bucketmanager

### 说明：

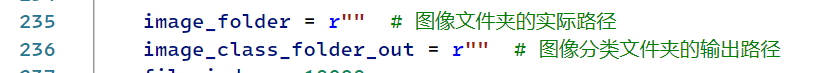
可将源文件夹中所有的图片按照大致比例进行分类并输出到目标文件夹

### 依赖安装：

pip install numpy

### 输入、输出文件夹：

修改代码文件235、236行的图片源文件夹、图片分类输出文件夹路径，例：r"E:\pic"或者r”/data/pic”

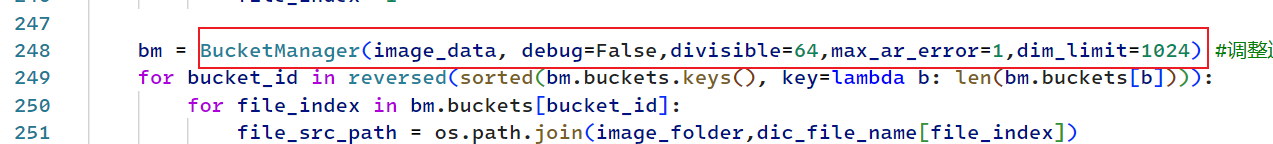


### 运行

终端(如bash、cmd)路径切换到bucketmanager.py所在的文件夹，并输入python bucketmanage.py即可运行

### BM参数：

修改248行BucketManager()函数参数可控制buckets的数量以及每个bucket包含的图片数量



### BM参数参考：

* 1. max\_size: 最大图像分辨率，在代码内仅用于计算 max\_tokens，固定为 max\_tokens = max\_size[0] / 8 \* max\_size[1] \* 8
  2. divisible: 推测为生成桶的循环过程中，w或h每次递增的值，单位为像素，需要能被 8 整除
  3. max\_ar\_error: 给图片分配桶时允许的最大误差，算法为 abs(bucket\_aspects — image\_aspect)，即桶宽高比与图片宽高比允许的最大误差
  4. seed: 种子
  5. step\_size: 无实际用处，推测原设计中为 self.f 的值
  6. min\_dim: 未知
  7. base\_res: 未知，从代码看是提供了一个参考分辨率，作为基准分辨率，即 w/h = 1，并将自己添加到 arb 桶中
  8. bsz: 未知，与 world\_size 共同作用
  9. world\_size: 未知
  10. global\_rank: 未知
  11. dim\_limit: 未知