



(12) 发明专利

(10) 授权公告号 CN 116746887 B

(45) 授权公告日 2023.12.01

(21) 申请号 202311042494.5

G06F 18/213 (2023.01)

(22) 申请日 2023.08.18

G06N 3/0442 (2023.01)

(65) 同一申请的已公布的文献号

G06N 3/08 (2023.01)

申请公布号 CN 116746887 A

G06F 18/24 (2023.01)

(43) 申请公布日 2023.09.15

G10L 25/03 (2013.01)

(73) 专利权人 智慧眼科技股份有限公司

G10L 25/24 (2013.01)

地址 410205 湖南省长沙市长沙高新开发区尖山路39号长沙中电软件园一期14栋

G10L 25/30 (2013.01)

G10L 25/66 (2013.01)

(72) 发明人 刘伟华 沈梦强 左勇

(56) 对比文件

(74) 专利代理机构 北京集佳知识产权代理有限公司 11227

CN 115802938 A, 2023.03.14

CN 113012719 A, 2021.06.22

专利代理师 郭庆玲

CN 113974607 A, 2022.01.28

(51) Int. Cl.

WO 2023128713 A1, 2023.07.06

CN 113724737 A, 2021.11.30

A61B 5/00 (2006.01)

审查员 廖怡芳

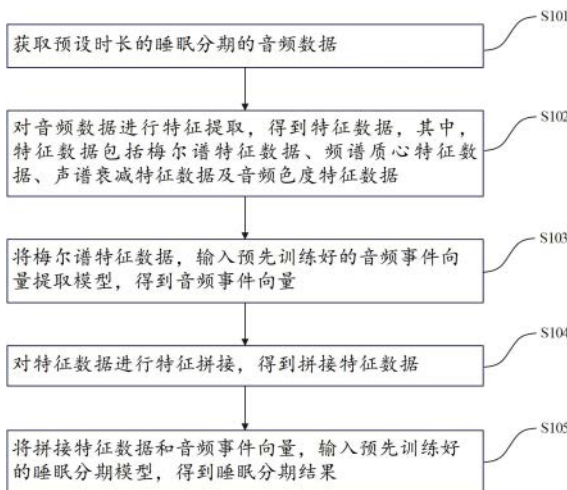
权利要求书3页 说明书11页 附图5页

(54) 发明名称

一种基于音频的睡眠分期的方法、系统、终端及存储介质

(57) 摘要

本申请提供一种基于音频的睡眠分期的方法、系统、终端及存储介质,涉及音频领域,特别是涉及一种基于音频的睡眠分期的方法,所述方法包括:获取预设时长的睡眠分期的音频数据;对所述音频数据进行特征提取,得到特征数据,其中,所述特征数据包括梅尔谱特征数据、频谱质心特征数据、声谱衰减特征数据及音频色度特征数据;将所述梅尔谱特征数据,输入预先训练好的音频事件向量提取模型,得到音频事件向量;对所述特征数据进行特征拼接,得到拼接特征数据;将所述拼接特征数据和所述音频事件向量,输入预先训练好的睡眠分期模型,得到睡眠分期结果。本申请能够提高睡眠分期的准确性。



1. 一种基于音频的睡眠分期的方法,其特征在于,包括:

获取预设时长的睡眠分期的音频数据;

对所述音频数据进行特征提取,得到特征数据,其中,所述特征数据包括梅尔谱特征数据、频谱质心特征数据、声谱衰减特征数据及音频色度特征数据;

将所述梅尔谱特征数据,输入预先训练好的音频事件向量提取模型,得到音频事件向量;

对所述特征数据进行特征拼接,得到拼接特征数据;

将所述拼接特征数据和所述音频事件向量,输入预先训练好的睡眠分期模型,得到睡眠分期结果;

其中,所述将所述梅尔谱特征数据,输入预先训练好的音频事件向量提取模型,得到音频事件向量之前,还包括建立所述训练好的音频事件向量提取模型,所述建立所述训练好的音频事件向量提取模型包括:

采集音频事件数据,按照所述预设时长对所述音频事件数据进行截取,得到音频事件样本;

根据预设事件类别,对所述音频事件样本进行标注,得到音频事件标注样本;

对所述音频事件标注样本,进行梅尔谱特征提取,得到音频事件特征样本;

构建睡眠音频事件分类模型,利用所述音频事件特征样本,对所述睡眠音频事件分类模型进行训练,得到训练好的睡眠音频事件分类模型,其中,所述睡眠音频事件分类模型包括第一特征提取网络模块和第一分类层;

将所述训练好的睡眠音频事件分类模型的第一分类层去除,并对所述训练好的睡眠音频事件分类模型的第一特征提取网络模块的输出做归一化处理,得到所述训练好的音频事件向量提取模型。

2. 如权利要求1所述的方法,其特征在于,还包括建立所述训练好的睡眠分期模型,所述建立所述训练好的睡眠分期模型包括:

采集睡眠音频数据,按照所述预设时长对所述睡眠音频数据进行截取,得到睡眠音频样本;

根据预设睡眠类别,对所述睡眠音频样本进行标注,得到睡眠音频标注样本;

对所述睡眠音频标注样本,进行特征提取,得到特征样本,其中,所述特征样本包括梅尔谱特征样本、频谱质心特征样本、声谱衰减特征样本及音频色度特征样本;

将所述梅尔谱特征样本,输入所述训练好的音频事件向量提取模型,得到音频事件样本向量;

对所述特征样本进行特征拼接,得到拼接特征样本;

构建睡眠分期模型,利用所述拼接特征样本和所述音频事件样本向量,对所述睡眠分期模型进行训练,得到所述训练好的睡眠分期模型。

3. 如权利要求2所述的方法,其特征在于,所述睡眠分期模型,包括第二特征提取网络模块、长短时记忆神经网络模块和第二分类层,所述构建睡眠分期模型,包括:

以所述拼接特征样本为输入,对所述拼接特征样本的进行特征编码得到的第一特征样本为输出,构建第二特征提取网络模块;

以所述第一特征样本和所述音频事件样本向量进行重复拼接得到的第二特征样本为

输入,对所述第二特征样本进行特征处理得到的第三特征样本为输出,构建长短时记忆神经网络模块;

以所述第三特征样本为输入,所述第三特征样本所属的所述预设睡眠类别为输出,构建第二分类层。

4.如权利要求3所述的方法,其特征在于,所述将所述拼接特征数据和所述音频事件向量,输入预先训练好的睡眠分期模型,得到睡眠分期结果,包括:

将所述拼接特征数据,输入所述训练好的睡眠分期模型的第二特征提取网络模块,得到第一特征数据;

将所述第一特征数据和所述音频事件向量进行重复拼接,得到的第二特征数据;

将所述第二特征数据,输入所述训练好的睡眠分期模型的长短时记忆神经网络模块,得到第三特征数据;

将所述第三特征数据,输入所述训练好的睡眠分期模型的第二分类层,得到睡眠分期结果,其中,所述睡眠分期结果包括所述第三特征数据所属的所述预设睡眠类别。

5.如权利要求1所述的方法,其特征在于,所述预设事件类别包括:咳嗽、鼾声、雨声、风声、梦呓、踢被子声、翻身、平静睡眠。

6.如权利要求2所述的方法,其特征在于,所述预设睡眠类别包括:清醒、快速动眼睡眠、浅睡和深睡。

7.一种基于音频的睡眠分期的系统,其特征在于,包括:

音频获取单元,用于获取预设时长的睡眠分期的音频数据;

特征提取单元,用于对所述音频数据进行特征提取,得到特征数据,其中,所述特征数据包括梅尔谱特征数据、频谱质心特征数据、声谱衰减特征数据及音频色度特征数据;

向量提取单元,用于将所述梅尔谱特征数据,输入预先训练好的音频事件向量提取模型,得到音频事件向量;

特征拼接单元,用于对所述特征数据进行特征拼接,得到拼接特征数据;

睡眠分期单元,用于将所述拼接特征数据和所述音频事件向量,输入预先训练好的睡眠分期模型,得到睡眠分期结果;

所述向量提取单元,还用于建立所述训练好的音频事件向量提取模型,所述向量提取单元,在执行所述建立所述训练好的音频事件向量提取模型时,具体用于:

采集音频事件数据,按照所述预设时长对所述音频事件数据进行截取,得到音频事件样本;

根据预设事件类别,对所述音频事件样本进行标注,得到音频事件标注样本;

对所述音频事件标注样本,进行梅尔谱特征提取,得到音频事件特征样本;

构建睡眠音频事件分类模型,利用所述音频事件特征样本,对所述睡眠音频事件分类模型进行训练,得到训练好的睡眠音频事件分类模型,其中,所述睡眠音频事件分类模型包括第一特征提取网络模块和第一分类层;

将所述训练好的睡眠音频事件分类模型的第一分类层去除,并对所述训练好的睡眠音频事件分类模型的第一特征提取网络模块的输出做归一化处理,得到所述训练好的音频事件向量提取模型。

8.一种基于音频的睡眠分期的终端,其特征在于,包括:存储介质和处理器;

所述存储介质内存储计算机执行指令；

所述处理器执行所述存储介质存储的计算机执行指令,以实现如权利要求1至6中任一项所述的方法。

9.一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中存储有计算机执行指令,所述计算机执行指令被处理器执行时用于实现如权利要求1至6中任一项所述的方法。

一种基于音频的睡眠分期的方法、系统、终端及存储介质

技术领域

[0001] 本申请涉及音频领域,特别是涉及一种基于音频的睡眠分期的方法、系统、终端及存储介质。

背景技术

[0002] 当代社会人们面临着诸多压力,导致睡眠质量不断降低,而低质量的睡眠会对人们的日常生活产生巨大的负面影响,监测睡眠质量能够帮助人们了解自己的睡眠状态,提高睡眠质量。现行的睡眠质量监测金标准为多导睡眠图(Polysomnography, PSG),但是PSG睡眠监测技术对监测条件要求较高,难以实现人们日常长期的应用,而基于音频的睡眠质量监测方式,具有系统简单、非接触等优点,非常适合用于人们日常的睡眠监测。

[0003] 基于音频进行睡眠分期时,分类的准确率往往会受到睡眠过程中的各种音频事件所影响,而这些事件不属于分期类别,比如,自身发出的音频以及外部环境的音频,这些音频事件都可能造成睡眠分期结果的错误。对于上述问题,现有的解决方式:对睡眠音频做降噪处理,然后对降噪后的音频进行分类,判断当前睡眠状态。但是该方法存在较多问题:睡眠过程中噪声种类繁多,构建降噪模型所需的数据量也较难构建,此外容易出现降噪效果较差问题,还有可能在降噪的同时,将睡眠的音频信息破坏,这些都会对分类的准确率产生影响。

[0004] 因此,如何提高睡眠分期的准确性,是本领域技术人员亟待解决的技术问题。

发明内容

[0005] 为解决上述技术问题,本申请提供一种基于音频的睡眠分期的方法,能够提高睡眠分期的准确性。本申请还提供一种基于音频的睡眠分期的系统、终端及存储介质,具有相同的技术效果。

[0006] 本申请的第一个目的为提供一种基于音频的睡眠分期的方法。

[0007] 本申请的上述申请目的一是通过以下技术方案得以实现的:

[0008] 一种基于音频的睡眠分期的方法,包括:

[0009] 获取预设时长的睡眠分期的音频数据;

[0010] 对所述音频数据进行特征提取,得到特征数据,其中,所述特征数据包括梅尔谱特征数据、频谱质心特征数据、声谱衰减特征数据及音频色度特征数据;

[0011] 将所述梅尔谱特征数据,输入预先训练好的音频事件向量提取模型,得到音频事件向量;

[0012] 对所述特征数据进行特征拼接,得到拼接特征数据;

[0013] 将所述拼接特征数据和所述音频事件向量,输入预先训练好的睡眠分期模型,得到睡眠分期结果。

[0014] 优选地,所述基于音频的睡眠分期的方法中,还包括建立所述训练好的音频事件向量提取模型,所述建立所述训练好的音频事件向量提取模型包括:

- [0015] 采集音频事件数据,按照所述预设时长对所述音频事件数据进行截取,得到音频事件样本;
- [0016] 根据预设事件类别,对所述音频事件样本进行标注,得到音频事件标注样本;
- [0017] 对所述音频事件标注样本,进行梅尔谱特征提取,得到音频事件特征样本;
- [0018] 构建睡眠音频事件分类模型,利用所述音频事件特征样本,对所述睡眠音频事件分类模型进行训练,得到训练好的睡眠音频事件分类模型,其中,所述睡眠音频事件分类模型包括第一特征提取网络模块和第一分类层;
- [0019] 将所述训练好的睡眠音频事件分类模型的第一分类层去除,并对所述训练好的睡眠音频事件分类模型的第一特征提取网络模块的输出做归一化处理,得到所述训练好的音频事件向量提取模型。
- [0020] 优选地,所述基于音频的睡眠分期的方法中,还包括建立所述训练好的睡眠分期模型,所述建立所述训练好的睡眠分期模型包括:
- [0021] 采集睡眠音频数据,按照所述预设时长对所述睡眠音频数据进行截取,得到睡眠音频样本;
- [0022] 根据预设睡眠类别,对所述睡眠音频样本进行标注,得到睡眠音频标注样本;
- [0023] 对所述睡眠音频标注样本,进行特征提取,得到特征样本,其中,所述特征样本包括梅尔谱特征样本、频谱质心特征样本、声谱衰减特征样本及音频色度特征样本;
- [0024] 将所述梅尔谱特征样本,输入所述训练好的音频事件向量提取模型,得到音频事件样本向量;
- [0025] 对所述特征样本进行特征拼接,得到拼接特征样本;
- [0026] 构建睡眠分期模型,利用所述拼接特征样本和所述音频事件样本向量,对所述睡眠分期模型进行训练,得到所述训练好的睡眠分期模型。
- [0027] 优选地,所述基于音频的睡眠分期的方法中,所述睡眠分期模型,包括第二特征提取网络模块、长短时记忆神经网络模块和第二分类层,所述构建睡眠分期模型,包括:
- [0028] 以所述拼接特征样本为输入,对所述拼接特征样本的进行特征编码得到的第一特征样本为输出,构建第二特征提取网络模块;
- [0029] 以所述第一特征样本和所述音频事件样本向量进行重复拼接得到的第二特征样本为输入,对所述第二特征样本进行特征处理得到的第三特征样本为输出,构建长短时记忆神经网络模块;
- [0030] 以所述第三特征样本为输入,所述第三特征样本所属的所述预设睡眠类别为输出,构建第二分类层。
- [0031] 优选地,所述基于音频的睡眠分期的方法中,所述将所述拼接特征数据和所述音频事件向量,输入预先训练好的睡眠分期模型,得到睡眠分期结果,包括:
- [0032] 将所述拼接特征数据,输入所述训练好的睡眠分期模型的第二特征提取网络模块,得到第一特征数据;
- [0033] 将所述第一特征数据和所述音频事件向量进行重复拼接,得到的第二特征数据;
- [0034] 将所述第二特征数据,输入所述训练好的睡眠分期模型的长短时记忆神经网络模块,得到第三特征数据;
- [0035] 将所述第三特征数据,输入所述训练好的睡眠分期模型的第二分类层,得到睡眠

分期结果,其中,所述睡眠分期结果包括所述第三特征数据所属的所述预设睡眠类别。

[0036] 优选地,所述基于音频的睡眠分期的方法中,所述预设事件类别包括:咳嗽、鼾声、雨声、风声、梦呓、踢被子声、翻身、平静睡眠。

[0037] 优选地,所述基于音频的睡眠分期的方法中,所述预设睡眠类别包括:清醒、快速动眼睡眠、浅睡和深睡。

[0038] 本申请的第二个目的为提供一种基于音频的睡眠分期的系统。

[0039] 本申请的上述申请目的二是通过以下技术方案得以实现的:

[0040] 一种基于音频的睡眠分期的系统,包括:

[0041] 音频获取单元,用于获取预设时长的睡眠分期的音频数据;

[0042] 特征提取单元,用于对所述音频数据进行特征提取,得到特征数据,其中,所述特征数据包括梅尔谱特征数据、频谱质心特征数据、声谱衰减特征数据及音频色度特征数据;

[0043] 向量提取单元,用于将所述梅尔谱特征数据,输入预先训练好的音频事件向量提取模型,得到音频事件向量;

[0044] 特征拼接单元,用于对所述特征数据进行特征拼接,得到拼接特征数据;

[0045] 睡眠分期单元,用于将所述拼接特征数据和所述音频事件向量,输入预先训练好的睡眠分期模型,得到睡眠分期结果。

[0046] 本申请的第三个目的为提供一种基于音频的睡眠分期的终端。

[0047] 本申请的上述申请目的三是通过以下技术方案得以实现的:

[0048] 一种基于音频的睡眠分期的终端,包括:存储介质和处理器;

[0049] 所述存储介质内存储计算机执行指令;

[0050] 所述处理器执行所述存储介质存储的计算机执行指令,以实现上述基于音频的睡眠分期的方法中任一所述方法。

[0051] 本申请的第四个目的为提供一种计算机可读存储介质。

[0052] 本申请的上述申请目的四是通过以下技术方案得以实现的:

[0053] 一种计算机可读存储介质,所述计算机可读存储介质中存储有计算机执行指令,所述计算机执行指令被处理器执行时用于实现上述基于音频的睡眠分期的方法中任一所述方法。

[0054] 上述技术方案,通过获取预设时长的睡眠分期的音频数据;在基于音频数据进行睡眠分期时,首先对音频数据进行特征提取,得到特征数据,其中,特征数据包括梅尔谱特征数据、频谱质心特征数据、声谱衰减特征数据及音频色度特征数据;然后将梅尔谱特征数据,输入预先训练好的音频事件向量提取模型,得到音频事件向量;对特征数据进行特征拼接,得到拼接特征数据;将拼接特征数据和音频事件向量,输入预先训练好的睡眠分期模型,得到睡眠分期结果,其中,音频事件向量可以作为当前睡眠音频环境的提示信息,使得睡眠分期模型具备“感知”当前睡眠环境的能力,最终输出音频数据所对应的睡眠分期结果,以此提升睡眠分期模型分类的准确率。

附图说明

[0055] 为了更清楚地说明本申请实施例或现有技术中的技术方案,下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图仅仅是本

申请中记载的一些实施例,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0056] 图1为本申请实施例中提供的一种基于音频的睡眠分期的方法的流程示意图;

[0057] 图2为本申请实施例中提供的睡眠音频事件分类模型的一种结构示意图;

[0058] 图3为本申请实施例中提供的音频事件向量提取模型的一种结构示意图;

[0059] 图4为本申请实施例中提供的睡眠分期模型的一种结构示意图;

[0060] 图5为本申请实施例中提供的一种基于音频的睡眠分期的系统的结构示意图;

[0061] 图6为本申请实施例中提供的一种基于音频的睡眠分期的终端的结构示意图。

具体实施方式

[0062] 为了使本领域的技术人员更好地理解本申请中的技术方案,下面将对本申请实施例中的技术方案进行清楚、完整地描述,显然,所描述的实施例仅仅是本申请的一部分实施例,而不是全部的实施例。基于本申请中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本申请保护的范围。

[0063] 在本申请所提供的实施例中,应该理解到,所揭露的方法和系统,可以通过其它的方式实现。以下所描述的系统实施例仅仅是示意性的,例如,模块的划分,仅仅为一种逻辑功能划分,实际实现时可以有另外的划分方式,如:多个模块或组件可以结合,或可以集成到另一个系统,或一些特征可以忽略,或不执行。另外,所显示或讨论的各组成部分相互之间的耦合、或直接耦合、或通信连接可以是通过一些接口,设备或模块的间接耦合或通信连接,可以是电性的、机械的或其它形式的。

[0064] 应当理解,本申请中如若使用了“系统”、“装置”、“单元”和/或“模块”,仅是用于区分不同级别的不同组件、元件、部件、部分或装配的一种方法。然而,如果其他词语可实现相同的目的,则可通过其他表达来替换该词语。

[0065] 此外,术语“第一”、“第二”仅用于描述目的,而不能理解为指示或暗示相对重要性或者隐含指明所指示的技术特征的数量。由此,限定有“第一”、“第二”的特征可以明示或者隐含地包括一个或者更多该特征。在本申请的描述中,“多个”、“若干个”的含义是两个或两个以上,除非另有明确具体的限定。

[0066] 本申请中如若使用了流程图,则该流程图是用来说明根据本申请的实施例的系统所执行的操作。应当理解的是,前面或后面操作不一定按照顺序来精确地执行。相反,可以按照倒序或同时处理各个步骤。同时,也可以将其他操作添加到这些过程中,或从这些过程移除某一步或数步操作。

[0067] 还需要说明的是,在本文中,诸如术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的物品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种物品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括上述要素的物品或者设备中还存在另外的相同要素。

[0068] 本申请实施例采用递进的方式撰写。

[0069] 如图1所示,本申请实施例提供一种基于音频的睡眠分期的方法,包括:

[0070] S101.获取预设时长的睡眠分期的音频数据;

[0071] 在S101中,具体地,可以获取音频设备按照预设时长采集的睡眠者的睡眠分期的音频数据,其中,睡眠分期的音频数据可以包括音频设备在睡眠者睡眠过程中采集到的所有音频数据,预设时长可以根据实际应用需求确定,例如,在进行整晚的睡眠分期的时候,可以按照每3秒采集一次音频数据,然后根据当前采集的音频数据,通过后续步骤分析确定当前的睡眠分期结果。睡眠分期的音频数据还可以通过其他合理方式获取,本申请对此不作限制。

[0072] S102.对音频数据进行特征提取,得到特征数据,其中,特征数据包括梅尔谱特征数据、频谱质心特征数据、声谱衰减特征数据及音频色度特征数据;

[0073] 在S102中,对音频数据进行4种不同的特征提取,得到梅尔谱特征数据、频谱质心特征数据、声谱衰减特征数据及音频色度特征数据,可以提升后续模型分类的效果。

[0074] 其中,梅尔谱特征,也称梅尔频率倒谱系数(Mel-frequency cepstral coefficients,MFCC),是一种常用于音频信号处理的特征提取方法,梅尔谱特征的提取过程包括以下步骤:首先将音频信号分帧,对每一帧进行傅里叶变换得到频谱,然后再将频谱转换到梅尔频率上,通过对每一段梅尔频率上的频谱进行加权平均,得到该帧的梅尔谱特征。梅尔谱特征具有良好的区分性和鲁棒性,能够提取出音频信号的关键特征,提高模型的准确率和稳定性。

[0075] 频谱质心特征(Spectral Centroid)是描述音色属性的重要物理参数之一,是频率成分的重心,是在一定频率范围内通过能量加权平均的频率,其单位是Hz。它是声音信号的频率分布和能量分布的重要信息。在主观感知领域,频谱质心描述了声音的明亮度,具有阴暗、低沉品质的声音倾向有较多低频内容,频谱质心相对较低,具有明亮、欢快品质的多数集中在高频,频谱质心相对较高。

[0076] 声谱衰减特征(Spectral Roll-off)是音频信号处理中的一种特征提取方法。它表示音频信号频谱中的一个特定点,该点之前的频率成分累积到了总能量的一定比例(通常为85%或90%)。声谱衰减特征可以用于识别音频信号中的不同声音,声谱衰减特征与其他特征(如MFCC等)结合使用,以提高音频处理任务的性能。计算声谱衰减特征的基本步骤如下:(1)分帧:将音频信号分成短时帧,通常每帧的长度为20-30ms,帧与帧之间有一定的重叠。(2)快速傅里叶变换(Fast Fourier Transformation,FFT):对每个帧进行FFT,以将信号从时域转换到频域。(3)计算声谱衰减特征:对于每个帧,找到一个频率点,使得该点之前的频率成分累积到了总能量的一定比例(例如85%或90%)。

[0077] 音频色度特征(Chroma features)是音频信号处理中的一种特征提取方法。它表示音频信号中的音高信息,可以用于识别音乐中的和声、旋律和音调。音频色度特征与其他特征(如MFCC、声谱衰减特征等)结合使用,以提高音频处理任务的性能。计算音频色度特征的基本步骤如下:(1)分帧:将音频信号分成短时帧,通常每帧的长度为20-30ms,帧与帧之间有一定的重叠。(2)快速傅里叶变换:对每个帧进行FFT,以将信号从时域转换到频域。(3)计算音频色度特征:将频谱映射到12个色度区间(代表12个半音),然后计算每个区间的能量。

[0078] S103.将梅尔谱特征数据,输入预先训练好的音频事件向量提取模型,得到音频事件向量;

[0079] 在S103中,音频事件可以是睡眠过程中出现的咳嗽、鼾声、梦呓、风声、雨声等,这

些事件不属于睡眠分期类别,这些音频事件都可能造成睡眠分期结果的错误。在本步骤中利用训练好的音频事件向量提取模型,作为提取先验知识的模型,根据梅格普特征数据,输出音频事件向量(先验知识),其中,音频事件向量可以作为当前睡眠音频环境的提示信息,可以提高后续睡眠分期的准确性。

[0080] S104.对特征数据进行特征拼接,得到拼接特征数据;

[0081] 在S104中,利用特征拼接的方式,将上述4种特征数据进行融合,以增强睡眠分期模型输入特征中包含的信息。在一些实施例,在步骤S102提取梅尔谱特征数据、频谱质心特征数据、声谱衰减特征数据及音频色度特征数据时,可以设定统一的帧长和帧移,以便于特征拼接。需要说明的是,S103和S104的执行顺序可以互换,也可以同时执行,其不影响本实施例的实现。

[0082] S105.将拼接特征数据和音频事件向量,输入预先训练好的睡眠分期模型,得到睡眠分期结果。

[0083] 在S105中,将拼接特征数据和音频事件向量,输入预先训练好的睡眠分期模型,其中,音频事件向量可以作为当前睡眠音频环境的提示信息,使得睡眠分期模型具备“感知”当前睡眠环境的能力,最终输出音频数据所对应的睡眠分期结果,以此提升睡眠分期模型分类的准确率。

[0084] 基于音频进行睡眠分期时,分类的准确率往往会受到睡眠过程中的各种音频事件所影响,而这些事件不属于分期类别,比如,自身发出的音频以及外部环境的音频,这些音频事件都可能造成睡眠分期结果的错误。对于上述问题,现有的解决方式:对睡眠音频做降噪处理,然后对降噪后的音频进行分类,判断当前睡眠状态。但是该方法存在较多问题:睡眠过程中噪声种类繁多,构建降噪模型所需的数据量也较难构建,此外容易出现降噪效果较差问题,还有可能在降噪的同时,将睡眠的音频信息破坏,这些都会对分类的准确率产生影响。

[0085] 上述实施例,通过获取预设时长的睡眠分期的音频数据;在基于音频数据进行睡眠分期时,首先对音频数据进行特征提取,得到特征数据,其中,特征数据包括梅尔谱特征数据、频谱质心特征数据、声谱衰减特征数据及音频色度特征数据;然后将梅尔谱特征数据,输入预先训练好的音频事件向量提取模型,得到音频事件向量;对特征数据进行特征拼接,得到拼接特征数据;将拼接特征数据和音频事件向量,输入预先训练好的睡眠分期模型,得到睡眠分期结果,其中,音频事件向量可以作为当前睡眠音频环境的提示信息,使得睡眠分期模型具备“感知”当前睡眠环境的能力,最终输出音频数据所对应的睡眠分期结果,以此提升睡眠分期模型分类的准确率。综上可知,上述实施例能够提高睡眠分期的准确性。

[0086] 在本申请的其他实施例中,所述基于音频的睡眠分期的方法中,还包括建立训练好的音频事件向量提取模型的步骤,建立训练好的音频事件向量提取模型的步骤的其中一种实现方式,包括:

[0087] S201.采集音频事件数据,按照预设时长对音频事件数据进行截取,得到音频事件样本;

[0088] 在S201中,具体地,可以采集睡眠过程中所有可能出现的音频事件数据,如睡眠过程中出现的咳嗽、鼾声、梦呓、风声、雨声等音频数据,按照预设时长对音频事件数据进行截

取,得到等长的音频事件样本,其中,音频事件数据可以通过音频设备进行采集,预设时长可以根据实际应用需求确定,例如,可以设置为3秒,音频事件样本用于后续模型的训练。

[0089] S202根据预设事件类别,对音频事件样本进行标注,得到音频事件标注样本;

[0090] 在S202中,预设事件类别,可以根据实际应用需求设置,在一些实施例中,预设事件类别可以包括:咳嗽、鼾声、雨声、风声、梦呓、踢被子声、翻身、平静睡眠等。按照预设事件类别,对音频事件样本进行标注后所得到的每个音频事件标注样本,可以携带有对应预设事件类别的标签信息。

[0091] S203.对音频事件标注样本,进行梅尔谱特征提取,得到音频事件特征样本;

[0092] 在S203中,可以对每个音频事件标注样本,按照16kHz进行采样,提取梅尔谱特征,得到音频事件特征样本,其中,音频事件特征样本,用于后续睡眠音频事件分类模型的训练。

[0093] S204.构建睡眠音频事件分类模型,利用音频事件特征样本,对睡眠音频事件分类模型进行训练,得到训练好的睡眠音频事件分类模型,其中,睡眠音频事件分类模型包括第一特征提取网络模块和第一分类层;

[0094] 在S204中,具体地,可以通过如下方式构建睡眠音频事件分类模型:以音频事件特征样本为输入,对音频事件特征样本进行特征编码得到的编码后的特征样本为输出,构建第一特征提取网络模块,以编码后的特征样本为输入,编码后的特征样本所属的预设事件类别为输出,构建第一分类层。

[0095] 其中,第一特征提取网络模块,用于对音频事件特征样本进行特征编码,得到编码后的特征样本;第一特征提取网络模块包括特征提取网络,特征提取网络是指用于进行特征提取的神经网络层,该神经网络层可以是卷积神经网络层、循环神经网络层、全连接神经网络层、长短时记忆神经网络层(Long Short-Term Memory,LSTM)、前馈神经网络层、池化神经网络层等,在此不进行具体限定。在一些实施例中,第一特征提取网络模块包括多层级联的特征提取网络层,其中,不同的特征提取网络层可以是相同的类型的神经网络层,也可以是不同类型的神经网络层,例如第一特征提取网络模块中的各特征提取网络层均为卷积神经网络层,又例如,第一特征提取网络模块的部分特征提取网络层为卷积神经网络层,部分特征提取网络层为全连接神经网络层。

[0096] 第一分类层,用于根据第一特征提取网络模块输出的编码后的特征样本,进行音频事件分类,输出编码后的特征样本所属的预设事件类别。需要说明的是,第一特征提取网络模块中各特征提取网络层针对音频事件特征样本所输出的编码后的特征信息均需要进行音频事件分类,而不只是将第一特征提取网络模块中最后一层特征提取网络层所输出的编码后的特征信息进行音频事件分类。在一些实施例中,第一分类层可以通过分类函数来进行音频事件分类,其中,分类函数可以为softmax函数、sigmoid函数等。分类函数,例如softmax函数,将第一特征提取网络模块中各特征提取网络层所输出的编码后的特征信息进行音频事件分类,得到编码后的特征样本对应于各预设事件类别的概率,并将最大概率对应的预设事件类别作为编码后的特征样本所属的预设事件类别,即作为音频事件样本对应的音频事件分类结果。

[0097] 在一些实施例中,可以使用交叉熵作为目标函数,利用音频事件特征样本,对睡眠音频事件分类模型进行训练,得到训练好的睡眠音频事件分类模型。其还可以合理采用其

他类型的音频分类的训练方法,本申请不限于此。训练好的睡眠音频事件分类模型的一结构示意图,可以参考图2,其中,第一特征提取网络模块包括n层级联的特征提取网络层:第一特征提取网络层、第二特征提取网络层、...、第n特征提取网络层;第一分类层与第一特征提取网络模块连接,第一分类层输出音频事件分类结果。

[0098] S205.将训练好的睡眠音频事件分类模型的第一分类层去除,并对训练好的睡眠音频事件分类模型的第一特征提取网络模块的输出做归一化处理,得到训练好的音频事件向量提取模型。

[0099] 在S205中,在睡眠音频事件分类模型训练完成之后,去除掉第一分类层,将第一特征提取网络模块的输出做归一化处理,调整之后的模型便是训练好的音频事件向量提取模型。利用训练好的音频事件向量提取模型,对特征数据进行处理之后的结果作为音频事件向量。训练好的音频事件向量提取模型的一结构示意图,可以参考图3,其中,第一特征提取网络模块包括n层级联的特征提取网络层:第一特征提取网络层、第二特征提取网络层、...、第n特征提取网络层;第一特征提取网络模块的输出做归一化处理后输出音频事件向量。

[0100] 在本实施例中,考虑到训练好的睡眠音频事件分类模型的第一分类层输出是音频事件类别信息,即该音频属于哪一类预设事件类别,而类别信息比较单一,为了获取到更丰富的先验知识,将去除掉第一分类层后的第一特征提取网络模块作为先验知识的语义向量表达,以此建立训练好的音频事件向量提取模型。基于训练好的音频事件向量提取模型可以提取音频事件向量,作为当前睡眠音频环境的提示信息,使得睡眠分期模型具备“感知”当前睡眠环境的能力,针对特定的睡眠事件再进行睡眠状态的分类,以此可以提升睡眠分期模型分类的准确率。

[0101] 在本申请的其他实施例中,所述基于音频的睡眠分期的方法中,还包括建立训练好的睡眠分期模型的步骤,建立训练好的睡眠分期模型的步骤的其中一种实现方式,包括:

[0102] S301.采集睡眠音频数据,按照预设时长对睡眠音频数据进行截取,得到睡眠音频样本;

[0103] 在S301中,具体地,可以采集睡眠过程中所有可能出现的睡眠音频数据(包含音频事件数据),按照预设时长对睡眠音频数据进行截取,得到等长的睡眠音频样本,其中,睡眠音频数据可以通过音频设备进行采集,预设时长可以根据实际应用需求确定,例如,可以设置为3秒,睡眠音频样本用于后续模型的训练。

[0104] S302根据预设睡眠类别,对睡眠音频样本进行标注,得到睡眠音频标注样本;

[0105] 在S302中,预设睡眠类别,可以根据实际应用需求设置,在一些实施例中,预设事件类别可以包括:清醒、快速动眼睡眠(Rapid Eye Movement,REM)、浅睡和深睡等。按照预设睡眠类别,对睡眠音频样本进行标注后所得到的每个睡眠音频标注样本,可以携带有对应预设睡眠类别的标签信息。

[0106] S303.对睡眠音频标注样本,进行特征提取,得到特征样本,其中,特征样本包括梅尔谱特征样本、频谱质心特征样本、声谱衰减特征样本及音频色度特征样本;

[0107] 在S303中,其具体实施详情,可以参考上述S102。

[0108] S304.将梅尔谱特征样本,输入训练好的音频事件向量提取模型,得到音频事件样本向量;

[0109] 在S304中,具体地,可以将梅尔谱特征样本输入由上述S201-S205建立的训练好的

音频事件向量提取模型,得到音频事件样本向量。

[0110] S305.对特征样本进行特征拼接,得到拼接特征样本;

[0111] 在S305中,其具体实施详情,可以参考上述S104。

[0112] S306.构建睡眠分期模型,利用拼接特征样本和音频事件样本向量,对睡眠分期模型进行训练,得到训练好的睡眠分期模型。

[0113] 在S306中,具体地,睡眠分期模型,可以包括第二特征提取网络模块、长短时记忆神经网络模块和第二分类层,可以通过如下方式构建睡眠分期模型:以拼接特征样本为输入,对拼接特征样本的进行特征编码得到的第一特征样本为输出,构建第二特征提取网络模块;以第一特征样本和音频事件样本向量进行重复拼接得到的第二特征样本为输入,对第二特征样本进行特征处理得到的第三特征样本为输出,构建长短时记忆神经网络模块;以第三特征样本为输入,第三特征样本所属的预设睡眠类别为输出,构建第二分类层。

[0114] 其中,第二特征提取网络模块,用于根据拼接特征样本进行特征编码,得到第一特征样本;第二特征提取网络模块的具体设置,可以参考上述S204中的第一特征提取网络模块;第二特征提取网络模块,可以基于特征提取网络,对拼接特征样本的进行特征编码,得到第一特征样本;

[0115] 进一步,为了将音频事件样本向量嵌入到网络当中,先对第一特征样本和音频事件样本向量进行重复拼接,得到第二特征样本。其中,考虑到第一特征样本和音频事件样本向量可能存在维度不同的问题,音频事件样本向量嵌入的方式为重复拼接。然后将第二特征样本输入长短时记忆神经网络模块,其中,长短时记忆神经网络模块包括用于进行特征提取的长短时记忆神经网络层,长短时记忆神经网络模块,可以基于长短时记忆神经网络层,对第二特征样本进行特征处理,得到第三特征样本;

[0116] 第二分类层,用于根据长短时记忆神经网络模块输出的第三特征样本,进行睡眠分类,输出第三特征样本所属的预设睡眠类别;第二分类层的具体设置,可以参考上述S204中的第一分类层。

[0117] 在一些实施例中,可以使用交叉熵作为目标函数,利用拼接特征样本和音频事件样本向量,对睡眠分期模型进行训练,得到训练好的睡眠分期模型。其还可以合理采用其他类型的音频分类的训练方法,本申请不限于此。训练好的睡眠分期模型的一结构示意图,可以参考图4,其中,第二特征提取网络模块可以包括多层级联的特征提取网络层,长短时记忆神经网络模块以第二特征提取网络模块的输出与音频事件向量进行重复拼接后得到的数据为输入,进行特征处理,第二分类层再根据长短时记忆神经网络模块的输出进行睡眠分期,输出睡眠分期结果。

[0118] 在本实施例中,通过构建包括第二特征提取网络模块、长短时记忆神经网络模块和第二分类层的睡眠分期模型,利用长短时记忆神经网络模块可以将音频事件样本向量嵌入到网络当中;根据拼接特征样本和音频事件样本向量,对睡眠分期模型进行训练,得到训练好的睡眠分期模型。

[0119] 在本申请的其他实施例中,将拼接特征数据和音频事件向量,输入预先训练好的睡眠分期模型,得到睡眠分期结果的步骤的其中一种实现方式,包括:

[0120] S401.将拼接特征数据,输入训练好的睡眠分期模型的第二特征提取网络模块,得到第一特征数据;

[0121] 在S401中,利用训练好的睡眠分期模型的第二特征提取网络模块,对拼接特征数据进行特征编码,得到第一特征数据。

[0122] S402.将第一特征数据和音频事件向量进行重复拼接,得到的第二特征数据;

[0123] 在S402中,为了将音频事件向量嵌入到网络当中,对第一特征数据和音频事件向量进行重复拼接,得到第二特征数据。其中,考虑到第一特征数据和音频事件向量可能存在维度不同的问题,音频事件向量嵌入的方式为重复拼接。

[0124] S403.将第二特征数据,输入训练好的睡眠分期模型的长短时记忆神经网络模块,得到第三特征数据;

[0125] 在S403中,利用训练好的睡眠分期模型的长短时记忆神经网络模块,对第二特征数据进行特征处理,得到第三特征数据。

[0126] S404.将第三特征数据,输入训练好的睡眠分期模型的第二分类层,得到睡眠分期结果,其中,睡眠分期结果包括第三特征数据所属的预设睡眠类别。

[0127] 在S404中,利用训练好的睡眠分期模型的第二分类层,根据第三特征数据,进行睡眠分类,输出第三特征数据所属的预设睡眠类别。

[0128] 在本实施例中,通过训练好的睡眠分期模型的长短时记忆神经网络模块,将音频事件向量嵌入到网络当中,使得音频事件向量可以作为当前睡眠音频环境的提示信息,使得睡眠分期模型具备“感知”当前睡眠环境的能力,以此可以提升睡眠分期模型分类的准确率。

[0129] 如图5所示,在本申请的另一实施例中,还提供一种基于音频的睡眠分期的系统,包括:

[0130] 音频获取单元10,用于获取预设时长的睡眠分期的音频数据;

[0131] 特征提取单元11,用于对音频数据进行特征提取,得到特征数据,其中,特征数据包括梅尔谱特征数据、频谱质心特征数据、声谱衰减特征数据及音频色度特征数据;

[0132] 向量提取单元12,用于将梅尔谱特征数据,输入预先训练好的音频事件向量提取模型,得到音频事件向量;

[0133] 特征拼接单元13,用于对特征数据进行特征拼接,得到拼接特征数据;

[0134] 睡眠分期单元14,用于将拼接特征数据和音频事件向量,输入预先训练好的睡眠分期模型,得到睡眠分期结果。

[0135] 如图6所示,在本申请的另一实施例中,还提供一种基于音频的睡眠分期的终端,包括:存储介质20和处理器21;

[0136] 所述存储介质20内存储计算机执行指令;

[0137] 所述处理器21执行所述存储介质20存储的计算机执行指令,以实现上述基于音频的睡眠分期的方法中任一所述方法。

[0138] 其中,处理器21可以包括一个或者多个处理核心。处理器21通过运行或执行存储在存储介质20内的指令、程序、代码集或指令集,调用存储在存储介质20内的数据,执行本申请的各种功能和处理数据。处理器21可以为特定用途集成电路、数字信号处理器、数字信号处理装置、可编程逻辑装置、现场可编程门阵列、中央处理器、控制器、微控制器和微处理器中的至少一种。可以理解地,对于不同的设备,用于实现上述处理器21功能的电子器件还可以为其它。

[0139] 其中,存储介质20可用于存储指令、程序、代码、代码集或指令集。存储介质20可以包括存储程序区和存储数据区,其中存储程序区可存储用于实现操作系统的指令、用于至少一个功能的指令以及用于实现上述任一基于音频的睡眠分期的方法的指令等;存储数据区可存储上述任一基于音频的睡眠分期的方法中涉及到的数据等。

[0140] 在本申请的另一实施例中,还提供一种计算机可读存储介质,所述计算机可读存储介质中存储有计算机执行指令,所述计算机执行指令被处理器执行时用于实现上述基于音频的睡眠分期的方法中任一所述方法。

[0141] 其中,所述计算机可读存储介质,可以为U盘、移动硬盘、只读存储器、随机存取存储器或者光盘等各种可以存储程序代码的介质。

[0142] 对所公开的实施例的上述说明,使本领域专业技术人员能够实现或使用本发明。对这些实施例的多种修改对本领域的专业技术人员来说将是显而易见的,本文中所定义的一般原理可以在不脱离本发明的精神或范围的情况下,在其它实施例中实现。因此,本申请将不会被限制于本文所示的这些实施例,而是要符合与本文所公开的原理和新颖特点相一致的最宽的范围。

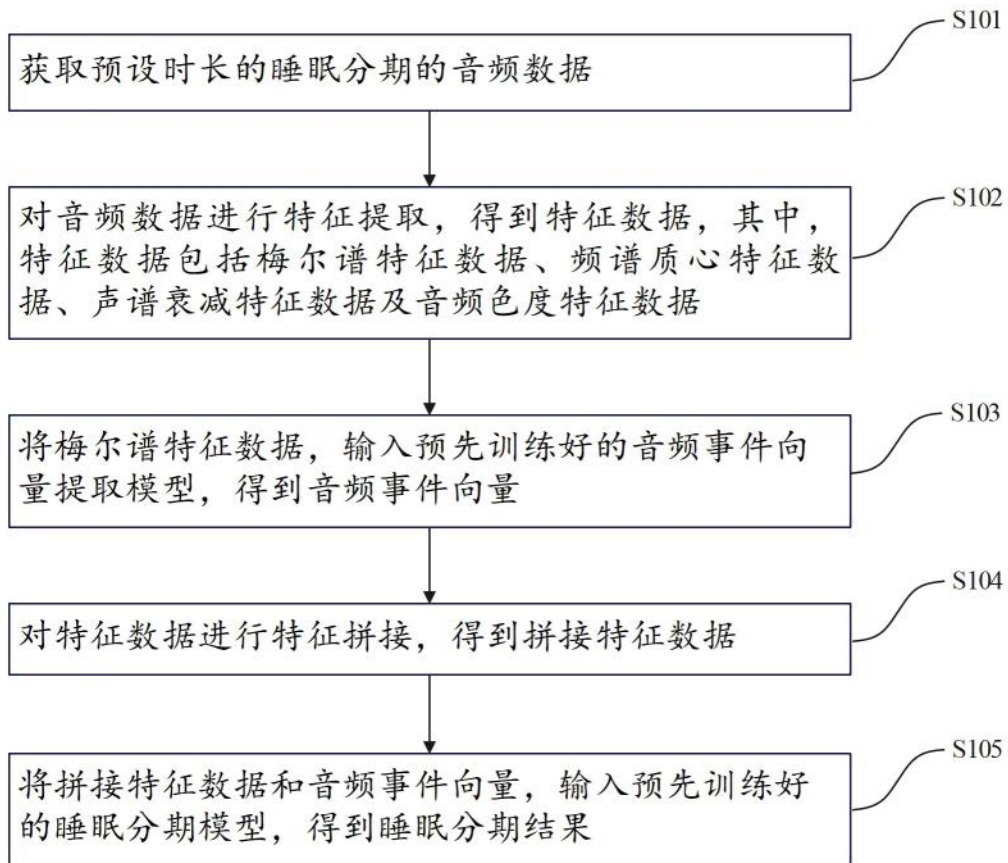


图 1

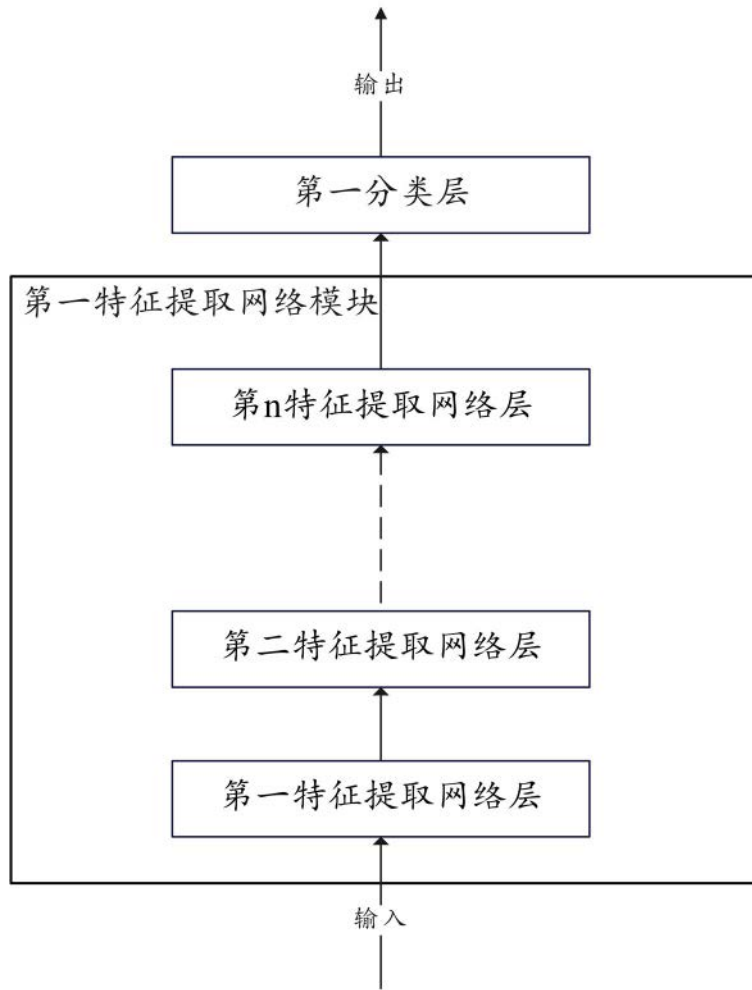


图 2

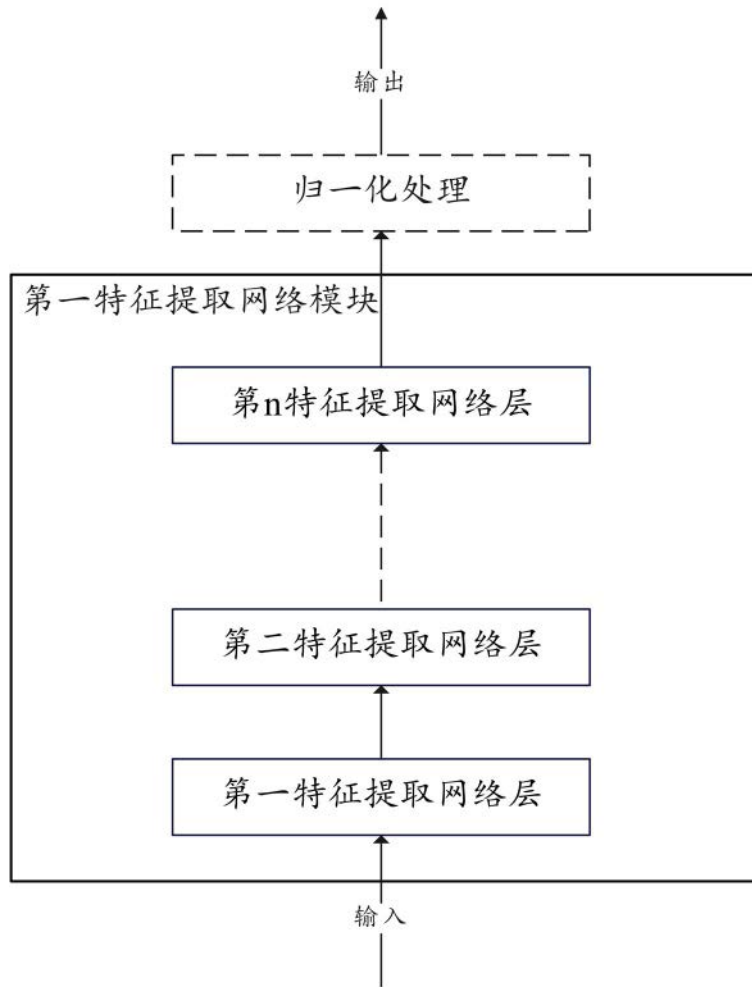


图 3

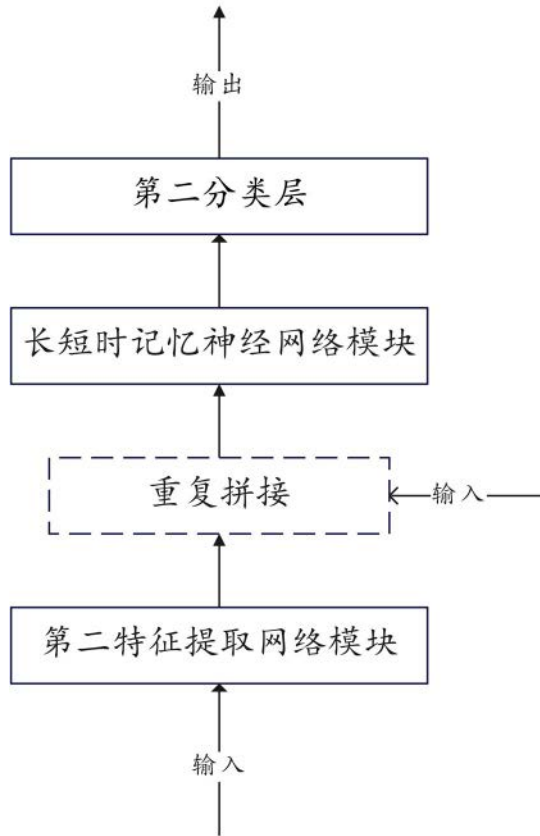


图 4

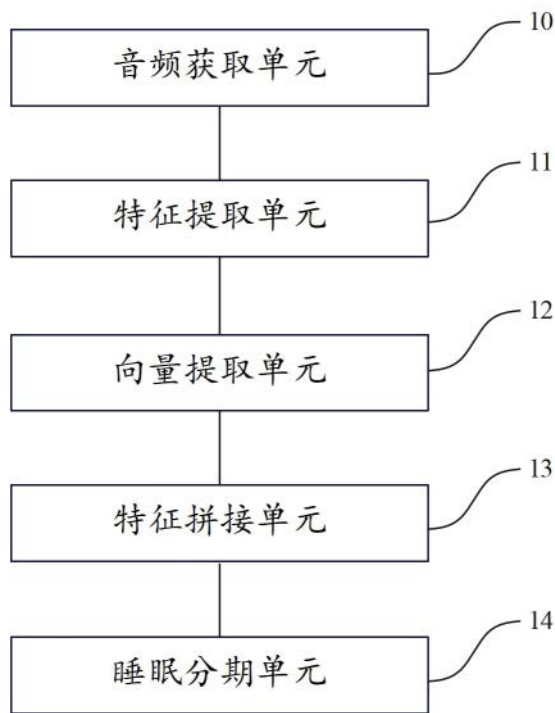


图 5



图 6