



(12) 发明专利申请

(10) 申请公布号 CN 115955932 A

(43) 申请公布日 2023.04.11

(21) 申请号 202180048396.5

(22) 申请日 2021.07.08

(30) 优先权数据

10-2020-0085449 2020.07.10 KR

10-2021-0089014 2021.07.07 KR

(85) PCT国际申请进入国家阶段日

2023.01.06

(86) PCT国际申请的申请数据

PCT/KR2021/008710 2021.07.08

(87) PCT国际申请的公布数据

W02022/010282 KO 2022.01.13

(71) 申请人 伊莫克有限公司

地址 韩国首尔

(72) 发明人 李准荣 高炫雄

(74) 专利代理机构 北京英赛嘉华知识产权代理有限公司 11204

专利代理师 王达佐 王艳春

(51) Int.Cl.

A61B 5/00 (2006.01)

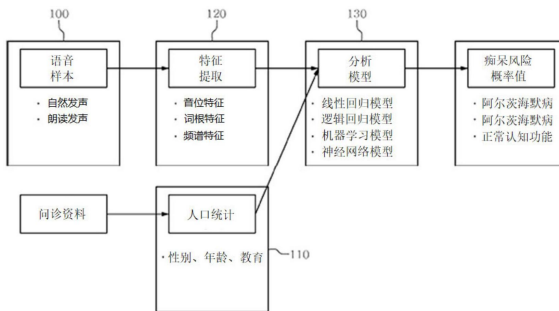
权利要求书2页 说明书8页 附图4页

(54) 发明名称

基于语音特征的阿尔茨海默病预测方法和装置

(57) 摘要

本发明提供一种基于语音特征的阿尔茨海默病预测方法和装置。根据实施例的阿尔茨海默病预测装置包括：语音输入部，被配置为录制受试者的语音，以生成语音样本；数据输入部，被配置为接收所述受试者的人口统计信息；语音特征提取部，从生成的所述语音样本中提取语音特征；以及预测模型，以能够基于所述语音特征和所述人口统计信息预测所述受试者是否患有阿尔茨海默病的方式预先学习。



1. 一种阿尔茨海默病预测装置,其特征在于,包括:
语音输入部,被配置为录制受试者的语音,以生成语音样本;
数据输入部,被配置为接收所述受试者的人口统计信息;
语音特征提取部,从生成的所述语音样本中提取语音特征;以及
预测模型,以能够基于所述语音特征和所述人口统计信息预测所述受试者是否患有阿尔茨海默病的方式预先学习。

2. 根据权利要求1所述的阿尔茨海默病预测装置,其特征在于,
所述人口统计信息包括受试者的年龄、性别和受教育年限。

3. 根据权利要求1所述的阿尔茨海默病预测装置,其特征在于,
所述语音特征提取部提取语音的基频、发声速度、发声时间、发声长度、停顿程度、停顿次数、停顿区间长度、振幅微扰、抖动、共振峰、谐波噪声比、响度、频谱质心梅尔频率倒谱系数、身份向量、发音速度、过零率、发音概率、线谱对、周期扰动、振幅扰动商、劲度、能量、强度(声音大小)、熵中的至少一种作为所述语音特征。

4. 根据权利要求3所述的阿尔茨海默病预测装置,其特征在于,
所述语音特征提取部包括人工神经网络模型,所述人工神经网络模型执行从所述语音样本筛选人的语音的预处理,
所述语音特征提取部从预处理的语音样本提取所述语音特征。

5. 根据权利要求1所述的阿尔茨海默病预测装置,其特征在于,
所述预测模型包括线性回归模型、逻辑回归模型、机器学习模型和神经网络模型中的至少一种分析模型。

6. 根据权利要求5所述的阿尔茨海默病预测装置,其特征在于,
所述逻辑回归模型为多变量逻辑回归模型,所述多变量逻辑回归模型被配置为如下式1所示:

[式1]

$$\log \left(\frac{p(X)}{1-p(X)} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_p X_p$$

其中, X_1 至 X_p 是作为自变量输入到预测模型的输入值,分别对应于 p 个语音特征和人口统计信息, β_1 至 β_p 对应于作为自变量的回归系数的常数值, β_0 对应于初始常数值, $p(X)$ 对应于痴呆风险概率值。

7. 根据权利要求6所述的阿尔茨海默病预测装置,其特征在于,
所述预测模型输出痴呆风险概率值,并输出基于所述痴呆风险概率值评估患者状态的状态信息,

所述状态信息被确定为阿尔茨海默病或正常认知功能。

8. 一种基于语音特征的阿尔茨海默病预测方法,其特征在于,包括以下步骤:

录制受试者的语音,以生成语音样本;

接收所述受试者的人口统计信息;

从生成的所述语音样本中提取语音特征;以及

将所述语音特征和所述人口统计信息输入到预先学习的预测模型,以预测所述受试者

是否患有阿尔茨海默病。

9. 一种存储计算机可读指令的计算机可读记录介质,其特征在于,
当所述指令由至少一个处理器执行时,使所述至少一个处理器执行以下步骤:
录制受试者的语音,以生成语音样本;
接收所述受试者的人口统计信息;
从生成的所述语音样本中提取语音特征;以及
将所述语音特征和所述人口统计信息输入到预先学习的预测模型,以预测所述受试者是否患有阿尔茨海默病。

基于语音特征的阿尔茨海默病预测方法和装置

技术领域

[0001] 本发明涉及一种作为用于诊断阿尔茨海默病 (Alzheimer's disease, AD) 的方法, 该疾病在痴呆症的致病性疾病中占最高比例, 通过使用患者的语音的非侵入性方法来帮助医务人员判断是否患有阿尔茨海默病的方法, 可以提供一种能够通过非侵入性方法预测具有高痴呆风险的人由于阿尔茨海默氏病引起的痴呆状态的方法和装置。

背景技术

[0002] AI扬声器的使用比例随着第四产业逐渐增加, 但是至今尚未对大众普及。有关AI扬声器的普及滞后的原因有很多, 但大多数消费者认为AI扬声器的内容不够丰富。即, 需要开发使用AI扬声器的各种内容。

[0003] 其中, AI扬声器和医疗保健行业被公认为兼顾盈利和公益的“一石二鸟”, 国内外许多研究团队正在使用AI扬声器开发与医疗保健相关的各种内容。

[0004] 截至2000年, 韩国65岁以上的老年人口占总人口的7.1%, 预计到2026年将超过20%。以具有典型的老年病痴呆症为例, 2000年统计为28万名, 占65岁以上老年人的8.3%, 预计2010年将达到43万名, 占老年人总数的8.6%, 并呈现上升趋势。由于随着老年人口的增加预期寿命增加, 痴呆症的患病率呈现出稳步上升的趋势。

[0005] 阿尔茨海默病 (Alzheimer's disease, AD) 是最常见的致病性疾病, 约占所有痴呆症的70%, 预计2020年将有约60万名阿尔茨海默病患者出现。阿尔茨海默病是一种典型的退行性痴呆疾病, 它通过大脑神经细胞的退化导致认知功能和日常生活能力下降, 是一种导致死亡的破坏性疾病, 不仅给患者带来痛苦, 而且给照顾患者的家人带来痛苦。由于阿尔茨海默病是一种不可逆转的疾病, 尚未开发出可治愈的治疗方法, 因此对阿尔茨海默病进行早期诊断和早期干预是目前最好的解决方案。随着阿尔茨海默病患病率增加, 人民保健和医疗费用支出方面也受到了巨大影响。

[0006] 语音和语言障碍是由于认知功能和神经退化引起的痴呆症的伴随症状之一, 经常被用作区分正常衰老和导致痴呆症的疾病的标准, 并且经常作为早期症状的标志物出现。为此, 在自动语音识别的研究中, 正在进行各种研究以发现用于诊断痴呆症的语音和语言标记。典型地, 主要进行来自DementiaBank的Pitt语料库 (corpus) 的数据 (例如, Cookie窃取插图) 预测痴呆症的研究。

[0007] 考虑到用于准确诊断阿尔茨海默病 (一种典型的导致痴呆症的疾病) 的脑脊液检查或正电子发射断层扫描 (Positron emission tomograph, PET) 被专科医生限制在特定的空间和时间, 需要一种可以使用AI扬声器和智能设备预测对话受试者是否患有阿尔茨海默病的方法和装置。

发明内容

[0008] 本发明要解决的技术问题

[0009] 本发明涉及一种利用语音的声学特征诊断阿尔茨海默病的方法, 通过语音识别麦

克风并使用AI扬声器或智能设备来收集对话受试者的应答,从而能够预测对话受试者是否患有阿尔茨海默病和可能性方法和装置。

[0010] 技术方案

[0011] 根据本发明的实施例的基于语音特征的阿尔茨海默病预测装置包括:语音输入部,被配置为录制受试者的语音,以生成语音样本;数据输入部,被配置为接收所述受试者的人口统计信息;语音特征提取部,被配置为从生成的所述语音样本中提取语音特征;以及预测模型,以能够基于所述语音特征和所述人口统计信息预测所述受试者是否患有阿尔茨海默病的方式预先学习。

[0012] 根据本发明的实施例的基于语音特征的阿尔茨海默病预测方法包括以下步骤:录制受试者的语音,以生成语音样本;接收所述受试者的人口统计信息;从生成的所述语音样本中提取语音特征;以及将所述语音特征和所述人口统计信息输入到预先学习的预测模型,预测所述受试者是否患有阿尔茨海默病。

[0013] 根据本发明的实施例的记录介质是一种存储计算机可读指令的计算机可读记录介质,当所述指令由至少一个处理器执行时,使所述至少一个处理器执行以下步骤:录制受试者的语音,以生成语音样本;接收所述受试者的人口统计信息;从生成的所述语音样本中提取语音特征;将所述语音特征和所述人口统计信息输入到预先学习的预测模型,预测所述受试者是否患有阿尔茨海默病。

[0014] 有益效果

[0015] 根据本发明的实施例的基于语音特征的阿尔茨海默病预测方法和装置是一种利用AI扬声器和智能设备来预测对话受试者是否患有阿尔茨海默病的方法和装置,有望帮助克服伴随现有的检查的时间、空间、资源上的限制,可以支援医务人员诊断阿尔茨海默病。

[0016] 通过利用本发明提供的语音的声学特征的阿尔茨海默病风险度预测,可以支持为不断增加的痴呆人群制定适当的干预计划。

附图说明

[0017] 图1是示出了根据本发明的实施例的基于语音特征的阿尔茨海默病预测装置的配置的框图。

[0018] 图2示例性示出了根据本发明的实施例的基于语音特征的阿尔茨海默病预测装置的操作过程。

[0019] 图3是示出了根据本发明的实施例的基于语音特征的阿尔茨海默病预测方法的顺序的框图。

[0020] 图4是示出了考虑到人口统计学和语音特征的阿尔茨海默病预测结果的图表。

具体实施方式

[0021] 下面将结合附图对本发明的优选实施例进行描述。参考附图中示出的实施例说明了本发明,但仅是作为一个实施例进行了说明,并不会因此而限制本发明的技术思想、其核心构成及作用。

[0022] 图1是示出本发明的实施例的阿尔茨海默病预测装置的配置的框图。图2示例性示出了根据本发明的实施例的阿尔茨海默病预测装置的操作过程。

[0023] 参见图1和图2,根据本发明的实施例的阿尔茨海默病预测装置10包括语音输入部100、数据输入部110、语音特征提取部120、预测模型130和数据存储部140。

[0024] 阿尔茨海默病预测装置10可以具有完全是硬件或具有部分是硬件、部分是软件的方面。例如,本说明书的阿尔茨海默病预测装置10和包括在其中的各部分可以统称为用于以电子通信方式发送和接收具有特定形式和内容的数据的装置和与此相关的软件。本说明书中的“部”、“模块(module)”、“服务器(server)”、“装置”、“设备”或“终端”等术语是指硬件和由相应硬件驱动的软件的组合。例如,这里的硬件可以是包括CPU或其他处理器(processor)的数据处理设备。另外,由硬件驱动的软件可以是指执行中的程序、受试者(object)、执行文件(executable)、执行线程(thread of execution)、程序(program)等。

[0025] 另外,构成阿尔茨海默病预测装置10的每个模块不一定旨在指代物理上区分的单独的组件。在图1中,语音输入部100、数据输入部110、语音特征提取部120、预测模型130和数据存储部140被示为彼此区分的单独的块,但这是将构成阿尔茨海默病预测装置10的装置仅在功能上由相应装置执行的操作划分的。因此,根据实施例,语音输入部100、数据输入部110、语音特征提取部120、预测模型130和数据存储部140中的一部分或全部可以集成在同一个装置内。例如,阿尔茨海默病预测装置10可以通过具有如AI扬声器的语音识别和数据处理能力的装置实现。但不限于此,所述构成中一个以上可以实现为与其他部分在物理上区分的单独的装置,或者也可以是在分布式计算环境中相互通信连接的组件。

[0026] 语音输入部100生成受试者的语音样本。语音输入部100可以被配置为通过以预定频率录制受试者的语音来生成语音样本。语音输入部100可以包括电容式麦克风和其控制装置,并且可以通过以16Hz以上的频率录制受试者的语音来生成语音样本。受试者将在可录音的安静的房间里进行图片描述、标准段落发声和故事回忆任务中的至少一项,并进行自然发声或朗读发声,并且语音输入部100通过录制这些受试者的发声来生成语音样本。

[0027] 数据输入部110接收受试者的人口统计信息。人口统计信息至少包括受试者的年龄、性别、受教育年限(受教育的学历程度)。通过数据输入部110输入的资料是指可以确认受试者的年龄、性别、受教育年限或者是指可以提取这些信息的数据。示例性地,人口统计信息可以通过受试者的问诊资料获取。问诊资料可以包括与患者的年龄、性别、受教育年限相关的信息,并且对应于利用经过培训的专科医生的医学判断而生成的资料,也可以是在其他医务人员和患者监护人的管理下收集的资料。数据输入部110可以通过所述问诊资料接收人口统计信息,但不限于此。在一些实施例中,受试者的人口统计信息可以是受试者直接输入的信息。

[0028] 通过语音输入部100和数据输入部110输入的数据可以存储在数据存储部140中。数据存储部140可以被配置为存储已输入的数据或提供后述的预测模型的数据处理所需的临时或临时存储空间。

[0029] 语音特征提取部120可以从已输入的受试者的语音样本提取患者的语音特征。语音特征提取部120可以提取与语音样本的音位特征、词根特征、频谱特征相关的语音特征。具体地,语音特征提取部120可以提取受试者语音的基频(fundamental frequency)、与发声有关的信息(发声速度、发声时间、发声长度)、与停顿相关的信息(停顿程度、停顿次数、停顿区间长度)、振幅微扰(Shimmer)、抖动(Jitter)、共振峰(formant)、谐波噪声比(harmonic-to-noise ratio)、响度(loudness)、频谱质心(spectral centroid)梅尔频率

倒谱系数(Mel Frequency Cepstral Coefficients,MFCC)、身份向量(i-vector)、发音速度、过零率(zero-crossing rate,zcr)、发音概率(voicing probability,vp)、线谱对(Line spectral paris,LSP)、周期扰动(Period perturbation)、振幅扰动商(Amplitude perturbation quotient,APQ)、劲度(Stiffness)、能量(Energy)、强度(声音大小, Intensity)、熵(Entropy)中的至少一种作为语音特征。

[0030] 其中,语音特征提取部120可以先执行用于量化语音样本的预处理过程。可以通过预处理过程将语音样本的时间、频率等调整为恒定。另外,可以通过预处理过程区分人声和非人声。语音特征提取部120可以包括人工神经网络模型(例如,卷积神经网络),其通过学习以从输入语音样本中仅筛选人声,并且可以通过学习的人工神经网络模型执行用于仅筛选人声的预处理。语音特征提取部120可以从预处理的语音样本提取包括语音的基频(fundamental frequency)、发声速度、发声时间、发声长度、停顿程度、停顿次数、停顿区间长度、振幅微扰、抖动(Jitter)、共振峰(formant)、谐波噪声比(harmonic-to-noise ratio)、响度(loudness)、频谱质心(spectral centroid)梅尔频率倒谱系数(Mel Frequency Cepstral Coefficients,MFCC)、身份向量(i-vector)、发音速度、过零率(zero-crossing rate,zcr)、发音概率(voicing probability,vp)、线谱对(Line spectral paris,LSP)、周期扰动(Period perturbation)、振幅扰动商(Amplitude perturbation quotient,APQ)、劲度(Stiffness)、能量(Energy)、强度(声音大小, Intensity)、熵(Entropy)中的至少一种的语音特征。语音特征提取部120还可以使用已公开的语音特征提取程序(例如,Praat)来提取对应于音韵、词根、频谱区域的语音特征。

[0031] 预测模型130可以处于预先学习的状态以通过语音特征提取部120提取的语音特征和通过数据输入部110提供的人口统计信息来预测受试者是否患有阿尔茨海默病。预测模型130可以包括线性回归模型、逻辑回归模型、机器学习模型和神经网络模型中的至少一种分析模型。

[0032] 预测模型130包括多变量逻辑回归模型,多变量逻辑回归模型可以配置为如下式1所示。

[0033] [式1]

$$[0034] \quad \log \left(\frac{p(X)}{1-p(X)} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_p X_p$$

[0035] (其中, X_1 至 X_p 是作为自变量输入到预测模型的输入值,分别对应于p个语音特征和人口统计信息, β_1 至 β_p 对应于作为自变量的回归系数的常数值, β_0 对应于初始常数值,p(X)对应于痴呆风险概率值)

[0036] 基于所述式1构成的预测模型130可以输出痴呆风险概率值,并输出基于所述痴呆风险概率值评估患者状态的状态信息。痴呆风险概率值是指被诊断为痴呆症的概率,状态信息是根据专科医生诊断标准对患者进行诊断的,患者的状态可以被确定为阿尔茨海默病或正常认知功能。例如,当计算出的痴呆风险概率值p(X)为0.5以上时,预测模型130可以将患者的状态确定为阿尔茨海默病,当计算出的痴呆风险概率值小于0.5时,将患者的状态确定为正常认知功能。

[0037] 但本发明的预测模型130并不限于此。预测模型130可以由多项多变量逻辑回归模型组成,预测模型130还可以基于计算出的痴呆风险概率值将患者的状态确定为阿尔茨海

默病、轻度认知障碍和正常认知功能之一。

[0038] 下面将说明根据本发明的另一实施例的基于语音特征的阿尔茨海默病预测方法。

[0039] 图3是根据实施例的基于语音特征的阿尔茨海默病预测方法流程图。本方法可以由根据图1和图2的预测装置执行,并且对于本实施例的说可以参考图1、图2和相关说明。

[0040] 参见图3,根据实施例的基于语音特征的阿尔茨海默病预测方法包括:录制受试者的语音,以生成语音样本的步骤S100;接收所述受试者的人口统计信息的步骤S110;从生成的所述语音样本中提取语音特征的步骤S120;以及通过将所述语音特征和所述人口统计信息输入到预先学习的预测模型来预测所述受试者是否患有阿尔茨海默病的步骤S130。

[0041] 根据实施例的基于语音特征的阿尔茨海默病预测方法的各步骤中的步骤S100和步骤S110是为了便于说明,按顺序进行记载和说明的,并非局限于记载顺序进行。在一些实施例中,步骤S110可以在步骤S100之前执行。另外,执行所述预测方法之前,可以先执行学习预测模型的步骤。

[0042] 录制受试者的语音,以生成语音样本S100。

[0043] 受试者将在可录音的安静的房间里进行图片描述、标准段落发声和故事回忆任务中的至少一项,并且进行自然发声或朗读发声,并通过录制这些受试者的发声来生成语音样本。语音输入部100可以包括电容式麦克风和其控制装置,并且可以通过以16Hz以上的频率录制受试者的语音来生成语音样本。

[0044] 接收所述受试者的人口统计信息S110。

[0045] 所述人口统计信息包括所述受试者的年龄、性别和受教育年限。示例性地,人口统计信息可以通过受试者的问诊资料获取。问诊资料可以包括与患者的年龄、性别、受教育年限相关的信息,并且对应于利用经过培训的专科医生的医学判断而生成的资料,也可以是在其他医务人员和患者监护人的管理下收集的资料。数据输入部110可以通过所述问诊资料接收人口统计信息,但不限于此。在一些实施例中,受试者的人口统计信息可以是受试者直接输入的信息。

[0046] 接下来,从生成的语音样本提取语音特征S120。

[0047] 可以从输入的受试者的语音样本提取患者的语音特征。用于量化患者的语音样本的预处理过程中可以使用人工神经网络模型。可以通过预处理过程将语音样本的时间、频率等调整为恒定。另外,当存在多个输入的语音样本时,可以执行用于筛选语音样本的预处理过程。语音特征提取部120可以从预处理的语音样本提取包括语音的基频(fundamental frequency)、发声速度、发声时间、发声长度、停顿程度、停顿次数、停顿区间长度、振幅微扰(Shimmer)、抖动(Jitter)、共振峰(formant)、谐波噪声比(harmonic-to-noise ratio)、响度(loudness)、频谱质心(spectral centroid)梅尔频率倒谱系数(Mel Frequency Cepstral Coefficients,MFCC)、身份向量(i-vector)、发音速度、过零率(zero-crossing rate,zcr)、发音概率(voicing probability,vp)、线谱对(Line spectral paris,LSP)、周期扰动(Period perturbation)、振幅扰动商(Amplitude perturbation quotient,APQ)、劲度(Stiffness)、能量(Energy)、强度(声音大小,Intensity)、熵(Entropy)中的至少一种语音特征。

[0048] 接下来,将所述语音特征和所述人口统计信息输入到预先学习的预测模型,以预测所述受试者是否患有阿尔茨海默病S130。

[0049] 预测模型130可以处于预先学习的状态以通过语音特征提取部120提取的语音特征和通过数据输入部110提供的人口统计信息来预测受试者是否患有阿尔茨海默病。预测模型130可以包括线性回归模型、逻辑回归模型、机器学习模型和神经网络模型中的至少一种分析模型。

[0050] 预测模型130包括多变量逻辑回归模型,所述多变量逻辑回归模型可以配置为如下式1所示。

[0051] [式1]

$$[0052] \quad \log \left(\frac{p(X)}{1-p(X)} \right) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_p X_p$$

[0053] (其中, X_1 至 X_p 是作为自变量输入到预测模型的输入值,分别对应于p个语音特征和人口统计信息, β_1 至 β_p 对应于作为自变量的回归系数的常数值, β_0 对应于初始常数值, $p(X)$ 对应于痴呆风险概率值)。

[0054] 基于所述式1构成的预测模型130可以输出痴呆风险概率值,并输出基于所述痴呆风险概率值评估患者状态的状态信息。痴呆风险概率值是指被诊断为痴呆症的概率,状态信息是根据专业医生诊断标准对患者进行诊断的,患者的状态可以被确定为阿尔茨海默病或正常认知功能。例如,当计算出的痴呆风险概率值为0.5以上时,预测模型130可以将患者的状态确定为阿尔茨海默病,当计算出的痴呆风险概率值小于0.5时,将患者的状态确定为正常认知功能。

[0055] 根据这些实施例的基于语音特征的阿尔茨海默病预测方法,可以被实现为应用程序或被实现为可以通过各种计算机组件执行的程序指令的形式,并记录在计算机可读记录介质中。所述计算机可读记录介质可以以单独或组合形式包括程序指令、数据文件、数据结构等。

[0056] 计算机可读记录介质的例子包括如硬盘、软盘和磁带的磁介质、如CD-ROM、DVD的光记录介质、如软盘(floptical disk)的磁光介质(magneto-optical media)以及如ROM、RAM、闪存等的专门配置用于存储和执行程序指令的硬件设备。

[0057] 程序指令的例子不仅包括如由编译器制作的机器语言代码,还包括可以使用解释器等由计算机执行的高级语言代码。所述硬件设备可以被配置为作为一个以上的软件模块来操作以执行根据本发明的处理,反之亦然。

[0058] 到目前为止,本发明针对优选的实施例进行了研究。本发明所属技术区域的普通技术人员可以理解,在不脱离本发明的本质特征的情况下,可以以修改的形式实施本发明。因此,所公开的实施例应被认为是说明性的而非限制性的。本发明的范围体现在权利要求范围内而不是上述说明,凡在与其同等范围内的所有差异均应被解释为包含在本发明中。

[0059] 实验例

[0060] 进行了实验以构建根据上述实施例的基于语音特征的阿尔茨海默氏病预测方法和装置的预测模型,并验证所构建的预测模型。

[0061] 用于构建预测模型的受试者数据是从总共210名患者中获得的,语音和诊断信息是从访问首尔波拉美(Boramae)医院的患者和在铜雀区(Dongjak-gu)痴呆症安心中心登记的患者中获得的。在总受试者中,阿尔茨海默病组有106名,正常组有104名。为了收集受试者的语音,在与检查者与受试者互动时收集了语音。

[0062] 为了量化受试者的语音,首先利用人工神经网络模型(卷积神经网络)对语音样本进行了预处理。即,当输入的语音样本数量较多时,首先执行用于筛选人的语音样本的预处理过程,从而防止噪音数据输入到学习中。对预处理后的语音,采用自动化语音特征提取方法提取了语音的语音特征。分别提取了如语音的基频(fundamental frequency(f0) mean,f0 std)、与发声相关的信息(发声速度、发声时间、发声长度)、与停顿(pause)相关的信息(停顿率(pause rate)、停顿次数(pause count)、停顿时长平均值(pause duration mean)、停顿时长标准差(pause duration standard deviation(std)))、振幅微扰(Shimmer)、抖动(Jitter)、共振峰(formant)、谐波噪声比(harmonic-to-noise ratio)、响度(loudness)、频谱质心(频谱质心平均值(spectral centroid mean)、频谱质心标准差(spectral centroid std))的语音特征。

[0063] 将人口统计信息(年龄、性别、受教育年限)和提取的每个语音特征作为输入值,并且阿尔茨海默病预测模型输出受试者的痴呆风险概率值,基于所述痴呆风险概率值,构建了输出评估患者状态的状态信息的预测模型。预测模型被实现为多变量逻辑回归模型,并被构建为输出痴呆风险概率值。当计算出的痴呆风险概率值为0.5以上时,预测模型可以将患者的状态确定为阿尔茨海默病,当计算出的痴呆风险概率值小于0.5时,将患者的状态确定为正常认知功能。

[0064] 通过构建的预测模型测试了预测性能。具体地,针对预测模型,计算出预测性能的指数加载器的受试者操作特征(receiver operating characteristic,ROC)曲线(curve)的曲线下面积(area under the curve,AUC)。AUC是通过计算ROC曲线下面积得到的值,是表示预测模型整体性能的代表性指数,越接近1表示性能越好。图4是示出构建的预测模型的结果的图表,根据约登指数(Youden Index)的最佳截止分数(optimal cutoff score),表现出如下表1所示的性能。

[0065] 表1

[0066]	截止分数 (Cutoff)	AUC	AUC.se	灵敏度 (Sensitivity)	特异度 (Specificity)	阳性预 测值 (PPV)	阴性预测 值(NPV)
	0.437	0.816	0.029	0.802	0.699	0.733	0.774

[0067] 预测能力为AUC=0.816,灵敏度(Sensitivity)为0.802,特异度(Specificity)为0.802。阳性预测值(positive predictive value,PPV)为0.733,阴性预测值(negative predictive value,NPV)为0.774。

[0068] 以往,用于诊断阿尔茨海默病的检查是使用了利用对脑内沉积的淀粉样蛋白的PET和脑脊液检查的方法。使用淀粉样蛋白PET的使用成本高,除了三级医院等开设专科医疗中心的医院外很难使用,而且还存在如暴露于辐射的风险因素。对于脑脊液分析,存在使用侵入性腰椎穿刺术、操作和分析需要劳动、可靠性因机构而异的局限性。

[0069] 相比之下,本发明的基于语音特征的阿尔茨海默病预测方法和装置通过使用非侵入性方法的检查来诊断阿尔茨海默病,并且可以通过AI扬声器/智能手机/平板电脑/计算机等进行的超越时间、空间、专科医生的限制的精神科筛选检查,可以在家中或非医院环境

中舒适地进行检查。另外,可以在一级、二级诊所确定阿尔茨海默病的危险程度,并且可以最大限度地减少假阳性,具有节省成本的效果,可以扩展到未来的治疗方案中。

[0070] 即,考虑到用于准确诊断阿尔茨海默病(一种典型的导致痴呆症的疾病)的脑脊液检查或正电子发射断层扫描(Positron emission tomograph,PET)被专科医生限制在特定的空间和时间,利用如本发明的AI扬声器以及智能设备的针对对话受试者进行的阿尔茨海默病测试方法有望帮助克服现有测试的时间、空间和资源限制。

[0071] 因此,采用语音的声学特征进行阿尔茨海默病风险诊断,有望为不断增加的痴呆症人群提供合适的干预方案。

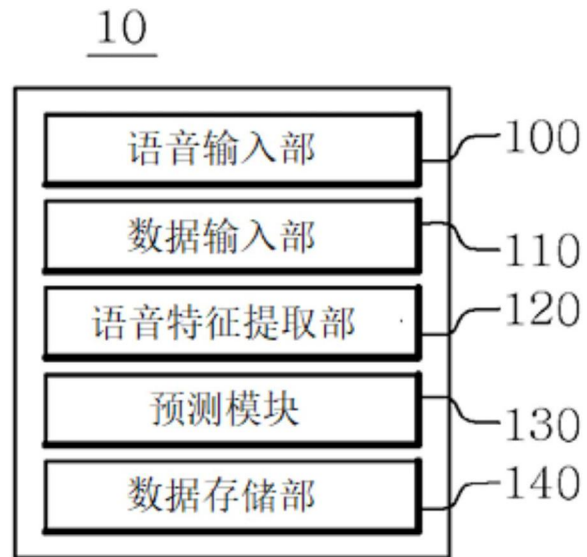


图1

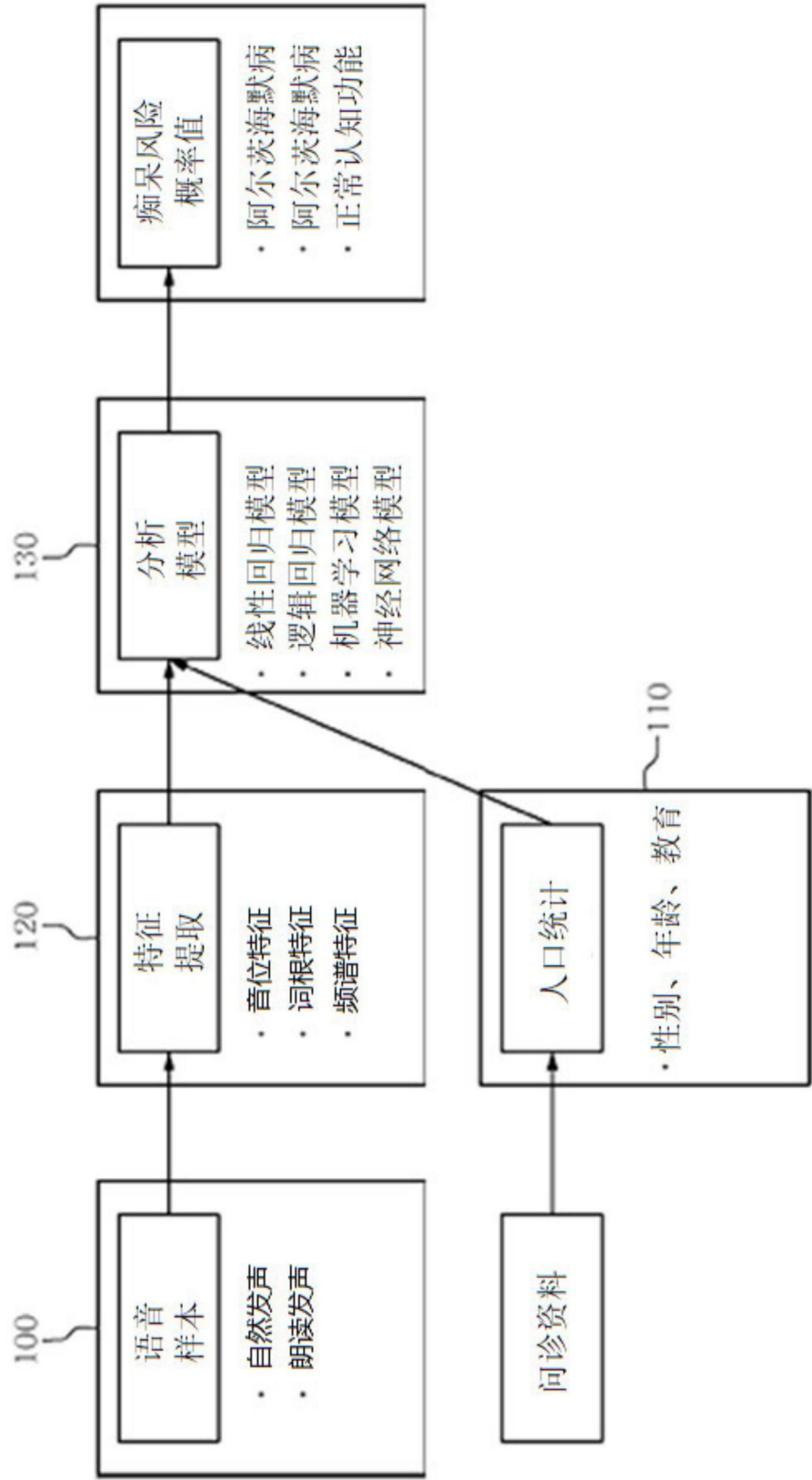


图2

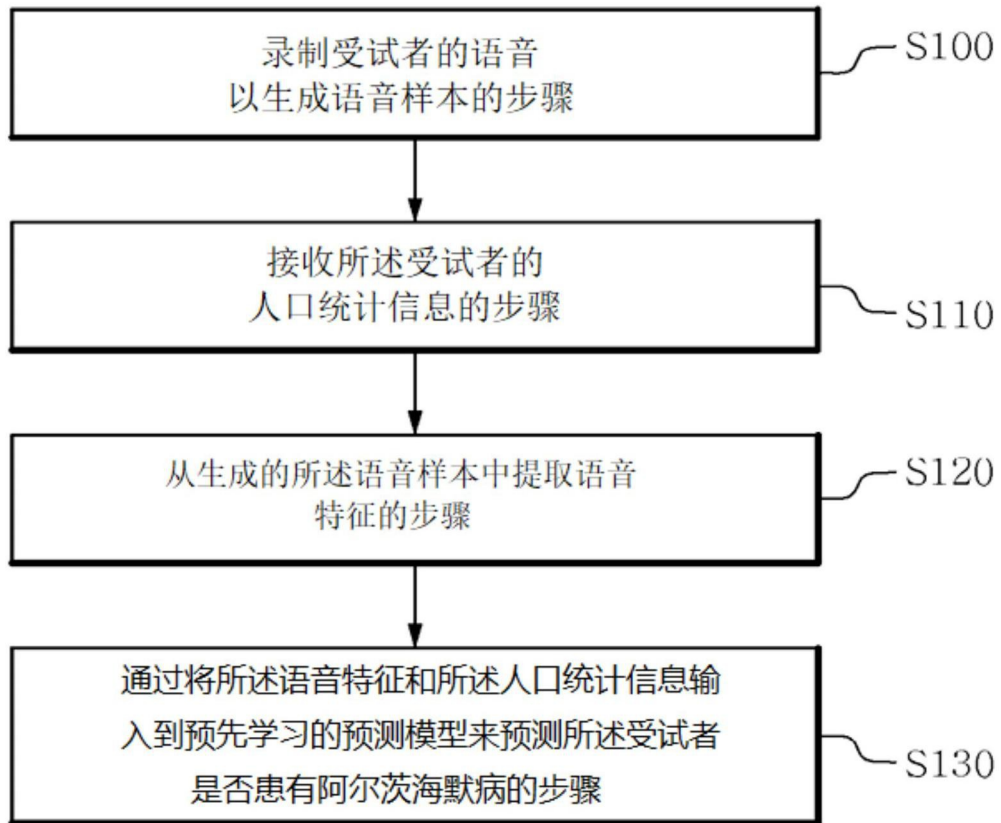


图3

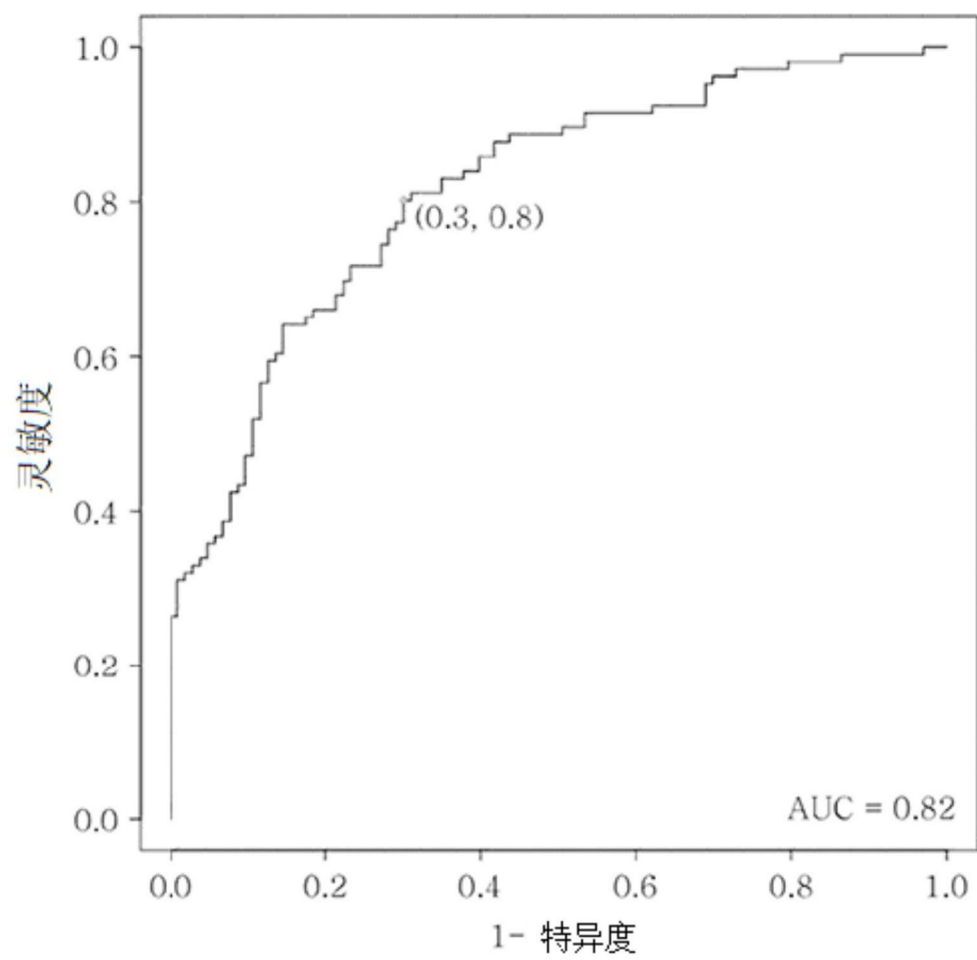


图4