

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2021 年 9 月 30 日 (30.09.2021)



(10) 国际公布号
WO 2021/189903 A1

(51) 国际专利分类号:
G10L 25/66 (2013.01) *G10L 25/27* (2013.01)
G10L 25/18 (2013.01)

(21) 国际申请号: PCT/CN2020/131983

(22) 国际申请日: 2020 年 11 月 27 日 (27.11.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202011074898.9 2020 年 10 月 9 日 (09.10.2020) CN

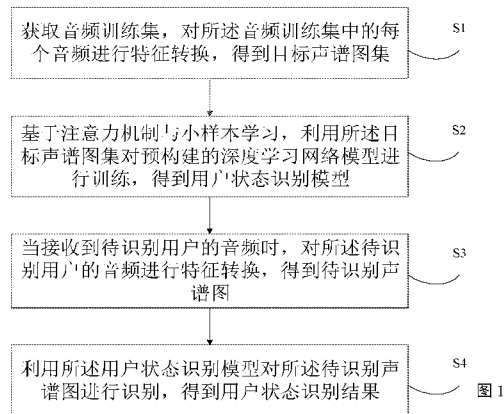
(71) 申请人: 平安科技(深圳)有限公司(PING AN TECHNOLOGY (SHENZHEN) CO.,LTD.) [CN/CN];
中国广东省深圳市福田区福田街道福

安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。

(72) 发明人: 魏文琦(WEI, Wenqi); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。王健宗(WANG, Jianzong); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。贾雪丽(JIA, Xueli); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。张之勇(ZHANG, Zhiyong); 中国广东省深圳市福田区福田街道福安社区益田路 5033 号平安金融中心 23 楼, Guangdong 518000 (CN)。程宁(CHENG, Ning); 中国广东省深圳市

(54) Title: AUDIO-BASED USER STATE IDENTIFICATION METHOD AND APPARATUS, AND ELECTRONIC DEVICE AND STORAGE MEDIUM

(54) 发明名称: 基于音频的用户状态识别方法、装置、电子设备及存储介质



S1 Acquire an audio training set, and perform feature conversion on each piece of audio in the audio training set so as to obtain a target sonogram set
S2 On the basis of an attention mechanism and small sample learning, train a pre-constructed deep learning network model by using the target sonogram set so as to obtain a user state identification model
S3 When audio of a user to be subjected to identification is received, perform feature conversion on the audio of said user so as to obtain a sonogram to be subjected to identification
S4 Identify said sonogram by using the user state identification model so as to obtain a user state identification result

(57) Abstract: An audio-based user state identification method and apparatus, and an electronic device and a computer-readable storage medium. The method comprises: acquiring an audio training set, and performing feature conversion on each piece of audio in the audio training set so as to obtain a target sonogram set (S1); on the basis of an attention mechanism and small sample learning, training a pre-constructed deep learning network model by using the target sonogram set so as to obtain a user state identification model (S2); when audio of a user to be subjected to identification is received, performing feature conversion on the audio of said user so as to obtain

福田区福田街道福安社区益田路5033号平安金融中心23楼, Guangdong 518000 (CN)。

(74) 代理人: 深圳市沃德知识产权代理事务所(普通合伙)(SHENZHEN WORLD INTELLECTUAL PROPERTY AGENCY (GENERAL PARTNERSHIP)); 中国广东省深圳市福田区梅林街道中康路128号卓越城1期2栋301B于志光, Guangdong 518000 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

- 包括国际检索报告(条约第21条(3))。
- 在修改权利要求的期限届满之前进行, 在收到该修改后将重新公布(细则48.2(h))。
- 根据申请人的请求, 在条约第21条(2)(a)所规定的期限届满之前进行。

a sonogram to be subjected to identification (S3); and identifying said sonogram by using the user state identification model so as to obtain a user state identification result (S4). In addition, the present application further relates to blockchain technology, and the audio training set can be stored in a blockchain. By using the method, the consumption of data resources is reduced, and the practicability of a model is enhanced.

(57) 摘要: 一种基于音频的用户状态识别方法和装置、电子设备以及计算机可读存储介质, 方法包括: 获取音频训练集, 对所述音频训练集中的每个音频进行特征转换, 得到目标声谱图集(S1); 基于注意力机制与小样本学习, 利用所述目标声谱图集对预构建的深度学习网络模型进行训练, 得到用户状态识别模型(S2); 当接收到待识别用户的音频时, 对所述待识别用户的音频进行特征转换, 得到待识别声谱图(S3); 利用所述用户状态识别模型对所述待识别声谱图进行识别, 得到用户状态识别结果(S4)。此外还涉及一种区块链技术, 所述音频训练集可以存储在区块链中。利用上述方法降低了数据资源消耗, 增强了模型的实用性。

基于音频的用户状态识别方法、装置、电子设备及存储介质

[0001] 本申请要求于2020年10月9日提交中国专利局、申请号为CN202011074898.9、名称为“基于音频的用户状态识别方法、装置及存储介质”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

技术领域

[0002] 本申请涉及人工智能领域，尤其涉及一种基于音频的用户状态识别方法、装置、电子设备及存储介质。

[0003]

背景技术

[0004] 随着智慧生活的理念逐渐普及，用户状态成为了智慧生活的核心关注点，因此用户状态的识别成为了一件非常重要的事情，如识别用户当前的健康状态，尤其在传染病大肆流行的时候，时刻了解每个人的健康状态很重要。一般情况下，需要用户到医院找医生进行身体检查才能了解其健康情况，医院本身就充满了各种病菌，去医院检查存在着被感染的风险。

发明概述

技术问题

[0005] 发明人意识到，目前，通常利用大量的用户的医学图像（如胸部X光）来训练机器学习模型实现用户状态识别来确定用户的健康状态，但是大量的用户的医学图像耗费了大量的数据资源，且用户的医学图像获取门槛较高导致实用性不强无法更好地推广普及。

问题的解决方案

技术解决方案

[0006] 本申请提供一种基于音频的用户状态识别方法，包括：

[0007] 获取音频训练集，对所述音频训练集中的每个音频进行特征转换，得到目标声

谱图集；

[0008] 基于注意力机制与小样本学习，利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型；

[0009] 当接收到待识别用户的音频时，对所述待识别用户的音频进行特征转换，得到待识别声谱图；

[0010] 利用所述用户状态识别模型对所述待识别声谱图进行识别，得到用户状态识别结果。

[0011] 本申请还提供一种基于音频的用户状态识别装置，所述装置包括：

[0012] 模型生成模块，用于获取音频训练集，对所述音频训练集中的每个音频进行特征转换，得到目标声谱图集；基于注意力机制与小样本学习，利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型；

[0013] 状态识别模块，用于当接收到待识别用户的音频时，对所述待识别用户的音频进行特征转换，得到待识别声谱图；利用所述用户状态识别模型对所述待识别声谱图进行识别，得到用户状态识别结果。

[0014] 本申请还提供一种电子设备，所述电子设备包括：

[0015] 存储器，存储至少一个指令；及

[0016] 处理器，执行所述存储器中存储的指令以实现如下步骤：

[0017] 获取音频训练集，对所述音频训练集中的每个音频进行特征转换，得到目标声谱图集；

[0018] 基于注意力机制与小样本学习，利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型；

[0019] 当接收到待识别用户的音频时，对所述待识别用户的音频进行特征转换，得到待识别声谱图；

[0020] 利用所述用户状态识别模型对所述待识别声谱图进行识别，得到用户状态识别结果。

[0021] 本申请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至少一个指令，所述至少一个指令被电子设备中的处理器执行时实现如下步骤：

[0022] 获取音频训练集，对所述音频训练集中的每个音频进行特征转换，得到目标声

谱图集；

[0023] 基于注意力机制与小样本学习，利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型；

[0024] 当接收到待识别用户的音频时，对所述待识别用户的音频进行特征转换，得到待识别声谱图；

[0025] 利用所述用户状态识别模型对所述待识别声谱图进行识别，得到用户状态识别结果。

发明的有益效果

对附图的简要说明

附图说明

[0026] 图1为本申请一实施例提供的基于音频的用户状态识别方法的流程示意图；

[0027] 图2为本申请一实施例提供的基于音频的用户状态识别方法中得到目标声谱图集的详细流程示意图；

[0028] 图3为本申请一实施例提供的基于音频的用户状态识别装置的模块示意图；

[0029] 图4为本申请一实施例提供的实现基于音频的用户状态识别方法的电子设备的内部结构示意图；

[0030] 本申请目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

发明实施例

本发明的实施方式

[0031] 应当理解，此处所描述的具体实施例仅仅用以解释本申请，并不用于限定本申请。

[0032] 本申请提供一种基于音频的用户状态识别方法。参照图1所示，为本申请一实施例提供的基于音频的用户状态识别方法的流程示意图。该方法可以由一个装置执行，该装置可以由软件和/或硬件实现。

[0033] 在本实施例中，基于音频的用户状态识别方法包括：

[0034] S1、获取音频训练集，对所述音频训练集中的每个音频进行特征转换，得到目标声谱图集；

[0035] 本申请实施例中，所述音频训练集为包含初始标签的音频的集合，较佳地，所述初始标签为用户的疾病情况，例如：急性支气管炎、慢性咽炎、百日咳、发烧；进一步地，由于不同疾病情况下用户的咳嗽音频具有对应的声音特征，因此，较佳地，所述音频训练集为不同的疾病情况对应的咳嗽音频的集合，其中，所述声音特征为咳嗽音频的频域特征可用声谱图表示。

[0036] 进一步地，为了后续模型更好的所述音频训练集中的每个音频的特征更加直观清晰，本申请实施例对所述音频训练集进行特征变换，得到所述目标声谱图集，包括：

[0037] S11、对所述音频训练集中的每个音频进行重采样，得到对应的数字语音信号；

[0038] 本申请实施例中，为了便于对所述音频训练集中的每个音频进行数据处理，对所述音频训练集中的每个音频进行重采样，得到对应的数字语音信号，较佳地，本申请实施例利用数模转换器对所述音频训练集中的每个音频进行重采样。

[0039] S12、对所述数字语音信号进行预加重，得到标准数字语音信号；

[0040] S13、汇总所有的所述标准数字语音信号，得到语音信号集；

[0041] 本申请实施例中，为了对所述音频训练集获取过程中导致的音频信息丢失进行补偿，对所述音频训练集中的每个音频进行预加重操作，

[0042] 详细地，本申请实施例中，所述对所述音频训练集中的每个音频进行预加重操作，包括：将所述音频训练集中的每个音频进行重采样，得到对应的数字语音信号；将所述数字语音信号进行预加重，得到标准数字语音信号，汇总所有的所述标准数字语音信号，得到语音信号集。

[0043] 详细地，本申请实施例利用如下公式进行所述预加重操作：

[0044] $y(t) = x(t) - \mu x(t-1)$

[0045] 其中， $x(t)$ 为所述数字语音信号， t 为时间， $y(t)$ 为所述标准数字语音信号， μ 为所述预加重操作的预设调节值，较佳地， μ 的取值范围为[0.9, 1.0]。

[0046] S14、对所述语音信号集中包含的每个标准数字语音信号进行特征转换，得到目标声谱图集。

- [0047] 本申请实施例中，所述语音信号集中的标准语音信号只能体现音频在时域上的变化，不能体现所述标准语音信号的音频特征，为了体现所述标准语音信号的音频特征，使音频特征更加直观和清晰，对所述语音信号集中的每个标准数字语音信号进行特征转换。
- [0048] 详细地，本申请实施例中对所述语音信号集中的每个标准数字语音信号进行特征转换，包括：利用预设声音处理算法，将所述语音信号集中的每个标准数字语音信号映射在频域，得到对应的目标声谱图，汇总所有的所述目标声谱图得到所述目标声谱图集。
- [0049] 较佳地，本申请中所述声音处理算法为梅尔滤波算法。
- [0050] 本申请实施例中，上述步骤只是对所述音频训练集的每个音频进行特征转换，不会影响所述音频训练集的每个音频对应的所述初始标签，所以所述目标声谱图集中每个目标声谱图都有对应的初始标签。
- [0051] S2、基于注意力机制与小样本学习，利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型；
- [0052] 本申请实施例中，由于所述音频训练集中的样本数量过少，为了保证后续模型的训练精度及鲁棒性，基于注意力机制与小样本学习，利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到基于音频的用户状态识别模型。
- [0053] 详细地，本申请实施例中，所述利用所述目标声谱图集对预构建的深度学习网络模型进行训练，包括：
- [0054] 步骤A：将所述目标声谱图集划分为训练集及测试集；
- [0055] 本申请实施例中，由于所述目标声谱图集集中的样本数据较少且不易获取，直接将所述目标声谱图集作为训练集会使得后续模型的鲁棒性较差，因此，本申请实施例将所述目标声谱图集划分为训练集及测试集，通过利用测试集不断的对训练模型进行测试调整增强模型的鲁棒性，所述将所述目标声谱图集划分为训练集及测试集，包括：将所述目标声谱图集集中的每个目标声谱图按照对应的初始标签进行分类，得到对应的分类目标声谱图集；从所述分类目标声谱图集中随机取出预设数量的目标声谱图作为测试子集，将所述分类声谱图集中所述训练子集的补集作为训练子集；汇总所有的所述训练子集的训练集，汇总所有的

所述测试子集得到测试集，较佳地，本申请实施例中所述预设数量为1。

[0056] 步骤B：利用所述训练集对所述深度学习网络进行训练得到初始识别模型，根据所述测试集对所述初始识别模型进行测试得到损失值，当所述损失值大于预设阈值时返回步骤A，当所述损失值小于或等于预设阈值时，将所述初始识别模型作为用户状态识别模型。

[0057] 较佳地，本申请实施例中所述深度学习网络为卷积神经网络。

[0058] 本申请实施例中，由于所述音频训练集中的音频时间可能不一致，导致所述目标声谱图集集中的图像大小存在差异，进而导致所述深度学习网络模型在训练的过程中提取的所述目标声谱图集集中的目标声谱图的特征维度不同，无法进行统一训练，因此，为了更好地利用所述音频训练集中的数据，本申请实施例利用所述训练集对所述深度学习网络前，需要在所述深度学习网络模型的全连接层前增加注意力机制处理层进行图像特征对齐，其中，所述注意力机制处理层根据图像特征维度不同进行特征对齐的网络，例如：目标声谱图A的在所述深度学习网络模型进行特征提取的图像特征a为 $D \times T1$ 维矩阵，目标声谱图B的在所述深度学习网络模型进行特征提取的图像特征b为 $D \times T2$ 维矩阵，所述注意力机制处理层将图像特征a乘上 $T1 \times 1$ 的预设权重矩阵转化为D维矩阵，将图像特征b乘上 $T2 \times 1$ 的预设权重矩阵转化为D维矩阵，实现图像特征a及图像特征b的特征对齐。

[0059] 进一步地，由于所述训练集的中的样本数量较少，所以本申请实施例需要对所述初始识别模型进行以验证该模型的识别能力方便对模型进行训练调整。

[0060] 详细地，本申请实施例中所述初始识别模型的识别类别和所述目标声谱图集集中的初始标签的类别相同，例如：所述目标声谱图集中共有两种初始标签为慢性咽炎、发烧，那么所述初始识别模型中的识别类别也有相同的两种为慢性咽炎、发烧。进一步地，本申请实施例中所述根据所述测试集对所述初始识别模型进行测试得到损失值，包括：提取所述初始识别模型中每种所述初始标签对应的特征向量，得到目标特征向量；利用所述初始识别模型对所述测试子集中的每个目标声谱图进行特征提取，得到测试特征向量；计算每种所述初始标签对应的所述目标特征向量与所述测试特征向量的距离，得到损失距离值；计算所有所述损失距离值的平均值，得到所述损失值。较佳地，本申请实施例采用欧

氏距离计算方法计算每种所述初始标签对应的所述目标特征向量与所述测试特征向量的距离。

[0061] 进一步地，本领域技术人员可知所述初始模型的不同识别类别连接对应不同的全连接层节点，全连接层节点有对应的顺序，本申请实施例获取初始识别模型的每种识别类别对应的全连接层节点的输出值并按照对应的全连接层节点的顺序进行组合，得到对应的所述目标特征向量；进一步地，本申请实施例将所述测试子集中的每个目标声谱图输入所述初始识别模型，根据所述测试子集中的每个目标声谱图对应的初始标签，获取所述初始识别模型中对应识别类别的全连接层节点的输出值并按照对应的全连接层节点的顺序进行组合，得到所述测试特征向量。

[0062] 本申请的另一实施例中，为了保证数据的隐私性，所述音频训练集可以存储在区块链节点中。

[0063] S3、当接收到待识别用户的音频时，对所述待识别用户的音频进行特征转换，得到待识别声谱图；

[0064] 本申请实施例中，所述待识别用户的音频与所述音频训练集中的音频的类别相同，较佳地，本申请实施例中，所述待识别用户的音频为用户的咳嗽音频。音频训练集

[0065] 进一步地，本申请实施例中对所述待识别用户的音频进行特征转换的方法与上述对所述音频训练集的每个音频进行特征转换的方法相同。

[0066] S4、利用所述用户状态识别模型对所述待识别声谱图进行识别，得到用户状态识别结果。

[0067] 本申请实施例中，所述用户状态识别结果为用户的健康状态，例如：急性支气管炎、慢性咽炎、百日咳、发烧。

[0068] 本申请实施例中，对所述音频训练集中的每个音频进行特征转换，得到目标声谱图集，使所述音频训练集中音频中的特征更加清晰直观，增加了后续模型训练的精度；基于注意力机制与小样本学习，利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型，增强了小样本训练集下的模型的鲁棒性及训练精度；对所述待识别用户的音频进行特征转换，得到待

识别声谱图，使所述待识别用户的音频特征更加清晰直观，提高了后续模型的识别精度；利用所述用户状态识别模型对所述待识别声谱图进行识别，得到用户状态识别结果，利用少量的更易得的音频数据进行模型的训练，降低了模型的训练的数据资源消耗，只需要用户的音频就可以对用户状态进行识别，增强了模型的实用性。

[0069] 如图3所示，是本申请基于音频的用户状态识别装置的功能模块图。

[0070] 本申请所述基于音频的用户状态识别装置100可以安装于电子设备中。根据实现的功能，所述基于音频的用户状态识别装置可以包括模型生成模块101、状态识别模块102。本发所述模块也可以称之为单元，是指一种能够被电子设备处理器所执行，并且能够完成固定功能的一系列计算机程序段，其存储在电子设备的存储器中。

[0071] 在本实施例中，关于各模块/单元的功能如下：

[0072] 所述模型生成模块101用于获取音频训练集，对所述音频训练集中的每个音频进行特征转换，得到目标声谱图集；基于注意力机制与小样本学习，利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型。

[0073] 本申请实施例中，所述音频训练集为包含初始标签的音频的集合，较佳地，所述初始标签为用户的疾病情况，例如：急性支气管炎、慢性咽炎、百日咳、发烧；进一步地，由于不同疾病情况下用户的咳嗽音频具有对应的声音特征，因此，较佳地，所述音频训练集为不同的疾病情况对应的咳嗽音频的集合，其中，所述声音特征为咳嗽音频的频域特征可用声谱图表示。

[0074] 进一步地，为了后续模型更好的所述音频训练集中的每个音频的特征更加直观清晰，本申请实施例所述模型生成模块101利用如下手段对所述音频训练集进行特征变换，得到所述目标声谱图集，包括：

[0075] 对所述音频训练集中的每个音频进行重采样，得到对应的数字语音信号；

[0076] 本申请实施例中，为了便于对所述音频训练集中的每个音频进行数据处理，对所述音频训练集中的每个音频进行重采样，得到对应的数字语音信号，较佳地，本申请实施例利用数模转换器对所述音频训练集中的每个音频进行重采样

- 。
- [0077] 对所述数字语音信号进行预加重，得到标准数字语音信号；
- [0078] 汇总所有的所述标准数字语音信号，得到语音信号集；
- [0079] 本申请实施例中，为了对所述音频训练集获取过程中导致的音频信息丢失进行补偿，对所述音频训练集中的每个音频进行预加重操作，
- [0080] 详细地，本申请实施例中，所述对所述音频训练集中的每个音频进行预加重操作，包括：对所述音频训练集中的每个音频进行重采样，得到对应的数字语音信号；将所述数字语音信号进行预加重，得到标准数字语音信号，汇总所有的所述标准数字语音信号，得到语音信号集。
- [0081] 详细地，本申请实施例所述模型生成模块101利用如下公式进行所述预加重操作：
- [0082] $y(t) = x(t) - \mu x(t-1)$
- [0083] 其中， $x(t)$ 为所述数字语音信号， t 为时间， $y(t)$ 为所述标准数字语音信号， μ 为所述预加重操作的预设调节值，较佳地， μ 的取值范围为 $[0.9, 1.0]$ 。
- [0084] 对所述语音信号集中包含的每个标准数字语音信号进行特征转换，得到目标声谱图集。
- [0085] 本申请实施例中，所述语音信号集中的标准语音信号只能体现音频在时域上的变化，不能体现所述标准语音信号的音频特征，为了体现所述标准语音信号的音频特征，使音频特征更加直观和清晰，对所述语音信号集中的每个标准数字语音信号进行特征转换。
- [0086] 详细地，本申请实施例中所述模型生成模块101利用如下手段对所述语音信号集中的每个标准数字语音信号进行特征转换，包括：利用预设声音处理算法，将所述语音信号集中的每个标准数字语音信号映射在频域，得到对应的目标声谱图，汇总所有的所述目标声谱图得到所述目标声谱图集。
- [0087] 较佳地，本申请中所述声音处理算法为梅尔滤波算法。
- [0088] 本申请实施例中，上述步骤只是对所述音频训练集的每个音频进行特征转换，不会影响所述音频训练集的每个音频对应的所述初始标签，所以所述目标声谱图集中每个目标声谱图都有对应的初始标签。

- [0089] 本申请实施例中，由于所述音频训练集中的样本数量过少，为了保证后续模型的训练精度及鲁棒性，基于注意力机制与小样本学习，利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到基于音频的用户状态识别模型。
- [0090] 详细地，本申请实施例中，所述模型生成模块101利用如下手段对预构建的深度学习网络模型进行训练，包括：
- [0091] 步骤A：将所述目标声谱图集划分为训练集及测试集；
- [0092] 本申请实施例中，由于所述目标声谱图集集中的样本数据较少且不易获取，直接将所述目标声谱图集作为训练集会导致后续模型的鲁棒性较差，因此，本申请实施例将所述目标声谱图集划分为训练集及测试集，通过利用测试集不断的对训练模型进行测试调整增强模型的鲁棒性，所述将所述目标声谱图集划分为训练集及测试集，包括：将所述目标声谱图集集中的每个目标声谱图按照对应的初始标签进行分类，得到对应的分类目标声谱图集；从所述分类目标声谱图集中随机取出预设数量的目标声谱图作为测试子集，将所述分类声谱图集中所述训练子集的补集作为训练子集；汇总所有的所述训练子集的训练集，汇总所有的所述测试子集得到测试集，较佳地，本申请实施例中所述预设数量为1。
- [0093] 步骤B：利用所述训练集对所述深度学习网络进行训练得到初始识别模型，根据所述测试集对所述初始识别模型进行测试得到损失值，当所述损失值大于预设阈值时返回步骤A，当所述损失值小于或等于预设阈值时，将所述初始识别模型作为用户状态识别模型。
- [0094] 较佳地，本申请实施例中所述深度学习网络为卷积神经网络。
- [0095] 本申请实施例中，由于所述音频训练集中的音频时间可能不一致，导致所述目标声谱图集集中的图像大小存在差异，进而导致所述深度学习网络模型在训练的过程中提取的所述目标声谱图集集中的目标声谱图的特征维度不同，无法进行统一训练，因此，为了更好地利用所述音频训练集中的数据，本申请实施例利用所述训练集对所述深度学习网络前，需要在所述深度学习网络模型的全连接层前增加注意力机制处理层进行图像特征对齐，其中，所述注意力机制处理层根据图像特征维度不同进行特征对齐的网络，例如：目标声谱图A的在所述深度学习网络模型进行特征提取的图像特征a为 $D \times T1$ 维矩阵，目标声谱图B的在所述深

度学习网络模型进行特征提取的图像特征b为 $D \times T2$ 维矩阵, 所述注意力机制处理层将图像特征a乘上 $T1 \times 1$ 的预设权重矩阵转化为D维矩阵, 将图像特征b乘上 $T2 \times 1$ 的预设权重矩阵转化为D维矩阵, 实现图像特征a及图像特征b的特征对齐。

[0096] 进一步地, 由于所述训练集的中的样本数量较少, 所以本申请实施例需要对所述初始识别模型进行以验证该模型的识别能力方便对模型进行训练调整。

[0097] 详细地, 本申请实施例中所述初始识别模型的识别类别和所述目标声谱图集中的初始标签的类别相同, 例如: 所述所述目标声谱图集中共有两种初始标签为慢性咽炎、发烧, 那么所述初始识别模型中的识别类别也有相同的两种为慢性咽炎、发烧。进一步地, 本申请实施例中所述模型生成模块101利用如下手段得到损失值, 包括: 提取所述初始识别模型中每种所述初始标签对应的特征向量, 得到目标特征向量; 利用所述初始识别模型对所述测试子集中的每个目标声谱图进行特征提取, 得到测试特征向量; 计算每种所述初始标签对应的所述目标特征向量与所述测试特征向量的距离, 得到损失距离值; 计算所有所述损失距离值的平均值, 得到所述损失值。较佳地, 本申请实施例采用欧氏距离计算方法计算每种所述初始标签对应的所述目标特征向量与所述测试特征向量的距离。

[0098] 进一步地, 本领域技术人员可知所述初始模型的不同识别类别连接对应不同的全连接层节点, 全连接层节点有对应的顺序, 本申请实施例所述模型生成模块101获取初始识别模型的每种识别类别对应的全连接层节点的输出值并按照对应的全连接层节点的顺序进行组合, 得到对应的所述目标特征向量; 进一步地, 本申请实施例所述模型生成模块101将所述测试子集中的每个目标声谱图输入所述初始识别模型, 根据所述测试子集中的每个目标声谱图对应的初始标签, 获取所述初始识别模型中对应识别类别的全连接层节点的输出值并按照对应的全连接层节点的顺序进行组合, 得到所述测试特征向量。

[0099] 本申请的另一实施例中, 为了保证数据的隐私性, 所述音频训练集可以存储在区块链节点中。

[0100] 所述状态识别模块102用于当接收到待识别用户的音频时, 对所述待识别用户的音频进行特征转换, 得到待识别声谱图; 利用所述用户状态识别模型对所述

待识别声谱图进行识别，得到用户状态识别结果。

[0101] 本申请实施例中，所述待识别用户的音频与所述音频训练集中的音频的类别相同，较佳地，本申请实施例中，所述待识别用户的音频为用户的咳嗽音频。音频训练集

[0102] 进一步地，本申请实施例中对所述待识别用户的音频进行特征转换的方法与上述对所述音频训练集的每个音频进行特征转换的方法相同。

[0103] 本申请实施例中，所述用户状态识别结果为用户的疾病情况，例如：急性支气管炎、慢性咽炎、百日咳、发烧。

[0104] 如图44所示，是本申请实现基于音频的用户状态识别方法的电子设备的结构示意图。

[0105] 所述电子设备1可以包括处理器10、存储器11和总线，还可以包括存储在所述存储器11中并可在所述处理器10上运行的计算机程序，如基于音频的用户状态识别程序。

[0106] 其中，所述存储器11至少包括一种类型的可读存储介质，所述可读存储介质包括闪存、移动硬盘、多媒体卡、卡型存储器（例如：SD或DX存储器等）、磁性存储器、磁盘、光盘等。所述存储器11在一些实施例中可以是电子设备1的内部存储单元，例如该电子设备1的移动硬盘。所述存储器11在另一些实施例中也可以是电子设备1的外部存储设备，例如电子设备1上配备的插接式移动硬盘、智能存储卡（Smart Media Card， SMC）、安全数字（Secure Digital， SD）卡、闪存卡（Flash Card）等。进一步地，所述存储器11还可以既包括电子设备1的内部存储单元也包括外部存储设备。所述存储器11不仅可以用于存储安装于电子设备1的应用软件及各类数据，例如基于音频的用户状态识别程序的代码等，还可以用于暂时地存储已经输出或者将要输出的数据。

[0107] 所述处理器10在一些实施例中可以由集成电路组成，例如可以由单个封装的集成电路所组成，也可以是由多个相同功能或不同功能封装的集成电路所组成，包括一个或者多个中央处理器（Central Processing unit， CPU）、微处理器、数字处理芯片、图形处理器及各种控制芯片的组合等。所述处理器10是所述电子设备的控制核心（Control Unit），利用各种接口和线路连接整个电子设

备的各个部件，通过运行或执行存储在所述存储器11内的程序或者模块（例如基于音频的用户状态识别程序等），以及调用存储在所述存储器11内的数据，以执行电子设备1的各种功能和处理数据。

[0108] 所述总线可以是外设部件互连标准（peripheral component interconnect，简称PCI）总线或扩展工业标准结构（extended industry standard architecture，简称EISA）总线等。该总线可以分为地址总线、数据总线、控制总线等。所述总线被设置为实现所述存储器11以及至少一个处理器10等之间的连接通信。

[0109] 图4仅示出了具有部件的电子设备，本领域技术人员可以理解的是，图4示出的结构并不构成对所述电子设备1的限定，可以包括比图示更少或者更多的部件，或者组合某些部件，或者不同的部件布置。

[0110] 例如，尽管未示出，所述电子设备1还可以包括给各个部件供电的电源（比如电池），优选地，电源可以通过电源管理装置与所述至少一个处理器10逻辑相连，从而通过电源管理装置实现充电管理、放电管理、以及功耗管理等功能。电源还可以包括一个或一个以上的直流或交流电源、再充电装置、电源故障检测电路、电源转换器或者逆变器、电源状态指示器等任意组件。所述电子设备1还可以包括多种传感器、蓝牙模块、Wi-Fi模块等，在此不再赘述。

[0111] 进一步地，所述电子设备1还可以包括网络接口，可选地，所述网络接口可以包括有线接口和/或无线接口（如WI-FI接口、蓝牙接口等），通常用于在该电子设备1与其他电子设备之间建立通信连接。

[0112] 可选地，该电子设备1还可以包括用户接口，用户接口可以是显示器（Display）、输入单元（比如键盘（Keyboard）），可选地，用户接口还可以是标准的有线接口、无线接口。可选地，在一些实施例中，显示器可以是LED显示器、液晶显示器、触控式液晶显示器以及OLED（Organic Light-Emitting Diode，有机发光二极管）触摸器等。其中，显示器也可以适当的称为显示屏或显示单元，用于显示在电子设备1中处理的信息以及用于显示可视化的用户界面。

[0113] 应该了解，所述实施例仅为说明之用，在专利申请范围上并不受此结构的限制。

- [0114] 所述电子设备1中的所述存储器11存储的基于音频的用户状态识别程序12是多个指令的组合，在所述处理器10中运行时，可以实现：
- [0115] 获取音频训练集，对所述音频训练集中的每个音频进行特征转换，得到目标声谱图集；
- [0116] 基于注意力机制与小样本学习，利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型；
- [0117] 当接收到待识别用户的音频时，对所述待识别用户的音频进行特征转换，得到待识别声谱图；
- [0118] 利用所述用户状态识别模型对所述待识别声谱图进行识别，得到用户状态识别结果。
- [0119] 具体地，所述处理器10对上述指令的具体实现方法可参考图1对应实施例中相关步骤的描述，在此不赘述。
- [0120] 进一步地，所述电子设备1集成的模块/单元如果以软件功能单元的形式实现并作为独立的产品销售或使用，可以存储在一个计算机可读取存储介质中，所述计算机可读存储介质可以是易失性，也可以是非易失性。所述计算机可读介质可以包括：能够携带所述计算机程序代码的任何实体或装置、记录介质、U盘、移动硬盘、磁碟、光盘、计算机存储器、只读存储器（ROM，Read-Only Memory）。
- [0121] 所述计算机可读存储介质中存储有计算机程序，其中，所述计算机程序被处理器执行时实现如下步骤：
- [0122] 获取音频训练集，对所述音频训练集中的每个音频进行特征转换，得到目标声谱图集；
- [0123] 基于注意力机制与小样本学习，利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型；
- [0124] 当接收到待识别用户的音频时，对所述待识别用户的音频进行特征转换，得到待识别声谱图；
- [0125] 利用所述用户状态识别模型对所述待识别声谱图进行识别，得到用户状态识别结果。

[0126] 进一步地，所述计算机可用存储介质可主要包括存储程序区和存储数据区，其中，存储程序区可存储操作系统、至少一个功能所需的应用程序等；存储数据区可存储根据区块链节点的使用所创建的数据等。

[0127] 在本申请所提供的几个实施例中，应该理解到，所揭露的设备，装置和方法，可以通过其它的方式实现。例如，以上所描述的装置实施例仅仅是示意性的，例如，所述模块的划分，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。

[0128] 所述作为分离部件说明的模块可以是或者也可以不是物理上分开的，作为模块显示的部件可以是或者也可以不是物理单元，即可以位于一个地方，或者也可以分布到多个网络单元上。可以根据实际的需要选择其中的部分或者全部模块来实现本实施例方案的目的。

[0129] 另外，在本申请各个实施例中的各功能模块可以集成在一个处理单元中，也可以是各个单元单独物理存在，也可以两个或两个以上单元集成在一个单元中。上述集成的单元既可以采用硬件的形式实现，也可以采用硬件加软件功能模块的形式实现。

[0130] 对于本领域技术人员而言，显然本申请不限于上述示范性实施例的细节，而且在不背离本申请的精神或基本特征的情况下，能够以其他的具体形式实现本申请。

[0131] 因此，无论从哪一点来看，均应将实施例看作是示范性的，而且是非限制性的，本申请的范围由所附权利要求而不是上述说明限定，因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本申请内。不应将权利要求中的任何附关联图标记视为限制所涉及的权利要求。

[0132] 本申请所指区块链是分布式数据存储、点对点传输、共识机制、加密算法等计算机技术的新型应用模式。区块链(Blockchain)，本质上是一个去中心化的数据库，是一串使用密码学方法相关联产生的数据块，每一个数据块中包含了一批次网络交易的信息，用于验证其信息的有效性(防伪)和生成下一个区块。区块链可以包括区块链底层平台、平台产品服务层以及应用服务层等。

[0133] 此外，显然“包括”一词不排除其他单元或步骤，单数不排除复数。系统权利

要求中陈述的多个单元或装置也可以由一个单元或装置通过软件或者硬件来实现。第二等词语用来表示名称，而并不表示任何特定的顺序。

[0134] 最后应说明的是，以上实施例仅用以说明本申请的技术方案而非限制，尽管参照较佳实施例对本申请进行了详细说明，本领域的普通技术人员应当理解，可以对本申请的技术方案进行修改或等同替换，而不脱离本申请技术方案的精神和范围。

权利要求书

- [权利要求 1] 一种基于音频的用户状态识别方法，其中，所述方法包括：
- 获取音频训练集，对所述音频训练集中的每个音频进行特征转换，得到目标声谱图集；
- 基于注意力机制与小样本学习，利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型；
- 当接收到待识别用户的音频时，对所述待识别用户的音频进行特征转换，得到待识别声谱图；
- 利用所述用户状态识别模型对所述待识别声谱图进行识别，得到用户状态识别结果。
- [权利要求 2] 如权利要求1所述的基于音频的用户状态识别方法，其中，所述对所述音频训练集中的每个音频进行特征转换，得到目标声谱图集，包括：
- 将所述音频训练集中的每个音频进行重采样，得到对应的数字语音信号；
- 对所述数字语音信号进行预加重，得到标准数字语音信号；
- 汇总所有的所述标准数字语音信号，得到语音信号集；
- 对所述语音信号集中包含的每个标准数字语音信号进行特征转换，得到目标声谱图集。
- [权利要求 3] 如权利要求2所述的基于音频的用户状态识别方法，其中，所述对所述语音信号集中包含的每个标准数字语音信号进行特征转换，得到目标声谱图集，包括：
- 利用预设声音处理算法，将所述语音信号集中的每个标准数字语音信号映射在频域，得到对应的目标声谱图；
- 汇总所有的所述目标声谱图得到所述目标声谱图集。
- [权利要求 4] 如权利要求1所述的基于音频的用户状态识别方法，其中，所述利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型，包括：

将所述目标声谱图集随机划分为训练集及测试集；

利用所述训练集对所述深度学习网络模型进行训练得到初始识别模型

；

根据所述测试集对所述初始识别模型进行测试得到损失值；

当所述损失值大于预设阈值时返回所述将所述目标声谱图集随机划分为训练集及测试集步骤；

当所述损失值小于或等于预设阈值时，将所述初始识别模型作为用户状态识别模型。

[权利要求 5]

如权利要求4所述的基于音频的用户状态识别方法，其中，所述将所述目标声谱图集随机划分为训练集及测试集，包括：

将所述目标声谱图集集中的每个目标声谱图按照对应的初始标签进行分类，得到对应的分类目标声谱图集；

从所述分类目标声谱图集中随机取出预设数量的目标声谱图作为测试子集，将所述分类声谱图集中所述测试子集的补集作为训练子集；

汇总所有的所述训练子集得到训练集；

汇总所有的所述测试子集得到测试集。

[权利要求 6]

如权利要求5所述的基于音频的用户状态识别方法，其中，所述根据所述测试集对所述初始识别模型进行测试得到损失值，包括：

提取所述初始识别模型中每种所述初始标签对应的特征向量，得到目标特征向量；

利用所述初始识别模型对所述测试集中的每个目标声谱图进行特征提取，得到对应的测试特征向量；

计算每种所述初始标签对应的所述目标特征向量与所述测试特征向量的距离，得到损失距离值；

计算所有所述损失距离值的平均值，得到所述损失值。

[权利要求 7]

如权利要求1至6中任何一项所述的基于音频的用户状态识别方法，其中，所述音频训练集为不同的疾病情况对应的咳嗽音频的集合。

[权利要求 8]

一种基于音频的用户状态识别装置，其中，所述装置包括：

模型生成模块，用于获取音频训练集，对所述音频训练集中的每个音频进行特征转换，得到目标声谱图集；基于注意力机制与小样本学习，利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型；

状态识别模块，用于当接收到待识别用户的音频时，对所述待识别用户的音频进行特征转换，得到待识别声谱图；利用所述用户状态识别模型对所述待识别声谱图进行识别，得到用户状态识别结果。

[权利要求 9]

一种电子设备，其中，所述电子设备包括：

至少一个处理器；以及，

与所述至少一个处理器通信连接的存储器；其中，

所述存储器存储有可被所述至少一个处理器执行的指令，所述指令被所述至少一个处理器执行，以使所述至少一个处理器能够执行如下步骤：

获取音频训练集，对所述音频训练集中的每个音频进行特征转换，得到目标声谱图集；

基于注意力机制与小样本学习，利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型；

当接收到待识别用户的音频时，对所述待识别用户的音频进行特征转换，得到待识别声谱图；

利用所述用户状态识别模型对所述待识别声谱图进行识别，得到用户状态识别结果。

[权利要求 10]

如权利要求9所述的电子设备，其中，所述对所述音频训练集中的每个音频进行特征转换，得到目标声谱图集，包括：

将所述音频训练集中的每个音频进行重采样，得到对应的数字语音信号；

对所述数字语音信号进行预加重，得到标准数字语音信号；

汇总所有的所述标准数字语音信号，得到语音信号集；

对所述语音信号集中包含的每个标准数字语音信号进行特征转换，得

到目标声谱图集。

[权利要求 11] 如权利要求10所述的电子设备，其中，所述对所述语音信号集中包含的每个标准数字语音信号进行特征转换，得到目标声谱图集，包括：利用预设声音处理算法，将所述语音信号集中的每个标准数字语音信号映射在频域，得到对应的目标声谱图；
汇总所有的所述目标声谱图得到所述目标声谱图集。

[权利要求 12] 如权利要求9所述的电子设备，其中，所述利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型，包括：
：
将所述目标声谱图集随机划分为训练集及测试集；
利用所述训练集对所述深度学习网络模型进行训练得到初始识别模型；
；
根据所述测试集对所述初始识别模型进行测试得到损失值；
当所述损失值大于预设阈值时返回所述将所述目标声谱图集随机划分为训练集及测试集步骤；
当所述损失值小于或等于预设阈值时，将所述初始识别模型作为用户状态识别模型。

[权利要求 13] 如权利要求12所述的电子设备，其中，所述将所述目标声谱图集随机划分为训练集及测试集，包括：
将所述目标声谱图集集中的每个目标声谱图按照对应的初始标签进行分类，得到对应的分类目标声谱图集；
从所述分类目标声谱图集中随机取出预设数量的目标声谱图作为测试子集，将所述分类声谱图集中所述测试子集的补集作为训练子集；
汇总所有的所述训练子集得到训练集；
汇总所有的所述测试子集得到测试集。

[权利要求 14] 如权利要求13所述的电子设备，其中，所述根据所述测试集对所述初始识别模型进行测试得到损失值，包括：
提取所述初始识别模型中每种所述初始标签对应的特征向量，得到目

标特征向量；

利用所述初始识别模型对所述测试集中的每个目标声谱图进行特征提取，得到对应的测试特征向量；

计算每种所述初始标签对应的所述目标特征向量与所述测试特征向量的距离，得到损失距离值；

计算所有所述损失距离值的平均值，得到所述损失值。

[权利要求 15] 如权利要求9至14中任何一项所述的电子设备，其中，所述音频训练集为不同的疾病情况对应的咳嗽音频的集合。

[权利要求 16] 一种计算机可读存储介质，存储有计算机程序，其中，所述计算机程序被处理器执行时实现如下步骤：

获取音频训练集，对所述音频训练集中的每个音频进行特征转换，得到目标声谱图集；

基于注意力机制与小样本学习，利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型；

当接收到待识别用户的音频时，对所述待识别用户的音频进行特征转换，得到待识别声谱图；

利用所述用户状态识别模型对所述待识别声谱图进行识别，得到用户状态识别结果。

[权利要求 17] 如权利要求16所述的计算机可读存储介质，其中，所述对所述音频训练集中的每个音频进行特征转换，得到目标声谱图集，包括：

将所述音频训练集中的每个音频进行重采样，得到对应的数字语音信号；

对所述数字语音信号进行预加重，得到标准数字语音信号；

汇总所有的所述标准数字语音信号，得到语音信号集；

对所述语音信号集中包含的每个标准数字语音信号进行特征转换，得到目标声谱图集。

[权利要求 18] 如权利要求17所述的计算机可读存储介质，其中，所述对所述语音信号集中包含的每个标准数字语音信号进行特征转换，得到目标声谱图

集，包括：

利用预设声音处理算法，将所述语音信号集中的每个标准数字语音信号映射在频域，得到对应的目标声谱图；

汇总所有的所述目标声谱图得到所述目标声谱图集。

[权利要求 19] 如权利要求16所述的计算机可读存储介质，其中，所述利用所述目标声谱图集对预构建的深度学习网络模型进行训练，得到用户状态识别模型，包括：

将所述目标声谱图集随机划分为训练集及测试集；

利用所述训练集对所述深度学习网络模型进行训练得到初始识别模型；

根据所述测试集对所述初始识别模型进行测试得到损失值；

当所述损失值大于预设阈值时返回所述将所述目标声谱图集随机划分为训练集及测试集步骤；

当所述损失值小于或等于预设阈值时，将所述初始识别模型作为用户状态识别模型。

[权利要求 20] 如权利要求19所述的计算机可读存储介质，其中，所述将所述目标声谱图集随机划分为训练集及测试集，包括：

将所述目标声谱图集集中的每个目标声谱图按照对应的初始标签进行分类，得到对应的分类目标声谱图集；

从所述分类目标声谱图集中随机取出预设数量的目标声谱图作为测试子集，将所述分类声谱图集中所述测试子集的补集作为训练子集；

汇总所有的所述训练子集得到训练集；

汇总所有的所述测试子集得到测试集。

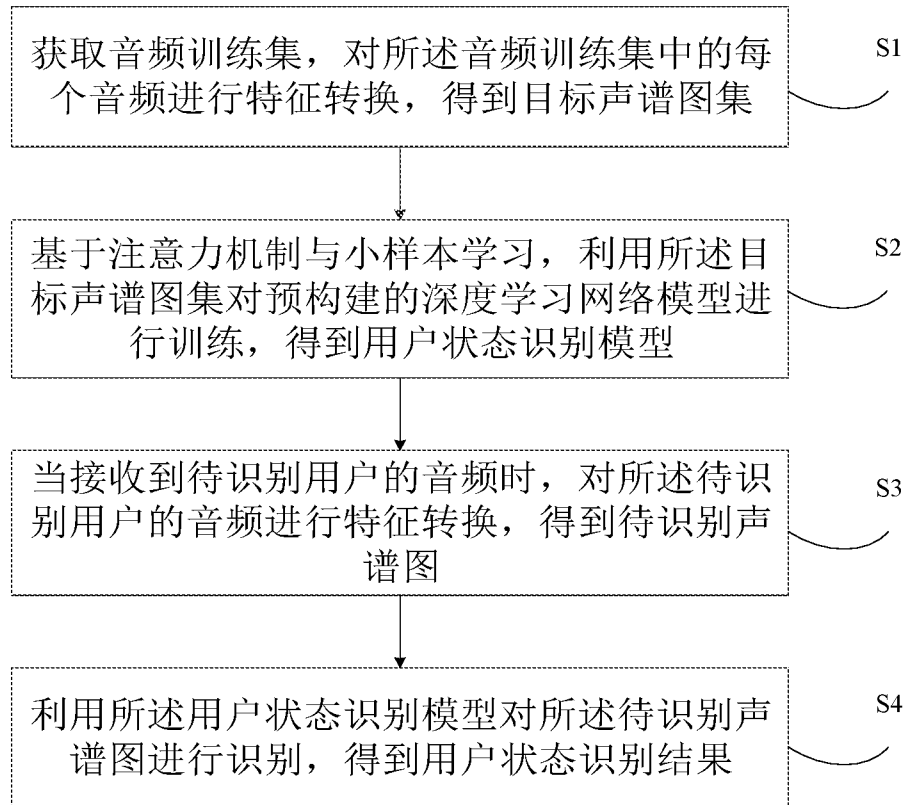


图 1

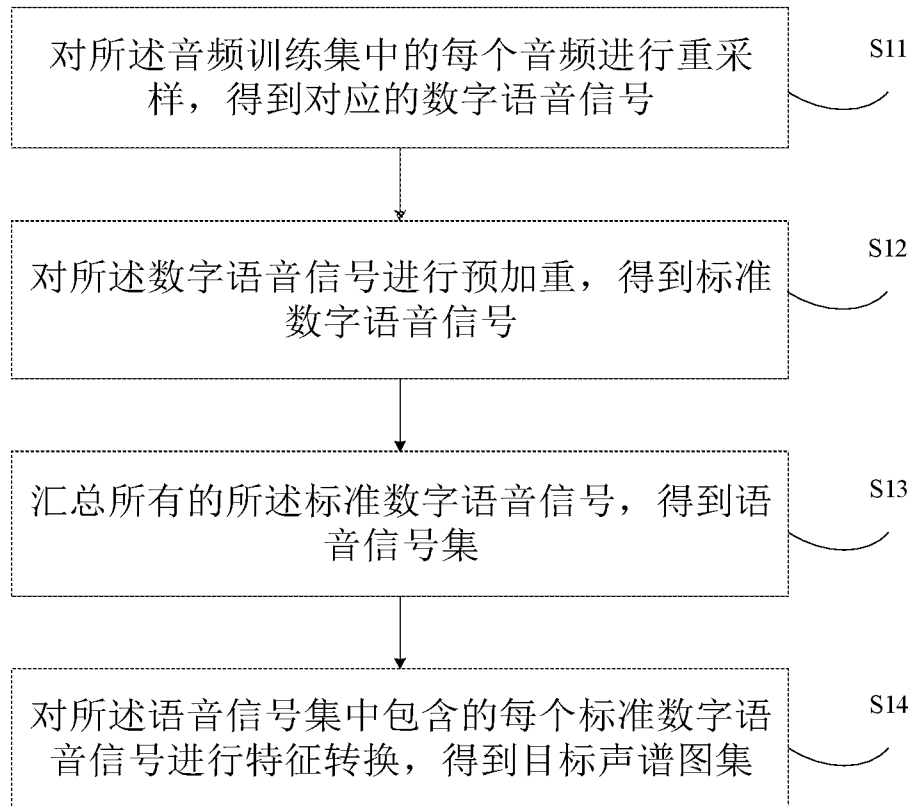


图 2

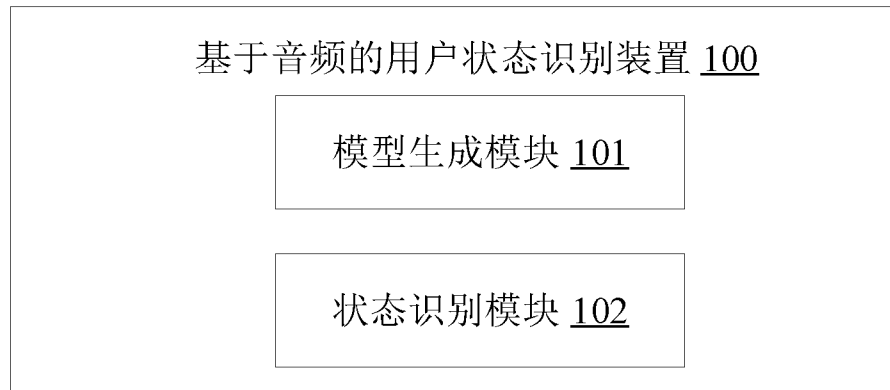


图 3



图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2020/131983

A. CLASSIFICATION OF SUBJECT MATTER

G10L 25/66(2013.01)i; G10L 25/18(2013.01)i; G10L 25/27(2013.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNPAT, IEEE, CNKI, 百度: 音频, 声音, 语音, 特征, 模型, 训练, 咳嗽, 疾病, 诊断, 注意力机制, voice, audio, speech, feature, model, train, cough, disease, illness, diagnosis, Attention Mechanism

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	CN 104321015 A (THE UNIVERSITY OF QUEENSLAND) 28 January 2015 (2015-01-28) description paragraphs [0090]-[0222], [0287], [0290], claims 13, 15	1-20
A	CN 106073706 A (INSTITUTE OF SOFTWARE, CHINESE ACADEMY OF SCIENCES) 09 November 2016 (2016-11-09) entire document	1-20
A	CN 106202952 A (NANJING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) 07 December 2016 (2016-12-07) entire document	1-20
A	CN 106847262 A (HUAZHONG AGRICULTURAL UNIVERSITY) 13 June 2017 (2017-06-13) entire document	1-20
A	WO 2019023879 A1 (SHENZHEN H&T SMART HOME TECHNOLOGY CO., LTD.) 07 February 2019 (2019-02-07) entire document	1-20
A	WO 2019119050 A1 (THE UNIVERSITY OF QUEENSLAND) 27 June 2019 (2019-06-27) entire document	1-20



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

11 May 2021

Date of mailing of the international search report

12 July 2021

Name and mailing address of the ISA/CN

China National Intellectual Property Administration (ISA/
CN)
No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2020/131983

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	104321015	A	28 January 2015	JP	2015514456	A	21 May 2015
				KR	20140142330	A	11 December 2014
				AU	2013239327	B2	23 August 2018
				US	2015073306	A1	12 March 2015
				WO	2013142908	A1	03 October 2013
				EP	2830496	A1	04 February 2015
				AU	2013239327	A1	06 November 2014
				EP	2830496	A4	16 December 2015
				JP	6435257	B2	05 December 2018
				US	10098569	B2	16 October 2018
				CN	110353685	A	22 October 2019
				KR	2081241	B1	25 February 2020
CN	106073706	A	09 November 2016	CN	106073706	B	20 August 2019
CN	106202952	A	07 December 2016	None			
CN	106847262	A	13 June 2017	None			
WO	2019023879	A1	07 February 2019	CN	108701469	A	23 October 2018
WO	2019119050	A1	27 June 2019	EP	3727148	A1	28 October 2020
				EP	3727148	A4	03 March 2021
				KR	20200122301	A	27 October 2020
				AU	2018386721	A1	16 July 2020
				CN	111629663	A	04 September 2020
				JP	2021506486	A	22 February 2021
				IN	202047029281	A	28 August 2020

国际检索报告

国际申请号

PCT/CN2020/131983

A. 主题的分类

G10L 25/66(2013.01)i; G10L 25/18(2013.01)i; G10L 25/27(2013.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G10L

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称, 和使用的检索词(如使用))

WPI, EPDOC, CNPAT, IEEE, CNKI, 百度:音频, 声音, 语音, 特征, 模型, 训练, 咳嗽, 疾病, 诊断, 注意力机制, voice, audio, speech, feature, model, train, cough, disease, illness, diagnosis, Attention Mechanism

C. 相关文件

类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
X	CN 104321015 A (昆山大学) 2015年 1月 28日 (2015 - 01 - 28) 说明书第[0090]-[0222], [0287], [0290]段, 权利要求13, 15	1-20
A	CN 106073706 A (中国科学院软件研究所) 2016年 11月 9日 (2016 - 11 - 09) 全文	1-20
A	CN 106202952 A (南京邮电大学) 2016年 12月 7日 (2016 - 12 - 07) 全文	1-20
A	CN 106847262 A (华中农业大学) 2017年 6月 13日 (2017 - 06 - 13) 全文	1-20
A	W0 2019023879 A1 (深圳和而泰智能家居科技有限公司) 2019年 2月 7日 (2019 - 02 - 07) 全文	1-20
A	W0 2019119050 A1 (THE UNIVERSITY OF QUEENSLAND) 2019年 6月 27日 (2019 - 06 - 27) 全文	1-20

☐ 其余文件在C栏的续页中列出。☒ 见同族专利附件。

* 引用文件的具体类型:

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件(如具体说明的)

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性

“&” 同族专利的文件

国际检索实际完成的日期

2021年 5月 11日

国际检索报告邮寄日期

2021年 7月 12日

ISA/CN的名称和邮寄地址

中国国家知识产权局(ISA/CN)

中国北京市海淀区蓟门桥西土城路6号 100088

传真号 (86-10)62019451

受权官员

李易玮

电话号码 86-(10)-53961395

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2020/131983

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	104321015	A	2015年 1月 28日	JP	2015514456	A	2015年 5月 21日
				KR	20140142330	A	2014年 12月 11日
				AU	2013239327	B2	2018年 8月 23日
				US	2015073306	A1	2015年 3月 12日
				WO	2013142908	A1	2013年 10月 3日
				EP	2830496	A1	2015年 2月 4日
				AU	2013239327	A1	2014年 11月 6日
				EP	2830496	A4	2015年 12月 16日
				JP	6435257	B2	2018年 12月 5日
				US	10098569	B2	2018年 10月 16日
				CN	110353685	A	2019年 10月 22日
				KR	2081241	B1	2020年 2月 25日
CN	106073706	A	2016年 11月 9日	CN	106073706	B	2019年 8月 20日
CN	106202952	A	2016年 12月 7日	无			
CN	106847262	A	2017年 6月 13日	无			
WO	2019023879	A1	2019年 2月 7日	CN	108701469	A	2018年 10月 23日
WO	2019119050	A1	2019年 6月 27日	EP	3727148	A1	2020年 10月 28日
				EP	3727148	A4	2021年 3月 3日
				KR	20200122301	A	2020年 10月 27日
				AU	2018386721	A1	2020年 7月 16日
				CN	111629663	A	2020年 9月 4日
				JP	2021506486	A	2021年 2月 22日
				IN	202047029281	A	2020年 8月 28日