**高光谱图像数据分析部分**

V 1.0.1

2023-5-2

**第一阶段的目标：**分析单株盆栽水稻的高光谱图像数据，获取叶绿素含量、叶片氮素含量、叶片光合速率、叶片颜色，植物干旱胁迫和盐胁迫程度表型的量化，植物病虫害表型的量化。

**产品技术参数的要求：**

**可测量植物表型：**植物冠层光谱反射率，可见光及近红外波段的全部光谱 反射指数包括 NDVI 等，由光谱反射数据可用于反演叶绿素含量，叶片氮含量，次生代谢物含量，以及光合速率参数，生物及非生物胁迫表型。

**高光谱图像分析软件：**高光谱数据查看，光谱反射率计算，数据滤波处理，分析计算全部常用光谱指数，反演计算叶绿素含量，叶片氮含量，次生代谢物含量，以及光合速率参数，生物及非生物胁迫表型。

**数据分析过程：**

1. 读取高光谱文件，显示RGB图像，保存图像。
2. 计算NDVI，根据NDVI的阈值去除非植物部分，显示去除背景后的RGB图像，保存图像。

NDVI阈值 0/0.2

1. 计算植物部分的每个像素的光谱总辐射的亮度值，根据亮度值设置的阈值去除荫蔽叶片（这些叶片接收到的入射光较低，与反射率参考板的入射光强差别过大，无法准确计算这些荫蔽叶片的反射率）。显示去除荫蔽叶片后的RGB图像，保存图像。

亮度值阈值：1000

1. 计算光照叶片部分的全部像素的光谱平均曲线，显示曲线，保存图像并且保存曲线数据到csv表格。

----k, b计算正常， 从光谱值到反射率的校正，计算过程有误，仍待解决

1. 根据光照叶片的光谱平均曲线，计算光谱指数，计算叶绿素含量，光合速率等表型参数。（具体计算公式参考文档最后部分，以及来自张美玲同学的python代码）。（先计算平均光谱曲线，再计算表型参数）

使用gridSearchCV预测叶绿素含量与AC

1. 根据上述第3步的结果，光照叶片部分的每个像素单独计算“光谱指数，计算叶绿素含量，光合速率等表型参数”，可以考虑生成伪彩图并保存图像；还可以将这些参数按照像素保存到csv表格（非光照叶片部分的像素位置对应参数可以默认为NA）。对于每个参数计算平均值（先单个像素计算表型参数，再平均）。
2. 聚类，根据原始数据聚类，或者根据去除背景后数据聚类。聚类的目的是进一步区分正常叶片和衰老叶片（受胁迫后的枯黄叶片），或者区分正常叶片和病斑叶片（受病虫害后的黄色，白色，红色或者褐色等颜色）。聚类后，判断哪一类的像素点表示衰老叶片区域或者病斑叶片区域（之后，我们通过扫描一定量的枯黄叶片和病斑叶片数据，建立特征谱库用于辅助这里的分类和判断）。最后，计算枯叶类和病斑类的像素点占总正常叶片像素点的比例，这一比例作为衡量胁迫表型的参数。
3. 对上述枯叶类或者病斑类的像素区域计算平均光谱，并且将光谱保存到csv表格。
4. 最后，再考虑水稻开花结穗后，通过光谱区分稻穗和叶片。通过稻穗的光谱特征计算其籽粒灌浆成熟度。（这部分需要查阅文献，或者单独的采集数据构建统计回归等模型）。

树上的叶子

低可信度描述已自动生成

图1， 单株水稻的示意图。盆栽水稻，仅能看到地上部分。

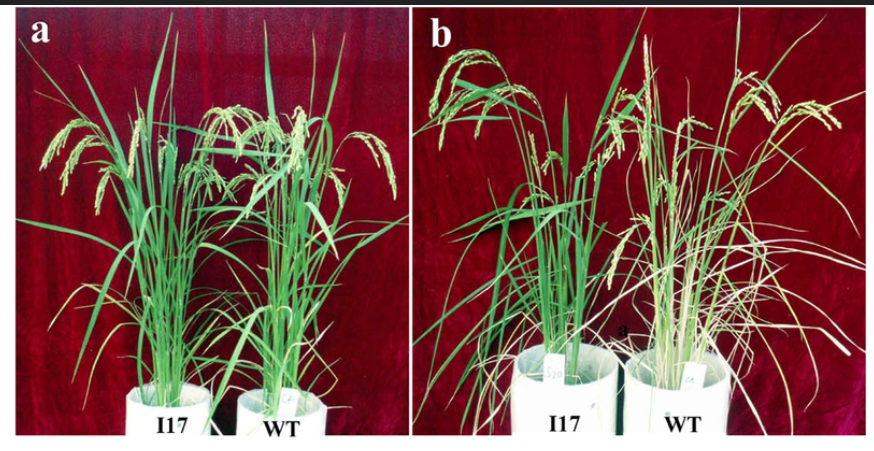
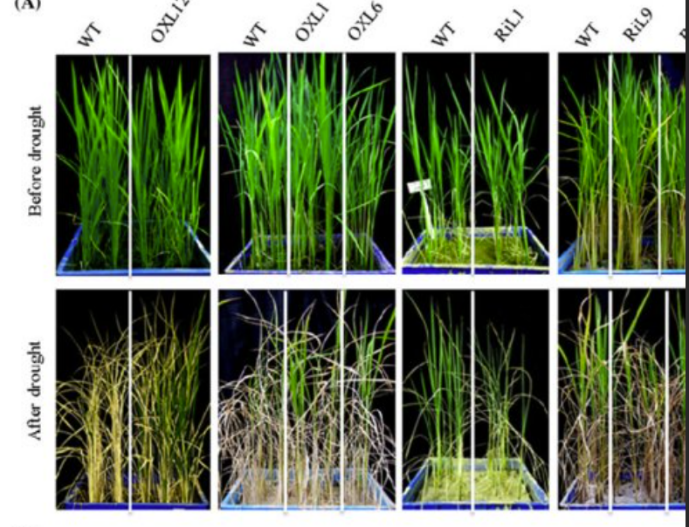


图2，干旱胁迫的单株水稻（盆栽）的表型，胁迫严重的植株会有明显的颜色浅绿，黄色或者白色的部分比较多。仅用于参考。

 图示

描述已自动生成

**图3**，干旱胁迫的水稻植株（苗期）图片。我们目前不用考虑一个图像里有多个植株的情况。仅用于参考。

图示

描述已自动生成

图4，苗期水稻（多株盆栽）的干旱处理实验。仅用于参考。

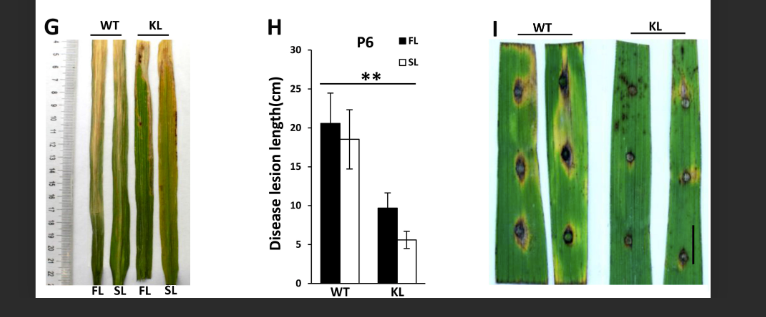


图5， 水稻叶片病斑，仅用于参考。各种不同的病斑会有很大差别，需要有具体的数据才能开发这些功能。

**常见光谱指数：**

还可以再查阅一些文献，整理更多的光谱指数。

**NDVI**

Normalized difference vegetation index

(R800 - R680)/(R800 + R680)

Plant structure, greenness (Rouse, 1974)

**OSAVI**

Optimized soil adjusted vegetation index

(1 + 0.16) · (R800 − R670)/(R800 + R670 +  0.16)

Plant structure, greenness (Rondeaux et al., 1996; (Haboudane et al., 2002)

**PSSRa**

Pigment-specific simple ratio a

(R800/R680)

Concentration of chlorophyll a (Blackburn, 1998)

**PSSRb**

Pigment-specific simple ratio b

(R800/R635)

Concentration of chlorophyll b (Blackburn, 1998)

**PRI**

Photochemical reflectance index

(R570 - R531)/(R570 + R531)

Photosynthetic activity (Gamon et al., 1992)

**MTVI2**

Modified triangular vegetation index - improved

1.5· [1.2(R800 - R550) - 2.5 · (R670 - R550)]/√[(2R800 +1)2 - (6R800 - 5√ R670) – 0.5]

Green leaf area index (Haboudane et al., 2004)

**RS-ETRi （这个指数目前不能计算，SIF是太阳诱导荧光参数）**

Remote-Sensing Electron Transport Rate index

SIF 687/(SIF 687 + SIF 760)·NDVI·PAR

Photosynthetic electron transport rate (ETR) index. (Liran et al., 2020)

**使用PLSR模型预测SPAD，A1200，N，Ca，Cb：**

表型参数：

SPAD：叶绿素含量（使用手持的SPAD叶绿素计快速测量的）；

A1200：光合速率（在光强1200的条件下测量的二氧化碳固定速率）；

N: 叶片氮素含量 （使用元素质谱仪分析得到的叶片干样品内的N元素比例）；

Ca: Chlorophyll a 叶绿素a的含量（使用分光光度计的方法测量）；

Cb: Chlorophyll b 叶绿素b的含量（使用分光光度计的方法测量）;

注：上面的SPAD单位与Ca或Cb的单位是不一样的。

张美玲给的code 和csv文件，可以使用其里面的PLSR模型拟合csv里面的数据得到拟合后的PLSR参数值。

我们可以试验一下，直接使用这个拟合的PLSR模型及参数值是否能预测高光谱图像的相关表型参数。

**对数据采集的改进意见：**

**高光谱相机数据采集软件输出的hdr文件，每个文件都应包含wavelength数据，这个是高光谱相机出厂定标的数据。**

**高光谱相机拍摄每一幅图像的时候，都应该在固定位置包括3%及30%参考板。分析软件可以直接从参考板的图像中的位置来识别参考板。**