**方法：**

**思路：**

采用级联分类器的方法，使用60s的滑动窗口数据训练一个随机森林分类器，使用该分类器实现对呼吸暂停低通气事件的粗筛，再使用训练好的20s窗口数据分类器实现对具体每一次事件的定位。之所以采用级联分类器的思想是因为直接采用短时间窗口单分类器的方法对睡眠呼吸暂停低通气事件的定位结果精准率较低，而60s窗口的分类器可以在保证正阳率的情况下筛除掉大部分阴性片段，进而保证最后的对于事件的分类结果的精准率与正阳率。

**特征提取：**

目前效果比较好的特征有三个：

1. 在短时间域上的血氧方差。因为伴随着呼吸暂停或者低通气事件的出现，人体血氧会发生明显的波动。
2. 在短时间域上的低通气次数。通过自适应的方式计算当前窗口之前30s内的流量信号峰值，计算低于峰值流量70%的呼吸次数作为特征。为了避免人体睡眠姿势改变等其他外界因素造成的对呼吸流量的干扰，在计算的过程中取得是之前30s时间内峰值流量的%95处的值。
3. 在短时间域上的血氧欠饱和次数。通过自适应的方式计算当前窗口之前30s内的血氧信号峰值，计算低于血氧峰值98%的时长作为特征。

以上1、3号特征均为从血氧信号中提取，通常会在呼吸暂停与低通气事件之后的一分钟内出现变化，存在一定的延迟，在实验的过程中为了抵消这一影响，采用主动对齐的方式将血氧信号主动提前τs的时长，但是仍然会存在不同的被试中τ不同的情况，所以1、3号特征会受到τ取不同值的影响，进而影响最后分类器的分类结果。

**训练方法：**

选择的分类方法为随机森林的方法，这里我测试了包括CART决策树、SVM、KNN、贝叶斯方法、随机森林方法等多种分类方法，最后从计算量与结果的可解释性角度出发选择的是随机森林的方法，随机森林的元分类器选择的是CART决策树，元分类器的个数设置为10个。

采用K折交叉训练的方式，将原始数据切割为K段，每次训练从中取出一段作为测试集，余下的K-1段作为训练集，使用随机森林的方法对训练集中切割得到的20s数据与60s数据进行训练，使用训练好的分类器在测试集上进行测试，使用正确率、正阳率、精准率对20s窗口数据的分类结果进行评价，同时使用精准率与正阳率对事件的分类结果进行评价。

在具体的训练过程中使用的是基于个体的训练方式，对于每一个被试数据训练出对应的分类器，选择数据库中的2、3、8、10、12、14、15、22号被试进行训练，得到的分类结果如下。

**结果：**

表格 1. 级联分类器对于20s窗口数据的分类结果

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 级联分类器1 | | | 级联分类器2 | | | 总分类器 | | |
| 被试 | 正确率 | 正阳率 | 精准率 | 正确率 | 正阳率 | 精准率 | 正确率 | 正阳率 | 精准率 |
| 1 | 0.796189 | 0.803825 | 0.611107 | 0.774097 | 0.843847 | 0.554259 | 0.904388 | 0.727093 | 0.584822 |
| 2 | 0.850008 | 0.818337 | 0.789181 | 0.764499 | 0.772672 | 0.571941 | 0.871776 | 0.779278 | 0.587673 |
| 3 | 0.852517 | 0.867047 | 0.724854 | 0.772489 | 0.905837 | 0.548008 | 0.91466 | 0.777488 | 0.594128 |
| 8 | 0.861482 | 0.894881 | 0.713406 | 0.74879 | 0.873576 | 0.531111 | 0.903657 | 0.712179 | 0.559562 |
| 10 | 0.808438 | 0.853384 | 0.78995 | 0.779244 | 0.865846 | 0.636504 | 0.874655 | 0.772241 | 0.677747 |
| 12 | 0.763587 | 0.751578 | 0.392124 | 0.708235 | 0.865309 | 0.281273 | 0.915878 | 0.7 | 0.33719 |
| 14 | 0.831548 | 0.872843 | 0.585804 | 0.645999 | 0.913612 | 0.345706 | 0.893266 | 0.797281 | 0.37139 |
| 15 | 0.824746 | 0.768783 | 0.508306 | 0.725621 | 0.833679 | 0.425054 | 0.928177 | 0.666667 | 0.453637 |
| 22 | 0.803763 | 0.74854 | 0.80961 | 0.753292 | 0.843476 | 0.62411 | 0.841623 | 0.573305 | 0.690729 |
| Ave | 0.821364 | 0.819913 | 0.65826 | 0.741363 | 0.857539 | 0.501996 | 0.893165 | 0.71529 | 0.539653 |

级联分类器2的分类结果会更偏向于阳性样本，这是因为经过级联分类器1筛选之后的片段主要包含的是睡眠呼吸暂停与低通气事件以及事件起始或者结束阶段在内的易混淆的片段。通过调整各级分类器的类权重最后的分类结果会有10个百分点左右的浮动。

表格 2. 级联分类器对于事件的分类结果

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 被试 | 精准率 | 召回率 | 错误事件 | 遗漏事件 | 总事件 | 人工标注AHI | 模型计算AHI |
| 1 | 0.657895 | 0.813008 | 52 | 23 | 123 | 20.22841 | 24.99772 |
| 2 | 0.860627 | 0.817881 | 40 | 55 | 302 | 42.53688 | 40.42412 |
| 3 | 0.79397 | 0.84492 | 41 | 29 | 187 | 29.84704 | 31.76236 |
| 8 | 0.756303 | 0.775862 | 29 | 26 | 116 | 18.06619 | 18.53342 |
| 10 | 0.746544 | 0.89011 | 55 | 20 | 182 | 29.19266 | 34.80663 |
| 12 | 0.446281 | 0.80597 | 67 | 13 | 67 | 10.47057 | 18.90953 |
| 14 | 0.520468 | 0.864078 | 82 | 14 | 103 | 14.8838 | 24.70999 |
| 15 | 0.654762 | 0.753425 | 29 | 18 | 73 | 11.98632 | 13.79247 |
| 22 | 0.831276 | 0.653722 | 41 | 107 | 309 | 43.96143 | 34.57161 |
| Ave | 0.726303 | 0.791382 | 436 | 305 | 1462 | 24.98422 | 27.22288 |

以上结果可能会随着随机森林的参数选择不同而产生变化，表现在上述评价指标上可能会有上下10个百分点以内的波动。整体来说事件的分类结果正阳率超过75%的水平同时精准率达到70%的水平，意味着该种方法可以检测出数据中存在的大部分睡眠呼吸暂停与低通气事件，但是同时会误报一部分错误事件。同时对于数据库中的不同被试，事件的分类结果不尽相同，其中个性化分类方法最好的是2号被试，对于12号被试结果则不太理想，这是因为对于不同的被试提取的以上特征对于呼吸暂停与低通气事件的表征作用还不尽相同，如下表所示。

表格 3. 不同特征的AUC面积

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 被试 | 特征1 | 特征2 | 特征3 | 被试 | 特征1 | 特征2 | 特征3 |
| 1 | 0.797352 | 0.842207 | 0.840074 | 13 | 0.814829 | 0.605702 | 0.680839 |
| 2 | 0.861465 | 0.861001 | 0.814548 | 14 | 0.863823 | 0.895681 | 0.793058 |
| 3 | 0.819802 | 0.907763 | 0.774735 | 15 | 0.857253 | 0.83803 | 0.8321 |
| 4 | 0.734671 | 0.802135 | 0.703591 | 16 | 0.801434 | 0.753877 | 0.760553 |
| 5 | 0.687132 | 0.73615 | 0.636067 | 17 | 0.693041 | 0.615396 | 0.579909 |
| 6 | 0.841909 | 0.683271 | 0.820077 | 18 | 0.801353 | 0.555823 | 0.796268 |
| 7 | 0.765061 | 0.717432 | 0.720017 | 19 | 0.73762 | 0.801508 | 0.725129 |
| 8 | 0.761391 | 0.877456 | 0.801917 | 20 | 0.603729 | 0.689446 | 0.6463 |
| 9 | 0.808188 | 0.782966 | 0.795831 | 21 | 0.782552 | 0.818095 | 0.749489 |
| 10 | 0.73668 | 0.854199 | 0.812303 | 22 | 0.893533 | 0.831511 | 0.857112 |
| 11 | 0.861675 | 0.692419 | 0.817558 | 23 | 0.743514 | 0.775025 | 0.749716 |
| 12 | 0.868199 | 0.892567 | 0.853143 |  |  |  |  |

对于不同的被试此处选择的τ值为统一的15s，但是对于下图所示的7号与19号被试来说，15s的τ值可能就不太合适，这也影响了1号与3号特征的表现。

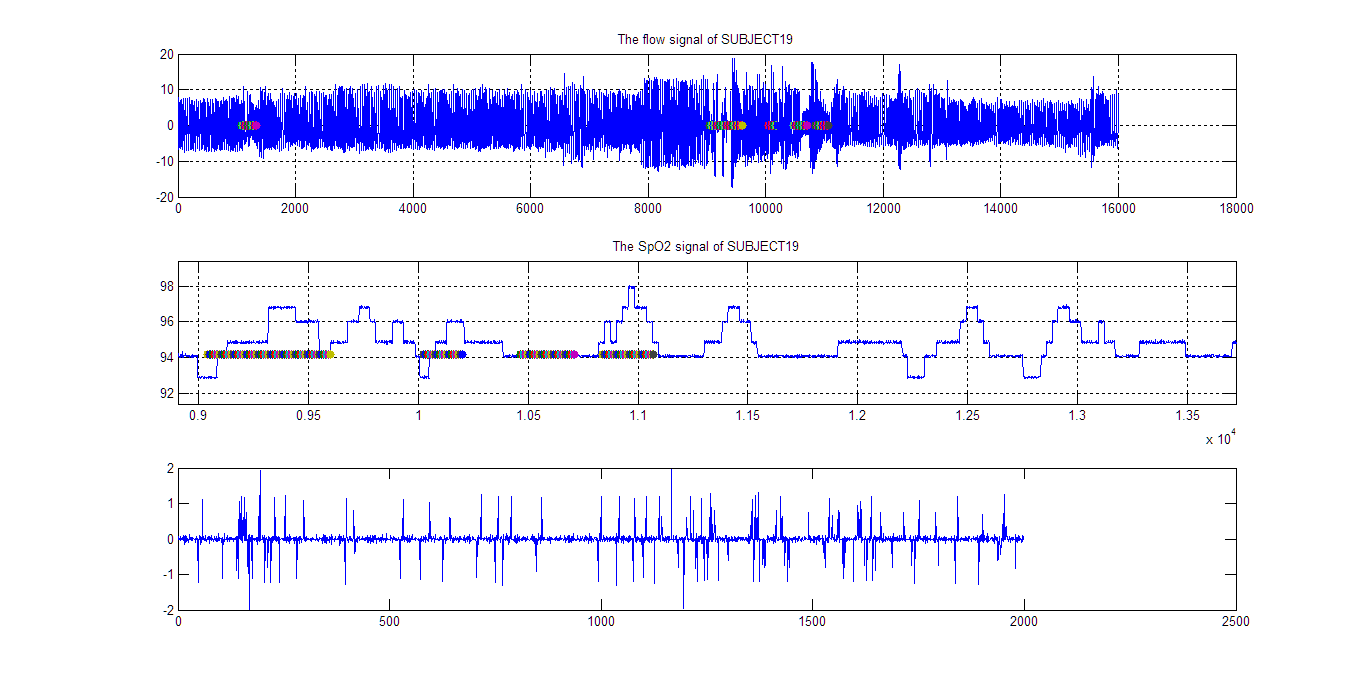
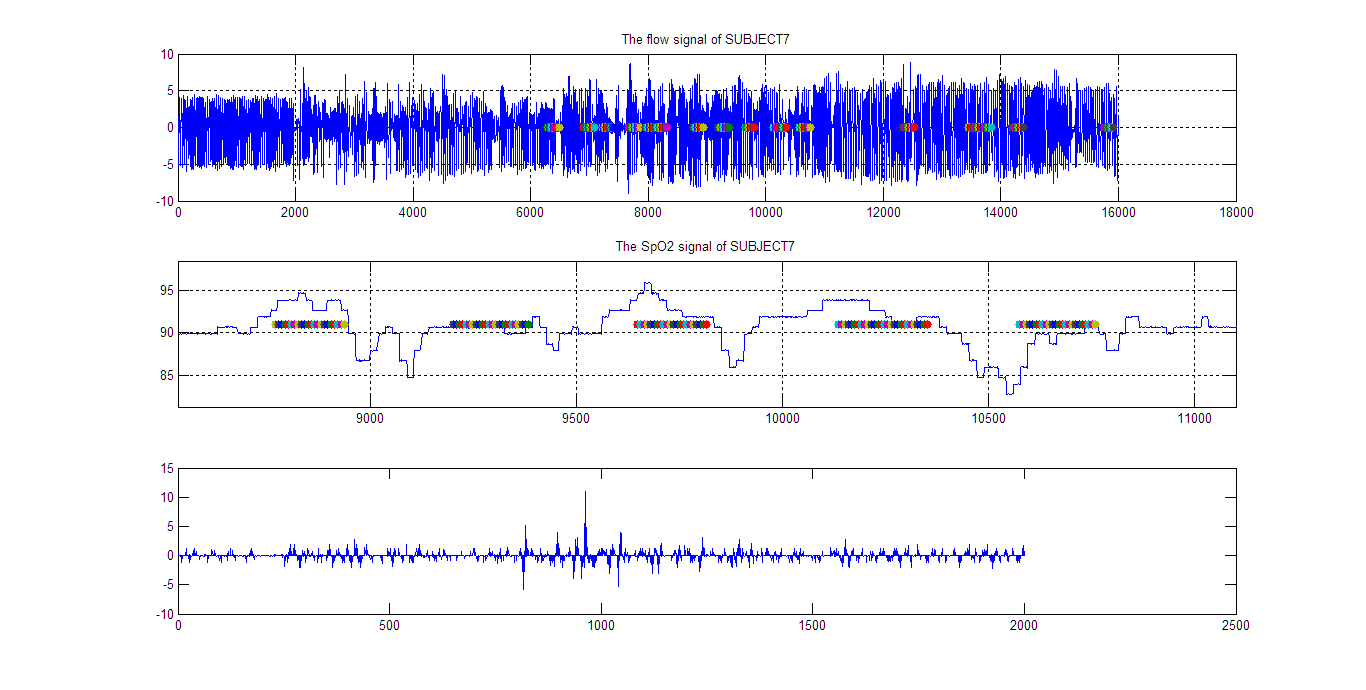


图 1. 分别为7号被试与19号被试的血氧信号片段，图中圆形标注的为呼吸暂停低通气事件。

**讨论：**

以上的结果均为基于个体化模型得到，时间分辨率为20s，也就是可以分辨出间隔在15s以上的睡眠呼吸暂停与低通气事件，事件的平均精准率与正阳率分别为0.683015，0.79661。目前研究存在的主要问题在于：

1. 正常状态与睡眠呼吸暂停与低通气状态的边界不是特别明显。以上研究的标签数据均为根据人工标签得到，如果当前片段中包含有超过5s钟的发病片段即认为该片段为阳性片段，而实际上根据这种方法得到的阴性样本与阳性样本在边界处区别并不明显，这也是造成级联分类器2的分类结果不好的主要原因。
2. 会存在某些即使肉眼观察也难以分辨的片段如下图所示。

b

a

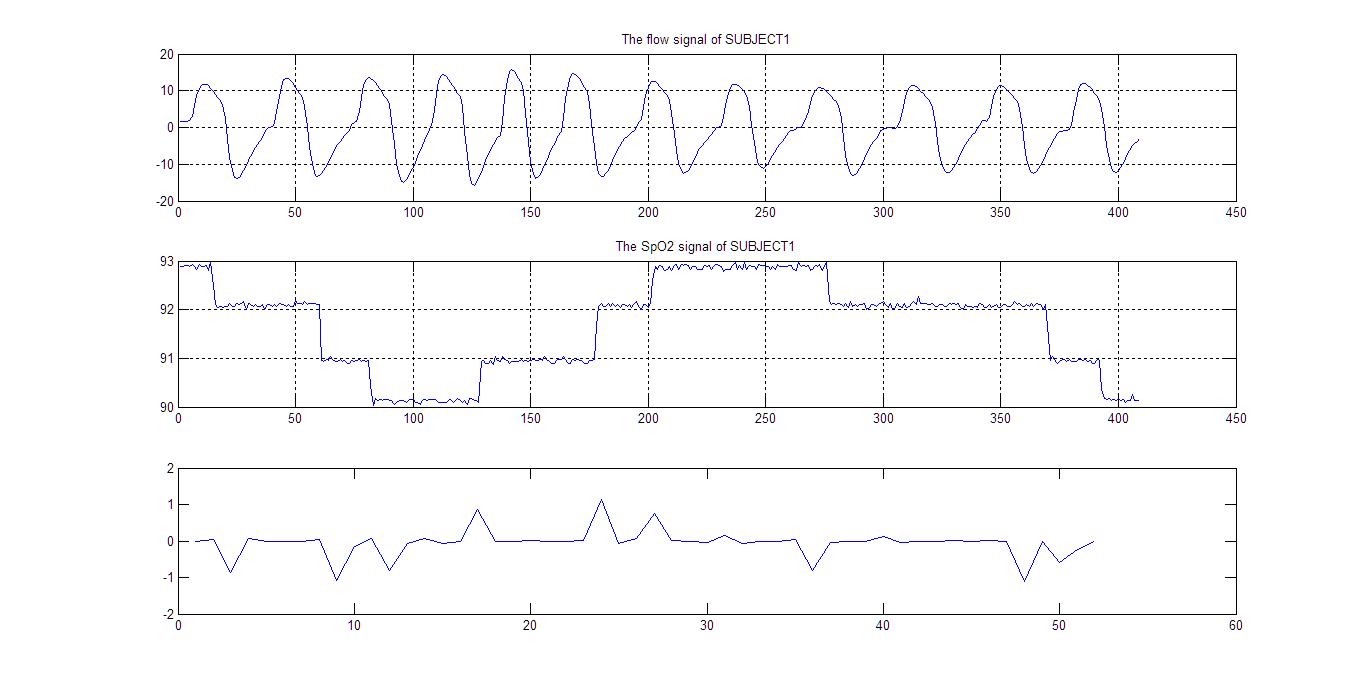
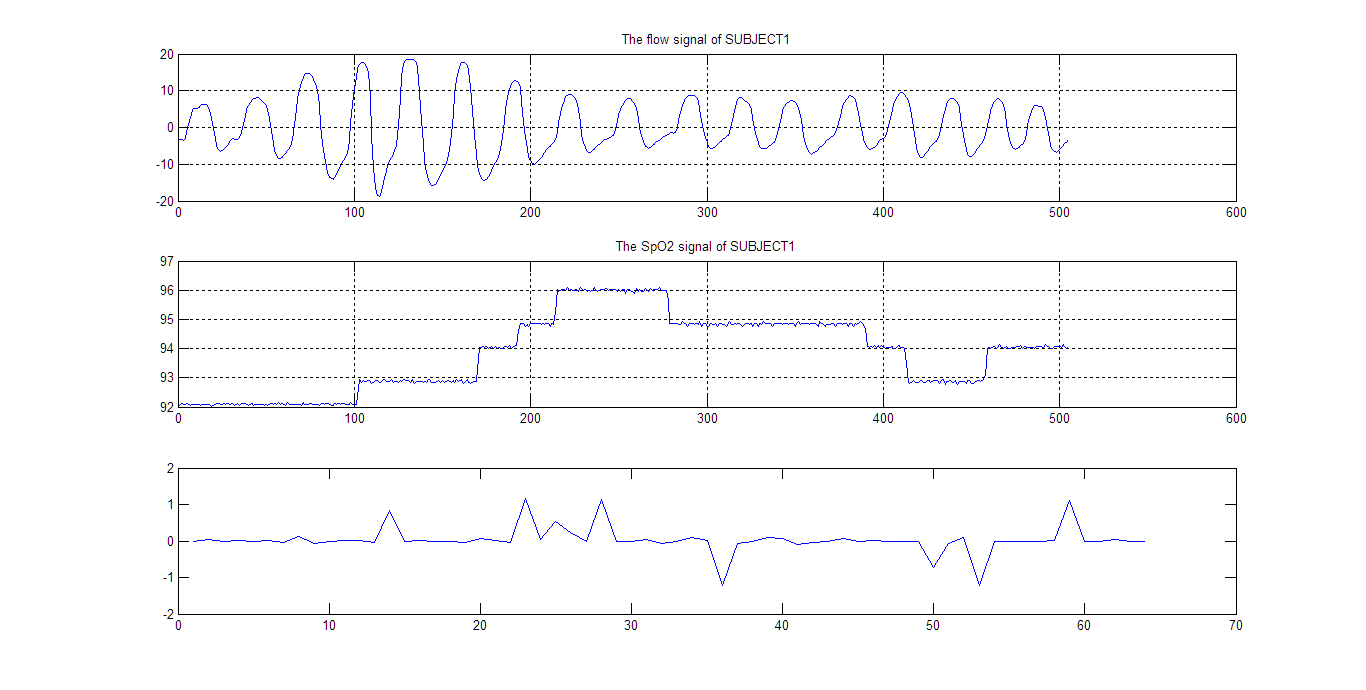


图 2. 分类结果中的假阳性片段，上为流量信号，下为血氧信号

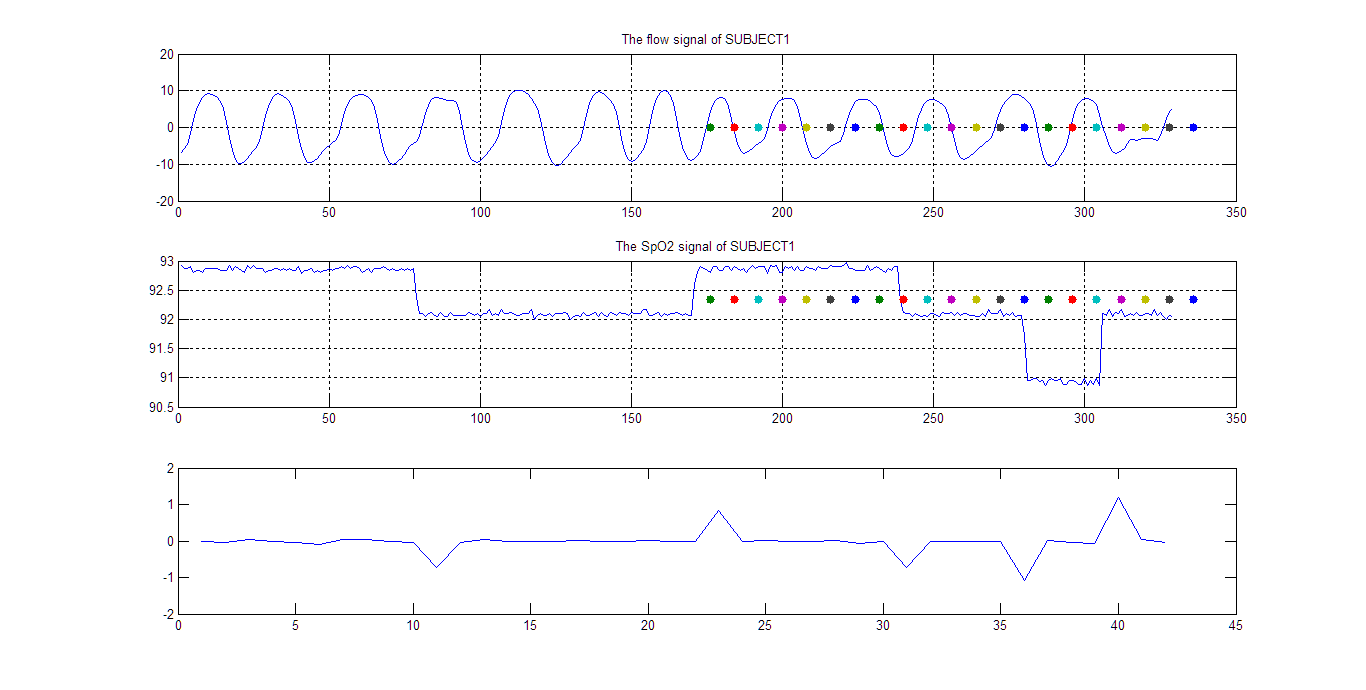
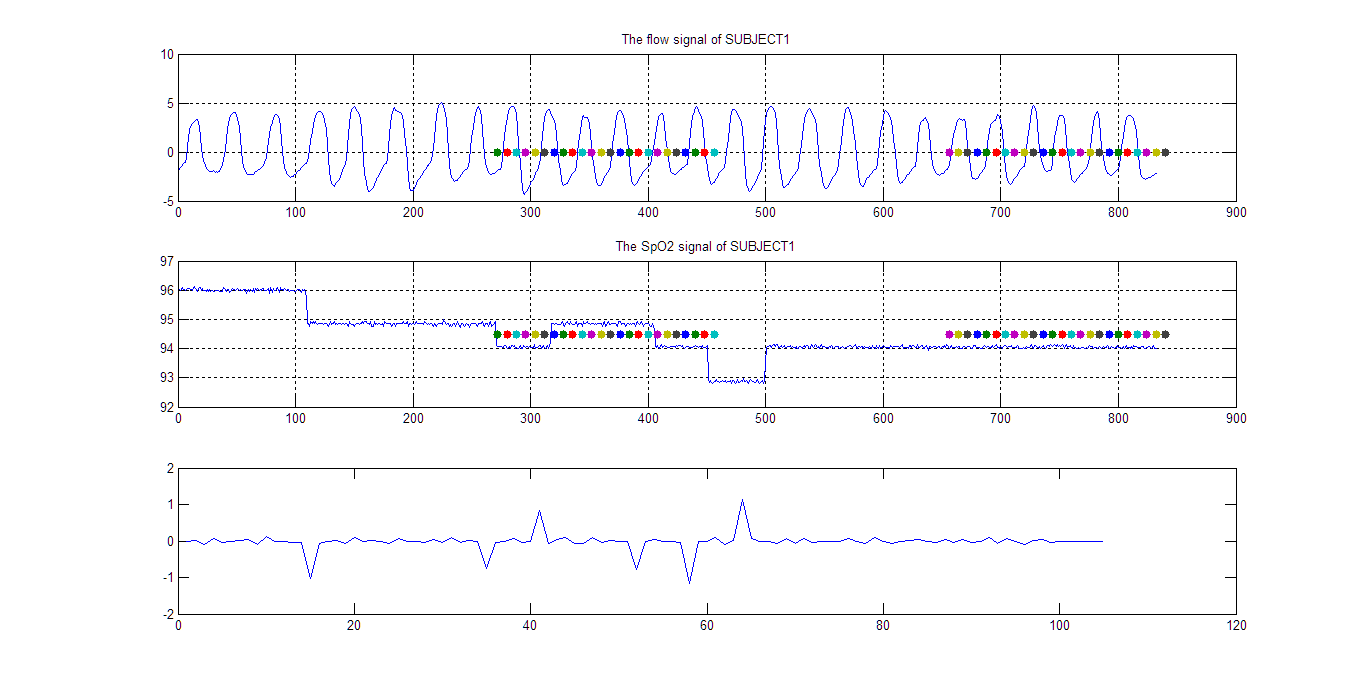
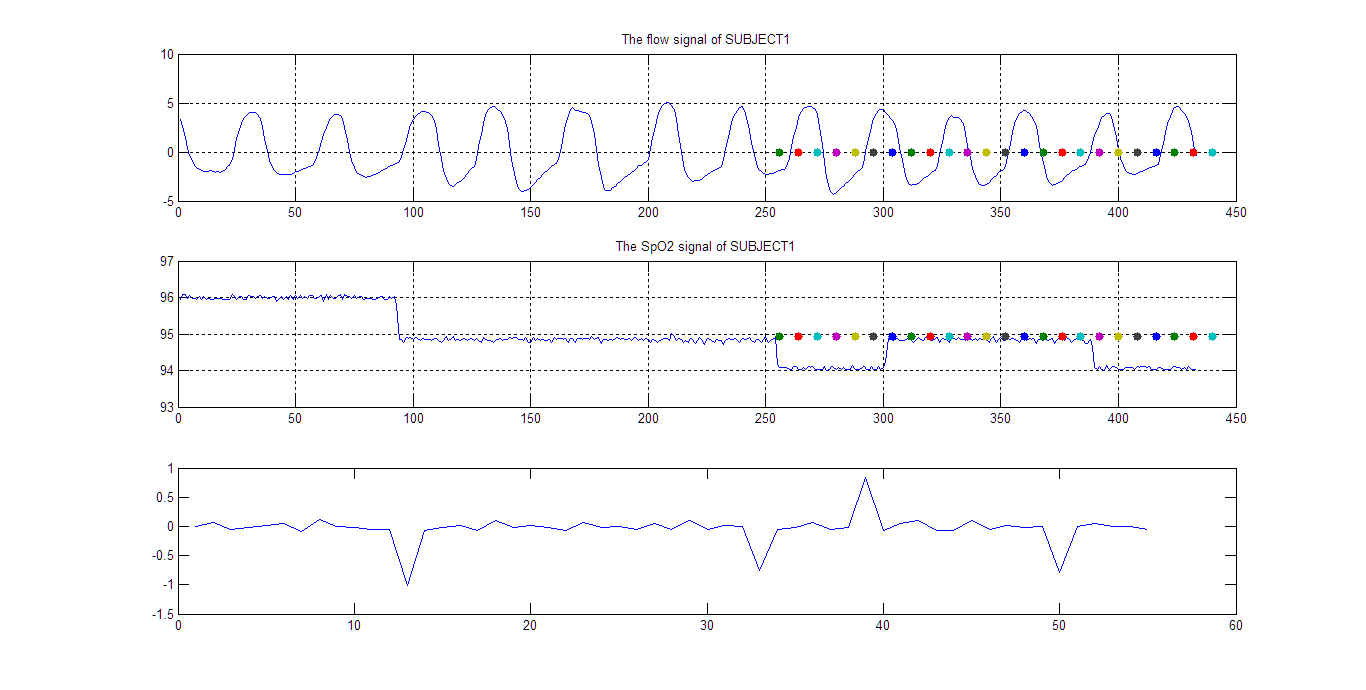
1. τ的选取对最后的分类结果会产生影响，改善τ的选取方法可能会改良最后的分类结果。

c

b

a

图 3. 分类结果中的假阴性片段，上为流量信号，下为血氧信号



|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 方法 | 片段分类结果 | | | 事件分类结果 | | 年份 |
| 卷积神经网络与10s滑动窗 | 96.6 | 81.1 | 98.5 | 98.1 | 89.1 | 2018 |
| 基于血氧形态学特征的一分钟片段检测 | 91.0 | 83.5 | 89.1 |  |  | 2018 |
| 基于口鼻压力形态学特征的检测 |  | 86.4 |  |  |  | 2016 |
| 基于血氧特征的检测 |  | 92.4 |  |  |  | 2015 |
|  |  |  |  |  |  |  |