1. **方案一**

**https://imatge-upc.github.io/activitynet-2016-cvprw/**

采用Large Scale Activity Recognition Challenge(CVPR2016)的ActivityNet Challenge第一名的模型。该模型将视频剪辑为16帧作为模型的输入，运用3D卷积抓取视频的特征，训练RNN对视频进行分类。



**分析：优点** 给出了源代码和相关文献，模型相对简单，便于研究和实施

**缺点 1.**在分类和检测的任务中，精度分别为0.5874和0.2237，可能精度达不到现在的要求2. 视频需要单个动作剪辑处理训练，耗时较长

1. **方案二**

**多流双向RNN的细微行为检测**

在以往，视频行为检测的成功案例更多的是采用二流CNN运用于光流和图像帧。在此方法中，采用定位在人的边界框的追踪算法，这不仅为图像和移动提供了参考框架，也抑制了不在边界框的噪声。文献称模型 在MPII Cooking2 和 a new Shopping 两个数据集的表现胜过当前行为分析的任何方法。

视频行为分析Demo链接：

https://www.youtube.com/watch?v=IIHKEs9m3WM

**分析：优点：**检测精度高，行为分类细致

**缺点：** 模型性对复杂，没有给出源代码，需要细细研究