全方位的全局特征。上述过程将特征图细化并输入到 Transformer 中，从而实现了更加细微和精确的全局特征提取。

(3)跳跃连接能够有效地将低层网络中丰富的边缘信息传递到高层特征，弥补高层特征损失的细节信息，从而帮助恢复图像的细粒度结构。然而 TransUNet 在实现跳跃连接时直接将编码器的输出特征与解码器的输入特征进行连接，不足以充分融合不同尺度下的语义信息并限制了其全局建模能力[27]。此外，浅层和深层特征之间距离较远，语义差距较大，单纯依赖简单的通道合并无法弥补下采样过程中损失的边缘信息。双注意力桥模块(Dual attention bridge block，DAB)结合了空间注意力和通道注意力机制，能够在跳跃连接中有效地增强不同层次特征的融合[28]。CBAM 模块通过先后在通道和空间维度上分别加权特征，主要侧重于局部特征的显著性提升，与 CBAM 模块不同，DAB 模块在此基础上进一步加强了跨尺度特征的融合能力，通过在不同尺度上细化特征表达，在编码器和解码器之间建立了更强的语义联系，从而在特征融合时有效弥补了编码器和解码器之间的语义差距。此外，DAB 模块的双重注意力机制通过整合来自不同层次的特征，进一步强化了全局依赖关系，提升模型的全局建模能力。因此将双注意力桥模块集成到跳跃连接部分以增强特征融合的能力，通过在跳跃连接处融合来自编码器的特征，不仅能够有效地整合不同尺度的特征信息，还弥补了编码器和解码器之间的语义差距，进一步强化了全局依赖关系。

(4)U－Net 结构在上采样和下采样过程中会难以避免地损失图像的细节位置信息。跨通道注意力融合模块(Cross-channel attention fusion block，CCAF)通过通道注意力机制有效融合高层语义和低层细节特征，在解码过程中借助低层图像中更丰富的位置信息来提高边缘分割的精度，同时能够避免上采样过程中的信息丢失[28]。因此通过 CCAF 模块融合来自跳跃连接和编码器的特征，以弥补上采样过程中位置信息损失。

通过结合 CNN 网络、Transformer 架构、CBAM 模块、DAB 模块和 CCAF 模块，构建了一种具有高效协同机制的 CDC－TransUNet 网络，其结构如图 6 所示。图中，ReLU 为激活函数，Channel Attention 为通道注意力机制，Spatial Attention 为空间注意力机制，Map Flattening 为映射扁平化，Map Reversion 为映射反转，FFN 为前馈网络，Hidden Feature 为隐藏特征，Linear Projection 为线性投射，LayerNorm 为归一化，Cross Attention 为跨注意力机制，MSA 为多头注意力，MLP 为多层感知器，n 为 Transformer 层的堆叠数量，n_{patch} 为图像分块数，D 为分块嵌入维度，C 为图像通道维度，H 和 W 为图像高度与宽度。

1.4　评价指标

为了测试和评估分类模型和语义分割模型各项性能，采用准确率、查准率、查全率和 F1 值作为分类模型评价指标，采用准确率、精准率、召回率和平均交并比作为语义分割模型评价指标。

1.5　试验环境配置与超参数设置

分类模型和语义分割模型的训练和验证均在同一台设备上完成，设备采用 Windows 10 操作系统，CPU 配置为 Intel(R) Xeon(R) W－2223 CPU @ 3.60 GHz，显卡配置为 NVIDIA GeForce RTX 3090(显存 24 GB)，Python 版本为 3.7，采用 TensorFlow－GPU 2.5.0 深度学习框架，CUDA 版本 11.2。分类模型和语义分割模型训练的迭代次数均为 100，优化器为 Adam，学习率设置为 0.001，损失函数为交叉熵损失函数，分类模型训练批次大小为 32，语义分割模型训练批次大小为 16。

2　试验与结果分析

2.1　分类模型

2.1.1　多尺度特征提取试验

多尺度空洞卷积由空洞率进行调控，为了验证不同空洞率的空洞卷积对分类准确率的影响，在相同的试验环境下，以 DenseNet121 模型为基准，采用田口试验方法对 L9 正交表中的 9 组空洞卷积组合在数据增强后的作业场景图像数据集上进行试验，以确定最佳空洞卷积组合。结果如表 2 所示。

从表 2 可以看出，采用空洞率为[2，1，2]空洞卷积组合，模型各项指标表现最优，准确率、查准率、查全率和 F1 值分别达到 92.52%、85.71%、90.81% 和 87.45%。因此选取空洞率为[2，1，2]组合的多尺度串联空洞卷积替换原始 DenseNet121 模型第 1 层 7×7 普通卷积，以进一步强化模型特征提取能力。

2.1.2　基于 DenseNet121 的消融试验

为验证本文方法对残膜回收机复杂作业环境下作业场景的分类效果，在相同的试验环境下，以 DenseNet121 模型为基准，引入通道注意力机制和空洞卷积对本研究数据增强后的作业场景图像数据集进行消融试验。试验结果如表 3 所示。

由表 3 可以看出，在原始 DenseNet121 模型的基础上引入通道注意力机制和空洞卷积都获得了比原始 DenseNet121 模型更优的性能指标，清楚地证

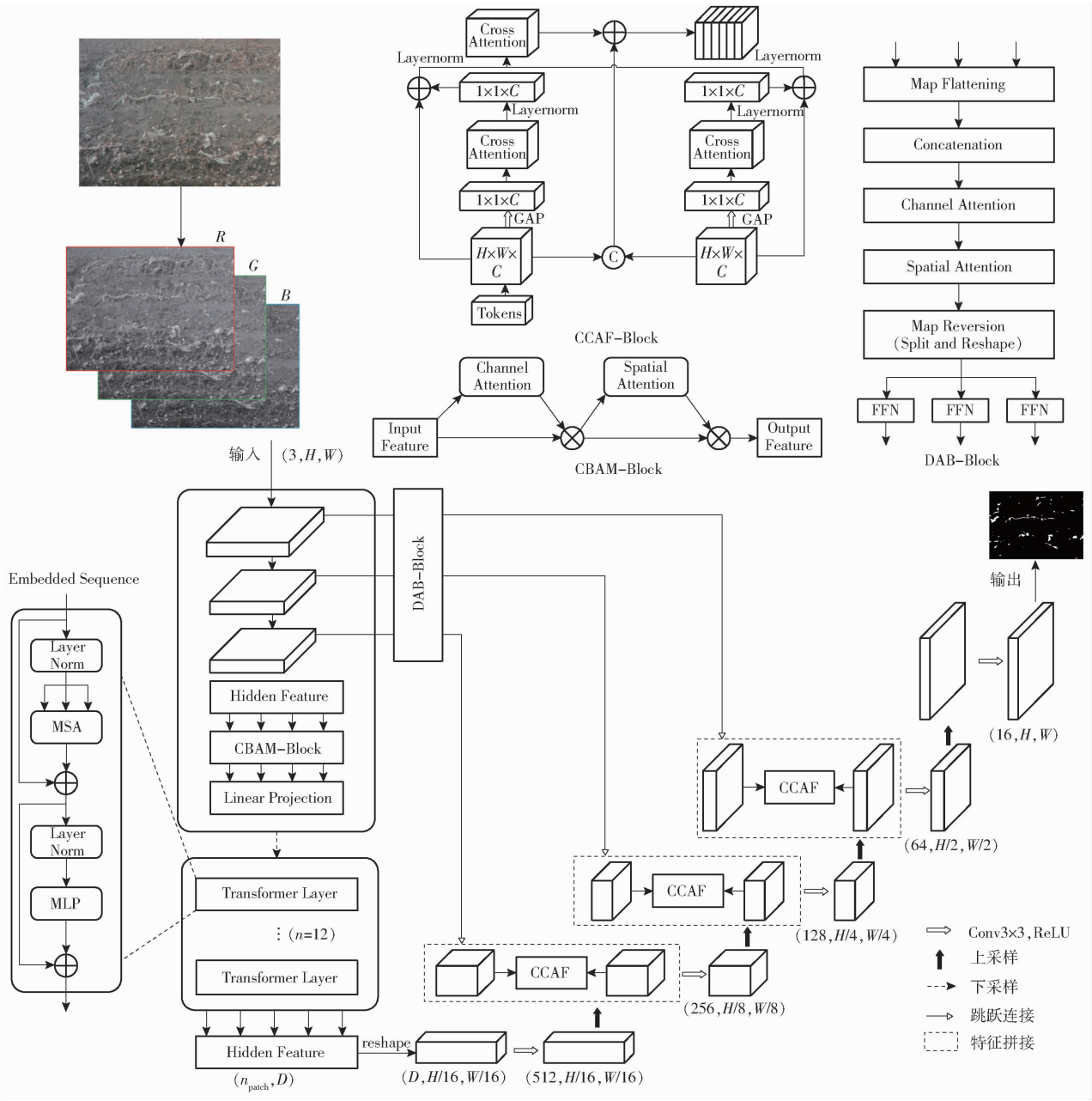

图 6 CDC－TransUNet 分割模型结构图

Fig. 6 Structure diagram of CDC－TransUNet segmentation model

表 2 不同空洞率空洞卷积下模型性能对比

Tab. 2 Comparison of model performance at different dilation rates of dilated convolution %

空洞率 r	准确率	查准率	查全率	F1 值
[1,1,1]	86.02	79.06	84.28	81.17
[1,2,2]	81.79	76.45	81.63	77.84
[1,4,4]	88.77	83.34	87.56	84.98
[2,1,2]	92.52	85.71	90.81	87.45
[2,2,4]	80.30	75.11	80.06	76.38
[2,4,1]	79.04	71.83	78.02	72.45
[4,1,4]	90.87	85.29	89.59	87.06
[4,2,1]	90.02	84.69	90.38	86.68
[4,4,2]	85.72	78.86	85.56	81.32

表 3 基于 DenseNet121 的消融试验结果

Tab. 3 Ablation experiments based on DenseNet121 %

序号	通道注意力机制	空洞卷积	准确率	查准率	查全率	F1 值
1	×	×	89.78	82.41	88.05	83.50
2	×	√	92.52	85.71	90.81	87.45
3	√	×	93.27	86.25	91.51	87.98
4	√	√	96.26	91.54	94.49	92.83

注：“√”表示在模型中加入此模块，“×”表示在模型中未加入此模块，下同。

明了通道注意力机制和空洞卷积对本研究作业场景分类任务的有效性。通道注意力机制使模型准确率、查准率、查全率和 F1 值分别提高 3.49、3.84、

3.46、4.48 个百分点，表明通道注意力机制通过增强有效特征信息通道的权重，能够提高分类的准确性。空洞卷积同样使得分类模型的各项指标均有提升，表明空洞卷积通过增大感受野的同时保持对细节信息的敏感度能够提升模型的分类效果。基于集成了通道注意力机制和空洞卷积的 DenseNet121 模型，本文提出的 SE－DenseNet－DC 模型结合了通道注意力机制和空洞卷积的双重优势，在测试集上最佳模型性能优于其他对比模型，最终准确率、查准率、查全率和 F1 值分别达到 96.26%、91.54%、94.49% 和 92.83%，相比原始 DenseNet121 模型分别提升 6.48、9.13、6.44、9.33 个百分点。

2.1.3　SE－DenseNet－DC 分类模型可视化评价

混淆矩阵可视化地展示了残膜回收机复杂作业环境下作业场景分类任务中真实类和预测类各自所包含的实例数目[29]，据此推测正确分类和错误分类的图像样本数目，以作为混淆因素分析结果的有力数据对比。为进一步验证本研究所提出改进模型的有效性，对原模型 DenseNet121 和改进后模型 SE－DenseNet－DC 的测试集分类结果绘制混淆矩阵。如图 7 所示。

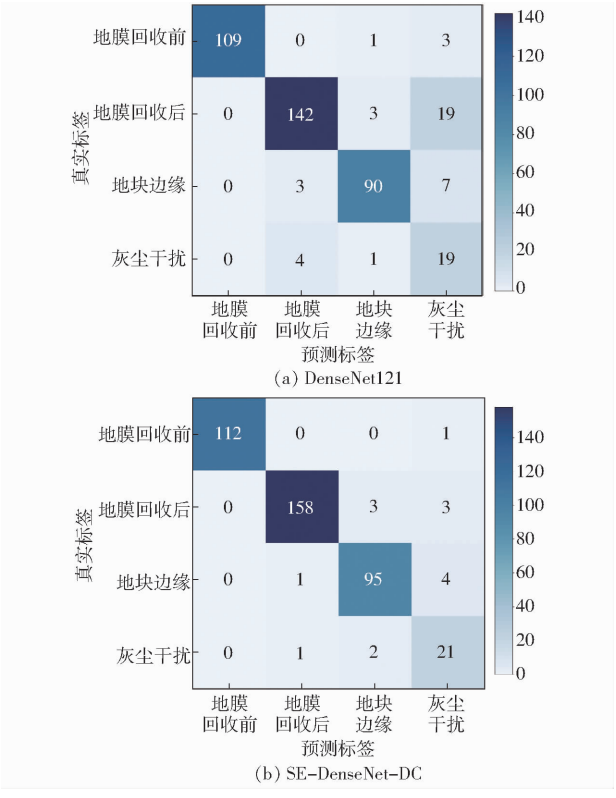

图 7　测试集分类的混淆矩阵

Fig. 7　Confusion matrix for classification of test set

由图 7a 可以看出，因地膜回收后目标场景图像与灰尘干扰场景图像之间特征较为相似，原始 DenseNet121 模型将目标场景图像误分为灰尘干扰场景图像的数量最多，且灰尘干扰场景图像被误分为目标场景图像的比率最大。而 SE－DenseNet－DC 模型（图 7b）极大改善了标签之间互相混淆的情况，错分图像总数由改进前的 41 幅下降到 15 幅；将目标场景图像识别为灰尘干扰场景图像出错数量从 19 幅降低到 3 幅；将灰尘干扰场景图像识别为目标场景图像出错数量从 4 幅降低到 1 幅。表明 SE－DenseNet－DC 分类模型能够更全面地学习残膜回收机实际作业过程中可能遇到的各种场景特征，针对识别目标图中存在多种相似非目标场景干扰，能够准确区分不同类别之间的差异，实现了对目标场景图像的有效提取，改善了非目标场景易被误识别的问题。

因此以主要混淆因素灰尘干扰图像样本为例，通过 Grad－cam 绘制热力图展示模型对图像特征的关注程度[30－31]，以进一步验证模型改进的有效性。如图 8 所示，原始 DenseNet121 模型难以有效区分不同类别之间的细微差异，导致难以准确聚焦到灰尘区域。而结合了通道注意力机制和空洞卷积的 SE－DenseNet－DC 模型具有强大的特征提取能力，可以捕捉到丰富的局部特征和全局上下文信息，聚焦区域集中在灰尘区域且更关注区域的细节特征。

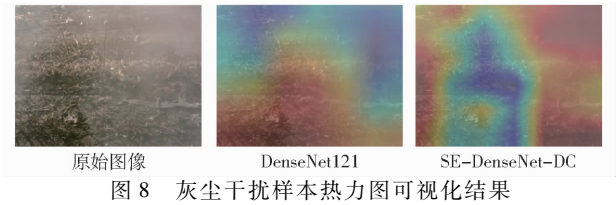

图 8　灰尘干扰样本热力图可视化结果

Fig. 8　Results of heat map visualization of dust interference samples

2.1.4　不同分类模型对比试验

为进一步验证 SE－DenseNet－DC 模型在作业场景分类任务中的有效性，在相同试验环境下，将 SE－DenseNet－DC 模型与经典分类模型 SwinTransformerV2[32]、MobileNetV3[33]、DenseNet201[34] 在数据增强后作业场景图像数据集上进行对比试验，试验结果如表 4 所示。

表 4　不同分类模型在测试集上的试验结果

Tab. 4　Test results of different classification models on test sets

模型	准确率/%	查准率/%	查全率/%	F1 值/%	单幅图像推理时间/s
SwinTransformerV2	58.38	55.14	63.52	56.29	0.199 7
MobileNetV3	64.84	60.60	67.68	60.92	0.067 8
DenseNet201	83.79	77.45	83.29	78.08	0.172 3
SE－DenseNet－DC	96.26	91.54	94.49	92.83	0.108 7

从表 4 可以看出,本文提出的 SE - DenseNet - DC 模型准确率达到 96.26%,查准率、查全率和 F1 值也超过其他网络模型。深层网络 DenseNet201 模型在本数据集上的性能表现较好,而 MobileNetV3 模型性能表现较差,说明浅层网络很难提取到本研究作业场景图像更深层次的特征,而 DenseNet201 模型网络复杂度较高、参数量较大,在数据量较少情况下过度拟合训练数据而不能达到 DenseNet121 模型的分类效果。由于作业场景图像数据量有限,视觉 Transformer 无法充分利用图像固有的尺度、平移不变性和特征局部性等先验知识,导致 SwinTransformerV2 模型各项指标明显低于 CNN 分类模型。此外,在实际应用中模型计算效率对识别任务有很大影响,本研究将模型分类单幅图像时间作为计算效率评估指标,SE - DenseNet - DC 模型平均耗时为 0.108 7 s,相比于 SwinTransformerV2 和 DenseNet201 模型分别低 0.091 s 和 0.063 6 s,与 MobileNetV3 相比,高 0.040 9 s,但准确率、查准率、查全率和 F1 值相比于 MobileNetV3 模型分别提高 31.42、30.94、26.81、31.91 个百分点。综合上述指标参数的评价结果,进一步表明本研究所提出的模型更适用于残膜回收机作业场景分类任务。

2.2 语义分割模型

2.2.1 基于 TransUnet 模型的消融试验

为验证各模块对模型的改进作用,在相同的试验环境下,以 TransUnet 模型为基准,对引入编码器部分的 CBAM 模块、跳跃连接中的 DAB 模块和解码器部分的 CCAF 模块,在本研究数据增强后的回收后地表残膜分割图像数据集上进行消融试验。试验结果如表 5 所示。

表 5 基于 TransUnet 模型的消融试验

Tab.5 Ablation tests based on the TransUnet model

序号	CBAM 模块	DAB 模块	CCAF 模块	准确率/%	精准率/%	召回率/%	MIOU/%
1	×	×	×	97.16	67.12	74.55	73.34
2	√	×	×	97.41	68.61	75.81	74.79
3	×	√	×	97.36	70.41	75.40	74.83
4	×	×	√	97.34	68.15	74.68	74.39
5	√	√	×	97.48	70.63	76.49	75.48
6	√	×	√	97.42	68.66	76.37	75.08
7	×	√	√	97.65	70.66	77.29	76.24
8	√	√	√	97.98	71.74	78.09	77.17

根据表 5 可以看出,与 TransUnet 模型相比,通过在编码器部分引入 CBAM 模块,实现了更加细微和精确的全局特征提取,准确率、精准率、召回率和 MIOU 分别提高 0.25、1.49、1.26、1.45 个百分点;在跳跃连接部分加入 DAB 模块,不仅能够增强不同尺度特征之间的关联,同时弥补了编码器和解码器之间的语义差距,准确率、精准率、召回率和 MIOU 分别提高 0.20、3.29、0.85、1.49 个百分点;在解码器部分加入 CCAF 模块,对特征进行通道级加权的同时有效融合了高层语义和低层细节特征,准确率、精准率、召回率和 MIOU 分别提高 0.18、1.03、0.13、1.05 个百分点;集成了 CBAM 模块、DAB 模块和 CCAF 模块的 TransUnet 模型,即本文提出的 CDC - TransUnet 模型在测试集上性能最佳,最终准确率、精准率、召回率和 MIOU 分别达到 97.98%、71.74%、78.09%、77.17%,相比原始 TransUnet 模型分别提升 0.82、4.62、3.54、3.83 个百分点。

2.2.2 不同分割模型对比试验

为验证本研究所构建分割模型在残膜图像分割方面的优势,在相同的试验环境下将 CDC - TransUnet 与 TransUnet[25]、HiFormer[35] 以及 TransCeption[36] 在数据增强后的回收后地表残膜分割图像数据集上进行比较,模型各项指标如表 6 所示。

表 6 不同分割模型在测试集上的试验结果

Tab.6 Trial results of different segmentation models on test set

模型	准确率/%	精准率/%	召回率/%	MIOU/%	单幅图像推理时间/s
TransUnet	97.16	67.12	74.55	73.34	0.412 7
HiFormer	97.05	67.01	73.62	72.02	0.403 5
TransCeption	97.11	66.51	74.26	72.86	0.387 2
CDC - TransUnet	97.98	71.74	78.09	77.17	0.432 6

从表 6 可以看出,CDC - TransUnet 模型准确率为 97.98%,相较于 TransUnet、HiFormer 和 TransCeption 模型分别提高 0.82、0.93、0.87 个百分点;精准率为 71.74%,较 TransUnet、HiFormer 和 TransCeption 模型分别提高 4.62、4.73、5.23 个百分点;召回率为 78.09%,相比于 TransUnet、HiFormer 和 TransCeption 模型分别提高 3.54、4.47、3.83 个百分点;MIOU 为 77.17%,相比于 TransUnet、HiFormer 和 TransCeptiont 模型分别提高 3.83、5.15、4.31 个百分点。CDC - TransUnet 模型平均耗时为 0.432 6 s,相比于其余模型略有增加,但由于在残膜图像分割任务中,分割精度要求比分割时间更关键,因此牺牲小部分分割速度选取最高分割精度的 CDC - TransUnet 模型。综上所述,本文提出的 CDC - TransUnet 模型在残膜图像分割方面具有更高的准确度和更好的特征提取能力,更适用于残膜图像分割任务。

为了准确评估不同分割模型在残膜回收机回收后地表残膜图像分割任务中的性能，进行了分割效果可视化对比，如图 9 所示。其中白色表示残膜部分，黑色表示背景部分。

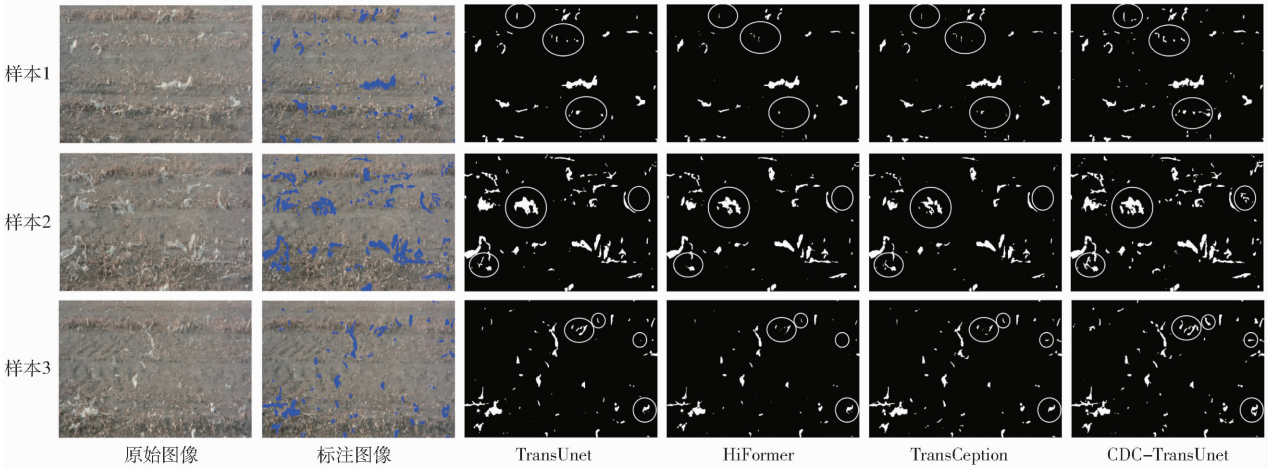

图 9　不同分割模型残膜分割效果

Fig. 9　Residual membrane segmentation effect of different segmentation models

从图 9 可以看出，所有算法均能够分割出地表残膜大致轮廓，但由于地表残膜尺寸各异、形状复杂多变，且存在大量性状相似的反光土块、棉铃壳和棉花等因素干扰，TransUnet、HiFormer 和 TransCeption 模型都出现了部分小目标漏识别、干扰因素误识别和较大块残膜部分边缘过分割和欠分割问题。相比之下，CDC－TransUnet 模型通过编码器部分的 CBAM 模块提取更细微和精确的全局特征，更好地保留了残膜边缘信息；通过跳跃连接部分的 DAB 模块融合多尺度特征增强了对残膜的整体和边缘细节定位能力；利用 CCAF 模块弥补上采样丢失的位置信息，对于残膜分割更为准确。3 个模块的集成使得本文所提出的分割模型在残膜的整体定位和边缘分割的情况更能吻合标注真值的分割标准。

2.2.3　残膜覆盖率预测结果验证

为验证本文所提出方法有效性，选取 10 块农田采集的共 200 幅图像数据进行模型验证，包含各类作业场景图像各 50 幅，串联改进的分类模型和语义分割模型，实现 200 幅图像数据中目标场景图像的有效提取并计算表层残膜覆盖率。基于语义分割模型对像素识别分类的结果，计算残膜回收机收后地表残膜覆盖率，即残膜像素点总数与图像总像素数之比[37]。

试验结果表明，50 幅目标场景图像中 48 幅被正确提取，模型分类准确率达到 96%，残膜覆盖率预测值与人工标注真实值决定系数（R^2）为 0.92，均方根误差（RMSE）为 0.23%，平均相对误差为 2.95%。单幅目标场景图像经过提取并计算表层残膜覆盖率评估时间平均为 0.54 s。残膜覆盖率预测值与真实值线性拟合结果如图 10 所示。结果表明，该方法在残膜回收机收后地表残膜覆盖率监测评估中具有较快的推理速度和较高的准确率，为残膜回收机回收质量实时准确评估提供一定技术支撑。

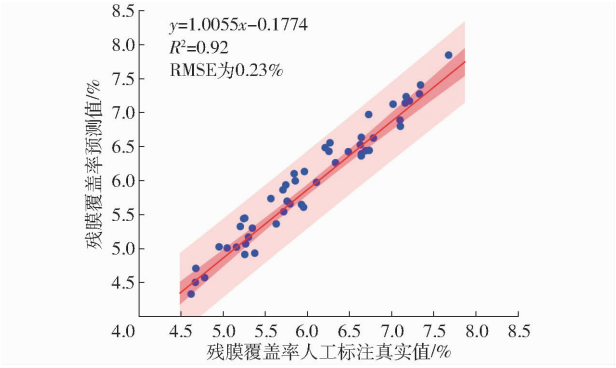

图 10　残膜覆盖率预测值与人工标注真实值拟合结果

Fig. 10　Result of fitting predicted value of residual film pollution with manual value

3　讨论

因图像数据采集的不确定性和作业环境的复杂性，3 类非目标场景干扰图像中均存在与目标场景图像特征相似的情况，常常存在非目标场景图像被误分类为目标场景图像的情况，图像中的灰尘、地块边缘和未收地膜严重影响残膜覆盖率的准确评估。传统模型难以准确区分不同类别之间的细微差异，导致分类错误率较高。因此本研究通过引入通道注意力机制和空洞卷积提高模型的特征提取能力。通过这 2 种方法的有效结合，模型能够更全面地学习残膜回收机实际作业过程中可能遇到的各种场景特征，从而显著提高了分类准确性。

对 TransUnet、HiFormer、TransCeption 和 CDC－TransUnet 模型在测试集上的性能进行了对比分析，分割结果准确率较高但 MIOU 相对较低。这主要是由于残膜回收机收后地表图像中残膜的像素数量相对较少，而背景像素（土地、秸秆、棉花和棉铃壳等）数量较

多,模型在预测背景时表现良好,导致高准确率。但由于 MIOU 是对所有类的平均交并比,模型在预测少数类(残膜)时表现不足导致整体 MIOU 值相对较低。

在数据采集过程中获取了地膜回收后场景图像在不同光照强度、角度及少量灰尘干扰下的图像样本,但并未详细讨论光照和灰尘干扰对图像的具体影响。下一步将重点分析这些环境因子对残膜覆盖率评估准确性的影响,并探索相应的补偿机制,以进一步完善研究。

残膜回收机回收后地表残膜在自然状态下会呈现不同程度的卷曲,对基于车载视觉成像的残膜覆盖率评估具有显著影响。在残膜覆盖率预测结果验证时并没有对地表残膜的展开面积进行详细讨论,下一步将引入卷曲指数来表征残膜卷曲程度,并构建预测模型,以提高残膜覆盖率评估的精度。

使用的图像数据集采集自覆膜棉田,但尚未对模型在其他作物覆膜农田中的泛化能力进行评估,下一步将拓展到其他作物的覆膜农田,收集更多多样化的数据来增强模型鲁棒性和适应性。

4 结论

(1)基于车载视觉成像系统,针对残膜回收机实际作业过程中存在多种相似非目标场景干扰,目标场景图像背景复杂且地表残膜尺寸小、破碎度大、无固定轮廓导致残膜覆盖率难以准确评估的问题,提出了一种基于车载成像和深度卷积神经网络的地表残膜识别方法,实现了作业场景的精准分类和收后地表残膜覆盖率的准确评估。

(2)利用通道注意力机制和空洞卷积融合深层神经网络构建的多重特征增强的 SE - DenseNet - DC 分类模型可以有效提取目标场景图像,改善非目标场景易被误分类的问题,分类准确率可达 96.26%。

(3)基于细节信息增强和多尺度特征融合的 CDC - TransUnet 分割模型可以精准分割目标场景图像复杂背景中尺寸小、破碎度大、无固定轮廓的地表残膜,MIOU 可达到 77.17%。

(4)残膜覆盖率预测结果表明,串联改进的分类模型和语义分割模型在残膜回收机收后地表残膜覆盖率监测评估中具有较快的推理速度和较高的准确率,决定系数 R^2 为 0.92,RMSE 为 0.23%,平均相对误差为 2.95%,单幅图像评估时间平均为 0.54 s。可为残膜回收机回收质量实时准确评估提供一定技术支撑。

参 考 文 献

[1] 中国纤维检验局. 中国棉花质量分析报告(2017/2018 年度)[J]. 中国棉麻产业经济研究, 2018(3): 10 - 18.
China Fiber Inspection Bureau. China Cotton Quality Analysis Report (2017/2018) [J]. Research on the Economic Development of China's Cotton and Hemp Industry, 2018(3): 10 - 18. (in Chinese)

[2] LI W, ZHAO J, ZHANG Z, et al. Uptake and effect of carboxyl-modified polystyrene microplastics on cotton plants[J]. Journal of Hazardous Materials, 2024, 466: 133581.

[3] 温浩军, 陈学庚, 陈浩, 等. 农田地膜回收机械应用现状与发展[J]. 农业环境科学学报, 2024, 43(6): 1271 - 1277.
WEN Haojun, CHEN Xuegeng, CHEN Hao, et al. Application status and development of farmland mulch recycling machinery [J]. Journal of Agro-Environment Science, 2024, 43(6): 1271 - 1277. (in Chinese)

[4] 梁荣庆, 陈学庚, 张炳成, 等. 新疆棉田残膜回收方式及资源化再利用现状问题与对策[J]. 农业工程学报, 2019, 35(16): 1 - 13.
LIANG Rongqing, CHEN Xuegeng, ZHANG Bingcheng, et al. Problems and countermeasures of recycling methods and resource reuse of residual film in cotton fields of Xinjiang[J]. Transactions of the CSAE, 2019, 35(16): 1 - 13. (in Chinese)

[5] 吴雪梅, 梁长江, 张大斌, 等. 基于无人机遥感影像的收获期后残膜识别方法[J]. 农业机械学报, 2020, 51(8): 189 - 195.
WU Xuemei, LIANG Changjiang, ZHANG Dabin, et al. Identification method of plastic film residue based on UAV remote sensing images[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(8): 189 - 195. (in Chinese)

[6] 宁纪锋, 倪静, 何宜家, 等. 基于卷积注意力的无人机多光谱遥感影像地膜农田识别[J]. 农业机械学报, 2021, 52(9): 213 - 220.
NING Jifeng, NI Jing, HE Yijia, et al. Convolutional attention based plastic mulching farmland identification via UAV multispectral remote sensing image[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 213 - 220. (in Chinese)

[7] YANG Q, LIU M, ZHANG Z, et al. Mapping plastic mulched farmland for high resolution images of unmanned aerial vehicle using deep semantic segmentation[J]. Remote Sensing, 2019, 11(17): 2008.

[8] 翟志强, 陈学庚, 邱发松, 等. 基于像素块和机器学习的播前棉田地表残膜覆盖率检测[J]. 农业工程学报, 2022, 38(6): 140 - 147.
ZHAI Zhiqiang, CHEN Xuegeng, QIU Fasong, et al. Detecting surface residual film coverage rate in pre-sowing cotton fields using pixel block and machine learning[J]. Transactions of the CSAE, 2022, 38(6): 140 - 147. (in Chinese)

[9] 张学军, 黄爽, 靳伟, 等. 基于改进 Faster R - CNN 的农田残膜识别方法[J]. 湖南大学学报(自然科学版), 2021, 48(8): 161 - 168.

ZHANG Xuejun, HUANG Shuang, JIN Wei, et al. Identification method of agricultural film residue based on improved Faster R－CNN[J]. Journal of Hunan University(Natural Sciences), 2021, 48(8): 161－168. (in Chinese)

[10] 江水泉，张海东，华英杰. 基于计算机视觉的农田残膜定位研究[J]. 中国农机化学报，2016，37(11)：150－154.
JIANG Shuiquan, ZHANG Haidong, HUA Yingjie. Research on agricultural film residue localization based on computer vision [J]. Journal of Chinese Agricultural Mechanization, 2016, 37(11): 150－154. (in Chinese)

[11] 王宝梁，沈文龙，张宝玉. 农用无人车环境感知技术发展现状及趋势分析[J]. 中国农机化学报，2021，42(11)：214－221.
WANG Baoliang, SHEN Wenlong, ZHANG Baoyu. Development status and trend of environmental perception technology for unmanned agricultural vehicles[J]. Journal of Chinese Agricultural Mechanization, 2021, 42(11): 214－221. (in Chinese)

[12] WANG S, ZHANG W, WANG X, et al. Recognition of rice seedling rows based on row vector grid classification[J]. Computers and Electronics in Agriculture, 2021, 190: 106454.

[13] RAHIM U F, UTSUMI T, MINENO H. Deep learning-based accurate grapevine inflorescence and flower quantification in unstructured vineyard images acquired using a mobile sensing platform[J]. Computers and Electronics in Agriculture, 2022, 198: 107088.

[14] BLOK P M, VAN EVERT F K, TIELEN A P M, et al. The effect of data augmentation and network simplification on the image-based detection of broccoli heads with Mask R－CNN[J]. Journal of Field Robotics, 2021, 38(1): 85－104.

[15] VAYSSADE J A, JONES G, GEE C, et al. Pixelwise instance segmentation of leaves in dense foliage[J]. Computers and Electronics in Agriculture, 2022, 195: 106797.

[16] MASSAH J, VAKILIAN K A, SHABANIAN M, et al. Design, development, and performance evaluation of a robot for yield estimation of kiwifruit[J]. Computers and Electronics in Agriculture, 2021, 185: 106132.

[17] 刘莫尘，高甜甜，马宗旭，等. 基于 MSRCR－YOLOv4－tiny 的田间玉米杂草检测模型[J]. 农业机械学报，2022，53(2)：246－255，335.
LIU Mochen, GAO Tiantian, MA Zongxu, et al. Target detection model of corn weeds in field environment based on MSRCR algorithm and YOLOv4－tiny[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(2): 246－255, 335 . (in Chinese)

[18] 权龙哲，夏福霖，姜伟，等. 基于 YOLO v4 卷积神经网络的农田苗草识别研究[J]. 东北农业大学学报，2021，52(7)：89－98.
QUAN Longzhe, XIA Fulin, JIANG Wei, et al. Research on recognition of maize seedlings and weeds in maize field based on YOLO v4 convolutional neural network[J]. Journal of Northeast Agricultural University, 2021, 52(7): 89－98. (in Chinese)

[19] HUANG G, LIU Z, VAN L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4700－4708.

[20] 李子茂，徐杰，郑禄，等. 基于改进 DenseNet 的茶叶病害小样本识别方法[J]. 农业工程学报，2022，38(10)：182－190.
LI Zimao, XU Jie, ZHENG Lu, et al. Small sample recognition method of tea disease based on improved DenseNet[J]. Transactions of the CSAE, 2022, 38(10): 182－190. (in Chinese)

[21] 宋玲，曹勉，胡小春，等. 基于 YOLOX 的复杂背景下木薯叶病害检测方法[J]. 农业机械学报，2023，54(3)：301－307.
SONG Ling, CAO Mian, HU Xiaochun, et al. Detection of cassava leaf diseases under complicated background based on YOLOX[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(3): 301－307. (in Chinese)

[22] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2018: 7132－7141.

[23] 马钦，万传峰，卫建，等. 基于 U－Net 和特征金字塔网络的秸秆覆盖率计算方法[J]. 农业机械学报，2023，54(1)：224－234.
MA Qin, WAN Chuanfeng, WEI Jian, et al. Calculation method of straw coverage based on U－Net network and feature pyramid network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(1): 224－234. (in Chinese)

[24] 张寅，朱桂熠，施天俊，等. 基于特征融合与注意力的遥感图像小目标检测[J]. 光学学报，2022，42(24)：140－150.
ZHANG Yin, ZHU Guiyi, SHI Tianjun, et al. Small object detection in remote sensing images based on feature fusion and attention[J]. Acta Optica Sinica, 2022, 42(24): 140－150. (in Chinese)

[25] CHEN J, LU Y, YU Q, et al. Transunet: transformers make strong encoders for medical image segmentation[J]. arXiv Preprint, arXiv:2102.04306, 2021.

[26] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3－19.

[27] SUN G, PAN Y, KONG W, et al. DA－TransUNet: integrating spatial and channel dual attention with transformer U－net for medical image segmentation[J]. Frontiers in Bioengineering and Biotechnology, 2024, 12: 1398237.

[28] 赵亮，刘晨，王春艳. 位置信息增强的 TransUnet 医学图像分割方法[J/OL]. 计算机科学与探索. https://link.cnki.net/urlid/11.5602.tp.20241016.1348.002.
ZHAO Liang, LIU Chen, WANG Chunyan. Positional enhancement TransUnet for medical image segmentation[J/OL]. Journal of Frontiers of Computer Science and Technology. https://link.cnki.net/urlid/11.5602.tp.20241016.1348.002. (in Chinese)

(下转第 70 页)

[27] 赵淑红,陈君执,王加一,等. 精量播种机 V 型凹槽拨轮式导种部件设计与试验[J]. 农业机械学报, 2018, 49(6): 146 - 158.
ZHAO Shuhong, CHEN Junzhi, WANG Jiayi, et al. Design and experiment on V-groove dialing round type guiding-seed device [J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(6): 146 - 158. (in Chinese)

[28] 丁幼春,朱凯,王凯阳,等. 薄面激光-硅光电池中小粒径种子流监测装置研制[J]. 农业工程学报, 2019, 35(8): 12 - 20.
DING Youchun, ZHU Kai, WANG Kaiyang, et al. Development of monitoring device for medium and small size seed flow based on thin surface laser-silicon photocell[J]. Transactions of the CSAE, 2019, 35(8): 12 - 20. (in Chinese)

[29] 丁幼春,王凯阳,杜超群,等. 高通量小粒径种子流检测装置设计与试验[J]. 农业工程学报, 2020, 36(13): 20 - 28.
DING Youchun, WANG Kaiyang, DU Chaoqun, et al. Design and experiment of high-flux small-size seed flow detection device[J]. Transactions of the CSAE, 2020, 36(13): 20 - 28. (in Chinese)

[30] 丁幼春,陈礼源,董万静,等. 油菜宽幅播种作业监测系统设计与试验[J]. 农业工程学报, 2021, 37(13): 38 - 45.
DING Youchun, CHEN Liyuan, DONG Wanjing, et al. Design and experiment of the sowing monitoring system for a wide-width rapeseed planter[J]. Transactions of the CSAE, 2021, 37(13): 38 - 45. (in Chinese)

[31] 李道亮,杨昊. 农业物联网技术研究进展与发展趋势分析[J]. 农业机械学报, 2018, 49(1): 1 - 20.
LI Daoliang, YANG Hao. State-of-the-art review for internet of things in agriculture[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(1): 1 - 20. (in Chinese)

[32] 毛文华,韩少云,赵博,等. 基于 Netty 和 Marshalling 的青饲机工况远程监测系统研究[J]. 农业机械学报, 2020, 51(8): 145 - 151.
MAO Wenhua, HAN Shaoyun, ZHAO Bo, et al. Study of remote monitoring system for silage harvester working condition based on Netty and Marshalling[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(8): 145 - 151. (in Chinese)

[33] 毛天宇,王星宇,常瑞,等. 面向 Java 语言生态的软件供应链安全分析技术[J]. 软件学报, 2023, 34(6): 2628 - 2640.
MAO Tianyu, WANG Xingyu, CHANG Rui, et al. Software supply chain analysis techniques for Java ecosystem[J]. Journal of Software, 2023, 34(6): 2628 - 2640. (in Chinese)

[34] 张玉清,贾岩,雷柯楠,等. HTML5 新特性安全研究综述[J]. 计算机研究与发展, 2016, 53(10): 2163 - 2172.
ZHANG Yuqing, JIA Yan, LEI Ke'nan, et al. Survey of HTML5 new features security[J]. Journal of Computer Research and Development, 2016, 53(10): 2163 - 2172. (in Chinese)

(上接第 37 页)

[29] 张彦斐, 魏鹏, 宫金良, 等. 复杂环境中苹果树识别与导航线提取方法[J]. 农业机械学报, 2022, 53(10): 220 - 227.
ZHANG Yanfei, WEI Peng, GONG Jinliang, et al. Apple tree recognition and navigation line extraction in complex environment[J]. Transactions of the Chinese Society for Agricultural Machinery, 2022, 53(10): 220 - 227. (in Chinese)

[30] SELVARAJU R R, COGSWELL M, DAS A, et al. Grad - CAM: visual explanations from deep networks via gradient-based localization[J]. International Journal of Computer Vision, 2020, 128(2): 336 - 359.

[31] 高霁月, 倪建功, 杨昊岩, 等. 基于数据平衡和深度学习的开心果品质视觉检测方法[J]. 农业机械学报, 2021, 52(7): 367 - 372.
GAO Jiyue, NI Jiangong, YANG Haoyan, et al. Pistachio visual detection based on data balance and deep learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7): 367 - 372. (in Chinese)

[32] LIU Z, HU H, LIN Y, et al. Swin transformer v2: scaling up capacity and resolution[C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2022: 12009 - 12019.

[33] 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.

[34] HUANG G, LIU Z, VAN L, et al. Densely connected convolutional networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4700 - 4708.

[35] HEIDARI M, KAZEROUNI A, SOLTANY M, et al. Hiformer: hierarchical multi-scale representations using transformers for medical image segmentation[C]// Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2023: 6202 - 6212.

[36] AZAD R, JIA Y, AGHDAM E K, et al. Enhancing medical image segmentation with TransCeption: a mul-ti-scale feature fusion approach[J]. arXiv Preprint, arXiv:2301.10847, 2023.

[37] YANG J, ZHAI Z, LI Y, et al. Design and research of residual film pollution monitoring system based on UAV[J]. Computers and Electronics in Agriculture, 2024, 217: 108608.