1.抽帧需求：10s 1帧

2.类别标签：多人场景对目标标框，必须确保目标上半身基本都在框内



**单人多标签：**

headDown：低头

headUp ：抬头

raiseHand：有明显举手行为；

usePhone：画面中有手机，且明显低头或者抬头看手机的行为；

sleep（趴伏）：明显趴在座位上睡觉行为；

stand：课间帧中出现站立、走动的行为；

takeNote：画面中有手握笔的行为；

lookBook ：画面中有书，低头看书行为；

sleep（坐着）：明显坐着睡觉行为；

listenClass ：抬头看像黑板那方的行为；

daze：发呆行为

（若不确定目标是否为以上行为，只对目标标定抬头或低头的单标签即可）

**多人多标签：**

talk：明显两人侧身交流行为

3.量：1个教室1个视频

4.视频：

3A102\_1(60min)、3A102\_2(19s)

3C102\_2(60min)、3C102\_6(19s)

5104\_1(60min)、5104\_3(15s)

3C103(30min）、

3C303(30min)

GT-B106(30min)