分析task2：

本质上是image segmentation 的问题

Method：

预处理：

1 滤波

获取图像后，首先需要对图像进行滤波预处理，降低噪声。 由于噪声属于高频信号，本任务选用高斯滤波器及中位滤波器一同去除噪声，其中高斯滤波 传递函数的形式如公式（１）所示

高斯公式

中位公式

2 图片色彩格式：

预处理完成后，紧接着进行植物提取。 也就意味着将植物 从图像中分割出来。常用的分割方法有聚类法、基于区域的方 法等。聚类法是对图像中的每一个像素点的色彩特征进行聚类 分析；基于区域的方法包括区域生长、区域分裂和聚合等10 。 这些方法运算时间较长，不能满足高通量环境下的高速运算需 要，且由于图像中包含多种特征区域，分割算法不稳定。

1 RGB归一化

考虑到植物的色彩表现为绿色，可以在 ＲＧＢ色彩空间中， 将绿色作为特征，通过使用阈值的方法将图像分割为绿色和非 绿色部分。随后便可以实现植物的提取。在 ＲＧＢ空间中，任一 像素点的颜色可以看作红、绿、蓝三个分量不同比例的组合，一 幅彩色图像即可以看成是红、绿、蓝三个灰度平面组合而成，如 果某点处的三个分量为灰度的最大值则该点表现为白色。但是 由于 ＲＧＢ色彩空间受光照影响很大，为了减少光照的影响，首 先在 ＲＧＢ空间中进行归一化处理。 即从图像中提取出绿分量 ｇ、红分量 ｒ和蓝分量 ｂ，归一化后为 Ｒ、Ｇ、Ｂ（见式（２））。为进行绿 色特征的判断，考虑使用［－１　２　 －１］模板对图像各点处（Ｒ、 Ｇ、Ｂ）分量进行处理， 如式（３）所示。 该模板强调绿色分量，压 制红色和蓝色分量，且模板和为 ０。模板具有对称性，且长度为 奇数，因此具有线性相位，保证时不变要求。 理想情况下，对于 蓝色和红色显著的区域，使用模板运算后的结果必为负值，而白 色区域使用模板运算后的结果为 ０，只有绿色区域在模板运算 后表现为大于 ０的结果。 其他颜色区域模板运算结果必为负 值，这个特点非常重要，只要使用大于 ０作为阈值判断的依据便 可以很方便地将绿色分割出来，避免了其他算法中阈值的选择， 体现了算法对图像的无关性。由于从归一化到模板运算是对每 个像素点使用加法和乘法，因此该算法可以在数字信号处理器 ＤＳＰ上高效实现，在保证 ＤＳＰ性能发挥的同时，也可以实现高 速处理，满足了高通量环境下对处理速度的要求。

A picture containing table

Description automatically generated

2. 基于HSV 色彩格式的图像分割

RGB 颜色空间利用三个颜色分量的线性组合来表示颜色，任何颜色都与这三个分量有关，而且这三个分量是高度相关的，所以连续变换颜色时并不直观，想对图像的颜色进行调整需要更改这三个分量才行。通过观察本task的输入数据，可以发现照片中的自然光照、遮挡和阴影等情况的影响较大，而且图片不同部分亮度差异明显。

RGB 颜色空间中，三个色彩分量都与亮度密切相关。也就是说，只要亮度改变，它们都会随之相应地改变，而没有一种更直观的方式来表达。但是人眼对于这三种颜色分量的敏感程度是不一样的，在单色中，人眼对红色最不敏感，蓝色最敏感，所以 RGB 颜色空间是一种均匀性较差的颜色空间。如果颜色的相似性直接用欧氏距离来度量，其结果与人眼视觉会有较大的偏差。对于某一种颜色，很难推测出较为精确的三个分量数值来表示。

基于上述理由，在图像处理中使用较多的是 HSV 颜色空间，它比 RGB 更接近人们对彩色的感知经验。非常直观地表达颜色的色调、鲜艳程度和明暗程度，方便进行颜色的对比。在 HSV 颜色空间下，比 BGR 更容易跟踪某种颜色的物体，常用于分割指定颜色的物体。

针对本项目而言，植物的叶子本身和背景会有较大的差异度，叶片直接的差异也仅仅体现在颜色色调、明暗程度和饱和程度上，而进一步的识别叶面面积，成熟程度，受病害虫害程度也需要依赖这些关键参数，所以在本项目中对两种不同的颜色空间格式对后续的分类预测准确度等指标进行了横向对比。

3 形态学算法：

1 形态学去除杂散点：

经过以上方法处理后的图像会都会一些杂散点，这时可以采用形态学算法中的开运算去除10:

结构元 Ｂ对图像 Ａ的开运算记作 Ａ 0Ｂ， 如式（４）所示，对图  像先腐蚀后膨胀。腐蚀可以去除杂散点，但同时也使叶片区域 变小，故需通过膨胀操作来恢复。

A picture containing logo

Description automatically generated

腐蚀记作 Ａ Ｂ，如式（５）所示，其中，Ａ 是 Ａ的补集， 是 Θ  空集。A picture containing text

Description automatically generated

膨胀记作 Ａ Ｂ，如式（６）所示，其中，Ｂ是 Ｂ关于原点对称 的映像， 是空集。

A picture containing text

Description automatically generated

实验步骤：

DSC 通过估计两个轮廓的相互重叠部分体积 （面积）与其总体积（面积）的百分比来描述它们之间 的相似程度。计算公式为：

A close up of a clock

Description automatically generated

其中，A和B分别表示两个目标区域；\union表示取二者共 同的部分；+表示取二者之和。这种计算方法是运用得最多的比较方法［8］ 。实际上，DSC是Kappa系数中 的一个特例［9］，在极限情况下，Kappa 系数的计算公 式可以被简成现在的 DSC 计算公式。早在 1945 年， Dice［10］ 就在研究生态学上不同物种之间数量时提到 这个参数。

8 SHARP G, FRITSCHER K D, PEKAR V, et al. Vision 20/20: perspectives on automated image segmentation for radiotherapy［J］. Med Phys, 2014, 41(5): 4871620.

9 ZIJDENBOS A P, DAWANT B M, MARGOLIN R A, et al. Morphometric analysis of white matter lesions in MR images: method and validation［J］. IEEE Trans Med Imaging, 1994, 13(4): 716-724.

10 DICE L R. Measures of the amount of ecologic association between species［J］. Ecology, 1945, 26: 297-302.

通过式（1）计算得到的DSC是一个取值范围在0 到 1 之间的参数。数值为 1 的时候表示两个目标区 域完全重合，数值为 0 则表示两个目标区域完全分离。

DSC对位置和大小的差异性都很敏感，例如,大小相等的两个图形，把它们的一半区域重叠，彼此只有一半面积，导致 DSC=0.5；或者一个区域完全覆盖 在另一个比它小一半上的区域上，DSC 就只有三分之二。在 DSC 测量图像相似性的过程中，位置的差 异相比大小的差异更能反映出不同，重合相同体积(面积)的图像，其所在位置不同会导致最后的 DSC 不同。这样能够反映出直观的感觉，两个区域(其中一个完全包含另一个区域)比两个部分重叠区域更 加相似。结合实际轮廓图像处理来看，虽然 DSC 在 计算过程中具有简洁明了的优点，但它却不能描述 所有的情况。