实验内容总结

1. 收集数据：

中式垃圾桶，中式剪纸为业务类别

每种为常规类别与细粒度类别数据各一半

通过爬虫工具获取，获取后改为ImageNet格式



1. 对于每个类，进行文字编码和图片编码，**计算CLIP余弦分数，给定一个候选阈值 t，把相似度分数转成 0/1 预测：y^​i​(t)=1[s\_i​≥t]**，由此计算出P，R，F1

脚本对每个类别独立地在一串阈值上遍历，记录**每个 t的 P，R，F1**，然后选择 F1 最高的那个阈值 t作为该类别的“最佳阈值”。

1. 使用CLIP: ViT-B/32对于已有数据基于最佳的f1 score进行卡阈值测试：
2. Siglip: 由于模型本身特殊性，输出的余弦相似度都在0附近，所以卡阈值测试中阈值的设定需要做出相应的调整