



河北大学
HEBEI UNIVERSITY

密 级:

分 类 号:

学校代码: 10075

学 号: 20207014014

硕士学位论文

无人机航拍图像中的罂粟植株语义分割算法

学 位 申 请 人: 罗梓菲

指 导 教 师: 杨文柱 教授

郑新永 高级工程师

专业学位类别: 电子信息硕士

专业学位领域: 计算机技术

院 系 名 称: 网络空间安全与计算机学院

答 辩 日 期: 二〇二三年五月

Classified Index:

CODE:10075

U.D.C:

No:20207014014

Thesis for the Degree of Master

Semantic Segmentation of Poppy Plants in UAV Images

Candidate: Luo Zifei

Supervisor: Prof. Yang Wenzhu

SE. Zheng Xinyong

**Category of Academic Degree: Master of Electronic and Information
Engineering**

Specialty: Computer Technology

College: School of Cyber Security and Computer

Date of Oral Defense: May, 2023

摘要

罂粟属于监管药物，严禁非法种植；由于非法种植地往往较为隐蔽，使得监管难度极大；利用无人机航拍实现对罂粟非法种植监管是目前执法部门常用的手段。为提高在航拍图像中识别罂粟植株的准确率，设计了无人机航拍场景下罂粟图像的语义分割算法。主要研究内容包括：

（1）基于敏感特征的数据增广方法

针对航拍图像中含罂粟植株的样本过少问题，以罂粟植株的显著特征（叶片纹理、果实形状等）为依据，混合随机窗口切割和旋转增广方法，利用 ExG（超绿指数）、CIVE（植被颜色指数）以及 GB、COM（混合植被因子）对 RGB 图像进行增广处理。经过敏感特征增广后的数据集数据量达 1536 张。实验结果表明，基于敏感特征增广算法在 U-Net 网络框架上 Dice 指数为 0.74，远高于基于切割、旋转的增广算法的指标。

（2）基于混合注意力的编解码语义分割算法

针对无人机拍摄的罂粟图像的尺度变化问题，设计了基于 Transformer 的改进 U 型网络--TAU-Net。在 Encoder 阶段将特征图序列化，线性投影后输入由多个 Transformer 层组成的混合注意力模块，保留全局信息，通过权重矩阵关注关键信息，抑制无关信息，使网络关注输出相应像素级的 labelmap。针对罂粟果实与植株根茎之间的特殊植物学形态的问题，在 U-Net 的编码阶段使用相对位置编码（RPE）对多头注意力模块(MHA)进行了改进。利用依靠可学习参数的查找表对 tokens 之间的相对距离进行编码，捕捉远程依赖关系。实验结果表明，TAU-Net 网络的 Dice 指数达到 0.77，高于 U-Net、DeepLabV3+ 模型在罂粟数据集上的指标。

关键词 非法种植监管 语义分割 小样本数据增广 编解码网络 自注意力机制

Abstract

Poppy is a regulated drug and illicit cultivation is strictly prohibited. Because illegal cultivation sites are often more hidden, making supervision extremely difficult. The use of UAV aerial photography to achieve regulation of illicit poppy cultivation is a common tool currently used by enforcement agencies. In order to improve the accuracy of identifying poppy plants in aerial images, a semantic segmentation algorithm for poppy images in UAV aerial photography scenes is proposed. The main research includes the following algorithms.

(1) Data augmentation method based on hybrid strategy

To address the problem of small poppy plant samples in aerial photography images, RGB images are processed using ExG (Excess Green Index), CIVE (Color Index of Vegetation Extraction) and COM vegetation index based on the distinctive features of poppies (leaf texture, fruit shape, etc.), mixed with random window cutting and rotational augmentation methods. The volume of data in the dataset after the hybrid strategy is 1536 images. The experimental results show that the Dice index of the hybrid strategy-based broadening algorithm is 0.74 on the U-Net network framework, which is much higher than the index of the cut and rotation-based augmentation algorithm.

(2) Hybrid attention-based encoder-decoder semantic segmentation algorithm

For the scale variation problem of poppy images captured by UAV, a Transformer-based U-shaped network TAU-Net is designed to solve the scale variation. The feature map is serialized in Encoder stage and input into a hybrid attention module consisting of multiple Transformer layers after linear projection, preserving the global information. By focusing on key information and suppressing irrelevant information through the weight matrix, the network focuses on outputting the pixel-level labelmap. in the experimental results. The Dice index of TAU-Net network reaches 0.77, which is higher than the index of U-Net and DeepLabV3+ models on the poppy dataset.

For the problem of special botanical morphology between poppy fruits and plant rhizomes, the multi-headed attention module (MHA) was improved using relative position encoding (RPE) in the encoding stage of U-Net. The relative distance between tokens is encoded using a look-up table that relies on learnable parameters to capture remote dependencies. The improved hybrid attention module using relative position encoding improved the network Dice index by 0.03.

Keywords Illegal species monitoring Semantic segmentation Small sample data augmentation Encoder-decoder network Self-attentive mechanism

目 录

第一章 绪论.....	1
1.1 研究背景及意义.....	1
1.2 国内外研究现状.....	2
1.2.1 小样本数据农业图像增广技术研究现状	2
1.2.2 农业图像语义分割算法研究现状.....	3
1.2.3 注意力算法研究现状.....	11
1.3 研究内容及创新点.....	13
1.3.1 主要研究内容.....	13
1.3.2 创新点.....	15
1.4 文章的组织结构.....	15
第二章 相关理论	17
2.1 小样本数据增广	17
2.2 基于编解码结构的语义分割算法	18
2.3 注意力机制.....	19
第三章 数据增广方法	21
3.1 引言.....	21
3.2 基于敏感特征的数据增广方法	21
3.2.1 基于 ExG、CIVE 植被因子增广方法.....	23
3.2.2 基于 GB、COM 植被因子增广方法.....	24
3.3 实验结果与分析.....	25
3.3.1 数据集.....	25
3.3.2 评价指标.....	26
3.3.3 实验设计.....	26
3.3.4 实验结果与讨论.....	26
3.4 本章小结.....	30
第四章 基于注意力的编解码语义分割算法	31
4.1 引言.....	31
4.2 系统的整体架构.....	31
4.3 基于 Transformer 的编解码网络.....	32
4.4 多头注意力模块.....	33
4.5 实验结果与分析.....	35
4.5.1 数据集.....	35
4.5.2 评价指标.....	36
4.5.3 实验设计.....	36

4.5.4 实验结果与讨论.....	37
4.6 本章小结.....	40
第五章 总结与展望	41
5.1 研究总结.....	41
5.2 工作展望.....	42
参考文献.....	45

第一章 绪 论

1.1 研究背景及意义

罂粟是一种一年生草本，从蒴果上提取的汁液，可加工成鸦片、吗啡、海洛因。因此，鸦片罂粟成为世界上毒品的重要根源。如今，违法种植罂粟的违法现象在我国并未断绝，为了减少私种罂粟给社会带来的不良影响及极大的潜在风险，发现并惩戒违法罂粟种植迫在眉睫。无人机在协助公安机关打击违法种植罂粟行动中发挥了重要作用。当前，有关警用无人机在禁种铲毒领域的研究尚不深入，难以回应实践层面的理论需求。因此，有必要在各地利用无人机探查罂粟种植情况，加强警用无人机打击非法种植罂粟效能，更好地为禁毒实战服务。使用无人机从高空拍摄高清图像，隐蔽性高，实现了隐蔽侦查。

从机器学习到实现人工智能，深度学习技术不可或缺。深度学习起源于多层感知器，根据不同的分类任务自主地提取整合低级特征到抽象的高级特征，实现机器对数据的分布式特征认知，这使它们能够处理复杂和多样化的应用场景^[1]。在其他领域的巨大成功促使人们将传统方法与深度学习方法相结合，以解决农业领域的问题。

图像语义分割是机器视觉领域中的一个重要方向，是对每个像素进行标签分类的一项基础任务，它是机器视觉应用的一个重要组成部分。机器视觉图像语义分割技术被广泛应用于各个领域，可以分为三种类型：基于传统方法的语义分割、基于深度学习的语义分割、将深度学习与传统方法相结合的语义分割。其中，深度学习与传统方法相结合的语义分割算法在精确度等方面远超于传统的图像分割方法^[2]。本研究利用深度学习技术，结合图像分割技术，对无人机所拍摄的高空无人机图像中的处于花期及结果期的罂粟植株图像使用改进的深度网络模型进行训练，分割出罂粟和其他合法种植植物，有效判别无人机拍摄区域是否存在非法种植罂粟的违法行为，如果存在罂粟的非法种植情况，通过利用深度网络得到的分割结果估计罂粟的种植规模。图像中干扰物造成的分割错误问题一直以来都是图像分割领域的难点问题。图像中存在干扰物，如树冠、房檐或其他相似植物等^[3]，这导致网络对罂粟植株辨识度低，容易导致分割错误的问题。因此不断提升图像分割的正确率是本次研究的研究重点。

1.2 国内外研究现状

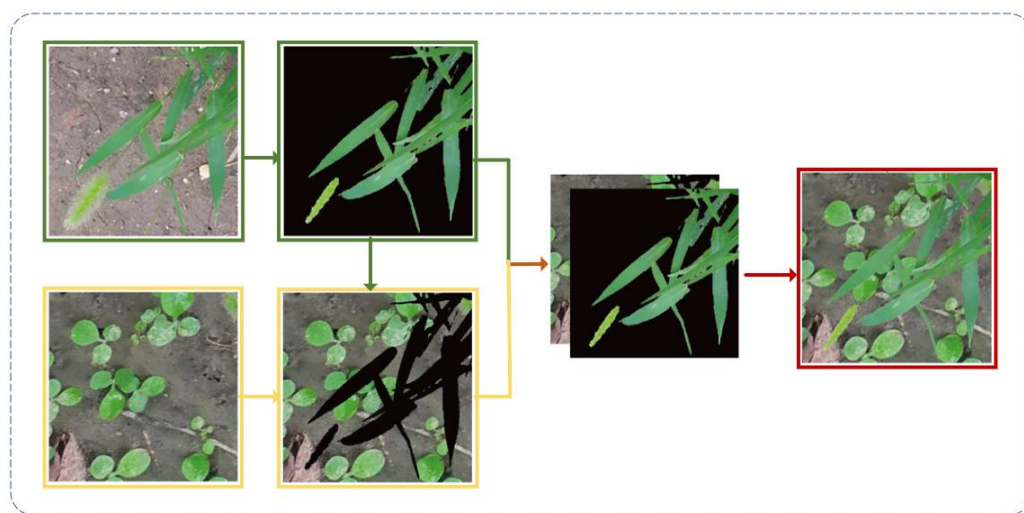
近年来,目标检测、图像分类、图像分割以及识别等方面的应用越来越广泛,其中,语义分割可以保留细粒度图像中目标区域边缘空间信息,在像素尺度上,以相似性为原则进行分类。在以场景理解为基础的任务中,语义分割算法不可或缺。例如,自动驾驶车辆的车道分析和无人机车辆的地理定位^[4]。随着语义分割技术的不断发展,它也被广泛应用用于解决多样化和数据丰富的遥感问题^[5]。包括环境监测^[6],作物覆盖与类型分析^[7],森林管理^[8],城市空间的建筑分类与土地利用分析^{[9][10]}。

1.2.1 小样本数据农业图像增广技术研究现状

在农业图像任务中使用监督方法的缺点之一是缺乏足够的标记数据集,这往往会影响网络的训练过程^[11]。例如,在植物病害和虫害的分割中,由于病例的发病条件不同,一些病变图像样本稀少,在实际项目中很难有足够的数据来支持训练。

在小样本问题上,传统方法通常在现有数据上进行几何变换或颜色变换,然而,它们并没有实质性地增加数据集。为解决这个问题,Potena 等人^[12]提出一个基于 RGB+NIR 图像的多步骤视觉系统用于作物和杂草分类的多步骤视觉系统。系统使用两种不同的 CNN 架构,一个浅层网络执行植被探测,然后使用一个更深的网络执行像素分类,将检测到的植被进一步区分为农作物和杂草。Lottes 等人^[13]通过使用作物排列信息解决小样本问题,该方法利用作物排列信息,例如从播种的信息作为先验知识,并利用半监督的随机森林方法来解决泛化问题。图像合成作为生成模型的经典思想,也可以用来增加输入数据。对于杂草的分割^[14],杂草图像被用作"前景",其他图像被用作"背景"。这些带有"前景"和"背景"的合成图像被用于预训练,以增加数据集并减少人工标注的样本数量,如图 1-1 所示。

小样本的问题也可以用基于生成对抗网络的图像生成方法^[15]来解决。生成对抗网络 (GAN)^[16]是一种无监督对抗学习方法,由一个生成网络和一个判别网络组成。生成网络判别网络进行对抗学习训练,目标是产生与真实数据分布一致的样本,扩大现有的数据集,混合数据集可以增强模型的学习能力。与现有的 GANs 相比,其变异模型,如:DCGAN (深度卷积生成对抗网络)^[17]、C-GAN (条件生成对抗网络)^[18]、AR-GAN (激活重构生成对抗网络)^[19]有更好的表现。

图 1-1 杂草数据集的合成过程^[14]

当 C-GAN 被用来增强番茄的叶子图像数据时^[18]，过拟合现象可以得到有效抑制。此外，为了合成更真实的西红柿叶子图像，AR-GAN^[19]可以保留掩膜的语义信息。上述图像增强策略也可用于其他小样本应用场景，以达到更好的分割效果。

1.2.2 农业图像语义分割算法研究现状

在过去的几年里，农业自动化发展迅速，基于机器视觉的方法已经推动了农业产业的进步。语义分割是机器视觉中实现像素级分类的基本任务之一。与对图像进行预测不同，语义分割产生了嵌入在物体空间信息中的像素级描述。随着语义分割方法的进步，基于视觉的技术已被用于解决多样性和数据丰富的农业问题，如作物覆盖和类型分析^{[20]-[22]}、森林种植^[23]、杂草分割^[24]、预测性农业^[25]、虫害和疾病识别^[26]，这有助于降低劳动成本和时间消耗。

语义分割技术已经取代了传统的人工观察和测量表型数据，在现代农业中发挥了极其重要的作用。例如，在温室耕作中^[27]，语义分割技术被用来监测农作物的生长状态。通过预测叶冠面积和植被高度，从水果表面提取像素点以确定水果的成熟度，监测农作物播种过程以及生长状态。在正确的时间收获作物以达到最佳成熟度来提高作物生产质量。在害虫分割方面^[28]，通过提取图像中昆虫的纹理、形状和大小来实现害虫识别是可行的，这有助于控制杀虫剂的使用。此外，农场分割对于自动导航是至关重要的^[29]，但由于地面和近地面环境的非均匀性问题，随着空间尺度的增大，在一些宏观信息的获取和参数的区域化方面存在很多困难，语义分割技术可以有效地提取地面信息。

然而,由于环境复杂、季节和光照变化等原因,特征提取和图像分割工作更加复杂。在深度学习开始研究之前,传统的分割方法对复杂环境的鲁棒性通过色彩空间转换和色彩通道的组合得到了极大的提升。对于极端的光照条件以及尖锐的阴影边缘,基于植被指数的方法分别利用独立的颜色通道来处理像素的颜色和亮度^[30]。其他方法使用基于机器学习的分类技术,如决策树^{[31]-[33]}、支持向量机^[34]。但传统的分割方法泛化能力弱,深度学习方法根据不同的分类任务自主地提取特征,这使它们能够处理复杂和多样化的应用场景。

1.2.2.1 基于传统方法的语义分割算法研究现状

在神经网络模型兴起之前,有许多传统的方法被设计用来解决农业图像分类和语义分割的问题。具有代表性的传统算法包括基于阈值的方法、基于聚类的方法、小波变换和随机森林等^{[35]-[38]}。本节将介绍一些典型的农业图像传统语义分割方法的主要思路和应用,如基于阈值的分割、基于区域的分割等。近年来被广泛用于农业自动化的传统语义分割过程如图 1-2 所示。

(1) 基于阈值的语义分割

阈值法是图像分割的经典方法之一,它利用分割目标与背景之间的灰度差异,设置不同的灰度阈值,对图像灰度直方图进行分类。通过判断图像中各像素点的特征属性,认为灰度值在同一灰度范围内的像素属于同一类别,具有一定的相似性。

阈值分割方法的关键是找到最佳阈值来准确分割图像。大津(Otsu)方法^[39]是一种全局阈值方法,是利用阈值进行图像分割的一个经典算法。大津算法中用来衡量差异的标准是最常见的类间差异。当阈值分割使类间方差最大时,意味着错误分类的概率最小。然而,当光照不均匀时,很难找到一个合适的单一阈值来准确分割图像。自适应阈值方法^[40]在一定程度上解决了这个问题,它根据图像不同区域的亮度分布,计算邻域的均值、中值和高斯加权平均值来确定局部阈值。

在过去的十年中,许多工作都是利用机器学习技术进行农业图像分割的。一方面,在植被图像分割中,各种基于阈值的方法遵循植被和图像中其他物体之间的平等设定,这在植物生长的不同阶段并不完全有效^[41]。另一方面,将 RGB 值转换为灰度的彩色指数技术并不总是能精准地辨别灰度图像,这是由照明条件对 RGB 图像值的影响造成的。

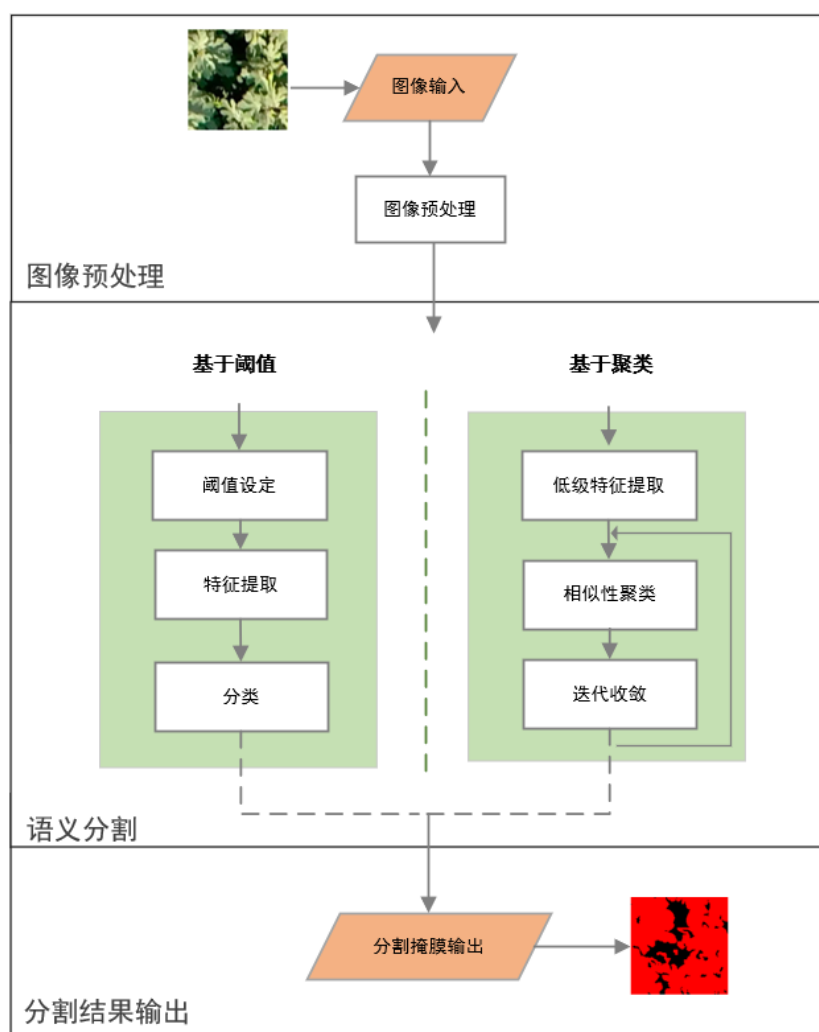


图 1-2 基于传统方法的语义分割过程

因此，尽管基于颜色指数和阈值指数的方法提供了简单和先进的技术，但由于光照变化条件对取得的植被分割结果的影响，它们有一定的局限性，特别是在晴天和阴天的情况下图像分割差别更大。

对于柑橘类水果图像的分割，在将原始 RGB 图像转换为 HSV 图像后，使用了分水岭算法，得到的 R^2 值为 0.93^[42]。为了检测适当的阈值，使用了 HSV 颜色空间，然后提取色调通道以生成色调直方图，使用过滤的色调直方图和色调直方图的拟合高斯曲线来检测适当的阈值^[43]。该方法可以产生准确和稳定的植被分割性能，平均准确率为 87.29%，标准偏差为 12.5%^[41]。

(2) 基于聚类的语义分割方法

基于聚类的语义分割方法是将具有相似特征的像素点聚集到同一区域，经过反复选

代和聚类直至收敛,得到分割结果。从计算机视觉的角度来看,超像素属于过分割,它是由具有相似特征(颜色、纹理等)的相邻像素组成的像素块。像素特征之间的相似性被用来对像素进行分组,用超像素来代表图像特征。这样就可以保留边缘信息,大大降低图像后期处理的复杂性,提高计算速度。超像素经常被用作图像分割算法的预处理步骤。这项技术已被广泛用于计算机视觉应用,包括姿势估计、目标识别和图像分割等。

简单线性迭代聚类(SLIC)算法^[44]是由 K-means 聚类算法扩展而来的高效超像素生成算法,近年来,人们研究了更多基于 SLIC 的超像素算法,这些算法提高了边界附着能力^{[45][46]}。此外,一个自适应的 SLIC (ASLIC)^[45]方法使用每个聚类的最大距离来动态归一化颜色和空间接近度,在速度上表现良好,但在弱边界上表现不佳。

为了更好地适应弱边界,对 SLIC 进行多级分层处理,用基于局部信息的简单线性迭代聚类方法(LI-SLI)对原始图像进行分割^[47]。属于同一物体的相邻超级像素根据概率分布的相似性进行合并,并减少超级像素的数量。与其他一些聚类方法相比,LI-SLI 的鲁棒性得到了改善,但速度较慢。

当 SLIC 方法用于实际的农业应用场景时,在自然环境中获得的图像的阴影和不均匀的光照往往会给分割带来很大的困难。例如,在对植物病害图像进行分割时,土壤的颜色与植物感染早/晚疫区的颜色几乎相同,而纹理则不同,这种特性给分割过程带来了负面影响^[48]。为了解决这个问题,SLIC 对原始图像进行了颜色均衡和超像素运算,以消除光照不均的影响,并在植物-村落数据集上取得了 97.2%的准确率^[49]。

为了消除各种光照的影响,SLIC 与基于图形的分割算法相结合,被用来生成区域建议,并提出了针对每种强度的集合神经网络的颜色模型,达到了 74.3%的平均准确率^[50]。超像素技术不仅可以作为图像预处理阶段的优化方法,还可以与深度学习方法相结合进行图像分割。

近年来,结合超级像素的深度学习方法得到了广泛的研究,如 Segmentation Aware Loss (SEAL)^[51]、超像素采样网络(SSN)^[52]和全卷积网络的超级像素分割(FCN)^[53]。当使用 SLIC 算法与深度学习技术相结合对动物胴体进行分割时,首先使用 SLIC 算法对动物胴体图像形成超像素。然后,这些超像素被手动标记为不同类别的物体(胴体、背景),这种超像素+CNN 的结构在胴体图像数据集中获得了优秀的分割结果。然而,超像素数量 k 的选择对分割结果有很大影响。例如,当 k=100 和 1000 时,像素准确率分

别达到 84.3%和 96.1%，IoU（Intersection over Union）分别为 77.3%和 92.2%^[54]。

1.2.2.2 基于深度学习方法的分割算法研究现状

由于缺乏泛化能力，传统的分割方法很难在实际应用中适应模型。深度学习方法根据不同的分类任务自主地提取特征，这使它们能够处理复杂和多样化的应用场景。在其他领域的巨大成功促使人们将传统的分割方法与深度学习方法相结合，以解决农业领域的问题^[55]。新开发的深度学习模型，包括 VGG、FCN^[56]、SegNet、U-Net、DeepLab 等，现在经常被用来实现像素级分割。深度学习方法由于其优越的性能，在农业视觉领域被广泛使用。它们的应用包括分类、检测和语义分割等。本文将基于深度学习的语义分割过程总结为三个阶段：特征提取、语义分割和后处理，如图 1-3 所示。本节介绍了基于深度学习方法的代表性分割网络架构。

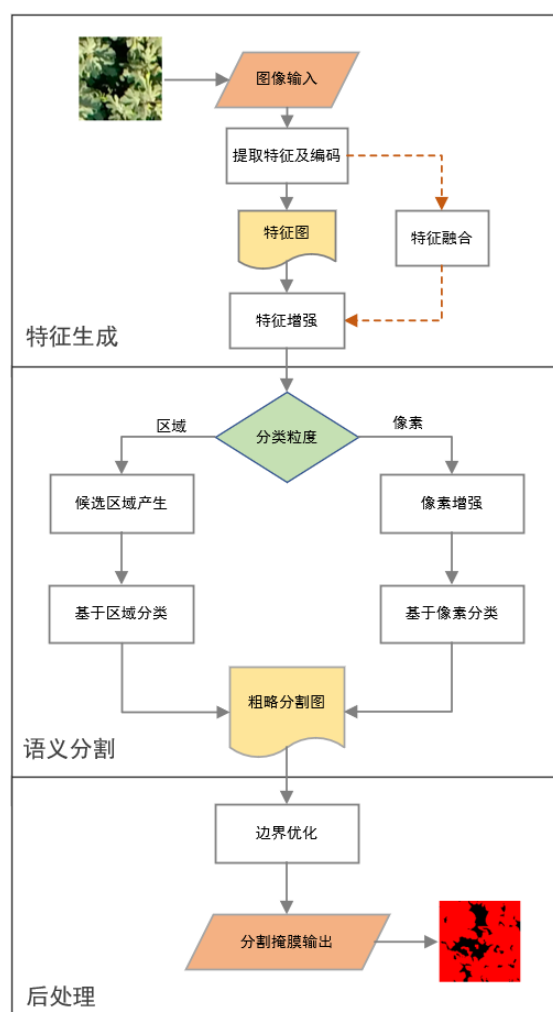


图 1-3 基于深度学习的语义分割过程

(1) SegNet

SegNet^[57]是在人工智能领域中表现出色的语义分割深度网络，是编解码器网络的经典模型之一。它遵循 FCN 架构，骨干网络由 VGG16 组成。SegNet 的结构如图 1-4 所示。

SegNet 有两个版本，即 SegNet^[58]和 Bayesian SegNet^[59]。SegNet 的作者还根据网络的深度提供了一个浅显的 BASIC 版本。从图 1-4 可以看出，SegNet 的结构与 U-Net 的结构非常相似，不同的是，SegNet 使用池化索引来恢复特征图的原始位置。

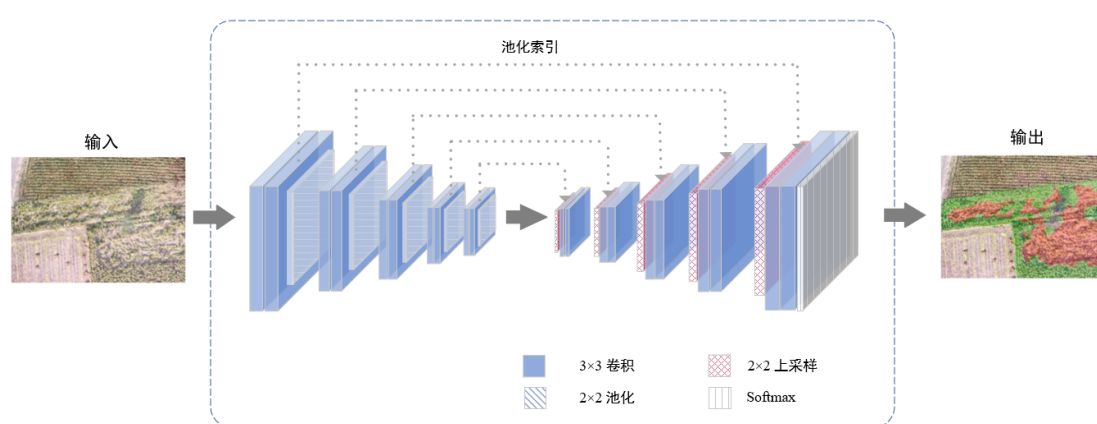


图 1-4 SegNet 及其在向日葵无人机图像分割中的应用^[58]

表 1-1 总结了使用 SegNet 作为骨干网络的分割方案，以及它们的改进和实验结果。虽然 SegNet 收敛速度快，但它并没有充分考虑像素与像素之间的关系。在农业领域，SegNet 在高分辨率遥感图像（如耕地图像）的大体积目标提取方面更具优势^[60]，也成功地应用于果园三维可视化和水果采摘视觉系统、点云等领域。

表 1-1 分割场景及其改进和使用 SegNet 作为骨干网络的实验结果

应用场景	方法	结果
向日葵种植区的划分 ^[58]	条件随机场（CRF）和图像融合被用来提高模型性能	89.8% (IoU)
Landsat-8 卫星图像分割 ^[74]	通过对初始化的 CNN 预训练模型进行有效的迁移学习程序，对森林图像进行微调以提高其泛化能力	82.1% (IoU)

(2) DeepLab

DeepLab^{[61]-[64]}通过采用空洞卷积来改进 FCN。空洞卷积通过增加空洞来增加感受野，因此卷积输出包含更大范围的信息并保留了图像的空间特征。

DeepLabV1^[61]和 DeepLabV2^[62]是具有卓越性能的端到端框架。它们采用 VGG

(DeepLabV1) 和 ResNet (DeepLabV2) 作为编码器, 同时使用反转卷积和全连接条件随机场 (CRF) 来提高模型捕捉细节的能力。

最初的 DeepLabV3^{[63][64]}继承了 DeepLabV1、DeepLabV2, 仍然采取 ResNet^[65]作为其主干。然而, ResNet 的最后一个区块被修改, 增加了 Atrous 卷积, 使用 ASPP (Atrous Spatial Pyramid Pooling) 而不是 CRF 来捕捉不同尺度的背景信息。ASPP 增强了提取密集特征的能力。然而, 由于池化层和卷积层的存在, 被分割目标的边界信息被严重丢失。相对而言, DeepLabV3+则从 Xception 模型中提取的深度可分离卷积, 对每个通道的空间进行单独卷积, 并取得了更高的精度。DeepLabV3+^[66]的结构如图 1-5 所示。

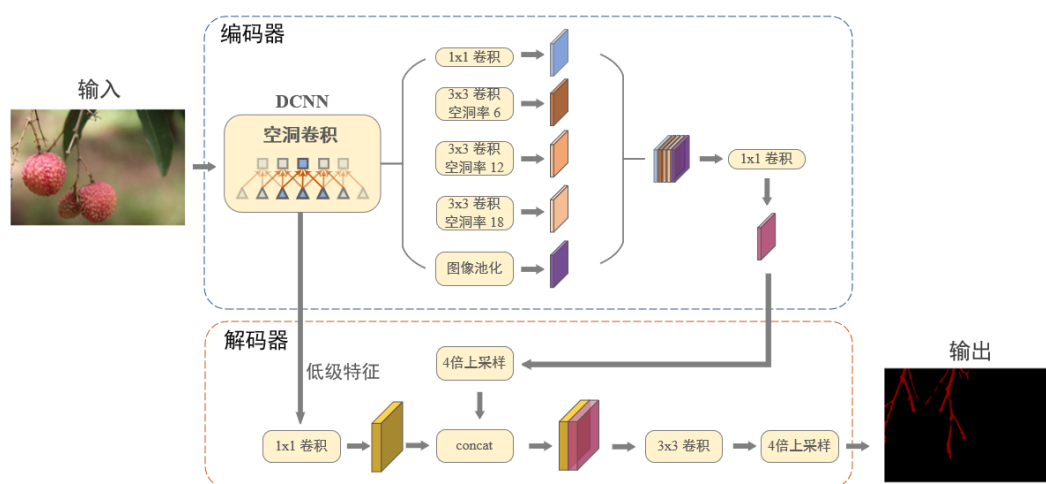


图 1-5 DeepLabV3+ 及其在荔枝枝干图像分割中的应用^[66]

DeepLabV3+ 网络在提取更加密集的特征方面表现出色^[67], 同时, DeepLabV3+ 在很大程度上扩大了感受野, 获得了多尺度的全局背景。表 1-2 总结了使用 DeepLabV3+ 作为骨干网络的分割方案, 以及它们的改进和实验结果。

表 1-2 分割场景及其改进和使用 DeepLabV3+ 作为骨干网络的实验结果

应用场景	方法	结果
荔枝枝条分割 ^[66]	DeepLabV3+ 使用的 Xception 特征提取模型在不同层次上得到了改进	76% (IoU)
猪群分割 ^[54]	DeepLabV3+ 深度学习方法引入了轻量级的基于行和列的注意力模块, 可以加快特征计算的效率	74.62% (IoU)
葡萄叶黑腐病斑的分割 ^[68]	在 DeepLabV3+ 编码器中加入了一个基于特征金字塔网络的特征融合分支	84.8% (IoU)

卷积层的不同扩展率可以充分利用高分辨率图像对植物进行多尺度测量和分割^[68], 这将有助于克服植物生长阶段形状的巨大变化对分割精度的影响, 但仍然受到局部区域的限制。其次, 巨大的参数数量带来了巨大的计算负担。现在 DeepLabV3+网络多用于实现高时空分辨率、高光谱遥感图像的分割^{[67][68]}。结合机器学习技术和遥感数据^{[69][70]}, DeepLabV3+网络可以识别不同尺度的土地覆盖、作物类型和气象数据等。

(4) 其他

VGG 网络、AlexNet、GoogLeNet, 以及其他一些深度学习模型^{[71]-[90]}由于其出色的性能, 经常被用作农业图像语义分割的编解码器网络的骨干。

VGG 网络^[73]和 Inception^[88]经常被模块化, 应用于 Encoder-Decoder 网络中的下采样部分, 并用于替换原始编码器或与原始编码器形成不同的编码阶段。在上采样阶段, 使用反卷积可以将 Inception 模块和上采样巧妙地结合起来, 如图 1-6 所示。增加的 Inception 模块对上采样的卷积层进行叠加和分解, 以捕捉多尺度信息。

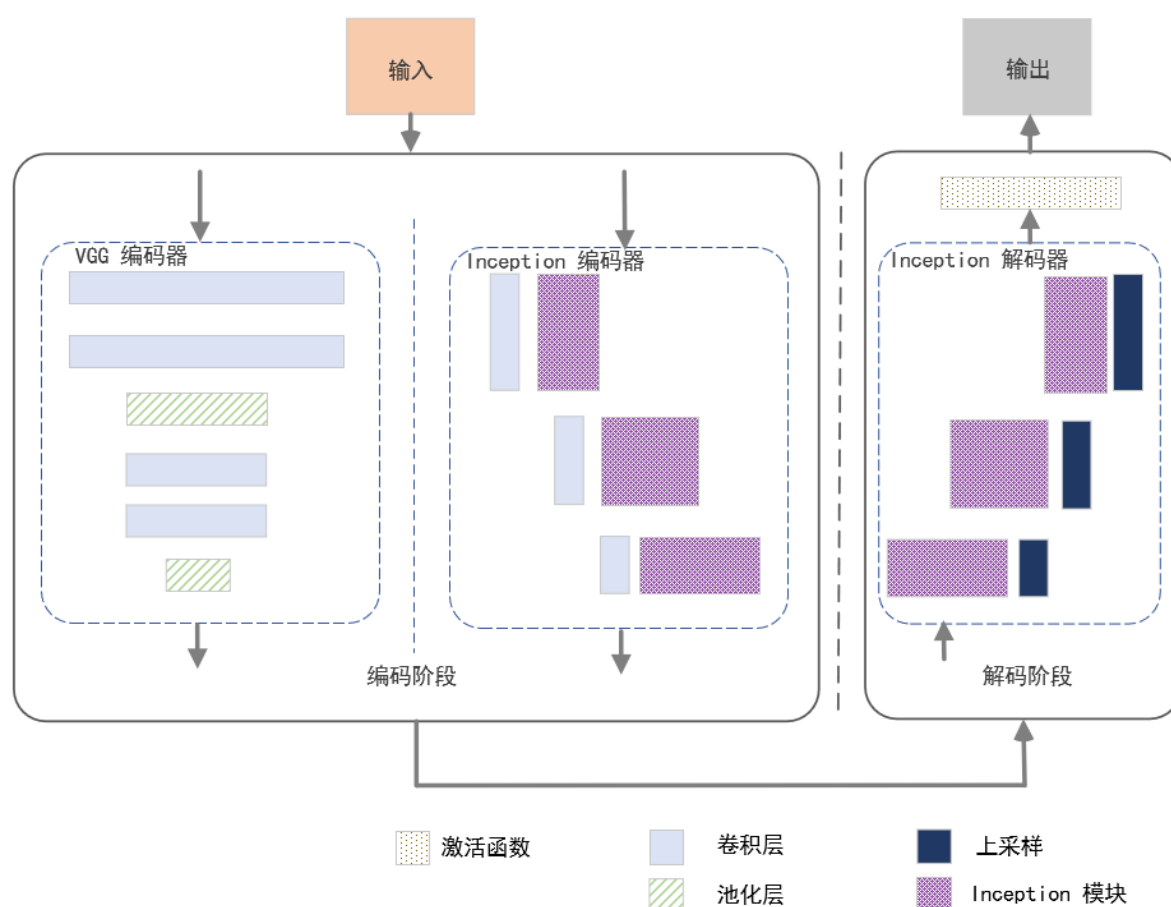


图 1-6 以 VGGNet、GoogLeNet 作为骨架的语义分割模型

上述 CNN 网络架构在农业图像的语义分割任务中表现良好。由于其出色的性能,

CNN 已被广泛研究，如密集连接卷积网络（DenseNet）^[71]和 ShuffleNet^[72]，它们也被用于农业图像的分割。

1.2.2.3 农业图像分割多尺度策略

无人机图像和遥感图像被广泛用于农业，它们有共同的特点：高分辨率。卷积神经网络的高性能在很大程度上取决于精细的空间细节和足够的背景信息，这两者都会引发高计算成本，这严重地限制了它们的实时处理能力。在农业遥感图像分割的一些场景中，如耕地分割^[73]，水产品监测^[74]等，由于高分辨率图像的特点，如地面物体细节丰富、类别差异大，现有的卷积神经网络图像分割方法普遍存在分割精度低、地面物体边界不准确的问题。

多尺度卷积神经网络已被用于解决这些问题。除了 DeepLab 网络中使用的 ASPP（Atrous Spatial Pyramid Pooling）模块外，DDCN（Dynamic Dilated Convolutional Network）^[75]也可用于多尺度策略。DDCN 的关键思想仍然是保持图像的分辨率，具体来说，该模型在训练过程中不会对动态生成的输入数据（原始图像和可能的输入尺寸的概率分布）进行下采样。高空间分辨率带来的丰富的纹理特征引发了同一作物的图像特征的巨大差异。1.2.2.2 节中介绍的网络本身就具有多尺度的优势，插入 ASPP、DDCN 等模块将进一步增强网络的功能。DeepLabV3+和 DDCN 结合多尺度策略的方法在对柑橘园的无人机图像进行语义分割时获得了 95.15%、95.46%的精度，高于 U-Net、SegNet 的 94.96%、94.96%^[76]。同时，卷积层中不同的扩展率可以充分利用高分辨率图像对麦穗进行多尺度测量和分割，这将有助于克服植物生长阶段形状的巨大变化对分割精度的影响^[77]。

1.2.3 注意力算法研究现状

当人眼观看图像时，只能看到高分辨率的一小块，这个小块称为中央凹。为了识别整个场景，眼睛基于中央凹进行特征提取。眼睛移动到图像的不同部分，直到从中央凹获得注意力焦点。在生物学上，人类利用有限的注意力资源，处理动态信息到视野子集的能力称为视觉注意力。眼动建模研究是许多人工智能任务的灵感来源，受人类视觉系统启发的不同类型的注意机制已经开发多年，深度神经网络的成功一直处于这些图像识别、机器翻译等人工智能任务的最前沿。研究人员一直在研究用注意力机制加强神经网络架构的方法。这些研究的主要目的是减少计算负担并提高模型性能。

受人类视网膜中央凹启发的第一个实施系统在图像分类^[78]中得到应用,该系统联合训练了受限玻尔兹曼机(RBM)和称为注视控制器的注意力组件。深度注意选择性网络(DasNet)提出了一种具有反馈连接的深度神经网络,通过强化学习将选择注意力引导到从图像中提取的关键特征上^[79]。

在 2014 年,注意力机制首次应用在机器视觉中,给后续的神经机器翻译注入了新的思想。注意力机制包括自注意力机制、软注意力机制、强注意力机制,他们的共同本质就是在长序列中对关键信息进行选择关注,降低无关信息的权重。

从 2015 年开始,注意力机制被用于不同的任务,并提出了应用注意力机制的新型神经网络架构。神经图灵机^[80]的引入允许网络在无监督情况下通过使用原内容进行端到端训练。基于软注意机制^[81],人们又提出了端到端记忆网络^[82],这是一种基于循环注意机制的记忆网络形式。之后,自注意力机制已在长短期记忆网络(LSTMN)^[83]中成功实现,它通过替换记忆单元来修改标准 LSTM 结构。记忆和注意力被添加到 LSTMN 中的序列编码器,为了计算序列的表示,自注意力机制用于描述关联的不同位置,并首次应用于自然语言处理^[84]。在 2017 年,提出了一种完全基于自注意力的新型神经网络架构——Transformer,这种架构的特点是它既不包含递归也不包含卷积。Transformer 通过用自注意力替换用于 NMT 的编码器-解码器架构中的传统循环层实现特征提取的工作,在机器翻译任务上也取得了不错的成绩。

最新的深度学习模型已经在注意力机制方面进行了探索,以提高性能。对于视觉而言,注意力机制按照维度可以分为通道注意力、空间注意力、时间注意力和分支注意力,由多个注意力模块堆叠而成的残差注意力沙漏网络^[92]能够快速收集图像全局信息,并融合全局信息与原始特征图,但由于网络深度大,计算负担重。SENet 网络^[93]是注意力机制应用的优秀自适应网络之一,SENet 通过压缩—激励方法使通道间建立相互依赖关系,从而调整权重。也有方法将空间注意力模型和通道注意力模型结合,在采样阶段并联工作^[94]。

将注意力机制融合至用于语义分割的深度模型中,利用其优势,使网络对重点区域更加敏感。ViT^[95]是一个只有注意模块堆积的编解码器网络,它通过分解 patch 来实现图像序列化。在此基础上,DeiT^[96]的效率更高。已有研究表明,自注意力层可以代替卷积运算实现高精度的语义分割,但它带来了巨大的计算成本。

1.3 研究内容及创新点

1.3.1 主要研究内容

在农业生产中，传统的植物产量估算方法往往利用人工方法，严重耗费劳动力。与普通农作物不同，罂粟通常是小规模种植，需要更精确的分割。罂粟的主要识别特征是植物的形态以及雄蕊和果实的特征。罂粟的叶子很大，边缘呈不规则的波浪状，叶子的基部以抱茎的方式包围着茎。然而，由于品种不同，罂粟的花瓣颜色从粉红色到红色不等，所以花色并不是一个好的特征。并且，在拍摄过程中拍摄角度与尺度存在差异大的问题，除此之外，复杂的自然环境使采集到的图像光照条件影响了罂粟植株的特征表达。到目前为止，还没有针对于罂粟图像的分割方法。如图 1-7 所示，罂粟花的图像在比例、角度和光照强度上有较大差异。



图 1-7 罂粟图像在（a）拍摄角度、（b）光照强度和（c）不同尺度方面存在差异

本文工作的主要目标是对罂粟图像进行准确分割，以区分罂粟和其他作物。为了实现罂粟的准确分割，本文需要解决以下困难：

- （1）罂粟不是一种常见的作物，很难收集，而且样本数量少，无法支持高精度度网络的训练；
- （2）RGB 图像大多能保留罂粟和其他农作物的形态特征，但图像中会有干扰。类似的树冠或其他植物的干扰很容易导致分割错误的问题；
- （3）拍摄高度的变化带来了复杂的尺度特征；
- （4）罂粟的植物形态比较特殊，同一株罂粟的果实和茎干之间往往存在较长的距离，会导致边缘信息的丢失。

由于传统方法缺乏处理图像丰富语义信息的能力，人们利用深度学习技术在农业产量估算方面做了很多工作。作为图像处理的一个重要工具，卷积神经网络可以充分利用

图像的语义信息。然而，由于卷积操作在建模长距离关系方面仍然存在局限性。因此，这些架构通常产生较弱的性能，特别是对于结构纹理、形状特征差异较大的罂粟图像。在这方面，本文以 CNN 为骨干，整合了具有优秀注意力机制的 Transformer 技术。与 CNN 的结合有效地避免了仅由变压器构建的网络所需的高计算能力。分割系统主要由以下两部分组成：

（1）基于敏感特征的数据增广方法研究

由于农业图像领域公共数据集较少，在实际应用场景中采用的数据集一般只包含数百张高清图像，达不到训练一个高精度网络的要求。在这种小样本问题中，数据集的增广至关重要。本文采用色彩空间转换，对图像进行通道分离，增强罂粟的颜色，形状特征。通过多个通道归一化后的 RGB 图像训练得到更高的分割精度。与此同时，将裁剪、旋转等其他增广数据集的方法进行分割结果比较，得出最优结果。

（2）基于注意力的编解码语义分割算法研究

在一些农业无人机图像分割场景中，由于高分辨率图像的特点，如地面物体细节丰富、类别差异大，现有的卷积神经网络图像分割方法普遍存在分割精度低、地面物体边界不准确的问题。除此之外，在对罂粟无人机图像进行分割时，由于相似植物干扰（如具有相似特征的大葱植株）、间接种植等问题的存在，提升了分割难度，产生了分割边界不清晰，分割错误像素，分割精度过低的问题。本文采用以 U-Net 为骨干并融合自注意力机制的网络来解决上述问题。

U-Net 是深度学习语义分割领域的一种代表性方法，它包含编码器和解码器以及连接二者的跳跃连接三部分。由于其精简的结构，U-Net 可以很好地应用于样本较少的农业图像。此外，它能很好地结合图像的全局和局部细节特征，融合具有规则性的低分辨率信息和具有模糊边界和复杂梯度的高分辨率信息，可以实现对一些农业图像的精确分割。通过对 U-Net 深度语义分割模型进行改进，研究用于罂粟的语义分割方法。在跳跃连接中实现的多头注意力模块将有效地抑制不相关区域中的激活，从而减少冗余特征的数量。以像素点的相关性为依据，从多个尺度上融合特征信息，保留上下文特征，处理边缘模糊问题。将改进的网络模型与其他网络模型（如原始 U-Net、DeepLabV3+）在制备的罂粟数据集上进行比较，以验证本项目所建网络的性能。

1.3.2 创新点

针对无人机拍摄的罂粟图像语义分割问题，文章从以下两个方面进行改进：

（1）小样本数据增强问题

首先，通过对图像分离通道的研究，探索解决罂粟语义分割问题中样本过小的问题。本文使用 ExG（超绿因子）、CIVE（植被颜色因子）和组合植被因子 COM（组合因子）来处理原始图像，使用了原始输入的进一步表示。转换后的图像与通道化的输入 RGB 图像相连接，并建立起归一化的输入图像，而不是将原始 RGB 图像作为网络输入，以强化罂粟的独特特征（叶子纹理、果实形状）。这些额外的表征不仅有助于学习权重参数，导致网络有更好的泛化特性，而且还获得了更好的分离植被的性能。这对于样本少、类间干扰大的罂粟语义分割任务很有帮助。

（2）图像尺度特征复杂、罂粟形态学特征复杂，分割边界不清晰问题

由于实际应用场景下，拍摄时的光照变化、存在遮挡物等问题，会严重干扰分割结果，同时，分割过程中会受到相似类别目标的干扰。U-Net 使用跳跃连接将来自下采样路径的空间信息与来自上采样路径的空间信息结合起来。然而，这带来了许多冗余的底层特征提取，因此在初始层中特征表示很差。本文利用 Transformer 改进了一个 U 型网络——TAU-Net，用于无人机拍摄的罂粟图像的语义分割任务。TAU-Net 主干网络同时包含了 CNN 网络和自注意力机制。由于罂粟的特征在不同尺度上差异很大，与仅由自注意模块或仅由卷积模块构建的网络不同，TAU-Net 利用 Transformer 的优势来感知整个图像，而不是感知区域有限的原始 U-Net。改进后的网络提高了网络对尺度变化的鲁棒性，而不需要承担更高的计算量。其次，在跳跃连接中实现的多头注意力模块将有效地抑制不相关区域（非罂粟区域）中的激活，从而减少冗余特征的数量。在获得多尺度特征信息的同时，全局信息可以引导选择低级特征，两个分支结合实现更好的分割效果。此外，由无人机采集的罂粟图像具有高分辨率和高空间结构的像素。罂粟的果实和根茎之间存在着远程依赖关系。在本文中，在编码阶段使用的 Transformer 融合了对位置编码方法，相对位置编码方法学习了语义分割任务产生的大量 tokens 之间的空间关系，以保持更准确的果实位置信息。

1.4 文章的组织结构

论文一共分为五章，具体安排如下：

第一章是文章的绪论。首先阐述了小样本数据农业图像增广技术国内外研究现状及农业图像语义分割算法国内外研究现状，其中，农业图像语义分割算法研究现状包括基于传统方法的语义分割算法及基于深度学习方法的分割算法国内外研究现状；然后描述了本文的研究内容、创新点；最后概况了文章的组织结构。

第二章是文章的相关工作。分别介绍了小样本数据增强、基于编解码结构的语义分割算法、自注意力机制。

第三章设计了为解决小样本问题的基于敏感特征的数据增广方法。方法针对于样本过少的难题，设计了融合了旋转、随机窗口、通道分离的增广方法，将罂粟图像进行随机切割划分和旋转操作，并计算不同的植被索引因子，分离通道。处理后的图像作为额外的表征支持 CNN 的训练。同时进行多次实验验证，验证使用数据增广算法后对分割网络的分割精度产生的影响。

第四章设计了一个混合注意力机制的端到端语义分割网络对输入数据进行语义标注。方法利用 Transformer 改进了一个 U 型网络——TAU-Net，用于无人机捕获的罂粟图像的语义分割任务，实现罂粟图像的精准分割。本章在对 U-Net 算法进行改进后结合第三章所提方法，对所提算法进行实验验证，并与一些经典算法的整体性能进行比较和可视化分析。

第五章是总结与展望。主要是对本文所做工作进行总结，并对今后的研究进行说明。最后是参考文献和致谢。

第二章 相关理论

2.1 小样本数据增广

在农业图像分割任务中使用监督方法的缺点之一是缺乏足够的标记数据集，这往往会影响网络的训练过程^[97]。深度学习模型的鲁棒性和通用性受到训练数据的多样性和数据量的影响。在小样本问题上，传统方法通常在现有数据上进行几何变换或颜色变换，然而，它们并没有实质性地增加数据集。现有其他的想法来解决这个问题。其中一个方法是利用从图像的分离出来的多个通道^[57]（原始 RGB 数据、植被指数、HSV 颜色通道和边缘检测器）被用作 CNN 的输入数据，这在有限的训练数据下增强了模型的泛化，方法如图 2-1 所示。

在植被图像分割中，各种基于阈值的方法遵循植被和图像中其他物体之间的平等设定，这在植物生长的不同阶段并不完全有效^[41]。除此之外，将 RGB 值转换为灰度的彩色指数技术并不总是能精准地辨别灰度图像。这是由照明条件对 RGB 图像值的影响造



图 2-1 使用多通道图像的编解码器网络的分割结果作为输入数据

成的。因此，尽管基于颜色指数和阈值指数的方法提供了简单和先进的技术，但由于光照变化条件对取得的植被分割结果的影响，它们有一定的局限性，特别是在晴天和阴天的情况下。对于柑橘类水果图像的分割，在将原始 RGB 图像转换为 HSV 图像后，使用了分水岭算法，得到的 R2 值为 0.93^[42]。为了检测适当的阈值，使用了 HSV 颜色空间，然后提取色调通道以生成色调直方图。使用过滤的色调直方图和色调直方图的拟合高斯曲线来检测适当的阈值^[43]。该方法可以产生准确和稳定的植被分割性能，平均准确率为 87.29%，标准偏差为 12.5%^[41]。在对前景和背景进行分割时，基于植物指数的方法被用于近似的预分割，以计算分割阈值。当它与 Otsu 算法相结合时，植物图像的语义分割精度达到了 97.4%，当图像使用 NCIVE 指数计算时，图像在真实世界环境中的分割鲁棒

性得到了明显的改善。此外，基于指数的方法的优点是，它们可以在降低计算成本的同时实现与复杂网络相当的分割性能。因此，基于植物指数的方法被广泛用于植被分割任务。

2.2 基于编解码结构的语义分割算法

传统图像分割方法大多利用图像的表层信息，对于需要大量语义信息的分割任务则不适用，无法应对实际的需求。面对包含多源信息的图像场景以及愈发复杂的任务要求，众多基于深度学习的图像语义分割方法被提出。

无人机图像和遥感图像被广泛用于农业，它们有共同的特点：高分辨率。卷积神经网络的高性能在很大程度上取决于精细的空间细节和足够的背景信息，这两者都会引发高计算成本，这严重地限制了它们的实时处理能力。其次，在一些农业遥感图像分割场景中，由于高分辨率图像的特点，如地面物体细节丰富、类别差异大，现有的卷积神经网络图像分割方法普遍存在分割精度低、地面物体边界不准确的问题。目前基于编解码结构的卷积神经网络可以很好地提升分割精度。

全卷积网络（Fully Connected Network, FCN）^[56]是深度学习用于语义分割的开创之作，确立了图像语义分割（即对目标进行像素级别的分类）通用网络模型框架。作为最早的基于 CNN 的高影响的分割技术之一，全卷积网络为编解码器网络的发展提供了关键思路。FCN 的重要思想是输入任意大小的图像，并通过高效的模型学习和推理得到相同大小的输出。池化操作使特征图发生了尺寸转换，为了在分割过程中保持与原始图像相同的尺寸，反卷积层被用来对特征图进行双线性上采样，以对每个像素进行预测，并且可以保留原始输入图像中的空间信息。最后，FCN 用卷积层代替了全连接层，使用 1×1 卷积对每个像素点进行预测性分类。

基于 FCN 的 U-Net 结构^[98]，最初被应用于医学图像分割，并表现出优异的分割性能，是编解码网络代表性网络框架。U-Net 网络结构由三个部分组成。（1）逐级抽象特征的下采样阶段（2）重建特征的上采样阶段（3）实现分类的最后卷积层。该架构如图 2-2 所示。

下采样阶段由经典的卷积网络架构组成，承担了网络的特征提取任务。下采样由重

复的卷积操作和池化操作模块组成，每个模块包含 2 个 3×3 卷积层、一个步长为 2 的 2×2 最大池化层和非线性 ReLU 层，并在每一个模块采样后对增加特征通道至采样前两倍。上采样阶段由反卷积网络架构组成，承担网络的解码、恢复分辨率任务。下采样由重复的反卷积操作和 concat 操作模块组成，每个模块包含架构中一个上采样的反卷积层，两个 3×3 的卷积层，与下采样模块相连的跳跃连接和非线性 ReLU 层。在每一个模块采样后将特征通道数量减少至采样前一半，特征图大小恢复至采样前两倍。跳跃连接是 concat 操作，将下采样的特征图与经过反卷积后的特征图进行拼接融合，融合后的特征图再进行两次 3×3 卷积。最后一层为卷积层，卷积核大小为 1×1 ，对特征图按照特定类别数量进行分类。

由于其精简的结构，U-Net 可以很好地应用于样本较少的农业图像。此外，U-Net 能很好地结合图像的全局和局部细节特征，融合具有规律性的低分辨率信息和具有模糊边界和复杂梯度的高分辨率信息，可以实现对一些农业图像的精确分割。但通常情况下，像素块的大小要比整个图像尺寸小得多，该算法只能提取一些局部特征信息，这就导致了分类性能的限制。其次，由于相邻像素块的重复，计算机需要做大量的重复性计算。近年来，U-Net 已成功应用于植物表型观察和病虫害检测等任务。

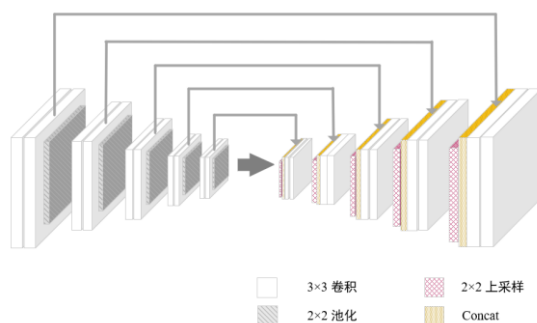


图 2-2 U-Net 网络结构

2.3 注意力机制

注意力机制的本质是对输入数据的特定选择，它通过突出输出上的关键输入而抑制非重要的信息，集中于大量数据中最关键的信息，从而提高网络的效率。自注意力是注意力机制的一个变种，它较少依赖外部信息，更善于捕捉数据或特征的内部相关性。

Transformer 起源于自然语言处理领域，是具有代表性的基于自注意力机制的算法，在视

觉领域也获得了巨大的成功^[95]。Transformer 由编码器-解码器堆栈组成，每个堆栈本身都有六个相同的层，每个堆栈仅包括注意力机制和前馈神经网络。

Transformer 自注意力在计算时为每个单词使用三个向量，即查询、键和值。这些向量是通过将输入与权重矩阵相乘来计算的 W^q 、 W^k 和 W^v ，并在训练中更新学习权重。通常，每个值都由具有相应键的查询函数加权，输出计算为值的加权和。

在 Transformer 中使用的是缩放点积(Scaled Dot-product)注意力。点积(Dot-product)方法是 Scaled Dot-product 的基础运算。具体算法如下：输入两个向量 a_1 、 a_2 ； a_1 、 a_2 向量乘上不同的权重矩阵 W^q 、 W^k ，得到 q 和 k ； q 与 k 做点积，得到输出关联程度 α 。与此同时，注意力机制通过学习去更新权重矩阵 W^q 和 W^k 。Dot-product 算法如图 2-3 所示。



图 2-3 Dot-product 算法

Scaled Dot-product 注意力是多头注意力模块的组成部分。在 Scaled Dot-product 机制中输入数据包括查询信息 Q ，键、值信息 K 、 V 三个向量。为了梯度的稳定，Transformer 对分数除以 $\sqrt{d_k}$ 进行归一化。再经 softmax 函数激活后，通过点乘得到输入向量的加权评分，加权评分相加后得到输出结果。算法公式如公式 2.1 所示：

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (2.1)$$

由于此架构不包含任何循环层或卷积层，因此在编码器和解码器的开头使用位置编码给出有关输入序列中相对或绝对位置的信息。在 Transformer 应用在视觉中时，Transformer 编码器中没有针对 tokens 位置信息的处理，故需要在 embedding 层后加入位置编码器，将位置不同可能会产生不同语义的信息加入到嵌入张量中，用来弥补位置信息的缺失。

第三章 数据增广方法

3.1 引言

为了减少对原始图像分布的影响,RGB 图像通常被直接用作神经网络的输入,但由于罂粟与普通农作物不同,通常是小规模种植,样本数量有限,这种方法不足以训练出一个具有良好性能的分割网络。本文的分割系统针对样本过少的难题,将罂粟图像进行切割划分和旋转操作,并计算不同的植被索引因子,分离通道。处理后的图像作为额外的表征支持 CNN 的训练。下面章节提供了该方法的整体架构与实验内容。

3.2 基于敏感特征的数据增广方法

植被因子主要是基于 RGB、Lab 以及 HSV 颜色空间构建。其中,由 3 个颜色通道组成,分别是红(R)、绿(G)、蓝(B)。特征鲜明,计算简单。在 RGB 颜色空间的基础上,为了提高无人机影像对罂粟与非植被物体、其他种类植物的识别精度,借鉴常用的增强影像中植被信息的植被因子构造符合罂粟特征的增强因子。

绿色植物在光线照射下,不同的波段的反射和吸收特性不同,如图 3-1 所示。以绿色植物的光反射和光吸收特性为依据,统合光谱信息,对将不同波段的反射率进行组合数学运算,实现增强绿色植物区域信息的同时,使非植被信息关注度变低。

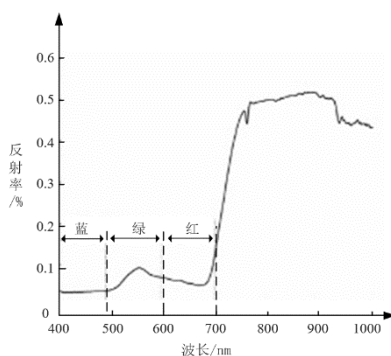


图 3-1 绿色植被光谱

罂粟茎高 30-80 厘米,花蕾呈卵球形,植株分支较多,分离度较差。影像获取时罂粟正处于开花期或者结果期,开花期为 4 至 6 月,结果期为 6 至 8 月,期间植被茂盛,绿植覆盖度高。罂粟的颜色特征全年表现为绿色,与其他作物相比,常见绿色农作物颜

色较暗，罂粟颜色偏向淡绿色。

虽然罂粟植被的可见光吸收率往往大于非植被，在村落中，由于遮挡物多、场景中信息复杂，罂粟植被的绿光反射效果不佳，导致植被识别效果较差。本文利用光谱曲线工具对比验证得到罂粟在绿波段高反射，但与其他植物存在反射数值差距的结论。罂粟图像在不同颜色波段的表现如图 3-2 所示。

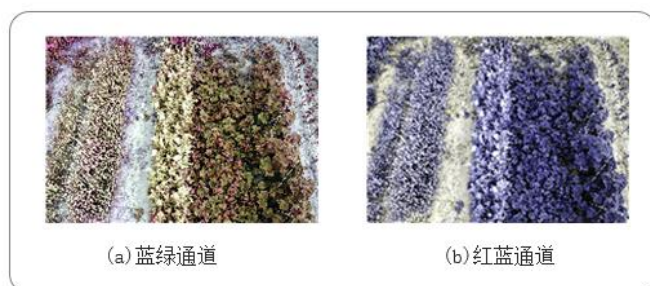


图 3-2 罂粟通道分离图像

综上特征，基于可见光影像构建 RGB 三波段植被因子，使识别提取罂粟植株的可行性增强。目前对于仅基于可见光波段的植被因子较少，主要有过红因子 ExR (Excess Red) 及红绿比值因子等，常见的植被因子的计算方式如下表 3-1 所示：

表 3-1 常用植被因子及其计算公式

植被因子	计算公式
ExR	$1.3 \times R - G$
ExGR	$ExG - ExR$
MExG	$1.262 \times G - 0.884 \times R - 0.311 \times B$
NDI	$128 \times (((G - R) / (G + R)) + 1)$
VEG	$G / (R^a \times B^{(1-a)}), a=0.667$
ExG	$2 \times G - B - R$
CIVE	$0.441 \times R - 0.811 \times G + 0.385 \times B + 18.78745$
GB	$G - B$
COM	$R + B$

有限的样本数量导致这种方法不足以训练出一个准确度良好的分割网络，为了解决这个问题，本文借鉴了表 3-1 中在农业领域具有出色表现的植被因子方法，从原始 RGB 图像中导出额外的表征。在罂粟图像中，罂粟植株与其他植物之间有着显著差别的特征包含：叶片特殊的锯齿状纹理特征；果实的圆形纹理特征和颜色特征；雄蕊的颜色特征。本文选择了四个对罂粟特征敏感的因子：ExG、CIVE、GB、COM 用于图像增强。

3.2.1 基于 ExG、CIVE 植被因子增广方法

正常光线条件下，绿色植物在绿光波段反射性强，其他波段吸收性强植被因子，植被因子方法通过强化绿光反射的方式区别植被和非植被。针对于罂粟图像，ExG、CIVE 植被因子可以有效地增强植物与其他类别（土壤、屋檐）之间的特征差异。超绿因子 ExG 是提取图像中绿色目标的常用算法，通过 R、G、B 分量的线性组合，弱化阴影和土壤区域，强化绿色区域，对绿色植被有较好的识别效果。除此之外，红光与蓝光波段对土壤与绿色植物的信息也有补偿作用。因此，相对于仅利用绿光与红光波段，或者仅用绿光与蓝光波段构造的植被因子，CIVE 因子可以同时处理三波段，可以实现植物与其它区域的明显分割。公式 3.1 与 3.2 分别表示 ExG、CIVE 植被因子的计算公式。

$$I_{\text{ExG}} = 2 * I_G - I_B - I_R \quad (3.1)$$

$$I_{\text{CIVE}} = 0.441 * I_R - 0.811 * I_G + 0.385 * I_B + 18.78745 \quad (3.2)$$

其中， I_R 、 I_G 和 I_B 分别代表相应波段的归一化像素值。 I_{ExG} 、 I_{CIVE} 表示 ExG、CIVE 植被因子。ExG、CIVE 植被因子进行处理的具体算法如算法 1 所示。

算法 1 使用 ExG、CIVE 植被因子进行图像增广

输入：无人机航拍的 RGB 原始图像 I 、增广类型 aType 。

说明： aType 为 ExG 或 CIVE

输出：增广罂粟图像 aImage 。

```

1:  Input RGB image  $I$ ;
2:   $B, G, R = \text{Split}(I)$ ;
3:  switch ( $\text{aType}$ )
4:      case ExG:
5:           $\text{index} = 2 * G - B - R$ ;
6:          break;
7:      case CIVE:
8:           $\text{index} = 0.441 * R - 0.811 * G + 0.385 * B + 18.78745$ ;
9:          break;
10: end

```

```
11: Convert index to gray image: aImage;
```

```
12: return aImage;
```

如图 3-3 所示，转换后的图像与通道化的输入 RGB 图像相连接，并建立起归一化的输入图像。

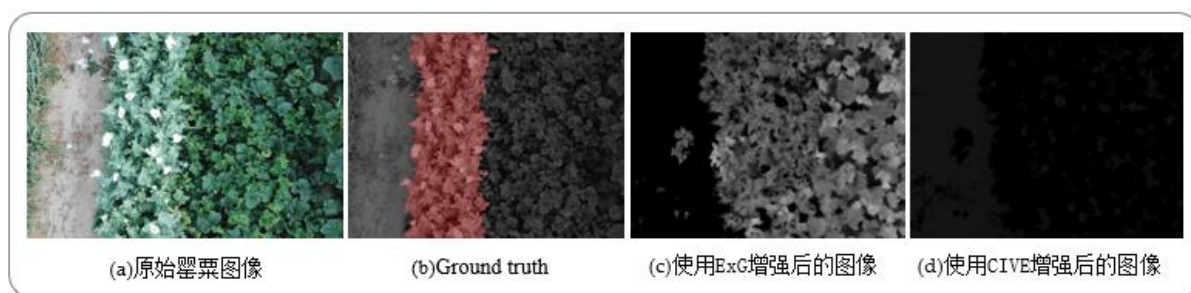


图 3-3 原图以及增强后的图像

3.2.2 基于 GB、COM 植被因子增广方法

GB 的作用是提取植物纹理特征，增强罌粟与其他植物的特征差。同时，本文在罌粟植被可见光反射特点的基础上，提出一种基于 RGB 颜色空间的植被因子 COM。利用每个因素的优势，把它们组合成共同植被因子，进一步扩大数据集。 I_{GB} 、 I_{COM} 表示 GB、COM 植被因子，计算公式如公式 3.4，3.4 所示。

$$I_{GB} = I_G - I_B \quad (3.3)$$

$$I_{COM} = I_R + I_B \quad (3.4)$$

使用 GB、COM 植被因子进行图像增广的具体算法如算法 2 所示。

算法 2 使用 GB、COM 植被因子进行图像增广

输入：无人机航拍的 RGB 原始图像 I、增广类型 aType。

说明：aType 为 GB 或 COM

输出：增广的罌粟图像 aImage。

```
1: Input RGB image I;
```

```
2: B, G, R = Split(I);
```

```
3: switch (aType)
```

```
4:   case GB:
```

```

5:     index = G-B;
6:     break;
7:     case COM:
8:         index = R+B;
9:         break;
10: end
11: Convert index to gray image: aImage;
12: return aImage;

```

如图 3-4 所示, 转换后的图像与通道化的输入 RGB 图像相连接, 并建立起归一化的输入图像。

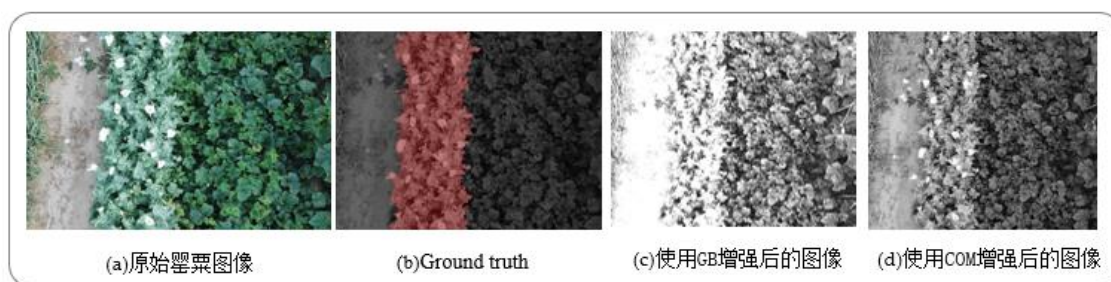


图 3-4 原图以及增强后的图像

这些额外的表征不仅强化了纹理特征, 有助于学习权重参数, 导致网络有更好的泛化特性, 而且还获得了更好的分离罌粟植被的性能。

3.3 实验结果与分析

3.3.1 数据集

为了保证实验可行性, 本研究所采用的无人机图像是在罌粟的开花期和结果期收集的。这些图像包含 32 张 3000×4000 分辨率的近地面图像。为了验证实验效果, 从不同尺度对罌粟进行了拍摄。这些图像是在不同程度的遮光和作物交错的情况下拍摄的。本文使用 labelme 软件将图像标记为两类, 即罌粟或背景, 并按照 4: 1 的比例划分训练集和测试集。

3.3.2 评价指标

在这项研究中，第一个指标是使用增广策略增广后数据集的数据量。第二个评价指标用来比较数据增广的实验效果的评价指标为 Dice 系数。Dice 系数，这是一个用来评估两个样本相似度的度量函数，大多用于语义分割任务。其公式如下：

$$Dice = \frac{2*TP}{2*TP+FP+FN} \quad (3.5)$$

3.3.3 实验设计

本研究提出的算法的实验使用单个 NVIDIA RTX A4000 GPU 进行。基于敏感特征的数据增广算法在罂粟花期、结果期图像数据集上进行了主要实验。本研究所提出的算法采用深度学习框架 PyTorch-1.12.1，编程语言为 Python_3.8，在 Windows10 操作系统下实现，GPU 型号为 NVIDIA GTX A4000，CUDA 版本为 11.3。

此实验是以经典分割网络 U-Net 作为骨干网络，按照相同的比例划分不同的数据集进行训练。在本实验中使用的数据集包含：（1）滑动窗口切割的 RGB 图像，窗口大小为 512；（2）随机窗口切割的 RGB 图像；（3）旋转+随机窗口切割的 RGB 图像；（4）经随机窗口切割后使用基于敏感特征的策略增广后的图像。batch size 大小设置为 8，输入图片大小为 512×512。epoch 分别为 100。学习率被设定为 0.01。权重衰减为 10^{-4} 。

3.3.4 实验结果与讨论

所提算法在 32 张高清罂粟图像上进行了实验，并将滑动窗口切割的 RGB 图像数据集、随机窗口切割的 RGB 图像数据集、旋转+随机窗口切割的 RGB 图像数据集、基于敏感特征的策略增广后的图像数据集在 U-Net 网络框架上进行对比实验。经随机窗口切割后的基于敏感特征增广方法与滑动窗口切割增广方法、随机窗口切割增广方法、旋转+随机窗口切割增广方法在数据量以及分割精度上进行了对比，其对比结果如表 3-1 所示。从表中可以看出基于敏感特征增广后的图像数据集在相同的网络框架上的分割精度优于其他三种算法。

农业图像语义分割中常用的评价指标主要是 Dice 指数，所提算法与当前常见的处理农业图像的算法在增广后的数据量以及 Dice 指数上进行了比较，其实验结果如表 3-1

所示。

表 3-2 所提四种增广算法对罂粟数据集增广后在 U-Net 网络框架上的实验结果

增广方法	数据量 (张)	Dice
滑动窗口	800	0.58
随机窗口	778	0.59
旋转+随机窗口	1624	0.65
基于敏感特征的策略	1536	0.74

从表中可以看出，所提算法增广后的数据量达到了 1536 张，Dice 分割精度达到了 0.74，是此对比实验的最高精度。另外，滑动窗口切割算法增广后的数据量达到了 800 张，Dice 分割精度达到了 0.58。随机窗口切割算法增广后的数据量达到了 778 张，Dice 分割精度达到了 0.59。旋转+随机窗口切割算法增广后的数据量达到了 1624 张，Dice 分割精度达到了 0.65。

滑动窗口切割算法对罂粟图像进行切割时，出现了大量切割图片中不包含目标区域的样本的问题，网络无法很好地学习罂粟特征。从图 3-5 中可以看出，基于滑动窗口切割增广算法的分割结果中，罂粟样本区域得到了较为完整的分割，与此同时，当拍摄角度发生轻微变化时，此算法对罂粟植株的识别率优于基于随机窗口切割的增广算法以及基于随机窗口+旋转切割的增广算法。但此方法的假阳率最高，将其他种类的绿色植被也作为罂粟目标分割，分割效果如 3-5 所示。

使用基于随机窗口切割算法对罂粟图像进行切割时，对不包含目标的样本数目多的问题有一定改善。从图 3-6 可以看出，基于随机窗口切割增广算法的分割结果中，罂粟样本区域的分割结果优于滑动窗口切割算法，假阳率低，并且对罂粟边缘细节敏感，果实的分割准确率也高于滑动窗口切割算法。但该算法对拍摄角度变化的鲁棒性较弱，当拍摄角度发生轻微变化时，此算法对罂粟植株的识别率远低于其他方法。

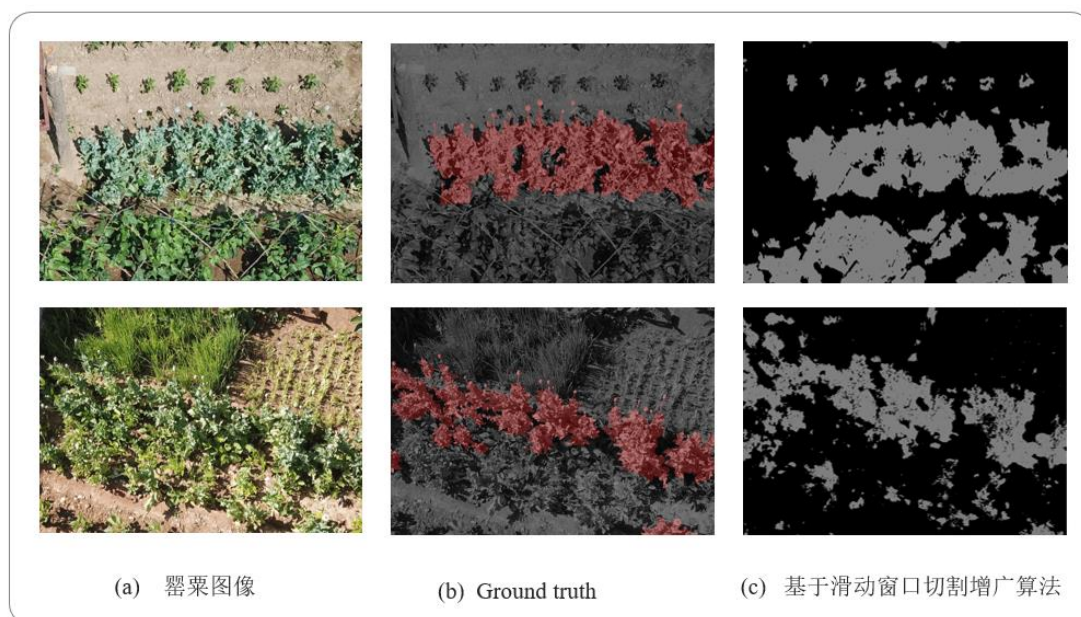


图 3-5 基于滑动窗口切割增广算法的分割结果

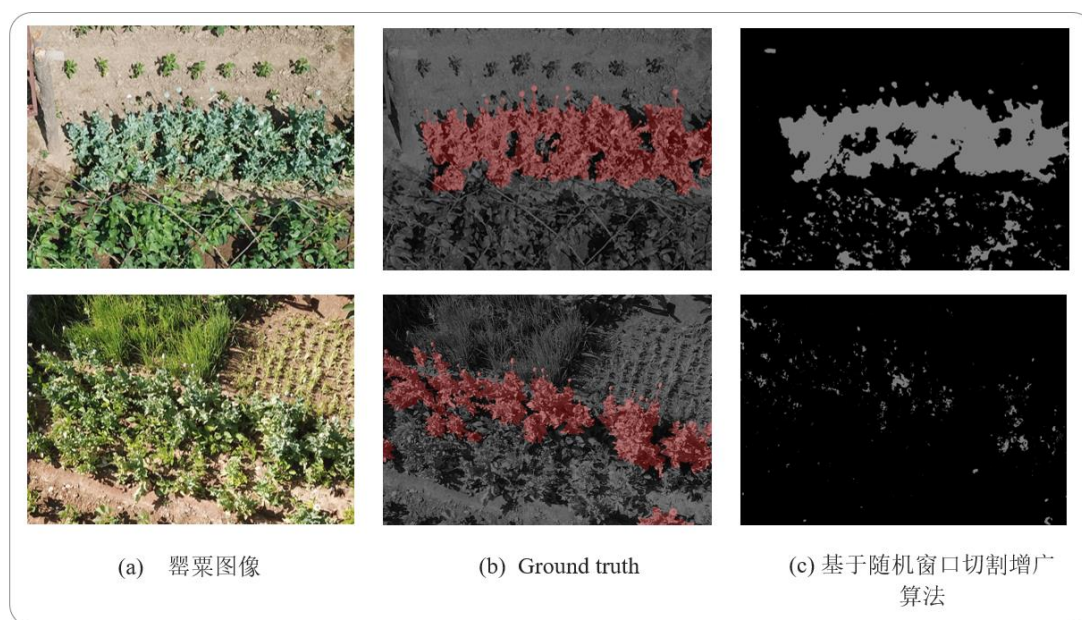


图 3-6 基于随机窗口切割增广算法的分割结果

如图 3-7 所示,基于旋转+随机窗口切割增广算法在对正向拍摄的图像进行分割时,分割精度出现大幅度下降,但对角度变化的鲁棒性强于仅使用随机窗口的算法。样本进行筛选后使用旋转+随机窗口方法进行增广,指标得到了明显的提升。

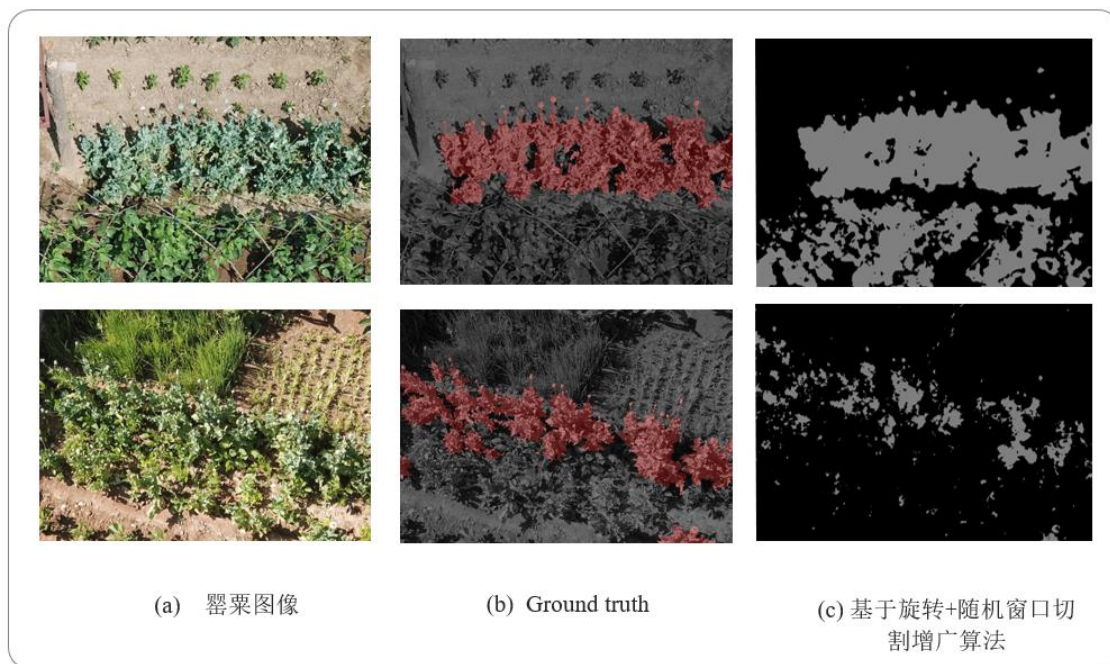


图 3-7 基于随机窗口+旋转切割增广算法的分割结果

本文提出的基于敏感特征的增广算法得到了最好的分割结果，如图 3-8 所示，土壤与植株的分割边缘清晰，假阳率低，并且对角度变化的鲁棒性也优于前三种算法。

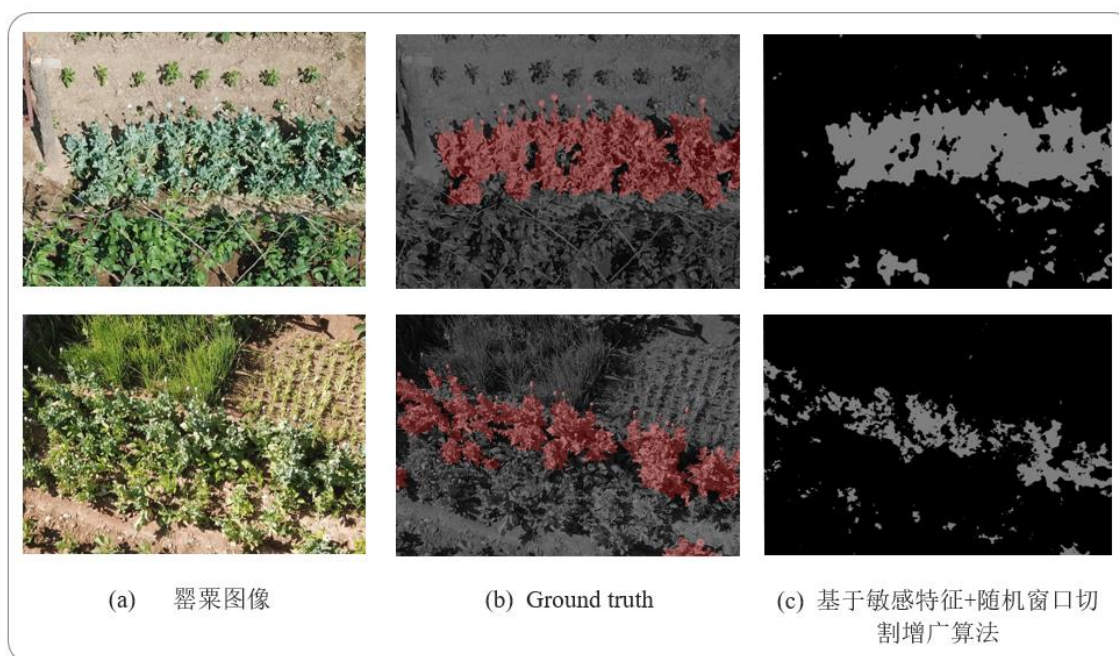


图 3-8 基于敏感特征+随机窗口切割的增广算法的分割结果

3.4 本章小结

本章主要针对于罂粟图像语义分割任务中过小的样本量问题，设计了一种用于数据增广的敏感特征方法，设计了一个将罂粟图像进行切割划分和旋转操作，并计算不同的植被索引因子，分离通道的增广算法。经实验验证，所提算法将原有的无人机图像样本量增至 1536，在 U-Net 模型上得到 0.74 的 Dice 指数，高于基于滑动窗口、随机窗口、旋转+随机窗口的增广算法的分割精度，验证了所提算法的有效性。

第四章 基于注意力的编解码语义分割算法

4.1 引言

无人机图像的语义分割是一项计算密集型任务，尤其是在分割具有复杂特征和类间干扰的罂粟图像时。此外，罂粟数据集的样本量很小。因此，结构简单、参数数量少、在以往小样本农业图像分割任务中表现良好的 U-Net 网络是这项任务的首选。U-Net 网络既有捕捉背景信息的收缩路径，也有允许精确定位的对称扩展路径。这些路径使网络能够将上下文信息传播到更高的分辨率，并融合罂粟图像中不同尺度的特征信息。然而，U-Net 有局限性：特征冗余，难以处理罂粟的果实和植物根茎之间存在的远程依赖关系。本文设计了一个改进的编解码器语义分割网络 TAU-Net，其中包含了注意力机制，对输入数据进行语义注释。

4.2 系统的整体架构

本文实现的无人机航拍图像中的罂粟植株语义分割系统由无人机图像拍摄、数据增广、分割网络三部分组成，如图 4-1 所示。

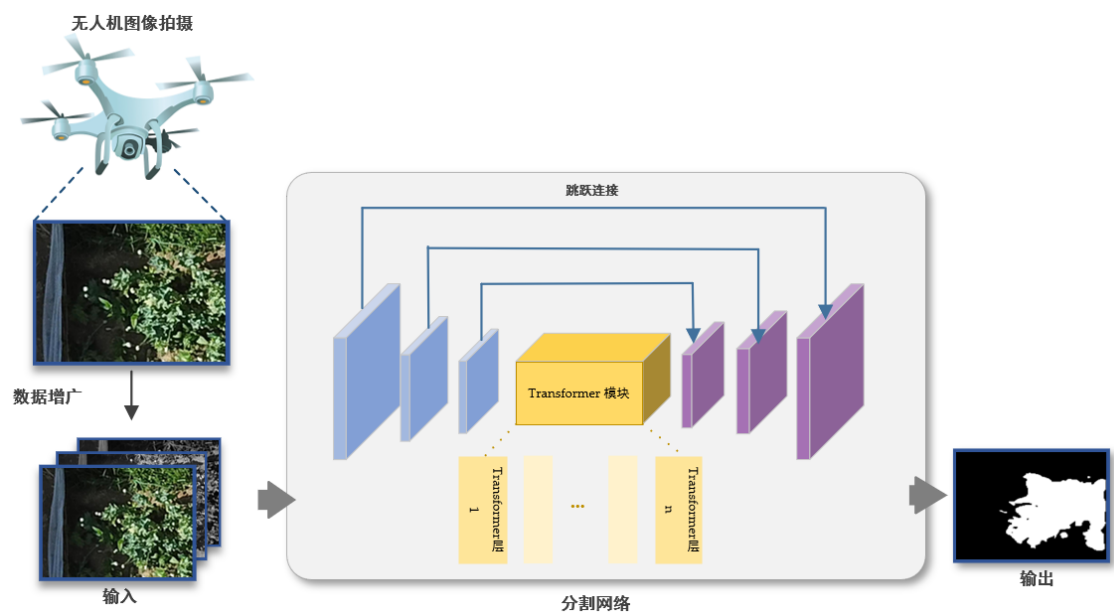


图 4-1 罂粟植株分割系统

首先，无人机在不同高度拍摄可能存在罂粟非法种植现象的航拍图像，经人工标注

后,使用第三章中所介绍的基于敏感特征的数据增广算法对原数据集进行扩充,增广后的数据集作为分割网络的输入。分割网络的骨干网络为经典分割网络 U-Net,分割网络由卷积模块、Transformer 模块、跳跃连接组成,其中 Transformer 模块中包含了多个混合注意力层。图像经过特征提取、特征融合、分辨率恢复后得到最终分割结果。

4.3 基于 Transformer 的编解码网络

无人机拍摄高度的变化导致罂粟植株在尺度发生改变时,罂粟的特征表示不同。针对尺度变化问题,提出了 CNN 与 Transformer 结合的编码器,结合池化和全局序列化的优势提升网络的尺度鲁棒性。本文以 U-Net 为骨干网络,提出了一个端到端的编解码网络:TAU-Net (Transformer Attention U-Net)网络。TAU-Net 利用 Transformer 的优势,可以感受到整个图像,而不是原始 U-net 有限制的感受野,适用于在不同尺度下特征差异较大的罂粟分割任务。网络模型图如图 4-2 所示。

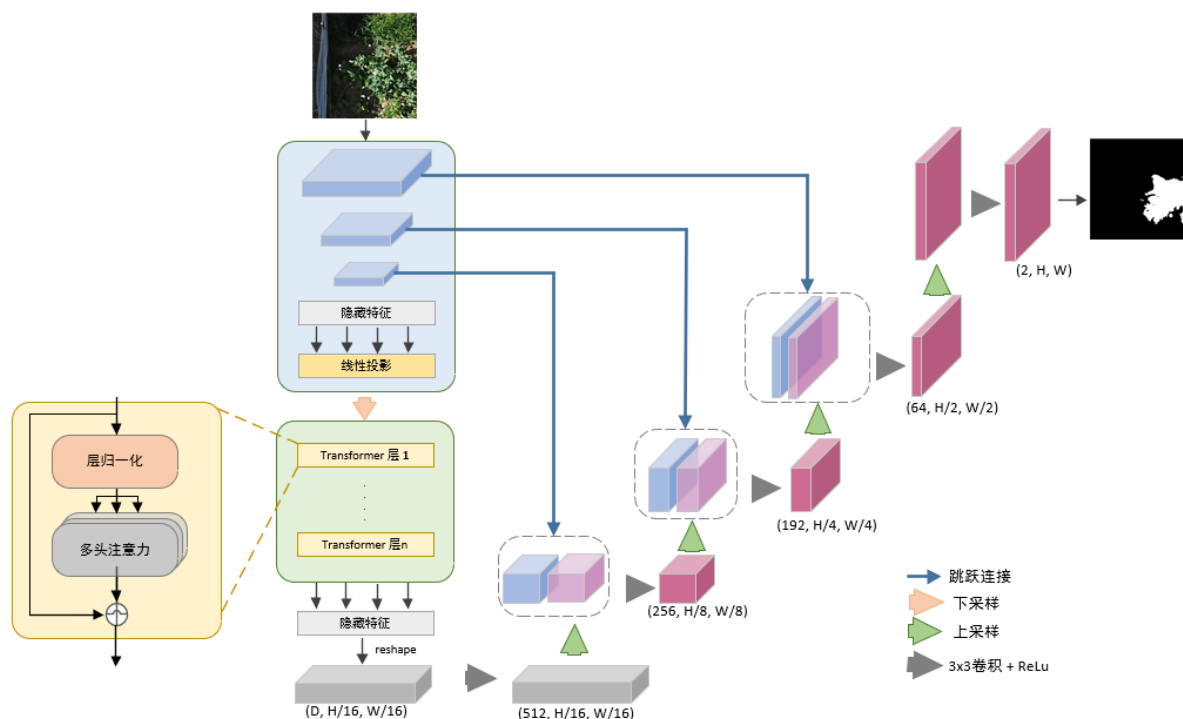


图 4-2 TAU-Net 网络模型图

(1) 输入阶段

本文使用第三章敏感特征增强后的标准化数据集作为 TAU-Net 网络的输入。为了

对齐, 输入图像的大小被调整为 $x \in \mathbb{R}^{H \times W \times C}$, 其中空间分辨率表示为 $H \times W$, 通道数为 C 。调整后的图像作为第一个卷积层的输入。

(2) 基于 CNN 的编码器

本文在编解码器网络的不同阶段使用不同的卷积模块, 在下采样阶段使用 ResNet 将输入图像编码为高级特征表示, 并通过最大池化层实现多尺度优势。图像输入至 ResNet 的第一个模块, 模块包含 7×7 卷积核, 最大池化层, 激活函数, stride 设置为 2。其余模块建立残差连接, 解决梯度消失问题。

(3) 基于 Transformer 的编码器

基于 Transformer 的编码器通过多尺度 transformer 模块和跳跃连接有效地整合全局特征和局部特征, 通过自注意力模块通过图像序列化, 更新权重, 抑制不相关的和干扰的特征, 在每个下采样阶段有效提取特征图的全局信息。将 Transformer 作为编码器的第一步是对输入进行序列化: 本文将输入的 x 重塑为一个平面的二维序列 $\{x_p^i \in \mathbb{R}^{P^2 C} | i = 1, \dots, N\}$ 。patch 的大小是 $P \times P$, 输入序列的长度是 $N = \frac{HW}{P^2}$, 每个 patch 的维度 $d_x = P^2 C$ 。第二步是 patch 嵌入: Patch x_p 在线性投影部分被映射到一个维度为 D 的嵌入空间。第三部分是 Transformer 层: Transformer 层由 n 层归一层和多头注意模块组成, 序列输入随后被归一化, 并输入到多头注意模块 (MHA)。

(4) 解码器

解码器用于对隐藏的特征进行解码。解码阶段包含反卷积、 3×3 卷积、ReLU。同时, 跳跃连接保留了下采样阶段的多尺度高分辨率特征信息。最终的掩码是在重复的解码模块之后得到的。这种设计提高了学习率, 同时保留了边缘信息, 并补偿了由于 Transformer 造成的低级细节特征的损失。

4.4 多头注意力模块

在对罂粟进行语义分割的任务中, tokens 的数量很大。同时, 罂粟图像的像素是高度空间结构化的, 而且罂粟特殊的植物形态导致果实和植物根茎之间存在远程依赖关系。相对位置编码 (RPE) 通常是通过一个具有可学习参数的查找表来计算的, 这些参数与

自注意力模块中的查询键交互。这样的方法可使模块能够结合不同嵌入特征的相对位置编码对输入 tokens 之间的相对距离进行编码，捕捉到 tokens 之间很远的依赖关系。改进后的编码模块能够处理更长的序列，保持语义分割所需的翻译不变性，并且能够进一步提高模型的表示能力。

偏置模式的相对位置编码已被添加到 MHA 模块中，改进后的 MHA 如图 4-3 所示。

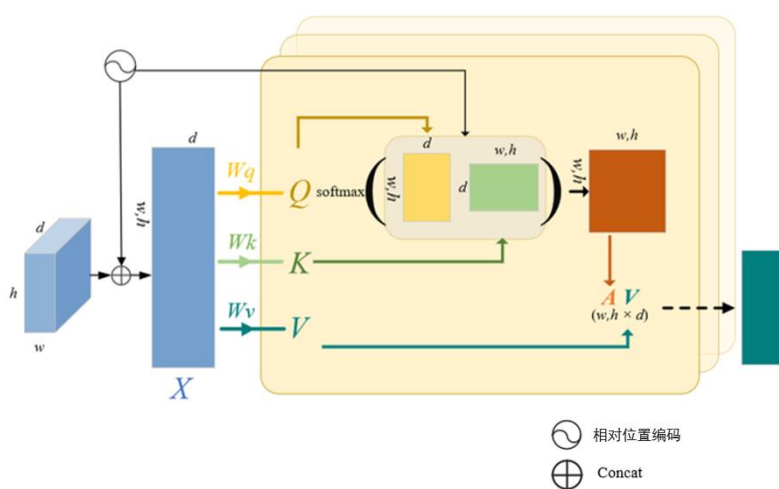


图 4-3 融合相对位置编码的 MHA 模型

MHA 模块通过学习每个像素之间的关系获得一个包含整个图像的感受野，该模块有三个输入：查询矩阵 Q 、关键矩阵 K 和值矩阵 V 。在该模块中， Q 、 K 、 V 矩阵的大小相等，对应于三个嵌入矩阵 W^Q 、 W^K 、 W^V 。注意力矩阵 A 的一行代表 Q 的一个元素与 K 的所有元素的相关性，并在多个头中实现注意力计算，值 V 的一列表示 A 中加权关注的特征图。 x 表示特征图的不同嵌入。自注意力机制将一个查询 Q 和一组键值映射到一个等长的序列输出。输出序列为 $z = (z_1, \dots, z_n)$ ，其中 $z_i \in \mathbb{R}^{d_z}$ ，表示为输入参数的线性转换加权求和。

$$z_i = \sum_{j=1}^n \alpha_{ij} (x_i W^V) \quad (4.1)$$

每个加权系数 α_{ij} ，使用 softmax 函数计算。

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^n \exp(e_{ik})} \quad (4.2)$$

在上述方程中， e_{ij} 使用缩放点积注意计算。

$$e_{ij} = \frac{(x_i W^Q)(x_j W^K)^T}{\sqrt{d_z}} \quad (4.3)$$

其中，投影 $W^Q, W^K, W^V \in \mathbb{R}^{d_x \times d_z}$ 是参数矩阵，每层都是唯一的。模块对于每一个 patch 之间的相对位置关系进行编码，并且不参与权重更新，如图 4-4 所示。

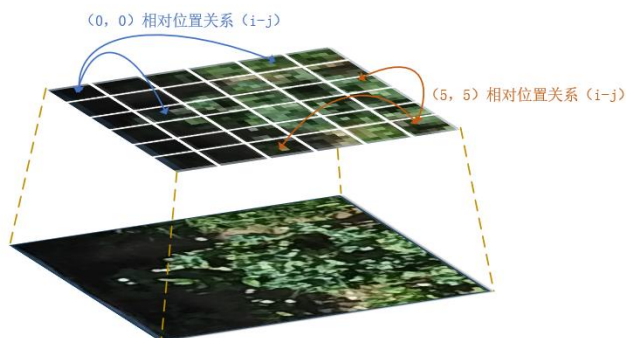


图 4-4 相对位置关系

在加入相对位置编码后，自注意力可以表达为：

$$z_i = \sum_{j=1}^n \alpha_{ij} (x_i W^V + p_{ij}^V) \quad (4.4)$$

$$e_{ij} = \frac{(x_i W^Q + p_{ij}^Q)(x_j W^K + p_{ij}^K)^T}{\sqrt{d_z}} \quad (4.5)$$

在上述方程中， $p_{ij}^V, p_{ij}^Q, p_{ij}^K \in \mathbb{R}^{d_z}$ ，其中 $d_z = d_x$ 。

将 $p_{ij}^V, p_{ij}^Q, p_{ij}^K$ 分别编码添加到查询向量 Q、键向量 K 和值向量 V 中的相对位置，用来表示两个标记之间的相对位置关系，并在自注意力的计算过程中加入。第 1 层的输出表示为公式 (4.6)：

$$z_\ell = MSA(LN(z_{\ell-1})) + z_{\ell-1} \quad (4.6)$$

4.5 实验结果与分析

4.5.1 数据集

无人机图像是在罂粟的开花期和结果期收集的。这些图像包含 54 张 3000×4000 分辨率的近地面图像。本文使用 labelme 软件将图像标记为两类，即罂粟或背景。在第四章的研究中，使用了第三章中表现良好的基于敏感特征的数据增广算法，对 54 张近地图像增广后作为本章的数据集，并按照 4: 1 的比例划分训练集和测试集。图像预处理

的细节在上节给出。

4.5.2 评价指标

在这项研究中，有两个指标被用来比较罂粟花分割的实验效果：Dice 系数，每张图片的训练时间。Dice 系数用来评估两个样本相似度，与第三章验证数据增广所用评价指标相同。

4.5.3 实验设计

本研究提出的算法相关实验使用单个 NVIDIA RTX A4000 GPU 完成。操作系统为 Windows 10，采用深度学习框架 PyTorch-1.12.1，编程语言为 Python_3.8，GPU 型号为 NVIDIA GTX A4000，CUDA 版本为 11.3。

（1）比较实验

对于混合编码器的设计，本文选择 ResNet-50 作为 CNN 中的编码器部分。骨干网络在 ImageNet 上进行了预训练。由于 Transformer 的输入要求，本文将输入分辨率调整为 224×224 ，并在解码器中使用反卷积将其恢复至完整分辨率。

本文将 batch 大小分别设置为 2、4、8、16，epoch 分别为 100、150。由于样本量小，学习率被设定为 0.01。权重衰减为 10^{-4} 。经过实验对比，当 batch 为 4，epoch 为 100，学习率为 0.01 时，获得了最好的分割结果。

本文在无人机罂粟图像数据集上进行了主要实验，在同一数据集上将改进的 TAU-Net 与高级分割网络 U-Net 和 DeepLabV3+进行了比较。使用的损失函数与第三章中基于敏感特征的数据增广方法实验中使用的损失函数不同，使用的损失函数是 Dice Loss 与 BCE 相结合的损失函数。相比单独使用 Dice Loss 损失函数或 BCE 损失函数，DiceBCE 损失具有以下优点：

①罂粟图片中出现罂粟样例只占了很少的区域时，这属于前后景极不均衡情况。如一个 512×512 的图片只有一个 10×10 的分割样例。BCE 全蒙背景依旧可以觉得效果很好，此时 BCE 无法解决这种极度不均衡情况。Dice Loss 却不受前景大小的影响。

②罂粟图片中大块的罂粟样例和小块罂粟样例同时出现时，这属于分割内容不均衡情况时。如一个 512×512 的图片有一个 10×10 和一个 200×200 的分割样例。Dice Loss

此时会趋向于学习大的块，而忽略小样本。但 BCE 依旧会对小样本进行学习。

(2) 消融研究

对 TAU-Net 进行了消融实验，观察不同设置对实验结果的影响，分析添加相对位置编码 MHA 模块的效果。本实验将 batch size 大小设置为 4， epoch 分别为 100，学习率被设定为 0.01。权重衰减为 10^{-4} 。本次实验中使用相同的超参数和设置进行训练了不同的网络模型。

4.5.4 实验结果与讨论

在表 4-1 中展示了 TAU-Net 在经过敏感特征增广后罂粟数据集上的分割结果。分割评价指标是 5 次实验的平均值。此外，表 4-1 还包含了 U-Net 和 DeepLabV3+在同一数据集上的比较实验结果。

表 4-1 U-Net 算法、DeepLabV3+算法和 TAU-Net 算法在增广罂粟数据集上的实验结果

网络模型	Dice 指数	处理时间（秒）
U-Net	0.74	1.4
DeepLabV3+	0.66	2.03
TAU-Net	0.77	1.87

结合表 4-1 和图 4-5 可以看出，在 U-Net 的分割结果中，假阳性的情况比较多。第一行，在拍摄角度有小幅改变改变的图像中，U-Net 的分株能力强，大范围内其他植物的正确预测率却较高；第二行，在拍摄高度较高的图像中，小块的杂草被错误地分割为罂粟；第三行，在小面积种植的罂粟图像中，U-Net 只能识别出光照条件较好的区域，分割空洞区域多。

结合表 4-1 和图 4-6 可以看出，DeepLabV3+与 U-Net 都能较好地预测土壤-罂粟的边界。第一行的结果表明，当拍摄角度发生变化时，DeepLabV3+的预测精度变得很低，这导致 DeepLabV3+的平均指标低于 U-Net 网络；第二行与第三行的结果表明，对于罂粟种植密集区的垂直拍摄，DeepLabV3+的分割结果的准确性优于 U-Net。但是，与此同时，DeepLabV3+参数数量很大，训练速度也比 U-Net 慢。

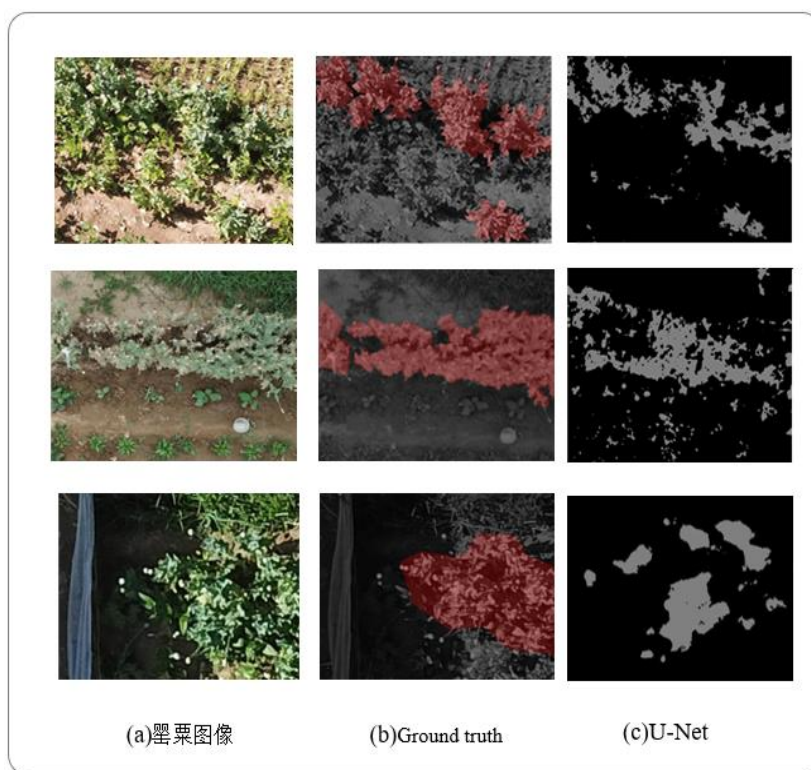


图 4-5 U-Net 对比实验分割结果

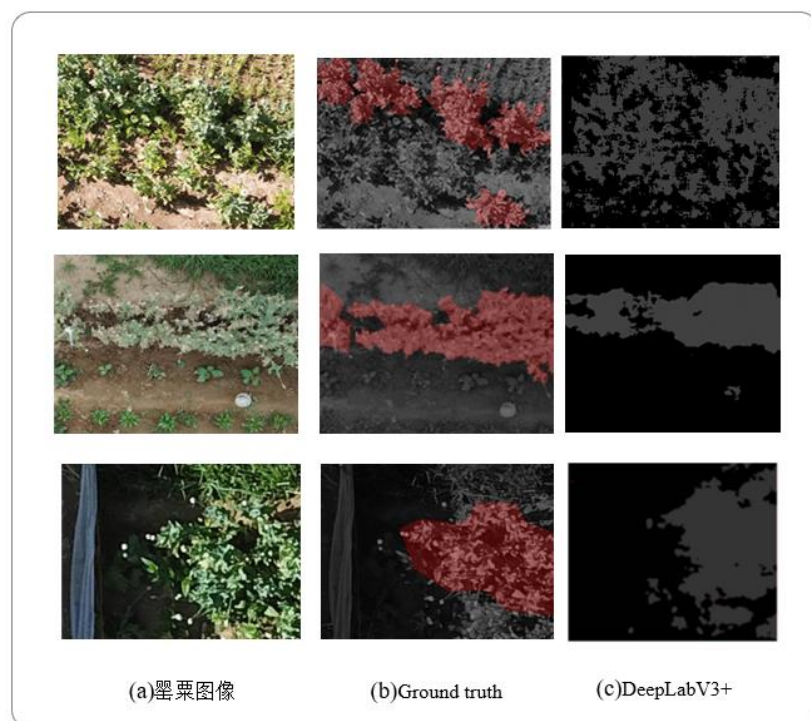


图 4-6 DeepLabV3+对比实验分割结果

结合表 4-1 和图 4-7 可以看出, TAU-Net 不仅能很好地预测土壤和罂粟的边界, 而

且假阳性较少,这表明 TAU-Net 在抑制噪声方面比其他方法有优势。第一行,在拍摄角度有小幅改变图像中,TAU-Net 的分株能力不如 U-Net,但在人工标记错误的区域也有正确的预测;第二行和第三行,在拍摄高度较高的图像和小面积种植的罂粟图像中,TAU-Net 对细节的分割更精确,精准度也高于 U-Net、DeepLabV3+。同时,Transformer 的引入并没有给网络带来过多的计算负担。

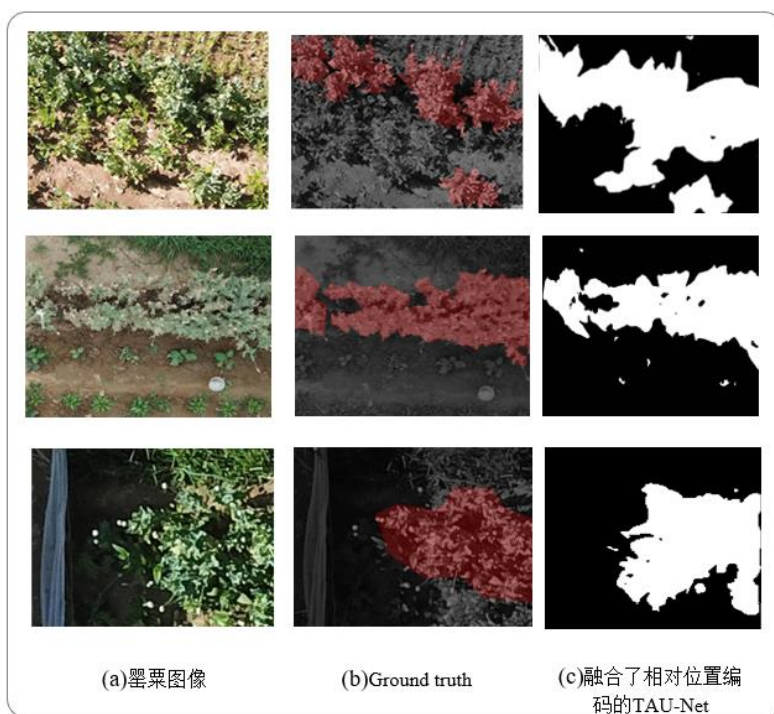


图 4-7 TAU-Net 对比实验分割结果

表 4-2 展示关于相对位置编码 MHA 模块的消融实验数据结果。结合表 4-2 和图 4-8 可以看出,使用相对位置编码改进后的 MHA 模块有效地改善了罂粟边缘分割,保留了更多的果实信息。

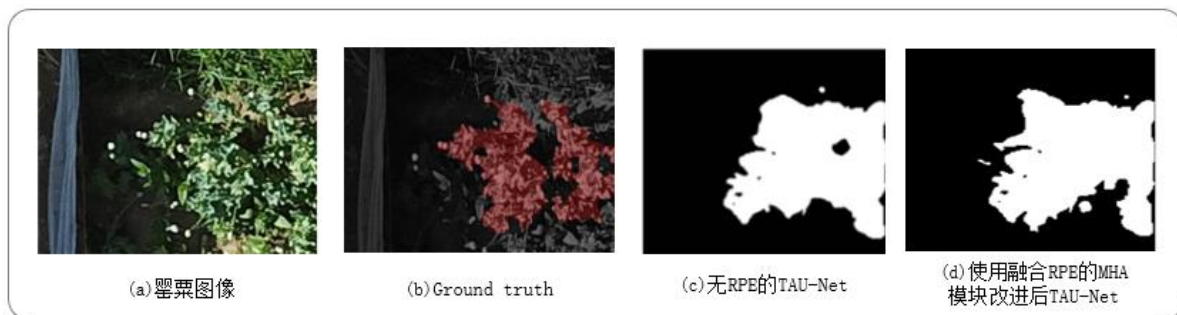


图 4-8 使用融合 RPE 的 MHA 模块改进后 TAU-Net 的消融实验分割结果

与此同时，融合相对位置编码后并未对网络增加过重的计算负担，处理时间由 1.84 秒增至 1.87 秒，适用于罂粟语义分割任务的高精度要求。但两个网络对罂粟植株的分株能力都比较弱，不能很好的识别植株间的空洞区域。

表 4-2 相对位置编码的 MHA 模块的消融性能

方法	Dice 指数	处理时间（秒）
TAU-Net（无 RPE 模块）	0.74	1.84
TAU-Net（RPE 模块）	0.77	1.87

4.6 本章小结

本章主要介绍了为罂粟无人机图像任务设计的一个端到端的编解码网络 TAU-Net。该方法利用编解码网络的特性融合罂粟图像中不同尺度的特征信息，并使用自注意力机制中相对位置编码处理罂粟的果实和植物根茎之间存在的远程依赖关系。所提算法在增广后的罂粟数据集上进行了验证，在 U-Net 模型上得到 0.74 的 Dice 指数，在 DeepLabV3+ 模型上得到 0.66 的 Dice 指数，在 TAU-Net 模型上得到 0.77 的 Dice 指数。在 MHA 模块的消融实验中，未添加相对位置编码的网络得到 0.74 的 Dice 指数，低于改进后的 TAU-Net 模型，验证了本次实验的有效性。

第五章 总结与展望

5.1 研究总结

违法种植罂粟的现象在我国并未完全杜绝，在山区尤为严重。为了降低私种罂粟给社会带来的不良影响及极大的潜在风险，发现并惩戒违法罂粟种植迫在眉睫。使用无人机从高空拍摄高清图像，隐蔽性好、效率高。本研究对无人机航拍收集到的图像使用改进的深度网络模型进行训练，分割出罂粟和其他合法种植植物，有效判别无人机拍摄区域是否存在非法种植罂粟的违法行为。

由于罂粟非法种植规模小且经常与其它植物混种，因此主要识别特征是植物的形态以及雄蕊和果实的特征，叶子边缘呈不规则的波浪状，茎长。特殊的形态意味着需要更精确的分割。本文利用 Transformer 改进了一个 U 型网络 TAU-Net，用于无人机捕获的罂粟图像的语义分割任务。该骨干网络包含了 CNN 网络 and 自注意力机制。罂粟的特征在不同尺度上差异很大。与仅由自注意模块或仅由卷积模块构建的网络不同，TAU-Net 利用 Transformer 的优势，从特征中去除不相关或嘈杂的区域，更好地提取罂粟信息，同时也解开了原始 U-Net 感受域的限制。改进后的网络提高了对规模变化的鲁棒性，而不需要高计算量。无人机采集的罂粟图像具有很高的分辨率，像素具有很高的空间结构，而且罂粟的果实和植物根茎之间存在着远程依赖关系。相对位置方法学习了 tokens 之间的关系，保留了更准确的位置信息，并实现了分割边界信息的优化。

主要工作包含以下几个方面：

(1) 首先阐述了小样本数据农业图像增广技术国内外研究现状及农业图像语义分割算法国内外研究现状，其中，农业图像语义分割算法研究现状包括基于传统方法的语义分割算法及基于深度学习方法的分割算法国内外研究现状。其次阐述此次研究的图像特征、分割难点、技术路线。并且，对此次研究所涉及的相关工作及原理进行了总结，内容包括农业图像小样本数据增强技术、基于编解码结构的语义分割算法、自注意力机制。接着总结了罂粟图像的特点，介绍了算法的主体框架，并在罂粟数据集上对所提算法进行了验证和分析研究。

(2) 设计了为解决小样本问题的基于敏感特征的数据增广方法。方法针对于样本

过少的难题，设计了融合了旋转、随机窗口、通道分离的增广方法，将罂粟图像进行随机切割划分和旋转操作，并计算不同的植被索引因子，分离通道。处理后的图像作为额外的表征支持 CNN 的训练。同时进行多次实验验证使用数据增广算法后对分割网络的分割精度产生的影响。

(3) 设计了一个融合自注意力机制的端到端语义分割网络——TAU-Net，对输入数据进行语义标注，用于无人机捕获的罂粟图像的语义分割任务，实现罂粟图像的精准分割。网络由 CNN 网络框架与 Transformer 模块组成，Transformer 模块代替原始网络的编码器，将特征图序列化，通过学习 patch 之间的关系，使网络更加关注重点区域。最后，将自注意力模块处理后的特征图进行上采样，并于跳跃连接的特征图进行融合，得到分割结果。

(4) 设计了一种结合相对位置编码的多头注意力模块，该算法以多头注意力为基础，在位置编码阶段计算 tokens 之间的空间信息，利用额外的 tokens 位置关系处理罂粟特殊植物学形态（茎、果实）信息，优化分割掩膜的边缘。

(5) 分别对基于敏感特征的数据增广办法、基于自注意力的编解码语义分割算法及两个算法相结合的语义分割算法在无人机采集的罂粟数据集上进行实验验证，其分割精度均比原始 U-Net 算法的要好，比原始 DeepLabV3+ 算法训练速度、分割速度更快。综合二者优点，减少训练负担的同时实现了比 CNN 原始网络更高的指标。并通过分割结果表明本研究所提算法罂粟数据集上得到了明显的分割提升，证明了本方法的有效性。

5.2 工作展望

所提方法在针对罂粟植株进行分割时，尽管比已有方法有了明显提升，但仍存在部分不完善之处，需要进一步深入研究。

- (1) 实验结果中存在部分假阳性，降低了实验的准确性；
- (2) 所提出的方法在处理作物间作方面能力不强；
- (3) 在不同的光照强度下，罂粟花的特征表现差异很大，这对网络的学习能力产生很大影响；
- (4) 网络对照片角度变化的鲁棒性也需要加强。

除此之外，针对于其他农业图像分割任务时，现有算法仍有较大的提升空间：

- (1) 监督方法总是需要大量的标记样本，这对农业图像分割任务是一个巨大的负担。

农业图像往往具有复杂的纹理和复杂的场景。在未来的工作中，弱监督或无监督的学习方法将有很大的潜力，迁移学习也可以用来减少训练样本的数量，将学到的知识转移到新构建的网络中。

- (2) 多模式融合方法在用于语义图像分割时，来自不同信息源的特征根据其特点被融合到分割网络的不同阶段，利用多种信息源产生比单一信息源更准确的预测结果。由于农业图像拍摄环境的复杂性，可以采用多模式融合的方法进行农业语义分割，对于同一分割场景，提供多类信息，相互补充。

参考文献

- [1] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. nature, 2015, 521(7553): 436-444.
- [2] Minaee S, Boykov Y Y, Porikli F, et al. Image segmentation using deep learning: A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.
- [3] Milioto A, Lottes P, Stachniss C. Real-time semantic segmentation of crop and weed for precision agriculture robots leveraging background knowledge in CNNs[C]//2018 IEEE international conference on robotics and automation (ICRA). IEEE, 2018: 2229-2235.
- [4] Qiao Y, Truman M, Sukkarieh S. Cattle segmentation and contour extraction based on Mask R-CNN for precision livestock farming[J]. Computers and Electronics in Agriculture, 2019, 165: 104958.
- [5] Riehle D, Reiser D, Griepentrog H W. Robust index-based semantic plant/background segmentation for RGB-images[J]. Computers and electronics in agriculture, 2020, 169: 105201.
- [6] Yang W, Wang S, Zhao X, et al. Greenness identification based on HSV decision tree[J]. Information Processing in Agriculture, 2015, 2(3-4): 149-160.
- [7] Mortensen A K, Dyrmann M, Karstoft H, et al. Semantic segmentation of mixed crops using deep convolutional neural network[C]//CIGR-AgEng conference. 2016: 26-29.
- [8] Zhuang S, Wang P, Jiang B. Segmentation of green vegetation in the field using deep neural networks[C]//2018 13th World Congress on Intelligent Control and Automation (WCICA). IEEE, 2018: 509-514.
- [9] He K, Gkioxari G, Dollár P, et al. Mask R-CNN[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2961-2969.
- [10] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[J]. Advances in neural information processing systems, 2015, 28.
- [11] Tassis L M, de Souza J E T, Krohling R A. A deep learning approach combining instance and semantic segmentation to identify diseases and pests of coffee leaves from in-field images[J]. Computers and Electronics in Agriculture, 2021, 186: 106191.
- [12] Potena C, Nardi D, Pretto A. Fast and accurate crop and weed identification with summarized train sets

- for precision agriculture[C]//Intelligent Autonomous Systems 14: Proceedings of the 14th International Conference IAS-14 14. Springer International Publishing, 2017: 105-121.
- [13] Lottes P, Hoferlin M, Sander S, et al. An effective classification system for separating sugar beets and weeds for precision farming applications[C]//2016 IEEE international conference on robotics and automation (ICRA). IEEE, 2016: 5157-5163.
- [14] Sa I, Chen Z, Popović M, et al. weednet: Dense semantic weed classification using multispectral images and mav for smart farming[J]. IEEE robotics and automation letters, 2017, 3(1): 588-595.
- [15] Zhang J, Rao Y, Man C, et al. Identification of cucumber leaf diseases using deep learning and small sample size for agricultural Internet of Things[J]. International Journal of Distributed Sensor Networks, 2021, 17(4): 15501477211007407.
- [16] Espejo-Garcia B, Mylonas N, Athanasakos L, et al. Combining generative adversarial networks and agricultural transfer learning for weeds identification[J]. Biosystems Engineering, 2021, 204: 79-89.
- [17] Wu Q, Chen Y, Meng J. DCGAN-based data augmentation for tomato leaf disease identification[J]. IEEE Access, 2020, 8: 98716-98728.
- [18] Abbas A, Jain S, Gour M, et al. Tomato plant disease detection using transfer learning with C-GAN synthetic images[J]. Computers and Electronics in Agriculture, 2021, 187: 106279.
- [19] Nazki H, Yoon S, Fuentes A, et al. Unsupervised image translation using adversarial networks for improved plant disease recognition[J]. Computers and Electronics in Agriculture, 2020, 168: 105117.
- [20] Yang S, Chen Q, Yuan X, et al. Adaptive coherency matrix estimation for polarimetric SAR imagery based on local heterogeneity coefficients[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(11): 6732-6745.
- [21] Kussul N, Lavreniuk M, Skakun S, et al. Deep learning classification of land cover and crop types using remote sensing data[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(5): 778-782.
- [22] Jadhav J K, Singh R P. Automatic semantic segmentation and classification of remote sensing data for agriculture[J]. Mathematical Models in Engineering, 2018, 4(2): 112-137.
- [23] Dechesne C, Mallet C, Le Bris A, et al. Semantic segmentation of forest stands of pure species as a global optimization problem[J]. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2017, 4: 141.

- [24] Zou K, Chen X, Wang Y, et al. A modified U-Net with a specific data argumentation method for semantic segmentation of weed images in the field[J]. Computers and Electronics in Agriculture, 2021, 187: 106242.
- [25] Anand T, Sinha S, Mandal M, et al. AgriSegNet: Deep aerial semantic segmentation framework for IoT-assisted precision agriculture[J]. IEEE Sensors Journal, 2021, 21(16): 17581-17590.
- [26] Singh P, Verma A, Alex J S R. Disease and pest infection detection in coconut tree through deep learning techniques[J]. Computers and electronics in agriculture, 2021, 182: 105986.
- [27] Ma J C, Zheng F X, Zhang L X, et al. Disease recognition system for greenhouse cucumbers based on deep convolutional neural network[J]. Transactions of the Chinese Society of Agricultural Engineering, 2018, 34(12): 186-192.
- [28] Yuan Y, Chen L, Wu H, et al. Advanced agricultural disease image recognition technologies: A review[J]. Information Processing in Agriculture, 2022, 9(1): 48-59.
- [29] Liu G, Bai L, Zhao M, et al. Segmentation of wheat farmland with improved U-Net on drone images[J]. Journal of Applied Remote Sensing, 2022, 16(3): 034511.
- [30] Riehle D, Reiser D, Griepentrog H W. Robust index-based semantic plant/background segmentation for RGB-images[J]. Computers and electronics in agriculture, 2020, 169: 105201.
- [31] Yang W, Wang S, Zhao X, et al. Greenness identification based on HSV decision tree[J]. Information Processing in Agriculture, 2015, 2(3-4): 149-160.
- [32] Mortensen A K, Dyrmann M, Karstoft H, et al. Semantic segmentation of mixed crops using deep convolutional neural network[C]//CIGR-AgEng conference. 2016: 26-29.
- [33] Zhuang S, Wang P, Jiang B. Segmentation of green vegetation in the field using deep neural networks[C]//2018 13th World Congress on Intelligent Control and Automation (WCICA). IEEE, 2018: 509-514.
- [34] Wang A, Xu Y, Wei X, et al. Semantic segmentation of crop and weed using an encoder-decoder network and image enhancement method under uncontrolled outdoor illumination[J]. Ieee Access, 2020, 8: 81724-81734.
- [35] Gilles J. Empirical wavelet transform[J]. IEEE transactions on signal processing, 2013, 61(16): 3999-4010.

- [36] Țălu Ș. Mathematical methods used in monofractal and multifractal analysis for the processing of biological and medical data and images[J]. ABAH Bioflux, 2012, 4(1): 1-4.
- [37] Gao J, Wang B, Wang Z, et al. A wavelet transform-based image segmentation method[J]. Optik, 2020, 208: 164123.
- [38] Guijarro M, Riomoros I, Pajares G, et al. Discrete wavelets transform for improving greenness image segmentation in agricultural images[J]. Computers and Electronics in Agriculture, 2015, 118: 396-407.
- [39] Otsu N. A threshold selection method from gray-level histograms[J]. IEEE transactions on systems, man, and cybernetics, 1979, 9(1): 62-66.
- [40] Wu Z, Gao Y, Li L, et al. Semantic segmentation of high-resolution remote sensing images using fully convolutional network with adaptive threshold[J]. Connection Science, 2019, 31(2): 169-184.
- [41] Hassanein M, Lari Z, El-Sheimy N. A new vegetation segmentation approach for cropped fields based on threshold detection from hue histograms[J]. Sensors, 2018, 18(4): 1253.
- [42] Dorj U O, Lee M, Yun S. An yield estimation in citrus orchards via fruit detection and counting using image processing[J]. Computers and electronics in agriculture, 2017, 140: 103-112.
- [43] Castillo-Martínez M Á, Gallegos-Funes F J, Carvajal-Gámez B E, et al. Color index based thresholding method for background and foreground segmentation of plant images[J]. Computers and Electronics in Agriculture, 2020, 178: 105783.
- [44] Achanta R, Shaji A, Smith K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J]. IEEE transactions on pattern analysis and machine intelligence, 2012, 34(11): 2274-2282.
- [45] Giraud R, Ta V T, Papadakis N. Robust superpixels using color and contour features along linear path[J]. Computer Vision and Image Understanding, 2018, 170: 1-13.
- [46] Di S, Liao M, Zhao Y, et al. Image superpixel segmentation based on hierarchical multi-level LI-SLIC[J]. Optics & Laser Technology, 2021, 135: 106703.
- [47] Khan S, Narvekar M. Novel fusion of color balancing and superpixel based approach for detection of tomato plant diseases in natural complex environment[J]. Journal of King Saud University-Computer and Information Sciences, 2022, 34(6): 3506-3516.
- [48] Mohanty S P, Hughes D P, Salathé M. Using deep learning for image-based plant disease detection[J]. Frontiers in plant science, 2016, 7: 1419.

-
- [49] Lu H, Cao Z, Xiao Y, et al. Region-based colour modelling for joint crop and maize tassel segmentation[J]. Biosystems Engineering, 2016, 147: 139-150.
- [50] Tu W C, Liu M Y, Jampani V, et al. Learning superpixels with segmentation-aware affinity loss[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018: 568-576.
- [51] Jampani V, Sun D, Liu M Y, et al. Superpixel sampling networks[C]//Proceedings of the European Conference on Computer Vision (ECCV). 2018: 352-368.
- [52] Yang F, Sun Q, Jin H, et al. Superpixel segmentation with fully convolutional networks[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 13964-13973.
- [53] Gonçalves D N, de Moares Weber V A, Pistori J G B, et al. Carcass image segmentation using CNN-based methods[J]. Information Processing in Agriculture, 2021, 8(4): 560-572.
- [54] Zheng Y Y, Kong J L, Jin X B, et al. CropDeep: The crop vision dataset for deep-learning-based classification and detection in precision agriculture[J]. Sensors, 2019, 19(5): 1058.
- [55] Ale L, Sheta A, Li L, et al. Deep learning based plant disease detection for smart agriculture[C]//2019 IEEE Globecom Workshops (GC Wkshps). IEEE, 2019: 1-6.
- [56] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3431-3440.
- [57] 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.
- [58] Song Z, Zhang Z, Yang S, et al. Identifying sunflower lodging based on image fusion and deep semantic segmentation with UAV remote sensing imaging[J]. Computers and Electronics in Agriculture, 2020, 179: 105812.
- [59] Mukhopadhyay P, Mallick S. Bayesian deep learning for seismic facies classification and its uncertainty estimation[M]//SEG Technical Program Expanded Abstracts 2019. Society of Exploration Geophysicists, 2019: 2488-2492.
- [60] Wang J, Liu W, Gou A. Numerical characteristics and spatial distribution of panoramic Street Green

- View index based on SegNet semantic segmentation in Savannah[J]. *Urban Forestry & Urban Greening*, 2022, 69: 127488.
- [61] Chen L C, Papandreou G, Kokkinos I, et al. Semantic image segmentation with deep convolutional nets and fully connected crfs[J]. *arXiv preprint arXiv:1412.7062*, 2014.
- [62] 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.
- [63] Chen L C, Papandreou G, Schroff F, et al. Rethinking atrous convolution for semantic image segmentation[J]. *arXiv preprint arXiv:1706.05587*, 2017.
- [64] 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.
- [65] He K, Zhang X, Ren S, et al. Identity mappings in deep residual networks[C]//*Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14*. Springer International Publishing, 2016: 630-645.
- [66] Peng H, Xue C, Shao Y, et al. Semantic segmentation of litchi branches using DeepLabV3+ model[J]. *IEEE Access*, 2020, 8: 164546-164555.
- [67] Alzu'bi A, Alsmadi L. Monitoring deforestation in Jordan using deep semantic segmentation with satellite imagery[J]. *Ecological Informatics*, 2022, 70: 101745.
- [68] Yuan H, Zhu J, Wang Q, et al. An improved DeepLab v3+ deep learning network applied to the segmentation of grape leaf black rot spots[J]. *Frontiers in Plant Science*, 2022, 13.
- [69] Sun K, Wang X, Liu S, et al. Apple, peach, and pear flower detection using semantic segmentation network and shape constraint level set[J]. *Computers and Electronics in Agriculture*, 2021, 185: 106150.
- [70] Jo Y, Lee S, Lee Y, et al. Semantic Segmentation of Cabbage in the South Korea Highlands with Images by Unmanned Aerial Vehicles[J]. *Applied Sciences*, 2021, 11(10): 4493.
- [71] Huang G, Liu Z, Van Der Maaten L, et al. Densely connected convolutional networks[C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017: 4700-4708.
- [72] Zhang X, Zhou X, Lin M, et al. Shufflenet: An extremely efficient convolutional neural network for

- mobile devices[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 6848-6856.
- [73] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [74] Potlapally A, Chowdary P S R, Shekhar S S R, et al. Instance segmentation in remote sensing imagery using deep convolutional neural networks[C]//2019 International Conference on contemporary Computing and Informatics (IC3I). IEEE, 2019: 117-120.
- [75] Pahlevan N, Chittimalli S K, Balasubramanian S V, et al. Sentinel-2/Landsat-8 product consistency and implications for monitoring aquatic systems[J]. Remote sensing of Environment, 2019, 220: 19-29.
- [76] Nogueira K, Dalla Mura M, Chanussot J, et al. Dynamic multicontext segmentation of remote sensing images based on convolutional networks[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(10): 7503-7520.
- [77] Osco L P, Nogueira K, Marques Ramos A P, et al. Semantic segmentation of citrus-orchard using deep neural networks and multispectral UAV-based imagery[J]. Precision Agriculture, 2021, 22(4): 1171-1188.
- [78] Ma J, Li Y, Liu H, et al. Improving segmentation accuracy for ears of winter wheat at flowering stage by semantic segmentation[J]. Computers and Electronics in Agriculture, 2020, 176: 105662.
- [79] Larochelle H, Hinton G E. Learning to combine foveal glimpses with a third-order Boltzmann machine[J]. Advances in neural information processing systems, 2010, 23.
- [80] Stollenga M F, Masci J, Gomez F, et al. Deep networks with internal selective attention through feedback connections[J]. Advances in neural information processing systems, 2014, 27.
- [81] Graves A, Wayne G, Danihelka I. Neural turing machines[J]. arXiv preprint arXiv:1410.5401, 2014.
- [82] Goodfellow I, Bengio Y, Courville A. Deep learning[M]. MIT press, 2016.
- [83] Sukhbaatar S, Weston J, Fergus R. End-to-end memory networks[J]. Advances in neural information processing systems, 2015, 28.
- [84] Cheng J, Dong L, Lapata M. Long short-term memory-networks for machine reading[J]. arXiv preprint arXiv:1601.06733, 2016.
- [85] Parikh A P, Täckström O, Das D, et al. A decomposable attention model for natural language inference[J].

- arXiv preprint arXiv:1606.01933, 2016.
- [86] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [87] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9.
- [88] He K, Zhang X, Ren S, et al. Identity mappings in deep residual networks[C]//Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14. Springer International Publishing, 2016: 630-645.
- [89] Chollet F. Xception: Deep learning with depthwise separable convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 1251-1258.
- [90] Shehab L H, Fahmy O M, Gasser S M, et al. An efficient brain tumor image segmentation based on deep residual networks (ResNets)[J]. Journal of King Saud University-Engineering Sciences, 2021, 33(6): 404-412.
- [91] Mnih V, Heess N, Graves A. Recurrent models of visual attention[J]. Advances in neural information processing systems, 2014, 27.
- [92] Wang F, Jiang M, Qian C, et al. Residual attention network for image classification[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 3156-3164.
- [93] 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.
- [94] Park J, Woo S, Lee J Y, et al. Bam: Bottleneck attention module[J]. arXiv preprint arXiv:1807.06514, 2018.
- [95] Dosovitskiy A, Beyer L, Kolesnikov A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. arXiv preprint arXiv:2010.11929, 2020.
- [96] Touvron H, Cord M, Douze M, et al. Training data-efficient image transformers & distillation through attention[C]//International conference on machine learning. PMLR, 2021: 10347-10357.
- [97] Kamilaris A, Prenafeta-Boldú F X. A review of the use of convolutional neural networks in agriculture[J]. The Journal of Agricultural Science, 2018, 156(3): 312-322.
- [98] Ronneberger O, Fischer 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, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer International Publishing, 2015: 234-241.