

本周把 caffe 的 R-C3D 程序跑通了，用 UCF101 上的 C3D 作为预训练的参数，训练得到的模型在 threshold 为 0.5 的条件下进行测试，得到的 mAP 为 0.286，同时这份代码给了一个训练好的 R-C3D 的 caffemodel，但下载了好几次都失败了，就还没有测试，他的模型给出的在 threshold=0.5 时 mAP 是 0.288，论文中给出的 mAP 是 0.289。在训练的时候代码默认的网络存储地址不存在，会在训练完保存的时候报错，需要手动建立文件夹。

关于自己的代码，我参考 faster rcnn 那份代码的 trainval_net.py 的训练文件，按顺序改成自己的文件，所使用的 config 文件用的 caffe 的 R-C3D 的，将视频以 25FPS 转换成图片后，得到视频含有其标记信息 roidb 以字典的方式存储在 pkl 文件中。这种把 dataloader 函数相关的代码完成了，其中 sampler 采用了自己的定义，迭代的时候对输入数据的顺序做了随机处理，此处可能是为了确保维度上的问题自己写了 sampler。在 dataset 的处理上，faster rcnn 做了 resize 和 crop 以及翻转来数据增强的处理，在 R-C3D 里面只做了 resize 处理，在 ground truth 的映射上只需要在 D 维上修改，相对 faster rcnn 来说更简单。

现在正在看 model 的构建这部分的代码，也是准备参照着 fasterRCNN 这个类来写自己的 RC3D 类，主要是搞懂 RPN 和 ROI 两部分的逻辑，下周的目标是完成这部分，我考试时间大概集中在元旦之后了，趁前面这段时间尽量完成自己的代码。