一、

拿到原始视频流先对根据拍摄时间把视频切分成日间和夜间两份大部分，然后再进行抽帧，之所以在抽帧前就先区分开是想对两种情况用不同的抽帧率。

夜间的数据可以分为两种，一种是RGB的低照度帧，另一种是类似灰度图像的红外帧，这两种情况是摄像头根据实际环境亮度自动切换的。

二、

拿到图像帧比较多，所以接下来就需要进行制定一些标准进行筛选，首先是筛掉一些明显异常的帧，比如说这种，目前是根据图像文件的大小来判断的。然后需要将图片送入一个预标注模型来得到伪标签，这些标签可以帮助我们排除掉大量的空帧（也就是没有检测到物体的帧），和重复帧（也就是相邻的相似度很高的帧），这样保留下来的都是对训练比较有价值的帧了，最后将这些帧的标签进行人工校对和标注，修正误检和漏检，就得到了有真是标签的数据集了

三、

接下来需要将整理完的数据集细分成多个目标域，比如从光照情况分就有低照度、红外的，从检测距离分就有0~5M，5M~10M和10M~30M，根据模型的目标场景可以把这些域进行组合和配比，构建相对应的训练集和验证集。

后续计划

1. 日高抽帧率，挖掘更多的红外数据；
2. 确定10M训练集及benchmark数据集的构建方案；
3. 确定30M的训练集和benchmark扩充方案；