一、实验目的

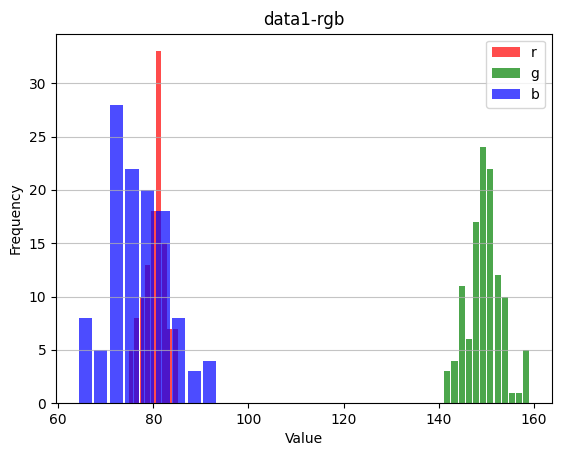
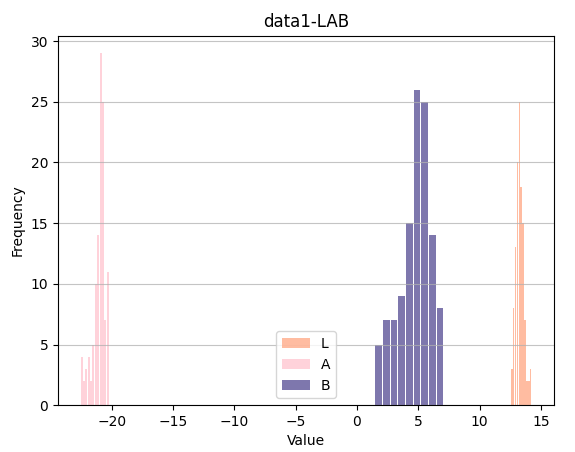
蔡司绿膜数据 rgb值预测lab值

1. 实验数据
2. 第一批数据, 共116张 data1
3. 第二批数据, 共78张 data2
4. 实验结果

实验1:

|  |  |
| --- | --- |
| 数据(116张) | 交叉验证精度(lab三个值的误差均在0.5以内) |
| Data1 | 0.955 |

可视化data1的rgb lab值分布情况:

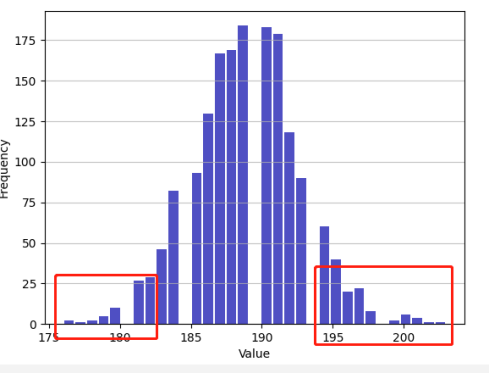
(data1 r g b三个通道的像素值分布) (data1 L A B三个数值分布)

data1 rgb值都分布集中, 对应lab界限分明.

实验2:

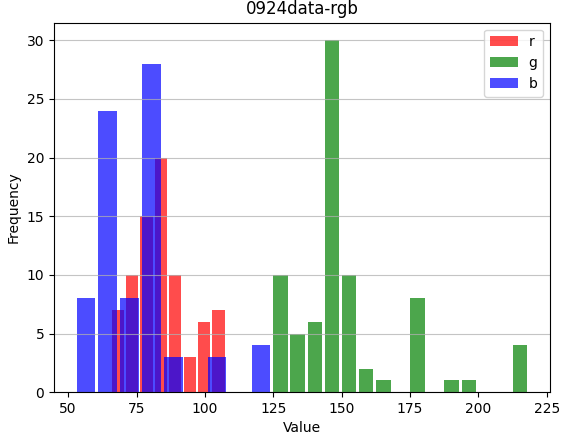
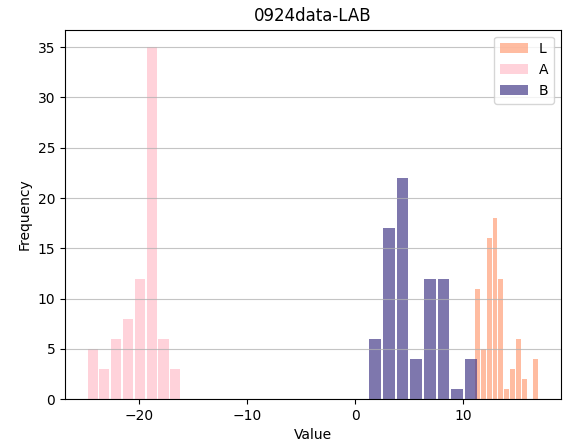
|  |  |
| --- | --- |
| Data2(78张) | |
| Method | 交叉验证精度 |
| Baseline | 0.73 |
| Roi重叠区域均值滤波(3x3) | 0.862 |
| Roi重叠区域均值滤波(5x5) | 0.845 |
| Roi重叠区域中值滤波(3x3) | 0.861 |
| Rgb值分布掐头去尾, 保留”中心”部分像素并取均值[1] | 0.872 |

实验描述[1]: 对每一样本, 提取到重叠区域的41x41矩形. 统计这1681个像素点的rgb三通道的分布情况. 如下图所示为r通道像素分布情况. 可将红色框出的少数量像素值剔除, 保留41x41区域内的”主流”像素值, 并取均值作为本例样本的rgb值.



r通道像素值分布

可视化data2的rgb lab值分布情况:

(data2 r g b三个通道的像素值分布) (data2 L A B三个数值分布)

data2, rgb值分布都不算集中, l b界限不明显. a值范围更广, 猜测可能原因是RGB中RB值分布分散. b值相比data1 偏大且分散.

1. 初步结论
2. data1样本量数量较多且异常数据较少, 精度基本达标; data2数据量少且存在异常样本, 精度还需优化.
3. 实验发现, 对样本的Roi重叠区域做一些预处理可明显提升预测精度. 通过滤波, 剔除头尾分布像素值等方法，均可一定程度上排除成像时背景带来的”噪点”干扰。
4. 样本成像是底盘带来的影响还在排查和优化..
5. 后续
6. 下一批验证数据的数量
7. 非特别异常的样本, 期望回归到的一个lab值范围(蓝绿膜区分开)