

学号： 2019012812



**西北农林科技大学**  
NORTHWEST A&F UNIVERSITY

## 2023 届本科生毕业论文

# 大豆叶片分割与面积估算系统设计与实现

学	院：	<u>信息工程学院</u>
专	业：	<u>软件工程</u>
年 级 班 级：		<u>2019级02班</u>
学 生 姓 名：		<u>杨治梁</u>
指 导 教 师：		<u>冯明晨</u>
协助指导教师：		<u></u>
完 成 日 期：		<u>2023年06月</u>



## 本科生毕业论文(设计)的独创性声明

本人声明：所呈交的本科毕业论文(设计)是我个人在导师指导下独立进行的研究工作及取得的研究结果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究结果，也不包含其他人和自己本人已获得西北农林科技大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同事对本研究所做的任何贡献均已在论文的致谢中作了明确的说明并表示了谢意。如违反此声明，一切后果与法律责任均由本人承担。

本科生签名：

时间：        年        月        日

## 关于本科生毕业论文(设计)知识产权的说明

本毕业论文(设计)的知识产权归属西北农林科技大学。本人同意西北农林科技大学保存或向国家有关部门或机构送交论文的纸质版和电子版，允许论文被查阅和借阅。

本人保证，在毕业离开西北农林科技大学后，发表或者使用本毕业论文(设计)及其相关的工作成果时，将以西北农林科技大学为第一署名单位，否则，愿意按《中华人民共和国著作权法》等有关规定接受处理并承担法律责任。

任何收存和保管本论文各种版本的其他单位和个人(包括作者本人)未经本论文作者的导师同意，不得有对本论文进行复制、修改、发行、出租、改编等侵犯著作权的行为，否则，按违背《中华人民共和国著作权法》等有关规定处理并追究法律责任。

本科生签名：

时间：        年        月        日

指导教师签名：

时间：        年        月        日



# 大豆叶片分割与面积估算系统设计与实现

**摘 要：**叶面积是植物整个生命周期中重要的生长特性。如何有效、无损、低成本地计算面积是一项复杂且具有挑战性的工作，需要进一步的研究。论文使用图像处理方法探索了叶面积估算问题，首先使用图像分割的方法对叶片进行分割；然后利用像素与面积之间的关系对叶片面积进行估算；最后集成所提出的分割算法和面积估计模型，开发了叶片分割和面积估算软件。具体研究工作如下：（1）收集并构建了大豆叶片图像数据集，其中包含具有不同拍摄高度和环境的大豆叶子图像。（2）针对复杂场景下植物叶片分割不准确的问题，提出了极值点的叶片分割方法，利用极值点将目标聚焦在感兴趣的叶片上，然后对目标区域进行了分割；（3）探索了像素点与面积之间的关系，较准确地估算出叶子的面积。

在实验中对比了其他几种分割算法，包括 U-Net、R2-UNet、Attention-UNet 和 FCN-32s，结果表明，本文提出的算法在叶片图像分割方面优于其他算法。通过本文的研究，实现了大豆叶片分割与面积估算系统，能够实现快速、准确地对大豆叶片进行分割，同时不对叶片本身造成伤害。这些成果将有助于植物生理学和农艺学研究，使得研究人员更好地了解植物生长状况，并为阐明不同植物物种之间的竞争、预测作物产量和协助决策提供理论依据。

**关键词：**叶片图像分割；极值点；叶面积估算；卷积神经网络



# Design and Realization of Soybean Leaf Segmentation and Area Estimation System

**Abstract:** Leaf area is an important growth characteristic of plants throughout their entire life cycle. The effective, non-destructive, and low-cost calculation of leaf area is a complex and challenging task that requires further research. In this paper, we explored the problem of leaf area estimation using image processing methods. We first segmented the leaves using an image segmentation method; then, we estimated the leaf area using the relationship between pixels and area, and finally, we integrated the proposed segmentation algorithm and area estimation model to develop leaf segmentation and area estimation software. The specific research work is as follows: (1) we collected and constructed a soybean leaf image data set, which contains soybean leaf images captured at different heights and under different environmental conditions; (2) we addressed the problem of inaccurate plant leaf segmentation in complex scenes using an extreme point-based leaf segmentation method. We used extreme points to focus on the interested leaves and then segment the target area; (3) we explored the relationship between pixels and area and accurately estimated the leaf area.

In this paper, we compared our proposed framework with other segmentation algorithms, including U-Net, R2-UNet, Attention-UNet, and FCN-32s. The results show that our proposed algorithm is superior to other algorithms in leaf image segmentation. Through this research, we have developed a soybean leaf segmentation and area estimation system that can quickly and accurately segment soybean leaves without causing damage to the leaves themselves. These achievements will be beneficial to the research of plant physiology and agronomy, enabling researchers to gain a better understanding of plant growth conditions and providing a theoretical basis for elucidating competition among different plant species, predicting crop yield, and assisting in decision-making.

**Keywords:** Leaf image segmentation; Extreme points; Leaf area estimation; Convolutional neural network





# 目 录

第 1 章 绪论.....	- 1 -
1.1 研究的背景与意义 .....	- 1 -
1.1.1 研究背景.....	- 1 -
1.1.2 研究意义.....	- 1 -
1.2 国内外研究现状 .....	- 2 -
1.2.1 叶片图像分割方法.....	- 2 -
1.2.2 叶片面积估算方法.....	- 4 -
1.3 研究内容 .....	- 4 -
第 2 章 相关技术与方法.....	- 7 -
2.1 基于深度学习的图像分割算法 .....	- 7 -
2.1.1 U-Net 分割算法.....	- 7 -
2.1.2 R2-UNet 分割算法 .....	- 8 -
2.1.3 Attention-UNet 分割算法.....	- 9 -
2.1.4 FCN-32s 分割算法 .....	- 11 -
2.2 基于极值点分割算法的相关方法 .....	- 12 -
2.2.1 CNN 网络 .....	- 12 -
2.2.2 ResNet 网络 .....	- 12 -
2.3 基于二值化的面积估算方法 .....	- 13 -
2.4 PyQt5 .....	- 14 -
第 3 章 基于极值点的分割算法的实现.....	- 17 -
3.1 数据采集 .....	- 17 -
3.2 数据预处理 .....	- 17 -
3.3 算法模型及实现 .....	- 19 -
3.5 对比算法的实现 .....	- 21 -
3.6 实验结果对比 .....	- 22 -
第 4 章 基于像素点的面积估算方法实现.....	- 25 -
4.1 面积估算 .....	- 25 -
4.2 估算结果对比 .....	- 26 -

第 5 章 基于 PyQt5 的交互式叶片分割与面积计算软件实现.....	- 29 -
5.1 可视化系统模块框架 .....	- 29 -
5.2 可视化系统设计与实现 .....	- 30 -
5.2.1 系统可视化界面.....	- 30 -
5.2.2 具体功能实现.....	- 30 -
第 6 章 总结与展望 .....	- 35 -
6.1 研究总结 .....	- 35 -
6.2 研究展望 .....	- 35 -
参考文献 .....	- 37 -
攻读本科学位期间发表的学术论文和参加科研情况 .....	- 41 -
致    谢 .....	- 43 -

## 第 1 章 绪论

### 1.1 研究的背景与意义

#### 1.1.1 研究背景

植物在农业、工业、医疗等领域占据重要地位。而植物表型是指植物可测量的特征和性状，是植物受自身基因表达、环境影响相互作用的结果，也是决定农作物产量、品质和抗逆性等性状的重要因素（翁杨 2019）。通过图像处理对植物叶片进行分割是研究植物表型性状的基础。而叶子作为植物最为重要的器官，是植物进行光合作用、呼吸作用以及蒸腾作用的场所。叶片面积的大小将会对这些指标产生重要的影响，因此，对于一些交叉学科比如生态学、遗传学、植物学、环境科学等，研究叶片面积大小具有重要的科学意义。

图像分割作为图像分析的第一步，是一种像素级的视觉任务，包括语义分割：预测每个像素的含义和实例分割：标识图像中每个对象的每个实例。传统的图像分割法有基于阈值和边缘等信息的分割，其中使用了大量人工干预算法和专业知识。得益于深度学习神经网络的发展，U-Net、FCN 等模型的出现大大提高了图像分割的效率。

叶面积计算在图像分割的基础上进行，传统的叶片面积测量主要有以下几种途径：一种是将叶子摘下，将叶子轮廓画在坐标纸上，然后在网格上计算叶片的面积。或者是将叶子画在均匀的透明纸上，根据单位面积纸重来估算叶子的面积；第二种方法是使用测量仪器进行测量，这种方法也需要将植物的叶片剪下来，然后进行测量。这几种传统方法具有明显的缺陷，一是对于有锯齿状的树叶测量精度不高，二是在测量过程中对植物叶片具有明显的损伤，三是使用过程中成本高。

#### 1.1.2 研究意义

大豆一直是我国重要的经济作物，是人类优质蛋白的重要来源。大豆是关系国计民生的重要基础性、战略性物资，因此对大豆表型的研究对农业生产以及工业制造有重要意义。叶片作为其重要器官，它的表型数据如面积，能极大程度地反映作物的生长状况，因此大豆叶片面积是研究大豆生长状态的理想化指标。研究发现，大豆植株高 30-90 厘米，形态低矮，而大豆种植具有“密植”的特点（任天佑 1982），所以导致在自然环境下叶片间相互遮挡，采集的图像中叶片边缘特征不明显，对叶片分割效果造成很大的障碍。传统的检测方法需要严格的图像采集标准或较多的人工处理流程。要实现快速、准确、自动化地检测大豆叶片面积的主要难点是复杂背景下大豆叶片图像分割识别与面积

估算。

因此针对自建大豆叶片数据集，本文认为可以基于深度学习和数字图像处理的相关知识，利用 Pytorch 框架创建基于极值点的分割模型对数据集进行分割，比较 U-Net 和 FCN 等不同模型的效果。优化模型算法，得到一种适合本数据集，能实现对复杂环境下大豆叶片的分割的模型；针对叶片分割结果，实现基于像素点的大豆叶片面积估算方法研究。最终，实现叶片面积的无损估算。为后期进行植物生长状态诊断、生长预测、检测病虫害和作物优化提供了理论支持，具有重要的研究意义。

## 1.2 国内外研究现状

为了解决生命科学领域中使用传统方法测量叶片面积存在的问题，国内外有越来越多学者对叶片面积的估算方法进行了研究，方法主要集中在对植物进行非接触式的测量与检测，其中尤其以使用图像处理和模式识别技术的方法为代表性研究。深度学习理论的出现为这一研究领域提供了新的解决途径与思路，国内外的研究方法如下文所述。

### 1.2.1 叶片图像分割方法

关于图像分割的文献非常广泛，它是一门前沿学科，从传统的基于图像的方法到基于深度学习的模型都有涉及。传统方法主要利用图像处理、拓扑学和数学理论对图像进行分割。主要算法有基于阈值的分割（Zhu S et al. 2007; Kittler J et al. 1985）、基于区域的分割（Stephen Gould et al. 2009）、基于边缘的分割（Iannizzotto G 2000）和特定工具组合分割（Gao J et al. 2020; Bhanu B 1995; Sagiv C 2006）。

与上述算法相比，深度学习的出现大大提高了图像分割的准确性。有许多方法依赖于深度学习框架来分割图像。对于基于特征编码器的方法，VGGNet（Simonyan K 2014）和 ResNet（He K 2016）在特征提取方面取得了很大的成功。区域建议是计算机视觉领域中非常常用的一种算法，在此基础上，研究人员提出了 R-CNN（Girshick R 2014）、Fast R-CNN（Ross Girshick 2015）、Faster R-CNN（Ren S 2016）和 Mask R-CNN（He K 2017）。同时，FCN（Evan Shelhamer 2017）解决了语义分割中的许多问题，并在此基础上提出了 U-Net（Ronneberger O 2015）和 SegNet（Badrinarayanan V 2017）。

Yang et al.（2017）构建了无损番茄叶面积在线检测系统，解决了传统叶面积检测方法繁琐、过分依赖人力、价格高、准确度低、易损坏叶片等缺点。Wei Ke（2020）基于 CNN，提出了多重锚学习（MAL），取得了新的最先进的检测性能。Mingxing Tan（2020）开发了一个新的目标检测器系列，称为 EfficientDet，实现了比现有技术更好的效果。Qi Fan（2020）不同于已有的方法，即学习从检测到分割的传递函数，或者学习聚类形状先

验来分割新的类型，他们提出学习潜在的阶级不可知论的共性，结合形状和外观的共性，他们的模型在 COCO 数据集的实例分割的部分监督设置和少数镜头设置上显著优于先进方法。

Maninis et al. (2018) 利用 CNN 和图像中的极值点，提出了一种用于图像分割的交互式标注框架。Rupinder et al. (2014) 提出了一种改进的 FCM 分水岭算法和 DBMF 算法用于图像分割。Thanammal et al. (2014) 利用直方图阈值技术进行自然图像分割。Richard et al. (2013) 提出分水岭变换用于模糊彩色图像分割。Waseem (2013) 做了一个关于图像分割的调查。在此基础上，考虑到 CNN 在图像处理上的优异性能以及本文的实验结果，本文最终选择 CNN 作为分割模型。Paprocki (2012) 对棉花植株使用 3DSOM 建立其三维模型，并对其使用新的混合网格分割算法进行分割，得到叶片的叶宽长度等参数。SANTOS T (2014) 使用运动恢复结构算法对株高大于 50cm 大豆和向日葵进行三维重建，并对整株植株进行了分割。

Zhi Tian (2020) 提出了一个简单而有效的实例分割框架，称为 CondInst (条件卷积实例分割)，他们采用基于实例的动态实例感知网络，而不是使用实例级 ROI 作为固定权值的网络输入。在 COCO 数据集上，性能优于一些最近的方法，实验结果如表 1-1 所示。

表 1-1 Zhi Tian (2020) CondInst 与其他算法对比

method	backbone	aug.	sched.	AP	AP <sub>50</sub>	AP <sub>75</sub>	AP <sub>S</sub>	AP <sub>M</sub>	AP <sub>L</sub>
Mask R-CNN	R-50-FPN		1×	34.6	<b>56.5</b>	36.6	15.4	36.6	<b>49.7</b>
<b>CondInst</b>	R-50-FPN		1×	<b>35.4</b>	56.4	<b>37.6</b>	<b>18.4</b>	<b>37.9</b>	46.9
Mask R-CNN*	R-50-FPN	✓	1×	35.5	57.0	37.8	19.5	37.6	46.0
Mask R-CNN*	R-50-FPN	✓	3×	37.5	59.3	40.2	21.1	39.6	48.3
TensorMask	R-50-FPN	✓	6×	35.4	57.2	37.3	16.3	36.8	49.3
<b>CondInst</b>	R-50-FPN	✓	1×	35.9	56.9	38.3	19.1	38.6	46.8
<b>CondInst</b>	R-50-FPN	✓	3×	37.8	59.1	40.5	21.0	40.3	48.7
<b>CondInst w/sem.</b>	R-50-FPN	✓	3×	38.8	60.4	41.5	21.1	41.1	51.0
Mask R-CNN	R-101-FPN	✓	6×	38.3	61.2	40.8	18.2	40.6	<b>54.1</b>
Mask R-CNN*	R-101-FPN	✓	3×	38.8	60.9	41.9	21.8	41.1	50.5
YOLACT-700	R-101-FPN	✓	4.5×	31.2	50.6	32.8	12.1	33.3	47.1
TensorMask	R-101-FPN	✓	6×	37.1	59.3	39.4	17.4	39.1	51.6
<b>CondInst</b>	R-101-FPN	✓	3×	39.1	60.9	42.0	21.5	41.7	50.9
<b>CondInst w/sem.</b>	R-101-FPN	✓	3×	<b>40.1</b>	<b>62.1</b>	<b>43.1</b>	<b>21.8</b>	<b>42.7</b>	52.6

### 1.2.2 叶片面积估算方法

为了估算叶面积，人们开发了各种分析设备和软件。Pandey et al. (2011) 提出了一种具有成本效益的替代方法，称为毫米方格纸法，用于计算单个叶片的面积。Zhang et al. (2010) 研究了叶片尺寸和高程的系数，并提出了一种基于叶片宽度的单变量模型进行叶面积估算的方法。这个单变量模型是通过测量到的叶片宽度 ( $W$ ) 和实际叶面积 ( $A$ ) 进行统计分析得到的。但是由于叶片的形态、纹理等因素可能会影响叶片宽度与叶面积之间的关系，因此在不同的植物物种和环境下，单变量模型的相关性和准确性可能会有所不同。

Yau et al. (2021) 利用 RGB-D 传感器和无监督聚类方法对叶面积进行分割和估计。Suárez 等人利用叶片的三个小叶来拟合回归模型，提高了叶面积估计的精度。Constantino (2018) 开发出水稻株高和分蘖数自动测量系统。AngTian Wang (2020) 用在训练期间通过边界框注释分割上下文的方法，扩展了基于部件的投票方案，对物体的边界框角进行投票，使模型能够可靠地估计部分遮挡物体的边界框。Paulo E. Linares Otoy and Sixto R. Prado Gardini (2020) 提出了一种通过从深度图像中获取三角形网格来估计波浪状单叶片表面积的新方法。其利用立体摄像机拍摄一组天竺葵的照片，构建三角形网格作为叶片面积表示，最后将得到的所有三角形面积之和作为估计的叶片面积。而陈柱(2022)则为了提高植物叶片面积测量的准确度，提出了一种植物叶片三维重建补偿方法。在多角度拍摄生成图像三维点云的基础上，对点云进行去噪、分割、填补、三角网格化处理。最后，对叶片面积进行估测。然而，这些方法成本高、对叶片有损伤，并且精度和效率有限。因此本文进行了一系列的实验来计算不同分辨率和不同场景下的叶片面积。

综上所述，本文使用基于极值点的分割算法分割大豆叶片图像，然后基于像素点估算叶片面积。

## 1.3 研究内容

如图 1-1 所示是本文的研究内容，本文主要包含三个内容，其中基于极值点的叶片分割是研究工作的第一步，基于像素点的面积估算在分割结果的基础上进行，最后封装以上方法实现了交互式的叶片分割与面积计算软件。

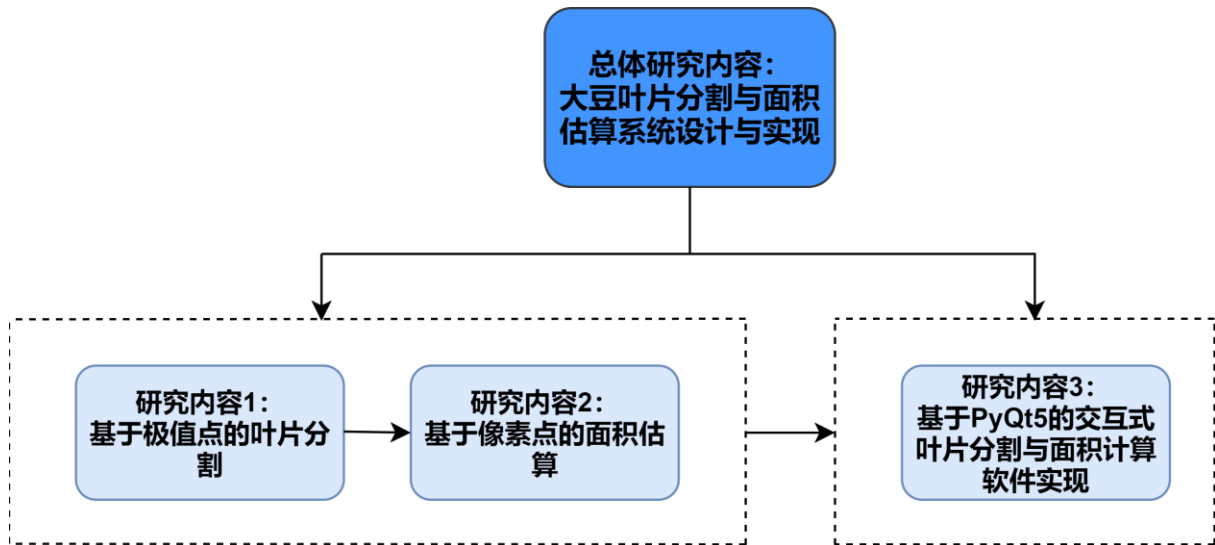


图 1-1 本文研究内容系统框架图。深色框图表示总体研究内容，浅色框图表示三个子研究内容。

### （1）基于极值点的叶片分割方法

本文旨在通过训练分割模型，预测分割复杂场景即自然状态下的叶片，进而无损地测量叶片面积。大豆叶片无需脱离本身所处的环境，因此叶片的颜色会与背景的颜色存在一定的相似性，另外多个叶片之间相互遮挡也会增加这一复杂性。如何在这一复杂的场景下实现大豆叶片的准确检测，是本文的重要研究内容。基于极值点的分割方法中，将目标对象的四个极值点作为参考信息，对较小物体和互相遮挡物体分割效果较好。基于极值点分割算法，对大豆叶片进行准确无损地分割，是本文的目标。

### （2）基于像素点的叶片面积估算方法

通过非自然场景的分割，得到了一组像素值和一组面积的对应，探索不同采集高度下，面积与像素的对应关系，设计出一种大豆叶片面积估算模型是本文的一个研究内容。

### （3）基于 PyQt5 的交互式叶片分割与面积计算软件实现

本文构建大豆叶片数据集，对叶片进行分割，通过研究像素点与面积之间的对应关系得出一种面积估算方法。在进行以上工作后，将基于 PyQt5 开发交互式的叶片分割与面积估算软件。用户导入待分割图片，软件将显示图片的尺寸数据；分割过程中，软件将根据用户提供的目标对象的极值点，调用分割模型进行分割，并显示在屏幕上；软件还提供分割结果的保存功能；对图片进行分割后，软件将根据分割结果计算叶片面积。





## 第2章 相关技术与方法

### 2.1 基于深度学习的图像分割算法

#### 2.1.1 U-Net 分割算法

U-Net 是一种新型的深度卷积神经网络模型，它被广泛用于医学影像分割任务。生物医学图像分割任务的数据量庞大，且具有多样性和复杂性，因此提高分割准确率和效率是一项重要的挑战。U-Net 网络结构如图 2-1 所示，模型结构由对称的编码器和解码器组成，并通过跳跃连接将不同层级的特征图合并起来进行细节处理。编码器部分由多个卷积层和池化层组成，用于将输入图像压缩为低分辨率的特征图，并提取不同层级和尺度的特征信息。随后，解码器通过反卷积操作将特征图恢复为原始大小，并结合跳跃连接将特定层级的特征图与解码器中对应的特征图进行直接连接，从而实现特征信息的精细处理。

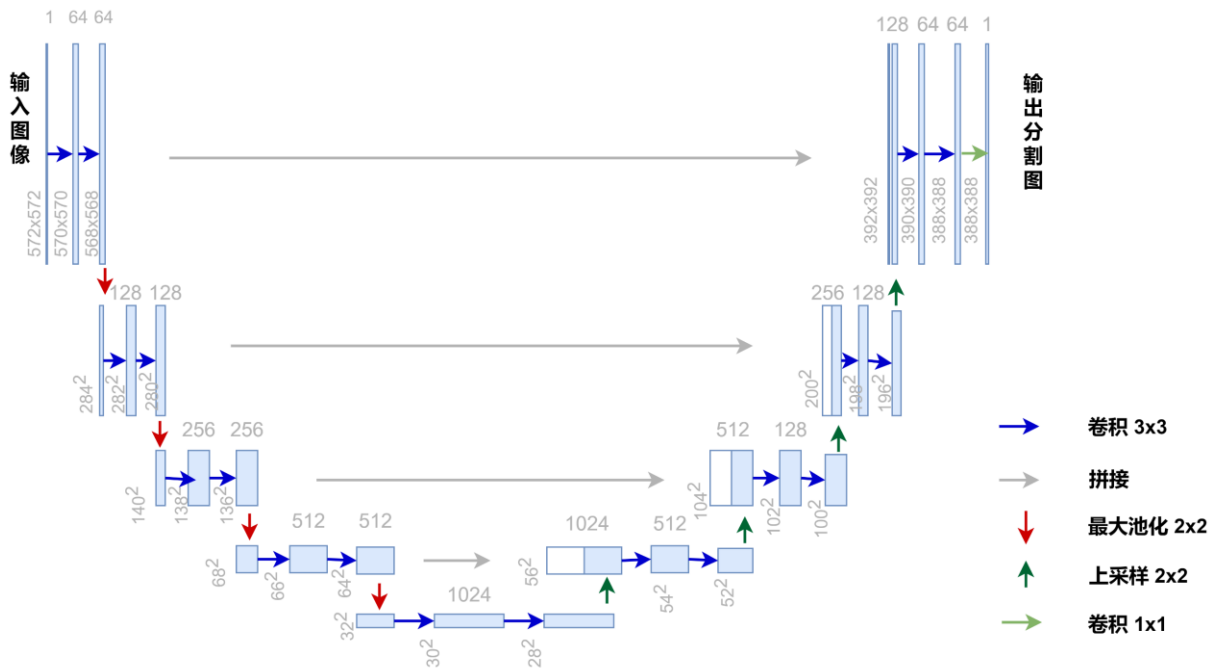


图 2-1 U-Net 网络结构

U-Net 的网络输出是像素级的 softmax 分类，像素级的 softmax 计算公式如下：

$$p_k(X) = \exp(a_k(X)) / (\sum_{k'=1}^K \exp(a_{k'}(X))) \quad (2-1)$$

其中， $X$  为像素位置， $k$  为类别数， $a_k(X)$  为输出像素中  $X$  所对应的第  $k$  个通道的值。使用带边界权值的损失函数：

$$E = \sum_{X \in \Omega} \omega(X) \log(p_{l(X)}(X)) \quad (2-2)$$

其中， $\omega(X)$ 为像素点的权值，为了给图像中贴近边界的像素点更高的权值。

U-Net 跳跃连接的设计非常关键，可以让解码器利用编码器中的高阶特征来实现更精细的分割效果，在生物学图像分割任务中具有更高的分割准确率和更快的收敛速度的优势。实验结果表明，在医学影像分割任务中，U-Net 具有更高的分割准确率和更快的收敛速度，能够为自动化的医学影像分析提供更加有效的解决方案。除了医学影像分割任务，U-Net 在其他领域也有广泛的应用，如视网膜图像分割、道路图像分割和海洋大气科学中的气象雷达显著特征提取。这些领域的应用均证明了 U-Net 模型在图像分割任务中的优越性。

### 2.1.2 R2-UNet 分割算法

与其它图像处理任务相比，语义分割需要对每个像素进行标注，即将像素分类为属于某一个物体或者背景。在这个问题上，传统的 U-Net 模型已经被广泛地使用，但是仍然存在着一些问题，其中之一就是在缺少标注数据时精度不高。为了解决这个问题，R2-UNet 被提出来了。

R2-UNet 是 U-Net 的一个改进版本，它利用了残差网络和注意力机制来增强 U-Net 的性能。网络模型如图 2-2 所示。

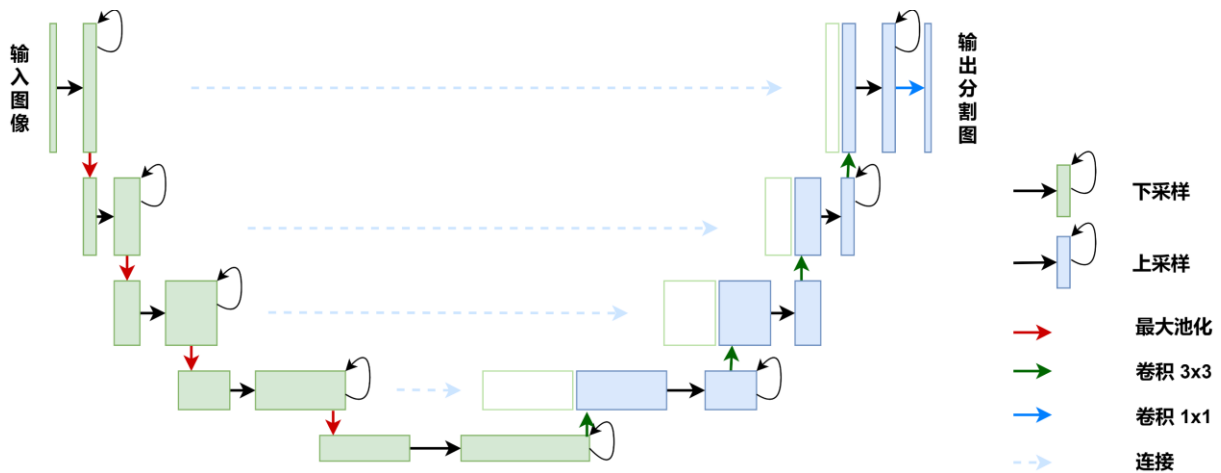


图 2-2 R2-UNet 网络模型

在 R2-UNet 中，首先构建一个多尺度特征图的编码器，然后对这些特征图进行解码。在编码器和解码器之间添加残差连接来帮助网络更好地学习具有不同分辨率的图像特征，以此来提高网络的性能。此外，在各个解码层中添加了注意力模块，使得网络可以对重要的特征进行更好的关注，并且抑制无关的特征。在 R2-UNet 中，RCNN 单元的最

终输出通过残差单元。假设 RRCNN 块的输出为  $x_{l+1}$ ，则可以计算如下：

$$x_{l+1} = x_l + F(x_l, \omega_l) \quad (2-3)$$

其中， $x_l$  表示 RRCNN 块的输入样本。 $x_{l+1}$  样本是 R2U-Net 模型的编码和解码卷积单元中紧接着的子采样层或上采样层的输入。该结构的主要优点是在保持 U-Net 显著优异的性能的同时，通过引入残差卷积来生成了更强的特征表达，从而有助于提高模型的精度。

另外该模型通过添加一个类似于 GAN 训练器的生成器和判别器来生成假数据，并将其与真实数据一起参与训练。这种方法不仅可以帮助模型克服标注数据不足的问题，而且能够提高模型在输入图像中的从容性。研究人员使用了 Cityscapes、CamVid 和 PASCAL VOC 2012 数据集对改进后的 R2-UNet 进行了测试，并与一些经典的语义分割模型进行了比较。实验结果表明，相较于普通的 U-Net 模型，R2-UNet 具有更高的精度和鲁棒性。同时，R2-UNet 在不同数据集上均达到了最佳语义分割性能。

R2-UNet 分割技术在语义分割任务上具有很好的表现，优点在于它能够处理缺少标注数据或者不充分的标记数据，具有很高的精度和鲁棒性。值得注意的是，该模型还具有可扩展性，可以被应用到许多不同的语义分割领域中，并且可以通过进一步的改进来增强其性能。

### 2.1.3 Attention-UNet 分割算法

Attention-UNet 是一种新兴的语义分割模型，它结合了 U-Net 和注意力机制，能够很好地解决传统 U-Net 在空间上缺少灵活性的问题。在 U-Net 中，特征提取与上采样的过程是通过下采样和上采样单元紧密相连来实现的，即硬编码的跨层连接。这种设计虽然简单有效，但是存在一个潜在问题，即当每个下采样分支都通过同样的方式计算时，整个网络的学习能力有所限制。因此，在 Attention-UNet 中引入了一种自适应的注意力机制（attention gate）来调整下采样和上采样之间的连接，如图 2-3 所示。

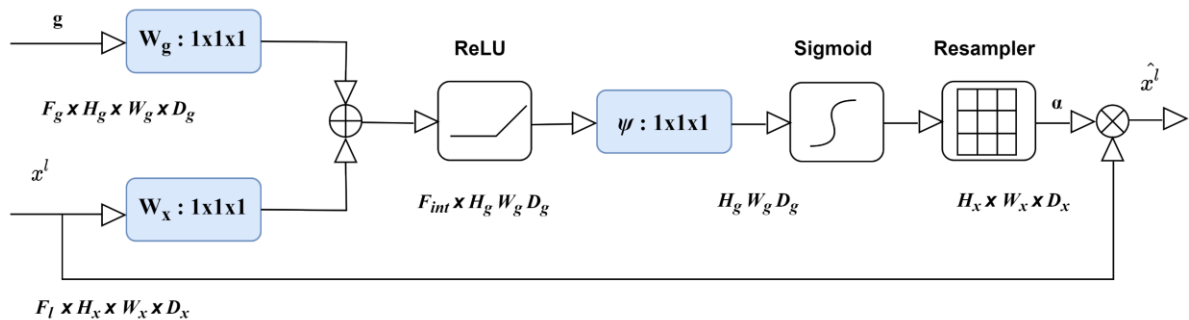


图 2-3 attention gate 原理

具体来说, Attention-UNet 的网络结构由两部分组成: 编码器和解码器。编码器由多个下采样模块组成, 每个模块包含两个卷积层和一个下采样操作。在每个下采样模块的输出的地方, 对输入进行空间和通道注意力机制的融合, 以提高模型对目标区域的感知能力。注意门 (attention gate) 的输出是输入特征图和注意系数逐一相乘, 使用以下公式来计算:

$$\hat{x}_{i,c}^l = x_{i,c}^l * \alpha_i^l \quad (2-4)$$

并且使用以下公式来计算注意系数:

$$q_{att}^l = \psi^T \left( \sigma_1 (W_x^T x_i^l + W_g^T g_i + b_g) \right) + b_\psi \quad (2-5)$$

$$\alpha_i^l = \sigma_2 (q_{att}^l(x_i^l, g_i; \theta_{att})) \quad (2-6)$$

其中, 其中  $\sigma_2(x_i, c) = \frac{1}{1 + \exp(-x_{i,c})}$  对应 sigmoid 激活函数。 $\theta_{att}$  表示 AG 的一系列参数:

线性变换  $W_x$ 、 $W_g$  和偏置项  $b_g$ 、 $b_\psi$ 。

在解码器中, 每个上采样模块都包含一个卷积层、一个上采样操作和一个跨层连接, 以帮助网络更好地学习不同分辨率下的特征。值得注意的是, Attention-UNet 还引入了一种全局上下文信息的注意力机制来进一步提高模型的精度。在此机制中, 使用全局平均池化操作来提取全局上下文特征, 然后通过注意力机制将全局上下文特征与卷积特征进行融合。这种机制的实现可以增加模型的灵活性, 并且有助于提高模型对图像特征的捕捉能力。

研究人员使用 ISBI Challenge 2012 和 2015 数据集对 Attention-UNet 进行了测试, 并将其与其他常用语义分割模型进行了比较。实验结果表明, Attention-UNet 在各项性能指标上均优于其他模型, 并且非常有效地解决了 U-Net 在空间上缺少灵活性的问题。此外, Attention-UNet 在处理病理图像方面表现尤为优异, 这也证明了它在医学图像领域中的应用潜力。

综上所述, Attention-UNet 是一种非常有效的语义分割模型, 它的网络结构和注意力机制使得模型能够更好地学习特定区域的信息。该模型不仅在图像分割任务中表现优异, 在医学图像处理方面也具有很好的潜力。

## 2.1.4 FCN-32s 分割算法

FCN 是一种非常经典的全卷积神经网络，用于处理图像语义分割任务。该网络最初由 Jonathan Long 等人于 2014 年提出，它的主要思想是把传统的 CNN 模型转化为全卷积网络，从而使 CNN 也能够处理变长输入数据。具体来说，FCN 通过将最后一个全连接层替换为卷积层，将池化层的输出尺寸下采样至与输入相同的尺寸，并且使用反卷积层将特征图上采样到原始分辨率，从而实现了端到端的语义分割。其中，FCN-32s 是将预测结果上采样了 32 倍，还原回原图大小，网络结构如图 2-4 所示。

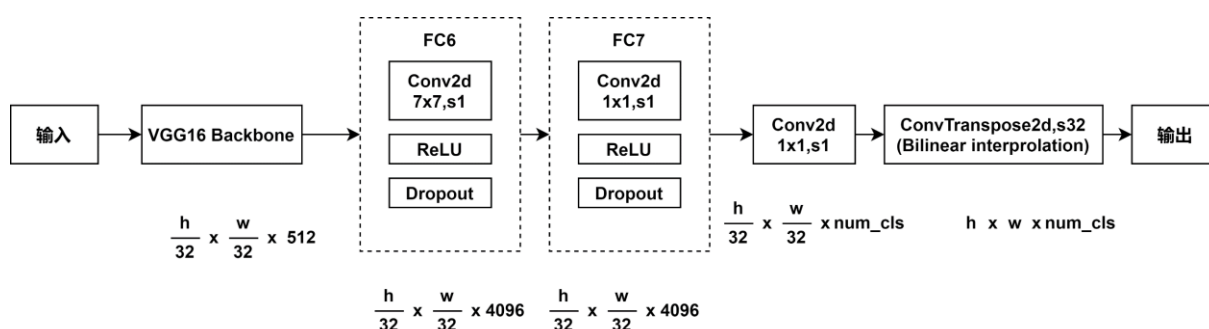


图 2-4 FCN-32s 网络结构

FCN 的结构由两个部分构成：编码器和解码器。其中编码器的作用是将图像的高层次特征提取出来，用来捕捉整体的语义信息；解码器则利用高层次特征生成对每个像素点的预测结果。编码器采用了预训练的 VGGNet 卷积结构，该网络拥有多层卷积层来学习图像的多种特征，因此可以很好地为解码器提供多种特征提取。而解码器则使用反卷积层将编码器输出的低分辨率特征图上采样，逐渐还原到原始图像尺寸，并最终得到每个像素点的类别预测结果。

在 FCN 结构中，并没有引入显式的跳跃连接或者注意力机制，因此它对图像局部信息的感知能力较弱，对于具有大量细节和复杂结构的图像可能会受到限制。但是，FCN 作为全卷积网络的代表，依然在语义分割任务中表现优异，特别是对于简单而规则的图像分割有着很好的效果。除此之外，FCN-32s 还可以扩展到更深的网络结构上，例如 FCN-16s 和 FCN-8s，这些模型引入了跳跃连接或者注意力机制，以提高模型在解决复杂图像分割任务中的性能。

FCN 将传统的 CNN 网络转化为全卷积网络，通过编码器和解码器结构对图像进行特征提取和像素级别的预测，具有推理速度快、效果优秀等优点，被广泛应用于实际场景中。虽然 FCN 对于复杂的图像处理可能存在局部信息感知不足的问题，但它的网络结构可以很好地扩展到更深的网络上，以进一步提高模型性能，为图像分割领域带来更多的可能性。

## 2.2 基于极值点分割算法的相关方法

### 2.2.1 CNN 网络

卷积神经网络(CNN)是一种应用广泛的深度学习算法，特别适用于图像识别和计算机视觉领域。CNN 通过多个卷积层和池化层的加工处理，从原始图像数据中提取出高级抽象的特征表示，再利用全连接层进行分类或回归等任务。

CNN 的核心部分是卷积层和池化层，其中卷积操作负责从输入数据中提取出特征信息，池化操作则用于缩小图像大小并减少特征的相互依赖性。卷积操作通常包括卷积核和偏置项，卷积核的大小可以根据具体任务而定。当卷积核在图像上滑动时，它会进行加权求和，并添加一个偏置项，以产生新的输出特征图。由于卷积操作能够充分利用输入数据的空间关系，它能够强化模型对图像特征的感知能力，从而提高分类效果。除了卷积层和池化层，CNN 还包括全连接层和激活函数。全连接层负责将特征图展开成一维向量，并计算输出结果，通常用于分类任务。激活函数则主要负责引入非线性因素，以提高模型的表达能力。

CNN 的训练过程是通过反向传播算法来实现的，它可以根据误差或损失反向更新每个神经元的权重和偏置。CNN 的目标是让模型在训练集上达到最佳效果，避免出现过拟合和欠拟合现象。为了减轻过拟合的影响，通常采用正则化方法，如 L1 正则化和 L2 正则化等。此外，还可以通过 Dropout 等技术随机关闭一部分神经元，以减少神经网络中的冗余信息和复杂度。

CNN 的应用涉及到许多领域，如图像识别、手写数字识别、物体检测、人脸识别、语音识别等。在图像识别领域，CNN 已经取得了重大成就，甚至能够超越人类的视觉识别水平。例如，在 ImageNet 大规模视觉识别竞赛(ILSVRC)中，CNN 已取得了优异的成绩，并成为开拓深度学习新时代的重要突破。

在其他应用方面，CNN 也有着广泛的应用前景。例如，在人脸识别领域，CNN 已经成为最先进的技术之一，能够有效地实现人脸检测、人脸特征提取和人脸匹配等任务。另外，在自然语言处理领域，CNN 也被应用于文本分类任务中，能够对输入的文本信息进行特征提取，从而实现文本分类和情感分析等任务。随着计算机计算能力的不断提升和深度学习发展的加速，CNN 将在图像识别、语音识别、自然语言处理等领域中发挥越来越重要的作用。

### 2.2.2 ResNet 网络

ResNet 是深度学习领域中具有开创性意义的卷积神经网络，其特点是在神经网络中

引入残差块，大幅提升了深层神经网络训练的稳定性和准确性，是深度学习发展历程中的一个重要里程碑。ResNet 的核心是残差块（Residual Block），它是由两个大小相同的卷积层和一个跳跃连接构成的，如图 2-5 所示。

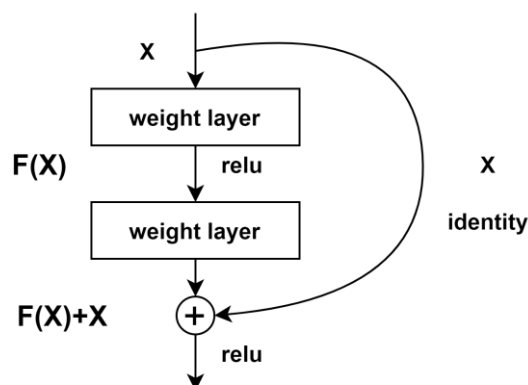


图 2-5 ResNet 残差块

其中，每个卷积层后面都跟着 Batch Normalization（批量归一化）和 ReLU（线性修正单元）激活函数。跳跃连接则是直接将输入数据与输出相加，以此形成残差块的结果。通过这样的设计，ResNet 可以使得训练非常深的神经网络变得可行，同时也有效缓解了梯度消失的问题。另外，ResNet 还在第一个卷积层之前引入了  $1 \times 1$  卷积层，这样能够降低输入数据的通道数，使得后续卷积层的计算量更小。

ResNet 的训练使用了传统的随机梯度下降（SGD）算法，但是在优化上采用了批量归一化技术和残差学习。批量归一化通过对网络的每个输入数据做一次归一化，增加了网络层之间的稳定性，有效地缓解了梯度消失问题。而残差学习则是通过直接将输入数据作为输出与残差块的结果相加，使得网络可以更容易地训练。

由于 ResNet 能够高效地提取图像特征，研究人员将它运用到医学图像分析领域，能够提高医学图像的分割准确性，从而帮助医生更好地检测疾病。ResNet 是一个具有里程碑意义的深度学习模型，在深度学习领域的发展历程中占据了重要的地位。

## 2.3 基于二值化的面积估算方法

在过去，叶面积估算多依赖人工测量，费时费力，而且精度也难以保证。然而，随着计算机视觉技术的发展和普及，基于图像处理技术的叶片面积估算方法得到了广泛应用。其中，基于二值化的方法是最常用的之一。

基于二值化的叶片面积估算方法，首先需要将叶片图像进行预处理，将其转换为二值图像，去掉背景噪声。二值化的方法可以采用简单全局阈值法或自适应阈值法，根据具体情况选择不同的方法。此时，图片中像素的灰度值只有黑白两种状态，即“0”或

“1”。这样就得到了一个由像素点组成的二值图像，其中“1”的像素点代表叶片的一部分，可以计算这些像素点的数量，并与一个已知的标准像素值进行比例计算，从而得到叶片的面积。

对于全局阈值的方法，其选择要考虑叶片与背景颜色的对比度。如果叶片颜色与背景颜色不同，则采用全局阈值方法计算叶片面积可能会产生误差。此时，可以选择自适应阈值方法，该方法可以根据图像的局部变化来自适应地选择阈值。自适应阈值的方法，将图片划分成一个个小的分块，在每个小的分块内进行背景噪声的去除和阈值的计算。这样可以更好地处理叶片颜色较为均一的情况。

基于图像处理的叶片面积估算方法是一种非常实用的技术，可以在极短的时间内计算大量叶片的面积。在选择具体方法时，需要根据具体情况综合考虑各种因素，并在实验中进行验证，从而得到更加准确的结果。

## 2.4 PyQt5

PyQt5 是一种基于 Python 编程语言的 GUI 工具包。它能够帮助开发人员创建各种不同类型的窗口应用程序，包括桌面应用、数据库应用、网络应用等。PyQt5 将 Qt C++ 应用程序框架嵌入到 Python 中。Qt 是一种跨平台的应用程序开发框架，它可以通过一系列 API 和工具来创建高性能、低延迟、可移植、可扩展性强的应用程序。PyQt5 对 Qt 框架的封装为开发者提供了一种快速、便捷的方式来创建跨平台 GUI 应用程序。PyQt5 可以在 Linux、Windows、MacOS 等操作系统上运行，并且具有良好的可移植性和兼容性。

PyQt5 提供了许多不同的模块，用于支持不同类型的应用程序开发。其中最常用的模块包括：QtCore、QtGui 和 QtWidgets。QtCore 模块是 PyQt5 库的核心模块，提供了一组数据类型、文件操作、I/O、事件处理、线程等核心功能。QtGui 模块则提供了图形系统和窗口系统的基础支持，包括窗口、标签、按钮、菜单等基本控件。最后，QtWidgets 模块提供了一组用于创建丰富的 GUI 应用程序的高级控件。

在 PyQt5 中，开发者可以使用 Qt Designer 进行界面设计，也可以纯编程方式来创建自定义界面。Qt Designer 是 Qt IDE 的一部分，它提供了一个可视化的界面设计工具，允许开发者通过拖拽控件和编辑属性来设计界面。使用 Qt Designer，开发者可以很容易地创建自定义的用户界面，并将其导出为 XML 格式的文件。然后，可以使用 PyQt5 的 uic 模块将这些 XML 文件转换为 Python 代码，并将其集成到应用程序中。

总之，PyQt5 是一种优秀的 GUI 工具包，它能够帮助开发者快速、便捷地创建跨平台的 GUI 应用程序。PyQt5 可以通过 Qt Designer 进行界面设计，也可以使用纯编程方



式来创建自定义界面。使用 **PyQt5**，开发者可以利用丰富的控件库和模块集合，创建各种类型的窗口应用程序。不仅如此，**PyQt5** 还具有良好的兼容性和可移植性，因此在实际开发中，它是非常受欢迎的一个 **GUI** 工具包。



## 第3章 基于极值点的分割算法的实现

### 3.1 数据采集

为了进行实验，本文收集了西北农林科技大学实验田的大豆叶片图像。对于同一片叶子，采用不同的拍摄策略，包括不同的场景，分别为自然场景和毫米网格纸场景；然后采用了两种拍摄高度，分别为 10 厘米和 15 厘米，作为对照。使用 1200 万像素的手机收集图像，因此原始图像的像素分辨率为  $3000 \times 4000$ ，并以 JPG 格式保存。

在获得原始数据后，对网格纸上带有非自然场景的图像，即大豆叶片进行裁剪。收集大豆叶片时使用的网格纸规格为  $4\text{mm} \times 4\text{mm}$ ，裁剪后的图像边缘与网格线重合，通过计算网格数量得到整张图像的面积。图像采集细节如图 3-1 所示。

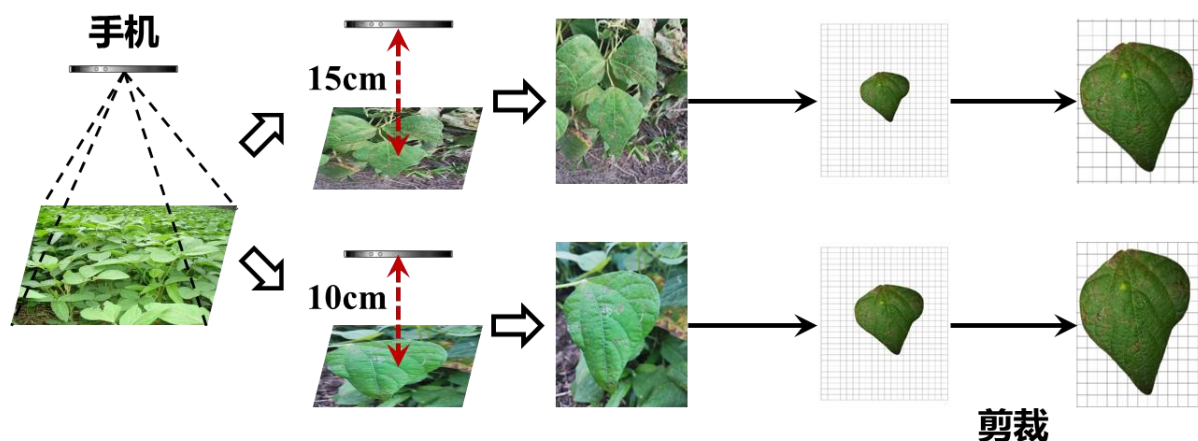


图 3-1 图像收集和预处理

### 3.2 数据预处理

图像二值化的方法，是将图像的灰度值转换为黑白二值图像，将相邻像素之间的差别提高到相对明显的状态，从而有效地将图像分成两个区域：黑色和白色。简单来说，就是将图像中的目标和背景分离出来。方法主要有两种：全局阈值法和自适应阈值法。

本文将非自然状态下的大豆叶片图像转化为二值图像，将图像分割问题转化为像素点的二分类问题。建立数据集时采用全局阈值法，全局阈值法是指将整个图像看作一个整体，然后对图像中每个像素的灰度值进行阈值处理。设定一个阈值，将灰度值高于该阈值的像素设为白色，低于该阈值的像素设为黑色。原始数据转化为二值图像后如图 3-2 所示。

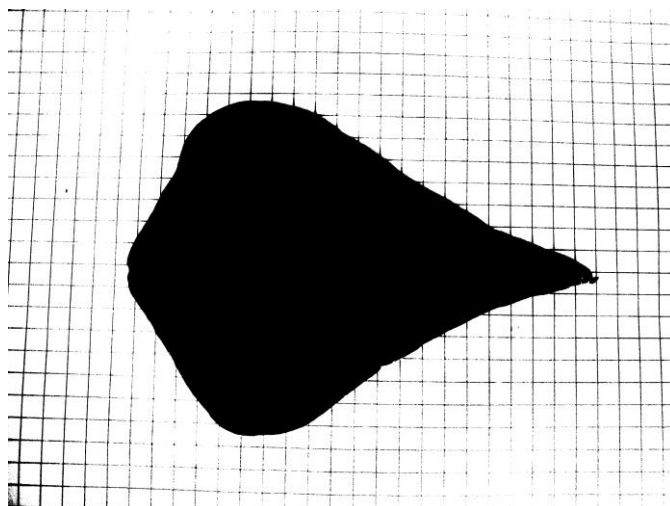


图 3-2 二值图像

在图像分割任务中，训练集中的 mask 集是非常重要的部分。mask 是指针对每张图片所标注出的像素级别的标签。通过将一个图像和其对应的 mask 相结合，可以训练出一个机器学习模型，用于识别和区分不同的物体或者背景。使用 mask 集进行训练可以帮助机器学习模型更好地理解图像中的不同物体边界和轮廓。这对于图像分割任务非常关键，在很多场合下，只有具有像素级别的标签才能够准确区分出物体和背景。所以本文制作了 mask 集作为标签，如图 3-3 所示。



图 3-3 mask 集标签

一方面，mask 集为机器学习模型提供了非常精细的标注信息。通过使用这些数据，机器学习模型可以更好地理解物体间的复杂关系，并且更好地学习到杂乱无序场景下的物体分类和分割。另一方面，mask 集还可以帮助机器学习模型学习到新的物体类别。对于一些新的物体类别，使用人工标注的 mask 集可以使得机器学习模型更快地适应这些

新的类别，并且学会正确分类和分割。在一些需要大量样本数据进行训练的任务中，使用 **mask** 集也可以使得机器学习模型更好地从有限的数据中学习到更丰富的特征。

裁剪后的图像大小不一，由于像素大小不一致的图像无法训练图像分割模型，所以本文将图像统一调整为 $512 \times 512$ 像素。本文使用 MATLAB 处理图像数据，MATLAB 提供不同的尺寸调整的函数，**imresize** 函数可以采用两种不同的缩放因子进行调整，但最重要的参数是插值方法，用于定义如何进行目标图像与原始图像的插值。插值方法的选取可能会影响到图像的质量与计算效率。MATLAB 提供了以下五种插值方法：**nearest**(最近邻)、**bilinear**(双线性)、**bicubic**(双立方)、**lanczos2**（高级滤波技术）和 **lanczos3**。本文使用了双线性插值，在目标像素周围的四个最近邻像素位置进行加权平均。这种方法可以平滑图像的边缘，保留图像的重要信息，适用于大多数应用场景。

### 3.3 算法模型及实现

算法的核心思想是利用极值点的信息，在深度学习模型的帮助下实现精准的目标分割。该算法框架主要由两个阶段组成，即生成候选框和迭代细化。

在第一阶段中，算法根据用户提供的极值点生成多个大小不同的矩形框，这些矩形框以极值点为中心并按大小排序。然后，算法在每个矩形框上运行一个深度学习模型，通过卷积神经网络来推断该矩形框是否包含目标物体。这里使用的深度学习模型采用残差网络（ResNet）和特征金字塔网络（FPN）的结构，可以有效地提取图像特征，并且在处理多尺度图像时具有良好的性能。

在第一阶段完成后，算法会在所有候选框中进行非极大值抑制，以排除重叠的较小框并保留最佳的候选框，可以大大减少后续迭代细化的计算量。在第二阶段，算法执行迭代细化过程，从而更精细地分割目标。在每次迭代中，算法会根据距离该候选框最近的极值点，对矩形框进行微调，并重新生成一个更为精细的矩形框。同时，算法也会更新与该矩形框相对应的分割掩码。这个过程重复多次，直到获得最后的精细分割结果。

该算法的成功应用，使得用户能够很容易地进行目标分割，并节省了大量手动处理数据的时间。同时，该算法在多个目标分割数据集上都取得了出色的分割效果，尤其对于复杂场景下的交互式分割任务有着明显的优势。因此，在学术界和工业界都受到了广泛的关注和应用。

对于带有网格纸的叶片图像，使用传统的深度学习模型进行分割非常方便。但当涉及到自然场景时，这些模型的表现并不好，没有得到预期的结果。解决这个问题最常见的方法是在目标区域周围画一个边界框。但是获得边界框的过程对我们的要求很高，因为它需要用户点击物体内外的点，然后沿着对角线拖动方框，并进行多次调整以获得准

确的边界框，因此这也是一项耗时的工作。Papadopoulos et al. (2017) 提出了一种基于极值点的方法，该方法利用了物体的最上、最下、最左、最右四个点。这四个点可以很容易地形成一个边界框。本文使用这四个极值点来捕获自然图像中期望分割的叶子，网络模型如图 3-4 所示。

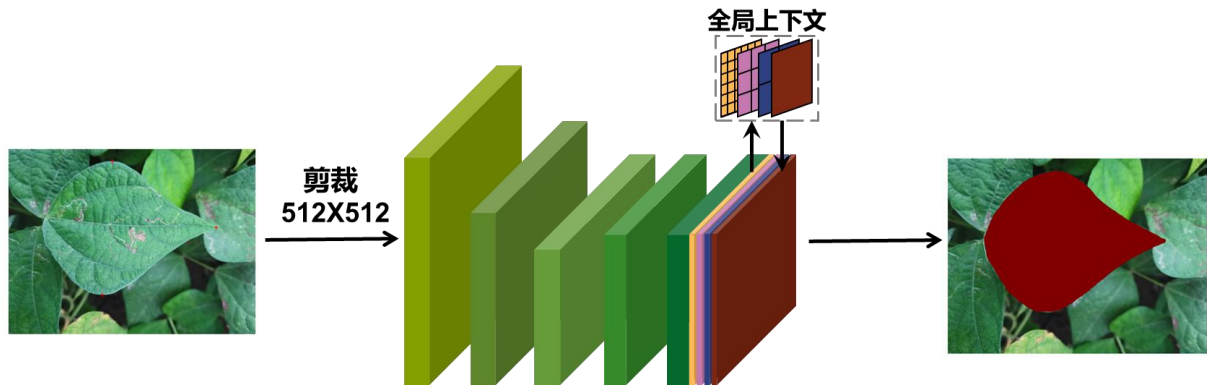


图 3-4 网络模型

对于自然状态下的图像，背景信息会干扰分割。为了聚焦在目标叶子上，本文使用一个由四个极值点组成的  $512 \times 512$  的边界框来裁剪图像，边界框内包含的区域可以被判别为包含当前目标的区域。围绕每个点创建热力图来感知图像中物体的位置信息，热力图反映了目标物体的边缘和内部的空间分布特征。而针对这个特征，卷积神经网络就能够很好地把握图中物体的位置信息。在每个极值点周围的热力图中，不同位置的像素对于它们在目标分割中的重要性是不同的。例如，一个处于目标边缘的极值点周围的像素比一个在目标内部的极值点周围的像素更重要。因此，通过针对每个极值点建立热力图，然后将热力图与输入图像的 RGB 通道连接，形成 CNN 的 4 通道输入，输入到模型中。CNN 就能够更加准确地感知图像中物体的位置信息，进而绘制目标对象的轮廓并进行填充，达到分割效果。

本文将 ResNet 作为主干网络，在最后两个阶段去掉了 FCN 层和最大池化层，以保持密集预测的分辨率。网络结构包含了四个阶段，每个阶段由多个基础块构成，其中空洞卷积应用于第三和第四阶段。

最后一步，利用金字塔场景解析模块 (Zhao H et al. 2017)，将全局上下文聚合到特征映射中。在该模块中，首先将输入图像分别缩小为四个不同尺度的图像，然后对每个尺度的图像进行卷积操作，得到响应图。接着，对于每个响应图，将其分为若干个区域，然后分别对每个区域中的像素点进行最大池化操作，得到局部特征描述。最后，将所有局部特征描述拼接起来，再通过一个  $1 \times 1$  的卷积层进行融合，得到一个全局上下文信息聚合到的特征映射。

网络的初始权值来自 ImageNet 和 PASCAL 的训练结果。网络的输出是一个概率图，它表示一个像素是否属于一个对象。为了有效地分割，引入了以对象为中心的裁剪策略，属于前景的样本比属于背景的样本多，因此有必要使用平衡损失。本文使用损失函数(3-1)，使熵损失最小化。

$$L = \sum_{j \in Y} \omega_{yj} C(y_j, \tilde{y}_j), j \in 1, 2, \dots, |Y| \quad (3-1)$$

其中， $\omega_{yj}$ 为权重，其依赖于像素 $j$ 的 $y_j$ ； $C(\cdot)$ 表示真实值与预测值之间的熵差。

### 3.5 对比算法的实现

本文使用其他几种分割算法进行对比试验，包括 U-Net、R2-UNet、Attention-UNet 和 FCN-32s 四种算法。算法具体的网络模型构建和参数设置如下文所述。

#### (1) U-Net

下采样编码器部分包括四个连续的模块，每个模块都是两个卷积层（DoubleConv）和一个最大池化层（MaxPool2d）。每个模块的卷积核大小不变，通道数以 32, 64, 128, 256 的倍数逐渐增加。最后的输出是经过四次下采样处理的特征图。上采样解码器使用反卷积运算（ConvTranspose2d），上采样大小为 2 倍，对应下采样操作的最大池化。每个上采样解码器同样包含两个卷积层，通道数逐渐减小。

在解码器部分，每个解码层将其上一个卷积层的输出与其上一个卷积层的输入切合起来，从而实现了跨层连接。并且，使用 concatenation 的方式将两个特征图简单连接起来。最后，整个网络的输出经过一个 sigmoid 函数激活将其映射到[0, 1]的范围内，得到二分类的概率，表示像素点属于目标对象的可能性。

#### (2) R2-UNet

编码器部分由 5 个 RRCNN 模块和每两个模块之间的一个下采样操作(MaxPool2d)构成。RRCNN 模块包括两个卷积层和一个循环卷积核块，该卷积块增强了特征之间的交互，提高了准确性和鲁棒性，其通道数以 64、128、256、512、1024 的间隔增加。

解码器部分由 5 个上采样模块和 5 个与之对应的 RRCNN 模块构成。每个上采样模块都是一个反卷积层(Upsample)和一个卷积层组成，其通道数需要和下采样对应层的通道数相等。解码器部分使用一个 up\_conv 模块，这个模块包括一个卷积层和一个反卷积层，与 U-Net 一样，它是用来上采样低分辨率的特征图像素，使其大小与下采样网络的同一层特征图相同。

### (3) Attention-UNet

网络结构包含了四个下采样层(Encoder),四个上采样层(Decoder)和四个 Attention 模块。其中,下采样层使用最大池化层,上采样层使用反卷积层(Up-Convolution)实现尺度的放大,而 Attention 模块则引入了注意力机制,帮助网络更好地自适应地关注感兴趣的特征。

网络在编码器中包含了 5 个卷积块,每个卷积块包含一个卷积层和一个激活函数,卷积层中包括 2 个卷积核大小为 $3 \times 3 \times 3$ 的卷积层,使得网络可以对输入图像的低层次和高层次语义信息都有有效地提取。在解码器中,每个卷积块都包含了一个反卷积层、一个注意力模块和一个卷积层。上采样层加上连通模块后,再通过反卷积将图像逐渐恢复到原始大小,并逐渐融合上采样前的特征图和其对应的下采样网络中提取的高层次特征。注意力模块则通过学习注意力系数,自适应地选取感兴趣的特征,并将其与上采样特征进行加权融合。最后输出 $1 \times 1$ 卷积层的 Sigmoid 概率分布,进行图像分割。

### (4) FCN-32s

FCN-32s 网络共包含四个反卷积层和一个卷积层,其中反卷积层用于将特征图按照前面的池化层的倍数放大。

第一个反卷积层(deconv)的输入是预训练网络的第四个池化层,输出通道数为 512,反卷积核大小为 $3 \times 3$ ,步长为 2, padding 为 1;第二个反卷积层的输入是第一层反卷积层的输出,输出通道数为 256,反卷积核大小为 $3 \times 3$ ,步长为 2, padding 为 1;第三个反卷积层和第四个反卷积层同理,输出通道数分别为 128 和 64。最后,通过一层 $1 \times 1$ 卷积层将通道数转换为分类数,再通过 Sigmoid 激活函数,得到每个像素点属于类别的概率分布。

在前向传播过程中,先将输入数据通过预训练模型前向传播得到特征图,然后通过多个反卷积层将特征图逐步上采样,最终将通道数转换为分类数,得到图像分割预测结果。

## 3.6 实验结果对比

本文首先使用不同拍摄高度非自然场景图像探索不同模型的分割结果。这些模型包括 Attention-UNet、FCN-32s、R2-UNet 和 U-Net。如图 3-5 和图 3-6 所示,可以清楚地看到,对于以网格纸为背景的叶片图像,所有的分割模型都表现良好,但本文的模型可以突出更多的叶片细节,所以本文利用该模型的结果来计算叶片的面积。



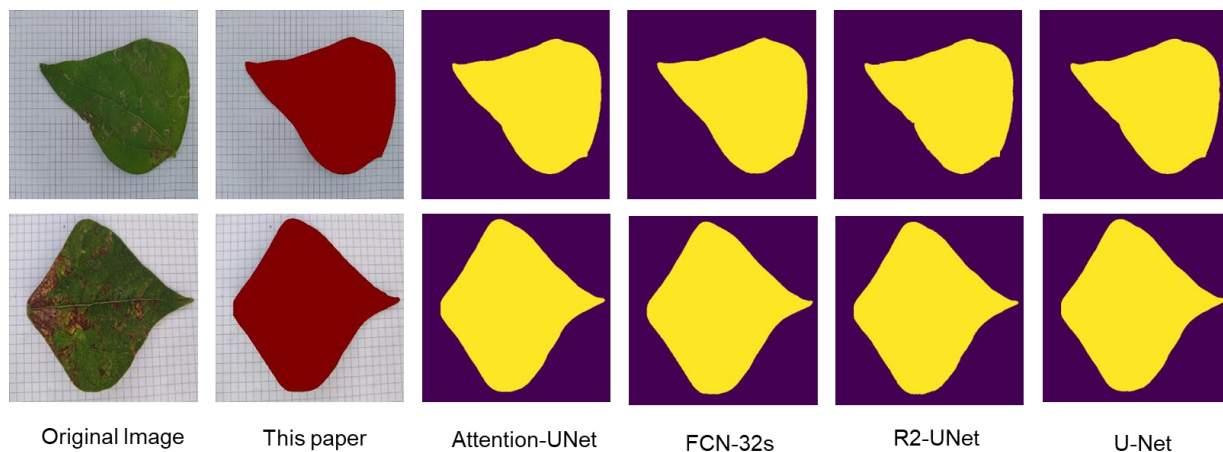


图 3-5 非自然场景图像分割结果对比(10cm 高度)

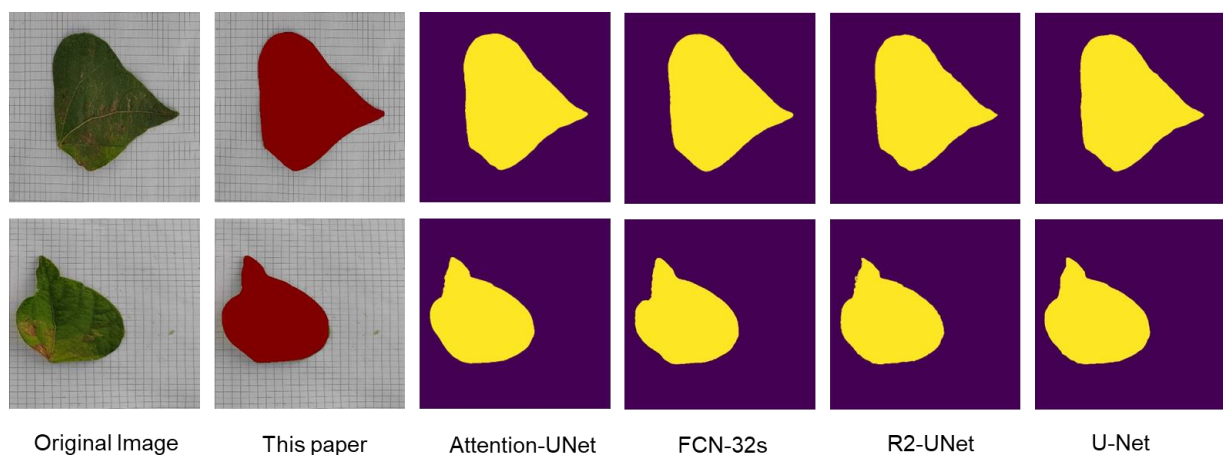


图 3-6 非自然场景图像分割结果对比(15cm 高度)

此外，本文还对自然场景图像进行了实验，并对上述模型进行了比较。如图 3-7 和图 3-8 所示，可以得出结论，本文的方法优于其他 4 种算法。并且由于 15cm 高度的图像包含更多的噪声信息，10cm 高度的分割效果优于 15cm 高度。

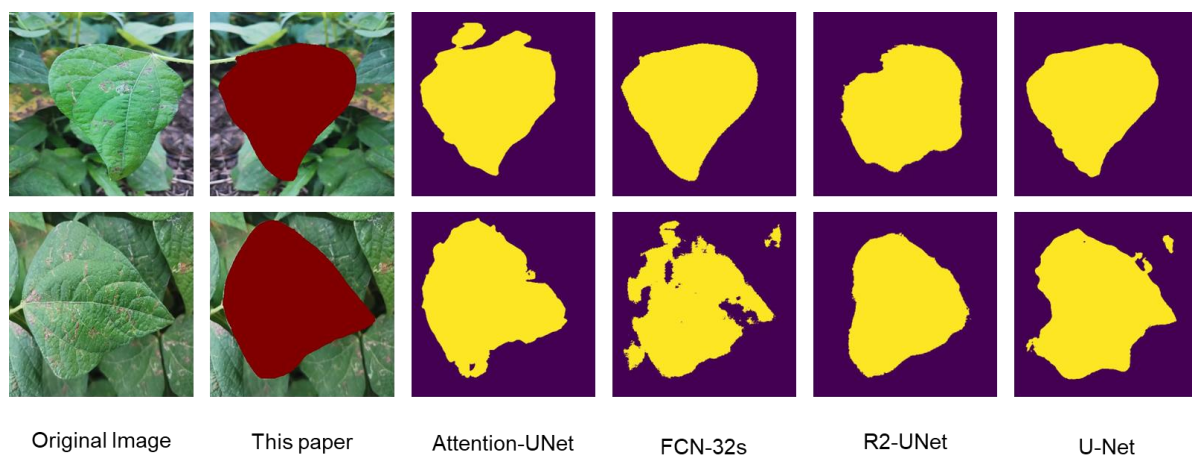


图 3-7 自然场景图像分割结果对比(10cm 高度)

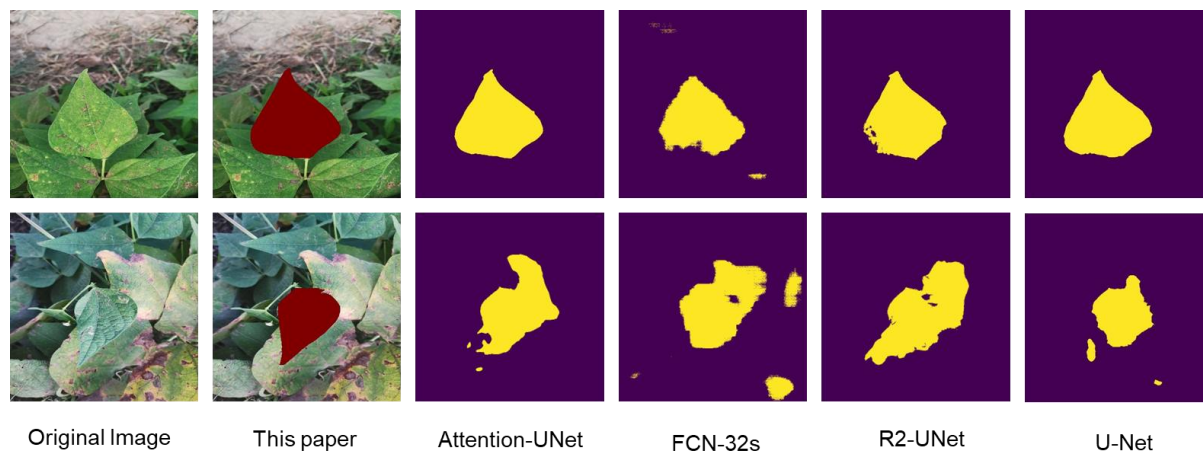


图 3-8 自然场景图像分割结果对比(15cm 高度)

## 第4章 基于像素点的面积估算方法实现

### 4.1 面积估算

基于像素点的面积估算方法是一种数字图像处理技术，可以应用于各种需要测量二维物体面积的场所。相比传统的实物测量方法，这种方法具有操作简单、速度快、精度高等优点，因此在现代工业、农业、医学、生物等领域得到了广泛应用。

叶片面积估算方法中，基于二值化的方法是一种常用的方法。其基本思路为：将叶片图像转换为二值图像，去掉背景噪声后，计算所得像素个数和已知像素之间的比例，从而得到叶片的面积。陈家茜和刘昌华(2021)提出了一种基于分水岭算法对甘蓝型油菜叶片进行分割的方法，最后通过统计目标区域的像素点个数计算叶片面积。

本文的面积估算方法与其类似，收集的大豆叶片图像经过 U-Net 等网络分割后，得到如图 4-1 所示的预测图像。由于数字图像是由像素阵列组成的，而每个像素只能表示一个固定的灰度值或颜色值。因此，当图像中出现亮度或颜色发生突然变化的边界时，邻近的像素点灰度值或颜色值也会发生相应的变化，这就导致了边缘处像素值的阶梯状变化。考虑到分割图像中叶子和背景边界处的像素值呈阶梯状变化，本文采用阈值法计算叶片所占的像素数量。

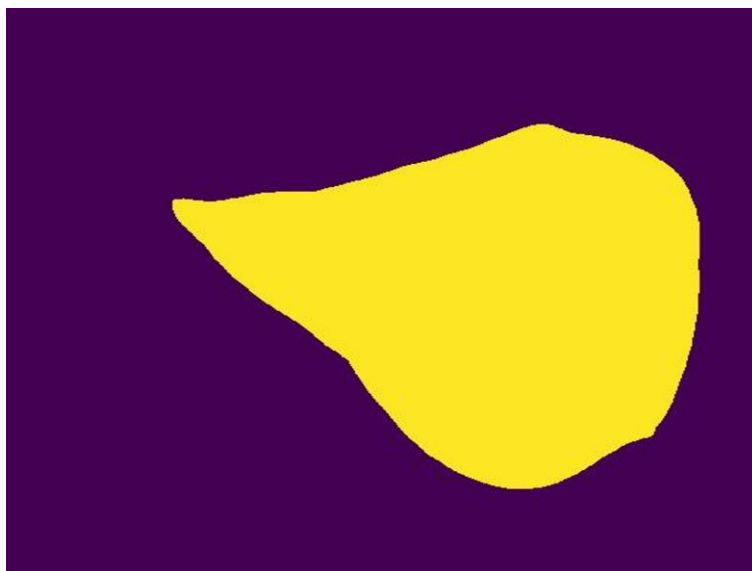


图 4-1 预测图像

如图 3-1 所示，对于自然场景中的每一片叶子，也将其记录在网格纸下，使用基于像素的方法辅助计算其面积。而对于每一张非自然场景图片，本文先对其进行裁剪，方便计算整张图像的面积，然后根据目标叶子的像素占整张纸的比例得到叶子的面积，可

以概括如下：

$$S_{leaf} = \frac{P_{leaf}}{P_{paper}} S_{paper} \quad (4-1)$$

其中， $S_{leaf}$ 和 $S_{paper}$ 表示目标叶子和纸张的面积， $P_{leaf}$ 和 $P_{paper}$ 表示目标叶子和纸张的像素点数量。

## 4.2 估算结果对比

抽取 10 张非自然场景下，拍摄距离为 10cm 的大豆叶片图像，对分割结果进行面积对比，如图 4-2 所示。可以看到由于非自然场景下大豆叶片位于方格纸上，叶片和背景有着明显的区别，分别使用多种分割算法对叶片进行分割后，面积估算结果基本一致。

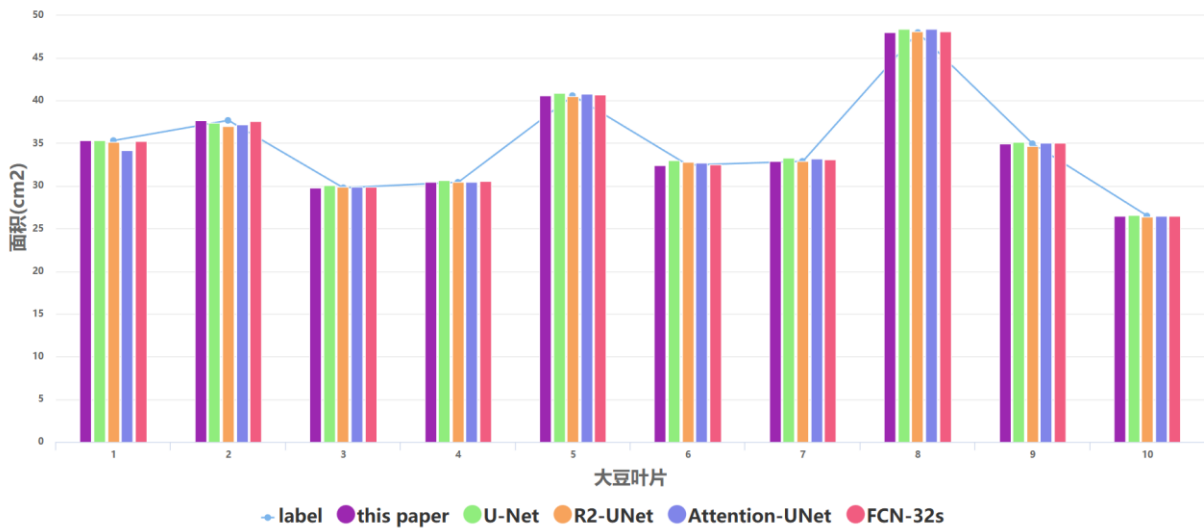


图 4-2 非自然(10cm)面积对比

抽取 10 张自然场景下，拍摄距离为 10cm 的大豆叶片图像，对分割结果进行面积对比，如图 4-3 所示。可以看到由于自然场景下叶片间相互遮挡，采集的图像中叶片边缘特征不明显，分别使用多种分割算法对叶片进行分割后，面积估算结果存在极大的差异。其中基于极值点的分割算法得到的叶片，经过面积估算后得到的结果与标签最为接近。可见基于极值点的分割算法取得了最佳的分割效果。

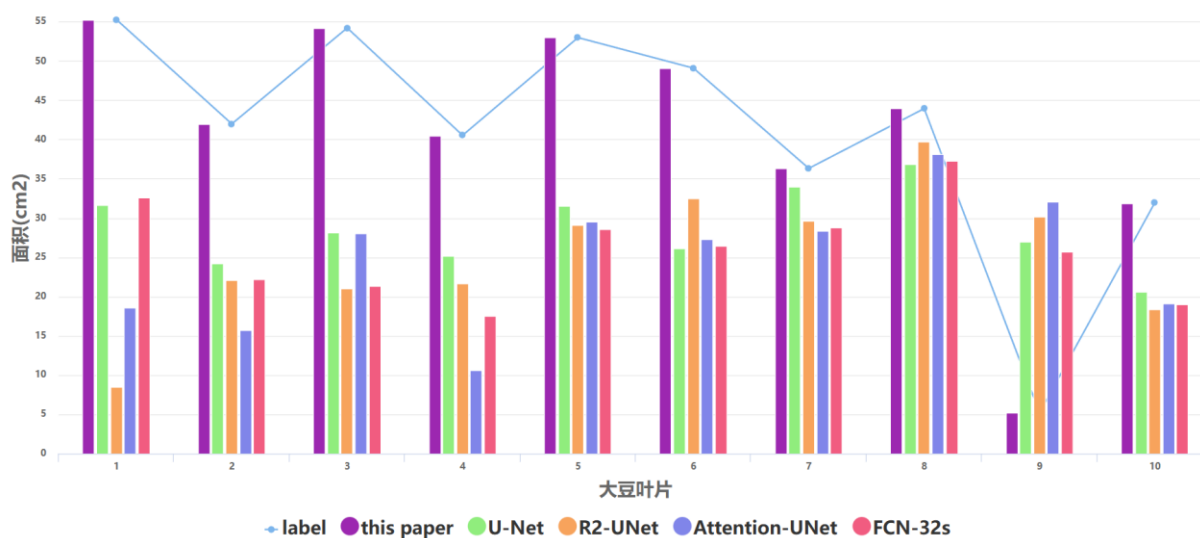


图 4-3 自然(10cm)面积对比

成功分割目标叶片后,利用公式(4-1)计算叶片的面积,在采集高度为 10cm 和 15cm 的情况下分别拟合了两条曲线。两条拟合曲线如 (4-2) 和 (4-3) 所示,其中  $x$  表示像素点数量,  $y$  表示叶片面积:

(1) 10cm 高度

$$y = 9.56 \times 10^{-6}x + 5.152 \quad (4-2)$$

采集高度为 10cm 的情况下,曲线图如图 4-4 所示。

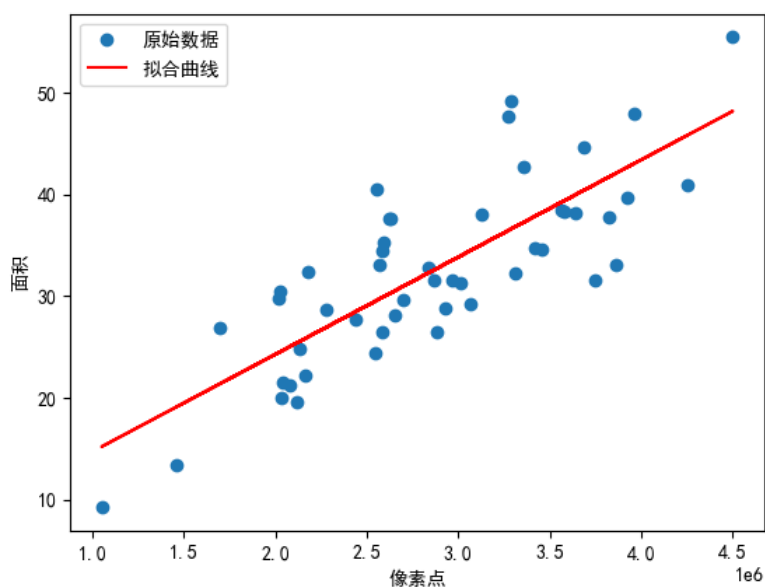


图 4-4 10cm 高度拟合曲线

## (2) 15cm 高度

$$y = 2.94 \times 10^{-5}x + 0.796 \quad (4-3)$$

采集高度为 15cm 的情况下，曲线图如图 4-5 所示。

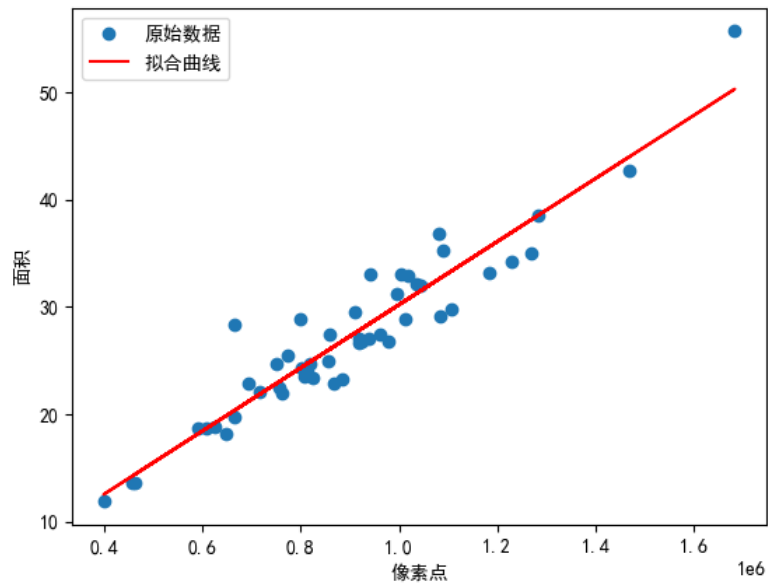


图 4-5 15cm 高度拟合曲线

## 第 5 章 基于 PyQt5 的交互式叶片分割与面积计算软件实现

基于极值点的图像分割的结果作为植物叶片等研究的数据基础，方便快捷地获得分割结果是后续工作的必要保障。为了便于分割结果的使用，采用 PyQt5 开发一款交互式叶片分割与面积计算软件，提供读取图像、分割图像、结果保存和面积估算功能。通过封装基于极值点的深度学习分割算法和基于像素点的面积估算方法，为用户提供一个便于操作的界面。

### 5.1 可视化系统模块框架

交互式叶片分割与面积计算软件共分为三个模块，包括图像文件操作模块、图像分割模块和面积估算模块。如图 5-1 所示为交互式叶片分割与面积估算系统功能模块的具体分类。

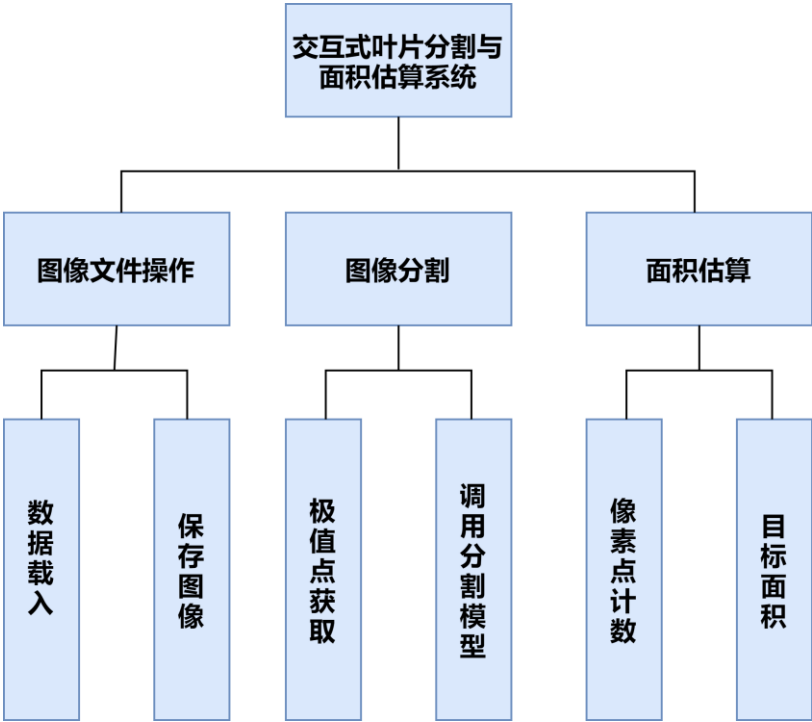


图 5-1 系统模块分类

该系统的详细功能如下：

(1) 图像文件操作模块

该模块可直接读取 jpg 和 png 格式的图像并显示，在读取图像的同时显示图像的规格信息作为图像分析的辅助信息。还有对图像分割结果保存的功能。

(2) 图像分割模块

该模块用于获取极值点，提示用户点击物体最上、最下、最左、最右四个极值点。将极值点信息作为网络输入的指导信号，然后进行图像分割，并将分割结果显示在系统界面上。

### （3）面积估算模块

该模块结合极值点分割算法，以图像分割结果为基础，计算目标物体的像素点数量，进而得到面积估算结果，并显示在系统界面上。

## 5.2 可视化系统设计与实现

### 5.2.1 系统可视化界面

通过 PyQt5 进行上述模块功能的可视化实现，系统主界面如图 5-2 所示。共包括三个部分：界面的主体部分，即为图像显示区域，用于显示原始图形和分割结果；界面下方左侧为功能操作区域，用于实现关于图像载入、分割等功能；右侧为信息区域，用于显示图像的规格和面积估算结果。



图 5-2 系统主界面

### 5.2.2 具体功能实现

#### （1）图像文件操作

通过功能按钮实现图像文件操作，单击功能操作区域中“打开”按钮即可选择本地图片。被选择的图片将会在原始图像位置处显示，同时将显示图片的像素规格（px），如图 5-3 所示。



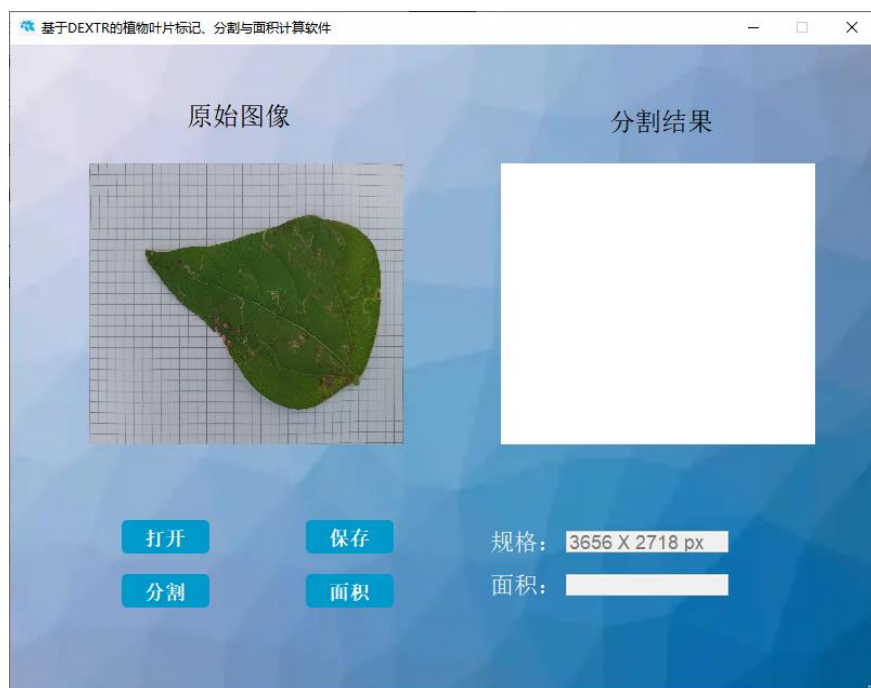


图 5-3 载入图像

图像在 QLabel 控件中显示，需要设置控件的 `pixmap` 属性，即传入图像的路径构建一个 `QPixmap` 对象。将读取的图像显示在左侧框中，并且通过 `LineEdit` 控件显示原始图像的规格大小。

## （2）图像分割

对第三章介绍的基于极值点的分割算法封装，单击 `Button` 控件“分割”，系统将读取到的图片路径作为算法的输入参数，去调用深度学习网络模型。

系统首先会根据已经训练好的权值文件，创建算法的深度学习网络模型，用一系列参数去实例化模型。然后弹出如图 5-4 所示的窗口，等待用户分别选择目标对象的四个极值点（顶部，底部，最左侧和最右侧部分）后，将带注释的极值点作为网络输入的指导信号即可完成相应的图像分割，然后将分割结果显示在上方右侧的框内，如图 5-5 所示。



图 5-4 获取极值点



图 5-5 显示分割结果

单击 Button 控件“保存”，若已选中图像但未进行图像分割，则显示“未分割图片”。若成功分割图像，则弹出“另存为”对话框，让用户选择保存文件的路径和文件名。当用户在“另存为”对话框中选择文件并确认后，将分割结果保存为指定格式的图像文件（.png 或.jpg），提示“保存成功”，如图 5-6 所示。

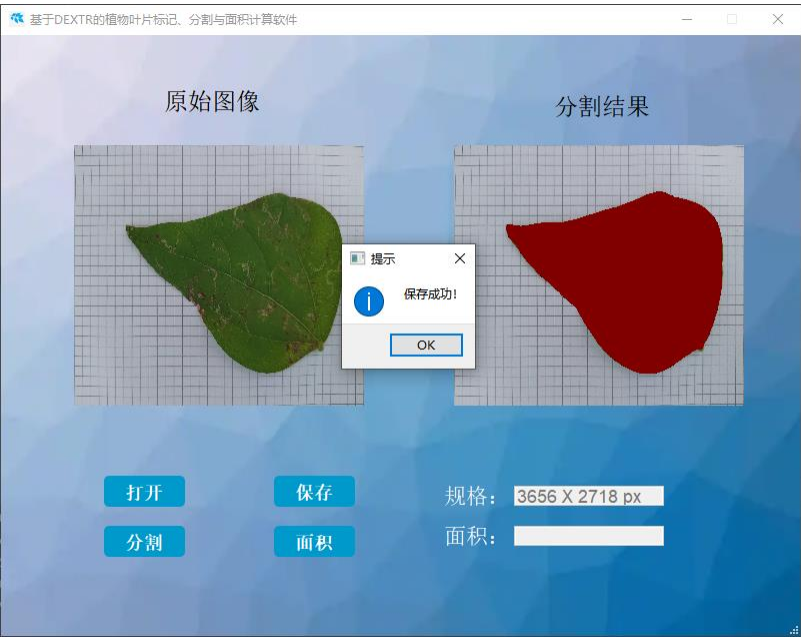


图 5-6 保存分割结果

### (3) 面积估算

单击 **Button** 控件“面积”，将数组中所有像素值乘以 255，将图像灰度值映射到 0~255 的整数范围，并转换为 8 位无符号整数 (uint8)。然后在图像中搜索属于叶片的像素点，得到叶片像素点数，从而计算叶片面积。此后，将面积估算结果显示到界面下方右侧的 **QLineEdit** 对象“面积”处，单位为平方厘米，如图 5-7 所示。

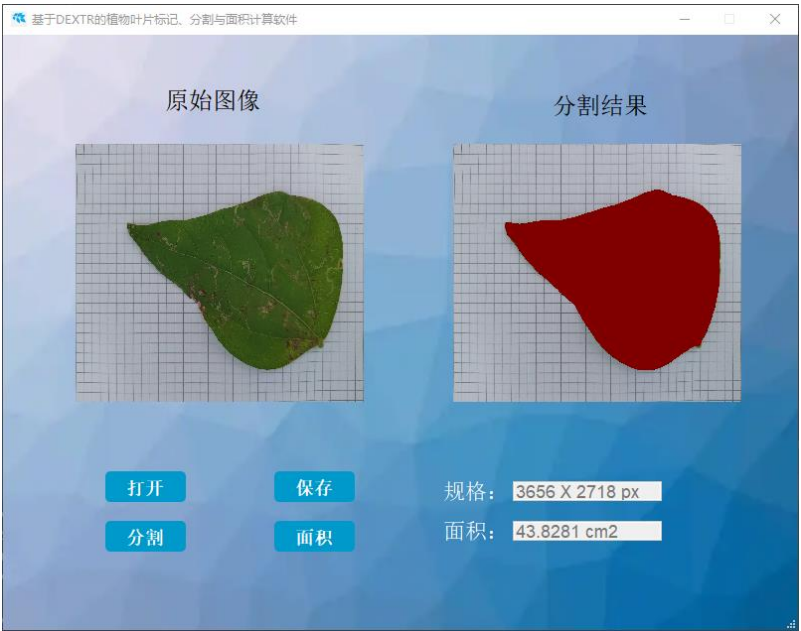


图 5-7 叶片面积计算



## 第6章 总结与展望

### 6.1 研究总结

本文基于 Python 语言实现了大豆叶片分割与面积估算系统设计与实现。基于这两个重要研究内容，即叶片分割和面积估算，研究总结如下：

(1) 大豆叶片分割部分，首先由于目前缺乏专门针对于大豆叶片表型特征的数据集，所以人工采集大量大豆叶片图片建立数据集，填补了这一空缺。本文基于 Python 语言，建立起了四种常见的图像分割算法模型 U-Net、R2-UNet、Attention-UNet 和 Fcn-32s，并进行了训练，然后实现了基于极值点的叶片分割算法。最后比较了上述算法模型对大豆叶片数据集的分割效果。结果表明，基于极值点的叶片分割算法取得了最佳的分割效果。(2) 面积估算部分，本文实现了基于像素点的面积估算方法。利用极值点分割算法得到的分割图像，在非自然场景下分别计算整张图像和目标叶片所占的像素点数量，得到叶片在整张图像中所占的比例，进而估算叶片的面积。

通过本文的研究，完成了大豆叶片分割与面积估算系统的设计与实现，通过对比不同算法的分割效果，进而得出更加适合本文数据集的极值点分割算法模型，它同时适用于自然状态和非自然状态的植物叶片分割，能实现自动，快速，准确地对大豆叶片进行分割，同时不对叶片本身造成伤害。并整合分割算法和面积估算方法，基于 PyQt5 设计了交互式的叶片分割与面积计算软件。软件可作为数据标注、图像分割等工作的工具，为植物表型研究提供便利。本次研究对后期进行植物生长状态诊断、生长预测、检测病虫害和作物优化具有重要的研究意义。

### 6.2 研究展望

对于大豆叶片分割和面积估算系统的实现，在后续的研究中，仍然可以在本文的基础上进一步优化：

(1) 算法的执行速度较慢。由于算法需要进行多次迭代，如果处理的图像尺寸较大，会导致算法的执行时间较长，这在实际应用中可能会影响效率；(2) 对于一些形状不规则或具有缺陷的物体，算法的分割结果可能并不理想。例如，对于某些边界曲线较为复杂的物体，算法可能无法准确识别其边界，并产生一些虚假的分割线；(3) 算法在一些边缘处理问题上还有改进空间。例如，在处理图像边缘时，算法可能会检测到一些虚假的边缘，这可能会对分割结果造成影响；(4) 数据集较小，算法模型的性能和泛化能力受到一定的影响。这些问题希望在今后的研究中进一步完善。



## 参考文献

- 翁杨, 曾睿, 吴陈铭, 王猛, 王秀杰, 刘永进. 2019. 基于深度学习的农业植物表型研究综述. 中国科学: 生命科学, 49(06):698-716.
- 任天佑, 马桂荣, 郭金香. 1982. 春播大豆高产品种的产量构成和形态特征. 山西农业科学, 12
- 陈柱, 杨君. 2022. 基于三维点云的植物叶片重建及其面积估算. 传感技术学报, 35(03):349-354.
- 陈家茜, 刘昌华. 2021. 基于植物图像的甘蓝型油菜叶片面积测量方法. 武汉轻工大学学报. 40(02): 18-23
- Zhu S, Xia X, Zhang Q, et al. 2007. An image segmentation algorithm in image processing based on threshold segmentation, 2007 third international IEEE conference on signal-image technologies and internet-based system. IEEE, Shanghai, China: 673-678.
- Kittler J, Illingworth J. 1985. On threshold selection using clustering criteria. IEEE transactions on systems, man, and cybernetics, vol.5: 652-655.
- Stephen Gould, Tianshi Gao, and Daphne Koller. 2009. Region-based segmentation and object detection. In Proceedings of the 22nd International Conference on Neural Information Processing Systems (NIPS'09). Curran Associates Inc., Red Hook, NY, USA, 655-663.
- Iannizzotto G, Vita L. 2000. Fast and accurate edge-based segmentation with no contour smoothing in 2-D real images. IEEE Transactions on Image Processing, vol. 9: 1232-1237.
- Gao J, Wang B, Wang Z, et al. 2020. A wavelet transform-based image segmentation method. Optik, vol.208: 164123.
- Bhanu B, Lee S, Ming J. 1995. Adaptive image segmentation using a genetic algorithm. IEEE Transactions on systems, man, and cybernetics, vol. 25: 1543-1567.
- Sagiv C, Sochen N A, Zeevi Y Y. 2006. Integrated active contours for texture segmentation. IEEE transactions on image processing, vol.15: 1633-1646.
- Simonyan K, Zisserman A. 2014 Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
- He K, Zhang X, Ren S, et al. 2016. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas, NV, USA. 770-778
- Girshick R, Donahue J, Darrell T, et al. 2014. Rich feature hierarchies for accurate object detection and

- semantic segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition. Columbus, Ohio. 580-587
- Ross Girshick. 2015. Fast R-CNN. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV) (ICCV '15). IEEE Computer Society, USA, 1440–1448
- Ren S, He K, Girshick R, et al. 2016. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 39, 6: 1137-1149
- He K, Gkioxari G, Dollár P, et al. 2017. Mask r-cnn. Proceedings of the IEEE international conference on computer vision. 2961-2969
- Evan Shelhamer, Jonathan Long, and Trevor Darrell. 2017. Fully Convolutional Networks for Semantic Segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 39, 4, 640–651
- Ronneberger O, Fischer P, Brox T. 2015. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. Springer, Cham. 234-241
- Badrinarayanan V, Kendall A, Cipolla R. 2017 Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence, 39,12: 2481-2495
- Q. Yang, W. Fu, B. Pan, L. Zhang, B. Feng and L. Li. 2019.Design of Intelligent Acquisition System for Tomato LeafArea. 2019 IEEE 4th InternationalConference on Signal and Image Processing (ICSIP). 875-878
- Wei Ke, Tianliang Zhang, Zeyi Huang, Qixiang Ye, Jianzhuang Liu, and Dong Huang. 2020. Multiple Anchor Learning for Visual Object Detection. In IEEE CVPR. 10206-10215
- Mingxing Tan, Ruoming Pang, Quoc V.Le, Google Research, Brain Team. 2020. EfficientDet: Scalable and Efficient Object Detection. In IEEE CVPR. 10781-10790
- Qi Fan, Lei Ke, Wenjie Pei, Chi-Keung Tang, and Yu-Wing Tai. 2020. Commonality-Parsing Network across Shape and Appearance for Partially Supervised Instance Segmentation. In ECCV. 379-396
- Maninis K K, Caelles S, Pont-Tuset J, et al. 2018. Deep extreme cut: From extreme points to object segmentation. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.: 616-625
- Rupinder Kaur and Er. Garima Malik, 2014, An Image Segmentation Using Improved FCM Watershed



- Algorithm and DBMF, *Journal of Image and Graphics*, 2,2:106-112
- Thanammal K. K., Jayasudha J. S., Vijayalakshmi R. R., and Arumugaperumal S. 2014, Effective Histogram Thresholding Techniques for Natural Images Using Segmentation, *Journal of Image and Graphics*, 2,2: 113-116
- Noël Richard, Christine Fernandez-Maloigne, Cristian Bonanomi, and Alessandro Rizzi, 2013, Fuzzy Color Image Segmentation using Watershed Transform," *Journal of Image and Graphics*, 1,3: 157-160
- Waseem Khan, 2013, Image Segmentation Techniques: A Survey, *Journal of Image and Graphics*, 1, 4: 166-170
- Paproki A, Sirault X, Berry S, et al. 2012. A novel mesh processing based technique for 3d plant analysis. *BMC Plant Biol*, 12: 63–75
- SANTOS T T, KOENIGKAN L V, BARBEDO J G A, et al. 2014. 3D plant modeling: localization, mapping and segmentation for plant phenotyping using a single hand-held camera. *Computer Vision-ECCV 2014 Workshops*. Springer International Publishing. 10.1007/978-3-319-16220-1. Chapter 18(2015):247-263
- Tian Z, Shen C, Chen H. 2020. Conditional convolutions for instance segmentation. *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I* 16. Springer International Publishing. 282-298
- Pandey S K, Singh H. 2011. A simple, cost-effective method for leaf area estimation. *Journal of botany*. 1-6
- Zhang L, Liu X S. 2010. Non-destructive leaf-area estimation for *Bergenia purpurascens* across timberline ecotone, southeast Tibet. *Annales Botanici Fennici*. Finnish Zoological and Botanical Publishing Board, 47, 5: 346-352
- Yau W K, Ng O E, Lee S W. 2021. Portable device for contactless, non-destructive and in situ outdoor individual leaf area measurement. *Computers and Electronics in Agriculture*, 187: 106278
- Suárez J C, Casanoves F, Di Rienzo J. 2022. Non-Destructive Estimation of the Leaf Weight and Leaf Area in Common Bean. *Agronomy*, 12(3): 711
- Constantino K P, Gonzales E J, Lazaro L M, et al. 2018. Towards an automated plant height measurement and tiller segmentation of rice crops using image processing. In: *Mechatronics and Machine Vision in Practice 3*. Cham: Springer International Publishing, 155–168
- Angtian Wang, Yihong Sun, Adam Kortylewski, Alan Yuille, Johns Hopkins University. 2020. Robust Object

- Detection under Occlusion with Context-Aware CompositionalNets. In IEEE CVPR. 12645-12654
- P. E. L. Otoyá and S. R. P. Gardini. 2020. Real-Time Non-Invasive Leaf Area Measurement Method using Depth Images. IEEE ANDESCON. 1-6
- Alom M Z, Yakopcic C, Hasan M, et al. 2019. Recurrent residual U-Net for medical image segmentation. Journal of Medical Imaging, 6(1): 014006-014006
- Oktay O, Schlemper J, Folgoc L L, et al. 2018. Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999
- Papadopoulos D P, Uijlings J R R, Keller F, et al. 2017. Extreme Clicking for Efficient Object Annotation. 2017 IEEE International Conference on Computer Vision (ICCV). IEEE : 4940-4949
- Zhao H, Shi J, Qi X, et al. 2017. Pyramid scene parsing network. Proceedings of the IEEE conference on computer vision and pattern recognition. 2881-2890.

## 攻读本科学位期间发表的学术论文和参加科研情况

### 一、发表论文、著作情况

(1) ICMLC 2023 会议论文《Soybean Leaf Segmentation and Area Estimation Based on Extreme Points》(EI 索引, Oral presentation)

(2) 软件著作权《基于 DEXTR 的植物叶片标记、分割与面积计算软件》(登记号: 2022SR0719456)

### 二、参加科研情况

(1) 主持国家级大学生创新创业训练计划项目“基于深度学习的大豆叶片检测与面积计算方法研究”(已结题, 结题等级为良好)

### 三、所获奖励与荣誉

(1) 2023 15th ICMLC 国际会议“最佳报告奖”。



## 致 谢

行文至此，感慨颇深。不知不觉间四年大学生活即将结束，在西北农林科技大学的四年里，有过遗憾，有过喜悦，但更多的是收获。

在此，首先感谢我的指导老师冯明晨老师。在本科学习期间中，他一直给予我耐心指导和悉心关怀。从开题到正式撰写论文，从选题研究到实验设计，他的指导让我对毕业设计充满信心和热情。感谢他不断的关注和支持，他在学术上的深入指导，让我感悟到了研究的真谛。感谢他的无私奉献和深入的教学理念，让我在学术方面得以更快的成长。

其次，非常感谢韩亚敏老师在评阅论文时提出了许多宝贵的修改建议和意见，这些意见对于提高我的论文质量起到了极大的作用。她的专业知识和严谨态度一直是我学术上的榜样。在此，我要向韩老师表达我最诚挚的感谢和敬意。

另外，我也得到过很多同学们的支持和帮助。感谢同学们的友爱与包容，他们对我充满的信任和鼓励让我在学术和生活方面受益匪浅。

最后，感谢我的家人，他们是支持我不断前进，追求远大理想的坚强后盾。

学生姓名：杨治梁

日期：二〇二三年五月