**结果：**

表格 1. 级联分类器对于10s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 | 0.862908 | 0.773284 | 0.758625 | 0.784991 | 0.858159 | 0.594142 | 0.914745 | 0.758377 | 0.594142 |
| 2 | 0.848103 | 0.747794 | 0.806292 | 0.760824 | 0.813326 | 0.563718 | 0.868043 | 0.810641 | 0.563718 |
| 3 | 0.878884 | 0.820938 | 0.745737 | 0.777267 | 0.876083 | 0.51829 | 0.906469 | 0.806925 | 0.51829 |
| 8 | 0.819326 | 0.847757 | 0.600011 | 0.744952 | 0.893403 | 0.455888 | 0.900943 | 0.826977 | 0.455888 |
| 10 | 0.822977 | 0.809329 | 0.819481 | 0.787787 | 0.856682 | 0.650943 | 0.861547 | 0.799308 | 0.650943 |
| 12 | 0.779251 | 0.773454 | 0.404524 | 0.706032 | 0.834347 | 0.286159 | 0.882589 | 0.756422 | 0.286159 |
| 14 | 0.877333 | 0.847601 | 0.682111 | 0.657183 | 0.909793 | 0.362356 | 0.885145 | 0.846389 | 0.362356 |
| 15 | 0.842848 | 0.474151 | 0.341753 | 0.622448 | 0.736227 | 0.160697 | 0.904448 | 0.438849 | 0.160697 |
| 22 | 0.814662 | 0.738412 | 0.839044 | 0.703265 | 0.86814 | 0.599672 | 0.823589 | 0.771317 | 0.599672 |
| Ave | 0.838477 | 0.759191 | 0.666397 | 0.727194 | 0.849573 | 0.465763 | 0.880677 | 0.786178 | 0.465763 |

以上结果是两折交叉验证之后得到的结果，采用的模型训练方法为个体化模型训练方法，对于片段的分类结果，级联模型取得了88%的正确率，78%的正阳率与46%的精准率。而之前采用相同数据库的2012年的一篇文章最终实现的对于60s片段数据的预测结果正确率、正阳率、正阴率在82%左右[1]。

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 | AHI error |
| 1 | 0.756757 | 0.807692 | 27 | 20 | 104 | 21.3528 | 22.79001 | -1.43721 |
| 2 | 0.828358 | 0.863813 | 46 | 35 | 257 | 40.46182 | 42.19365 | -1.73183 |
| 3 | 0.763514 | 0.843284 | 35 | 21 | 134 | 25.65139 | 28.33138 | -2.68 |
| 8 | 0.609195 | 0.80303 | 34 | 13 | 66 | 14.35129 | 18.91761 | -4.56632 |
| 10 | 0.69 | 0.890323 | 62 | 17 | 155 | 27.50937 | 35.49596 | -7.98659 |
| 12 | 0.461538 | 0.80597 | 63 | 13 | 67 | 10.98861 | 19.18907 | -8.20046 |
| 14 | 0.581699 | 0.89899 | 64 | 10 | 99 | 14.78225 | 22.84529 | -8.06304 |
| 15 | 0.293103 | 0.515152 | 41 | 16 | 33 | 7.042087 | 12.377 | -5.33491 |
| 22 | 0.807273 | 0.810219 | 53 | 52 | 274 | 43.60357 | 43.76271 | -0.15914 |
| Ave | 0.700071 | 0.834315 | 425 | 197 | 1189 | 23.58113 | 28.103 | -4.52187 |

级联模型对于具体睡眠呼吸暂停与低通气事件的预测结果取得了70%的精准率与83%的召回率。AHIrefer与AHIest平均相差4.52events/h。说明该种方法能够检测并且定位出大部分的睡眠呼吸暂停与低通气事件，但是同时伴随着一定概率的虚警。以往的文章并没有对于睡眠呼吸暂停与低通气事件的统计结果，一般直接呈现的是对被试的睡眠呼吸暂停综合征严重程度的预测结果，2018年的一片文章采用卷积神经网络的方法实现了对正常、轻微、中等、严重四种严重程度69.8%、67.2%、76%、84.5%的正阳率[2]，但是在这里由于被试数目过少，并没与对该项指标进行计算。

图 1. AHIrefer与AHIest散点关系图

拟合直线的斜率为0. 8389，截距为8.1446。通过上图可以看出级联模型预测出的AHI会高于人工标注计算出的AHI，说明现在级联模型的主要问题在于预测的精准率不足，或者说虚警的情况较严重。



图 2. AHIrefer与AHIest的Bland-Altman图

AHIrefer与AHIest的平均差之为-2.2311events/h，95%的置信区间内AHIrefer与AHIest相差-5.36到0.90events/h。说明该种检测方法具有良好的泛化能力，对于不同的被试的预测结果表现出了良好的一致性。

**方法：**

**预处理：**

1. **去除无效数据**

原始数据中某些时段存在由于患者改变睡眠姿势造成血氧传感器接触不良造成的数据丢失的情况，主要有以下几种情况

1),标签中标注为Cheyne-Stokes、Periodic breathing事件的数据，研究中只针对睡眠呼吸暂停与低通气事件，所以将该部分数据也去掉。

2),血氧低于80%的数据，正常人血氧水平在95-99的水平，低于80的血氧数据可能是因为患者与传感器接触不良造成的，如图3所示。原始数据中大概存在10小时的血氧数据低于80%需要去除。

图 3. 血氧信息丢失片段

3),流量数据高于20的数据，根据数据库中的数据观察，流量信号在高于20的水平的时候出现失真，存在信息丢失的情况，如图4所示，在预处理过程中同样将该部分数据去掉。

图 4. 流量信息丢失片段

1. **切割窗口与主动对齐**

将原始整夜的数据采用滑动窗的方式切割成固定长度的数据片段，根据原始人工标注数据对切割后的数据片段添加标签，如果片段内包含有五秒以上的呼吸暂停与低通气事件则标注为阳性片段，反之标注为阴性片段。

由于血氧数据本身具有延迟效应，通常在呼吸暂停与低通气事件发生之后20-40s的时间内才会发生下降，如果采用比较长的窗口进行切割可以捕捉到呼吸暂停与低通气事件引起的血氧欠饱和的特征，但是采用比较长的窗口的话不能对距离很近的两次事件进行分辨，如果进一步缩短窗口长度的话又不能够保证抓住血氧欠饱和的这一特征，所以在使用短窗口对数据进行滑动切割的时候实现对血氧数据进行了主动对齐。

在实际处理的时候分别采用一个60s的窗口以及一个10s的窗口对原始数据进行切割，60s窗口切割的时候无需主动对齐，而使用10s窗口切割的时候τ设置为了23s进行主动对齐。

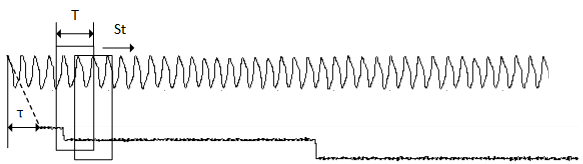


图 5. 主动对齐、滑动窗切割

**特征提取：**

流量通道的特征提取都是基于对流量呼气相与吸气相的检测基础上进行的，首先通过流量通道信号的导数信息提取得到吸气相与呼气相的起止点，如下图所示。

图 6. 呼吸信号和呼气相、吸气相起点

经过上述处理之后的流量片段数据可以用下面的公式表示：

(1)

睡眠呼吸暂停事件的定义是信号峰值比上一次事件的基线下降超过90%并且持续超过10s钟，呼吸低通气事件的定义是信号峰值比上一次事件的基线下降超过30%并且持续超过10s钟并且伴随着血氧下降超过3%或者觉醒。根据上述定义，流量基线与血氧基线计算公式如下，表示第个片段数据：

(1)

(2)

呼吸低通气阈值计算公式如下：

(3)

呼吸暂停阈值计算公式如下:

(4)

正常呼吸阈值计算公式如下：

(5)

血氧欠饱和阈值计算公式如下：

(6)

第k个60s窗口数据特征集

1. 血氧低于91的时长

当前数据片段内血氧低于91%的总的数据长度

1. 呼吸暂停次数

当前数据片段内低于的总的呼吸次数

1. 呼吸暂停与低通气次数

当前数据片段内低于的总的呼吸次数

1. 呼吸暂停与低通气次数所占比重

当前数据片段内低于的呼吸次数与总的呼吸次数的比值

1. 血氧欠饱和时长

当前数据片段内低于的总的数据长度

1. 血氧下降水平
2. 血氧极差
3. 血氧标准差

当前数据片段的血氧标准差

第k个10s窗口数据特征集

1. 呼吸暂停次数
2. 平均通气量
3. 正常呼吸次数所占比例

当前数据片段内高于的呼吸次数与总的呼吸次数的比值

1. 呼吸暂停与低通气次数
2. 血氧欠饱和时长
3. 呼吸暂停与低通气次数所占比重
4. 血氧下降水平

**训练与预测：**

1. **训练**

实时的个体化预测模型对分类器的训练时间与预测事件都有较高的要求，而之前研究中选用的诸如SVM、卷积神经网络的分类方法一方面计算复杂度较高，另外一方面分类器的超参数不好确定，所以这里选择的是随机森林的方法。

在训练的过程中分别采用相同一段事件切割得到的60s数据片段与10s数据片段对分类器进行训练，最后得到两个随机森林分类器。在这里随机森林的元分类器选用的是CART决策树，不同的是针对60s数据片段的随机森林包含10棵决策树，针对10s数据片段的随机森林包含20棵决策树。

1. **预测**

在这里采用两折交叉训练的方式，在训练的过程中使用一半数据训练出两个随机森林分类器，剩下一半的数据作为测试集进行预测，在预测的过程中首先使用针对60s数据片段的随机森林分类器输出60s数据片段的预测结果，在根据上述结果找出对应的10s数据片段作为第二个随机森林分类器的输入，输出最后的10s数据片段的预测结果。具体的预测过程如图7所示。

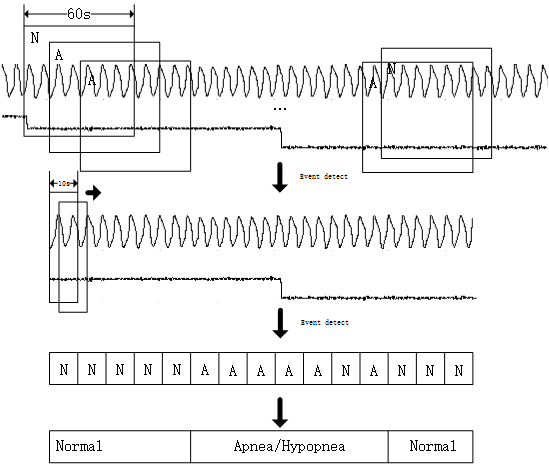


图 7. 预测过程

1. **事件检测**

需要注意的一点是每一个分类器输出预测结果以后都会经过事件检测处理，一方面是为了检测出当前预测结果中的具体每一次睡眠呼吸暂停与低通气事件的起点与终点，另外一方面是为了去除预测结果中不应该存在的情况。不应该存在的情况包括以下两种:

1. 预测结果中事件持续小于10s的情况。

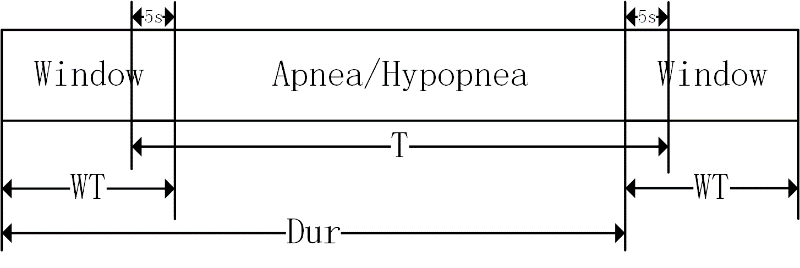
根据睡眠呼吸暂停与低通气事件的定义，一次事件的最短持续时长为10s，这对于预测结果中的单次事件的持续长度Dur产生了一定的约束条件，如图8所示。

图 8. 预测结果的单次事件持续长度Dur

根据图8可以计算出预测结果中单次事件的持续长度可以由以下公式计算得出

公式中T表示单次睡眠呼吸暂停与低通气事件的持续时长，所以可以知道预测结果中的单次事件持续长度*Dur*>*WT*，在事件检测器中会将出现的该种情况对应的阳性片段重置为阴性。

1. 预测结果中两次事件间隔小于5s的情况。

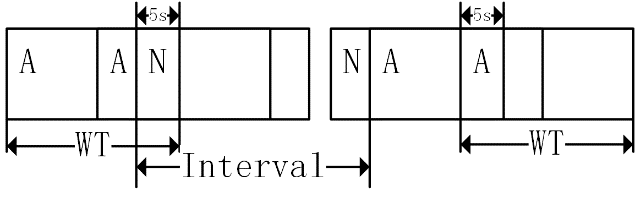
预测结果中两次相邻事件的间隔interval同样会存在一定的约束，如图9所示。

图 9. 预测结果的两次相邻事件的间隔interval

根据图9容易计算得到*interval*>5，所以在事件检测器中对于该种情况的阴性片段重置为阳性片段。

经过级联分类器与事件检测器我们可以得到事件的总次数与每次事件的起点与终点，在评价过程中如果预测出的事件与人工标注中的事件存在重合则认为这次事件预测正确，这样认为的原因在于我们关心的重点是该种方法能不能定位到具体某次事件的发生位置，预测结果中定位到的事件能不能涵盖人工标注中的事件，所以具体到某一次事件的起止点的微小偏移认为是可以忽略的。

1. **AHI预测**

在上述对于睡眠呼吸暂停与低通气事件的预测结果基础之上，我们可以计算出患者的AHI指数，进而对患者睡眠呼吸暂停综合征严重程度的判断。

[1] B. Xie and H. Minn, "Real-Time Sleep Apnea Detection by Classifier Combination," *IEEE Transactions on Information Technology in Biomedicine,* vol. 16, no. 3, pp. 469-477, 2012.

[2] S. H. Choi *et al.*, "Real-time apnea-hypopnea event detection during sleep by convolutional neural networks," *Computers in Biology and Medicine,* vol. 100, pp. 123-131, 2018/09/01/ 2018.