data1: data before 0812： 116

0924 data：78

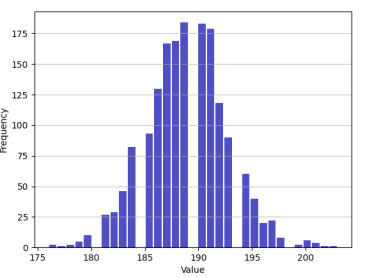
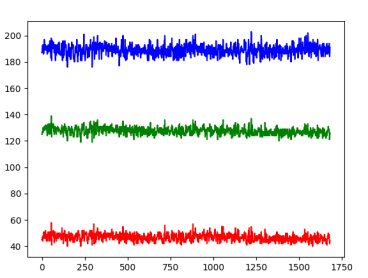
|  |  |  |  |
| --- | --- | --- | --- |
| Params参数搜索 | acc | | |
| Data1 + 0924data acc | Data1 acc | 0924data acc |
| Data1 + 0924data  gamma\_ed | 0.870 | 0.923 | 0.854 |
| Data1 + 0924data  no gamma | 0.731 | 0.928 | 0.755 |
| Data1 gamma\_ed | 0.823 | 0.955 | 0.861 |
| Data1 no gamma | 0.824(data1在0924data上有一点鲁棒性..) | 0.952 | 0.861 |
| 0924data  gamma\_ed | 0.7026 | 0.908 | 0.73 |
| 0924data no gamma | 0.7026 | 0.9023 | 0.732 |

还是需要gamma处理的, 然后, 尽量不要轻易混合数据训练模型, 取样带来的不同误差可能给模型带来异常数据, 使得模型学习紊乱...

0924的data较少且可能比较异常, 需要多一些数据支撑参数搜索[第一行的0.854]

1014 针对0924data的优化

单张样本的rgb分布



|  |  |  |  |
| --- | --- | --- | --- |
| 0924data rgb三个通道统计一下去头去尾选取中间的众数（top3众数） 把背景上的干扰看看能不能排除一下 （可能存在一个rgb组合对应多张样本的问题, rgb值尽量保证float比较好..） | | | |
| 0924data  gamma\_ed: 0.84  不大鲁棒.. | Data1 + 0924data  gammd\_ed:0.72 | （top4众数）  0924data gamma\_ed: 0.77 |  |
| 统计rgb分布然后掐头去尾, 保留出现次数最多的topk(456感觉都可以调下试试), 这些部分像素值再取均值 | | | |
| 0924data: 0.872 到顶... | | | |

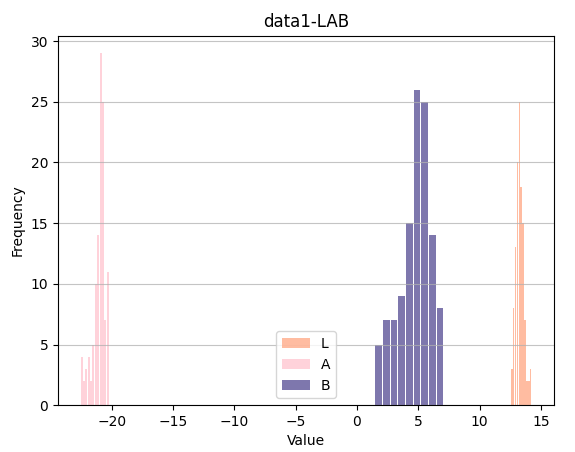
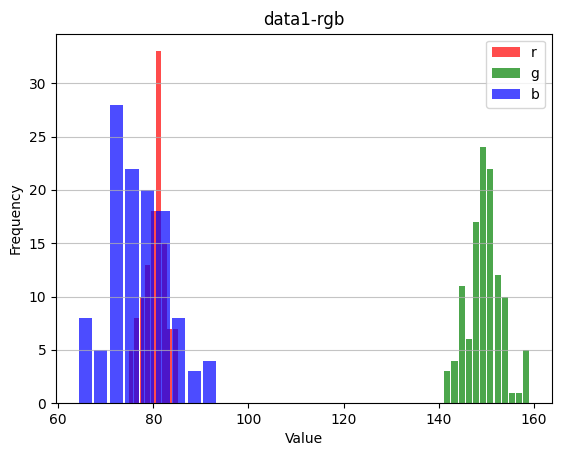
0924data的提取roi脚本是41x41的规整矩形, 所以可以尝试下各种滤波操作

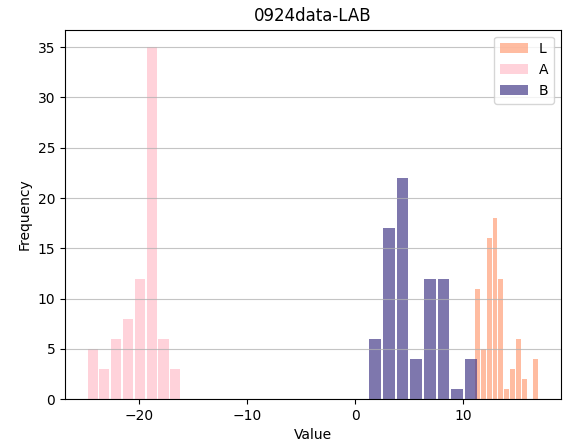
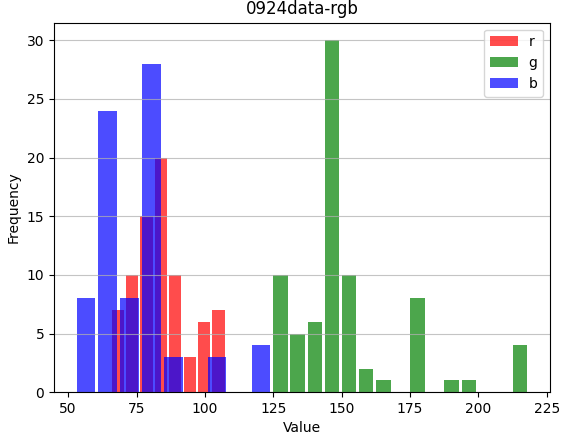
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 0924 data | 均值滤波3x3 | 均值滤波5x5 | 高斯滤波3x3 | 中值滤波3x3 |
|  | 0.862 | 0.845 | 0.721 | 0.8616 |

1015:

一、对比data1 和 0924 green数据

1. data1的lab rgb值写入csv 和0924的data对比. 画一下两批数据的rgb lab直方图





对比数据发现, data1 rgb值都分布集中, 对应lab界限分明.

0924 green data, rgb值分布都不算集中, ab界限不明显.