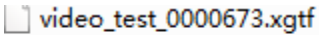


本周还是在写代码阶段，和卿志武交流了一下思路，关于拿数据集中一小组数据拿来跑自己的代码 debug，以及可以用相同的数据跑 caffe 的代码来逐步比较检测自己的代码出现的问题。于是取了一个视频的数据作为测试将数据输入部分的代码修改了一遍，去除了一些 pytorch 的 faster rcnn 代码中的一些可以通过输入的视频和 ground truth 能推理出来的变量删除了，在使用的时候再获取这个参数。将 dataloader 循环时候的输入量只取了输入的视频序列 (batch_size, channel, length, height, width)，以及 ground truth windows(start_time, end_time, class)。接下来是把 RPN 部分的代码完整过了一遍，在做正负样本取个数的时候有几个可选参数没搞懂为什么要做重复操作，准备在程序跑通以后再看看区别。现在正在看 ROI 处理的代码，应该下周能将训练部分的程序跑起来。

关于 THUMOS14 数据集中 Test 部分的问题，xgtf 文件夹里面一共 214 个文件，但有两个文件是  video_test_0000673.xgtf~，也就是有 213 个不同的视频，但其中有一个视频 0001292 在 annotation 中只存在于 Ambiguous_test 中，故将其舍去，于是只有 212 个测试的视频。然后做了抽帧后的 check 操作，由于我抽帧使用的是 ffmpeg，设置了 fps 为 25（在 roidb 中采用的 ground truth 的帧长为秒数*25），因此导致了得到的图片数量和视频帧数存在差别，针对不同 fps 的视频采取不同设置抽帧，最终保证了 opencv 得到的视频帧数和抽帧得到的图片数量是相等的。