



(12)发明专利申请

(10)申请公布号 CN 110141226 A

(43)申请公布日 2019.08.20

(21)申请号 201910455076.6

A61B 5/00(2006.01)

(22)申请日 2019.05.29

(71)申请人 清华大学深圳研究生院

地址 518055 广东省深圳市南山区西丽街道深圳大学城清华校区A栋二楼

申请人 东莞见达信息技术有限公司

(72)发明人 王兴军 徐子尧 邹庆言 冯发润
黄海平 李庆

(74)专利代理机构 北京三友知识产权代理有限公司 11127

代理人 周晓飞 王天尧

(51)Int.Cl.

A61B 5/0476(2006.01)

A61B 5/0496(2006.01)

A61B 5/0488(2006.01)

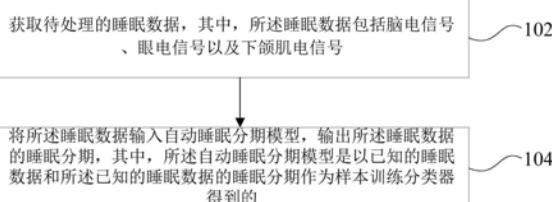
权利要求书3页 说明书9页 附图3页

(54)发明名称

自动睡眠分期方法、装置、计算机设备及计算机存储介质

(57)摘要

本发明实施例提供了一种自动睡眠分期方法、装置、计算机设备及计算机可读存储介质，其中，该方法包括：获取待处理的睡眠数据，其中，所述睡眠数据包括脑电信号、眼电信号以及下颌肌电信号；将所述睡眠数据输入自动睡眠分期模型，输出所述睡眠数据的睡眠分期，其中，所述自动睡眠分期模型是以已知的睡眠数据和所述已知的睡眠数据的睡眠分期作为样本训练分类器得到的。该方案的自动睡眠分期过程由于采用了脑电信号、眼电信号以及下颌肌电信号，与现有技术中只根据单通道脑电进行自动睡眠分期的方式相比，可以有利于提高自动睡眠分期的准确率。



1. 一种自动睡眠分期方法,其特征在于,包括:

获取待处理的睡眠数据,其中,所述睡眠数据包括脑电信号、眼电信号以及下颌肌电信号;

将所述睡眠数据输入自动睡眠分期模型,输出所述睡眠数据的睡眠分期,其中,所述自动睡眠分期模型是以已知的睡眠数据和所述已知的睡眠数据的睡眠分期作为样本训练分类器得到的。

2. 如权利要求1所述的自动睡眠分期方法,其特征在于,训练分类器得到所述自动睡眠分期模型,包括:

采用所有样本训练分类器,得到初步的自动睡眠分期模型;

针对每个样本,将该样本的睡眠数据输入所述初步的自动睡眠分期模型,将所述初步的自动睡眠分期模型中间层的输出作为该样本的高层表示样本,得到每个样本的高层表示样本;

针对每类睡眠分期,对属于该类睡眠分期的样本的高层表示样本进行密度聚类,得到属于该类睡眠分期的多个高层表示子集;得到属于每类睡眠分期的高层表示子集;

采用属于每类睡眠分期的高层表示子集训练分类器,得到最终的所述自动睡眠分期模型。

3. 如权利要求2所述的自动睡眠分期方法,其特征在于,对属于该类睡眠分期的样本的高层表示样本进行密度聚类,得到属于该类睡眠分期的多个高层表示子集,包括:

针对属于该类睡眠分期的各样本的高层表示样本,对维度大于预设维度阈值的高层表示样本进行降维,降维至维度等于所述预设维度阈值;

针对属于该类睡眠分期的所有高层表示样本,计算每两个高层表示样本之间的距离;

针对每个高层表示样本,计算与该高层表示样本的距离小于预设距离阈值的高层表示样本的个数,将该个数作为该高层表示样本的局部密度,将局部密度最大的高层表示样本作为该类睡眠分期的聚类中心;

针对属于该类睡眠分期的所有高层表示样本,按照与聚类中心的距离大小顺序,将属于该类睡眠分期的所有高层表示样本依次分成多个高层表示子集。

4. 如权利要求2所述的自动睡眠分期方法,其特征在于,还包括:

在采用属于每类睡眠分期的高层表示子集训练分类器之前,针对各类睡眠分期的相同序数的高层表示子集,在各类睡眠分期的相同序数的高层表示子集包括的高层表示样本数目不平衡的情况下,进行平衡处理,直至各类睡眠分期的相同序数的高层表示子集包括的高层表示样本数目相同。

5. 如权利要求2所述的自动睡眠分期方法,其特征在于,所述初步的自动睡眠分期模型中间层的输出包括所述初步的自动睡眠分期模型的一个或多个隐藏层的输出。

6. 如权利要求3所述的自动睡眠分期方法,其特征在于,采用属于每类睡眠分期的高层表示子集训练分类器,包括:

针对属于每类睡眠分期的高层表示子集,按照与聚类中心的距离由小到大的顺序,采用属于各类睡眠分期的各高层表示子集依次训练并微调分类器。

7. 如权利要求1至6中任一项所述的自动睡眠分期方法,其特征在于,还包括:

在以已知的睡眠数据和所述已知的睡眠数据的睡眠分期作为样本训练分类器之前,将

所述样本中所述已知的睡眠数据去除噪声、去除伪迹以及变换数据采样频率为同一固定频率；

对去除噪声、去除伪迹以及变换数据采样频率后的所述已知的睡眠数据进行归一化处理。

8. 一种自动睡眠分期装置，其特征在于，包括：

数据获取模块，用于获取待处理的睡眠数据，其中，所述睡眠数据包括脑电信号、眼电信号以及下颌肌电信号；

自动睡眠分期模块，用于将所述睡眠数据输入自动睡眠分期模型，输出所述睡眠数据的睡眠分期，其中，所述自动睡眠分期模型是以已知的睡眠数据和所述已知的睡眠数据的睡眠分期作为样本训练分类器得到的。

9. 如权利要求8所述的自动睡眠分期装置，其特征在于，所述自动睡眠分期模块，包括：

第一训练单元，用于采用所有样本训练分类器，得到初步的自动睡眠分期模型；

高层表示获取单元，用于针对每个样本，将该样本的睡眠数据输入所述初步的自动睡眠分期模型，将所述初步的自动睡眠分期模型中间层的输出作为该样本的高层表示样本，得到每个样本的高层表示样本；

密度聚类单元，用于针对每类睡眠分期，对属于该类睡眠分期的样本的高层表示样本进行密度聚类，得到属于该类睡眠分期的多个高层表示子集；得到属于每类睡眠分期的高层表示子集；

第二训练单元，用于采用属于每类睡眠分期的高层表示子集训练分类器，得到最终的所述自动睡眠分期模型。

10. 如权利要求9所述的自动睡眠分期装置，其特征在于，所述密度聚类单元，包括：

降维子单元，用于针对属于该类睡眠分期的各样的高层表示样本，对维度大于预设维度阈值的高层表示样本进行降维，降维至维度等于所述预设维度阈值；

距离计算子单元，用于针对属于该类睡眠分期的所有高层表示样本，计算每两个高层表示样本之间的距离；

局部密度确定子单元，用于针对每个高层表示样本，计算与该高层表示样本的距离小于预设距离阈值的高层表示样本的个数，将该个数作为该高层表示样本的局部密度，将局部密度最大的高层表示样本作为该类睡眠分期的聚类中心；

密度聚类子单元，用于按照与聚类中心的距离大小顺序，将属于该类睡眠分期的所有高层表示样本依次分成多个高层表示子集。

11. 如权利要求9所述的自动睡眠分期装置，其特征在于，还包括：

平衡模块，用于在采用属于每类睡眠分期的高层表示子集训练分类器之前，针对各类睡眠分期的相同序数的高层表示子集，在各类睡眠分期的相同序数的高层表示子集包括的高层表示样本数目不平衡的情况下，进行平衡处理，直至各类睡眠分期的相同序数的高层表示子集包括的高层表示样本数目相同。

12. 如权利要求9所述的自动睡眠分期装置，其特征在于，所述初步的自动睡眠分期模型中间层的输出包括所述初步的自动睡眠分期模型的一个或多个隐藏层的输出。

13. 如权利要求10所述的自动睡眠分期装置，其特征在于，所述第二训练单元具体用于，针对属于每类睡眠分期的高层表示子集，按照与聚类中心的距离由小到大的顺序，采用

属于各类睡眠分期的各高层表示子集依次训练并微调分类器。

14. 如权利要求8至13中任一项所述的自动睡眠分期装置，其特征在于，还包括：

预处理模块，用于在以已知的睡眠数据和所述已知的睡眠数据的睡眠分期作为样本训练分类器之前，将所述样本中所述已知的睡眠数据将所述样本中所述已知的睡眠数据去除噪声、去除伪迹以及变换数据采样频率为同一固定频率；对去除噪声、去除伪迹以及变换数据采样频率后的所述已知的睡眠数据进行归一化处理。

15. 一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，其特征在于，所述处理器执行所述计算机程序时实现权利要求1至7任一项所述的自动睡眠分期方法。

16. 一种计算机可读存储介质，其特征在于，所述计算机可读存储介质存储有执行权利要求1至7任一项所述的自动睡眠分期方法的计算机程序。

自动睡眠分期方法、装置、计算机设备及计算机存储介质

技术领域

[0001] 本发明涉及数据处理技术领域,特别涉及一种自动睡眠分期方法、装置、计算机设备及计算机可读存储介质。

背景技术

[0002] 睡眠分期是按照一定的标准对人在睡眠过程中所经历的阶段进行划分。研究表明健康者在睡眠过程中会经历若干个睡眠阶段周期性循环的过程,而存在睡眠问题的患者这种周期性的睡眠结构不够明显或者比较混乱。例如,失眠症患者的睡眠障碍主要体现在入睡过程较为困难,即从清醒阶段到睡眠阶段的时间过长。相反,嗜睡症患者则极易入睡,从清醒阶段到睡眠阶段的时间极短。又如,睡眠呼吸障碍疾病是重大源头疾病,与多种疾病具有高共患率,我国未诊人群及高危人群数量极大。准确的睡眠分期监测能够为睡眠问题的诊断提供有效的辅助工具,进而采取相应的措施进行调理和治疗。

[0003] 目前,睡眠监测的常见方式是先对患者实施整晚的PSG(多导睡眠图)睡眠参数监测,次日由技师对整晚数据曲线图进行人工分图和事件分析,再由医生根据事件进行分析诊断。除监控视频数据外,每个患者的整晚PSG监测一般在200MB以上,有经验的医师每天完成3例左右患者数据的完整分析和诊断报告。虽然基于规则的计算机辅助分图有了长足进展,但受准确性限制,手工分析仍是目前国内外普遍认可的方式。而受技师差异的影响,手工分图还存在高达5%~20%的变异率。随着睡眠监测的患者数量的增多,有限的睡眠分图技师越来越难以满足日益增长的睡眠分图需求。因此,自动睡眠分期越来越引起研究人员的兴趣。

[0004] 睡眠分期是PSG分图的基础。对患者整晚的睡眠监测记录进行睡眠分期,不仅可以为分析患者的睡眠结构提供依据,还是进一步分析患者是否发生微觉醒事件的参考依据,而睡眠分期和微觉醒事件,又是进一步分析患者是否发生呼吸暂停、低通气等呼吸事件的参考依据。自动睡眠分期需要从每帧PSG信号中提取特征,再根据特征进行分类。传统地,提取的特征包括时频特征和非线性动力学特征,时频特征的提取方法包括小波变换、经验模式分解等,非线性动力学特征包括多尺度熵、Lyapunov指数等。根据手工设计的特征,使用传统的机器学习模型,例如线性判别分析、随机森林、支持向量机、AdaBoosting,进行自动睡眠分期。基于特征工程的自动睡眠分期模型依赖于人工设计的特征,然而,每种人工设计的特征都存在其局限性,我们很难知道应该提取哪些特征。

发明内容

[0005] 本发明实施例提供了一种自动睡眠分期方法,以解决现有技术中由于只根据单通道脑电进行自动睡眠分期而存在的睡眠分期准确率低的技术问题。该方法包括:

[0006] 获取待处理的睡眠数据,其中,所述睡眠数据包括脑电信号、眼电信号以及下颌肌电信号;

[0007] 将所述睡眠数据输入自动睡眠分期模型,输出所述睡眠数据的睡眠分期,其中,所

述自动睡眠分期模型是以已知的睡眠数据和所述已知的睡眠数据的睡眠分期作为样本训练分类器得到的。

[0008] 本发明实施例还提供了一种自动睡眠分期装置，以解决现有技术中由于只根据单通道脑电进行自动睡眠分期而存在的睡眠分期准确率低的技术问题。该装置包括：

[0009] 数据获取模块，用于获取待处理的睡眠数据，其中，所述睡眠数据包括脑电信号、眼电信号以及下颌肌电信号；

[0010] 睡眠分期模块，用于将所述睡眠数据输入自动睡眠分期模型，输出所述睡眠数据的睡眠分期，其中，所述自动睡眠分期模型是以已知的睡眠数据和所述已知的睡眠数据的睡眠分期作为样本训练分类器得到的。

[0011] 本发明实施例还提供了一种计算机设备，包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序，所述处理器执行所述计算机程序时实现上述任意的自动睡眠分期方法，以解决现有技术中由于只根据单通道脑电进行自动睡眠分期而存在的睡眠分期准确率低的技术问题。

[0012] 本发明实施例还提供了一种计算机可读存储介质，所述计算机可读存储介质存储有执行上述任意的自动睡眠分期方法的计算机程序，以解决现有技术中由于只根据单通道脑电进行自动睡眠分期而存在的睡眠分期准确率低的技术问题。

[0013] 在本发明实施例中，提出了采用训练的自动睡眠分期模型来实现基于综合多个通道的信息(例如，包括脑电信号、眼电信号以及下颌肌电信号等多个通道的信息)的自动睡眠分期，该自动睡眠分期过程由于采用了脑电信号、眼电信号以及下颌肌电信号，与现有技术中只根据单通道脑电进行自动睡眠分期的方式相比，可以有利于提高自动睡眠分期的准确率。

附图说明

[0014] 此处所说明的附图用来提供对本发明的进一步理解，构成本申请的一部分，并不构成对本发明的限定。在附图中：

[0015] 图1是本发明实施例提供的一种自动睡眠分期方法的流程图；

[0016] 图2是本发明实施例提供的一种具体实施上述自动睡眠分期方法的流程图；

[0017] 图3是本发明实施例提供的一种计算机设备的结构框图；

[0018] 图4是本发明实施例提供的一种自动睡眠分期装置的结构框图。

具体实施方式

[0019] 为使本发明的目的、技术方案和优点更加清楚明白，下面结合实施方式和附图，对本发明做进一步详细说明。在此，本发明的示意性实施方式及其说明用于解释本发明，但并不作为对本发明的限定。

[0020] 在本发明实施例中，提供了一种自动睡眠分期方法，如图1所示，该方法包括：

[0021] 步骤102：获取待处理的睡眠数据，其中，所述睡眠数据包括脑电信号、眼电信号以及下颌肌电信号；

[0022] 步骤104：将所述睡眠数据输入自动睡眠分期模型，输出所述睡眠数据的睡眠分期，其中，所述自动睡眠分期模型是以已知的睡眠数据和所述已知的睡眠数据的睡眠分期

作为样本训练分类器得到的。

[0023] 由图1所示的流程可知,在本发明实施例中,提出了采用训练的自动睡眠分期模型来实现基于综合多个通道的信息(例如,包括脑电信号、眼电信号以及下颌肌电信号等多个通道的信息)的自动睡眠分期,该自动睡眠分期过程由于采用了脑电信号、眼电信号以及下颌肌电信号,与现有技术中只根据单通道脑电进行自动睡眠分期的方式相比,可以有利于自动提高睡眠分期的准确率。

[0024] 具体实施时,上述睡眠数据包括的脑电信号、眼电信号以及下颌肌电信号可以从PSG(多导睡眠图)信号中提取得到。

[0025] 具体实施时,采用已知的睡眠数据和已知的睡眠数据的睡眠分期作为样本训练分类器,其中,已知的睡眠数据也包括脑电信号、眼电信号以及下颌肌电信号,已知的睡眠数据也可以从已知的PSG信号中提取,已知的睡眠数据的睡眠分期可以是由医师或其他分析方式确定的,本申请对此不做限定。

[0026] 具体实施时,上述分类器可以包括但不限于朴素贝叶斯、决策树、逻辑回归、支持向量机、神经网络(例如,任意形式的卷积神经网络和循环神经网络)等。

[0027] 具体实施时,为了提高自动睡眠分期模型的准确率,在本实施例中,通过以下方式训练分类器得到所述自动睡眠分期模型:

[0028] 采用所有样本训练分类器,得到初步的自动睡眠分期模型;具体的,在训练分类器的过程中,可以将所有样本划分为训练集、验证集以及测试集,通过训练一个神经网络,得到初步的自动睡眠分期模型。

[0029] 针对每个样本,将该样本的睡眠数据输入所述初步的自动睡眠分期模型,将所述初步的自动睡眠分期模型中间层的输出作为该样本的高层表示样本,进而得到每个样本的高层表示样本;具体的,所述初步的自动睡眠分期模型中间层的输出包括所述初步的自动睡眠分期模型的一个或多个隐藏层的输出,即可以将初步的自动睡眠分期模型的任意一个隐藏层的输出或任意多个隐藏层组合的输出作为高层表示样本。

[0030] 针对每类睡眠分期,对属于该类睡眠分期的样本的高层表示样本进行密度聚类,得到属于该类睡眠分期的多个高层表示子集;得到属于每类睡眠分期的高层表示子集;具体的,为了克服综合多个通道的信息存在特征波形不够明显的问题,通过对高层表示样本进行密度聚类,使得属于每类睡眠分期的多个高层表示子集分别具有不同的特征波形明显度。

[0031] 采用属于每类睡眠分期的高层表示子集训练分类器,得到最终的所述自动睡眠分期模型。具体的,在采用属于每类睡眠分期的高层表示子集训练分类器的过程中,也可以将高层表示子集划分为训练集、验证集、测试集,进而训练并微调分类器得到最终的上述自动睡眠分期模型。

[0032] 具体实施时,在本实施例中,可以通过以下方式对属于该类睡眠分期的样本的高层表示样本进行密度聚类,得到属于该类睡眠分期的多个高层表示子集:

[0033] 针对属于该类睡眠分期的各样本的高层表示样本,对维度大于预设维度阈值的高层表示样本进行降维,降维至维度等于所述预设维度阈值;具体的,该维度可以是高层表示样本的数据点的个数,降维方法可以包括但不限于主成分分析、线性判别分析、等距映射、局部线性嵌入、拉普拉斯特征映射、局部保留投影等方法。

[0034] 针对属于该类睡眠分期的所有高层表示样本,计算每两个高层表示样本之间的距离;

[0035] 针对每个高层表示样本,计算与该高层表示样本的距离小于预设距离阈值的高层表示样本的个数,将该个数作为该高层表示样本的局部密度,将局部密度最大的高层表示样本作为该类睡眠分期的聚类中心;

[0036] 按照与聚类中心的距离大小顺序,将属于该类睡眠分期的所有高层表示样本依次分成多个高层表示子集。

[0037] 具体的,在将属于该类睡眠分期的所有高层表示样本依次分成多个高层表示子集的过程中,针对每类睡眠分期,按照与聚类中心的距离由小到大的顺序形成多个高层表示子集,包含高层表示样本的与聚类中心的距离越大的高层表示子集,特征波形越明显。

[0038] 例如,将属于每个睡眠分期类别的高层表示样本集合记作

[0039] $D_i = \{(x, y) | y = i\}$

[0040] 其中,x表示高层表示样本,y表示对应的睡眠分期类别, $i=1, 2, \dots, num_sleep_stages$, num_sleep_stages 表示睡眠分期的类别数。

[0041] 对 D_i 中的每个样本 $d_{ij} = (x_i^{(j)}, y^{(j)} = i)$,其中 $x_i (i=1, 2, \dots, num_sleep_stages)$ 表示属于第i个睡眠分期类别的高层表示样本,j表示样本序号, $j=1, 2, \dots, |D_i|$, $|D_i|$ 表示 D_i 集合的元素个数。

[0042] 对 D_i 集合中任意两个样本 d_{ij} 与 d_{ik} (d_{ik} 中的k也表示样本序号,只是k的数值不同于j),计算样本之间的距离 dis_{ijk}

[0043] $dis_{ijk} = |x_i^{(j)} - x_i^{(k)}|^2 \quad (y^{(j)} = y^{(k)} = i, 1 \leq j, k \leq |D_i|)$

[0044] 对 D_i 集合中每个样本 d_{ij} ,定义局部密度 ρ_{ij}

[0045] $\rho_{ij} = \sum_k I(dis_{ijk} < d_T)$

[0046] 其中, $I(\cdot)$ 为指示函数, d_T 为所有 dis_{ijk} 中按从小到大排列的位于T%分位点的数值,T $\in [0, 100]$ 为预先设定的参数。

[0047] 记 $d_{ic} = (x_i^{(c)}, y^{(c)} = i)$ 为第i类睡眠分期的聚类中心,其中 $c = \arg \max_j \rho_{ij}$ 。

[0048] 对 dis_{ijk} 按从小到大的顺序划分为N个集合,用 $D_{in} (n=1, 2, \dots, N)$ 表示 D_i 的第n个高层表示子集,用 $D_{im} (m=1, 2, \dots, N)$ 表示 D_i 的第m个高层表示子集, D_{in}, D_{im} 满足

[0049] $\bigcup_{i=1}^N D_{in} = D_i$

[0050] $D_{in} \cap D_{im} = \emptyset (1 \leq m, n \leq N, m \neq n)$

[0051] 用 $E_n = \bigcup_{i=1}^{num_sleep_stages} D_{in}$ 表示包含有各个睡眠分期类别的第n个高层表示子集。

[0052] 具体实施时,本申请发明人发现,如果不进行数据增强,在类别不平衡的数据集上训练,很容易使分类结果更偏向于样本数目更多的类别,造成分类精度下降,为了进一步提高睡眠分期模型的准确率,在本实施例中,对各睡眠分期的高层表示子集进行数据增强,例如,在采用属于每类睡眠分期的高层表示子集训练分类器之前,针对各类睡眠分期的相同

序数的高层表示子集,在各类睡眠分期的相同序数的高层表示子集包括的高层表示样本数目不平衡的情况下,进行平衡处理,直至各类睡眠分期的相同序数的高层表示子集包括的高层表示样本数目相同。

[0053] 具体的,针对各类睡眠分期的第n个高层表示子集 E_n ,针对每类睡眠分期的第n个高层表示子集 D_{in} ,统计该第n个高层表示子集 E_n 包括多少个该类睡眠分期的高层表示样本,如各类睡眠分期的第n个高层表示子集 E_n 所包含的高层表示样本数目 $|D_{in}|$ 不同,则进行平衡处理,直至各类睡眠分期的第n个高层表示子集所包含的高层表示样本数目相同,记对 D_{in} 进行平衡处理之后得到的集合为 D'_{in} ,即使 $|D'_{1n}| = |D'_{2n}| = \dots = |D'_{Sn}|$ ($S = \text{num_sleep_stages}$, $n = 1, 2, \dots, N$),得到N个各类别样本数目平衡后的高层表示子集

$$E'_n = \bigcup_{i=1}^{\text{num_sleep_stages}} D'_{in}。平衡方法可以包括但不限于简单上采样、简单下采样、SMOTE算法、训练$$

生成式对抗网络生成样本等方法。

[0054] 具体实施时,由于采集到的睡眠监测信号有时会受到较强噪声的干扰,有时会出现特征波不明显的情况,这些情况可能导致数据质量的不一致性,为了避免数据质量的不一致性对分类器训练产生影响,提升分类精度,在本实施例中,采用属于每类睡眠分期的高层表示子集训练并微调分类器,包括:

[0055] 针对属于每类睡眠分期的高层表示子集,按照与聚类中心的距离从小到大的顺序,采用属于该类睡眠分期的高层表示子集训练分类器,即先采用与聚类中心的距离相对小的高层表示子集训练分类器,再采用与聚类中心的距离相对大的高层表示子集训练分类器,以得到最终的自动睡眠分期模型。

[0056] 具体实施时,为了进一步避免数据受噪声影响,在本实施例中,在以已知的睡眠数据和已知的睡眠数据的睡眠分期作为样本训练分类器之前,将样本中已知的睡眠数据变换为同一固定频率;即将脑电信号EEG_raw、眼电信号EOG_raw以及下颌肌电信号EMG_raw的采样频率变换到某一固定频率,例如变换采样频率到100Hz,得到EEG_100hz, EOG_100hz, EMG_100hz。还可以对变换频率后的已知的睡眠数据进行归一化处理,对采样频率变换后的脑电信号EEG_100hz、眼电信号EOG_100hz、下颌肌电信号EMG_100hz进行数据归一化(数据归一化的方法可以包括但不限于线性函数归一化、零均值归一化等),分别得到EEG_norm, EOG_norm, EMG_norm, 将每一帧数据归一化后的脑电信号EEG_norm、眼电信号EOG_norm、下颌肌电信号EMG_norm与前后若干帧信号拼接起来,作为一个样本,从而得到一个脑电信号、眼电信号、下颌肌电信号联合的数据集ExG,进而采用数据集ExG来训练分类器。

[0057] 具体的,以下详细描述实施上述自动睡眠分期方法的过程,如图2所示,该过程包括以下步骤:

[0058] 步骤S1、从睡眠监测得到的PSG信号中,提取脑电信号、眼电信号、下颌肌电信号及医师标注的睡眠分期,以此作为样本,并对样本进行预处理,得到预处理后的脑电信号、眼电信号、下颌肌电信号;

[0059] 具体的,步骤S1可以具体包括S11、S12、S13以及S14:

[0060] 步骤S11、从睡眠监测记录文件中,根据文件格式协议,将脑电信号、眼电信号、下颌肌电信号提取出来,分别得到EEG_raw, EOG_raw, EMG_raw;

[0061] 步骤S12、将提取出的脑电信号EEG_raw、眼电信号EOG_raw、下颌肌电信号EMG_raw

的采样频率变换到某一固定频率,例如变换采样频率到100Hz,分别得到EEG_100hz,EOG_100hz,EMG_100hz;

[0062] 步骤S13、对采样频率变换后的脑电信号EEG_100hz、眼电信号EOG_100hz、下颌肌电信号EMG_100hz进行数据归一化,得到EEG_norm,EOG_norm,EMG_norm;

[0063] 步骤S14、将每一帧数据归一化后的脑电信号EEG_norm、眼电信号EOG_norm、下颌肌电信号EMG_norm与前后若干帧信号拼接起来,作为一个样本,从而得到一个脑电信号、眼电信号、下颌肌电信号联合的样本数据集ExG。

[0064] 步骤S2、将步骤S1得到的预处理后的脑电信号、眼电信号、下颌肌电信号组成的样本数据集划分为训练集、验证集、测试集,并训练一个神经网络,得到初步的自动睡眠分期模型;

[0065] 具体的,步骤S2具体可以包括S21、S22:

[0066] 步骤S21、将得到的脑电信号、眼电信号、下颌肌电信号联合的样本数据集,划分为训练集ExG_train、验证集ExG_validation、测试集ExG_test;

[0067] 步骤S22、根据划分的训练集、验证集、测试集,训练分类网络,得到初步的自动睡眠分期模型model01。所述的分类网络,包括任意形式的卷积神经网络和循环神经网络。

[0068] 步骤S3、对所有样本 ExG_i ($i = 1, 2, \dots, n_{samples}$), $n_{samples}$ 为样本数,以样本作为步骤S2得到的初步的自动睡眠分期模型model01的输入,将初步的自动睡眠分期模型model01中间层的输出 $f_i = feature(ExG_i)$,作为样本的高层表示样本,进而得到所有样本的高层表示样本;

[0069] 具体的,步骤S3具体可以包括S31、S32:

[0070] 步骤S31、将脑电信号、眼电信号、下颌肌电信号联合的样本数据集ExG,输入初步的自动睡眠分期模型model01,得到模型的中间层输出 $f_i = feature(ExG_i)$,作为数据集的高层表示。所述的中间层输出,可以是分类模型任意一个隐藏层的输出或任意几个隐藏层输出的组合;

[0071] 步骤S32、将样本的高层表示 $f_i = feature(ExG_i)$ 和睡眠分期标签 y_i 组合在一起,作为高层表示样本数据集 $D = \{(f_i, y_i), i = 1, 2, \dots, n_{samples}\}$ 。

[0072] 步骤S4、将属于每个睡眠分期类别的高层表示样本集合记作

[0073] $D_i = \{(x, y) | y = i\}$

[0074] 其中,x表示高层表示样本,y表示对应的睡眠分期类别, $i = 1, 2, \dots, num_sleep_stages$, num_sleep_stages 表示睡眠分期的类别数。

[0075] 对 D_i 中的每个样本 $d_{ij} = (x_i^{(j)}, y^{(j)} = i)$,其中 x_i ($i = 1, 2, \dots, num_sleep_stages$) 表示属于第i个睡眠分期类别的高层表示样本,j表示样本序号, $j = 1, 2, \dots, |D_i|$, $|D_i|$ 表示 D_i 集合的元素个数。

[0076] 对 D_i 集合中任意两个样本 d_{ij} 与 d_{ik} (d_{ik} 中的k也表示样本序号,只是k的数值不同于j),计算样本之间的距离 dis_{ijk}

[0077] $dis_{ijk} = |x_i^{(j)} - x_i^{(k)}|^2$ ($y^{(j)} = y^{(k)} = i, 1 \leq j, k \leq |D_i|$)

[0078] 对 D_i 集合中每个样本 d_{ij} ,定义局部密度 ρ_{ij}

$$[0079] \quad \rho_{ij} = \sum_k I(dis_{ijk} < d_T)$$

[0080] 其中, $I(\cdot)$ 为指示函数, d_T 为所有 dis_{ijk} 中按从小到大排列的位于 $T\%$ 分位点的数值, $T \in [0, 100]$ 为预先设定的参数。

[0081] 记 $d_{ic} = (x_i^{(c)}, y_i^{(c)} = i)$ 为第 i 类睡眠分期的聚类中心, 其中 $c = \arg \max_j \rho_{ij}$ 。

[0082] 对 dis_{ick} 按从小到大的顺序划分为 N 个集合, 用 D_{in} ($n = 1, 2, \dots, N$) 表示 D_i 的第 n 个高层表示子集, 用 $E_n = \bigcup_{i=1}^{num_sleep_stages} D_{in}$ 表示包含有各个睡眠分期类别的第 n 个高层表示子集。

[0083] 步骤 S5、对步骤 S4 得到的各类睡眠分期的第 j 个高层表示子集 E_j , 统计其中属于该类睡眠分期的样本数量, 进行各类别样本数目的平衡, 直至各类睡眠分期的第 j 个高层表示子集 E_j 所包含的高层表示样本数目相同, 即使 $|D'_{1j}| = |D'_{2j}| = \dots = |D'_{num_sleep_stages, j}|$ ($j = 1, 2, \dots, N$), 得到 N 个睡眠分期类别的样本数目平衡后的高层表示子集

$$D_{augmentation}^j = \bigcup_{i=1}^{num_sleep_stages} D'_{ij};$$

[0084] 具体的, 步骤 S5 具体可以包括 S51、S52:

[0085] 步骤 S51、对每类睡眠分期的第 j 个高层表示子集, 统计其中该睡眠分期类别的样本数目;

[0086] 步骤 S52、若各个睡眠分期类别的样本数目不平衡, 则进行类别平衡, 得到 N 个类别平衡后的高层表示子集。所述的类别平衡可以包括但不限于简单上采样、简单下采样、SMOTE 算法、训练生成式对抗网络生成样本等方法。

[0087] 步骤 S6、对步骤 S5 得到的 N 个各类别样本数目平衡后的高层表示子集, 对每个子集划分训练集、验证集、测试集, 并设计学习策略, 根据设计的学习策略分阶段地训练分类器, 得到最终的自动睡眠分期模型。

[0088] 具体的, 步骤 S6 具体可以包括 S61、S62:

[0089] 步骤 S61、对已得到的 N 个类别平衡后的高层表示子集中的每个子集 $D_{augmentation}^j$, 划分训练集 $D_{augmentation}^{j_train}$ 、验证集 $D_{augmentation}^{j_validation}$ 、测试集 $D_{augmentation}^{j_test}$;

[0090] 步骤 S62、设计学习策略, 根据设计的学习策略分阶段地训练分类器, 先在局部密度大的高层表示子集上训练、再在局部密度小的高层表示子集上训练, 得到最终的自动睡眠分期模型。

[0091] 具体实施时, 在本实施例中, 还提供了计算机设备, 如图 3 所示, 包括存储器 302、处理器 304 及存储在存储器上并可在处理器上运行的计算机程序, 所述处理器执行所述计算机程序时实现上述任意的自动睡眠分期方法。

[0092] 具体的, 该计算机设备可以是计算机终端、服务器或者类似的运算装置。

[0093] 具体实施时, 在本实施例中, 还提供了计算机可读存储介质, 所述计算机可读存储介质存储有执行上述任意的自动睡眠分期方法的计算机程序。

[0094] 具体的, 计算机可读存储介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块

或其他数据。计算机可读存储介质的例子包括,但不限于相变内存 (PRAM)、静态随机存取存储器 (SRAM)、动态随机存取存储器 (DRAM)、其他类型的随机存取存储器 (RAM)、只读存储器 (ROM)、电可擦除可编程只读存储器 (EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器 (CD-ROM)、数字多功能光盘 (DVD) 或其他光学存储、磁盒式磁带,磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质,可用于存储可以被计算设备访问的信息。按照本文中的界定,计算机可读存储介质不包括暂存电脑可读媒体 (transitory media),如调制的数据信号和载波。

[0095] 基于同一发明构思,本发明实施例中还提供了一种自动睡眠分期装置,如下面的实施例所述。由于自动睡眠分期装置解决问题的原理与自动睡眠分期方法相似,因此自动睡眠分期装置的实施可以参见自动睡眠分期方法的实施,重复之处不再赘述。以下所使用的,术语“单元”或者“模块”可以实现预定功能的软件和/或硬件的组合。尽管以下实施例所描述的装置较佳地以软件来实现,但是硬件,或者软件和硬件的组合的实现也是可能并被构想的。

[0096] 图4是本发明实施例的自动睡眠分期装置的一种结构框图,如图4所示,该装置包括:

[0097] 数据获取模块402,用于获取待处理的睡眠数据,其中,所述睡眠数据包括脑电信号、眼电信号以及下颌肌电信号;

[0098] 自动睡眠分期模块404,用于将所述睡眠数据输入自动睡眠分期模型,输出所述睡眠数据的睡眠分期,其中,所述自动睡眠分期模型是以已知的睡眠数据和所述已知的睡眠数据的睡眠分期作为样本训练分类器得到的。

[0099] 在一个实施例中,所述自动睡眠分期模块,包括:

[0100] 第一训练单元,用于采用所有样本训练分类器,得到初步的自动睡眠分期模型;

[0101] 高层表示获取单元,用于针对每个样本,将该样本的睡眠数据输入所述初步的自动睡眠分期模型,将所述初步的自动睡眠分期模型中间层的输出作为该样本的高层表示样本,得到每个样本的高层表示样本;

[0102] 密度聚类单元,用于针对每类睡眠分期,对属于该类睡眠分期的样本的高层表示样本进行密度聚类,得到属于该类睡眠分期的多个高层表示子集;得到属于每类睡眠分期的高层表示子集;

[0103] 第二训练单元,用于采用属于每类睡眠分期的高层表示子集训练分类器,得到最终的所述自动睡眠分期模型。

[0104] 在一个实施例中,所述密度聚类单元,包括:

[0105] 降维子单元,用于针对属于该类睡眠分期的各样本的高层表示样本,对维度大于预设维度阈值的高层表示样本进行降维,降维至维度等于所述预设维度阈值;

[0106] 距离计算子单元,用于针对属于该类睡眠分期的所有高层表示样本,计算每两个高层表示样本之间的距离;

[0107] 局部密度确定子单元,用于针对每个高层表示样本,计算与该高层表示样本的距离小于预设距离阈值的高层表示样本的个数,将该个数作为该高层表示样本的局部密度,将局部密度最大的高层表示样本作为该类睡眠分期的聚类中心;

[0108] 密度聚类子单元,用于按照与聚类中心的距离大小顺序,将属于该类睡眠分期的

所有高层表示样本依次分成多个高层表示子集。

[0109] 在一个实施例中,上述装置还包括:

[0110] 平衡模块,用于在采用属于每类睡眠分期的高层表示子集训练分类器之前,针对各类睡眠分期的相同序数的高层表示子集,在各类睡眠分期的相同序数的高层表示子集包括的高层表示样本数目不平衡的情况下,进行平衡处理,直至各类睡眠分期的相同序数的高层表示子集包括的高层表示样本数目相同。

[0111] 在一个实施例中,所述初步的自动睡眠分期模型中间层的输出包括所述初步的自动睡眠分期模型的一个或多个隐藏层的输出。

[0112] 在一个实施例中,所述第二训练单元具体用于,针对属于每类睡眠分期的高层表示子集,按照与聚类中心的距离由小到大的顺序,采用属于各类睡眠分期的各高层表示子集依次训练并微调分类器。

[0113] 在一个实施例中,上述装置还包括:

[0114] 预处理模块,用于在以已知的睡眠数据和所述已知的睡眠数据的睡眠分期作为样本训练分类器之前,将所述样本中所述已知的睡眠数据变换为同一固定频率;对变换频率后的所述已知的睡眠数据进行归一化处理。

[0115] 本发明实施例实现了如下技术效果:提出了采用训练的自动睡眠分期模型来实现基于综合多个通道的信息(例如,包括脑电信号、眼电信号以及下颌肌电信号等多个通道的信息)的自动睡眠分期,该自动睡眠分期过程由于采用了脑电信号、眼电信号以及下颌肌电信号,与现有技术中只根据单通道脑电进行自动睡眠分期的方式相比,可以有利于提高自动睡眠分期的准确率。

[0116] 显然,本领域的技术人员应该明白,上述的本发明实施例的各模块或各步骤可以用通用的计算装置来实现,它们可以集中在单个的计算装置上,或者分布在多个计算装置所组成的网络上,可选地,它们可以用计算装置可执行的程序代码来实现,从而,可以将它们存储在存储装置中由计算装置来执行,并且在某些情况下,可以以不同于此处的顺序执行所示出或描述的步骤,或者将它们分别制作成各个集成电路模块,或者将它们中的多个模块或步骤制作成单个集成电路模块来实现。这样,本发明实施例不限制于任何特定的硬件和软件结合。

[0117] 以上所述仅为本发明的优选实施例而已,并不用于限制本发明,对于本领域的技术人员来说,本发明实施例可以有各种更改和变化。凡在本发明的精神和原则之内,所作的任何修改、等同替换、改进等,均应包含在本发明的保护范围之内。

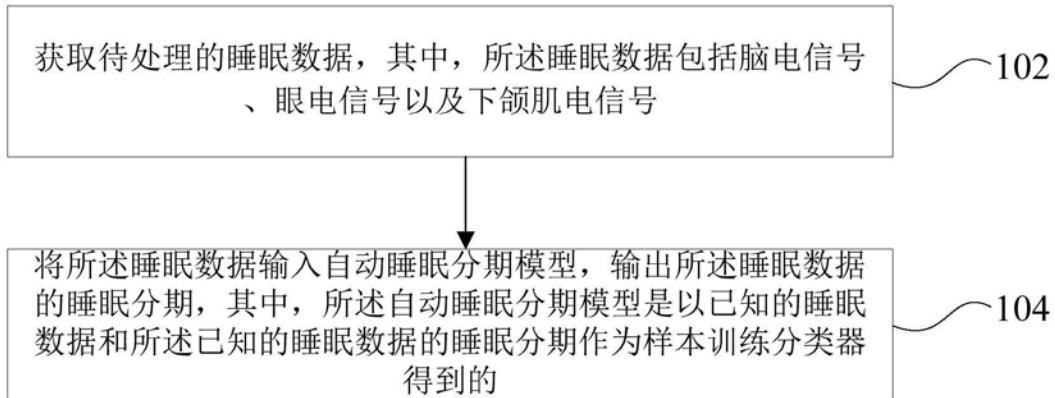


图1

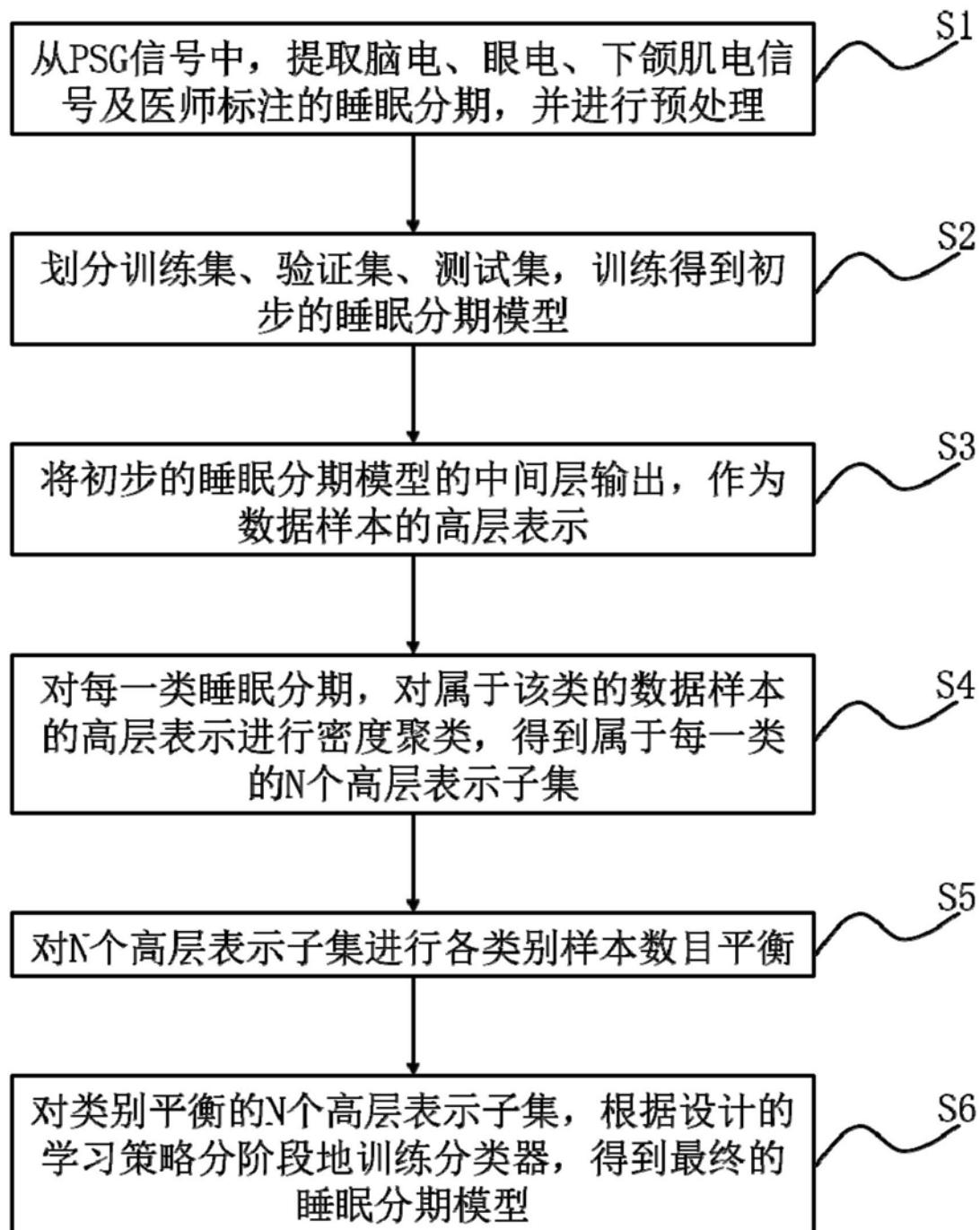


图2

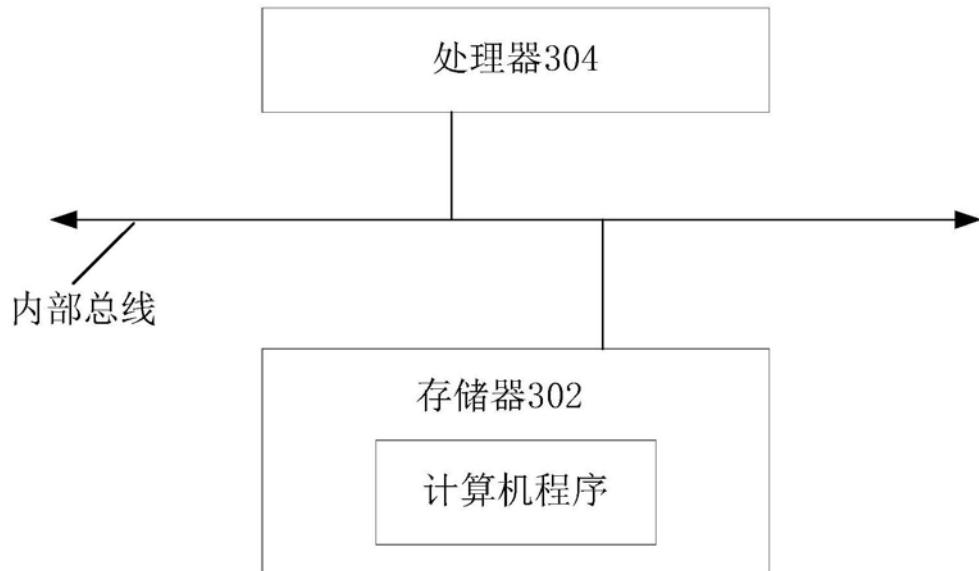


图3

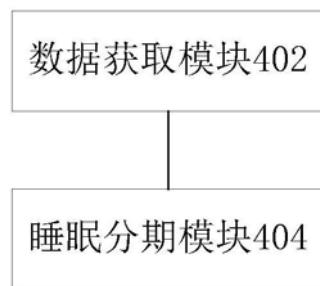


图4

专利名称(译)	自动睡眠分期方法、装置、计算机设备及计算机存储介质		
公开(公告)号	CN110141226A	公开(公告)日	2019-08-20
申请号	CN201910455076.6	申请日	2019-05-29
[标]申请(专利权)人(译)	清华大学深圳研究生院 东莞见达信息技术有限公司		
申请(专利权)人(译)	清华大学深圳研究生院 东莞见达信息技术有限公司		
当前申请(专利权)人(译)	清华大学深圳研究生院 东莞见达信息技术有限公司		
[标]发明人	王兴军 徐子尧 冯发润 黄海平 李庆		
发明人	王兴军 徐子尧 邹庆言 冯发润 黄海平 李庆		
IPC分类号	A61B5/0476 A61B5/0496 A61B5/0488 A61B5/00		
CPC分类号	A61B5/0476 A61B5/0488 A61B5/0496 A61B5/4806 A61B5/4812 A61B5/7235 A61B5/7267		
代理人(译)	周晓飞 王天尧		
外部链接	Espacenet Sipo		

摘要(译)

本发明实施例提供了一种自动睡眠分期方法、装置、计算机设备及计算机可读存储介质，其中，该方法包括：获取待处理的睡眠数据，其中，所述睡眠数据包括脑电信号、眼电信号以及下颌肌电信号；将所述睡眠数据输入自动睡眠分期模型，输出所述睡眠数据的睡眠分期，其中，所述自动睡眠分期模型是以已知的睡眠数据和所述已知的睡眠数据的睡眠分期作为样本训练分类器得到的。该方案的自动睡眠分期过程由于采用了脑电信号、眼电信号以及下颌肌电信号，与现有技术中只根据单通道脑电进行自动睡眠分期的方式相比，可以有利于提高自动睡眠分期的准确性。

获取待处理的睡眠数据，其中，所述睡眠数据包括脑电信号、眼电信号以及下颌肌电信号

102

将所述睡眠数据输入自动睡眠分期模型，输出所述睡眠数据的睡眠分期，其中，所述自动睡眠分期模型是以已知的睡眠数据和所述已知的睡眠数据的睡眠分期作为样本训练分类器得到的

104